Nanoscale Surface and Interface Mechanics of Elastic-Plastic Media with Smooth, Patterned, and Rough Surfaces

By

Xi Yin

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Engineering – Mechanical Engineering in the Graduate Division of the University of California, Berkeley

Committee in charge:
Professor Kyriakos Komvopoulos, Chair
Professor Hari Dharan
Professor Ronald Gronsky

Spring 2011
Nanoscale Surface and Interface Mechanics of Elastic-Plastic Media with Smooth, Patterned, and Rough Surfaces

Copyright © 2011

By

Xi Yin
Abstract

Nanoscale Surface and Interface Mechanics of Elastic-Plastic Media with Smooth, Patterned, and Rough Surfaces

by

Xi Yin

Doctor of Philosophy in Engineering-Mechanical Engineering

University of California, Berkeley

Professor Kyriakos Komvopoulos, Chair

The main objective of this dissertation was to develop both finite element and analytical models of contact, friction, and wear phenomena encountered at the nanoscale. This was achieved by the development of continuum mechanics and discrete dislocation models of the deforming homogeneous or layered media and the use of self-affine (fractal) geometry for the representation of the interface topography. The specific accomplishments of this work are as follows.

The contact problem of a rigid flat indenting an elastic-plastic semi-infinite medium with a sinusoidal surface profile was examined in light of a two-dimensional plane-strain finite element analysis. Numerical results of the dimensionless contact pressure, normal approach, average surface rise, center-line-average roughness, peak-to-valley roughness, cavity volume, and ratio of truncated-to-real contact area versus fractional contact area obtained for relatively compliant and stiff elastic-perfectly plastic half-spaces were compared with results obtained from a slip-line plasticity analysis. These results have direct application to metal working processes, such as rolling, and provide insight into the evolution of surface and subsurface contact deformation and asperity interaction of contacting surfaces exhibiting periodic waviness.

Mechanical failure of patterned alternating phase-shift mask (APSM) nanostructures due to dynamic pressure loadings caused by megasonic cleaning was examined in the context of simulation results obtained from a two-dimensional plane-strain finite element analysis. A parametric study of the effects of microstructure geometry and loading frequency on the subsurface stress state and mask structural integrity was performed for two typical chromium-quartz APSM patterns. Numerical results elucidate possible failure modes and effect of microstructure dimensions on
pattern damage during megasonic cleaning, and have direct implications to the design of extreme ultraviolet lithography masks and optimization of the megasonic cleaning process.

Analytical models were developed to study the friction, wear, energy dissipation, and plastic flow of surfaces exhibiting multi-scale roughness in both sliding and normal contacts. A contact mechanics study of friction, energy dissipation, and abrasive wear of a hard and rough (fractal) surface sliding against a soft and smooth substrate was developed based on the slip-line theory of plasticity. The slip-line model yields relationships of the deformation behavior and coefficient of friction of a fully plastic asperity microcontact in terms of the applied normal load and interfacial adhesion. The analysis of the rough surface contact provides insight into the dependence of global friction coefficient, energy dissipated during sliding contact, and abrasive wear rate and wear coefficient on the global interference (total normal load effect), interfacial friction conditions (adhesion effect), fractal parameters (roughness effect), and elastic-plastic material properties (deformation mode effect). Numerical results for representative contact systems illustrate the effects of interfacial adhesion, global interference (total normal load), topography parameters, and material properties on friction coefficient, dissipated frictional energy, and wear rate/coefficients.

The dependence of plastic deformation at asperity contacts and wear rate (coefficient) on global interference (total normal load), elastic-plastic material properties, topography (roughness), and work of adhesion of contacting surfaces was examined in a contact mechanics analysis of adhesive wear of rough (fractal) surfaces in normal contact. Loss of materials (wear) was presumed to originate from plastic contacting asperities, accounting for the contribution of interfacial adhesion to the normal load at each asperity microcontact. The effects of material properties, roughness, surface compatibility, and environmental conditions on the adhesive wear rate and wear coefficient were discussed in the context of numerical results for representative contact systems.

Plane-strain indentation of a single-crystal semi-infinite medium by a rigid indenter was analyzed by discrete dislocation plasticity. The profile of the rigid indenter was characterized by either a smooth (cylindrical) or a rough (fractal) surface. This is the first contact analysis based on discrete dislocations derived for crystalline materials indented by a surface exhibiting multi-scale roughness. Short-range dislocation interactions were modeled in accord to dislocation constitutive rules, while long-range dislocation interactions were modeled by the elastic stress fields of edge dislocations. Simulation results provided insight into the effects of contact load, dislocation source and obstacle densities, slip-plane orientation and distribution, indenter radius, topography (roughness) of fractal surface, and multi-scale asperity interactions on damage at the onset of yielding (emission of first dislocation dipole)
and plasticity evolution represented by the development of dislocation structures. Plastic deformation under the theoretical strength of the material was related to contact size effects.

The findings in this dissertation provide fundamental understanding of surface deformation behavior, evolution of subsurface stress field due to contact traction, and tribological characteristics of elastic-plastic media with patterned and rough surface profiles subject to contact and/or surface loadings. The obtained results have direct implications in various industry fields, such as metal working, semiconductor electronics packaging, magnetic storage recording, and microelectromechanical devices.
Dedications

To my father, mother, and sister
for their unconditional love, support, and understanding throughout my life
# Table of Contents

List of Figures ..................................................................................................................... v  
List of Tables ......................................................................................................................... xvi  
Acknowledgements ............................................................................................................. xvii  
Chapter 1  *Introduction* ........................................................................................................ 1  
Chapter 2  *Finite element analysis of an elastic-perfectly plastic half-space with a sinusoidal surface profile compressed by a rigid plane* ................................................. 6  
   2.1. Introduction .................................................................................................................. 6  
   2.2. Finite element model of a contacting sinusoidal surface ............................................ 7  
   2.3. Results and discussion .............................................................................................. 10  
      2.3.1. Apparent contact pressure .................................................................................. 10  
      2.3.2. Normal approach ............................................................................................. 13  
      2.3.3. Surface roughness ............................................................................................ 14  
      2.3.4. Deformed surface shape and average surface rise .......................................... 14  
      2.3.5. Cavity volume ................................................................................................. 17  
      2.3.6. Ratio of truncated contact area to real contact area .......................................... 17  
   2.4. Conclusions ................................................................................................................ 20  
Chapter 3  *Dynamic finite element analysis of failure in alternating phase-shift masks caused by megasonic cleaning* ................................................................. 21  
   3.1. Introduction ................................................................................................................ 21  
   3.2. Finite element modeling ............................................................................................ 23  
   3.3. Results and discussion .............................................................................................. 28  
   3.4. Conclusions ................................................................................................................ 43  
Chapter 4  *A slip-line plasticity analysis of sliding friction of rough surfaces exhibiting self-affine (fractal) behavior* .......................................................... 44  
   4.1. Introduction ................................................................................................................ 44  
   4.2. Single-asperity plasticity analysis ............................................................................. 46  
   4.3. Contact mechanics analysis ..................................................................................... 50  
      4.3.1. Rough surface contact model .......................................................................... 50  
      4.3.2. Friction analysis ............................................................................................... 54  
      4.3.3. Energy dissipation at the contact interface ....................................................... 55  
   4.4. Results and discussion .............................................................................................. 57  
      4.4.1. Single microcontact .......................................................................................... 57  
      4.4.2. Rough surface ................................................................................................. 58
### 4.5. Conclusions

**Chapter 5  A slip-line plasticity analysis of abrasive wear of rough (fractal) surfaces**

- 5.1. Introduction ................................................................. 67
- 5.2. Single-contact plasticity analysis ........................................ 68
- 5.3. Rough surface contact analysis .......................................... 72
  - 5.3.1. Contact model .......................................................... 72
  - 5.3.2. Wear rate ................................................................. 77
  - 5.3.3. Wear coefficient ....................................................... 79
- 5.4. Results and discussion ..................................................... 80
  - 5.4.1. Deformation mode at a single asperity contact ................ 80
  - 5.4.2. Wear rate ................................................................. 83
  - 5.4.3. Wear coefficient ....................................................... 86
- 5.5. Conclusions ................................................................. 87

**Chapter 6  An adhesive wear model of fractal surfaces in normal contact**

- 6.1. Introduction ................................................................. 90
- 6.2. Surface description ........................................................ 91
- 6.3. Contact mechanics analysis ............................................ 93
- 6.4. Results and discussion ..................................................... 101
  - 6.4.1. Deformation mode of asperity contacts ......................... 103
  - 6.4.2. Wear rate ................................................................. 107
  - 6.4.3. Wear coefficient ....................................................... 108
  - 6.4.4. Applicability and extension of the present analysis ........... 111
- 6.5. Conclusions ................................................................. 112

**Chapter 7  A discrete dislocation plasticity analysis of a single-crystal half-space indented by a rigid cylinder**

- 7.1. Introduction ................................................................. 114
- 7.2. Analytical model ............................................................ 116
- 7.3. Results and discussion ..................................................... 121
  - 7.3.1. Nucleation of dislocations and initiation of yielding ............ 122
  - 7.3.2. Multiplication of dislocations and evolution of dislocation structures 124
  - 7.3.3. Plastic flow under theoretical strength ............................. 130
- 7.4. Conclusions ................................................................. 133

**Chapter 8  A discrete dislocation plasticity analysis of a single-crystal semi-infinite medium indented by a rigid surface exhibiting multi-scale roughness**

- 8.1. Introduction ................................................................. 135
- 8.2. Analytical model ............................................................ 137
  - 8.2.1. Rough surface representation ...................................... 137
Fig. 2.1 Schematic showing indentation of a deformable half-space with a sinusoidal surface by a rigid plane. Because of symmetry, only the region ABCD (cell) was analyzed with the finite element method.

Fig. 2.2 Finite element model of the cell structure.

Fig. 2.3 Analytical and finite element simulation results for the dimensionless apparent contact pressure $Q$ versus the fractional contact area $\gamma$ during loading for relatively compliant ($E_1$) and stiff ($E_2 = 100 \ E_1$) half-spaces.

Fig. 2.4 Expansion of the plastic zone with increasing fractional contact area $\gamma$ for relatively compliant half-space ($E_1/\sigma_y = 1000$): (a) $\gamma = 0.1$, (b) $\gamma = 0.3$, (c) $\gamma = 0.5$, and (d) $\gamma = 0.7$.

Fig. 2.5 Analytical and finite element simulation results for the dimensionless normal approach $\Delta$ versus the fractional contact area $\gamma$ during loading for relatively compliant ($E_1$) and stiff ($E_2 = 100 \ E_1$) half-spaces.

Fig. 2.6 Analytical and finite element simulation results for (a) the dimensionless center-line-average roughness $R_a$, and (b) the dimensionless peak-to-valley roughness $H$ versus the fractional contact area $\gamma$ for relatively compliant ($E_1$) and stiff ($E_2 = 100 \ E_1$) half-spaces.

Fig. 2.7 Evolution of the surface profile of the relatively compliant elastic-perfectly plastic half-space ($E_1/\sigma_y = 1000$) with the dimensionless global interference: (a) during loading, and (b) after full unloading.

Fig. 2.8 Analytical and finite element simulation results for the normalized average surface rise $\bar{\Delta}_r$ versus the fractional contact area $\gamma$ after full unloading for relatively compliant ($E_1$) and stiff ($E_2 = 100 \ E_1$) half-spaces.

Fig. 2.9 Analytical and finite element simulation results for the normalized cavity volume $V$ versus fractional contact area $\gamma$ for relatively compliant ($E_1$) and stiff ($E_2 = 100 \ E_1$) half-spaces during loading.

Fig. 2.10 Ratio of truncated-to-real contact area (contact area ratio) $a'/a$ versus fractional contact area $\gamma$ for relatively compliant ($E_1$) and stiff ($E_2 = 100 \ E_1$) half-spaces during loading.
Fig. 3.1 Schematic of the megasonic cleaning process of EUV photomasks (not to scale). The pattern microstructures have been exaggerated for clarity. In reality, the fluid stream impinges onto a huge number (on the order of ~10^8) of microstructures.

Fig. 3.2 AFM image showing extensive damage of a microstructure on a patterned chromium-quartz photomask due to megasonic cleaning. The nonuniform damage extends into the quartz substrate, suggesting that failure occurred by an abrupt failure process.

Fig. 3.3 Cross-section schematics of typical chromium-quartz microstructures under a uniform pressure $p$ ($a =$ undercut depth, $d =$ zero-aperture depth, $b =$ line width, $t =$ chromium layer thickness, and $l =$ distance between zero-aperture and $\pi$-aperture): (a) pattern A and (b) pattern B. The region enclosed by a thick (blue) line is the cell analyzed by the FEM.

Fig. 3.4 Schematic of cavitation bubble collapse in the proximity of the photomask surface.

Fig. 3.5 Schematics of pressure loading functions used in the FEM analysis: (a) step dynamic pressure $p_c$ applied at time $t_0$, (b) cyclic dynamic pressure ($T =$ loading period, $\tau =$ time under maximum pressure $p_{\text{max}}$ during a loading period), and (c) static pressure ($T =$ loading period, $p_{\text{max}} =$ maximum pressure).

Fig. 3.6 FEM models of cell microstructures in (a) pattern A and (b) pattern B under a uniform normal pressure $p$. (The red nodes were used to calculate the residual stress in the chromium layer).

Fig. 3.7 Maximum first principal stress $\sigma_{i}^{\text{max}}$ in the cell microstructure of (a) pattern A ($t =$ 100 nm, $a =$ 40 nm, $d =$ 40 nm, $b =$ 160 nm) and (b) pattern B ($t =$ 100 nm, $b =$ 150 nm) due to dynamic pressure loading ($f =$ 1 MHz) as a function of the size of the smallest finite element $h_{\text{min}}$. 

vi
Fig. 3.8 Collaps pressure $p_c$ and equivalent plastic strain $\varepsilon_p$ in the cell microstructure of pattern A versus (a) undercut depth $a$ ($d = 80$ nm, $t = 100$ nm, $b = 200$ nm), (b) zero-aperture depth $d$ ($a = 40$ nm, $t = 100$ nm, $b = 160$ nm), (c) chromium layer thickness $t$ ($a = 50$ nm, $d = 80$ nm, $b = 200$ nm) and (d) line width $b$ ($a = 30$ nm, $d = 80$ nm, $t = 100$ nm).

Fig. 3.9 Collaps pressure $p_c$ and equivalent plastic strain $\varepsilon_p$ in the cell microstructure of pattern B versus (a) chromium layer thickness $t$ ($b = 130$ nm) and (b) line width $b$ ($t = 100$ nm).

Fig. 3.10 Maximum tensile stress $\sigma_{\text{I}}^{\text{max}}$ in the cell microstructure of pattern A due to static and dynamic pressure loading ($f = 1$, 5, and 10 MHz) versus (a) undercut depth $a$ ($d = 80$ nm, $t = 100$ nm, $b = 200$ nm), (b) zero-aperture depth $d$ ($a = 40$ nm, $t = 100$ nm, $b = 160$ nm), (c) chromium layer thickness $t$ ($a = 50$ nm, $d = 80$ nm, $b = 200$ nm) and (d) line width $b$ ($a = 30$ nm, $d = 80$ nm, $t = 100$ nm).

Fig. 3.11 Maximum tensile stress $\sigma_{\text{I}}^{\text{max}}$ in the cell microstructure of pattern B due to static and dynamic pressure loading ($f = 1$, 5, and 10 MHz) versus (a) chromium layer thickness $t$ ($b = 130$ nm) and (b) line width $b$ ($t = 100$ nm).

Fig. 3.12 First principal stress $\sigma_{\text{I}}$ in the cell microstructure of (a) pattern A ($t = 100$ nm, $a = 50$ nm, $d = 80$ nm, $b = 200$ nm) and (b) pattern B ($t = 100$ nm, $b = 50$ nm) due to dynamic pressure loading ($f = 10$ MHz). Locations of $\sigma_{\text{I}}^{\text{max}}$ are identified by arrows.

Fig. 3.13 Maximum equivalent plastic strain $\varepsilon_p^{\text{max}}$ in the cell microstructure of pattern A due to static and dynamic pressure loading ($f = 1$, 5, and 10 MHz) versus (a) undercut depth $a$ ($d = 80$ nm, $t = 100$ nm, $b = 200$ nm), (b) zero-aperture depth $d$ ($a = 40$ nm, $t = 100$ nm, $b = 160$ nm), (c) chromium layer thickness $t$ ($a = 50$ nm, $d = 80$ nm, $b = 200$ nm) and (d) line width $b$ ($a = 30$ nm, $d = 80$ nm, $t = 100$ nm). The vertical dashed lines indicate the critical microstructure dimensions.
Fig. 3.14 Maximum equivalent plastic strain $\varepsilon_p^{\text{max}}$ in the cell microstructure of pattern B due to static and dynamic pressure loading ($f = 1$, 5, and 10 MHz) versus (a) chromium layer thickness $t$ ($b = 130$ nm) and (b) line width $b$ ($t = 100$ nm). The vertical dashed lines indicate the critical microstructure dimensions.

Fig. 3.15 Equivalent plastic strain $\varepsilon_p$ in the cell microstructure of (a) pattern A ($t = 100$ nm, $a = 50$ nm, $d = 80$ nm, $b = 200$ nm) and (b) pattern B ($t = 100$ nm, $b = 50$ nm) due to dynamic pressure loading ($f = 10$ MHz). Locations of $\varepsilon_p^{\text{max}}$ are identified by arrows.

Fig. 3.16 Maximum equivalent plastic strain $\varepsilon_p^{\text{max}}$ in the cell microstructure of (a) pattern A ($t = 100$ nm, $a = 50$ nm, $d = 80$ nm, $b = 200$ nm) and (b) pattern B ($t = 100$ nm, $b = 90$ nm) due to dynamic pressure loading ($f = 10$ MHz) as a function of loading cycles $N$.

Fig. 3.17 Effect of average residual stress in the chromium layer $\bar{\sigma}_1^R$ on the first principal stress $\sigma_1$ in the cell microstructure of pattern A ($t = 100$ nm, $a = 50$ nm, $d = 80$ nm, $b = 200$ nm) due to dynamic pressure loading ($f = 10$ MHz): (a) $\sigma_1^{\text{max}}$ versus $\bar{\sigma}_1^R$, and $\sigma_1$ distributions for $\bar{\sigma}_1^R$ equal to (b) $-1$ GPa, (c) 0, and (d) 1 GPa. Locations of $\sigma_1^{\text{max}}$ are identified by arrows.

Fig. 3.18 Effect of average residual stress in the chromium layer $\bar{\sigma}_1^R$ on the first principal stress $\sigma_1$ in the cell microstructure of pattern B ($t = 100$ nm, $b = 50$ nm) due to dynamic pressure loading ($f = 10$ MHz): (a) $\sigma_1^{\text{max}}$ versus $\bar{\sigma}_1^R$, and $\sigma_1$ distributions for $\bar{\sigma}_1^R$ equal to (b) $-1.0$, (c) 0, and (d) 1.0 GPa. Locations of $\sigma_1^{\text{max}}$ are identified by arrows.
Fig. 4.1  (a) Slip-line field of a rigid spherical asperity (wear particle) plowing through a rigid-perfectly plastic surface and removing material by a microcutting process, (b) detailed view of the plastic region showing the orthogonal network of $\alpha$- and $\beta$-lines, and (c) hodograph of the slip-line field (Komvopoulos, 2011).

Fig. 4.2  Three-dimensional fractal surface generated from Eq. (4.14) for $G = 2.39 \times 10^{-4}$ nm, $D = 2.24$, $M = 10$, $\gamma = 1.5$, $L = 7.04 \times 10^{7}$ nm, and $L_x = 5$ nm ($\sigma = 290$ nm).

Fig. 4.3  Two-dimensional profile of a truncated rough (fractal) surface.

Fig. 4.4  Friction coefficient $\mu$ of a fully plastic microcontact as a function of dimensionless normal load $\Delta F_y/(kR^2)$ and ratio of interfacial shear strength to shear strength of the soft material $s/k$.

Fig. 4.5  (a) Total contact area of truncated asperities normalized by the apparent contact area $S'/S_a$ and (b) total contact area of truncated fully plastic asperities normalized by the total contact area of truncated asperities $S'_p/S'$ versus global interference $h$ for Al$_2$O$_3$/CrN contact system with material properties given in Table 4.1 ($\sigma = 1000$ nm).

Fig. 4.6  (a) Friction coefficient $f$ of Al$_2$O$_3$/CrN contact system versus global interference $h$ and ratio of interfacial shear strength to shear strength of the softer material $s/k$ for $\sigma = 1000$ nm; (b) friction coefficient $f$ of Al$_2$O$_3$/CrN, Al$_2$O$_3$/TiC, and AISI 1095/AISI 1020 contact systems versus fractal dimension $D$ and rms surface roughness for $h = 15$ nm, $G = 2.39 \times 10^{-4}$ nm, and $s/k = 0.5$; and (c) friction coefficient $f$ of Al$_2$O$_3$/CrN contact system versus rms surface roughness $\sigma$ and ratio of interfacial shear strength to shear strength of the softer material $s/k$ for $h = 15$ nm and $D = 2.24$. (The material properties of each contact system are given in Table 4.1.)
Fig. 4.7 (a) Energy ratio $\xi_c$ of $\text{Al}_2\text{O}_3$/CrN contact system versus global interference $h$ and ratio of interfacial shear strength to shear strength of the softer material $s/k$ for $\sigma = 1000$ nm; (b) energy ratio $\xi_c$ of $\text{Al}_2\text{O}_3$/CrN, $\text{Al}_2\text{O}_3$/TiC, and AISI 1095/AISI 1020 contact systems versus fractal dimension $D$ and rms surface roughness $\sigma$ for $h = 15$ nm, $G = 2.39 \times 10^{-4}$ nm, and $s/k = 0.5$; and (c) energy ratio $\xi_c$ of $\text{Al}_2\text{O}_3$/CrN contact system versus rms roughness $\sigma$ and ratio of interfacial shear strength to shear strength of the softer material $s/k$ for $h = 15$ nm and $D = 2.24$. (The material properties of each contact system are given in Table 4.1.)

Fig. 4.8 Ratios of elastic, plastic shearing, and material removal energies normalized by the total external energy $\xi_e$, $\xi_s$, $\xi_c$, respectively, of $\text{Al}_2\text{O}_3$/CrN contact system versus ratio of interfacial shear strength to shear strength of the softer material $s/k$ for $h = 15$ nm and $\sigma = 1.5 \times 10^4$ nm. (The material properties of the $\text{Al}_2\text{O}_3$/CrN contact system are given in Table 4.1.)

Fig. 5.1 (a) Slip-line field of a hard spherical asperity (or wear particle) removing material from a soft surface by a microcutting process, (b) detailed view of plastic region showing the network of orthogonal $\alpha$ and $\beta$ slip-lines, and (c) hodograph of the slip-line field (Komvopoulos, 2011).

Fig. 5.2 Three-dimensional fractal surface generated from Eq. (5.7) for $D = 2.24$, $G = 2.39 \times 10^{-4}$ nm, $\gamma = 1.5$, $M = 10$, $L = 7.04 \times 10^3$ nm, and $L_s = 5$ nm $(\sigma = 290$ nm).

Fig. 5.3 Two-dimensional profile of a truncated rough (fractal) surface.

Fig. 5.4 Degree of penetration $D_p$ versus dimensionless shear strength $s/k$. Discrete points are experimental data obtained at the transition from cutting to plowing/wedge formation of sliding surfaces (Hokkirigawa and Kato, 1988). The solid curve fitted through the experimental data defines the boundary between abrasive wear (cutting) and no wear (plowing/wedge formation) conditions.

Fig. 5.5 Ratio of penetration depth during sliding to indentation depth $d_p/d_a$ versus dimensionless normal load $P/(kR^3)$ and interfacial shear strength $s/k$ for a single asperity microcontact. The dashed curve and solutions for $s/k = 0$ and 1 define the regime of admissible slip-line solutions.
Fig. 5.6  (a) Pile-up thickness-to-asperity radius ratio \( t/R \) and (b) pile-up thickness-to-penetration depth ratio \( t/d_s \) versus dimensionless normal load \( P/(kR^3) \) and interfacial shear strength \( s/k \) for a single asperity microcontact. The dashed curve and solutions for \( s/k = 0 \) and \( 1 \) define the regime of admissible slip-line solutions.

Fig. 5.7  (a) Wear rate \( V/S \) of \( \text{Al}_2\text{O}_3/\text{CrN} \) contact system versus global interference \( h \) and dimensionless interfacial shear strength \( s/k \) for \( D = 2.24, \ G = 1.07 \times 10^{-4} \) nm, and \( \sigma = 100 \) nm; (b) Wear rate \( V/S \) of \( \text{Al}_2\text{O}_3/\text{CrN} \) contact system versus surface roughness \( \sigma \) and dimensionless interfacial shear strength \( s/k \) for \( D = 2.24 \) and \( h = 30 \) nm; (c) Wear rate \( V/S \) of \( \text{Al}_2\text{O}_3/\text{CrN}, \ \text{Al}_2\text{O}_3/\text{TiC}, \ \text{and AISI 1095/AISI 1020} \) contact systems versus fractal dimension \( D \) and surface roughness \( \sigma \) for \( G = 2.39 \times 10^{-4} \) nm, \( h = 30 \) nm, and \( s/k = 0.5 \). (Elastic-plastic material properties of each contact system are given in Table 5.1.)

Fig. 5.8  Critical contact area of truncated asperity for fully plastic microcontact deformation \( a'_c \) of \( \text{Al}_2\text{O}_3/\text{CrN}, \ \text{Al}_2\text{O}_3/\text{TiC}, \ \text{and AISI 1095/AISI 1020} \) contact systems versus fractal dimension \( D \) and surface roughness \( \sigma \) for \( G = 2.39 \times 10^{-4} \) nm. (Elastic-plastic material properties of each contact system are given in Table 5.1.)

Fig. 5.9  (a) Wear coefficient \( K \) of \( \text{Al}_2\text{O}_3/\text{CrN} \) contact system versus global interference \( h \) and dimensionless interfacial shear strength \( s/k \) for \( D = 2.24, \ G = 1.07 \times 10^{-4} \) nm, and \( \sigma = 100 \) nm; (b) Wear coefficient \( K \) of \( \text{Al}_2\text{O}_3/\text{CrN} \) contact system versus surface roughness \( \sigma \) and dimensionless interfacial shear strength \( s/k \) for \( D = 2.24 \) and \( h = 30 \) nm; (c) Wear coefficient \( K \) of \( \text{Al}_2\text{O}_3/\text{CrN}, \ \text{Al}_2\text{O}_3/\text{TiC}, \ \text{and AISI 1095/AISI 1020} \) contact systems versus fractal dimension \( D \) and surface roughness \( \sigma \) for \( G = 2.39 \times 10^{-4} \) nm, \( h = 30 \) nm, and \( s/k = 0.5 \). (Elastic-plastic material properties of each contact system are given in Table 5.1.)

Fig. 6.1  Three-dimensional fractal surface generated from Eq. (6.3) for \( D = 2.24, \ G = 2.39 \times 10^{-4} \) nm, \( M = 10, \ \gamma = 1.5, \ L = 7.04 \times 10^3 \) nm, and \( L_s = 5 \) nm \( (\sigma = 290 \) nm).

Fig. 6.2  Two-dimensional profile of a rough (fractal) surface truncated by a rigid plane.
Fig. 6.3 Normalized mean contact pressure $p_m/\sigma_Y$ versus normalized representative strain $E\delta/\sigma_Y r'$ for two constitutive models of asperity deformation.

Fig. 6.4 Schematics of truncation and deformation of an asperity by a rigid plane.

Fig. 6.5 Critical truncated contact area $a'_c$ versus environment (interfacial adhesion) for surface properties given in Tables 6.1 and 6.2 ($\sigma = 290$ nm).

Fig. 6.6 Critical truncated contact area $a'_c$ versus rms roughness of equivalent surface $\sigma$ ($D = 2.24$) for surface properties given in Tables 6.1 and 6.2.

Fig. 6.7 Ratio of total truncated contact area of fully-plastic asperity contacts to apparent sample area $S'/S_a$ versus global interference $h$ for surface properties given in Tables 6.1 and 6.2 ($\sigma = 290$ nm).

Fig. 6.8 Ratio of total truncated contact area of fully-plastic asperity contacts to total truncated contact area $S'_p/S'$ versus global interference $h$ for surface properties given in Tables 6.1 and 6.2 ($\sigma = 290$ nm).

Fig. 6.9 Wear rate $dV_p^h/dh$ versus global interference $h$ for surface properties given in Tables 6.1 and 6.2 ($\sigma = 290$ nm).

Fig. 6.10 Wear rate $dV_p^h/dh$ versus environment (interfacial adhesion) for surface properties given in Tables 6.1 and 6.2 ($\sigma = 290$ nm).

Fig. 6.11 Wear rate $dV_p^h/dh$ versus rms roughness of equivalent surface $\sigma$ ($D = 2.24$) for surface properties given in Tables 6.1 and 6.2.

Fig. 6.12 Wear coefficient $K$ versus global interference $h$ for surface properties given in Tables 6.1 and 6.2 ($\sigma = 290$ nm).

Fig. 6.13 Wear coefficient $K$ versus environment (interfacial adhesion) for surface properties given in Tables 6.1 and 6.2 ($\sigma = 290$ nm).

Fig. 6.14 Wear coefficient $K$ versus rms roughness of equivalent surface $\sigma$ ($D = 2.24$) for surface properties given in Tables 6.1 and 6.2.
Fig. 7.1  Schematic of a single-crystal half-space with parallel and equally spaced slip planes of fixed orientation indented by a rigid cylinder.

Fig. 7.2  Damage parameter $D_Y$ defined at the onset of yielding (emission of first dislocation dipole) versus (a) dimensionless slip-plane distance $d/b$, (b) dislocation source density $\rho_s$, and (c) dimensionless indenter radius $R/b$.

Fig. 7.3  Dimensionless dislocation density $\rho_d/\rho_s$ versus slip-plane orientation angle $\theta$ for contact load $P=12$ N/m corresponding to the onset of yielding in the least favorable slip-plane direction ($\theta=90^\circ$).

Fig. 7.4  Dimensionless dislocation density $\rho_d/\rho_s$ versus dimensionless contact load $P/P_Y$ and slip-plane orientation angle $\theta$.

Fig. 7.5  Dislocation density $\rho_d$ versus dimensionless contact load $P/P_Y$ and dislocation source density $\rho_s$ for slip-plane orientation angle (a) $\theta=45^\circ$ and (b) $\theta=0^\circ$.

Fig. 7.6  Dimensionless dislocation density $\rho_d/\rho_s$ versus dimensionless contact load $P/bE$ and dimensionless indenter radius $R/b$ for slip-plane orientation angle (a) $\theta=45^\circ$ and (b) $\theta=0^\circ$.

Fig. 7.7  Dislocation maps for different contact widths $2r$ (loads) and slip-plane orientation angle $\theta=45^\circ$. (Positive and negative edge dislocations are shown by red and blue symbols, respectively.)

Fig. 7.8  Dimensionless shear-band width $w/b$ versus dimensionless contact load $P/P_Y$.

Fig. 7.9  Contours of dimensionless total resolved shear stress $\tau/\tau_e$ for different contact widths $2r$ corresponding to the dislocation maps shown in Fig. 7.

Fig. 7.10  Dimensionless total and mobile dislocation densities $\rho_d/\rho_s$ versus dimensionless load $P/P_Y$ and dislocation obstacle density $\rho_o$ for slip-plane orientation angle (a) $\theta=45^\circ$ and (b) $\theta=0^\circ$.

Fig. 7.11  Effect of (a) dislocation source density $\rho_s$ and (b) dislocation obstacle density $\rho_o$ on the variation of dimensionless mean contact pressure $p_m/Y$ with representative macroscopic strain $\varepsilon$.

Fig. 7.12  Ratio of threshold radius of rigid indenter to slip-plane distance $R_{th}/d$ versus dimensionless slip-plane distance $d/b$ for slip-plane orientation angle $\theta=0^\circ$ and $45^\circ$. 

xiii
Fig. 8.1 Critical segment of length $L_C = 250 \text{ nm}$ obtained from a fractal surface profile generated from Eq. (8.1) for $L = 5379 \text{ nm}$, $L_S = 5 \text{ nm}$, $\gamma = 1.5$, $D = 1.54$, and $G = 1.14 \text{ nm}$ ($\sigma = 29.2 \text{ nm}$).

Fig. 8.2 Schematic of single-crystal semi-infinite medium containing Frank-Read dislocation sources, which are randomly distributed on parallel and equally spaced slip-planes of fixed orientation indented by a rigid rough (fractal) surface. ($h_{\text{max}}$ = maximum global interference; $h_{12}$ = critical global interference for the interaction between asperities (1) and (2); and $h_{\text{AB}}$ = critical global interference for interaction between asperities (A) and (B).)

Fig. 8.3 Piecewise linear contact pressure profile consisting of overlapping triangular pressure elements used in the numerical procedure.

Fig. 8.4 (a) Dimensionless dislocation density $\rho_d/\rho_s$ and (b) dimensionless shear-band width $w/b$ versus dimensionless contact load $P/bE$ for a single-crystal semi-infinite medium with slip-plane orientation angle $\theta = 45^\circ$ indented by a rigid, smooth cylindrical asperity with dimensionless radius of curvature $R/b = 2 \times 10^4$ and a rigid, rough (fractal) asperity with topography parameters $D = 1.25$ and $G = 9.46 \times 10^{-4} \text{ nm}$.

Fig. 8.5 (a) Dimensionless dislocation density $\rho_d/\rho_s$ versus dimensionless contact load $P/bE$ and slip-plane orientation angle $\theta$ for a single-crystal semi-infinite medium indented by a rigid, rough (fractal) surface with topography parameters $D = 1.54$ and $G = 2.93 \times 10^{-2}$ nm, and (b) dislocation density $\rho_d$ versus dimensionless contact load $P/bE$ and dislocation source density $\rho_s$ for a single-crystal semi-infinite medium with slip-plane orientation angle $\theta = 0^\circ$ indented by the same rigid, rough surface.

Fig. 8.6 Evolution of dislocation structures below the surface of a single-crystal semi-infinite medium with slip-plane orientation angle $\theta = 0^\circ$ for dimensionless global interference $h/d$ equal to (a) 0.10, (b) 0.26, (c) 0.58, and (d) 0.90. Contact widths between asperities of the rough surface and the crystalline medium are denoted by thick lines at the medium surface ($x_2 = 0$). Positive and negative edge dislocations are shown by red and blue symbols, respectively. Slip-planes are only shown in (a) for clarity. The region on the left of the vertical dashed line in each figure was used to obtain the dislocation density plotted in Fig. 8.7.
Fig. 8.7 Dimensionless dislocation density $\rho d/\rho s$ versus dimensionless global interference $h/b$ for a single-crystal semi-infinite medium with slip-plane orientation angle $\theta = 0^\circ$ indented by a rough surface, a relatively large and rough asperity (A), and a relatively small and smooth asperity (1). Asperities (A) and (1) are defined in Fig. 8.2. Vertical dashed lines indicate the critical dimensionless global interference for asperity interaction at different length scales.

Fig. 8.8 Contours of dimensionless total resolved shear stress $\tau / \tau_s$ for dimensionless global interference $h/d$ equal to (a) 0.10, (b) 0.26, (c) 0.58, and (d) 0.90 corresponding to the dislocation structures shown in Fig. 8.6. Contact widths between asperities of the rough surface and the crystalline medium are denoted by thick lines at the medium surface ($x_2 = 0$).

Fig. 8.9 Dimensionless dislocation density $\rho d/\rho s$ versus dimensionless contact load $P/bE$ and dimensionless standard deviation of slip-plane distance $s/d$ for a single-crystal semi-infinite medium with randomly distributed slip-planes of orientation angle $\theta = 0^\circ$ indented by a rigid, rough (fractal) surface with topography parameters $D = 1.54$ and $G = 2.93 \times 10^{-2}$ nm.

Fig. 8.10 Dimensionless dislocation density $\rho d/\rho s$ versus dimensionless contact load $P/bE$ for a single-crystal semi-infinite medium with slip-plane orientation angle $\theta = 0^\circ$ indented by rigid, rough (fractal) surfaces of different rms roughness $\sigma$: (a) $D = 1.54$ and $G = 2.93 \times 10^{-2}$ nm ($\sigma = 10.8$ nm), 1.14 nm ($\sigma = 29.2$ nm), and 5.48 nm ($\sigma = 101.3$ nm), and (b) $G = 1.14$ nm and $D = 1.24$ ($\sigma = 107.5$ nm), 1.54 ($\sigma = 29.2$ nm), and 1.84 ($\sigma = 9.6$ nm).

Fig. 8.11 Dimensionless dislocation density $\rho d/\rho s$ versus dimensionless contact load $P/bE$ and surface roughness-to-standard deviation of slip-plane distance ratio $\sigma s$ for a single-crystal semi-infinite medium with slip-plane orientation angle $\theta = 0^\circ$ indented by rigid, rough (fractal) surfaces with $D = 1.54$ and different $G$ values.

Fig. A.1 Slip-line geometry used to determine angles $\alpha_2$, $\alpha_3$, and $\alpha_5$ in Eqs. (4.12) and (4.13) (Komvopoulos, 2011).
List of Tables

Table 3.1 Microstructure dimensions used in the FEM analysis.
Table 3.2 Material properties used in the FEM analysis.
Table 4.1 Elastic-plastic material properties of contacting surfaces
Table 5.1 Elastic-plastic material properties of contact systems
Table 6.1 Properties of contact surfaces
Table 6.2 Composite compatibility index of contact surfaces at different environments
Table 7.1 Material properties of single-crystal copper.
Table 8.1 Material properties of single-crystal semi-infinite medium.
Acknowledgements

First and foremost, I would like to extend the most sincere gratitude to my advisor Professor Kyriakos Komvopoulos for his innovative and patient guidance, intellectual and financial support, and very professional editing of our research papers. The creative thinking, independent learning skills, and the excellent research trainings I obtained with him during the five years of my PhD study at Berkeley will certainly benefit my entire career.

I am very grateful to Prof. Hari Dharan and Prof. Ronald Gronsky for their helpful suggestions and encouragement for the successful accomplishments of my PhD qualification exam and dissertation. I’d like to thank the other two committee members of my qualification exam, Prof. Tarek I. Zohdi and Prof. David Steigmann, for their challenging and helpful comments of my research.

This research work was partially funded by Intel Corporation. I am also grateful to Dr. Ted Liang, Dr. Kevin Orvek, Dr. Seh-jin Park, and Dr. Kishore Chakravorty from Intel Mask Operations for many stimulating discussions and helpful visits to their advanced industrial labs.

Last but not least, I wish to thank all of my colleagues in the Computational Surface Mechanical Laboratory (CSML) and Surface Science and Engineering Laboratory (SSEL) at the Department of Mechanical Engineering, especially Dr. Hanshen Zhang, Taekwon Jee, Dr. Shannon J. Timpe, Zhichao Song, and Huaming Xu, for their invaluable assistance and mental support for my completion of this research.
Chapter 1

Introduction

Surface mechanics is an important subject of solid mechanics. It deals with the deformation behavior, stress state, and failure characteristics at the material surface and subsurface due to surface tractions and/or mechanical contact. As an efficient investigation tool of tribological behavior, surface contact mechanics allows for basic understanding of the fundamental mechanisms of friction, wear, energy dissipation, and plastic deformation at contact interfaces of solid bodies in relative motion, also providing the means to enhance the efficiency and durability of mechanical contact systems.

Originating from the preliminary study of Hertz for frictionless elastic contact of ellipsoids (Johnson, 1987), contact mechanics analysis has been extended to different contact geometries and various elastic-plastic constitutive rules, with the inclusion of friction force. There are many engineering applications of contact mechanics, from macroscopic contact systems and processes, such as mechanical seals, bearings, automobile gears, and rolling process to microscopic contact systems, such as semiconductor electronics packaging, head-disk interface of computer hard-disk drives, and microswitches in microelectromechanical systems (MEMS). Advances in nanotechnology and continuing miniaturization of MEMS devices have increased the demand for materials that can function effectively and reliably at the nanoscale. Therefore, reliability and failure properties of these nano-/micro-devices have become increasingly important.

Patterned layered media have been used extensively in nanoelectronics such as semiconductor integrated circuit and high-density data storage devices. Various nanoscopic surface patterns are usually produced by lithography-based techniques. Due to concentrated surface forces or contact loads, high stresses are generated in the proximity of material surfaces, which can ultimately lead to mechanical failures due to microcracking, localized plastic flow, or contact fatigue. The mechanics of patterned nanostructures under processing loadings is therefore of prime importance. Because of complex material constitutive relationships, surface topography, and surface loading conditions, closed-form solutions of the deformation and stress fields of patterned layered media are either impossible to obtain or computationally extremely cumbersome. Therefore, numerical techniques, such as the finite element method (FEM) used in the present dissertation, have been used to numerically simulate the mechanical responses of the patterned layered media under process conditions.

At the nanoscale, even macroscopically smooth surfaces exhibit random
multi-scale roughness, which cannot be ignored in contact mechanics studies of real engineering materials. Contact of two rough surfaces occurs between asperities of varying size and random distribution over the contact interface. Relationships between contact load, real contact area, elastic-plastic material properties at the asperity level, surface topography, interfacial adhesion, near-surface microstructure and resulting subsurface stresses and strains, friction coefficient, wear rate, and dissipated frictional energy are essential in applications such as hard-disk drives, surface micromachining, and semiconductor electronics packaging. Traditionally, rough surfaces have been characterized by conventional statistical parameters, such as mean height, mean slope, and mean curvature of asperity contacts. However, these rough surface models are scale-dependent and are based on simplifying assumptions about the asperity shape. Therefore, fractal geometry characterized by scale invariance and self-affine property has been used in recent contact mechanics studies to represent the multi-scale roughness demonstrated by many real surfaces. Consequently, fractal geometry was also used in this dissertation to model randomly rough surfaces.

As a major surface failure mode, wear can be detrimental to normal operation and functionality of many scientific instruments and engineering components. Abrasive and adhesive wear are two of the most common wear mechanisms of metal and ceramic materials. These wear mechanisms are encountered over a wide range of length scales. Depending on surface relative motion, these two types of wear can take place in either sliding or normal contact modes.

The microstructure of crystalline materials plays a significant role in nanoscale contact behavior. To obtain fundamental understanding of surface plasticity due to indentation loading, discrete dislocation plasticity is an effective analytical method. Unlike crystal plasticity, which is based on continuum mechanics and thus phenomenological constitutive relationships, discrete dislocation plasticity accounts for yielding and plastic flow from the crystallography standpoint and, hence, is a true mechanism-based method of studying plastic flow.

The main objective of this dissertation was to study contact nanomechanics through modeling and simulation. In Chapters 2 and 3, the FEM technique was used to analyze the mechanical response of sinusoidal periodic and patterned layered media, respectively. Although separate FEM meshes are needed for different body geometries, FEM modeling allows for irregular geometry, arbitrary loading and boundary conditions, and complex material constitutive relationships. In Chapters 4-8, analytical models of friction, wear, and surface plasticity are presented for randomly rough contacting surfaces. While this approach only allows for simple constitutive models and surface topographies to be analyzed, it provides solutions that directly reflect the relationship between interested quantities and other parameters. Hence, analytical methods show how properties and behaviors of interest depend on important parameters both qualitatively and quantitatively. The contents of Chapters
2-8 are summarized in the following paragraphs.

In Chapter 2, the indentation problem of a rigid plane in normal contact with a deformable semi-infinite solid with a sinusoidal surface profile was analyzed with the FEM technique. Results of dimensionless parameters are presented in terms of the degree of contact of relatively compliant and stiff half-space media exhibiting elastic-perfectly plastic behavior. FEM results are compared with analytical solutions derived from a contact model based on slip-line plasticity theory. It is shown that the hypothesis of a uniform surface rise is a reasonable assumption that can lead to good approximations of overall quantities, also providing an accurate description of changes in the surface profile for relatively large fractional contact areas. The results of this chapter provide insight into the evolution of surface and subsurface contact deformation and asperity interaction of contacting surfaces possessing periodic waviness and have direct implications in surface roughness variations encountered in metal-working processes.

The mechanical response of alternating phase-shift mask (APSM) microstructures subjected to dynamic pressure loadings relevant to those observed under megasonic cleaning conditions, a process widely used in photomask cleaning, was analyzed with FEM technique in Chapter 3. A parametric study of the effects of microstructure dimensions, pressure amplitude, and loading frequency on the mask structural integrity was performed for two typical chromium/quartz APSM patterns. Failure due to microfracture and plastic deformation processes that may occur during megasonic cleaning was examined for different loading frequencies. FEM results provide insight into possible failure modes and critical microstructure dimensions for instantaneous microstructure damage. Different failure scenarios revealed by the obtained FEM results are in qualitative agreement with experimental observations. The results of this chapter have direct implications to the design of extreme ultraviolet lithography masks and the optimization of the megasonic cleaning process.

When two surfaces are in sliding contact, various friction and wear mechanisms may arise at different length scales. A plasticity analysis of sliding for a rough (fractal) surface sliding against a smooth surface is presented in Chapter 4. The analysis is based on a slip-line model of a rigid spherical asperity (wear particle) plowing and cutting through a soft semi-infinite medium. Solutions for the fraction of fully plastic asperity microcontacts responsible for the evolution of friction, coefficient of friction, and energy dissipated during sliding contact are interpreted in terms of the total normal load (global interference), interfacial adhesion characteristics, topography (fractal) parameters, and elastic-plastic material properties by incorporating the slip-line model of a single microcontact into a fractal contact analysis of sliding friction of rough surfaces. Numerical results provide insight into the effects of global interference, surface roughness, interfacial adhesion, and material
properties on plastic deformation at the asperity contact level, friction coefficient, and frictional energy dissipated during sliding.

A generalized abrasive wear analysis of three-dimensional rough (fractal) surfaces sliding against relatively softer surfaces is presented in Chapter 5. This analysis is based on the same slip-line model of a rigid spherical asperity (or wear particle) presented in Chapter 4. This asperity-based slip-line model is integrated into a fractal mechanics analysis to obtain relationships of the abrasive wear rate and wear coefficient in terms of the interfacial shear strength, topography (fractal) parameters, elastic-plastic material properties, and total normal load. Analytical results from the slip-line analysis provide insight into the deformation behavior due to a single rigid asperity contact plowing through a rigid-perfectly plastic material under different normal loads and interfacial friction (adhesion) conditions. In addition, the effects of normal load (global interference), surface roughness, material properties, and interfacial shear strength (lubrication effect) on the abrasive wear coefficient (rate) of sliding fractal surfaces are interpreted in the context of numerical results.

In Chapter 6, a generalized adhesive wear analysis that takes into account the effect of adhesion force on the total contact load was developed for three-dimensional fractal surfaces in purely normal contact. A wear criterion based on the critical contact area of fully plastic asperity contacts is used to model the removal of material from the contact interface. The fraction of fully plastic asperity contacts and the wear rate and wear coefficient are expressed in terms of the total normal load (global interference), fractal (topography) parameters, elastic-plastic material properties, surface energy, material compatibility, and interfacial adhesion characteristics controlled by the environment of the interacting surfaces. Numerical results are presented for representative ceramic-ceramic, ceramic-metallic, and metal-metal contact systems to illustrate the dependence of asperity plastic deformation, wear rate, and wear coefficient on global interference, surface roughness, material properties, and work of adhesion (affected by the material compatibility and the environment of the contacting surfaces). The analysis presented in this chapter provides insight into the effects of surface material properties and interfacial adhesion on adhesive wear of rough surfaces in normal contact.

In the previous chapters, the analyzed media are all isotropic and homogeneous (except for the two-layered media studied in Chapter 3), and continuum mechanics has been used to analyze contact behaviors and failure mechanisms. Considering the microstructure characteristics of crystalline media, elastic-plastic indentation of a single-crystal half-space either by a rigid smooth cylinder (Chapter 7), a rigid rough asperity (Chapter 8), or a rigid fractal surface (Chapter 8) are analyzed by discrete dislocation plasticity. Short-range dislocation interactions are modeled by a set of constitutive rules of dislocation emission, glide, pinning (by obstacles), and annihilation. For the indentation by a rigid smooth cylindrical asperity presented in
Chapter 7, the occurrence of the first dislocation dipole, multiplication of dislocations, and evolution of subsurface stresses are discussed in terms of contact load, dislocation source density, slip-plane distance and orientation angle, and indenter radius. In the presence of dislocation sources, the critical load for dislocation initiation (onset of slip) is less than that of an ideally homogeneous medium and depends on dislocation source density, slip-plane distance, and indenter radius. The critical indenter radius resulting in deformation under the theoretical material strength is determined from numerical results, and the role of dislocation obstacles is interpreted in terms of their spatial density. Simulations provide insight into yielding and plastic deformation of indented single-crystal materials, and establish a basis for developing coarse-grained plasticity models of localized contact deformation in polycrystalline solids.

Chapter 8 presents the first discrete dislocation plasticity study of crystalline materials indented by a rigid surface possessing multi-scale roughness. Dislocation multiplication and the development of subsurface shear stresses due to asperity microcontacts forming between a single-crystal half-space medium and a rough surface are interpreted in terms of surface roughness and topography (fractal) parameters, slip-plane direction and spacing, dislocation source density, and contact load (surface interference). The effect of multi-scale interactions between asperity microcontacts on plasticity is demonstrated by results showing the development of dislocation structures. Numerical solutions provide insight into plastic flow of crystalline materials in normal contact with surfaces exhibiting multi-scale roughness.

Finally, Chapter 9 gives a summary of the most important findings reported in Chapters 2-8 and the main conclusions derived from the studies included in this dissertation.
Chapter 2

Finite element analysis of an elastic-perfectly plastic half-space

with a sinusoidal surface profile compressed by a rigid plane

2.1. Introduction

When two rough surfaces are brought into contact, with certain combinations of the roughness and elastic properties, the contacting surfaces can be flattened completely by elastic deformation in either two-dimensional (plane strain, Westergaard (1939)) or three-dimensional (Johnson et al., 1985) configurations. In the case that the rough surfaces cannot be completely flattened by elastic deformation, plastic deformation under higher loads tends to cause flattening of the surfaces to some extent, depending on various conditions. Pullen and Williamson (1972) applied very high loads to study the plastic deformation of a solid surface with the bulk material constrained in a relatively rigid container and observed the persistence of asperities, i.e., resistance to being flattened, and a nearly uniform rise of the material in the noncontacting regions after deformation. The persistence of asperities was attributed to the interaction between neighboring asperities (Childs, 1977).

In some metal working processes such as rolling, where the normal load is sufficiently high, the bulk material of a workpiece is plastically deformed during its contact with an "indenter" in a broad sense, such as a roll. In the case of rolling, the bulk deformation of the workpiece can be analyzed by using the slip-line field theory (Firbank and Lancaster, 1965) or the finite element method (Rao and Lee, 1989). Accompanying the bulk deformation, the surface roughness of the workpiece changes as it moves toward the exit during the entire contact process. Such roughness changes have been observed experimentally in rolling of steel by Atala and Rowe (1975).

During a contact process, the workpiece undergoes plastic deformation caused by the high contact pressure. However, the indenter and workpiece surfaces do not conform to each other completely after the plastic deformation (Pullen and Williamson, 1972; Williamson and Hunt, 1972; Wanheim, 1973). To better understand the qualitative and quantitative aspects of surface topography variations in plastic deformation, a systematic predictive approach to the roughness changes is desirable. An attempt was made by Bay and Wanheim (1976) to determine the center-line-average roughness of a deformed workpiece by using a crude approximation based on the assumption of serrated (saw-tooth-like) surfaces of the die or indenter. However, the more representative case of a sinusoidal profile for characterizing single-level roughness has not been studied under similar conditions,
and further confirmation for the plastic deformation analysis awaits more efforts.

The objectives of the present research are, therefore, to use the finite element method (FEM) to study the elastic-plastic contact between a rigid solid (indenter) and a deformable solid (workpiece) possessing a sinusoidal surface profile and to predict the surface texture of the workpiece after the plastic deformation. The obtained FEM results are compared with the analytical results based on an upper-bound analysis (Wang et al., unpublished work) and are used to verify the experimental observation of the uniform rise of the surface material by Pullen and Williamson (1972). Numerical results are presented to illustrate the effects of asperity interaction and subsurface deformation on the evolution of the real contact area, normal approach, average surface rise, center-line-average and peak-to-valley roughnesses, and cavity volume in terms of the global interference (normal load). The present study is useful in predicting the final roughness of a deformed workpiece.

2.2. Finite element model of a contacting sinusoidal surface

Because of the surface roughness, actual contact occurs only over an area \( A_r \), known as the real area of contact, which is smaller than the apparent or nominal contact area \( A_a \). The real-to-apparent contact ratio is termed a fractional contact area \( \gamma \), given by \( \gamma = A_r / A_a \). By definition, the possible range of \( \gamma \) is \( 0 \leq \gamma \leq 1 \).

When a normal load \( W \) is applied to the contact, the apparent pressure \( q \) is given by \( q = W / A_a \) and the dimensionless apparent pressure \( Q \) is given by

\[
Q = \frac{q}{2k}
\]

(2.1)

where \( k \) is the yield stress in pure shear. According to the Tresca yield criterion, \( k = \sigma_y / 2 \), where \( \sigma_y \) is the tensile yield strength of the softer solid (workpiece in this case). Thus, \( Q = q / \sigma_y \). The real pressure at the contact interface is given by \( p_r = W / A_r \). It follows that

\[
p_r = q / \gamma
\]

(2.2)

A two-dimensional (plane-strain) contact problem of a rigid flat surface (indenter) compressing a half-space with a sinusoidal surface profile (workpiece) is shown schematically in Fig. 2.1. The rough surface profile has an amplitude \( h/2 \) and a wavelength \( \lambda \). The contact between the rigid plane and the deformable medium was
assumed to be frictionless. The rigid plane is displaced perpendicular to the center-line of the rough surface profile by a finite displacement, known as an absolute normal.

\[ \delta_0 \]

Since \( \delta_0 \) contains the bulk elastic deformation, \( \delta_b \), represented by a downward displacement of a reference point \( R \) (Fig. 2.1), to compare the finite element analysis results with the analytical results, this bulk deformation needs to be subtracted from \( \delta_0 \) to give a relative normal approach \( \delta \), i.e.,

\[ \delta = \delta_0 - \delta_b \] (2.3)

The reference point, \( R \), was chosen to be located at a distance of \( 0.75\lambda \) below the valley level of the rough surface in its undeformed state, as shown in Fig. 2.1. The choice of such a distance is to allow the region between the valleys of the rough surface and the level of point \( R \) to completely contain the zone with major plastic deformation, but not to be too close to the boundary of the plastic zone. For typical roughness or waviness, the wavelength is much larger than the peak-to-valley height variation; thus, all simulations were performed for \( h/\lambda = 0.05 \).
In the finite element analysis, the average rise, $\delta_r$, was calculated as the difference between the bulk deformation and the average absolute displacement, $\delta_{ave}$, of the lower surface. As downward displacements are defined as positive, $\delta_b > 0$, $\delta_5 > 0$, $\delta > 0$, and $\delta_{ave} > 0$. The average absolute displacement, $\delta_{ave}$, is the average of the vertical displacements of chosen surface points, i.e., $\delta_{ave} = \frac{1}{n} \sum_{i=1}^{n} \delta_i$, where $n$ is the number of noncontacting surface points (nodes) and $\delta_i$ (downward as positive) is the absolute displacement of the $i$th surface point. Then, the average rise $\delta_r$ is given by

$$\delta_r = \delta_b - \delta_{ave} \tag{2.4}$$

A positive value of $\delta_r$ indicates a rise of the surface of the simulated half-space medium relative to the reference point $R$.

The dimensionless forms of $\delta$ and $\delta_r$ are defined as

$$\Delta = \frac{\delta}{h/2} \tag{2.5}$$

$$\Delta_r = \frac{\delta_r}{h/2} \tag{2.6}$$

Figure 2.2 shows the finite element mesh of the deformable half-space. Due to periodicity and symmetry considerations, only a segment of the structure, referred to as a cell, had to be modeled. The mesh consists of 4905 four-node, quadrilateral, plain-strain elements with a total of 5290 nodes. A two-by-two Gaussian integration scheme was used in the analysis. The mesh close to the sinusoidal surface was refined to enhance the accuracy and convergence. There are 400 nodes within a half wavelength of the sinusoidal profile. Due to symmetry, the nodes of boundaries $AB$ and $CD$ were constrained against the displacement in the $x$-direction, whereas the nodes of boundary $BC$ (model bottom) were constrained against displacement in the $y$-direction.

In the finite element analysis, the half-space medium with a sinusoidal surface profile was modeled as an elastic-nearly-perfectly plastic material, with an artificially set strain hardening coefficient (exponent of plastic strain) of a very small value, 0.005, to ensure convergence. To examine the effect of bulk elastic deformation on the development of plasticity at the sinusoidal surface, simulations were performed for relatively compliant and stiff half-space media with elastic modulus $E_1$ and $E_2$. 

9
respectively, with $E_2 = 100 \ E_1$. The same values of Poisson’s ratio (0.3) and yield strength $\sigma_y$ (such that $E_1/\sigma_y = 1000$) were assumed for the two materials. The incompressibility condition was imposed during plastic deformation. Quasi-static finite element simulations were performed with the multi-purpose finite element code ABAQUS/Standard. The finite element results and the analytical results from the upper-bound theoretical model (Wang et al., unpublished work), both in dimensionless form, are compared in this section for the same contact process of increasing the fractional contact area $\gamma$, or equivalently, increasing the normal load.

2.3. Results and discussion

FEM and analytical results, presented in the form of dimensionless parameters, are compared in this section with the increase of fractional contact area (normal load).

2.3.1. Apparent contact pressure

Both finite element and analytical results demonstrate a linear increase in the dimensionless apparent pressure $Q$ with increasing fractional contact area $\gamma$ up to some critical value, above which the slope of the curve $dQ/d\gamma$ increases, as shown in Fig. 2.3. This linear behavior is due to independent plastic deformation of individual
asperities, with a dimensionless slope in reasonable agreement with the slip-line field solution for the plastic compression of a single wedge (Hill et al., 1947). Because elastic behavior is considered in the finite element model besides plasticity, the simulation results include a state of nearly full contact ($\gamma \approx 1$) at $Q \approx 4.7$, or $p_r \approx 4.7\sigma$. This contact pressure is much higher than that needed for deforming an isolated asperity ($\approx 3\sigma$), which is in agreement with the results of Gao et al. (2006).

The incomplete contact under relatively high loads in the analytical results reveals the phenomenon of persistence of asperities (Williamson and Hunt, 1972; Childs, 1977). The increase in contact pressure above that for compressing a single

![Fig. 2.3 Analytical and finite element simulation results for the dimensionless apparent contact pressure $Q$ versus the fractional contact area $\gamma$ during loading for relatively compliant ($E_1$) and stiff ($E_2 = 100E_1$) half-spaces.](image)

wedge is due to the interaction between neighboring asperities (Pullen and Williamson, 1972), as reflected by the overlap of their plastic zones shown in Fig. 2.4. The evolution of the plastic zone in the half-space can be observed as the fractional contact area (or load) increases. Plastic zone interaction becomes significant when $\gamma \approx 0.7$ (Fig. 2.4(d)), in agreement with the critical value determined from Fig. 2.3. The increase in the slope may also be associated with the transition from a surface profile containing large cavities to a profile with small cavities between plastically
compressed asperities. Because of the constraints imposed by the overlapping plastic zones under neighboring asperity contacts, horizontal deformation is restricted and the only feasible deformation is elastic compression.

Fig. 2.4 Expansion of the plastic zone with increasing fractional contact area $\gamma$ for relatively compliant half-space ($E_i/\sigma_y = 1000$): (a) $\gamma = 0.1$, (b) $\gamma = 0.3$, (c) $\gamma = 0.5$, and (d) $\gamma = 0.7$.

The finite element results show a larger critical fractional contact area (0.875) than the analytical results do (0.80). The difference in the apparent contact pressure determined from the two approaches increases with an increase of the applied load (or fractional contact area). The lower contact pressure predicted by the finite element
analysis (Fig. 2.3) can be attributed to bulk elastic deformation, which is not taken into account in the analytical model. However, the finite element results for a two orders of magnitude stiffer material ($E_2 = 100E_1$) do not differ from those obtained with the relatively compliant half-space ($E_1$), presumably because the elastic modulus of the stiffer material is not sufficiently high to approximate rigid-plastic behavior.

2.3.2. Normal approach

Figure 2.5 shows an increase in the dimensionless normal approach, $\Delta$, with the fractional contact area (or normal load). The differences between the numerical and analytical results are attributed to the contribution of elastic deformation below contacting asperities in the finite element model. In the low-load range ($\gamma < 0.3$), the finite element results are close to the analytical solution because the effect of elastic deformation in the subsurface is secondary. Although the major elastic deformation effects reflected by the displacement of the reference point $R$ (Fig. 2.1) has already been subtracted, the elastic deformation of the lower material medium in the regions above the reference point can still affect the finite element results of the surface displacements. In the intermediate load range ($0.3 < \gamma < 0.6$), the normal approach predicted by the finite element model is less than that of the analytical model because of the significant effect of bulk elastic deformation in the finite element simulations. In the high-load range ($\gamma > 0.7$), fully plastic deformation together with elastic

![Fig. 2.5 Analytical and finite element simulation results for the dimensionless normal approach $\Delta$ versus the fractional contact area $\gamma$ during loading for relatively compliant ($E_1$) and stiff ($E_2 = 100E_1$) half-spaces.](image-url)
compression occurs below the contacting asperities and the two surfaces show almost complete conformity. In this situation, the finite element results for the stiffer medium \((E_2 = 100E_1)\) converge toward the analytical solution.

### 2.3.3. Surface roughness

The center-line-average roughness of the deformed surface profile is \(R'_a\), with its dimensionless form \(R_a\) defined as

\[
R_a = \frac{R'_a}{\bar{a}R_{a0}}
\]  

(2.7)

where \(R'_a\) refers to the center-line-average roughness of the initial undeformed profile.

During loading, the variation of \(R_a\) is shown as a function of the fractional contact area \(\gamma\) in Fig. 2.6(a). It is seen that \(R_a\) decreases with the increase of the fractional contact area. Three deformation regimes can be observed on this plot. For \(\gamma < 0.2\), the finite element results are close to the analytical solution because surface (asperity) plasticity is the main deformation mode. In the range \(0.2 < \gamma < 0.8\), the finite element model yields higher roughness values than the analytical model due to the significant effect of elastic compression below the asperity contacts. For \(\gamma > 0.8\), the difference between the finite element and analytical results decreases due to the decreased contribution of bulk elastic deformation in the total deformation accumulated in the vicinity of asperity contacts. A similar trend has been observed in comparing the finite element and analytical results of the dimensionless peak-to-valley roughness \(H\) (Fig. 2.6(b)), with a slightly greater difference in the middle range of \(\gamma\) and a much closer agreement for \(\gamma < 0.25\).

### 2.3.4. Deformed surface shape and average surface rise

The surface profiles at different values of the normal approach (global interference) \(\delta\) for the loading and full unloading conditions are shown in Figs. 2.7(a) and 2.7(b), respectively. For low surface interference \((\delta/h < 0.35)\), the tendency is for the material to expand in the horizontal direction, decreasing the free space or cavity between the contacting asperities. However, for relatively large surface interference \((\delta/h > 0.35)\), the bottom of the valley rises, and the rises of the surface nodes gradually become more uniform with increasing normal approach (interference). The difference in the rise between valley nodes and higher nodes decreases due to the intensifying effect of bulk elastic-plastic deformation with increasing interference. Comparing Fig. 2.7(b) to Fig. 2.7(a) reveals that each surface node of the unloaded profile has moved upward by the same distance and that the shape of the loaded profile has been maintained. This suggests a uniform elastic recovery of the surface.
Fig. 2.6 Analytical and finite element simulation results for (a) the dimensionless center-line-average roughness $R_a$, and (b) the dimensionless peak-to-valley roughness $H$ versus the fractional contact area $\gamma$ for relatively compliant ($E_1$) and stiff ($E_2 = 100E_1$) half-spaces.
Fig. 2.7 Evolution of the surface profile of the relatively compliant elastic-perfectly plastic half-space \((E_1/\sigma_y = 1000)\) with the dimensionless global interference: (a) during loading, and (b) after full unloading.
can also be seen that the increase in the maximum surface interference enhances the elastic recovery.

The variation of the normalized average surface rise after full unloading $\bar{\Delta}$, defined in Eq. (2.6) with $\delta$, calculated by using Eq. (2.4) for noncontacting nodes, is shown in Fig. 2.8 as a function of the fractional contact area $\gamma$. Because the average surface rise was calculated as the mean value of the vertical displacements of all surface nodes that did not contact the rigid plane at the maximum global interference, the number of averaged nodes decreased with the increase of the load. Both finite element and analytical results show an increase in average surface rise with fractional contact area. The less surface rise predicted by the finite element model is attributed to elastic deformation in the subsurface region.

2.3.5. Cavity volume

The cavity volume, $V'$, is the volume of the space per unit groove length between the two surfaces after deformation. Its dimensionless form $V$ is defined as

$$V = \frac{V'}{h \cdot \frac{\lambda}{2}}$$

It is observed that $V$ decreases with increasing fractional contact area (or load), as shown in Fig. 2.9. Three deformation regimes similar to those observed for roughness in Fig. 2.6(a) can be distinguished in Fig. 2.9. In the intermediate range $0.2 < \gamma < 0.8$, the relative deviation of the finite element results from the analytical ones is less than that in the center-line-average roughness shown in Fig. 2.6(a). This is because the expansion of the contact area (giving a greater $\gamma$) causes a reduction of the cavity volume $V$ (Fig. 2.7(a)) in the elastic-plastic finite element analysis. For a given normal approach, the simultaneous changes in $\gamma$ and $V$ merely shift the finite element data point to a location closer to an analytical data point with a slightly larger $\gamma$, thus maintaining the closeness of the two data sets.

2.3.6. Ratio of truncated contact area to real contact area

Figure 2.10 shows a comparison between finite element results of the truncated contact area to the real contact area ratio $a'/a$ versus the fractional contact area for relatively compliant ($E_1$) and stiff ($E_2 = 100 E_1$) half-spaces. Referring to the curve for $E_1$, elastic deformation is dominant when $\gamma < 0.02$ and $a'/a \approx 2.2$, similar to the compression of a hemispherical asperity in three-dimensions, for which $a'/a = 2$ according to the Hertz theory. For very small $\gamma$ values, the contact portion of the sinusoidal contacts can be accurately approximated by cylindrical asperities (Gao et al., 2006). However, for $\gamma > 0.2$ the results for sinusoidal asperities deviate from those
Fig. 2.8 Analytical and finite element simulation results for the normalized average surface rise $\bar{\Delta}_r$ versus the fractional contact area $\gamma$ after full unloading for relatively compliant ($E_1$) and stiff ($E_2 = 100 E_1$) half-spaces.

for cylindrical ones due to the increasing effect of the non-uniform curvature of the sinusoidal profile. As the load increases, the fractional contact area $\gamma$ increases and plastic deformation sets in to expand the asperity in the horizontal direction such that $a'/a$ decreases drastically in a range of small $\gamma$.

In the case of the stiffer material with $E_2$, however, Fig. 2.10 shows that the stage of plastic deformation comes earlier after a brief drop of the value of $a'/a$, as the fractional contact area $\gamma$ increases. The truncated contact area is similar to the real contact area even for $\gamma < 0.02$ because fully plastic deformation is dominant throughout the range of fractional contact area. The higher $a'/a$ values for the relatively compliant half-space in the range of small to moderate values of $\gamma$ can thus be attributed to the stronger effects of bulk elastic deformation.
Fig. 2.9  Analytical and finite element simulation results for the normalized cavity volume $V$ versus fractional contact area $\gamma$ for relatively compliant ($E_1$) and stiff ($E_2 = 100 \ E_1$) half-spaces during loading.

Fig. 2.10 Ratio of truncated-to-real contact area (contact area ratio) $a'/a$ versus fractional contact area $\gamma$ for relatively compliant ($E_1$) and stiff ($E_2 = 100 \ E_1$) half-spaces during loading.
2.4. Conclusions

The contact problem of a rigid flat indenting an elastic-plastic half-space medium with a sinusoidal surface profile was examined in light of a two-dimensional plane-strain FEM analysis. Numerical results of the dimensionless apparent contact pressure, normal approach, average surface rise, center-line-average roughness, peak-to-valley roughness, cavity volume, and ratio of truncated-to-real contact area versus fractional contact area obtained for relatively compliant and stiff elastic-perfectly plastic half-spaces were compared with results obtained from a slip-line plasticity analysis. This analysis has direct applications to metal working processes such as rolling. Based on the presented results and discussion, the following main conclusions can be drawn from the current study.

1) The apparent pressure $Q$ increases linearly with the fractional contact area $\gamma$ in the range of $0 < \gamma < 1/2$, and nonlinearly with increased rates in the range of $1/2 \leq \gamma < 1$ due to neighboring asperity interaction, which constrains subsurface deformation in the vicinity of the contacting asperities.

2) The hypothesis of a uniform rise of the surface material during plastic deformation based on the experimental observation by Pullen and Williamson (1972) has been demonstrated to be a reasonable assumption that can lead to good approximations to the overall behaviors of the rough surface. However, only when the fractional contact area is relatively large can this assumption serve as an accurate description of the changes in the surface profile shape.

3) For a half-space with a sinusoidal surface compressed by a rigid plane, when the load increases, the fractional contact area and the normal approach increase. Consequently, the average surface rise becomes larger, accompanied by a reduction of the cavity volume. However, the surface roughness decreases. Differences between FEM and analytical results were explained by considering the effect of bulk elastic deformation to the surface deformation behavior.

4) The initial deformation stage of the sinusoidal surface profile can be approximated by that of a surface consisting of cylindrical asperity contacts. However, at progressive stages of deformation (large global interference), the results for cylindrical and sinusoidal asperities deviate because of the shape profile effect on surface deformation.
Chapter 3

Dynamic finite element analysis of failure in alternating phase-shift masks caused by megasonic cleaning

3.1. Introduction

The explosive growth of semiconductor devices has been largely due to advances in optical lithography. Extreme ultraviolet (EUV) lithography is the predominant fabrication process in chip-printing technology, used to produce patterns of nanometer-sized microstructures on silicon wafers (Stix, 2001). One of the main challenges in EUV lithography is the removal of contaminant nanoparticles larger than 30 nm from the photomask surface. Preserving the photomask high printing efficiency by preventing surface damage during cleaning (Muralidharan et al., 2007) is critical in EUV lithography.

One of the most successful techniques for removing nanoparticle contaminants from EUV photomasks is megasonic cleaning. In this process, shown schematically in Fig. 3.1, acoustic waves of frequencies on the order of 1 MHz propagate in the

![Fig. 3.1 Schematic of the megasonic cleaning process of EUV photomasks (not to scale). The pattern microstructures have been exaggerated for clarity. In reality, the fluid stream impinges onto a huge number (on the order of \(10^8\)) of microstructures.](image)
cleaning medium (Busnaina et al., 1995), resulting in several physical phenomena, such as acoustic cavitation and streaming (Kapila et al., 2005). Each point along a wave oscillates between a pressure maximum (compression) and a pressure minimum (rarefaction). Cavitation bubbles form at sites where the pressure drops below the vapor pressure of the fluid, approaching the pressure minimum. These cavitation bubbles subsequently collapse as those sites transition from the pressure minimum to the pressure maximum. Rapid flow of surrounding fluid to fill the void created by a collapsed bubble produces an intense shock wave. The cavitation energy transferred to the surface due to the collapse of many bubbles close to the substrate surface activates an intense scrubbing process, which is effective in removing contaminants from the surfaces of devices, such as memory disks, LCD devices, and semiconductor wafers.

Although megasonic cleaning is effective in removing surface contaminants, excessive cavitation energy may cause extensive surface damage (Kapila et al., 2006), particularly to tiny surface features such as microstructures on patterned photomask surfaces. Figure 3.2 shows an atomic force microscope (AFM) image of a patterned chromium-quartz photomask obtained after megasonic cleaning. Although the microstructure on the right does not show any discernible damage, the microstructure on the left has been extensively damaged for no apparent reason. This transient cavitation damage becomes more pronounced with the decrease of the microstructure size and the increase of the cavitation energy that depends on the frequency and
amplitude of the megasonic wave. Since the increase of the frequency opposes bubble growth, the smaller bubbles generated at higher frequencies decrease the cavitation energy. Alternatively, higher wave amplitudes cause each point along the wave to oscillate over a larger pressure range (between rarefaction and compression), producing larger cavitation bubbles which, in turn, increase the cavitation energy. Thus, there is a direct correlation between the intensity of the megasonic wave, the pressure range that the fluid medium oscillates, and the cavitation energy transferred to the surface.

Preventing damage of pattern microstructures on photomask surfaces without sacrificing the effectiveness of the cleaning process is of paramount importance. The main objective of this study was to examine the structural integrity of alternating phase-shift mask (APSM) microstructures subjected to cyclic (dynamic) pressure loadings relevant to megasonic cleaning. To accomplish this objective, patterned APSMs under different dynamic pressure loadings were analyzed with the finite element method (FEM). Possible failure modes and damage mechanisms due to megasonic cleaning of chromium-quartz APSMs are interpreted in light of FEM results obtained for different patterns, dynamic pressure loadings, microstructure dimensions, and residual stresses. In addition, the critical microstructure dimensions resulting in plasticity-induced failure or instantaneous collapse of two typical APSM patterns are extracted from dynamic simulation results.

3.2. Finite element modeling

As mentioned previously, the main theme in this study was to examine the mechanical response of APSMs due to dynamic loadings relevant to those of megasonic cleaning, a process critical to EUV lithography (Chakravorty et al., 2006). The APSM architecture is based on a dual sided trenched configuration. Figure 3.3 shows cross-section schematics of two typical APSM designs with and without an undercut, hereafter referred to as patterns A and B, respectively. Under the conditions of the megasonic cleaning process, the predicted average size of the collapsing bubbles (Brennen, 1995) is much larger than the microstructure cross-section dimensions. Shock waves generated from collapsing bubbles in the proximity of the mask surface impact the pattern microstructures (Fig. 3.4). In view of the large width of the wave front and the very small cross-section dimensions of the microstructures, the shock waves apply predominantly compressive loads perpendicular to the mask surface. Thus, a uniform pressure was applied at the top of each pattern microstructure analyzed in the present study.

Structural failures due to microstructure collapse and/or cumulative fatigue were considered as possible failure mechanisms. Therefore, three types of pressure functions, shown in Fig. 3.5, were used in the FEM analysis. To study instantaneous failure
Fig. 3.3 Cross-section schematics of typical chromium-quartz microstructures under a uniform pressure $p$ ($a =$ undercut depth, $d =$ zero-aperture depth, $b =$ line width, $t =$ chromium layer thickness, and $l =$ distance between zero-aperture and $\pi$-aperture): (a) pattern A and (b) pattern B. The region enclosed by a thick (blue) line is the cell analyzed by the FEM.

(collapse), the dynamic pressure was applied as a step function (Fig. 3.5(a)). The collapse pressure was determined as the critical pressure at which deformation continuity was not satisfied in the analysis. This was accomplished by increasing the step dynamic pressure in an incremental fashion up to a critical value which resulted in deformation discontinuities that prevented convergence. This critical pressure is referred to as the collapse pressure. For cumulative fatigue failure, the pressure in the
dynamic FEM analysis was represented by a periodic function (Fig. 3.5(b)) and the total time was fixed at 1 µs. By varying the loading period $T$, dynamic FEM simulations were performed for megasonic frequency $f = 1, 5,$ and 10 MHz. A static FEM analysis was also performed for comparison (Fig. 3.5(c)).

The microstructure dimensions used in the parametric FEM study are given in Table 3.1. Only one dimension was varied in each series of simulations. To ensure 180° phase shift, a fixed distance between zero-aperture and π-aperture was used in all simulations (i.e., $l = 170$ nm). Because the cross-section dimensions are significantly smaller than the out-of-plane microstructure dimension and the relatively large average size of collapsing bubbles (~1 µm), two-dimensional FEM models were used to analyze microstructure failure. Figure 3.6 shows the FEM meshes of APSM patterns A and B. From symmetry and periodicity considerations, only a portion of each

![Collapsing bubble](image)

**Fig. 3.4** Schematic of cavitation bubble collapse in the proximity of the photomask surface.

<table>
<thead>
<tr>
<th>Pattern*</th>
<th>Parameter</th>
<th>Magnitude (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$a$</td>
<td>0 10 20 30 40 50</td>
</tr>
<tr>
<td></td>
<td>$d$</td>
<td>40 80 120 160</td>
</tr>
<tr>
<td></td>
<td>$t$</td>
<td>25 50 75 100</td>
</tr>
<tr>
<td></td>
<td>$b$</td>
<td>100 120 140 160 180 200</td>
</tr>
<tr>
<td>B</td>
<td>$t$</td>
<td>25 50 75 100</td>
</tr>
<tr>
<td></td>
<td>$b$</td>
<td>50 70 90 110 130 150</td>
</tr>
</tbody>
</table>

* $l = 170$ nm in both patterns A and B.
Fig. 3.5 Schematics of pressure loading functions used in the FEM analysis: (a) step dynamic pressure $p_c$ applied at time $t_0$, (b) cyclic dynamic pressure ($T =$ loading period, $\tau =$ time under maximum pressure $p_{\text{max}}$ during a loading period), and (c) static pressure ($T =$ loading period, $p_{\text{max}} =$ maximum pressure).
cross-section pattern (referred to as cell) was modeled. Both meshes consist of four-node, quadrilateral, plain-strain elements. A reduced Gaussian integration scheme was used throughout the analysis. Mesh refinement at sharp corners was used to enhance the convergence and accuracy. The nodes at boundaries AB and CD were constrained against displacement in the horizontal direction, while those at boundary BC were constrained against displacement in the vertical direction. The FEM meshes shown in Fig. 3.6 are sufficiently refined to produce convergent results. This was

![FEM models of cell microstructures](image)

Fig. 3.6 FEM models of cell microstructures in (a) pattern A and (b) pattern B under a uniform normal pressure $p$. (The red nodes were used to calculate the residual stress in the chromium layer).
verified by examining the effect of the smallest element size on convergence. A representative result from the mesh refinement study is shown in Fig. 3.7. The maximum first principal stress $\sigma_{i}^{\text{max}}$ in the microstructures of patterns A and B due to a cyclic dynamic pressure of 600 MPa amplitude and 1 MHz frequency (Fig. 3.5(b)) is shown to converge at a minimum element size of $h_{\text{min}} = 2.5 \text{ nm}$, which is the size of the smallest elements used in the FEM meshes shown in Fig. 3.6. This size of the smallest elements is above the threshold for continuum description.

In the FEM analysis, the chromium layer and the quartz substrate were modeled as homogeneous isotropic materials with properties given in Table 3.2. An elastic-perfectly plastic constitutive model was used for both materials. This model satisfies $J_2$ plastic flow theory based on the multiplicative decomposition of the deformation gradient into elastic and plastic components. Volume change due to plastic deformation was assumed to be negligible. Standard elastic constitutive models were used when $\sigma_{eq} < S_Y$ (where $\sigma_{eq}$ is the von Mises equivalent stress and $S_Y$ is the yield strength), while $J_2$ flow theory was used at yield material points that satisfied the von Mises yield criterion ($\sigma_{eq} = S_Y$). All of the simulations were performed with the multi-purpose FEM code ABAQUS/Standard. In view of geometric nonlinearities (large displacements) and nonlinear material behavior (plasticity), an updated Lagrangian formulation was used in all simulations.

### 3.3. Results and discussion

The von Mises yield criterion is commonly used to interpret material failure. However, this criterion does not differentiate between failures due to predominantly tensile and compressive stresses. Since the chromium layer is stronger (harder) and stiffer than the quartz substrate (Table 3.2), the microstructure resistance to plastic

---

**Table 3.2  Material properties used in the FEM analysis.**

<table>
<thead>
<tr>
<th>Material</th>
<th>Elastic modulus $E$ (GPa)</th>
<th>Yield strength $S_Y$ (GPa)</th>
<th>Poisson ratio $\nu$</th>
<th>Mass density $\rho$ (kg/m$^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromium</td>
<td>105.0</td>
<td>4.0</td>
<td>0.30</td>
<td>6900</td>
</tr>
<tr>
<td>Quartz</td>
<td>54.2</td>
<td>2.6</td>
<td>0.25</td>
<td>2600</td>
</tr>
</tbody>
</table>

*Source: Komvopoulos and Jee, unpublished work*
deformation due to tension is less than that due to compression. Therefore, microfracture due to high tensile stresses is a major failure mechanism of mask patterns subjected to megasonic cleaning. In the present FEM analysis, the propensity for microfracture is interpreted in terms of the maximum tensile (first principal) stress 

![Image](image.png)

Fig. 3.7 Maximum first principal stress $\sigma_{1}^{\text{max}}$ in the cell microstructure of (a) pattern A ($t = 100 \text{ nm}$, $a = 40 \text{ nm}$, $d = 40 \text{ nm}$, $b = 160 \text{ nm}$) and (b) pattern B ($t = 100 \text{ nm}$, $b = 150 \text{ nm}$) due to dynamic pressure loading ($f = 1 \text{ MHz}$) as a function of the size of the smallest finite element $h_{\text{min}}$.

because the susceptibility to microcracking of relatively hard coatings is strongly dependent on the magnitude of the maximum tensile stress. Another possible failure mechanism of the mask patterns is cumulative plasticity. Thus, the development of plastic deformation is discussed in the context of the evolution of the equivalent plastic strain.

The dependence of the collapse pressure $p_c$ on the microstructure dimensions of the masks with patterns A and B is shown in Figs. 3.8 and 3.9, respectively. For pattern A, $p_c$ decreases with the increase of the undercut depth (Fig. 3.8(a)) and the zero-aperture depth (Fig. 3.8(b)), while it increases with the increase of the chromium layer thickness (Fig. 3.8(c)) and the line width (Fig. 3.8(d)). In the case of pattern B, $p_c$ is practically independent of the chromium layer thickness (Fig. 3.9(a)) and the line width (Fig. 3.9(b)). In addition, pattern B demonstrates a higher load carrying capacity.
Fig. 3.8 Collapse pressure $p_c$ and equivalent plastic strain $\varepsilon_p$ in the cell microstructure of pattern A versus (a) undercut depth $a$ ($d = 80$ nm, $t = 100$ nm, $b = 200$ nm), (b) zero-aperture depth $d$ ($a = 40$ nm, $t = 100$ nm, $b = 160$ nm), (c) chromium layer thickness $t$ ($a = 50$ nm, $d = 80$ nm, $b = 200$ nm) and (d) line width $b$ ($a = 30$ nm, $d = 80$ nm, $t = 100$ nm).

(higher $p_c$) than pattern A. The results shown in Figs. 3.8 and 3.9 can be explained by considering the effect of dimensional variations on the dominant mode of deformation. Since the deformation of the microstructure of pattern A is predominantly due to bending, variations of the cross-section dimensions affected significantly the bending stiffness and, in turn, the microstructure deformation. Alternatively, because an undercut or zero-aperture depth does not exist in the microstructure of pattern B, the variations in the chromium layer thickness and the line width did not exhibit a first-order effect on bending deformation, whereas the effect of these variations on the axial deformation was secondary. Contours of the equivalent plastic strain $\varepsilon_p$ at the instant of microstructure collapse are also included in Figs. 3.8 and 3.9. Most simulation cases indicate that shear band formation in the substrate is a possible failure mode. Shear band formation originated from the plastic zones formed at the edges of the quartz cavities that acted as stress raisers. The plastic zones expanded toward each other, eventually merging to form a shear band, as shown in Figs. 3.8(a), 3.8(d), 3.9(a), and 3.9(b). This finding suggests that failure caused by the collapse pressure may be due to plastic deformation in the quartz substrate, resulting in the removal of the chromium layer and the upper region of the quartz substrate, in qualitative agreement with the damage shown in Fig. 3.2. In addition to failure due to shear band formation, the $\varepsilon_p$ contours reveal the existence of two other possible failure modes. As shown in
Fig. 3.9 Collapse pressure $p_c$ and equivalent plastic strain $\bar{\varepsilon}_p$ in the cell microstructure of pattern B versus (a) chromium layer thickness $t$ ($b = 130$ nm) and (b) line width $b$ ($t = 100$ nm).

Fig. 3.8(b), a high zero-aperture depth resulted in excessive bending (buckling), leading to the abrupt failure of the long and slim arm of the microstructure. Alternatively, Fig. 3.8(c) shows that a microstructure with a thin chromium layer may fail due to peeling off of the layer before the occurrence of failure due to shear band formation in the substrate.
Simulation results for cyclic dynamic pressure loading (Fig. 3.5(b)) and static pressure loading (Fig. 3.5(c)) are compared next. Both dynamic and static simulations were performed for $p_{\text{max}} = 600 \text{ MPa}$. This is a reasonable value considering that the maximum pressure at the center of a collapsing bubble is predicted to be $\sim 1 \text{ GPa}$ (Benjamin and Ellis, 1966). Microcracking and plastic flow can be interpreted in terms of $\sigma_{\text{i}}^{\text{max}}$ and the equivalent plastic strain $\bar{e}_p$. Figures 3.10 and 3.11 show the dependence of $\sigma_{\text{i}}^{\text{max}}$ on the microstructure dimensions and frequency of pressure.

Fig. 3.10 Maximum tensile stress $\sigma_{\text{i}}^{\text{max}}$ in the cell microstructure of pattern A due to static and dynamic pressure loading ($f = 1, 5, \text{ and } 10 \text{ MHz}$) versus (a) undercut depth $a$ ($d = 80 \text{ nm}$, $t = 100 \text{ nm}$, $b = 200 \text{ nm}$), (b) zero-aperture depth $d$ ($a = 40 \text{ nm}$, $t = 100 \text{ nm}$, $b = 160 \text{ nm}$), (c) chromium layer thickness $t$ ($a = 50 \text{ nm}$, $d = 80 \text{ nm}$, $b = 200 \text{ nm}$) and (d) line width $b$ ($a = 30 \text{ nm}$, $d = 80 \text{ nm}$, $t = 100 \text{ nm}$).
loading. For pattern A, $\sigma_i^{\text{max}}$ increases monotonically with the increase of the undercut depth (Fig. 3.10(a)) and the zero-aperture depth (Fig. 3.10(b)), is independent of the chromium layer thickness when the layer is thicker than 50 nm (Fig. 3.10(c)), and decreases with the decrease of the line width (Fig. 3.10(d)). The microstructures with the thinnest chromium layer (Fig. 3.10(c)) and the smallest line width (Fig. 3.10(d)) exhibit a high likelihood for collapse under dynamic pressure loading. However, in the case of pattern B, $\sigma_i^{\text{max}}$ is independent of the chromium layer thickness (Fig. 3.11(a)) and almost invariant of the line width for static and low-frequency dynamic loading,

![Fig. 3.11 Maximum tensile stress $\sigma_i^{\text{max}}$ in the cell microstructure of pattern B due to static and dynamic pressure loading ($f = 1$, 5, and 10 MHz) versus (a) chromium layer thickness $t$ ($b = 130$ nm) and (b) line width $b$ ($t = 100$ nm).](image)

while in the case of high-frequency dynamic loading it decreases with the increase of the line width (Fig. 3.11(b)). This trend may be attributed to interaction effects of waves reflected from the symmetric boundaries, which tend to intensify with the decrease of the line width and the increase of the loading frequency. In addition, for a fixed total cleaning time, all of the simulated cases show that $\sigma_i^{\text{max}}$ due to dynamic loading is much higher than that due to static loading. For $f = 10$ MHz, Figs. 3.10(a), 3.10(b), and 3.10(d) show that $\sigma_i^{\text{max}}$ is about an order of magnitude higher than that for static loading. This difference is due to the contribution of inertia effects in dynamic loading. Moreover, the increase of $\sigma_i^{\text{max}}$ with the loading frequency indicates a higher
probability of cavitation damage, attributed to the increase of the loading cycles during a certain simulation time with the increase of the loading frequency.

The propensity for mask cracking due to dynamic pressure loading at $f = 10$ MHz can be further interpreted in the light of $\sigma_I$ distributions in the microstructures of patterns A and B shown in Fig. 3.12. In the case of pattern A, $\sigma_I$ exhibits a discontinuity across the layer/substrate interface (Fig. 3.12(a)). The region in the chromium layer adjacent to the interface is in tension, while that in the quartz substrate is in compression. This is due to differences in the elastic-plastic mechanical properties of the layer and the substrate materials. The $\sigma_I^\text{max}$ stress arises at the inner corner of the shallow quartz cavity (stress concentration effect). However, in the case of pattern B, most areas around the interface are in tension (Fig. 3.12(b)), and $\sigma_I^\text{max}$ occurs at the stress concentration corner of the interface. Thus, it may be inferred that crack initiation is likely to occur at the layer/substrate interface. Therefore, there is a higher likelihood
for failure due to interfacial cracking than failure due to shear band formation in the substrate.

Fig. 3.13 Maximum equivalent plastic strain $\varepsilon_{pl}^{\max}$ in the cell microstructure of pattern A due to static and dynamic pressure loading ($f = 1, 5,$ and 10 MHz) versus (a) undercut depth $a (d = 80$ nm, $t = 100$ nm, $b = 200$ nm), (b) zero-aperture depth $d (a = 40$ nm, $t = 100$ nm, $b = 160$ nm), (c) chromium layer thickness $t (a = 50$ nm, $d = 80$ nm, $b = 200$ nm) and (d) line width $b (a = 30$ nm, $d = 80$ nm, $t = 100$ nm). The vertical dashed lines indicate the critical microstructure dimensions.
The effects of the microstructure dimensions and the loading frequency on the evolution of plasticity can be evaluated in the context of results for the maximum equivalent plastic strain $\varepsilon^\text{max}_p$ shown in Figs. 3.13 and 3.14. The cumulated $\varepsilon^\text{max}_p$ in the cell microstructure of pattern A increases with the increase of the undercut depth (Fig. 3.13(a)) and the zero-aperture depth (Fig. 3.13(b)) and decreases with the increase of the chromium layer thickness (Fig. 3.13(c)) and the line width (Fig. 3.13(d)). In the case of pattern B, more plasticity accumulates in the microstructure with the decrease of the chromium layer thickness (Fig. 3.14(a)) and/or the decrease of the line width (Fig. 3.14(b)). These findings are also in qualitative agreement with the damage trends.

Fig. 3.14 Maximum equivalent plastic strain $\varepsilon^\text{max}_p$ in the cell microstructure of pattern B due to static and dynamic pressure loading ($f = 1, 5,$ and $10 \text{ MHz}$) versus (a) chromium layer thickness $t$ ($b = 130 \text{ nm}$) and (b) line width $b$ ($t = 100 \text{ nm}$). The vertical dashed lines indicate the critical microstructure dimensions.
observed in megasonic cleaning experiments (Chakravorty et al., 2006). Moreover, all
dynamic simulations demonstrate a monotonic increase in cumulated $\varepsilon_p^{\text{max}}$ with
increasing loading frequency, indicating a greater propensity for microstructure failure
under high-frequency megasonic cleaning conditions. Static pressure loadings of the
same magnitude yielded insignificant plastic deformation and no collapse. Microstructure dimensions for preventing excessive plastic deformation and/or
collapse can be extracted from the results shown in Figs. 3.13 and 3.14. The critical
microstructure dimensions (indicated by vertical dashed lines) are also in good
agreement with experimental findings (Chakravorty et al., 2006).

The effect of the microstructure dimensions on the accumulation of plasticity
for a given dynamic pressure loading ($f = 10 \text{ MHz}$) can be interpreted by comparing the
$\varepsilon_p$ contours in the microstructures of patterns A and B shown in Fig. 3.15. For pattern
A, $\varepsilon_p^{\text{max}}$ occurs at the inner corners of the shallow quartz cavity (Fig. 3.15(a)), in
agreement with the deformation behavior and failure mode determined from Fig. 3.8,
whereas in the case of pattern B, plastic zones are produced at the edges of the
layer/substrate interface (Fig. 3.15(b)), indicating a higher probability for interfacial
failure. For both patterns, the locations of the cumulated $\varepsilon_p^{\text{max}}$ are the same as those of
$\sigma_f^{\text{max}}$ (Fig. 3.12).

The evolution of plasticity can be further examined by considering the variation
of $\varepsilon_p^{\text{max}}$ with the loading cycles $N$ shown in Fig. 3.16. The following power-law
relationship was obtained from curve fitting:

$$\varepsilon_p^{\text{max}} = \alpha N^{\beta}$$  \hspace{1cm} (3.1)

where material parameters $\alpha$ and $\beta$ can be determined from the intercept and the slope,
respectively, of the lines fitted through the FEM data. Therefore, Eq. (3.1) can be used
to determine the pressure loading cycles to failure (i.e., the megasonic cleaning time to
damage) in terms of the fracture strain of the microstructures of patterns A and B. Thus,
since $\varepsilon_p^{\text{max}}$ occurs either in the quartz substrate (pattern A) or the chromium/quartz
interface (pattern B), as seen in Fig. 3.15, the megasonic cleaning time to failure can be
obtained in terms of the fracture strain of the quartz substrate or the chromium/quartz
interface.
Residual stresses may also contribute to the microstructure failure during megasonic cleaning. To examine the effect of a residual stress in the chromium layer on microstructure damage, different thermal expansion coefficients were assigned to the finite elements of the chromium and quartz media, and the temperature of the entire mesh was changed uniformly to produce a residual stress in the chromium layer $\sigma_i^\varepsilon$. Hence, different levels of residual stress were generated in the chromium layer by adjusting the coefficient of thermal expansion mismatch and the temperature variation. This is an effective method for introducing a residual stress in FEM models and does not imply that the thermal expansion mismatch is the only source of residual stress.

Figures 3.17 and 3.18 show the effect of $\sigma_i^\varepsilon$ (calculated by averaging the $\sigma_i$ stress at the center nodes through the layer thickness (red nodes in Fig. 3.6) for a given temperature variation) on the distribution of $\sigma_i$ in the microstructures of patterns A and B, respectively, due to the same dynamic loading ($f = 10$ MHz). Figures 3.17(a) and 3.18(a) show that a compressive $\sigma_i^\varepsilon$ stress produces a higher $\sigma_i^{\max}$ stress in both
Fig. 3.16 Maximum equivalent plastic strain $\varepsilon^\text{max}_p$ in the cell microstructure of (a) pattern A ($t = 100$ nm, $a = 50$ nm, $d = 80$ nm, $b = 200$ nm) and (b) pattern B ($t = 100$ nm, $b = 90$ nm) due to dynamic pressure loading ($f = 10$ MHz) as a function of loading cycles $N$.

microstructures. For pattern A, a compressive $\sigma^p_t$ stress yields high tensile stress regions in the chromium layer, and $\sigma^\text{max}_t$ arises at the center of the layer (Fig. 3.17(b)). However, much higher tensile stresses were produced throughout the layer in the case of tensile $\sigma^p_t$ stress and $\sigma^\text{max}_t$ occurred at the bottom of the layer adjacent to the interface (Fig. 3.17(d)). Thus, $\sigma^p_t$ caused $\sigma^\text{max}_t$ in the microstructure of pattern A to shift from the inner corner of the shallow quartz cavity (Fig. 3.17(c)) to either the bottom or the bulk of the layer, depending on whether $\sigma^p_t$ is tensile or compressive, respectively, increasing the propensity for fracture and/or fatigue of the chromium layer. It is not possible to ascertain what type of residual stress would be more damaging based on the results shown in Fig. 3.17. Although a compressive residual stress increased $\sigma^\text{max}_t$ in the layer more than a tensile residual stress, the latter produced higher tensile stresses throughout the bulk of the layer, increasing the probability for failure in the layer by cohesive fracture or delamination from the substrate. A markedly
Fig. 3.17 Effect of average residual stress in the chromium layer \( \bar{\sigma}_l^R \) on the first principal stress \( \sigma_I \) in the cell microstructure of pattern A (\( t = 100 \, \text{nm}, a = 50 \, \text{nm}, d = 80 \, \text{nm}, b = 200 \, \text{nm} \)) due to dynamic pressure loading (\( f = 10 \, \text{MHz} \)): (a) \( \sigma_I^{\text{max}} \) versus \( \bar{\sigma}_l^R \), and \( \sigma_I \) distributions for \( \bar{\sigma}_l^R \) equal to (b) –1 GPa, (c) 0, and (d) 1 GPa. Locations of \( \sigma_I^{\text{max}} \) are identified by arrows.

Different behavior was observed with pattern B. For this pattern case, \( \sigma_I^{\text{max}} \) arose at the inner corner of the shallow quartz cavity regardless of the type of the \( \bar{\sigma}_l^R \) stress; however, \( \bar{\sigma}_l^R \) affected the overall stress distribution. A comparison of the stress distributions shown in Figs. 3.18(b)–(d) indicates that a compressive residual stress

41
Fig. 3.18 Effect of average residual stress in the chromium layer $\overline{\sigma}^R_l$ on the first principal stress $\sigma_I$ in the cell microstructure of pattern B ($t = 100$ nm, $b = 50$ nm) due to dynamic pressure loading ($f = 10$ MHz): (a) $\sigma_I^{\text{max}}$ versus $\overline{\sigma}^R_l$, and $\sigma_I$ distributions for $\overline{\sigma}^R_l$ equal to (b) $-1.0$, (c) 0, and (d) 1.0 GPa. Locations of $\sigma_I^{\text{max}}$ are identified by arrows.

produced higher tensile stresses in both the layer and the substrate. The lowest tensile stresses were obtained for zero residual stress. Unlike pattern A, the results for pattern B
indicate that a compressive residual stress would be most detrimental to the endurance of the pattern microstructure. Therefore, whether a tensile or compressive residual stress would be more destructive to the integrity of the pattern microstructure depends on the pattern geometry, material properties, and loading conditions.

3.4. Conclusions

Failure of patterned chromium-quartz APSMs due to dynamic pressure loading caused by megasonic cleaning was examined in the context of simulation results obtained from a two-dimensional plane-strain finite element analysis. Based on the obtained results and discussion, the following main conclusions can be drawn from the present study.

1) The collapse pressure of the examined pattern microstructures is generally higher than the pressure due to bubble cavitation. Therefore, it is unlikely that these mask patterns will undergo instantaneous failure during megasonic cleaning.

2) Microstructures with a larger undercut depth, larger zero-aperture distance, thinner chromium layer, and smaller line width are more likely to be damaged during megasonic cleaning. A higher load carrying capacity was determined for the microstructure without an undercut and zero-aperture cut (pattern B).

3) For a zero-aperture to π-aperture distance of 170 nm, the likelihood for microstructure failure due to plastic deformation increases significantly with the increase of the megasonic frequency, particularly for undercut depth and zero-aperture distance greater than 30 and 120 nm, respectively, and chromium layer thickness and line width less than 50 and 90 nm, respectively.

4) Simulation results revealed two different failure modes, namely shear band formation in the quartz substrate from stress raiser points (e.g., cavity corners) and cracking in the bulk of the chromium layer (cohesive fracture) or the layer/substrate interface (delamination).

5) A power-law relationship of the maximum equivalent plastic strain and the loading cycles was extracted from simulation results. This relationship can be used to estimate the critical cleaning time for microstructure failure in terms of the fracture strain of the quartz substrate (pattern A) or the chromium/quartz interface (pattern B).

6) A residual stress in the chromium layer enhances microstructure failure during megasonic cleaning.
Chapter 4

A slip-line plasticity analysis of sliding friction of rough surfaces

exhibiting self-affine (fractal) behavior

4.1. Introduction

Sliding friction due to surface plowing by relatively hard asperities and/or wear particles plays an important role in various industrial processes, such as surface micromachining and chemomechanical polishing (Lin et al., 2004). Sliding friction affects the durability of various machine elements possessing contact interfaces and the efficiency of polishing and grinding processes. Understanding of kinetic friction of sliding rough surfaces has been largely based on empirical knowledge derived from experimental findings (El-Sherbiny and Salem, 1984; Wong et al., 1998; Hisakado and Tani, 1999; Siu and Li, 2000; Dulias et al., 2002; Wang et al., 2007; Menezes et al., 2009). However, rigorous analyses of sliding friction providing accurate predictions of friction coefficient and frictional energy loss in terms of important parameters are essential for energy conservation and for designing efficient mechanical systems and material removal processes.

A significant number of friction studies have been devoted to the analysis of static friction between rough surfaces (Ogilvy, 1991; Karpenko and Akay, 2001; Yang and Komvopoulos, 2005; Lee and Polycarpou, 2007; Cohen et al., 2008) and the dependence of kinetic friction on the surface topography (El-Sherbiny and Salem, 1984; Ogilvy, 1991; Ford, 1993; Karpenko and Akay, 2001; Wang et al., 2007) or the applied normal load (Ogilvy, 1991; Tworzydlo et al., 1998; Ali and Sahoo, 2006). However, these studies have been based on semi-empirical approaches that rely on scale-dependent statistical topography parameters, such as mean and variance of the summit heights and slopes. To overcome problems with traditional approaches based on scale-dependent statistical surface parameters (Greenwood and Williamson, 1966) and to model real surfaces exhibiting multi-scale roughness, the surface topography in recent studies of rough surfaces was described by fractal geometry (Majumdar and Bhushan, 1990; Wang and Komvopoulos, 1995; Yan and Komvopoulos, 1998; Borri-Brunetto et al., 1999; Ciavarella et al., 2000; Persson et al., 2002; Komvopoulos and Yang, 2006; Yin and Komvopoulos, 2010). However, kinetic friction of rough surfaces demonstrating fractal behavior has received significantly less attention. Palasantzas (2004) investigated the adhesive characteristics of rubber sliding against a fractal surface on friction, Sahoo and Roy Chowdhury (2000) examined friction of adhesive fractal surfaces subjected to small-amplitude sliding, and Buzio et al. (2003) used the atomic force microscope to study the friction behavior of carbon films.
exhibiting fractal behavior. However, these studies considered only adhesion between elastically deformed asperities (Sahoo and Roy Chowdhury, 2000; Palasantzas, 2004) or did not examine the dependence of the friction coefficient on various factors (Buzio et al., 2003), particularly the effect of important parameters on the frictional energy dissipated at plastic asperity contacts.

A significant body of the friction literature includes studies dealing with surface plasticity induced by hard asperities or wear particles trapped at the contact interface. Because of relatively large plastic strains, deformation was assumed to be rigid-perfectly plastic, allowing for plane-strain deformation during plowing and cutting by hard asperities to be analyzed by the slip-line theory of plasticity (Challen and Oxley, 1979; Komvopoulos et al., 1986a,b; Hokkirigawa and Kato, 1988; Shi and Ramalingam, 1991; Kopalinsky and Black, 1995; Kopalinsky and Oxley, 1995; Black et al., 1997; Fang and Jawahir, 2002; Fang and Dewhurst, 2005; Bressan and Williams, 2009). Slip-line fields for surface plowing have been used to study deformation at the asperity scale by modeling a hard asperity (wear particle) by a rigid wedge plowing through the soft surface and pushing ahead a wave of the deformed material (Komvopoulos et al., 1986a; Kopalinsky and Black, 1995; Kopalinsky and Oxley, 1995; Black et al., 1997; Bressan and Williams, 2009). In addition, slip-line fields of a rigid asperity cutting through a soft material have been used to study friction and wear behavior at the asperity level (Komvopoulos et al., 1986b; Hokkirigawa and Kato, 1988; Shi and Ramalingam, 1991; Kopalinsky and Oxley, 1995; Fang and Jawahir, 2002; Fang and Dewhurst, 2005).

Although the above plasticity analyses have yielded important insight into the dependence of friction on various deformation modes at a single asperity contact (e.g., plowing, wave formation, and cutting), a comprehensive analysis of sliding friction of rough surfaces comprised of multi-wavelength asperity contacts has not been reported yet. Thus, the main objective of this study was two-fold. First, instead of using an empirical approach based on experimental evidence, a generalized friction model of a hard and rough surface plowing and cutting through a soft and smooth surface was derived based on the slip-line theory of plasticity. Second, the friction coefficient and energy dissipated during sliding were determined in terms of interfacial shear strength (indicative of the adhesion characteristics of the interacting surfaces), surface topography (fractal parameters) of the rough surface, total normal load (global interference), and elastic-plastic material properties of the deformable (soft) surface. The effects of the latter factors on friction coefficient and energy loss during sliding of surfaces exhibiting multi-scale roughness are discussed in the context of numerical results of representative ceramic-ceramic, ceramic-metallic, and metal-metal contact systems.
4.2. Single-asperity plasticity analysis

Contact between rough surfaces occurs at surface protrusions (asperities) and, possibly, between asperities and wear particles trapped between the sliding surfaces. The resulting microscopic contacts (hereafter referred to as the microcontacts) span a wide range of length scales, from a few nanometers to several hundred micrometers. Deformation at these microcontacts can be elastic or plastic, depending on the local surface topography, normal load distribution, and elastic-plastic material properties. Deformation and friction behavior of a single plastic microcontact have been analyzed by the slip-line theory of plasticity, assuming rigid and sharp asperities (wear particles) plowing and cutting through a soft surface modeled as a rigid-perfectly plastic material (Challen and Oxley, 1979; Petryk, 1987; Lacey and Torrance, 1991; Kopalinsky and Oxley, 1995; Torrance and Buckley, 1996; Fang and Jawahir, 2002; Dundur and Das, 2009). However, because hard asperities and wear particles possess finite curvatures at micrometer or submicrometer scales, the local geometry is more accurately represented by a spherical shape (Tsukizoe and Sakamoto, 1975). Challen and Oxley (1983) approximated the contact arc of a rigid cylindrical asperity with the plowed surface by its chord and used a slip-line field for wedge asperities (Challen and Oxley, 1979) to analyze the polishing process. Busquet and Torrance (2000) developed a slip-line model of a cylindrical slider plowing through a soft material without removing material. Fang (2003) presented a slip-line model of surface machining by a round-edge tool; however, this model is suitable for macroscopic machining and, therefore, cannot be used to analyze cutting-mode deformation at the asperity and wear particle scale.

In the present analysis it is assumed that the dominant mode of friction is plowing and that the deformed material is removed by a microcutting process. A hard (rigid) asperity (wear particle) of radius of curvature $R$ penetrates into a soft material to depth $d_s$ (Fig. 4.1(a)). For relatively shallow indentation ($d_s/R < 0.1$), plane-strain deformation of the soft material is a reasonable simplification (Komvopoulos et al., 1986a). Plastic shearing ahead of the hard asperity results in sliding of the plastically deformed material against the rigid asperity at a relative velocity $u_c$ (Fig. 4.1(a)). The deformed surface is modeled as an isotropic, rigid-perfectly plastic material, moving toward the rigid asperity at a constant velocity $u$.

The deformation field of the soft material is represented by a network of orthogonal shear lines, referred to as the $\alpha$- and $\beta$-lines (Fig. 4.1(b)). Because separation of the material sliding against the rigid asperity occurs at point B, boundary BD is a free surface (i.e., zero surface traction) and, therefore, the $\alpha$- and $\beta$-lines intersect BD at an angle of 45º. The angle $\eta$ between an $\alpha$-line and the tangent at any point along the contact interface BEG depends on the ratio of the interfacial shear
Fig. 4.1  (a) Slip-line field of a rigid spherical asperity (wear particle) plowing through a rigid-perfectly plastic surface and removing material by a microcutting process, (b) detailed view of the plastic region showing the orthogonal network of $\alpha$- and $\beta$-lines, and (c) hodograph of the slip-line field (Komvopoulos, 2011).
strength \( s \) to the shear strength of the plowed material \( k \), i.e.,

\[
\eta = \frac{1}{2} \cos^{-1}\left(\frac{s}{k}\right)
\]  

(4.1)

where \( 0 \leq s/k \leq 1 \). The dimensionless shear strength ratio \( s/k \) represents the interfacial adhesion characteristics, with the extreme values corresponding to the idealized conditions of adhesionless \((s/k = 0)\) and sticking \((s/k = 1)\) contact interfaces.

For a slip-line field to be admissible, it must satisfy mass conservation and certain geometric, trigonometric, and kinematic conditions. Since the plowed material is modeled as rigid-perfectly plastic (incompressible), mass conservation yields

\[
\Delta \dot{V} = l \cdot u_c = d_s \cdot u
\]  

(4.2)

where \( \Delta \dot{V} \) is the material removal rate at the microcontact level and \( l \) is the thickness of the removed material.

The following relationships are deduced from geometry considerations:

\[
R \cos \phi + t + d_s = R
\]  

(4.3)

\[
l \sin\left(\frac{\pi}{4} + \beta\right) = t = \frac{l \sin\left(\frac{\pi}{4} - \eta\right)}{\cos\left(\frac{\pi}{4} - \eta\right)}
\]  

(4.4)

\[
\beta + \phi + \eta = \frac{\pi}{2}
\]  

(4.5)

\[
d_s = R_1 (1 - \cos(\theta - \beta)) + \frac{l \sin(\theta - \beta)}{\sqrt{2} \cos\left(\frac{\pi}{4} - \eta\right)}
\]  

(4.6)

\[
R_1 \sin(\theta - \beta) + \frac{l \cos(\theta - \beta)}{\sqrt{2} \cos\left(\frac{\pi}{4} - \eta\right)} = R \sin \phi + \frac{l \cos\left(\frac{\pi}{4} + \beta\right)}{\cos\left(\frac{\pi}{4} - \eta\right)}
\]  

(4.7)

where \( t \) is the pile-up thickness, \( R_1 \) is the radius of curvature of arc IG, and \( \phi, \beta, \) and \( \theta \) are slip-line field angles (Figs. 4.1(a) and 4.1(b)).

Figure 4.1(c) shows the hodograph of the slip-line field shown in Fig. 4.1(a), where \( \mathbf{u}_{DI} \) is the relative velocity along boundary DI (Fig. 4.1(b)). Kinematic
admissibility considerations yield the following velocity condition in region BDF consisting of straight slip-lines:

\[ (u_\alpha^B)^2 + (u_\beta^B)^2 = u_c^2 \sin^2 \phi + (u_c \cos \phi + u)^2 \]  (4.8)

where \( u_\alpha^B \) and \( u_\beta^B \) are the velocities at point B along the \( \alpha \)- and \( \beta \)-lines, respectively, obtained from a numerical scheme described elsewhere (Komvopoulos, 2011).

For a given velocity ratio \( u_c/u \), Eqs. (4.1)-(4.7) were solved simultaneously to obtain \( \beta, \phi, \theta, \) and \( R_1 \) in terms of \( s/k \) and \( d_s/R \). If the obtained slip-line field did not satisfy the velocity condition given by Eq. (4.8), a different value of \( u_c/u \) was assumed, and the numerical procedure was repeated until a kinematically admissible slip-line field was obtained.

The coefficient of friction at a single fully plastic microcontact \( \mu \) can be expressed in terms of the resultant forces acting on boundaries BE and EG, i.e.,

\[ \mu = \frac{\Delta F_y^p}{\Delta F_x^p} = \frac{\Delta F_x^{BE} + \Delta F_y^{EG}}{\Delta F_x^{BE} + \Delta F_y^{EG}} \]  (4.9)

where \( \Delta F_x^p \) and \( \Delta F_y^p \) are resulting forces in the \( x \)- and \( y \)-direction. The normal and tangential stresses at the contact interface BEG were obtained from Henky’s plasticity equations, using the Mohr circle and the known boundary conditions, i.e., stress-free surface BD and Eq. (4.1). \( \Delta F_x^{BE}, \Delta F_y^{BE}, \Delta F_x^{EG}, \) and \( \Delta F_y^{EG} \) were obtained by integrating the stress components acting along the corresponding interface (Komvopoulos, 2011), i.e.,

\[ \Delta F_x^{BE} = \int_0^{\phi_0} \left[ 1 + 2\alpha_N + \sin(2\eta) \right] kR \sin(\phi - \alpha_N) d\alpha_N - sR \sin(\phi - \sin(\phi - \phi_0)) \]  (4.10)

\[ \Delta F_y^{BE} = \int_0^{\phi_0} \left[ 1 + 2\alpha_N + \sin(2\eta) \right] kR \cos(\phi - \alpha_N) d\alpha_N + sR \cos(\phi - \phi_0) - \cos \phi \]  (4.11)

\[ \Delta F_x^{EG} = \int_0^{\phi_0 - \phi_0} \left[ 1 + 2\phi_0 + 2(\alpha_3 - \alpha_5 + \alpha_2) + \sin(2\eta) \right] kR \sin \lambda d\lambda - sR \sin(\phi - \phi_0) \]  (4.12)

\[ \Delta F_y^{EG} = \int_0^{\phi_0 - \phi_0} \left[ 1 + 2\phi_0 + 2(\alpha_3 - \alpha_5 + \alpha_2) + \sin(2\eta) \right] kR \cos \lambda d\lambda + sR [1 - \cos(\phi - \phi_0)] \]  (4.13)

where angles \( \phi_0, \alpha_N, \lambda, \alpha_2, \alpha_3, \) and \( \alpha_5 \) are shown in Fig. 4.1(a). Angles \( \alpha_2, \alpha_3, \) and \( \alpha_5 \) are obtained from geometric and trigonometric relationships (Appendix), whereas angle \( \phi_0 \) is given by
4.3. Contact mechanics analysis

4.3.1. Rough surface contact model

The hard and rough surface is assumed to be isotropic and self-affine, with a profile given by (Yan and Komvopoulos, 1998)

\[ z(x, y) = L \left( \frac{G}{L} \right)^{(D-2)} \left( \frac{\ln \gamma}{M} \right)^{1/2} \sum_{m=1}^{M} \sum_{q=0}^{q_m} \gamma^{(D-3)q} \times \]

\[ \left\{ \cos \phi_{m,q} - \cos \left[ \frac{2 \pi \gamma^q (x^2 + y^2)^{1/2}}{L} \right] \cos \left[ \tan^{-1} \left( \frac{y}{x} \right) - \frac{\pi m}{M} \right] + \phi_{m,q} \right\} \]

where \( G \) and \( D \) are the fractal roughness and fractal dimension (2<\(D<3\)), respectively, both independent of spatial frequency in the regime where the surface exhibits fractal behavior, \( M \) is the number of superimposed ridges, \( L \) is the sample length, \( q \) is a spatial frequency index varied in the range \( 0 \leq q \leq \text{int} \left[ \log(L/L_s)/\log \gamma \right] \), where \( L_s \) is the a cut-off length (typically on the order of the material’s lattice dimension for continuum description to hold), \( \phi_{m,q} \) is a random phase evenly distributed in the range \([0, 2\pi]\), used to prevent the coincidence of different frequencies at each point of the surface profile, and \( \gamma (\gamma > 1) \) is a parameter controlling the frequency density in the surface profile with a typical value of 1.5 (Komvopoulos and Yan, 1997).

The fractal roughness \( G \) determines the wave amplitudes of all frequencies comprising the surface profile, whereas the fractal parameter \( D \) indicates the relative contributions of low- and high-frequency components in the profile. The surface roughness increases with increasing \( G \) or decreasing \( D \). Since \( G \) and \( D \) are scale-invariant, surface description at different length scales does not depend on the measurement range. The longest and shortest wavelengths are respectively set equal to the sample length \( L \) and the instrument resolution limit, which is above Thus, the lowest and highest frequencies in the surface profile are \( \omega_l = 1/L \) and \( \omega_h = 1/L_s \), respectively.

A typical three-dimensional (3D) fractal surface constructed from Eq. (4.14) is shown in Fig. 4.2. Equation (4.14) was used to generate the rough (fractal) surfaces in
Fig. 4.2 Three-dimensional fractal surface generated from Eq. (4.14) for \( G = 2.39 \times 10^{-4} \) nm, \( D = 2.24, M = 10, \gamma = 1.5, L = 7.04 \times 10^3 \) nm, and \( L_s = 5 \) nm (\( \sigma = 290 \) nm).

the analysis presented in the following sections. Because the rough (fractal) surface is assumed to be isotropic, any two-dimensional (2D) surface profile is a statistical representation of the 3D surface topography. Figure 4.3 shows schematically a 2D profile of a fractal surface (Eq. (4.14)) truncated by a rigid plane. Displacement of the rigid plane toward the rough surface by a global interference \( h \) produces several truncated segments, which are approximated by spherical asperities having a base area equal to the truncation area \( a' \), height equal to the local interference \( \delta \), and radius of curvature \( R \), given by (Yan and Komvopoulos, 1998)

\[
\delta = \frac{(a')^{(3-D)/2}}{2^{(D-4)} \pi^{(3-D)/2} G^{(2-D)} (\ln \gamma)^{1/2}} 
\tag{4.15}
\]

\[
R = \frac{(a')^{(D-1)/2}}{2^{(5-D)} \pi^{(D-1)/2} G^{(D-2)} (\ln \gamma)^{1/2}} 
\tag{4.16}
\]

The size distribution of truncated asperities is assumed to follow an island-distribution rule (Mandelbrot, 1975, 1983), as in previous contact mechanics analyses of fractal surfaces (Majumdar and Bhushan, 1990; Wang and Komvopoulos, 1994; Yan and Komvopoulos, 1998; Komvopoulos and Ye, 2001), i.e.,

\[
N(a') = \left( \frac{a'_l}{a'} \right)^{(D-1)/2} 
\tag{4.17}
\]

where \( N(a') \) is the number of truncated asperities with areas greater than \( a' \), and \( a'_l \) is the largest truncated asperity area obtained at a given global interference. The
distribution density function of truncated asperities $n(a')$ of a 3D surface profile is given by (Yan and Komvopoulos, 1998)

$$n(a') = -\frac{dN(a')}{da'} = \frac{(D-1)}{2a_L} \left( \frac{a_L}{a'} \right)^{(D+1)/2} (4.18)$$

For a given global interference, $a_L'$ can be obtained from the total contact area of truncated asperities $S'$ using the relationship (Komvopoulos and Ye, 2001):

$$S' = \int_{a_S'}^{a_L'} a'n(a')da'$$

(4.19)

where $a_S'$ is the smallest truncated asperity area, which, for continuum description to hold, is set equal to 5–6 times the lattice distance of the deformed material.

After substituting Eq. (4.18) into Eq. (4.19) and integrating,

$$S' = \frac{D-1}{3-D} \left[ 1 - \left( \frac{a_S'}{a_L'} \right)^{(3-D)/2} \right] a_L'$$

(4.20)

For a given global interference, $S'$ was obtained numerically as the sum of the areas of all truncated asperities (Komvopoulos and Ye, 2001), then $a_L'$ was determined as a function of $a_S'$, $D$, and $S'$ from Eq. (4.20), and, finally, the area range $[a_S', a_L']$ and spatial distribution of the truncated asperities were obtained from Eqs. (4.18) and (4.20).

The deformation mode at the microcontact level is determined by the critical contact area of truncated asperity demarcating the transition from elastic to fully plastic deformation (Yan and Komvopoulos, 1998). In the present analysis, the rough surface is assumed to be rigid, while the soft substrate material is assumed to be elastic-perfectly plastic, implying either fully plastic or elastic microcontact deformation. Whether the deformation mode at a microcontact is fully plastic ($a' \leq a'_C$) or elastic ($a' > a'_C$) depends on the critical contact area of truncated asperity $a'_C$, given by (Yan and Komvopoulos, 1998)

$$a'_C = \left[ \frac{2^{(11-2D)}}{9\pi^{(4-D)}} G^{(2D-4)} \left( \frac{E}{H} \right)^2 \ln \gamma \right]^{1/(D-2)}$$

(4.21)
where $E$ and $H$ are the elastic modulus and hardness of the soft surface, respectively. Equation (4.21) indicates that $a'_c$ is a function of the fractal parameters of the rough surface $D$ and $G$ and the elastic and plastic properties of the softer material $E$ and $H$, implying that $a'_c$ is an intrinsic parameter of a contact system, i.e., independent of global parameters such as total normal load and global interference.

The deformation mode at the microcontact level can be determined by comparing $a'_l$ with $a'_c$. Thus, when $a'_l > a'_c$, both elastic and fully plastic microcontacts co-exist at the contact interface, whereas when $a'_l \leq a'_c$, only fully plastic microcontacts exist at the contact interface.

![Fig. 4.3 Two-dimensional profile of a truncated rough (fractal) surface.](image)

The total truncated contact area of fully plastic microcontacts $S'_p$ is given by

$$S'_p = \int_{a'_l}^{a'_c} a'n(a')da'$$

(4.22)
Substituting Eq. (4.18) into Eq. (4.22) and integrating,

\[
S_p' = \left( \frac{D - 1}{3 - D} \right) a'_L \left( (D - 1)/2 \right) \left[ a'_C \left( (3 - D)/2 \right) - a'_S \left( (3 - D)/2 \right) \right]
\]  

(4.23)

4.3.2. Friction analysis

The global coefficient of friction \( f \) is defined as the ratio of the total horizontal force \( F_x \) to the total vertical force \( F_y \), i.e.,

\[
f = \frac{F_x}{F_y} = \frac{F_x^p + F_x^e}{F_y^p + F_y^e}
\]  

(4.24)

where superscripts \( p \) and \( e \) denote fully plastic and elastic microcontacts, respectively. These total force components are obtained as the sum of all microcontact forces acting in the particular direction, i.e.,

\[
F_x^p = \int_{a'_e} \left( \Delta F_{x, BE}^p + \Delta F_{x, EG}^p \right) n(a') da'
\]  

(4.25)

\[
F_x^e = \int_{a'_e} \Delta F_x^e n(a') da'
\]  

(4.26)

\[
F_y^p = \int_{a'_e} \left( \Delta F_{y, BE}^p + \Delta F_{y, EG}^p \right) n(a') da'
\]  

(4.27)

\[
F_y^e = \int_{a'_e} \Delta F_y^e n(a') da'
\]  

(4.28)

where \( \Delta F_x^e \) and \( \Delta F_y^e \) are the horizontal and vertical forces applied to a single elastic microcontact, respectively, obtained from Hertz theory (Johnson, 1987) as

\[
\Delta F_x = s a_e
\]  

(4.29)

where \( a_e \) is the real contact area of an elastic asperity, given by

\[
a_e = \frac{a'}{2}
\]  

(4.30)

and

\[
\Delta F_y = \left[ \frac{16R E^2 S^3}{9} \right]^{1/2}
\]  

(4.31)

From Eqs. (4.1)–(4.7), (4.10)–(4.13), (4.15), (4.16), (4.18), (4.20), (4.21), and (4.24)–(4.31), it follows that

\[
f = f(D, G, \frac{s}{k}, E, H, F_y)
\]  

(4.32)

Equation (4.32) indicates that the global friction coefficient depends on the
topography parameters of the rough surface \((D \text{ and } G)\), the interfacial shear strength \((s/k)\), the elastic and plastic material properties of the soft surface \((E \text{ and } H)\), and (indirectly) the total normal load \(F_y\) through the total truncated contact area \(S'\), which defines the magnitude of \(a'_L\) (Eq. (4.20)).

### 4.3.3. Energy dissipation at the contact interface

The external energy \(E_{ex}\) supplied to the contact system comprises reversible and irreversible energies \(E_{re}\) and \(E_{ir}\), respectively, where \(E_{re}\) represents the energy stored as elastic strain energy \(E_e\), while \(E_{ir}\) is the sum of the energy dissipated to shear plastically the soft material \(E_s\) and the energy dissipated to remove material by the microcutting process \(E_c\). Hence,

\[
E_{ex} = E_{re} + E_{ir} = E_e + (E_s + E_c)
\]

(4.33)

where

\[
E_{re} = \int_{a_c}^{d_c} \Delta E_e(a')n(a')da'
\]

(4.34)

\[
E_{ir} = \int_{a_s}^{d_c} \Delta E_p(a')n(a')da'
\]

(4.35)

where \(\Delta E_e\) is the energy stored at an elastic microcontact and \(\Delta E_p\) is the energy dissipated by plastic deformation at a fully plastic microcontact.

The energy dissipated to remove material \(E_c\) can be expressed as

\[
E_c = \int_{a_s}^{d_c} \Delta E_c(a')n(a')da'
\]

(4.36)

where \(\Delta E_c\) is the energy dissipated at a single fully plastic microcontact to remove material, given by \(\Delta E_c = \Delta E_p - \Delta E_s\), where \(\Delta E_s\) is the energy dissipated at a fully plastic microcontact to plastically shear the material.

Energy terms \(\Delta E_e\), \(\Delta E_p\), and \(\Delta E_c\) can be expressed as:

\[
\Delta E_e = \left(\Delta F_x^e \cdot \mathbf{u}\right) \cdot T
\]

(4.37)

\[
\Delta E_p = \left[\left(\Delta F_x^{BE} + \Delta F_x^{EG}\right) \cdot \mathbf{u}\right] \cdot T
\]

(4.38)

\[
\Delta E_c = \left[\left(\Delta F_x^{BE} + \Delta F_x^{EG}\right) \cdot \mathbf{u}_c + \left(\Delta F_y^{BE} + \Delta F_y^{EG}\right) \cdot \mathbf{u}_c\right] \cdot T
\]

(4.39)

where \(T\) is the total time of sliding.
The fractions of elastic energy $\xi_e$, plastic shearing energy $\xi_s$, and material removal energy $\xi_c$ are defined as

$$\xi_e = \frac{E_e}{E_{ex}}, \quad \xi_s = \frac{E_s}{E_{ex}}, \quad \xi_c = \frac{E_c}{E_{ex}} \quad (4.40)$$

Using Eqs. (4.2), (4.18), (4.29), (4.30), (4.33)–(4.40) and Fig. 4.1(a), these energy fractions can be expressed as

$$\xi_e = \frac{s(D-1)}{2(3-D)} \left[ 1 - \left(\frac{a'_c}{a'_L}\right)^{(3-D)/2} \right] a'_L \quad (4.41a)$$

$$\xi_s = \frac{\int_{a'_s}^{a'_c} \left( (\Delta F_{BE} + \Delta F_{EG}) \right) n(a') da' + \frac{s(D-1)}{2(3-D)} \left[ 1 - \left(\frac{a'_c}{a'_L}\right)^{(3-D)/2} \right] a'_L}{\int_{a'_s}^{a'_c} \left[ (\Delta F_{BE} + \Delta F_{EG}) \right] n(a') da' + \frac{s(D-1)}{2(3-D)} \left[ 1 - \left(\frac{a'_c}{a'_L}\right)^{(3-D)/2} \right] a'_L} \quad (4.41b)$$

$$\xi_c = \frac{\int_{a'_s}^{a'_c} \left[ (\Delta F_{BE} + \Delta F_{EG}) \sin\phi - (\Delta F_{BE} + \Delta F_{EG}) \cos\phi \right] n(a') da' + \frac{s(D-1)}{2(3-D)} \left[ 1 - \left(\frac{a'_c}{a'_L}\right)^{(3-D)/2} \right] a'_L \left(\frac{d_s}{l}\right)}{\int_{a'_s}^{a'_c} \left[ (\Delta F_{BE} + \Delta F_{EG}) \sin\phi - (\Delta F_{BE} + \Delta F_{EG}) \cos\phi \right] n(a') da' + \frac{s(D-1)}{2(3-D)} \left[ 1 - \left(\frac{a'_c}{a'_L}\right)^{(3-D)/2} \right] a'_L \left(\frac{d_s}{l}\right)} \quad (4.41c)$$

From Eqs. (4.1)–(4.7), (4.10)–(4.13), (4.18), (4.20), (4.21), and (4.41a-4.41c) it follows that

$$\xi_i = f(E, H, D, G, \frac{S}{k}, a'_s, F_y) \quad (i = e, s, \text{ or } c) \quad (4.42)$$

Equation (4.42) indicates that the fraction of energy dissipated to remove material by the microcutting process depends on the elastic and plastic material properties of the deformed surface, the topography of the plowing rough (fractal) surface, the interfacial shear strength (surface adhesion characteristics), the smallest area of truncated asperity (fixed at ~5-6 times the lattice distance of the soft material), and (indirectly) the total normal load through $S'$ which controls the magnitude of $a'_L$ (Eq. (4.20)).
4.4. Results and discussion

4.4.1. Single microcontact

Before presenting results for contacting rough surfaces it is instructive to consider friction at the microcontact level. For a fully plastic microcontact, $\Delta F_y^p = Ha_s = 6k$ (since $H = 6k$). Using Eq. (4.9) and relationship $\Delta F_y^p = 6ka_s$, the friction coefficient of a fully plastic microcontact $\mu$ is plotted as a function of the dimensionless normal load $\Delta F_y^p/(kR^2)$ and interfacial shear strength $s/k$, as shown in Fig. 4.4. Although the friction coefficient increases with both normal load and interfacial adhesion, the load effect is more pronounced than that of interfacial adhesion, the latter becoming significant only in the high-load range. For $\Delta F_y^p/(kR^2) = 2.5$, for example, the friction coefficient of adhesionless surfaces ($s/k = 0$) is higher than that of sticking surfaces ($s/k = 1$) by more than 80%. In view of the secondary effect of adhesion (i.e., $s/k$) on the magnitude of $\mu$, it may be inferred that the dominant contribution to the friction coefficient at the microcontact level is from plowing.

![Fig. 4.4 Friction coefficient $\mu$ of a fully plastic microcontact as a function of dimensionless normal load $\Delta F_y/(kR^2)$ and ratio of interfacial shear strength to shear strength of the soft material $s/k$.](image-url)
4.4.2. Rough surface

Because Eqs. (4.32) and (4.42) cannot be solved analytically, numerical results of the global friction coefficient $f$ and the material removal energy fraction $\xi_c$ are presented below for representative ceramic-metallic ($\text{Al}_2\text{O}_3$/CrN), ceramic-ceramic ($\text{Al}_2\text{O}_3$/TiC), and metal-metal (AISI 1095 steel/AISI 1020 steel) contact systems with material properties given in Table 4.1. All of the numerical solutions presented below are for a 3D surface profile generated from Eq. (4.14) for $L = 7.04 \times 10^3$ nm, $L_s = 5$ nm, $\gamma = 1.5$, $M = 10$, and $\phi_{m,q} = \pi/2$, and smallest area of truncated asperity $a'_S = \pi (r'_s)^2 = \pi (L_s / 2)^2 = 19.6$ nm$^2$. The effect of global interference (normal load) on the microcontact deformation mode is analyzed first, followed by results of the global friction coefficient and energy dissipated during sliding.

<table>
<thead>
<tr>
<th>Contact system</th>
<th>Material</th>
<th>$E$ (GPa)</th>
<th>$H$ (GPa)</th>
<th>$E/H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Al}_2\text{O}_3$/CrN</td>
<td>$\text{Al}_2\text{O}_3$</td>
<td>307</td>
<td>27.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CrN</td>
<td>103</td>
<td>14.8</td>
<td>6.96</td>
</tr>
<tr>
<td>$\text{Al}_2\text{O}_3$/TiC</td>
<td>$\text{Al}_2\text{O}_3$</td>
<td>307</td>
<td>27.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TiC</td>
<td>450</td>
<td>23.5</td>
<td>19.2</td>
</tr>
<tr>
<td>AISI 1095/AISI 1020</td>
<td>AISI 1095</td>
<td>200</td>
<td>6.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AISI 1020</td>
<td>200</td>
<td>1.71</td>
<td>117</td>
</tr>
</tbody>
</table>

(a)Sources: Suh (1986); Komvopoulos and Zhang (2008)

4.4.2.1. Microcontact deformation

Figure 4.5(a) shows the variation of the total area of truncated asperities $S'$, normalized by the apparent contact area $S_a$ ($S_a = L^2$), as a function of the global
Fig. 4.5 (a) Total contact area of truncated asperities normalized by the apparent contact area $S'/S_a$ and (b) total contact area of truncated fully plastic asperities normalized by the total contact area of truncated asperities $S'_p/S'$ versus global interference $h$ for Al$_2$O$_3$/CrN contact system with material properties given in Table 4.1 ($\sigma = 1000$ nm).
interference \( h \) for Al\(_2\)O\(_3\)/CrN contact system (\( \sigma = 1000 \) nm). Although \( S'/S_a \) increases approximately linearly with the global interference, it remains a very small fraction of the apparent contact area (on the order of \( 10^{-4} \)) despite the variation of the global interference (normal load) in a fairly wide range. Figure 4.5(b) shows the ratio of the total area of fully plastic truncated asperities to the total area of truncated asperities \( S'_p/S' \) as a function of global interference \( h \) for the same contact system and roughness. The very low magnitude of \( S'_p/S' \) suggests that fully plastic microcontacts are a small fraction of all microcontacts comprising \( S' \), suggesting that microcontact deformation is predominantly elastic. The rapid decrease of \( S'_p/S' \) with increasing \( h \) in the low-interference regime (\( h < 40 \) nm) indicates an enhancement of the dominance of elastic deformation at the microcontact level with increasing normal load.

From Eqs. (4.20) and (4.23), it follows that

\[
S'_p/S' = \frac{\left( \frac{a'_C}{a'_L} \right)^{(3-D)/2} - \left( \frac{a'_S}{a'_L} \right)^{(3-D)/2}}{1 - \left( \frac{a'_S}{a'_L} \right)^{(3-D)/2}}
\]

(4.43)

Since \( a'_S/a'_L << 1 \) and \( a'_C \) is only a function of topography (fractal) parameters and material properties (Eq. (4.21)), i.e., \( a'_C \) is fixed for a given contact system, Eq. (4.43) can be simplified as following

\[
\frac{S'_p}{S'} \propto \left( \frac{1}{a'_L} \right)^{(3-D)/2}
\]

(4.44)

Equation (4.44) provides an explanation for the decrease of \( S'_p/S' \) with increasing global interference (Fig. 4.5(b)), considering that the latter implies an increase in \( a'_L \). The small negative slope of the \( S'_p/S' \) curve for \( h > ~50 \) nm implies that the variation of the global interference in this range does not influence significantly the contribution of plastic deformation to the overall deformation at the contact interface.

4.4.2.2. Coefficient of friction

Figure 4.6(a) shows the friction coefficient \( f \) (Eq. (4.32)) of the Al\(_2\)O\(_3\)/CrN contact
system ($\sigma = 1000$ nm) as a function of global interference $h$ and dimensionless interfacial shear strength $s/k$. The results indicate that friction mainly depends on interfacial adhesion conditions and the effect of global interference (normal load) is secondary, in qualitative agreement with experimental observations. Figure 4.6(b) shows a comparison of friction coefficients of typical ceramic-metallic ($\text{Al}_2\text{O}_3/\text{CrN}$), ceramic-ceramic ($\text{Al}_2\text{O}_3/\text{TiC}$), and metal-metal (AISI 1095/AISI 1020) contact systems for fractal dimension $D$ varied between 2.2 and 2.8, or surface roughness $\sigma$ varied in the wide range of $(0.02–126) \times 10^4$ nm and fixed all other parameters ($G = 2.39 \times 10^{-4}$ nm, $s/k = 0.5$, and $h = 15$ nm). A trend of the friction coefficient to decrease with increasing $D$ is observed with all contact systems. For fixed fractal parameters (roughness), friction is controlled by the material properties of the softer surface. It is noted that $f$ scales with the inverse of the $E/H$ ratio of the soft material (Table 1), and the effect of the elastic and plastic properties of the soft surface becomes more significant with increasing roughness. A similar trend of friction coefficient $f$ to increase with surface roughness $\sigma$ is also observed in Fig. 4.6(c) for fixed material properties ($\text{Al}_2\text{O}_3/\text{CrN}$), especially with increasing interfacial shear strength (adhesion). For $s/k = 0$, friction is very low (typical of well-lubricated surfaces) and shows a weak dependence on surface roughness, in agreement with experimental observations.

4.4.2.3. Energy dissipation

Figure 4.7(a) shows the energy dissipation ratio $\xi_c$ versus global interference $h$ for fixed material properties ($\text{Al}_2\text{O}_3/\text{CrN}$) and surface roughness ($\sigma = 1000$ nm). In the ideal case of adhesionless contact interface ($s/k = 0$), $\xi_c \approx 0.5$ regardless of global interference. Since elastically deformed microcontacts do not dissipate energy irreversibly, it may be inferred that under this condition $\xi_c \approx \xi_p$, implying that half of the external energy is dissipated to remove material from the softer surface (i.e., CrN in this case) and the other half to deform plastically the microcontacts. For $s/k > 0$, $\xi_c$ decreases sharply with increasing $h$ (or normal load). This can be attributed to the rapid decrease of the fraction of fully plastic microcontacts with the increase of global interference (Fig. 4.5(b)), implying that most of the external energy is temporarily stored as elastic strain energy in elastically deformed microcontacts. For very small global interference ($h = 10$ nm), $\xi_c \approx 0.5$ regardless of $s/k$ because all microcontacts exhibit fully plastic deformation since $a'_l < a'_r$. The general trend for $\xi_c$ to increase with the decrease of $s/k$ illustrates the beneficial effect of lubrication on the removal of material due to minimization of energy loss to overcome interfacial adhesion at the microcontacts.

Figure 4.7(b) illustrates the effect of material properties of the soft surface and the fractal dimension (roughness) of the rough surface on $\xi_c$ for fixed all other parameters ($h = 15$ nm, $G = 2.39 \times 10^{-4}$ nm, and $s/k = 0.5$). Among the three contact systems, $\xi_c$ of the ceramic-ceramic contact system shows the greatest dependence on
Fig. 4.6 (a) Friction coefficient $f$ of Al$_2$O$_3$/CrN contact system versus global interference $h$ and ratio of interfacial shear strength to shear strength of the softer material $s/k$ for $\sigma = 1000$ nm; (b) friction coefficient $f$ of Al$_2$O$_3$/CrN, Al$_2$O$_3$/TiC, and AISI 1095/AISI 1020 contact systems versus fractal dimension $D$ and rms surface roughness for $h = 15$ nm, $G = 2.39 \times 10^{-4}$ nm, and $s/k = 0.5$; and (c) friction coefficient $f$ of Al$_2$O$_3$/CrN contact system versus rms surface roughness $\sigma$ and ratio of interfacial shear strength to shear strength of the softer material $s/k$ for $h = 15$ nm and $D = 2.24$. (The material properties of each contact system are given in Table 4.1.)

surface roughness. The fact that $\xi_c$ of the metal-metal contact system does not show a dependence on fractal dimension (roughness) is attributed to the largest $a'_c$ values of this contact system that has the highest $E/H$ ratio (Table 1), implying that most of the microcontacts are fully plastically deformed. Alternatively, the ceramic-metallic system demonstrates the lowest $\xi_c$ through the entire roughness (fractal dimension) range because of the dominance of elastic deformation of the microcontacts due to the lowest $a'_c$ value (lowest $E/H$ ratio) of this contact system.

Figure 4.7(c) shows the variation of $\xi_c$ with surface roughness $\sigma$ and $s/k$ for fixed material properties (Al$_2$O$_3$/CrN), $h = 15$ nm, and $D = 2.24$. The surface roughness was varied by several orders of magnitude by adjusting the magnitude of the fractal roughness $G$. The trend is for $\xi_c$ to increase with roughness and interfacial shear strength, except in the case of adhesionless surfaces ($s/k = 0$) for which $\xi_c \approx 0.5$ regardless of roughness. For low roughness ($\sigma \approx 100$ nm), $\xi_c$ shows significant variation between adhering ($s/k > 0$) and adhesionless surfaces ($s/k = 0$). For example, increasing $s/k$ from 0 to 0.05 produces a decrease in $\xi_c$ by $\approx 45\%$. This observation is in agreement with experimental results showing a significant effect of surface adhesion on the nanotribological properties of microelectromechanical devices (Timpe and Komvopoulos, 2005, 2006).

The trend of $\xi_c$ to decrease with $s/k$ seen in Fig. 4.7(c) is similar to that shown in Fig. 4.7(a). This can be explained by plotting the three energy ratios as functions of dimensionless interfacial shear strength, as shown in Fig. 4.8. While $\xi_c$ increases with $s/k$, both $\xi_s$ and $\xi_c$ decrease monotonically, indicating that energy dissipation due to plastic deformation and material removal decreases with the increase in interfacial adhesion due to the dominance of elastic deformation at the microcontacts of surfaces exhibiting high adhesion.
4.5. Conclusions

A contact mechanics analysis of friction and energy dissipation at the microcontact level was performed for hard (rigid) and rough (fractal) surfaces in sliding contact with soft and smooth surfaces based on the slip-line theory of plasticity. The slip-line analysis yields solutions of the friction coefficient of a single fully plastic microcontact in terms of the normal load and interfacial shear strength. The analysis of contacting surfaces demonstrating multi-scale roughness provides relationships of the friction coefficient and energy dissipated in the form of plastic
deformation and material removal in terms of global interference (normal load effect), interfacial shear strength (adhesion effect), fractal parameters (roughness effect), and elastic-plastic material properties of the soft surface (deformation effect). Numerical results for representative contact systems elucidated the effects of interfacial adhesion, normal load, topography parameters of the rough surface, and material properties of the deformed surface on the friction coefficient and energy dissipation at the contact interface. Based on the presented results and discussions, the following main conclusions can be drawn from this study.

1) The deformation mode at the microcontact level is controlled by the critical area of truncated asperity, which is a function of the elastic-plastic material properties of the soft surface and the topography (roughness) of the rough surface.

2) The friction coefficient of a fully plastic microcontact is mainly due to the plowing friction component; the effect of interfacial adhesion is secondary. This difference in the contribution of the plowing and adhesion friction components to the total friction coefficient increases with the normal load.

3) For sliding contact interfaces exhibiting fractal behavior, plastic microcontacts are a small fraction (less than ~10%) of the total number of microcontacts established at the contact interface, indicating that elastic deformation is the dominant deformation mode at the microcontact level over a wide range of global interference (normal load).

4) The friction coefficient of rough surfaces increases with interfacial adhesion, surface roughness of the rough surface, and reciprocal of elastic modulus-to-hardness ratio of the soft surface, showing a weak dependence on global interference (normal load).

5) The fraction of energy dissipated in the form of plastic deformation and material removal decreases with the increasing interfacial adhesion and global interference (normal load) and increases with the surface roughness and elastic modulus-to-hardness ratio of the soft material.

6) For adhesionless contact interfaces, both friction and energy dissipation due to plastic deformation and material removal are approximately independent of normal load and surface roughness.
Chapter 5

A slip-line plasticity analysis of abrasive wear of rough (fractal) surfaces

5.1. Introduction

Abrasive wear plays an important role in mechanical components involving contact interfaces, such as gears and bearings, as well as in mechanical surfacing of materials (Hisakado et al., 2001; Izciler and Tabur, 2006). This type of wear occurs over a wide range of length scales and is the prime cause of mechanical dysfunction under various industrial settings (Eyre, 1976). In abrasive wear, relatively hard asperities and/or wear particles trapped at the contact interface act as cutting tools, removing material by a combination of microscopic plowing and microcutting mechanisms, which are of fundamental importance in many manufacturing processes, such as surface micromachining and chemical-mechanical polishing (Lin et al., 2004). Abrasive wear also plays an important role in the life and material removal efficacy of machine tools and grinding/polishing components. Therefore, insight into material removal by abrasion over a wide range of length scales is critical to the optimization of surface machining processes and the enhancement of the operation efficiency and endurance of mechanical elements possessing contact interfaces.

Early studies of abrasive wear were mainly based on Archard’s classical wear model (Archard, 1953). Traditionally, surface wear resistance has been quantified by the magnitude of the wear coefficient defined in the context of Archard’s wear equation, reflecting the effects of the applied normal load and hardness of the worn material on the material removal rate. According to classical abrasive wear theory, the wear coefficient is proportional to the average slope (sharpness) of the abrading asperities and/or wear particles (Rabinowicz, 1995). However, this relationship between abrasive wear coefficient and average slope of abrading asperities/particles is an oversimplification. A rigorous analysis of the effect of surface topography on the abrasive wear coefficient is therefore essential.

In previous studies, wear coefficient relationships were either not obtained (Hisakado, 1977; Hisakado et al., 1987) or derived empirically from experimental results (Kato, 1992; Yang, 2003) and surface description was based on scale-dependent roughness parameters. To account for the multi-scale roughness demonstrated by most engineering materials, fractal geometry, characterized by the properties of scale invariance and self-affinity, was incorporated in contact mechanics studies of rough surfaces (Majumdar and Bhushan, 1991; Wang and Komvopoulos, 1994a,b; Sahoo and Roy Chowdhury, 1996; Yan and Komvopoulos, 1998; Persson et al., 2002; Komvopoulos and Gong, 2007; Yin and Komvopoulos, 2010). However,
abrasive wear of rough surfaces exhibiting fractal behavior has received significantly less attention compared to friction and contact deformation of fractal surfaces.

Quantitative abrasive wear models based on plastic deformation induced by hard asperities/particles sliding against a soft surface have been presented in earlier studies. A common approach has been to assume rigid-perfectly plastic behavior of the abraded material, and use the slip-line theory of plasticity to analyze the removal of material by a rigid asperity/particle (Challen and Oxley, 1979). Slip-line theory has been extensively used to analyze plowing and cutting friction mechanisms (Komvopoulos et al., 1986a,b; Hokkirigawa and Kato, 1988; Kopalinsky and Oxley, 1995). In slip-line models of plowing friction, a hard asperity/particle modeled as a rigid wedge pushes a wave of plastically deformed material ahead as it plows through the soft surface (Komvopoulos et al., 1986a; Kopalinsky and Oxley, 1995; Bressan and Williams, 2009). Alternatively, in slip-line fields of friction involving material removal by a cutting process, the plastically deformed material slides against the rigid wedge, resulting in the formation of a chip, which can be used to estimate the material removal rate (Komvopoulos et al., 1986b; Hokkirigawa and Kato, 1988; Kopalinsky and Oxley, 1995). However, integration of single-asperity/particle slip-line models into fractal mechanics analysis of abrasive wear to determine the abrasive wear rate and wear coefficient of sliding rough surfaces has not been presented yet.

The main objective of the present study was to develop a comprehensive plasticity analysis of the abrasive wear process occurring at sliding contact interfaces that exhibit multi-scale roughness. Instead of relying on empirical approaches based on experimental measurements, a three-dimensional abrasive wear model of a hard (relatively rigid) fractal surface plowing through a soft and smooth surface was developed using the slip-line theory of plasticity. Results of the single-contact model are presented first to elucidate the dependence of local deformation (microcontact scale) on normal load and interfacial adhesion. This model is then incorporated into a three-dimensional plasticity analysis of abrasive wear of rough (fractal) surfaces, and the abrasive wear rate and wear coefficient are obtained in terms of the total normal load (global interference), surface topography (fractal) parameters of the rough surface, elastic-plastic material properties of the worn surface, and interfacial friction (adhesion) characteristics. In addition, numerical results for typical ceramic/ceramic, ceramic/metallic, and metal/metal sliding contact interfaces are presented to illustrate the effects of normal load, roughness, material properties, and interfacial adhesion on abrasive wear of rough surfaces.

5.2. Single-contact plasticity analysis

Sliding contact of rough surfaces occurs between surface asperities of varying sizes, exhibiting elastic or elastic-plastic deformation. A generally accepted wear model is that of sliding surfaces undergoing plowing and microcutting by hard
asperities and/or wear particles trapped at the contact interface. Since wear implies the removal of material, only elastic-plastic asperity contacts where microcutting is the dominant deformation mode contribute to wear. Because most surfaces exhibit multi-scale roughness, understanding abrasive wear at the asperity level is fundamental for developing a generalized abrasive wear model of real rough surfaces. Therefore, a single-asperity contact model of the microcutting process based on the slip-line theory of plasticity is presented first in this section.

In previous slip-line models, the abrasive (or cutting) body was assumed to be rigid and perfectly sharp (Challen and Oxley, 1979; Petryk, 1987; Lacey and Torrance, 1991; Kopalinsky and Oxley, 1995; Torrance and Buckley, 1996). However, even macroscopically sharp asperities and acicular wear particles demonstrate finite local curvature. This motivated Tsukizoe and Sakamoto (1975) to approximate the local geometry of asperities and wear particles by a spherical shape. Therefore, it is necessary to use a slip-line field of locally smooth asperities (or wear particles) plowing through a soft surface. However, such a slip-line analysis has not been reported yet. Although Challen and Oxley (1983) considered a slip-line model of a cylindrical hard asperity, the contact arc between the asperity and the deformed material was approximated by its chord to enable the slip-line field developed by Challen and Oxley (1979) for wedge-shaped asperities to be used in the former analysis. More recently, Busquet and Torrance (2000) presented a slip-line analysis of a sliding rigid cylinder, but because their model does not account for cutting (chip formation), it cannot be used to model abrasive wear.

In the present analysis, it is presumed that material removal occurs by a microcutting process encountered at various asperity (wear particle) length scales. The cutting edge of the hard asperity (wear particle) is assumed to have a radius of curvature $R$ and to penetrate the soft material to depth $d_s$. For shallow plowing (i.e., $d_s/R < 0.1$), deformation of the soft surface may be approximated by the plane-strain condition. Material removal involves intense shearing of the soft material ahead of the hard (rigid) asperity, resulting in material sliding against the asperity surface at a relative velocity $u_c$, as shown in Fig. 5.1(a) (Komvopoulos, 2011). The soft material is assumed to be isotropic and rigid-perfectly plastic and to approach the rigid asperity at constant velocity $u$.

The deformation field of the soft material is represented by a network of orthogonal characteristic lines, known as $\alpha$ and $\beta$ shear lines (Fig. 5.1(b)). Since separation of the material sliding against the hard asperity occurs at point B, boundary BD is a free-surface (i.e., both normal and tangential tractions are zero) and, consequently, the $\alpha$- and $\beta$-lines intersect boundary BD at an angle of 45º. The angle $\eta$ between an $\alpha$-line and the tangent at any point of interface BG depends on the ratio

69
Fig. 5.1 (a) Slip-line field of a hard spherical asperity (or wear particle) removing material from a soft surface by a microcutting process, (b) detailed view of plastic region showing the network of orthogonal $\alpha$ and $\beta$ slip-lines, and (c) hodograph of the slip-line field (Komvopoulos, 2011).
of the interfacial shear strength \( s \) to the shear strength of the soft material \( k \), i.e.,
\[
\eta = \frac{1}{2} \cos^{-1}\left( \frac{s}{k} \right)
\]  
(5.1)
where \( 0 \leq s/k \leq 1 \).

For a slip-line field to be admissible, it must satisfy mass conservation, certain geometric/trigonometric relationships, and kinematic constraints. Since the soft material is assumed to be perfectly plastic (incompressible), mass conservation requires that the material removal (wear volume) rate \( \Delta \dot{V} \) (per unit length in the out-of-plane direction) at the asperity/wear particle scale is given by
\[
\Delta \dot{V} = l \cdot u_c = d_s \cdot u = d_s \cdot \left( \frac{S}{T} \right)
\]  
(5.2)
where \( l \) is the thickness of the removed material, \( S \) is the total distance of sliding, and \( T \) is the corresponding time of sliding.

Geometry considerations yield the following equations:
\[
R \cos \phi + t + d_s = R
\]  
(5.3)
\[
t = \frac{l \sin \left( \frac{\pi}{4} + \beta \right)}{\cos \left( \frac{\pi}{4} - \eta \right)}
\]  
(5.4)
\[
\beta + \phi + \eta = \frac{\pi}{2}
\]  
(5.5)
where \( t \) is the pile-up thickness, and slip-line angles \( \beta \) and \( \phi \) are shown in Fig. 5.1(b).

For fixed dimensionless penetration depth \( d_s/R \) and interfacial shear strength \( s/k \), the slip-line field is defined by Eqs. (5.3)-(5.5). The hodograph of this slip-line field is shown in Fig. 5.1(c), where \( u_{DI} \) is the relative velocity along boundary DI, and angle \( \theta \) is shown in Fig. 5.1(b). Kinematic admissibility requires that the following velocity condition is satisfied in region BDF consisting of a network of straight \( \alpha \)- and \( \beta \)-lines:
\[
u_{\alpha B}^2 + v_{\beta B}^2 = u_c^2 \sin^2 \phi + (u_c \cos \phi + u)^2
\]  
(5.6)
where \( u_{\alpha B} \) and \( u_{\beta B} \) are the velocities at point B along the \( \alpha \)- and \( \beta \)-lines, respectively, determined from a numerical scheme described elsewhere (Komvopoulos, 2011).
Assuming a dimensionless velocity \( u_c/u \), the slip-line field geometry was determined by solving numerically Eqs. (5.3)-(5.5). If the obtained solution did not satisfy the velocity condition given by Eq. (5.6), a different value of \( u_c/u \) was assumed, and the numerical procedure was repeated until a kinematically admissible slip-line field was obtained.

### 5.3. Rough surface contact analysis

#### 5.3.1. Contact model

The hard and rough surface is assumed to be isotropic and self-affine, having a three-dimensional (3D) surface profile (Yan and Komvopoulos, 1998):

\[
\begin{align*}
 z(x, y) &= L \left( \frac{G}{L} \right)^{(D-2)} \left( \frac{\ln \gamma}{M} \right)^{1/2} \sum_{m=1}^{M} \sum_{q=0}^{q_{\text{max}}} \gamma^{(D-3)q} \times \\
 & \quad \left\{ \cos \phi_{m,q} - \cos \left[ \frac{2\pi q^q (x^2 + y^2)^{1/2}}{L} \cos \left( \tan^{-1} \left( \frac{y}{x} \right) - \frac{\pi m}{M} + \phi_{m,q} \right) \right] \right\}
\end{align*}
\] (5.7)

where \( L \) is the profile length, \( D \) and \( G \) are the fractal dimension (2\(<D<3\)) and fractal roughness, respectively, both independent of spatial frequency in the range where the surface exhibits fractal behavior, \( \gamma (\gamma > 1) \) is a parameter that controls the density of frequencies in the surface profile, with a typical value of 1.5 (Komvopoulos and Yan, 1997), \( M \) is the number of superimposed ridges, \( q \) is a spatial frequency index with lower and upper limits \( q_{\text{min}} = 0 \) and \( q_{\text{max}} = \text{int} \left[ \log(L/L_s)/\log \gamma \right] \), respectively, and \( \phi_{m,q} \) is a random phase uniformly distributed in the range \([0, 2\pi]\) used to prevent the coincidence of different frequencies at any point of the surface profile.

Fractal parameter \( D \) determines the relative contributions of high- and low-frequency components in the surface profile, while fractal roughness \( G \) controls the wave amplitudes through the entire frequency range of the surface profile. The surface roughness increases with the decrease of \( D \) and the increase of \( G \). Because of the scale invariance of \( D \) and \( G \), fractal surface description at different length scales is independent of the scale of measurement. The smallest wavelength, corresponding to the instrument resolution limit, is set above the cutoff length \( L_s \) (typically a few times above the lattice distance) in order for continuum description to hold, while the largest wavelength is set equal to the profile length \( L \). Hence, the highest and lowest frequencies in the surface profile are \( \omega_n = 1/L_s \) and \( \omega_l = 1/L \), respectively. Figure 5.2 shows a typical 3D fractal surfaces constructed from Eq. (5.7), which was used to generate the rough (fractal) surfaces used in the elastic-plastic contact analysis presented below.
Fig. 5.2 Three-dimensional fractal surface generated from Eq. (5.7) for $D = 2.24$, $G = 2.39 \times 10^{-4}$ nm, $\gamma = 1.5$, $M = 10$, $L = 7.04 \times 10^3$ nm, and $L_s = 5$ nm ($\sigma = 290$ nm).

Since the abrading rough surface is assumed to be isotropic, any two-dimensional (2D) surface profile is statistically representative of the 3D surface topography. A 2D profile of a fractal surface (Eq. (5.7)) truncated by a plane is shown schematically in Fig. 5.3. Displacement of the plane toward the rough surface by global interference $h$ produces several truncated profile segments. Each truncated segment is approximated by an asperity with a spherical cap shape having a base area equal to the truncated area $a'$ and height equal to the local interference $\delta$ given by (Yan and Komvopoulos, 1998)

$$\delta = \frac{(a')^{(3-D)/2}}{2^{(D-4)\pi(3-D)/2}G^{(2-D)/2}(\ln \gamma)^{-1/2}} \quad (5.8)$$

The radius of curvature of the spherical asperity $R$ is given by (Yan and Komvopoulos, 1998)

$$R = \frac{(a')^{(D-1)/2}}{2^{(5-D)\pi(D-1)/2}G^{(D-2)/2}(\ln \gamma)^{1/2}} \quad (5.9)$$

The size distribution of truncated asperities follows an island-like distribution similar to that observed in geophysics (Mandelbrot, 1975, 1983) that has been used in previous contact mechanics studies of fractal surfaces (Komvopoulos and Ye, 2001; Majumdar and Bhushan, 1991; Wang and Komvopoulos, 1994a, 1994b; Yan and
Komvopoulos, 1998). According to this island-like distribution, truncated asperities follow the power-law relationship

\[ \frac{N(a')}{N(a''_L)} = \left( \frac{a''_L}{a'} \right)^{(D-1)/2} \]  

(5.10)

where \( N(a') \) is the number of asperities with truncated areas greater than \( a' \), and \( a''_L \) is the largest area of truncated asperity at a given global interference. The density function of truncated asperities \( n(a') \) of a 3D surface profile is given by (Yan and Komvopoulos, 1998)

\[ n(a') = -\frac{dN(a')}{d a'} = \frac{(D - 1)}{2a''_L} \left( \frac{a''_L}{a'} \right)^{(D+1)/2} \]  

(5.11)

For a given global interference \( h \), the largest area of truncated asperity \( a''_L \) can be determined from the total area of truncated asperities \( S' \) of the rough surface using the relationship (Komvopoulos and Ye, 2001):

\[ S' = \int_{a''_L}^{a'_L} a' n(a') da' \]  

(5.12)

where \( a'_s \) is the smallest area of truncated asperity, which is set larger than the lattice dimension for a continuum description to hold. For instance, the diameter of the smallest truncated asperity contact can be set equal to \(~5-6\) times the lattice dimension of the abraded soft material.

Substitution of Eq. (5.11) into Eq. (5.12) and integration gives

\[ S' = \left( \frac{D - 1}{3 - D} \right) \left[ 1 - \left( \frac{a'_s}{a''_L} \right)^{(3-D)/2} \right] a''_L \]  

(5.13)

At a given global interference, \( S' \) was first calculated numerically by summing up the contact areas of all truncated asperities on the rough surface (Komvopoulos and Ye, 2001), \( a'_L \) was then obtained from Eq. (5.13) as a function of \( a'_s, D, \) and \( S' \), and, finally, area range \([a'_s,a'_L]\) and spatial distribution of the truncated contacting asperities were determined from Eqs. (5.11) and (5.13).

Before presenting the slip-line analysis of the abrasive wear process, it is instructive to consider deformation at the asperity level. Yan and Komvopoulos (1998) derived a relationship of the critical contact area of truncated asperity
Fig. 5.3 Two-dimensional profile of a truncated rough (fractal) surface.

demarking the transition from elastic to fully plastic deformation. In the current study, the hard and rough surface is assumed to be rigid, while the soft and smooth substrate is modeled as elastic-perfectly plastic, implying either elastic or fully plastic deformation at the microcontact level, depending on the local interference and truncated asperity contact area. Microcontact deformation is modeled to be either elastic \( a' > a'_c \) or fully plastic \( a' \leq a'_c \), depending on the critical area of truncated asperity \( a'_c \) given by (Yan and Komvopoulos, 1998)

\[
a'_c = \frac{2^{(11-2D)}}{9\pi^{(4-D)}} G^{(2D-4)} \left( \frac{E}{H} \right)^2 \ln \gamma \right]^{1/(D-2)}
\]  

where \( E \) and \( H \) are the elastic modulus and hardness of the soft surface, respectively. Equation (5.14) shows that \( a'_c \) depends on the elastic and plastic properties of the soft surface \((E \text{ and } H)\) and the fractal parameters of the hard and rough countersurface \((D \text{ and } G)\).

The contact interface comprises elastic and fully plastic microcontacts \((a'_e > a'_c)\) or only fully plastic microcontacts \((a'_e \leq a'_c)\). Since wear implies material removal due to irreversible deformation, it may be inferred that only fully plastic microcontacts have the potential to contribute to the formation of wear debris. Therefore, a wear criterion must be introduced for fully plastic microcontacts.
Experimental studies (Hokkirigawa and Kato, 1988; Kato, 1992) suggest that wave formation (plowing), wedge formation, and cutting are dominant deformation modes at the asperity level, and that the occurrence of each mode depends on the penetration depth of the hard asperity/wear particle and the shear strength at the contact interface (interfacial adhesion effect). Since abrasive wear implies material removal by a microcutting process, the experimental data of the former studies were used to obtain a wear criterion for fully plastic microcontacts. Figure 5.4 shows a deformation map illustrating the dependence of cutting and plowing/wedge formation modes on the degree of penetration $D_p$ and interfacial adhesion, represented by $s/k$. The boundary between these two deformation regimes was obtained by curve fitting through experimental data (Hokkirigawa and Kato, 1988) corresponding to the transition from plowing/wedge formation to cutting. Thus, in the present analysis, only those fully plastic microcontacts with $D_p$ values above the curve shown in Fig. 5.4 were assumed to contribute to the removal of material. Using the equation of the fitted curve, a critical contact area of truncated asperity $a'_w$ demarcating the transition from plowing/wedge formation (no wear) to cutting (wear) was determined, as described in the Sect. 3.2. Thus, only fully plastic microcontacts ($a' \leq a'_c$) with $a' \leq a'_w$ were
assumed to contribute to abrasive wear.

5.3.2. Wear rate

To determine $a'_w$, it is assumed that the normal load carried by each microcontact during sliding is equal to that under indentation loading. In this situation, incompressibility condition yields

$$a' = a_n = a_s$$

where $a_n$ is the contact area of a single spherical asperity resulting from pure normal loading (indentation), and $a_s$ is the contact area of the same asperity during sliding contact. From Eq. (5.15) and the fact that contact during sliding is confined over the front half of the spherical asperity (Fig. 5.1(a)), it follows that

$$a_n = \pi r_n^2 = \pi R^2 / 2 = a_s \rightarrow \sqrt{2} r_n = r_s$$

where $r_n$ is the contact radius due to indentation and $r_s$ is the contact radius of a sliding spherical asperity (Fig. 5.1(a)).

From Eqs. (5.1)-(5.5) and (5.16) and relationship $r_s = R \sin \phi$ (Fig. 5.1(a)), it follows that

$$d_s = f\left(a_s, R, \frac{s}{k}\right)$$

Moreover, because of Eq. (5.15), Eq. (5.17) can be further expressed as

$$d_s = f\left(a', R, \frac{s}{k}\right)$$

The degree of penetration $D_p$ is given by (Hokkirigawa and Kato, 1988)

$$D_p = 0.8 \left(\frac{d}{r_s}\right)$$

In view of Eqs. (5.15), (5.16), (5.18), and (5.19), $D_p$ can be obtained as

$$D_p = f\left(a', R, \frac{s}{k}\right)$$

Equation (5.20) and deformation map shown in Fig. 5.4 indicate that at the asperity/wear particle (microcontact) scale,
Substituting Eq. (5.9) into Eq. (5.20) yields

\[ D_p = f\left(D, G, \frac{s}{k}, a'\right) \]  

(5.22)

From Eq. (5.22) and Fig. 5.4, it follows that for a rough (fractal) surface,

\[ a_w' = f\left(D, G, \frac{s}{k}\right) \]  

(5.23)

Equation (5.23) shows that the critical contact area of truncated asperity for material removal (wear) depends on the topography parameters of the rough surface and the interfacial friction condition controlled by the adhesion characteristics of the sliding surfaces.

The abrasive wear rate is defined as the ratio of the total wear volume generated from all fully plastic microcontacts for a total distance of sliding \( S \). Depending on \( a_c' \), \( a_w' \), and \( a_L' \), the abrasive wear rate \( V/S \) can be expressed as following:

(a) For \( a_L' > a_c' > a_w' \),

\[ \frac{V}{S} = \int_{a'_w}^{a'_L} \frac{\Delta V(a')}{S} n(a') da' \]  

(5.24a)

(b) For \( a_L' > a_w' \geq a_c' \),

\[ \frac{V}{S} = \int_{a'_w}^{a'_L} \frac{\Delta V(a')}{S} n(a') da' \]  

(5.24b)

where \( \Delta V/S \) is the wear volume per unit sliding distance at the microcontact level, obtained from Eq. (5.2) as

\[ \Delta V(a')/S = d_s(a') \]  

(5.25)

From Eqs. (5.9), (5.11), (5.13), (5.14), (5.18), and (5.23)-(5.25), it follows that

\[ \frac{V}{S} = f\left(D, G, \frac{s}{k}, E, \frac{k}{H}, P\right) \]  

(5.26)

where \( P \) is the total normal load.
Equation (5.26) indicates that the abrasive wear rate depends on the elastic and plastic properties of the abraded surface, the dimensionless interfacial shear strength (controlled by the adhesion of the interacting surfaces), the topography of the abrading rough surface, and (implicitly) the total normal load $P$ through the total area of truncated asperities $S'$, which determines the magnitude of $a'_L$ (Eq. (5.13)).

5.3.3. Wear coefficient

The normal load carried by a single fully plastic microcontact can be expressed as $\Delta P = Ha_s$. In addition, the hardness and shear strength of the abraded surface are related by $H = 6k$. Therefore,

$$\Delta P = 6ka_s \quad (5.27)$$

Based on the classical definition of the wear coefficient (Archard, 1953), the wear coefficient of a single asperity contact $\Delta K$ can be expressed as

$$\Delta K(a') = \frac{\Delta V(a')}{S} \cdot \frac{H}{\Delta P(a')} \quad (5.28)$$

Substituting Eqs. (5.15) and (5.25) and relationship $\Delta P = Ha_s$ into Eq. (5.28) yields

$$\Delta K(a') = d_s(a') / a' \quad (5.29)$$

The total wear coefficient $K$ is defined as the statistical average of the wear coefficients of all microcontacts established at the contact interface. Hence, depending on $a'_C$, $a'_w$, and $a'_L$, the abrasive wear coefficient $K$ can be obtained as following:

(a) For $a'_L > a'_C > a'_w$,

$$K = \frac{\int_{a'_w}^{a'_L} \Delta K(a')n(a')da'}{\int_{a'_w}^{a'_L} n(a')da'} \quad (5.30a)$$

(b) For $a'_L > a'_w \geq a'_C$,

$$K = \frac{\int_{a'_w}^{a'_C} \Delta K(a')n(a')da'}{\int_{a'_w}^{a'_C} n(a')da'} \quad (5.30b)$$
From Eqs. (5.9), (5.11), (5.13), (5.14), (5.18), (5.23), (5.29), and (5.30), it follows that

\[
K = f \left( D, G, \frac{s}{k}, \frac{E}{H}, P \right)
\]  

(5.31)

Similar to the wear rate, Eq. (5.31) indicates that the abrasive wear coefficient depends on the topography of the rough surface (through fractal parameters \(D\) and \(G\)), the interfacial shear strength \(s/k\) (affected by the adhesion of the sliding surfaces), the elastic and plastic material properties of the abraded surface \((E\) and \(H))\), and indirectly on the total normal through the total area of truncated asperities \(S'\), which controls the magnitude of \(a'_L\) (Eq. (5.13)).

5.4. Results and discussion

5.4.1. Deformation mode at a single asperity contact

Before presenting rough surface results, the deformation behavior of a single microcontact under cutting conditions is examined in this section in the context of slip-line plasticity solutions. The ratio of the penetration depth during sliding \(d_s\) to the indentation depth \(d_n\) can be used to examine the dependence of the deformation mode during abrasion on the applied normal load. For a spherical asperity of radius \(R\) penetrating a half-space to depth \(d_n\),

\[
R^2 = r_n^2 + (R - d_n)^2
\]

(5.32)

From Eqs. (5.15), (5.16), and (5.32), it is found that

\[
d_n = R - \left( R^2 - \frac{d'_s}{\pi} \right)^{1/2}
\]

(5.33)

Equation (5.33) indicates that \(d_n\) depends on \(R\) and \(a'_s\), but not on \(s/k\).

From Eqs. (5.15), (5.17), (5.27), and (5.33), it follows that \(d_s/d_n\) is an implicit function of the asperity radius \(R\), dimensionless normal load \(P/kR^2\), and interfacial shear strength \(s/k\), i.e.,

\[
\frac{d_s}{d_n} = f \left( R, \frac{P}{kR^2}, \frac{s}{k} \right)
\]

(5.34)

Since a closed-form solution for \(d_s/d_n\) cannot be obtained from Eq. (5.34), numerical results are presented in Fig. 5.5. The dashed line and solutions for \(s/k = 0\) and 1.0 define the regime where the present slip-line model is admissible. Outside this region slip-line models representative of other friction mechanisms, such as plowing and wedge formation, may be applicable. The analysis yields \(d_s/d_n > 1\) through the
Fig. 5.5 Ratio of penetration depth during sliding to indentation depth $d_s/d_n$ versus dimensionless normal load $P/(kR^2)$ and interfacial shear strength $s/k$ for a single asperity microcontact. The dashed curve and solutions for $s/k = 0$ and 1 define the regime of admissible slip-line solutions.

entire load range $P/(kR^2)$ where the slip-line is admissible. This trend becomes more pronounced with increasing normal load and decreasing interfacial shear strength, implying higher loads for inducing abrasive wear between strongly adhering surfaces, in accord with experimental observations (Hokkirigawa and Kato, 1988).

Further insight into the sliding conditions that are conducive to abrasive wear can be obtained by considering the dependence of the dimensionless pile-up thickness $t/R$ on dimensionless normal load $P/kR^2$ and interfacial shear strength $s/k$. Equations (5.1)-(5.5) give that $t/R$ is an implicit function of $d_s$, $R$, and $s/k$, i.e.,

$$
\frac{t}{R} = f\left(d_s, R, \frac{s}{k}\right) 
$$

(5.35)

Consequently, Eqs. (5.17), (5.27), and (5.35) give that

$$
\frac{t}{R} = f\left(R, \frac{P}{kR^2}, \frac{s}{k}\right) 
$$

(5.36)

Figure 5.6(a) shows that the pile-up thickness $t$ at microcontacts resulting in abrasive wear is less than 10% of the asperity radius of curvature $R$, showing a trend to increase with normal load and interfacial shear strength. The dashed curve
Fig. 5.6 (a) Pile-up thickness-to-asperity radius ratio $t/R$ and (b) pile-up thickness-to-penetration depth ratio $t/d_s$ versus dimensionless normal load $P/(kR^2)$ and interfacial shear strength $s/k$ for a single asperity microcontact. The dashed curve and solutions for $s/k = 0$ and $1$ define the regime of admissible slip-line solutions.

(determined from Eqs. (5.15), (5.21), and (5.27)) and solutions for $s/k = 0$ and $1$) define the range of the slip-line solutions of $t/R$. Figure 5.6(b) shows the variation of the pile-up thickness-to-penetration depth ratio $t/d_s$ with the dimensionless normal
5.4.2. Wear rate

Because it is not possible to obtain a closed-form solution for \( \frac{d_i}{h_i} \) (Eq. (5.23)) and, in turn, determine the deformation mode at each microcontact, numerical results of the wear rate and wear coefficient (Eqs. (5.26) and (5.31), respectively) were obtained.

### Table 5.1 Elastic-plastic material properties of contact systems

<table>
<thead>
<tr>
<th>Contact system</th>
<th>Material</th>
<th>Properties(^{(a)})</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>( E ) (GPa)</td>
<td>( H ) (GPa)</td>
<td>( E/H )</td>
<td></td>
</tr>
<tr>
<td>Al(_2)O(_3)/CrN</td>
<td>Al(_2)O _3</td>
<td>307</td>
<td>27.6</td>
<td>/</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CrN</td>
<td>103</td>
<td>14.8</td>
<td>6.96</td>
<td></td>
</tr>
<tr>
<td>Al(_2)O(_3)/TiC</td>
<td>Al(_2)O _3</td>
<td>307</td>
<td>27.6</td>
<td>/</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TiC</td>
<td>450</td>
<td>23.5</td>
<td>19.2</td>
<td></td>
</tr>
<tr>
<td>AISI 1095/AISI 1020</td>
<td>AISI 1095</td>
<td>200</td>
<td>6.08</td>
<td>/</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AISI 1020</td>
<td>200</td>
<td>1.71</td>
<td>117</td>
<td></td>
</tr>
</tbody>
</table>

\(^{(a)}\)Sources: Suh (1986); Komvopoulos and Zhang (2008)

in terms of global interference \( h \), root-mean-square surface roughness \( \sigma \) and fractal dimension \( D \) of the rough surface, and material properties \( E \) and \( H \) of the abraded surface. Table 5.1 gives the material properties of three representative contact systems: ceramic-ceramic (Al\(_2\)O\(_3\)/TiC), ceramic-metallic (Al\(_2\)O\(_3\)/CrN), and metal-metal (AISI
1095 steel/AISI 1020 steel). The numerical results presented in this section are for a 3D surface topography generated from Eq. (5.7) for $\gamma = 1.5$, $M = 10$, $L_s = 5$ nm, $L = 7.04 \times 10^3$ nm, and $\phi_{m,q} = \pi / 2$. In all simulation cases, $a'_s$ was obtained for $L_s = 5$ nm, i.e., $a'_s = \pi (r'_s)^2 = \pi (L_s / 2)^2 = 19.6 \text{nm}^2$.

Figure 5.7(a) shows the wear rate $V/S$ as a function of global interference $h$ and dimensionless interfacial shear strength $s/k$ for Al$_2$O$_3$/CrN contact system with fixed surface topography parameters ($D = 2.24$, $G = 1.07 \times 10^{-4}$ nm, and $\sigma = 100$ nm). The wear rate increases linearly with the global interference and with the decrease of $s/k$, suggesting that higher wear (material removal) rates can be achieved by reducing the interfacial shear strength (e.g., lubrication), a well-known trend in surface machining.

The dependence of the wear rate $V/S$ on surface roughness $\sigma$ can be interpreted in the context of the results shown in Fig. 5.7(b) for Al$_2$O$_3$/CrN contact system and fixed global interference ($h = 30$ nm). The roughness was varied in the range of 10–1000 nm by varying the fractal roughness $G$, while keeping the fractal dimension fixed ($D = 2.24$). The wear rate increases with increasing roughness and decreasing interfacial adhesion. The roughness effect is more significant in the low-roughness range ($\sigma = 10$–100 nm), while that of interfacial adhesion is significant only in the high-roughness range ($\sigma = 100$–1000 nm).

The effect of fractal dimension $D$ (or surface roughness $\sigma$) on wear rate $V/S$ is illustrated in Fig. 5.7(c) for Al$_2$O$_3$/CrN, Al$_2$O$_3$/TiC, and AISI 1095/AISI 1020 contact systems and fixed global interference ($h = 30$ nm), interfacial shear strength ($s/k = 0.5$), and fractal roughness ($G = 2.39 \times 10^{-4}$ nm). The decreasing trend of $V/S$ with increasing $D$ can be explained by considering the variation of $a'_c$ corresponding to the contact systems of Fig. 5.7(c) with $D$ shown in Fig. 5.8. All contact systems show an increase in $V/S$ with decreasing $D$, i.e., $V/S$ increases with the surface roughness. For fixed surface topography, $V/S$ depends on the ratio of the elastic modulus-to-hardness ratio $E/H$ of the softer (abraded) sliding surface (Table 5.1). This finding is supported by the fact that, for fixed topography parameters, $a'_c$ depends only on $E/H$ (Eq. (5.14)). The effect of material properties becomes significant for low $D$ values, i.e., in the high-roughness range $\sigma > 10^5$ nm.
Fig. 5.7 (a) Wear rate $V/S$ of $\text{Al}_2\text{O}_3$/CrN contact system versus global interference $h$ and dimensionless interfacial shear strength $s/k$ for $D = 2.24$, $G = 1.07 \times 10^{-4}$ nm, and $\sigma = 100$ nm; (b) Wear rate $V/S$ of $\text{Al}_2\text{O}_3$/CrN contact system versus surface roughness $\sigma$ and dimensionless interfacial shear strength $s/k$ for $D = 2.24$ and $h = 30$ nm; (c) Wear rate $V/S$ of $\text{Al}_2\text{O}_3$/CrN, $\text{Al}_2\text{O}_3$/TiC, and AISI 1095/AISI 1020 contact systems versus fractal dimension $D$ and surface roughness $\sigma$ for $G = 2.39 \times 10^{-4}$ nm, $h = 30$ nm, and $s/k = 0.5$. (Elastic-plastic material properties of each contact system are given in Table 5.1.)

Fig. 5.8 Critical contact area of truncated asperity for fully plastic microcontact deformation $a'_c$ of $\text{Al}_2\text{O}_3$/CrN, $\text{Al}_2\text{O}_3$/TiC, and AISI 1095/AISI 1020 contact systems versus fractal dimension $D$ and surface roughness $\sigma$ for $G = 2.39 \times 10^{-4}$ nm. (Elastic-plastic material properties of each contact system are given in Table 5.1.)

5.4.3 Wear coefficient

Figure 5.9(a) shows the wear coefficient $K$ as a function of global interference $h$ and interfacial adhesion $s/k$ for $\text{Al}_2\text{O}_3$/CrN contact system with fixed surface topography parameters ($D = 2.24$, $G = 1.07 \times 10^{-4}$ nm, and $\sigma = 100$ nm). The wear coefficient is on the order of $10^{-2}$, which is typical of two-body abrasive wear (Rabinowicz, 1995). The trend for $K$ to decrease with increasing $h$ is observed with all
interfacial shear strength conditions. The higher wear coefficients obtained with lower $s/k$ values indicate that reducing interfacial adhesion (e.g., by lubricating the sliding surfaces) enhances abrasive wear much more than decreasing the global interference (i.e., increasing the normal load). This is attributed to less energy loss to combat friction in well-lubricated or low-adhesion contact interfaces. Figure 5.9(a) suggests that light loads and low-adhesion conditions promote higher material removal rates, which is of particular interest in polishing and grinding processes.

The effect of surface roughness $\sigma$ on wear coefficient $K$ can be interpreted in light of results for Al$_2$O$_3$/CrN contact system and fixed global interference ($h = 30$ nm) shown in Fig. 5.9(b). Surface roughness was varied by two orders of magnitude (10–1000 nm) by changing the fractal roughness $G$ while keeping the fractal dimension fixed ($D = 2.24$). These results reveal a significantly more pronounced roughness effect on wear than interfacial adhesion. The trend for the wear coefficient to increase with surface roughness is in good agreement with experimental observations. Moreover, the wear coefficient increases slower in the low-roughness range (10–100 nm) than in the high-roughness range (100–1000 nm), which is also in agreement with differences in the material removal rate under polishing and grinding conditions, respectively.

Figure 5.9(c) shows a comparison of the abrasive wear coefficient $K$ of Al$_2$O$_3$/CrN, Al$_2$O$_3$/TiC, and AISI 1095/AISI 1020 contact systems for different values of fractal dimension $D$ (or surface roughness $\sigma$) and fixed fractal roughness ($G = 2.39 \times 10^{-4}$ nm), interfacial shear strength ($s/k = 0.5$), and global interference ($h = 30$ nm). The variation of $K$ with $D$ and $\sigma$ is similar to that of $V/S$ (Fig. 5.7(c)) and $a_c'$ (Fig. 5.8). The ceramic-metallic contact system exhibits the lowest wear coefficient through the entire range of fractal dimension (surface roughness), indicating higher abrasive wear resistance compared to the ceramic-ceramic and metal-metal contact systems.

5.5. Conclusions

A contact mechanics analysis of abrasive wear of a hard rough (fractal) surface sliding against a soft and smooth surface was derived based on the slip-line theory of plasticity. The slip-line analysis provides insight into the deformation mode at the asperity microcontact level in terms of the applied normal load and interfacial shear strength. Material removal (wear) was presumed to originate from fully plastic microcontacts by a microcutting process. Numerical solutions of the abrasive wear rate and wear coefficient were obtained in terms of interfacial shear strength (adhesion effect), fractal parameters of the rough surface (roughness effect), elastic-plastic material properties of the worn surface (deformation effect), and applied normal load. Numerical results for representative contact systems revealed the effects of normal
load, interfacial adhesion, topography of rough surface, and material properties of the abraded surface on the abrasive wear rate and wear coefficient. Based on the presented results and discussions, the following main conclusions can be drawn from the present analysis:

1) Material removal at the asperity/wear particle scale depends on the critical area of truncated asperity, which depends on the elastic-plastic material properties of the worn surface, topography parameters (roughness) of the abrading rough surface, and interfacial shear strength (adhesion characteristics) of the sliding surfaces.

2) The penetration depth of a sliding microcontact is always greater than the indentation depth of the same microcontact obtained under the same normal load. This effect becomes more pronounced with increasing normal load and decreasing interfacial shear strength.

3) The wear rate increases sharply with the roughness of the abrading surface and decreases with intensifying interfacial shear strength.

4) The wear coefficient decreases with the increase in global interference (normal load) and moderately with the increase in interfacial shear strength, whereas it increases rapidly with increasing roughness of the abrading surface.

5) Both the abrasive wear rate and the wear coefficient depend on the elastic modulus-to-hardness ratio of the abraded surface.

6) The abrasive wear coefficient is on the order of $10^{-2}$, depending on the global interference (normal load effect), topography of abrading surface (roughness effect), material properties of worn surface (deformation effect), and interfacial shear strength (adhesion effect) of the sliding surfaces.
Chapter 6

An adhesive wear model of fractal surfaces in normal contact

6.1. Introduction

Wear plays an important role in many fields of science and technology. The implications of wear can be either beneficial or detrimental to the performance of scientific instruments and engineering components possessing contact interfaces. Since the seminal study of adhesive wear by Archard (1953), several wear mechanisms have been proposed to explain the loss of material from sliding surfaces, including abrasion, corrosion, erosion, contact fatigue, and delamination (Kruschov, 1957; Suh, 1973, 1986). Among various wear mechanisms, adhesive wear is the most common process of material removal encountered over a wide range of length scales. This type of wear is responsible for the failure of many mechanical and electromechanical components whose functionality depends on the tribological properties of contact interfaces. Thus, accurate prediction of the adhesive wear rate in tribological systems is of great technological and scientific importance.

Significant research effort has been devoted to study the dependence of adhesive wear on various factors, such as normal load, sliding speed, interfacial adhesion/friction conditions, and material properties (Lisowski and Stolarski, 1981; Finkin, 1972; Paretkar et al., 1996; Yang, 2003a). Archard’s wear model has been used extensively to quantify the wear rate of sliding surfaces (Qureshi and Sheikh, 1997; Yang, 2004), develop adhesion models of single-asperity junctions (Rabinowicz, 1980), and perform energy-based analyses of adhering asperities (Warren and Wert, 1990). However, the majority of relationships between adhesive wear rate, sliding speed, and contact area reported in early studies were based on semi-empirical approaches and statistical topography parameters (e.g., mean and variance of the surface heights, slopes, and curvatures) that do not account for the scale dependence of topography parameters, a characteristic feature of multi-scale roughness of engineering surfaces.

To overcome shortcomings with scale-dependent statistical surface parameters (Greenwood and Williamson, 1966) and random process theory (Nayak, 1973) commonly used in contact mechanics, the surface topography in contemporary contact analyses was described by fractal geometry (Majumdar and Bhushan, 1990, 1991; Wang and Komvopoulos, 1994a, 1994b, 1995; Sahoo and Roy Chowdhury, 1996; Komvopoulos and Yan, 1998; Borri-Brunetto et al., 1999; Ciavarella et al., 2000; Komvopoulos and Ye, 2001; Persson et al., 2002; Yang and Komvopoulos, 2005; Gong and Komvopoulos, 2005a, 2005b; Komvopoulos and Yang, 2006; Komvopoulos
and Gong, 2007; Komvopoulos, 2008). Because fractal geometry is characterized by the properties of continuity, non-differentiability, scale-invariance, and self-affinity (Mandelbrot, 1983), it has been used in various fields of science and engineering to describe disordered phenomena, including changes in surface topography due to wear and fracture processes. For example, Zhou et al. (1993) used a fractal contact model to examine the dependence of the adhesive wear rate on fractal parameters and material properties, Shirong and Gouan (1999) developed a fractal model of adhesive wear for the running-in stage of sliding, and Sahoo and Roy Chowdhury (2002) studied the effect of adhesion between contacting asperities on the adhesive wear behavior of fractal surfaces subjected to light loads. Although the previous studies have provided insight into the effects of fractal dimension, material properties, and surface adhesion on the loss of material by adhesive wear, the developed wear models are extensions of Archard’s model and, therefore, can only be applied to sliding surfaces. However, experimental evidence (Martin et al., 2002) and molecular dynamics simulations (Bhushan et al., 1995) have shown that adhesive wear can occur even in the absence of relative slip between the contacting surfaces. Hence, a comprehensive adhesive wear theory of rough surfaces in normal contact is necessary to bridge this gap of knowledge.

The main objective of the present analysis is twofold. First, instead of an empirical approach based on experimental results and observed trends, an adhesive wear model of rough surfaces in normal contact is derived based on plasticity-induced wear behavior that accounts for adhesion between interacting asperities. Second, the adhesive wear rate and wear coefficient are obtained in terms of the total normal load (global interference), surface topography (fractal) parameters, elastic-plastic material properties, and interfacial adhesion characteristics that depend on the material compatibility and contact environment. Results for representative contact systems with fractal surface topographies reveal the effects of roughness, surface material properties, and interfacial adhesion on adhesive wear.

6.2. Surface description

Normal contact of two rough surfaces can be analyzed by an equivalent contact model consisting of a deformable rough surface with effective material properties and equivalent roughness in contact with a rigid plane (Greenwood and Williamson, 1966). The effective elastic modulus $E^*$ and hardness $H^*$ of the equivalent surface are given by

$$\frac{1}{E^*} = \frac{1 - v_A^2}{E_A} + \frac{1 - v_B^2}{E_B}$$  \hspace{1cm} (6.1)$$

$$H^* = \min[H_A, H_B]$$  \hspace{1cm} (6.2)
where subscripts A and B refer to the two surfaces in normal contact, and $E$, $\nu$, and $H$ denote elastic modulus, Poisson’s ratio, and hardness, respectively.

In the present analysis, the equivalent rough (fractal) surface is assumed to be isotropic and self-affine, and its three-dimensional (3D) surface profile $z(x,y)$ is given by (Yan and Komvopoulos, 1998)

$$z(x, y) = L \left( \frac{G}{L} \right)^{(D-2)} \left( \frac{\ln \gamma}{M} \right)^{1/2} \sum_{m=1}^{M} \sum_{q=0}^{q_{\text{max}}} \gamma^{(D-3)q} \times$$

$$\left\{ \cos \phi_{m,q} - \cos \left[ \frac{2\pi \gamma^q (x^2 + y^2)^{1/2}}{L} \times \cos \left( \tan^{-1} \left( \frac{y}{x} \right) - \frac{\pi m}{M} \right) + \phi_{m,q} \right] \right\}$$

where $D$ \( (2 \leq D < 3) \) and $G$ are the fractal dimension and fractal roughness, respectively, both independent of wavelength in the range where the surface exhibits fractal behavior, $L$ is the sample length, $M$ is the number of superimposed ridges, $\phi_{m,q}$ is a random phase uniformly distributed in the range $[0, 2\pi]$ by a random number generator to prevent the coincidence of different frequencies at any point of the surface profile, $q$ is a spatial frequency index, and $\gamma$ ($\gamma > 1$) controls the density of frequencies in the surface profile. Surface flatness and frequency distribution density considerations suggest that $\gamma = 1.5$ is typical for most surfaces (Komvopoulos and Yan, 1997).

The fractal parameter $D$ determines the relative contributions of high- and low-frequency components in the surface profile. The amplitudes of high-frequency components become comparable to those of low-frequency components with increasing $D$, whereas the amplitudes of wavelengths comprising the surface profile intensify with increasing $G$. Because of the scale invariance of $D$ and $G$, fractal surface description at different length scales is independent of measurement scale. The shortest wavelength corresponds to the instrument resolution limit and the longest wavelength is equal to the sample length. For a continuum description to hold, the shortest wavelength is set above a cut-off length $L_s$, typically on the order of the material’s interatomic distance. Thus, the highest and lowest frequencies in the surface profile are $\omega_s = 1/L_s$ and $\omega_l = 1/L$, respectively. The lower limit of $q$ is zero, while its upper limit is given by $q_{\text{max}} = \text{int}[\log(L/L_s)/\log \gamma]$.

A surface profile generated from Eq. (6.3) exhibits scale invariance in a finite range of length scale, outside of which the surface topography can be characterized by
a deterministic function (Wang and Komvopoulos, 1994b). A typical 3D fractal surface constructed from Eq. (6.3) is shown in Fig. 6.1. Equation (6.3) was used to generate the 3D rough (fractal) surfaces used in the elastic-plastic contact analysis presented in the following section.

6.3. Contact mechanics analysis

Because the equivalent rough surface is assumed to be isotropic, any two-dimensional (2D) surface profile is a statistical representation of the 3D surface topography. Figure 6.2 shows a 2D profile of a rough surface truncated by a rigid plane. The advancement of the rigid plane toward the rough surface by interference increments \( dh \) leads to the formation of truncated segments. Each truncated segment is approximated by an asperity with a spherical cap shape having a base radius \( r' \) equal to one-fourth of the asperity’s base wavelength and height equal to the local interference \( \delta \) given by (Yan and Komvopoulos, 1998)

\[
\delta = 2G^{(D-2)}(\ln \gamma)^{1/2} (2r')^{(3-D)}
\]  

(6.4)

Fig. 6.1 Three-dimensional fractal surface generated from Eq. (6.3) for \( D=2.24, G = 2.39\times10^{-4} \text{ nm}, M=10, \gamma=1.5, L = 7.04\times10^3 \text{ nm}, \) and \( L_S = 5 \text{ nm} (\sigma = 290 \text{ nm}) \).
The size distribution function of asperity contacts is a fundamental concept in contact mechanics. Asperity contacts resulting from the truncation of a rough surface by a rigid plane follow an island distribution similar to that observed in geophysics (Mandelbrot, 1975, 1983). This island-like distribution was used in earlier contact mechanics studies of fractal surfaces (Majumdar and Bhushan, 1991; Wang and Komvopoulos, 1994a, 1994b, 1995; Komvopoulos and Yan, 1997; Yan and Komvopoulos, 1998; Komvopoulos and Ye, 2001) and also the present analysis. This island distribution obeys the following power-law relationship,

\[ N(a') = \left( \frac{a'_L}{a'} \right)^{(D-1)/2} \quad (6.5) \]

where \( N(a') \) is the number of asperities with truncated areas larger than \( a' \), and \( a'_L \) is the largest truncated contact area at a given global interference \( h \). The contact size distribution of the truncated asperities \( n(a') \) of a 3D surface profile is given by (Yan and Komvopoulos, 1998)

\[ n(a') = -\frac{dN(a')}{da'} = \frac{(D - 1)}{2a'_L} \left( \frac{a'_L}{a'} \right)^{(D+1)/2} \quad (6.6) \]

Thus the number of truncated asperities with areas between \( a' \) and \( a' + da' \) is equal to \( n(a')da' \).

At a given global interference, \( a'_L \) can be determined from the total truncated area \( S' \) of the equivalent rough surface using the relationship (Komvopoulos and Ye, 2001):

\[ S' = \int_{a'_S}^{a'_L} a'n(a')da' \quad (6.7) \]

where \( a'_S \) is the smallest truncated contact area, which for a continuum description to hold must be greater than the atomic dimensions. For example, the diameter of the smallest truncated contact asperity can be set equal to ~5–6 times the lattice parameter of the softer material.

After substituting Eq. (6.6) into Eq. (6.7) and integrating, the total truncated area can be expressed as

\[ S' = \left( \frac{D - 1}{3 - D} \right) a'_L \left[ 1 - \left( \frac{a'_S}{a'_L} \right)^{(3-D)/2} \right] \quad (6.8) \]
The total truncated area at a given global interference was obtained numerically by summing up the contact areas of all truncated asperities on the equivalent rough surface (Komvopoulos and Ye, 2001). Then, \( a'_{L} \) was determined from Eq. (6.8) in terms of \( a'_{S} \), \( D \), and \( S' \). Finally, Eqs. (6.6) and (6.8) were used to obtain the area range \([a'_{S}, a'_{L}]\) and, in turn, the spatial distribution of the truncated asperity contacts at a given global interference.

![Equivalent fractal surface](image)

**Fig. 6.2** Two-dimensional profile of a rough (fractal) surface truncated by a rigid plane.

Before proceeding with the adhesive wear analysis of rough surfaces in normal contact it is necessary to consider surface deformation at the asperity level. Yan and Komvopoulos (1998) derived a relationship of the critical truncated contact area at the transition from elastic to fully-plastic asperity deformation, whereas Wang and Komvopoulos (1994b) included in their analysis the intermediate regime of elastic-plastic deformation and derived relationships for the critical truncated contact area for elastic-plastic and fully-plastic asperity deformation. In the present study, asperity deformation is assumed to follow an elastic-perfectly plastic material behavior, implying either elastic or fully-plastic deformation at asperity contacts, depending on the local surface interference and truncated asperity contact area.
Because wear implies the removal of material as a consequence of irreversible deformation, it is assumed that only fully-plastic asperity contacts contribute to material removal. The effect of neglecting the intermediate regime of elastic-plastic deformation included in previous contact mechanics studies (Komvopoulos and Ye, 2001; Kogut and Etsion, 2002; Kogut and Komvopoulos, 2004; Mukherjee et al., 2004; Komvopoulos, 2008) in the present analysis of adhesive wear can be interpreted in the context of Fig. 6.3. Since wear particles are not likely to form from asperities in the lower range of elastic-plastic deformation, the elastic-fully plastic behavior is a reasonable approximation that also results in significant enhancement of the computational efficiency.

Fully-plastic deformation of an asperity contact occurs when \( a' \leq a'_c \), whereas elastic deformation arises when \( a' > a'_c \), where \( a'_c \) is the critical truncated contact area for fully-plastic deformation, obtained from the classical definition of material hardness.
where $\Delta F_\epsilon$ is the normal load at an elastic asperity contact calculated for $\alpha' = a'_c$.

The equation of the normal load derived from Hertz theory, modified to include the effect of interfacial adhesion (Johnson et al., 1971), is given by

$$
\Delta F_\epsilon = \frac{4E^*r^3}{3R} - \left[8\pi W_{AB}E^*r^3\right]^{1/2}
$$

where $r$ is the contact radius of an elastically deformed spherical asperity, related to the truncated contact area by

$$
r = \left[\frac{a'}{2\pi}\right]^{1/2}
$$

$R$ is the asperity radius of curvature given by (Yan and Komvopoulos, 1998)

$$
R = \frac{(a')^{(D-1)/2}}{2^{(5-D)}\pi^{(D-1)/2}G^{(D-2)}(\ln \gamma)^{1/2}}
$$

and $W_{AB}$ is the work of adhesion of contacting surfaces A and B, expressed as (Rabinowicz, 1977)

$$
W_{AB} = c_m c_l (\Gamma_A + \Gamma_B)
$$

where $c_m$ is the material compatibility index (controlled by the solid solubility limit of materials A and B), $c_l$ is the lubrication compatibility index, which depends on the environment of the interacting surfaces, and $\Gamma_A$ and $\Gamma_B$ are the surface energies of materials A and B, respectively.

Figure 6.4 shows schematically the truncation and deformation of an asperity by a rigid plane. Equations (6.9)–(6.13) indicate that $a'_c$ is a function of effective material properties $E^*$ and $H^*$, fractal parameters of the equivalent rough surface $D$ and $G$, surface energies $\Gamma_A$ and $\Gamma_B$, and interfacial adhesion controlled by the magnitude of the composite compatibility index $c_m c_l$.

When $a'_e > a'_c$, elastic and fully-plastic asperity contacts co-exist at the interface. The total truncated contact area of elastic and fully-plastic asperity contacts, $S'_e$ and $S'_p$, respectively, can be written as
\[ S' = \int_{a'L}^{a'C} a' n(a') da' \] (6.14)

\[ S'_p = \int_{a'S}^{a'C} a' n(a') da' \] (6.15)

After substituting Eq. (6.6) into Eqs. (6.14) and (6.15) and integrating,

\[ S'_e = \left( \frac{D - 1}{3 - D} \right) a'_L \left[ 1 - \left( \frac{a'_C}{a'_L} \right)^{(3-D)/2} \right] \] (6.16)

\[ S'_p = \left( \frac{D - 1}{3 - D} \right) (D-1)/(2) a'_L (3-D)/2 - a'_S (3-D)/2 \] (6.17)

Hence, the total truncated contact area \( S' \) can be expressed as

\[ S' = S'_e + S'_p = \left( \frac{D - 1}{3 - D} \right) a'_L \left[ 1 - \left( \frac{a'_C}{a'_L} \right)^{(3-D)/2} \right] \] (6.18)

For relatively small global interference (light load), the average spacing of asperity contacts is much larger than the asperity contact sizes. Also, only the largest wavelength in the waveform of a truncated asperity is considered in the analysis, i.e., all the smaller asperities residing on the largest asperity are neglected. Based on these two facts, the effect of asperity interaction (Sahoo and Banerjee, 2005; Sahoo, 2006) was ignored in the present analysis as secondary.

Fig. 6.4 Schematics of truncation and deformation of an asperity by a rigid plane.

The total normal load \( F \) can be obtained as the sum of the total forces of elastic and fully-plastic asperity contacts, \( F_e \) and \( F_p \), respectively, i.e.,
\[ F = F_e + F_p \]  

(6.19)

\[ F_e = \int_{a_s}^{a_i} \Delta F_e(a')n(a')da' \]  

(6.20)

\[ F_p = \int_{a_s}^{a_i} \Delta F_p(a')n(a')da' \]  

(6.21)

where \( \Delta F_e \) is given by Eq. (6.10) and \( \Delta F_p \) is the normal load at a fully-plastic asperity contact given by (Roy Chowdhury and Pollock, 1981)

\[ \Delta F_p = a'H'^* - 2\pi W_{AB}R \]  

(6.22)

Equations (6.6), (6.9)–(6.13), and (6.19)–(6.22) indicate that the total normal load is a function of the elastic-plastic material properties, the topography of the equivalent rough (fractal) surface, the interfacial adhesion (controlled by the material compatibility and contact environment), and the largest and smallest contact areas of truncated asperities.

The wear volume at a given global interference \( h \) was calculated by the following numerical procedure. At each increment of global interference \( dh \), the volumes of spherical segments used to approximate fully-plastic asperity contacts were subtracted from the surface profile, as illustrated schematically in Fig. 6.2. The incremental truncation of the surface profile by the rigid plane resulted in the cumulative removal of material from fully-plastic asperity contacts. Wear particles of uniform shape (i.e., spherical segments) were produced from all fully-plastic asperity contacts by setting the interference increment \( dh \) equal to the minimum local interference \( \delta_{min} \) (Yan and Komvopoulos, 1998), i.e.,

\[ dh = \delta_{min} = 2G^{(D-2)}(\ln \gamma)^{1/2} \left(2r_s'\right)^{(3-D)} \]  

(6.23)

where \( r_s' \) is the truncated radius of the smallest asperity contact. The removed spherical segment has a sphere radius \( R \), large-base area \( a' \), large-to-small base distance \( dh \), and distance between the large base and the apex of the spherical cap at the bottom of the spherical segment equal to \( \delta \) (Fig. 6.2). Thus, the volume of a wear particle approximated by a spherical segment \( dV_{p,1} \) can be written as

\[ dV_{p,1} = \left( \frac{a' - \pi \delta^2}{2} + \pi R \delta \right) dh - \pi (R - \delta)dh^2 - \frac{\pi}{3} dh^3 \]  

(6.24)

Surface contact at a given global interference and corresponding adhesive wear were assumed to be time-independent processes, i.e., the present analysis is
applicable to materials exhibiting time-independent deformation and static loads. The incremental interference scheme described previously was only used to enhance the calculation accuracy of the total wear volume at a given global interference (by capturing the details of the surface profile at small interference increments) and does not imply wear accumulation due to repeated contact under different loads.

The wear rate at a given global interference is defined as the ratio of the total wear volume $dV_p^k$ calculated in the last ($k^\text{th}$) interference increment $dh (= h/k)$ to the global interference increment $dh$, where $dV_p^i$ is the sum of the volumes of all wear particles removed from the rough surface at the $k^\text{th}$ interference increment. Hence,

$$dV_p^k = \int_{a_i}^{a_{i+1}} dV_{p,i}(a') n(a',a_{Lk}) da'$$

where $a_{Lk}$ is the largest truncated contact area at the $k^\text{th}$ interference increment.

The wear coefficient $K$ is defined as the total volume of wear particles $V$ divided by the total volume of all contacting asperities $V_t$, i.e.,

$$K = \frac{V}{V_t} = \frac{V_p}{V_p + V_e}$$

where $V_e$ and $V_p$ are the total volumes of elastic and fully-plastic asperity contacts, respectively. The cumulated total wear volume $V_p$ is given by

$$V_p = \sum_{i=1}^{k} dV_p^i$$

where $dV_p^i$ is the wear volume at the $i^\text{th}$ increment, obtained as the sum of the volumes of all wear particles generated at this increment, i.e.,

$$dV_p^i = \int_{a_{C}}^{a_{C,i}} dV_{p,i}(a') n(a',a_{L,i}) da'$$

where $a_{L,i}$ is the largest truncated contact area at the $i^\text{th}$ increment. The total volume of elastically deformed asperity contacts at a given global interference (obtained in one numerical step) is

$$V_e = \int_{a_{C}}^{a_{C,i}} dV_{e,i}(a') n(a',a_{L,i}) da'$$

where $dV_{e,i}$ is the volume of an elastic asperity contact approximated by a spherical cap of base area $a'$ and height $\delta$ given by
Equations (6.4), (6.6), (6.8)–(6.13), and (6.23)–(6.30) indicate that both the wear rate \( \frac{dV_p}{dh} \) and the wear coefficient \( K \) depend on the elastic-plastic material properties, topography (fractal) parameters, surface energies, material compatibility, interfacial adhesion, and (indirectly) total normal load through the total truncated contact area \( S' \) that controls the magnitude of \( a'_L \) (Eq. (6.8)).

### 6.4. Results and discussion

A contact mechanics analysis of adhesive wear of rough surfaces exhibiting fractal behavior was presented in the previous sections. Because it is impossible to obtain a closed-form solution for the critical truncated contact area \( a'_c \) to determine the deformation mode at each asperity contact and then calculate the wear rate and wear coefficient, numerical results are presented in this section for typical material properties and different interfacial adhesion conditions and surface roughness. The contact systems selected for numerical analysis, associated material properties, and corresponding composite compatibility index are given in Tables 6.1 and 6.2, respectively. These contact systems are representative of ceramic-ceramic (Al\(_2\)O\(_3\)/TiC), ceramic-metallic (Al\(_2\)O\(_3\)/CrN), and metal-metal (AISI 1095/AISI 1020) contact interfaces. The numerical results presented in this section are for a 3D equivalent surface topography generated from Eq. (6.3) for \( \gamma = 1.5, M = 10, L_s = 5 \text{ nm}, \)

<table>
<thead>
<tr>
<th>Contact system</th>
<th>Material</th>
<th>( \Gamma ) (mN/m)</th>
<th>( E ) (GPa)</th>
<th>( \nu )</th>
<th>( H ) (GPa)</th>
<th>( E^* ) (GPa)</th>
<th>( H^* ) (GPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al(_2)O(_3)/TiC</td>
<td>Al(_2)O(_3)</td>
<td>740</td>
<td>307</td>
<td>0.25</td>
<td>27.6</td>
<td>192</td>
<td>23.5</td>
</tr>
<tr>
<td></td>
<td>TiC</td>
<td>900</td>
<td>450</td>
<td>0.18</td>
<td>23.5</td>
<td>84.0</td>
<td>14.8</td>
</tr>
<tr>
<td>Al(_2)O(_3)/CrN</td>
<td>Al(_2)O(_3)</td>
<td>740</td>
<td>307</td>
<td>0.25</td>
<td>27.6</td>
<td>109</td>
<td>1.71</td>
</tr>
<tr>
<td></td>
<td>CrN</td>
<td>44.3</td>
<td>103</td>
<td>0.30</td>
<td>14.8</td>
<td>109</td>
<td>1.71</td>
</tr>
<tr>
<td>AISI 1095/</td>
<td>AISI 1095</td>
<td>1700</td>
<td>200</td>
<td>0.29</td>
<td>6.08</td>
<td>109</td>
<td>1.71</td>
</tr>
<tr>
<td>AISI 1020/</td>
<td>AISI 1020</td>
<td>1700</td>
<td>200</td>
<td>0.29</td>
<td>1.71</td>
<td>109</td>
<td>1.71</td>
</tr>
</tbody>
</table>

(a) Sources: Suh (1986); Komvopoulos and Zhang (2008)
Table 6.2 Composite compatibility index of contact surfaces at different environments

<table>
<thead>
<tr>
<th>Contact system</th>
<th>Vacuum</th>
<th>Clean air</th>
<th>Poor lubricant</th>
<th>Fair lubricant</th>
<th>Good lubricant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al₂O₃/TiC</td>
<td>0.40</td>
<td>0.36</td>
<td>0.27</td>
<td>0.20</td>
<td>0.15</td>
</tr>
<tr>
<td>Al₂O₃/CrN</td>
<td>0.40</td>
<td>0.36</td>
<td>0.27</td>
<td>0.20</td>
<td>0.15</td>
</tr>
<tr>
<td>AISI 1095/</td>
<td>1.50</td>
<td>1.00</td>
<td>0.37</td>
<td>0.14</td>
<td>0.05</td>
</tr>
<tr>
<td>AISI 1020</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) Source: Rabinowicz (1977)

Fig. 6.5 Critical truncated contact area $a'_C$ versus environment (interfacial adhesion) for surface properties given in Tables 6.1 and 6.2 ($\sigma = 290$ nm).

$L = 7.04 \times 10^3$ nm, and $\phi_{m,q} = \pi / 2$. In all cases, $a'_S$ was scaled to the smallest wavelength existing in the surface profile, i.e., $a'_S = \pi (r'_S)^2 = \pi (L_S / 2)^2 = 19.6$ nm$^2$.

Unless otherwise stated, the fractal parameters of the 3D equivalent surface topography are $D = 2.24$ and $G = 2.39 \times 10^{-4}$ nm ($\sigma = 290$ nm). The effects of material properties, interfacial adhesion (i.e., effects of material compatibility and
environment), and surface topography (i.e., roughness effect) on the deformation mode of asperity contacts are discussed first, followed by the presentation of wear rate and wear coefficient results and a discussion of the applicability of the present analysis.

6.4.1. Deformation mode of asperity contacts

The effect of the work of adhesion, controlled by the surface energies of the interacting surfaces (Table 6.1) and the combined compatibility index \( c_{cmcl} \) (Table 6.2), on the deformation mode of the asperity contacts can be interpreted in terms of the dependence of \( \alpha'_c \) (Eq. (6.9)) on contact environment shown in Fig. 6.5. The trend is for \( \alpha'_c \) to decrease with the work of adhesion (or interfacial adhesion) resulting from the enhancement of the lubrication efficacy. The highest \( \alpha'_c \) corresponds to the metal-metal contact system in vacuum, and the lowest \( \alpha'_c \) to the ceramic-metallic system at any environment. The results shown in Fig. 6.5 reveal a strong effect of interfacial adhesion (effects of material compatibility and contact environment) and elastic-plastic material properties on the dominance of plasticity at the contact interface.

Figure 6.6 reveals the effect of surface roughness \( \sigma \) on the deformation mode of asperity contacts under vacuum conditions, i.e., highest \( c_{cmcl} \) value for each contact system (Table 6.2). The equivalent surface roughness \( \sigma \) was varied in the range of \( 1-10^4 \) nm by varying the fractal roughness \( G \) between \( 1.94 \times 10^{-14} \) and \( 4.88 \times 10^2 \) nm and fixing the fractal dimension \( (D = 2.24) \). Surface topographies with \( \sigma = 1 \) and \( 10^4 \) nm are typical of ultra-smooth surfaces (e.g., wafers) and well-polished bearing surfaces, respectively. In all cases, \( \alpha'_c \) tends to increase monotonically with surface roughness. The results show a strong dependence of \( \alpha'_c \) on both material properties and surface roughness. For example, \( \alpha'_c \) of the metal-metal contact system is less than that of the ceramic-ceramic contact system in the case of ultra-smooth surfaces \( (\sigma = 1 \text{ nm}) \), whereas the opposite is observed for relatively rough surfaces \( (\sigma > 10 \text{ nm}) \).

To examine the evolution of plasticity at the contact interface in terms of the global interference (normal load), the total truncated contact area of fully-plastic
Fig. 6.6 Critical truncated contact area $a'_C$ versus rms roughness of equivalent surface $\sigma$ ($D = 2.24$) for surface properties given in Tables 6.1 and 6.2.

asperity contacts $S'_p$ normalized by the apparent sample area $S_a$ ($S_a = L^2$) is shown in Fig. 6.7 as a function of global interference $h$ for fixed surface topography and environment (vacuum). In all cases, $S'_p / S_a << 1$ (despite the wide variation of the global interference) and $S'_p$ increases linearly with global interference. This trend can be explained by considering that both $a'_L$ and $S'_p$ increase with $S'$ (Eqs. (6.8) and (6.17), respectively) and that $S'$ increases with $h$. For a given global interference, Fig. 6.7 shows that $S'_p / S_a$ is a strong function of material properties, in accord with Eq. (6.17) and the results shown in Fig. 6.5 for vacuum.

The ratio of the total truncated contact area of fully-plastic asperities to the total truncated contact area $S'_p / S'$, shown in Fig. 6.8 as a function of global interference $h$ for fixed topography ($\sigma = 290$ nm) and environment (vacuum), can be
used to further examine the dominant deformation mode at the asperity level. The low-order magnitude of $S'_{p}/S'$ indicates that fully-plastic asperities represent a small percentage of all the asperity contacts comprising the total truncated contact area, implying the dominance of elastic deformation at the asperity level. Equations (6.17) and (6.18) and Fig. 6.5 indicate that, for fixed global interference, topography parameters (roughness), and contact environment, $S'_{p}/S'$ depends on $a'_C$. The rapid decrease in $S'_{p}/S'$ with increasing $h$ in the low-interference range ($h < \sim 70$ nm) reveals an enhancement of the dominance of elastic deformation at the asperity level with increasing global interference. This trend can be attributed to the decreasing trend of $S'_{p}/S'$ with increasing $a'_L$ (or global interference), deduced from Eqs. (6.17) and (6.18). The small negative slope of $S'_{p}/S'$ in the upper-interference range ($h > \sim 70$ nm) implies that the variation of the global interference within this range does not affect significantly the relative contribution of plastic deformation to the overall
Fig. 6.8 Ratio of total truncated contact area of fully-plastic asperity contacts to total truncated contact area $S'_{p} / S'$ versus global interference $h$ for surface properties given in Tables 6.1 and 6.2 ($\sigma = 290$ nm).

Fig. 6.9 Wear rate $dV^k_{p}/dh$ versus global interference $h$ for surface properties given in Tables 6.1 and 6.2 ($\sigma = 290$ nm).
deformation at the contact interface. The critical global interference (~70 nm) demarcating the low-interference range from the upper-interference range is equal to ~0.25 $\sigma$.

6.4.2. Wear rate

Figure 6.9 shows the wear rate $dV_p^k/dh$ as a function of global interference $h$ of the three contact systems in vacuum. The linear increase in wear rate is a consequence of the increase in $a'_L$ with global interference (Eq. (6.8)), which, for given $a'_s$ and $a'_C$, yields a denser asperity distribution $n(a')$ (Eq. (6.6)), implying a higher wear volume $dV_p^k$ (Eq. (6.25)). Figure 6.9 shows that, for fixed topography parameters, global interference (i.e., fixed $a'_L$), and work of adhesion, $dV_p^k/dh$ depends on the material properties, which control the magnitude of $a'_C$ (Eq. (6.25)).

The effect of the work of adhesion (i.e., material compatibility and/or contact environment) on $dV_p^k/dh$ for a given surface topography and fixed global interference can be interpreted in the context of the results shown in Fig. 6.10. The trend for $dV_p^k/dh$ to decrease with the work of adhesion (or interfacial adhesion) due to the improvement of lubrication is similar to that shown in Fig. 6.5, and can be attributed to the dependence of $dV_p^k/dh$ only on $a'_C$ (Eq. (6.25)) for fixed topography and global interference. The lowest wear rate corresponds to the ceramic-metallic system ($\text{Al}_2\text{O}_3$/CrN), i.e., the contact system with the lowest work of adhesion at any environment.

The dependence of the wear rate $dV_p^k/dh$ on the rms roughness of the equivalent surface $\sigma$ (or fractal roughness $G$ since $D = 2.24$) is shown in Fig. 6.11 for fixed global interference and environment (vacuum). The well-known trend of the wear rate to increase with surface roughness shown by all contact systems suggests that higher $G$ values promote plastic deformation at the asperity level, thus enhancing the wear rate. This finding can be explained by considering the dependence of $a'_C$ and $a'_L$ on the equivalent surface roughness $\sigma$. The trend in Fig. 6.11 is similar to that shown in Fig. 6.6 ($a'_C$ increases with $\sigma$) because, for fixed global interference and environment, $dV_p^k/dh$ depends mainly on $a'_C$ (Eq. (6.25)). Although the increase in $G$ (or $\sigma$) decreases $S'$, which, in turn, results in the decrease of both $a'_L$ and
Fig. 6.10 Wear rate $dV^k_p/dh$ versus environment (interfacial adhesion) for surface properties given in Tables 6.1 and 6.2 ($\sigma = 290$ nm).

$n(a')$ for fixed $D$ and $h$ (Eqs. (6.8) and (6.6), respectively), this effect is secondary (Eq. (6.25)) compared to that of $a'_C$. Thus, $dV^k_p$ is mostly affected by the magnitude of $a'_C$.

6.4.3. Wear coefficient

Figure 6.12 shows the wear coefficient $K$ as a function of global interference $h$ for fixed topography and environment (vacuum). In all cases, the wear coefficient is on the order of $10^{-4}$, which is typical of adhesive wear of clean surfaces (Rabinowicz, 1965, 1977), and decreases rapidly with the increase of the global interference (normal load), reaching a steady state at a critical global interference that depends on the elastic-plastic material properties. The relatively high order of magnitude of $K$ is typical of contacting surfaces in vacuum. As explained previously, for fixed global interference, topography (fractal parameters/roughness), and environment, $S'_p/S'$ is controlled by $a'_C$, implying a dependence of plastic deformation and, in turn, wear
Fig. 6.11 Wear rate $dV_p^k/dh$ versus rms roughness of equivalent surface $\sigma$ ($D = 2.24$) for surface properties given in Tables 6.1 and 6.2.

The effect of the work of adhesion (contact environment) on the wear coefficient can be interpreted in light of Fig. 6.13. All contact systems show an increase in wear coefficient with increasing work of adhesion (i.e., decreasing lubrication efficacy). The ceramic-ceramic and metal-metal contact systems exhibit similar wear coefficients; however, they are much higher than those of the ceramic-metallic contact system at any environment. This difference may be attributed to the much lower work of adhesion of the ceramic-metallic contact system, mainly attributed to the very low surface energy of the CrN surface (Table 6.1).

Figure 6.14 shows the variation of the wear coefficient $K$ with the rms roughness of the equivalent surface $\sigma$ for fixed global interference ($h = 150$ nm) and environment (vacuum). The results reveal the existence of two wear regimes. In the low-roughness regime ($\sigma < 100$ nm), the wear coefficient increases gradually with the surface roughness, assuming values on the order of $10^{-4}$, which are in the middle range of adhesive wear coefficient (Rabinowicz, 1965, 1977). However, in the high-roughness regime ($\sigma > 100$ nm), the wear coefficient increases more rapidly with the surface roughness, and for metal-metal and ceramic-ceramic interfaces with
Fig. 6.12 Wear coefficient $K$ versus global interference $h$ for surface properties given in Tables 6.1 and 6.2 ($\sigma = 290$ nm).

Fig. 6.13 Wear coefficient $K$ versus environment (interfacial adhesion) for surface properties given in Tables 6.1 and 6.2 ($\sigma = 290$ nm).
\( \sigma > 1000 \) nm it reaches values on the order of \( 10^{-3} \) that are usually associated with severe adhesive wear of clean surfaces (Rabinowicz, 1965, 1977). The lowest wear coefficient of the ceramic-metallic contact system throughout the examined roughness range is indicative of its superior antiwear characteristics. The trend for the wear coefficient to increase with surface roughness follows the trend shown in Fig. 6.6. Despite the decrease of \( a'_{lk} \) and \( a'_{li} \) with the roughness increase (or increase in \( G \) since \( D = 2.24 \)), the wear coefficient is mostly affected by the magnitude of \( a'_C \) (Eqs. (6.26)–(6.29)).

![Figure 6.14](image)

Fig. 6.14 Wear coefficient \( K \) versus rms roughness of equivalent surface \( \sigma \) (\( D = 2.24 \)) for surface properties given in Tables 6.1 and 6.2.

6.4.4. Applicability and extension of the present analysis

In the present analysis, the Hertzian equation of the normal load at a single elastic asperity was modified to include the effect of adhesion (Eq. (6.10)) following the JKR adhesion model (Johnson et al., 1971). However, the analysis is not restricted to a particular adhesion model. For example, the effect of adhesion on the normal load at a single asperity can also be represented by the DMT adhesion model (Derjaguin et al., 1975) without additional modification of the analysis. The applicability of the JKR or DMT adhesion models can be determined by a dimensionless parameter, referred to as the Tabor parameter (Tabor, 1977), given by
\[ \mu = \left( \frac{RW_{AB}^2}{E^{*2}z_0^3} \right)^{1/3} \]  

(6.31)

where \( z_0 \) is the equilibrium separation distance of the two surfaces. Johnson and Greenwood (1997) have argued that the DMT and JKR models are applicable when \( \mu < 0.1 \) and \( \mu > 5 \), respectively.

For the contact systems examined in this study, \( z_0 \) was determined based on the method proposed by Yu and Polycarpou (2004) and \( c_{nm}c_l \) values corresponding to good lubrication conditions (Table 6.2) were used to obtain the lowest \( W_{AB} \) values for each contact system. Since \( R \propto a' \), the smallest truncated area for elastic deformation, i.e., \( a'_c \) (Figs. 6.5 and 6.6), was used to calculate the asperity radius of curvature in each case. Substituting the values of \( z_0 \), \( W_{AB} \), \( R \), and \( E^{*} \) obtained for each contact system into Eq. (6.31), the Tabor parameter of the ceramic-ceramic, ceramic-metallic, and metal-metal contact systems was found equal to 6.77, 6.21, and 11.2, respectively. These values represent lower bounds of the Tabor parameter because the lowest values of \( W_{AB} \) and \( R \) of each contact system were used to calculate \( \mu \). Because \( \mu > 5 \) for all interfacial adhesion (lubrication) conditions and entire range of elastic asperity radius considered in the analysis, it is concluded that the JKR model is suitable for analyzing these contact systems. The present analysis can also be applied to surfaces exhibiting \( \mu < 0.1 \), provided the contact load equation (Eq. (6.10)) is replaced by that of the DMT adhesion model.

6.5. Conclusions

A contact mechanics analysis of adhesive wear of rough (fractal) surfaces was performed to elucidate the dependence of plastic deformation at asperity contacts and wear rate (coefficient) on global interference (normal load), elastic-plastic material properties, topography (roughness), and work of adhesion of the contacting surfaces. Material loss (wear) was presumed to originate from fully-plastic asperity contacts, accounting for the contribution of interfacial adhesion to the normal load at each asperity contact. Numerical results for representative contact systems revealed the effects of material properties, roughness, surface compatibility, and environmental conditions on the adhesive wear rate and wear coefficient. Based on the presented results and discussion, the following main conclusions can be drawn from this study.

1) Plastic deformation at asperity contacts is controlled by the critical truncated contact area, which depends on the elastic-plastic material properties, topography (roughness), and work of adhesion (affected by the material compatibility and contact environment or lubrication condition) of the contacting surfaces.

2) The number of plastically deformed asperity contacts increases rapidly with the
global interference. However, the plastic truncated contact area is much less (<1–2%) than the total truncated contact area, revealing the dominance of elastic deformation at the asperity level over a wide range of global interference.

3) The wear rate increases monotonically with the global interference, whereas the wear coefficient decreases rapidly to a steady state, showing a weak dependence on normal load.

4) Both the wear rate and the wear coefficient decrease with the work of adhesion (interfacial adhesion) and increase with the roughness of the contacting surfaces.

5) The adhesive wear coefficient may vary significantly (in the range of ~10^{-4}–10^{-3} for the contact systems examined), depending on the material properties, surface topography (roughness), and work of adhesion, which depends on the surface energies of the contacting surfaces and interfacial adhesion controlled by the material compatibility and contact environment.
Chapter 7

A discrete dislocation plasticity analysis of a single-crystal half-space indented by a rigid cylinder

7.1. Introduction

Plastic deformation of contacting solids controls surface mechanical properties, such as hardness, friction, wear, and contact fatigue. Deformation of asperity microcontacts established between rough surfaces under load has been the main objective of many contact analyses. In early studies, asperity deformation at crystalline solid surfaces was examined within the context of macroscopic contact theories, using concepts and approaches from continuum plasticity theory and contact mechanics. For typical crystalline materials, these studies provide an appropriate description of asperity deformation for contact sizes on the order of tens of micrometers or larger. However, the topography of most engineering surfaces comprises much smaller asperities, typically ranging from a few nanometers to several micrometers. As a consequence, contact deformation is confined within a small surface volume and the microcontact size is comparable to a characteristic microstructure scale, e.g., mean size of dislocation cells in the surface layer of the contacting solids. Since material behavior at such small scales is influenced by the discrete lattice structure, the crystal cannot be treated as a homogeneous elastic-plastic medium and classical (continuum) plasticity constitutive laws are not valid (Kuhlmann-Wilsdorf, 1981; Pollock, 1992). Indeed, nano-/micro-indentation experiments have shown a strong scale effect on plasticity (Ma and Clarke, 1995; Swadener et al., 2002; Gouldstone et al., 2007) that cannot be described by scale-invariant continuum plasticity.

Among various analytical studies dealing with scale effects on microscale plasticity of crystalline solids, nonlocal phenomenological crystal plasticity (Shu and Fleck, 1999; Acharya and Bassani, 2000; Gurtin, 2002) introduces a length scale to account for the effect of geometrically necessary dislocations in the presence of plastic strain gradients, such as those in deformation due to indentation loading. However, similar to classical plasticity theory, nonlocal formulations are phenomenological; therefore, they are not based on fundamental plasticity phenomena, such as the collective evolution of discrete dislocations. Considering microstructure and crystallography at very small scales characteristic of asperity microcontacts, discrete dislocation plasticity (Van der Giessen and Needleman, 1995; Polonsky and Keer, 1996a) is more appropriate for analyzing nonlinear plasticity problems. In this method, the discrete nature of plastic flow in the deformed medium is represented by the
collective action of discrete edge dislocations moving along specific slip planes. Hence, unlike continuum plasticity formulations where deformation is averaged over dislocations, discrete dislocation plasticity treats dislocations individually as they propagate through a continuum medium. This method allows for a true mechanism-based plasticity analysis in which plastic flow is not dictated by assumed constitutive models of the phenomenological material behavior.

Micro-/nano-indentation methods yield insight into surface and subsurface deformation behavior of contacting rough surfaces because they enable the study of single-contact events at the asperity scale. The link between these single-contact encounters and real asperity microcontacts includes parameters such as asperity (indenter) shape and interaction between neighboring asperities. Previous discrete dislocation analyses of single-crystal media deformed by a rigid indenter elucidated the effect of indenter shape on localized deformation (Polonsky and Keer, 1996a,b; Kreuzer and Pippan, 2004a; Widjaja et al., 2005). Indentation of a single-crystal material can be regarded as coarse-grained plastic flow in a polycrystalline solid. Discrete dislocation plasticity has been used to analyze wedge indentation of single-crystal media (Balint et al., 2006; Widjaja et al., 2007a); however, these analyses were based on finite element modeling and, therefore, cannot be used to study deformation phenomena under different indentation loading conditions. Although the effect of the indenter radius on plane-strain indentation has been examined (Widjaja et al., 2007b), a comprehensive study of the effects of other intrinsic material parameters, such as the orientation and distance of slip planes and the presence of dislocation obstacles, has not been reported yet. While indenters of arbitrary shape have been considered (Kreuzer and Pippan, 2004a,b), only the effect of dislocation source density on dislocation structures was examined. The focus in other studies has been the significance of slip band orientation on indentation size effect (Kreuzer and Pippan, 2007) and the mechanical behavior of thin films in the presence of a limited number of dislocation sources (Kreuzer and Pippan, 2005). Size-dependent deformation has been demonstrated by simulation results of indentation studies based on discrete dislocation plasticity (Kreuzer and Pippan, 2004a,b; Balint et al., 2006; Widjaja et al., 2007a,b).

An important limitation of previous discrete dislocation studies is that damage due to dislocation activity was not examined. Although microcontact sizes comparable to a characteristic microstructure length were considered (Polonsky and Keer, 1996b), indentation deformation under the theoretical strength of the material was not analyzed. The main objective of this publication is threefold. First, a discrete dislocation plasticity analysis of plane-strain indentation of a single-crystal half-space by a rigid cylinder is developed using a set of constitutive rules of dislocation emission, glide, pinning, and annihilation. Second, the effects of contact load, indenter radius (sharpness), crystal orientation, dislocation source density, and obstacle density on the evolution of plasticity are interpreted in the context of numerical results
revealing the evolution of dislocation structures and subsurface shear stresses. Third, the initiation of yielding and deformation under the theoretical strength of the material are discussed in terms of the indenter radius, dislocation source density, and slip-plane distance.

7.2. Analytical model

Because dislocation density and interaction mechanisms intensify rapidly with increasing contact load, simulating plastic deformation by real crystal dislocations is computationally intensive. Thus, to increase the computational efficiency, the single-crystal half-space is modeled as an initially stress-free, homogeneous, isotropic elastic medium with parallel and equally spaced slip planes. In addition, individual crystal dislocations are modeled as emitted dislocations, the existence of partial dislocations is neglected, and the magnitude of the Burgers vector of all dislocations is set equal to the lattice distance of the crystalline half-space.

Figure 7.1 shows schematically the indentation of a half-space medium by a rigid cylinder of radius $R$. Surface contact is assumed to be frictionless, i.e., energy dissipation occurs mainly by subsurface plastic deformation. Friction effects are not considered in this study because the focus is on subsurface plasticity due to indentation loading, not on surface interaction effects (e.g., adhesion) on deformation behavior. The half-space contains parallel slip planes at fixed distance $d$ and orientation angle with respect to the free surface $\theta$. Frank-Read dislocation sources (Van der Giessen and Needleman, 1995) randomly distributed on slip planes are also shown in Fig. 7.1. The location and density of dislocation sources $\rho_s$ (per unit area) are fixed in each simulation. At a critical shear stress, bowing of Frank-Read segments leads to generation of new dislocation loops. When the total resolved shear stress $\tau$ due to the applied contact load $P$ exceeds the dislocation emission stress $\tau_e$ (also referred to as the source strength) at the location of a specific dislocation source, a pair of edge dislocations (dislocation dipole) with anti-parallel Burgers vectors $\pm b$ is generated on the slip plane, with the source centered between the two edge dislocations (Van der Giessen and Needleman, 1995). Both the Burgers vectors of the formed dislocations and the dislocation lines lie in the slip planes. The sign of the dipole is determined by the sign of the resolved shear stress at the source location along the slip plane. The distance between two emitted dislocations $l_e$, obtained from the balance of the shear stress exerted by the two dislocations to each other and the total resolved shear stress, is given by (Van der Giessen and Needleman, 1995)

$$l_e = \frac{E}{4\pi(1-\nu^2)} \frac{b}{\tau_e}$$

(7.1)

where $E$ and $\nu$ are the elastic modulus and Poisson’s ratio, respectively. Displacement discontinuities between dislocation dipoles contribute to macroscopic plastic strain.
The dislocation emission stress characterizes the lattice resistance to dislocation motion (apart from elastic dislocation interactions) and is usually equal to a small fraction of the shear modulus (Kovács and Zsoldos, 1973). It is also assumed that dislocation activity in the bulk is the dominant plastic flow process, i.e., nucleation of dislocation dipoles occurs only from dislocation sources in the subsurface, whereas nucleation of dislocations from surface sources is neglected, as in a previous study (Nicola et al., 2007).

Damage in crystalline solids is characterized by the development of material discontinuities (Lemaitre and Dufailly, 1987). Because the initial structure of the half-space contains dislocation sources, the level of damage can be represented by a damage parameter at initial yield $D_Y$, defined as

$$D_Y = 1 - \left( \frac{P_{SY}}{P_Y} \right)$$  \hspace{1cm} (7.2)

where $P_{SY}$ is the contact load that produces the first dislocation dipole under the given simulation parameters, and $P_Y$ is the yield load determined by the von Mises yield criterion, given by (Johnson, 1987)

$$P_Y = \frac{\pi R}{E} p_{0,Y}^{2}$$ \hspace{1cm} (7.3)

where $p_{0,Y}$ is the maximum contact pressure at the instant of the onset of yielding, given by $p_{0,Y} = 1.79Y$, where $Y$ is the yield strength of the half-space material. Eq. (7.2) indicates that $0 \leq D_Y \leq 1$.

Nucleated edge dislocations can be modeled as line singularities in a linear elastic solid, and the stress field of each dislocation outside its core can be described by linear elasticity. The total resolved shear stress $\tau$ at the location of a dislocation source and along a slip-plane is obtained as the sum of the local shear stress predicted from Hertz analysis and the shear stress at the same location due to the long-range (elastic) dislocation stress fields of all existing dislocations, except the dislocation (if any) at that location, i.e.,

$$\tau = \tau^H + \sum_i \tau_i^d$$ \hspace{1cm} (7.4)
where $\tau^H$ is the Hertzian shear stress and $\tau^d_i$ is the shear stress due to the $i^{th}$ dislocation. The Hertzian contact stresses at a point $(x_1, x_2)$ can be expressed as (Johnson, 1987)

\[ \sigma^H_{11} = -\frac{P_0}{r} \left[ m \left( \frac{1}{n^2 + m^2} + \frac{n^2 + x^2}{n^2 + m^2} \right) + 2x_2 \right] \]  
(7.5)

\[ \sigma^H_{22} = -\frac{P_0}{r} \left[ m \left( \frac{1}{n^2 + m^2} - \frac{n^2 + x^2}{n^2 + m^2} \right) \right] \]  
(7.6)

\[ \tau^H_{12} = -\frac{P_0}{r} \left[ m \frac{m^2 - x^2}{n^2 + m^2} \right] \]  
(7.7)
where $\sigma_{11}''$ and $\sigma_{22}''$ are the normal stresses in the $x_1$- and $x_2$-directions, respectively, $\tau_{12}''$ is the shear stress, $r$ is the half-contact width, $p_0$ is the maximum contact pressure, given by

$$p_0 = \frac{2P}{\pi r} \quad (7.8)$$

and $m$ and $n$ are defined as

$$m^2 = \frac{1}{2} \left[ \left(r^2 - x_1^2 + x_2^2\right)^2 + 4x_1^2x_2^2 \right]^{1/2} + r^2 - x_1^2 + x_2^2 \quad (7.9a)$$

$$n^2 = \frac{1}{2} \left[ \left(r^2 - x_1^2 + x_2^2\right)^2 + 4x_1^2x_2^2 \right]^{1/2} - \left(r^2 - x_1^2 + x_2^2\right) \quad (7.9b)$$

where $\text{sign}(m) = \text{sign}(x_2)$ and $\text{sign}(n) = \text{sign}(x_1)$.

The dislocation shear stress field in a half-space can be obtained from a previous analysis (Kreuzer and Pippan, 2004b), where the shear stress field of an edge dislocation adjacent to the free surface was found by solving a boundary value problem using the method of complex functions. The stresses at point $z(x_1, x_2)$ ($z = x_1 + ix_2$) due to a dislocation at point $z_0$ are given by (Kreuzer and Pippan, 2004b)

$$\sigma_{11}^d + \sigma_{22}^d = 2[\phi(z) + \overline{\phi}(z)] \quad (7.10)$$

$$\sigma_{22}^d - i\tau_{12}^d = \phi(z) - \phi(\overline{z}) + (z - \overline{z})\overline{\phi}'(z) \quad (7.11)$$

where $\sigma_{11}^d$ and $\sigma_{22}^d$ are the normal stresses in the $x_1$- and $x_2$-directions, respectively, $\tau_{12}^d$ is the shear stress, and $\phi(z)$ is a complex potential given by

$$\phi(z) = \frac{2A}{z - z_0} - \frac{2A}{z - \overline{z}_0} - \frac{2A(z - \overline{z})}{(z - \overline{z}_0)^2} \quad (7.12)$$

where $A$ is a material constant, defined as

$$A = \frac{Gb}{2\pi(\kappa + 1)} \quad (7.13)$$

where $G$ is the shear modulus and $\kappa = 3 - 4\nu$ for plane-strain deformation. This
analytical approach for determining the dislocation stress field is convenient and more advantageous than the method of Van der Giessen and Needleman (1995) where one component of the dislocation stress field was obtained by solving a boundary value problem with the finite element method, a computationally intensive approach since it requires a separate finite element model for each dislocation structure generated during loading.

For simplicity, nucleating dislocations are assumed to glide along a fixed set of slip planes, and dislocation climb and cross slip are neglected. Hence, the crystalline half-space examined in this study exhibits anisotropic plasticity. The dislocation glide velocity $u_g$ is given by (Van der Giessen and Needleman, 1995)

$$u_g = \frac{b\tau}{K}$$

(7.14)

where $K$ is the drag coefficient. For metals, the main source of drag is due to phonons (Van der Giessen and Needleman, 1995). Dislocation glide along a slip plane of a crystal can be inhibited by various obstacles, such as small precipitates. In the present analysis, obstacles are modeled as material points where moving dislocations are pinned down (Van der Giessen and Needleman, 1995). Usually, pinned dislocations act as precursors of dislocation pile-ups. The resolved shear stress acting on the leading dislocation increases rapidly as dislocations accumulate in a pile-up and, eventually, the leading dislocation overcomes the obstacle by a thermally activated process. Obstacle bypassing by a dislocation is simulated by releasing the pinned dislocation as soon as the total resolved shear stress exceeds the obstacle strength $\tau_0$. Similar to dislocation sources, obstacles of density $\rho_o$ (per unit area) are randomly distributed in the half-space medium and all obstacles are assumed to have the same strength.

Two opposite dislocations on the same slip plane are annihilated and, thus, removed from the simulation, when their distance is less than the critical annihilation distance $l_a$, an intrinsic material parameter. Annihilation of multiple dislocations is treated as sequences of pair annihilations. In addition, all dislocations exiting the medium through the surface, including the contact region, are also removed from the analysis as no longer contributing to the total resolved shear stress. Due to the highly nonuniform dislocation stress field and complex interactions of the dislocations with other dislocations, obstacles, and the half-space surface, the dislocated body is not in thermodynamic equilibrium at any given deformation stage of the analysis.

To include history-dependent plasticity in the analysis and to model large deformations, an incremental approach that uses a prescribed time step and corresponding incremental load was used in all simulations. At each time step, the total resolved shear stress field was obtained and the nucleation of dislocations,
relocation of existing dislocations, pinning of dislocations by obstacles, annihilation of opposite dislocations, and exit of dislocations from the half-space surface were simulated based on the dislocation constitutive rules described above. The previous procedure was repeated at each load step to determine the stress field and evolving dislocation structures. The macroscopic strain $\varepsilon$ in the half-space was determined as the sum of a representative elastic strain due to Hertzian contact deformation $\varepsilon_e$ and a representative plastic strain due to dislocation multiplication $\varepsilon_p$. The representative elastic strain can be approximated as

$$\varepsilon_e = \frac{r}{R}$$

(7.15)

The representative plastic strain $\varepsilon_p$ at each load step was obtained incrementally using a plastic strain increment $\Delta\varepsilon_p$ defined as

$$\Delta\varepsilon_p = \rho_d b \bar{s}$$

(7.16)

where $\rho_d$ is the dislocation density (per unit area) at the particular step, and $\bar{s}$ is the average displacement of all dislocations within the time step, given by

$$\bar{s} = \bar{u}_g \Delta t$$

(7.17)

where $\bar{u}_g$ is the average glide velocity of all dislocations at that step, and $\Delta t$ is the time increment. Eqs. (7.16) and (7.17) indicate that the macroscopic plastic strain is a consequence of microscale dislocation activity.

Since the present analysis is carried out in the context of small deformation gradient, the effects of geometry changes on the momentum balance and lattice rotation on dislocation glide are neglected. However, the contact area between the half-space and the rigid indenter is determined from the deformed surface profile (Fig. 7.1) using Hertz theory (Johnson, 1987).

### 7.3. Results and discussion

The numerical results presented in this section are for a half-space medium with material properties typical of those of single-crystal copper (Table 7.1). Characteristic lengths are the half-contact width $r$, magnitude of Burgers vector $b$, and slip-plane distance $d$ (Fig. 7.1). For typical machined and/or worn surfaces, $d$ can be set approximately equal to the average size of dislocation cells or sub-grains in the uppermost material layer (Rigney, 1988). Unless otherwise stated, the results presented below are for $d/b = 100$ and $l_d/b = 10$. In view of the random distributions of dislocation sources and obstacles, a statistical analysis was performed for each set of parameters. Thus, numerical results are presented in the form of data points representing average values obtained from five simulations of different source and
obstacle distributions and identical all other parameters. The effect of loading rate on the deformation behavior was found to be insignificant; therefore, all simulation results given below are for a fixed loading rate of $7.2 \times 10^{23} \text{ N/m·s}$.

<table>
<thead>
<tr>
<th>$E$ (GPa)</th>
<th>$\nu$</th>
<th>$b$ (nm)</th>
<th>$\tau_e$ (MPa)</th>
<th>$K$ (Pa·s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>110</td>
<td>0.34</td>
<td>0.25</td>
<td>78</td>
<td>$1.7 \times 10^{-5}$</td>
</tr>
</tbody>
</table>

(a) Sources: Jassby and Vreeland, 1970; Ross, 1992; Van der Giessen and Needleman, 1995

7.3.1. Nucleation of dislocations and initiation of yielding

From a dislocation plasticity perspective, the emission of the first dislocation dipole (the two-dimensional equivalent of a dislocation loop) defines the onset of yielding. Due to preexisting dislocation sources, the single-crystal half-space is damaged as opposed to the ideally homogeneous half-space that deforms only elastically. Figure 7.2 shows the effects of dimensionless slip-plane distance $d/b$, dislocation source density $\rho_s$, and dimensionless indenter radius $R/b$ on the damage parameter at the onset of yielding $D_Y$. The damage parameter due to local yielding is $D_Y \approx 0.3$, implying a significant decrease of the material strength in the presence of dislocation sources. As expected, damage increases with decreasing slip-line distance (Fig. 7.2(a)) and/or increasing dislocation source density (Fig. 7.2(b)) due to the decrease in critical load for first dipole emission and enhancement of dislocation nucleation, respectively. Moreover, $D_Y$ increases sharply with the indenter radius, reaching a steady-state value of $\approx 0.35$ for a critical indenter radius $R_{cr}$ much larger than the slip-plane distance, i.e., $R_{cr} > 500d$ (Fig. 7.2(c)).

To determine the most favorable slip-plane orientation for the onset of yielding, the dimensionless dislocation density $\rho_d/\rho_s$ was obtained as a function of slip-plane orientation angle $\theta$, for a contact load ($P = 12 \text{ N/m}$) sufficiently high to generate a dislocation dipole for the least favorable slip-plane direction, i.e., $\theta = 90^\circ$ (Fig. 7.3). The maximum density of nucleating dislocations corresponds to $\theta = 45^\circ$, which is in excellent agreement with the onset of yielding predicted by continuum mechanics yield criteria (Johnson, 1987).
Fig. 7.2 Damage parameter $D_Y$ defined at the onset of yielding (emission of first dislocation dipole) versus (a) dimensionless slip-plane distance $d/b$, (b) dislocation source density $\rho_s$, and (c) dimensionless indenter radius $R/b$. 
7.3.2 Multiplication of dislocations and evolution of dislocation structures

The increase in dislocation density during loading provides insight into the evolution of plasticity after the onset of yielding. Figure 7.4 shows the effects of
dimensionless contact load $P/P_Y$ and slip-plane orientation angle $\theta$ on dimensionless dislocation density $\rho_d/\rho_s$. It can be seen that after the initiation of yielding, the dislocation density increases monotonically with the contact load for all slip-plane orientations; however, the most favorable slip-plane direction is parallel to the half-space surface ($\theta = 0^o$).

Figure 7.5 shows the dislocation density $\rho_d$ as a function of dimensionless contact load $P/P_Y$ and dislocation source density $\rho_s$ for slip-plane orientation angle (a) $\theta = 45^o$ and (b) $\theta = 0^o$.

Figure 7.5 shows the dislocation density $\rho_d$ as a function of dimensionless contact load $P/P_Y$ and dislocation source density $\rho_s$ for slip-plane directions $\theta = 0^o$ and
45°. The dislocation density increases nonlinearly with the density of Frank-Read dislocation sources in both slip-plane directions. However, a comparison of Figs. 7.5(a) and 7.5(b) indicates that the slip-plane direction exhibits a more pronounced effect on dislocation multiplication than the dislocation source density. Moreover, for a given dislocation source density, the dislocation density for θ = 45° is less than that for θ = 0°, in agreement with the results shown in Fig. 7.4.

**Fig. 7.6** Dimensionless dislocation density $\rho_d/\rho_s$ versus dimensionless contact load $P/bE$ and dimensionless indenter radius $R/b$ for slip-plane orientation angle (a) $\theta = 45°$ and (b) $\theta = 0°$. 
Figure 7.6 shows the dependence of dimensionless dislocation density $\rho_d/\rho_s$ on dimensionless contact load $P/bE$ and indenter radius $R/b$. The simulation results shown in Fig. 7.6 are for $R/d$ in the range of 2–200, corresponding to $R/b$ between $2 \times 10^2$ and $2 \times 10^4$. Because $P_f$ depends on $R$, which is a variable in this simulation case, the contact load is normalized by $bE$. Although dislocation multiplication is weakly dependent on the indenter radius for $\theta = 0^\circ$, for a given contact load, a sharper indenter generates more dislocations in slip-plane direction $\theta = 45^\circ$.

Fig. 7.7 Dislocation maps for different contact widths $2r$ (loads) and slip-plane orientation angle $\theta = 45^\circ$. (Positive and negative edge dislocations are shown by red and blue symbols, respectively.)
Figure 7.7 shows the effect of contact width $2r$ (or load) on the evolution of dislocations at slip planes oriented at $\theta = 45^\circ$ with respect to the half-space surface. (Slip planes are shown only in Fig. 7.7(a) for clarity.) Discrete dislocation activity is first encountered in the highly stressed region below the contact interface, indicated in each plot by a solid line of length $2r$. However, the location of initial dislocation dipoles depends on the random distribution of dislocation sources. As the contact width (load) increases, plasticity spreads throughout the subsurface region, resulting in the development of dislocation structures characterized by well-defined slip (shear) bands. Dislocation multiplication due to the load increase leads to the development of dislocation pile-ups and shear band expansion. A shear band is defined as the aggregate of slip planes with dislocations. The shear band width $w$ is equal to the product of the slip-plane distance $d$ and the number of slip planes with at least one dislocation. The dislocation-active region is below the contact interface and its width appears to scale to the contact width. Indeed, as shown in Fig. 7.8, the dimensionless shear band width $w/b$ increases continuously with the dimensionless contact load $P/P_Y$ up to a steady-state of $\sim 10^3$ for $P/P_Y > 6$, resembling strain hardening behavior.

Figure 7.9 shows contour maps of dimensionless total resolved shear stress $\tau/\tau_e$ corresponding to the dislocation maps shown in Fig. 7.7. The maps shown in Figs. 7.7 and 7.9 reveal intensification of the shear stress field with increasing contact load and simultaneous occurrence of large stress gradients at dislocation pile-ups. The evolution of these heterogeneous dislocation arrangements gives rise to local stress raisers that may act as precursors of microscopic failure mechanisms. This important
Fig. 7.9 Contours of dimensionless total resolved shear stress $\tau/\tau_e$ for different contact widths $2r$ corresponding to the dislocation maps shown in Fig. 7.7.

The physical phenomenon is intrinsic to the present discrete dislocation analysis and cannot be reproduced by phenomenological plasticity theories.

The effect of obstacle density on dislocation multiplication is illustrated in Fig. 7.10 for two characteristic slip-plane directions ($\theta = 0^\circ$ and $45^\circ$). Due to dislocation pinning when $\tau \leq \tau_o$ ($\tau_o = 3 \tau_e$), the densities of mobile dislocations are contrasted...
with the total dislocation densities. When the contact load is very low, most dislocations are not pinned by obstacles. For relatively low contact loads (e.g., \( P/P_Y < 2 \)), the effect of obstacle density on dislocation evolution is negligible. Dislocation pinning begins to affect dislocation activity for loads \( P/P_Y > 3 \). At high contact loads (e.g., \( P/P_Y > 6 \)), dislocation mobility is significantly restricted by the obstacles. For example, for \( \rho_o = \rho_s = 32 \, \mu m^{-2} \), about 50% of the dislocations generated at a load \( P/P_Y = 6 \) are pinned by obstacles (Fig. 7.10(a)). This trend is in qualitative agreement with the basic mechanism of dislocation hardening. Another interesting finding is the gradual decrease in mobile dislocation density with increasing contact load. This trend is attributed to the dominance of dislocation annihilation and pinning over dislocation emission at high loads.

Figure 7.11 shows the dimensionless mean contact pressure \( p_m/Y \) as a function of representative macroscopic strain \( \varepsilon \). The obstacle-free medium shows a softening effect with increasing dislocation source density (Fig. 7.11(a)), whereas the medium containing dislocation obstacles demonstrates a hardening effect with increasing obstacle density, for fixed dislocation source density (Fig. 7.11(b)). The softening behavior observed in Fig. 7.11(a) may be attributed to an intensifying effect of dislocation annihilation with the increase in dislocation source density. Alternatively, the hardening behavior seen in Fig. 7.11(b) is a consequence of the decrease in density of mobile dislocations with the increase of the obstacle density shown in Fig. 7.10.

**7.3.3. Plastic flow under theoretical strength**

Plastic flow of materials at stresses well below the theoretical strength \( \tau_c \) (Kovács and Zsoldos, 1973) is a consequence of the effect of dislocations. However, this is true only when the contact size is much larger than a characteristic microstructure length. Local plastic flow (defined as the formation of the first dislocation dipole) depends on the presence of dislocation sources in the region where the flow criterion \( \tau \geq \tau_c \) is first satisfied. Thus, in the absence of a dislocation source to generate slip, dislocations will not be generated even if the plastic flow criterion is locally satisfied. This situation may arise when the contact size is comparable to a characteristic length of the crystal microstructure (e.g., slip-plane distance). Because of the finite number of dislocation sources, the overall deformation behavior shows a strong dependence on the spatial distribution of dislocation sources. If at the instant of first dipole emission the maximum shear stress \( \tau_{\text{max}} \) in the loaded half-space (Johnson, 1987) is higher than the theoretical strength \( \tau_c \), plastic flow will occur under the theoretical strength by slip of a single atomic plane rather than the nucleation of dislocations, implying scale-dependent deformation of elastic-plastic asperity microcontacts at the contact interface of rough surfaces.

In sections 3.1 and 3.2, the characteristic microstructure length (i.e., slip-plane...
distance) was fixed at all contact widths (loads). However, it is well known that the near-surface microstructure is usually finer than the bulk material (Kuhlmann-Wilsdorf, 1981; Rigney, 1988). Thus, the characteristic microstructural length is likely to decrease with contact size. Therefore, it is of interest to consider the effect of simultaneously varying characteristic length and contact size on contact deformation. Thus, simulations were performed for fixed slip-plane distance and

![Graphs showing dimensionless total and mobile dislocation densities vs. dimensionless load](image)

**Fig. 7.10** Dimensionless total and mobile dislocation densities $\rho_d/\rho_s$ versus dimensionless load $P/P_Y$ and dislocation obstacle density $\rho_o$ for slip-plane orientation angle (a) $\theta = 45^\circ$ and (b) $\theta = 0^\circ$. 


gradually decreasing indenter radius to determine the threshold (critical) radius $R_{th}$ for plastic flow under the theoretical strength. Figure 7.12 shows the dimensionless threshold radius $R_{th}/d$ as a function of dimensionless slip-plane distance $d/b$ for two representative slip-plane directions ($\theta = 0^\circ$ and $45^\circ$). In all cases, $R_{th} < d$ and, more importantly, $R_{th}$ assumes much lower values for $\theta = 45^\circ$ than $\theta = 0^\circ$. In addition, $R_{th}$
approaches steady-state values of $0.5d$ and $-0.8d$ for $\theta = 45^\circ$ and $0^\circ$, respectively. The lower $R_{th}$ values for $\theta = 45^\circ$ are attributed to the higher probability of dislocation emission due to the intersection of the half-space surface by the

![Graph](image)

Fig. 7.12 Ratio of threshold radius of rigid indenter to slip-plane distance $R_{th}/d$ versus dimensionless slip-plane distance $d/b$ for slip-plane orientation angle $\theta = 0^\circ$ and $45^\circ$.

slip planes at $\theta = 45^\circ$. The higher $R_{th}$ predicted for single-crystal media with slip planes parallel to the surface reveals a microstructure effect on the scale-dependence of material deformation in the presence of dislocations.

### 7.4. Conclusions

Indentation of a single-crystal half-space by a rigid cylinder was analyzed by discrete dislocation plasticity. Long-range dislocation interaction was modeled by incorporating in the analysis the elastic stress fields of dislocations, while short-range dislocation interaction was represented by dislocation constitutive rules. Simulation results provided insight into the effects of dislocation source and obstacle densities, slip-plane orientation and distance, indenter radius, and contact load on damage at the onset of yielding (emission of first dipole) and plastic flow due to the development of dislocation structures. Plastic flow under the theoretical strength of the material was related to contact size effects. Based on the presented results and discussion, the following main conclusions can be drawn from this study.
1) Damage due to initial yielding associated with the emission of the first dislocation dipole is equal to ~30%, and increases with the dislocation source density and/or the decrease in slip-plane distance. The lowest contact load at which the first dislocation dipole commences corresponds to a 45° slip-plane orientation angle with respect to the surface normal, in accord with the onset of yielding predicted by continuum mechanics yield criteria.

2) After the onset of yielding, the dislocation density increases with the contact load, dislocation source density, and indenter sharpness. Although the most favorable slip-plane direction for first dislocation dipole emission is at 45° with respect to the half-space surface, the slip-plane direction parallel to the surface is the most favorable for dislocation multiplication. The shear-band width increases with contact load up to a steady-state value resembling macroscopic strain hardening.

3) Dislocation pile-ups generate local stress raisers that could play the role of precursors for microscopic deformation mechanisms. Dislocation pinning is augmented by an increase in obstacle density, in agreement with the basic dislocation hardening mechanism.

4) Plastic flow under the theoretical strength of the material may occur when the indenter radius is less than the slip-plane distance. The dependence of plastic flow on contact size is stronger for a slip-plane direction of 0° than 45° with respect to the half-space surface.
Chapter 8

A discrete dislocation plasticity analysis of a single-crystal semi-infinite medium indented by a rigid surface exhibiting multi-scale roughness

8.1. Introduction

Surface plasticity of contacting solid bodies is of fundamental importance in the study of friction, mechanical wear, contact fatigue, and thermoelectric conduction. A major source of energy dissipation in machine elements having contact interfaces is plastic deformation at asperity microcontacts. Plastic flow at contact interfaces could be detrimental to the reliability and performance of mechanical systems. Considerable effort has been devoted to develop analytical and numerical elastic-plastic models of surface and subsurface deformation behavior of contacting solids (Bhattacharya and Nix, 1988; Komvopoulos, 1989; Kral and Komvopoulos, 1996; Zisis and Giannakopoulos, 2009); however, these contact models rely on continuum mechanics approaches, which assume contact between homogeneous solids with ideally smooth surfaces. Although these approximations are generally acceptable at relatively large scales, contact-mode devices in various leading-edge technologies, such as microelectromechanical systems, magnetic recording, and semiconductor packaging, require contact analysis at submicrometer scales at which macroscopically smooth surfaces exhibit multi-scale roughness and contact is confined to asperity microcontacts of sizes typically ranging from a few nanometers to several micrometers. Consequently, contact deformation arises within a very small surface volume in the vicinity of microcontacts with sizes on the same order of magnitude as the characteristic microstructure scale, e.g., average size of dislocation cells in the near-surface region of the contacting solids. Because material behavior at such small scales is strongly affected by the discrete lattice structure, crystalline solids cannot be treated as homogeneous elastic-plastic media; thus classical (continuum) plasticity constitutive laws break down at such small scales (Kuhlmann-Wilsdorf, 1981; Pollock, 1992). Indeed, in a study of residual stresses due to contact between rough surfaces it was shown that the development of a near-surface layer under tension could not be analyzed by scale-invariant continuum plasticity (Wang et al., 2006).

A common approach among various analytical approaches dealing with scale effects on microscale plasticity of crystalline solids has been nonlocal phenomenological crystal plasticity (Shu and Fleck, 1999; Acharya and Bassani, 2000; Gurtin, 2002). These treatments consider the effect of geometrically necessary dislocations in deformation involving plastic strain gradients, such as those resulting from indentation. However, similar to classical plasticity, nonlocal formulations are
phenomenological and do not account for the fundamental basis of plasticity, i.e., the collective evolution of discrete dislocations. At very small scales typical of asperity microcontacts, microstructure and crystallography considerations suggest that discrete dislocation plasticity (Van der Giessen and Needleman, 1995; Polonsky and Keer, 1996a) is more appropriate for studying nonlinear plasticity. In this approach, the discrete nature of plastic flow is characterized by the collective motion of discrete edge dislocations along specific slip-planes, and dislocations are modeled as line singularities in a linear-elastic solid. Thus, unlike classical plasticity and nonlocal formulations in which deformation is averaged over dislocations, discrete dislocation plasticity treats dislocations individually as they move through a continuum medium. This approach allows for a true mechanism-based plasticity analysis in which plastic deformation does not follow constitutive rules derived from phenomenological material behavior.

The majority of previous discrete dislocation analyses of indented single-crystal media have been for a single asperity (indenter) (Kreuzer and Pippan, 2004b, 2007; Balint et al., 2006; Widjaja et al., 2007a; Yin and Komvopoulos, 2011). Indentation of a single-crystal material can be regarded as a coarse-grained model of plastic flow in a polycrystalline solid. Although a few multi-asperity indentation studies based on the discrete dislocation approach have been reported (Polonsky and Keer, 1996a; Nicola et al., 2007, 2008; Zhang et al., 2010), the indenter was described by Euclidean geometry and the contact interface was assumed to be perfectly smooth. While Polonsky and Keer (1996a) studied interactions between microcontacts, their analysis does not account for the interaction effect of different roughness wavelengths on deformation behavior because the microcontacts were modeled to be of the same shape and size. Nicola et al. (2007, 2008) and Zhang et al. (2010) simulated multi-asperity contact and nanoimprinting of flat single crystals, respectively, using discrete dislocation plasticity. However, these analyses are not applicable for real randomly rough surfaces because they are limited to a periodic array of identical flat indenters and are based on finite element modeling specific to certain indentation contact conditions.

Contact of real rough surfaces cannot be simply treated as a collection of single asperities. The importance of asperity interactions in the deformation behavior of contacting rough surfaces has been examined by Zhao and Chang (2001) and Ciavarella et al. (2006). The topography of rough surfaces has been commonly described by statistical parameters (Greenwood and Williamson, 1966; Jackson and Green, 2006). However, because these parameters depend on the sample length and the instrument resolution, they do not provide accurate representation of the multi-scale roughness of real surfaces. Fractal geometry has been used to characterize rough surfaces in contact mechanics (Majumdar and Bhushan, 1991; Persson et al., 2002; Willner, 2004; Komvopoulos and Yang, 2006; Komvopoulos and Gong, 2007; Yin and Komvopoulos, 2010) due to its properties of scale-invariance and self-affinity.
However, the former fractal studies rely on continuum mechanics approaches to model the deformation behavior of the contacting surfaces, which break down at small length scales.

Despite valuable insight into contact deformation of elastic-plastic rough surfaces exhibiting fractal behavior provided by previous studies, a discrete dislocation plasticity analysis of fractal surface contact that accounts for multi-scale asperity interaction has not been published yet. Therefore, the objective of this study was twofold. First, a discrete dislocation plasticity analysis of a single-crystal semi-infinite medium indented by a rough (fractal) rigid surface was developed based on constitutive rules of dislocation emission, glide, and annihilation. The effect of surface roughness of a single indenter on dislocation multiplication was examined in light of numerical results for rigid indenters with smooth and fractal surfaces. Second, the effects of contact load, slip-plane direction, dislocation source density, randomness of slip-plane spacing, and surface topography (fractal) parameters on the evolution of plasticity were analyzed in the context of analytical results revealing the development of dislocation structures and subsurface shear stresses under the effect of multi-scale asperity interactions.

8.2. Analytical model

8.2.1. Rough surface representation

To account for the self-affinity property of real surfaces over a wide range of length scales, the topography of the rough and rigid surface was characterized by fractal geometry. Thus, the profile of the rough surface is approximated by a truncated Weierstrass–Mandelbrot function \( z(x) \) (Berry and Lewis, 1980), which for dimensional consistency, can be written as (Wang and Komvopoulos, 1994a)\

\[
    z(x) = L \left( \frac{G}{L} \right)^{(D-1)n_{\text{max}}} \sum_{n=0}^{n_{\text{max}}} \frac{\cos(2\pi \gamma^n x / L)}{\gamma^{(2-D)n}}
\]  

where \( D \) and \( G \) are the fractal dimension (1<\( D < 2 \)) and fractal roughness, respectively, both independent of spatial frequency in the range where the rough surface exhibits fractal behavior, \( L \) is the fractal sample length, \( \gamma \) is a scaling parameter that controls the density of frequencies in the profile and has a typical value of 1.5 (Komvopoulos and Yan, 1997), and \( n \) is a frequency index with an upper limit \( n_{\text{max}} = \text{int}[\log(L/L_s)/\log\gamma] \), where \text{int}[...] indicates the integer part of the number in brackets, and \( L_s \) is the cut-off length.

The fractal parameter \( D \) indicates the relative contributions of low- and high-frequency components in the surface profile, and the fractal roughness \( G \)
controls the wave amplitudes over the entire range of frequencies comprising the surface profile. Lower $D$ and/or higher $G$ values yield rougher surfaces. A surface profile described by Eq. (8.1) is scale-invariant in a finite range of wavelengths, and the largest and the smallest wavelengths are determined by the fractal sample length $L$ and the cutoff length $L_S$, respectively. Equation (8.1) was used to generate the rough surfaces in the discrete dislocation plasticity analysis of indentation presented below.

To study the effect of multiscale asperity interactions on the evolution of plasticity and to reduce the computational time, a critical surface segment was determined by truncating the entire surface profile ($L = 5379$ nm) to a maximum interference of 11 nm. The critical profile segment obtained from this procedure (Fig. 8.1) has a length $L_C$ about 50 times larger than the typical local interference and contains the largest number of asperity microcontacts. The origin of the critical segment is denoted by $x_S$. The horizontal and vertical coordinates of the critical segment were normalized by its length.

8.2.2. Contact analysis

Figure 8.2 shows the critical segment of the rigid fractal surface indenting a single-crystal semi-infinite medium. As the crystalline medium interferes with the rough surface at a global interference $h$, asperity contacts are established within the apparent contact area. Due to the randomness of the fractal profile, a range of wavelengths contribute to the formation of asperity microcontacts for a given

Fig. 8.1  Critical segment of length $L_C = 250$ nm obtained from a fractal surface profile generated from Eq. (8.1) for $L = 5379$ nm, $L_S = 5$ nm, $\gamma = 1.5$, $D = 1.54$, and $G = 1.14$ nm ($\sigma = 29.2$ nm).
interference. The largest (base) wavelength in the waveform of a truncated asperity is equal to the truncated width $2r'$. It is presumed that the local contact force is due to the deformation of a spherical cap asperity of base diameter equal to the base wavelength (Yan and Komvopoulos, 1998), consistent with a previous study (Greenwood and Wu, 2001).

The contact interface is modeled by a deformable single-crystal surface indented by rigid cylindrical asperities, and radius of curvature of the $i^{th}$ asperity contact $R_i$ given by (Komvopoulos and Gong, 2007)

$$R_i = \frac{(r'_i)^D}{2^{(4-D)} G^{(D-1)}} \quad (8.2)$$

where $r'_i$ is the truncated half-width of the $i^{th}$ asperity contact.

From Hertzian analysis, the elastic deformation force at the $i^{th}$ asperity contact $\Delta P_i$ is expressed as (Johnson, 1987)

$$\Delta P_i = \frac{\pi r_i^2 E}{4R_i} \quad (8.3)$$

where $r_i = r'_i/2$ is the half-contact width due to elastic deformation of the $i^{th}$ asperity contact (Fig. 8.3).

The total deformation force $P$ at the contact interface was determined by summing up the forces generated at individual asperity contacts over the entire contact interface, i.e.,

$$P = \sum_{i=1}^{N} \Delta P_i \quad (8.4)$$

where $N$ is the total number of asperity contacts established across the contact interface for a given interference.

For a single asperity contact, the distribution of the contact pressure was determined from a numerical scheme in which the nodal contact pressure is proportional to the square root of the local interference. The resulting piecewise-linear distribution of the contact pressure, obtained from the superposition of overlapping triangular pressure elements (Johnson, 1987), is shown in Fig. 8.3. The maximum pressure $p'_j$ of the $j^{th}$ triangular pressure element at the $i^{th}$ asperity contact is given by (Komvopoulos and Gong, 2007)
Fig. 8.2 Schematic of single-crystal semi-infinite medium containing Frank-Read dislocation sources, which are randomly distributed on parallel and equally spaced slip-planes of fixed orientation indented by a rigid rough (fractal) surface. ($h_{\text{max}}$ = maximum global interference; $h_{12}$ = critical global interference for the interaction between asperities (1) and (2); and $h_{AB}$ = critical global interference for the interaction between asperities (A) and (B).)
\[ p_i' = \frac{(\delta_i^j)^{1/2}}{M_j} \frac{\Delta P_i}{\varepsilon} \] (8.5)

where \( \delta_i^j \) is the local interference at the \( j \)-th node of the \( i \)-th asperity contact, \( M_j \) is the total number of grid nodes at the \( i \)-th asperity contact, and \( \varepsilon \) is the grid size. It is noted that using a piecewise-linear distribution to approximate the contact pressure yields continuous surface displacements.

The contact stresses in the crystalline medium due to indentation loading were obtained from the superposition of the stress fields generated by the triangular distributions of all the pressure elements at each asperity contact and another superposition of the stress fields of all the asperity contacts formed at the contact interface. Therefore, the shear stress at a material point \((x_1, x_2)\) in the semi-infinite medium along a certain direction due to indentation loading (Hertzian contact) \( \tau^H \) can be written as

\[ \tau^H(x_1, x_2) = \sum_{i=1}^{N} \sum_{j=1}^{M_j-1} \tau^H_j(x_1, x_2) \] (8.6)
where $\tau_{ij}^H$ is the Hertzian shear stress along a certain direction due to the $j^{th}$ triangular contact pressure distribution at the $i^{th}$ asperity contact.

The Hertzian contact stresses at a point $(x_1, x_2)$ due to the $j^{th}$ triangular contact pressure distribution at the $i^{th}$ asperity contact can be expressed as (Johnson, 1987):

$$
\sigma_{11}^{Hj} = -\frac{p_j}{\varepsilon} m \left( 1 + \frac{n^2 + x_1^2}{n^2 + m^2} \right) + 2x_2
$$ (8.7)

$$
\sigma_{22}^{Hj} = -\frac{p_j}{\varepsilon} n \left[ 1 - \frac{n^2 + x_2^2}{n^2 + m^2} \right]
$$ (8.8)

$$
\tau_{12}^{Hj} = -\frac{p_j}{\varepsilon} n \left[ \frac{m^2 - x_2^2}{n^2 + m^2} \right]
$$ (8.9)

where $\sigma_{11}^{Hj}$ and $\sigma_{22}^{Hj}$ are the normal stresses in the $x_1$- and $x_2$-directions, respectively, $\tau_{12}^{Hj}$ is the shear stress, and $m$ and $n$ are defined as

$$
m^2 = \frac{1}{2} \left\{ \left( \varepsilon^2 - x_1^2 + x_2^2 \right)^2 + 4x_1^2x_2^2 \right\}^{1/2} + \left( \varepsilon^2 - x_1^2 + x_2^2 \right)
$$ (8.10a)

$$
n^2 = \frac{1}{2} \left\{ \left( \varepsilon^2 - x_1^2 + x_2^2 \right)^2 + 4x_1^2x_2^2 \right\}^{1/2} - \left( \varepsilon^2 - x_1^2 + x_2^2 \right)
$$ (8.10b)

where $\text{sign}(m) = \text{sign}(x_2)$ and $\text{sign}(n) = \text{sign}(x_1)$.

8.2.3. Discrete dislocation plasticity analysis

Because modeling plastic deformation by real crystal dislocations is computationally intensive, the single-crystal medium is modeled as an initially stress-free, homogeneous, isotropic elastic material containing parallel slip-planes with crystal dislocations represented by emitted subsurface dislocations. It is assumed that partial dislocations do not exist and that the Burgers vector of all dislocations is equal to the lattice distance of the crystalline solid.

The crystalline medium contains parallel slip-planes oriented at an angle $\theta$ with respect to the free surface, as shown in Fig. 8.2. It is further presumed that dislocations nucleate only from Frank-Read sources (Van der Giessen and Needleman, 1995), which are randomly distributed on slip-planes, and also that dislocation activity in the bulk is the dominant plastic deformation process, i.e., nucleation of dislocations...
from surface sources, as in the study of Nicola et al. (2007), is not considered in the present analysis. In each simulation, the location and density of dislocation sources \( \rho_s \) (per unit area) are fixed. When a critical shear stress is applied to a dislocation source, bowing of Frank-Read segments leads to the formation of new dislocation loops. If the total resolved shear stress \( \tau \) due to the applied total contact force \( P \) exceeds the dislocation emission stress \( \tau_e \) at a Frank-Read source, a pair of edge dislocations (dislocation dipole) with anti-parallel Burgers vectors \( \pm b \) develops on the slip-plane, with the source centered between the two emitted edge dislocations (Van der Giessen and Needleman, 1995). The dislocation emission stress indicates the lattice resistance to dislocation motion, and it is generally equal to a small fraction of the shear modulus (Kovács and Zsoldos, 1973). The Burgers vectors of the produced dislocations are parallel to the slip-planes, while the dislocation lines are perpendicular to the plane of deformation. The sign of the dipole is controlled by the sign of the total resolved shear stress at the source location along the slip-plane. From the balance of the shear stress exerted by the two dislocations on each other and the total resolved shear stress, the distance between the two emitted dislocations \( l_e \) is obtained as (Van der Giessen and Needleman, 1995)

\[
l_e = \frac{E}{4\pi(1-\nu^2)} \frac{b}{\tau_e}
\]

(8.11)

where \( b \) is the magnitude of the Burgers vector, and \( E \) and \( \nu \) are the elastic modulus and Poisson’s ratio, respectively. Macroscopic plastic strain is a manifestation of displacement discontinuities between dislocation dipoles. However, the stress field outside the core of each initiated edge dislocation is described by linear elasticity.

In the present analysis, contact is assumed to be frictionless contact, implying that energy dissipation is mainly due to subsurface plasticity. The effect of contact friction is not considered because the focus is on subsurface plastic flow due to indentation loading, not on the effects of surface contact phenomena on deformation. Therefore, the total resolved shear stress \( \tau \) at the location of a dislocation source on a slip-plane is determined as the sum of the Hertzian shear stress \( \tau^H \) due to the indentation of the semi-infinite medium by the rough surface (Eq. (8.6)) and the shear stress at the same location due to the long-range (elastic) stress fields of all existing dislocations, except for the dislocation (if any) at that location, i.e.,

\[
\tau = \tau^H + \sum_k \tau^d_k
\]

(8.12)

where \( \tau^d_k \) is the shear stress generated at the location of the particular dislocation source by the \( k^{th} \) dislocation. The dislocation shear stress field in the semi-infinite medium is obtained as the solution of a boundary value problem determined from the
complex function analysis of Kreuzer and Pippan (2004b). This method is more advantageous than that of Van der Giessen and Needleman (1995), which is computationally intensive because it uses a separate finite element model for each dislocation structure generated during loading to obtain the image component of the dislocation stress field.

At any deformation stage, the crystalline medium is not in thermodynamic equilibrium because of the highly nonuniform dislocation stress field and complex dislocation-dislocation and dislocation-surface interactions. Nonequilibrium may result in dislocation motion and lead to dislocation redistribution along the slip-planes. For simplicity, it is assumed that nucleated dislocations glide along existing slip-planes, while dislocation climb or cross slip are not modeled, implying crystalline medium exhibiting anisotropic plasticity. The dislocation glide velocity $u_g$ can be obtained from the linear drag relation, given by (Van der Giessen and Needleman, 1995)

$$u_g = \frac{b \tau}{K} \quad (8.13)$$

where $K$ is the drag coefficient. When the distance between two opposite dislocations on the same slip-plane is less than the critical annihilation distance $l_a$, an intrinsic material parameter, they are annihilated and removed from the simulation. Annihilation of multiple dislocations is treated as a sequence of annihilated pairs of dislocations. Dislocations exiting through the surface, including the contact region, are also removed from the simulation as no longer contributing to the total resolved shear stress field.

To perform a history-dependent plasticity analysis, an incremental formulation that uses a prescribed time step and corresponding incremental load applied to the indenting rough surface was used in all simulations. By capturing the details of the surface profile at small load (interference) increments, this approach accounts for the effects of smaller wavelengths on top of the base wavelength of a truncated asperity contact, even though for a given interference only the largest wavelength in the asperity waveform is considered in the contact force analysis. At each time step, the total resolved shear stress field was determined first, and nucleation of new dislocations, movement of existing dislocations, annihilation of opposite dislocations, and exodus of dislocations from the surface were then simulated based on the dislocation constitutive rules described above. This procedure was repeated at each load step to determine the evolving dislocation structures and corresponding shear stress field.

8.3. Results and discussion
Numerical results are presented in this section for fractal surface profiles generated from Eq. (8.1) for $L = 5379\, \text{nm}$, $L_S = 5\, \text{nm}$, $L_C = 250\, \text{nm}$, and $\gamma = 1.5$. The grid size $\varepsilon$ was set equal to one tenth of the cut-off length $L_S$ (i.e., $\varepsilon / L_S = 0.1$). The material properties of the semi-infinite medium, given in Table 8.1, are characteristic of those of single-crystal copper. For typical micromachined and/or nanofabricated surfaces, $d$ can be set approximately equal to the mean sub-grain size in the near-surface layer (Rigney, 1988). Unless otherwise stated, the results presented below are for $d/b = 50$ and $l_a/b = 5$. To provide generalized solutions, in most cases simulation results are presented in the form of dimensionless parameters. Due to the random distribution of the dislocation sources, a statistical analysis was carried out for each set of parameters. Hence, numerical results are shown in the form of data points representing mean values determined from five simulations of different source distributions and identical all other parameters. Since the effect of loading rate on the deformation behavior was found to be insignificant, all simulation results discussed below are for a constant loading rate ($7.2 \times 10^{23} \, \text{N/m} \cdot \text{s}$).

### Table 8.1 Material properties of single-crystal semi-infinite medium. *(a)*

<table>
<thead>
<tr>
<th>$b$ (nm)</th>
<th>$E$ (GPa)</th>
<th>$\nu$</th>
<th>$\tau_e$ (MPa)</th>
<th>$K$ (Pa·s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>110</td>
<td>0.34</td>
<td>78</td>
<td>$1.7 \times 10^{-5}$</td>
</tr>
</tbody>
</table>

*(a)* Sources: Jassby and Vreeland, 1970; Ross, 1992; Van der Giessen and Needleman, 1995

8.3.1. Indentation by a smooth or rough asperity

The increase in dislocation density $\rho_d$ (per unit area) during contact loading provides insight into the evolution of subsurface plasticity. To examine the roughness effect on the multiplication of dislocations due to single-asperity indentation, simulations were performed for rigid asperities with smooth (cylindrical) or rough (fractal) profiles indenting the same single-crystal medium. The smooth asperity has a radius of curvature $R = 20000b$, while the rough asperity has fractal parameters: $G = 9.46 \times 10^{-4}\, \text{nm}$ and $D = 1.25$. Figure 8.4(a) shows the dimensionless dislocation density $\rho_d / \rho_s$ as a function of dimensionless contact load $P/bE$ for smooth and rough (fractal) asperities indenting a single-crystal medium with slip-planes oriented at $\theta =$
45°. Although the dislocation density increases with contact load in both cases, higher dislocation densities were produced by the rough asperity through the entire load range, especially for high contact loads. This is attributed to the higher stresses generated in the semi-infinite medium by the rough asperity.

The multiplication of dislocations due to the increase of the contact load results in the formation of wider slip (shear) bands. A shear band is defined as a cluster of slip-planes with dislocations, and its width $w$ is equal to the number of slip-planes having at least one dislocation multiplied by the slip-plane distance $d$. Figure 8.4(b) shows that the dimensionless shear band width $w/b$ increases continuously with the contact load $P/bE$, following a trend similar to that of the dislocation density (Fig. 8.4(a)). The wider shear bands produced by the rough asperity suggest that roughness enhanced plastic flow in the crystalline medium.

8.3.2. Indentation by a rough surface

Figure 8.5(a) shows the dimensionless dislocation density $\rho_d/\rho_s$ as a function of dimensionless contact load $P/bE$ and slip-plane orientation angle $\theta$ due to the indentation of a single-crystal medium by a rigid fractal surface for fixed dislocation source density ($\rho_s = 32 \mu m^{-2}$) and surface topography parameters ($G = 2.93 \times 10^{-2}$ nm and $D = 1.54$). The dislocation density increases monotonically with the contact load for all slip-plane directions; however, the most favorable slip-plane orientation for dislocation multiplication is parallel to the surface ($\theta = 0^\circ$). Consequently, all simulation results presented next are for $\theta = 0^\circ$. Figure 8.5(b) shows the variation of the dislocation density $\rho_d$ with the dimensionless contact load $P/bE$ and the density of Frank-Read sources $\rho_s$ for same rough surface profile used to obtain the results shown in Fig. 8.5(a). It can be seen that the dislocation density increases nonlinearly with the contact load and the dislocation source density.

Figure 8.6 shows the effect of contact loading on the evolution of dislocation structures at slip-planes (shown only in Fig. 8.6(a) for clarity) of the semi-infinite medium indented by the rough surface profile shown in Fig. 8.1. Contact regions between the asperities of the rough surface and the indented medium are denoted by solid lines at the top boundary ($x_2 = 0$) of each plot. At a critical load (or global surface interference), discrete dislocations were generated in the highly stressed region below the first established asperity contacts (Fig. 8.6(a)). Upon further increase of the contact load, new asperity contacts are established and the dislocation density increases rapidly; however, the distance between asperity contacts is relatively large and dislocation-active regions (plastic zones) are not affected by the stress fields of neighboring contacts (Figs. 8.6(b) and 8.6(c)). At higher contact loads, the contact regions approach each other, and the resulting asperity interaction leads to rapid multiplication of dislocations and the spreading of plasticity throughout the subsurface region (Fig. 8.6(d)). It is interesting to note that the depth of the plastic zone is much
larger than the contact width, which is in sharp contrast with predictions based on continuum plasticity theories according to which, plastic deformation is confined within the vicinity of each asperity contact.

Fig. 8.4  (a) Dimensionless dislocation density $\rho_d/\rho_s$ and (b) dimensionless shear-band width $w/b$ versus dimensionless contact load $P/bE$ for a single-crystal semi-infinite medium with slip-plane orientation angle $\theta = 45^\circ$ indented by a rigid, smooth cylindrical asperity with dimensionless radius of curvature $R/b = 2 \times 10^4$ and a rigid, rough (fractal) asperity with topography parameters $D = 1.25$ and $G = 9.46 \times 10^{-4}$ nm.
The results presented in Fig. 8.7 can be used to further interpret the effect of multi-scale asperity interaction at rough contact interfaces on the evolution of

Fig. 8.5  (a) Dimensionless dislocation density $\rho_d/\rho_s$ versus dimensionless contact load $P/bE$ and slip-plane orientation angle $\theta$ for a single-crystal semi-infinite medium indented by a rigid, rough (fractal) surface with topography parameters $D = 1.54$ and $G = 2.93 \times 10^{-2}$ nm, and (b) dislocation density $\rho_d$ versus dimensionless contact load $P/bE$ and dislocation source density $\rho_s$ for a single-crystal semi-infinite medium with slip-plane orientation angle $\theta = 0^\circ$ indented by the same rigid, rough surface.
Fig. 8.6 Evolution of dislocation structures below the surface of a single-crystal semi-infinite medium with slip-plane orientation angle $\theta = 0^\circ$ for dimensionless global interference $h/d$ equal to (a) 0.10, (b) 0.26, (c) 0.58, and (d) 0.90. Contact widths between asperities of the rough surface and the crystalline medium are denoted by thick lines at the medium surface ($x_2 = 0$). Positive and negative edge dislocations are shown by red and blue symbols, respectively. Slip-planes are only shown in (a) for clarity. The region on the left of the vertical dashed line in each figure was used to obtain the dislocation density plotted in Fig. 8.7.

subsurface plasticity. Figure 8.7 shows the dimensionless dislocation density $\rho_d/\rho_s$ (corresponding to each region on the left of the vertical dashed lines shown in Fig. 8.6) versus the dimensionless global interference $h/d$ for three different surface profiles generated from the fractal surface shown in Fig. 8.2. Two different length scales of
asperity interaction occur with the increase of the global interference. Interaction between small and relatively smooth asperities (1) and (2) occurs at a global interference $h_{12}/d \approx 0.08$, as evidenced by the deviation of the dislocation density predicted for asperity (1) from that for the larger (and rougher) asperity (A), which contains asperities (1) and (2), and also the dislocation map shown in Fig. 8.6(b). As the global interference (load) increases, a new deviation in the dislocation density is observed at $h_{AB}/d \approx 0.43$ due to the interaction of the larger (and rougher) asperities (A) and (B), also shown in Fig. 8.6(c). For even larger global interference, the dislocation density is determined by the rough surface and the intensified asperity interaction induces plastic flow deep into the subsurface. Deviations between dislocation densities corresponding to indenting surface profiles of different length scales reflect interactions between the subsurface stress fields of neighboring asperity contacts, resulting in faster multiplication of dislocations than isolated asperity contacts for the same global interference. A comparison of the critical global interferences for asperity interaction with the corresponding distance of interacting asperities leads to the following relationship:

![Fig. 8.7 Dimensionless dislocation density $\rho_d/\rho_s$ versus dimensionless global interference $h/b$ for a single-crystal semi-infinite medium with slip-plane orientation angle $\theta = 0^\circ$ indented by a rough surface, a relatively large and rough asperity (A), and a relatively small and smooth asperity (1). Asperities (A) and (1) are defined in Fig. 8.2. Vertical dashed lines indicate the critical dimensionless global interference for asperity interaction at different length scales.](image)

150
where \( l_{12} \) is the distance between asperities (1) and (2) at a global interference \( h_{12} \) and \( l_{AB} \) is the distance between asperities (A) and (B) at a global interference \( h_{AB} \) (Fig. 8.2). Therefore, it may be inferred that in addition to the fractal surface profile, the initiation of interactions between dislocations under neighboring asperity contacts also demonstrates self-similarity at different asperity length scales.

Accurate estimation of the subsurface shear stress field due to indentation is essential for the analysis of microscale failure mechanisms where dislocations play a dominant role. Figure 8.8 shows isostress contours of the dimensionless total resolved shear stress in the slip-plane direction \( \tau / \tau_c \) corresponding to the dislocation structures shown in Fig. 8.6. The subsurface shear stress intensifies dramatically with the increase of the global interference. The dislocation maps and stress contours shown in Figs. 8.6 and 8.8, respectively, reveal high shear stress spikes due to the pile-up of dislocations. These high localized shear stresses may be precursors of microscale failure mechanisms, such as microcrack nucleation and propagation, in the vicinities of asperity contacts. This important physical characteristic captured in the present discrete dislocation analysis cannot be reproduced by continuum theories of plasticity.

To examine discrete dislocation plasticity of single-crystal microstructures with random slip-plane distributions, simulations were performed with semi-infinite media exhibiting random slip-plane distances indented by the same rough surface used to obtain the results shown in Fig. 8.5. In these simulations, the average value of the slip-plane distance was \( d = 50b \), while the standard deviation \( s \) was varied between 0 and 1.58\( d \). Figure 8.9 shows the dependence of the dimensionless dislocation density \( \rho_d / \rho_s \) on the dimensionless contact load \( P/bE \) for a single-crystal medium with dislocation source density \( \rho_s = 32 \, \mu m^{-2} \) and slip-plane orientation angle \( \theta = 0^\circ \) indented by a rigid fractal surface with surface topography parameters \( G = 2.93 \times 10^{-2} \) nm and \( D = 1.54 \). As in previous simulations, the dislocation density increases monotonically with the contact load; however, increasing the standard deviation of the slip-plane distance yielded significantly higher dislocation densities. This is attributed to the stronger dislocation interactions encountered in the media with larger variation in slip-plane distance, leading to intensification of the subsurface shear stress field and, consequently, nucleation of more dislocations.

The topography effect of the indenting rough surface on the evolution of plasticity in the indented medium can be further examined in the context of simulation results obtained with rough surfaces characterized by different fractal parameters, resulting in root-mean-square roughness \( \sigma \) in the range of 9.6–107.5 nm. Figure 8.10(a) shows the effect of varying the rms roughness \( \sigma \) by changing the fractal roughness \( G \) while keeping the fractal dimension \( D \) fixed (\( D = 1.54 \)) on the
Fig. 8.8  Contours of dimensionless total resolved shear stress $\tau/\tau_e$ for dimensionless global interference $h/d$ equal to (a) 0.10, (b) 0.26, (c) 0.58, and (d) 0.90 corresponding to the dislocation structures shown in Fig. 8.6. Contact widths between asperities of the rough surface and the crystalline medium are denoted by thick lines at the medium surface ($x_2 = 0$).
Fig. 8.9  Dimensionless dislocation density $\rho_d/\rho_s$ versus dimensionless contact load $P/bE$ and dimensionless standard deviation of slip-plane distance $s/d$ for a single-crystal semi-infinite medium with randomly distributed slip-planes of orientation angle $\theta = 0^\circ$ indented by a rigid, rough (fractal) surface with topography parameters $D = 1.54$ and $G = 2.93 \times 10^{-2} \text{ nm}$.

development of dislocations in a semi-infinite medium with $\rho_s = 32 \text{ } \mu\text{m}^{-2}$ and $\theta = 0^\circ$. For a given contact load, more dislocations were produced by the rougher surface (higher $G$ and $\sigma$ values) due to intensification of the subsurface shear stress field, especially at higher contact loads. Figure 8.10(b) shows the effect of the fractal dimension $D$ on the development of dislocations in the same medium for fixed fractal roughness ($G = 1.14 \text{ nm}$). The dislocation density shows a weak dependence on fractal dimension despite the variation of the surface roughness by an order of magnitude. Although the decrease in fractal dimension increases the surface roughness, it also increases the average distance between asperity contacts, which decreases the effect of asperity interaction on dislocation multiplication. Thus, the intensification of the shear stress fields below contacting asperities due to the roughness increase is offset by the weakening of the interaction of dislocations under different microcontacts due to the larger distance between asperity contacts.
Fig. 8.10  Dimensionless dislocation density $\rho_d/\rho_s$ versus dimensionless contact load $P/bE$ for a single-crystal semi-infinite medium with slip-plane orientation angle $\theta = 0^\circ$ indented by rigid, rough (fractal) surfaces of different rms roughness $\sigma$: (a) $D = 1.54$ and $G = 2.93 \times 10^{-2}$ nm ($\sigma = 10.8$ nm), 1.14 nm ($\sigma = 29.2$ nm), and 5.48 nm ($\sigma = 101.3$ nm), and (b) $G = 1.14$ nm and $D = 1.24$ ($\sigma = 107.5$ nm), 1.54 ($\sigma = 29.2$ nm), and 1.84 ($\sigma = 9.6$ nm).
The two characteristic lengths in the present study are the standard deviation of the slip-plane distances $s$ and the rms roughness (fractal parameters) of the indenting rough surface $\sigma$. The separate effects of these two parameters on dislocation multiplication were discussed in light of Figs. 8.9 and 8.10. The simultaneous effects of these parameters on the development of dislocations with the increase of the contact load are shown in Fig. 8.11. Because of the stronger effect of the fractal roughness on the formation of dislocations than the fractal parameter, the results shown in Fig. 8.11 were obtained by varying $G$ (in order to change $\sigma$), while keeping the fractal dimension constant ($D = 1.54$). For a given contact load, higher dislocation densities were obtained with increasing $\sigma/s$. Hence, although the dislocation density at a given contact load increases with both $s$ and $G$ (Figs. 8.9 and 8.10(a)), the roughness of the indenting surface plays a dominant role in determining the plastic deformation behavior of the crystalline medium due to indentation loading.

Fig. 8.11 Dimensionless dislocation density $\rho_d/\rho_s$ versus dimensionless contact load $P/bE$ and surface roughness-to-standard deviation of slip-plane distance ratio $\sigma/s$ for a single-crystal semi-infinite medium with slip-plane orientation angle $\theta = 0^\circ$ indented by rigid, rough (fractal) surfaces with $D = 1.54$ and different $G$ values.

8.4. Conclusions

Indentation of a single-crystal semi-infinite medium by a rigid rough (fractal) surface was analyzed by discrete dislocation plasticity. This is the first quantitative
microcontact analysis of crystalline materials indented by a surface exhibiting multi-scale roughness based on discrete dislocations. Short-range dislocation interactions were modeled using dislocation constitutive rules, while long-range dislocation interactions were modeled by the elastic stress fields of edge dislocations. The obtained results provide insight into the effects of contact load, topography (roughness) of indenting surface, dislocation source densities, slip-plane orientation and distribution, and multi-scale interactions of asperity microcontacts on plastic flow represented by dislocation structures. In view of the obtained results and discussion, the following main conclusions can be drawn from this study.

1) Higher dislocation densities and wider shear bands were obtained with a rough (fractal) asperity than a smooth cylindrical asperity under the same contact load due to higher contact stresses produced by the rough asperity.

2) The slip-plane direction parallel to the surface of the indented medium is the most favorable direction for dislocation multiplication.

3) The dislocation density increases with the contact load, dislocation source density, standard deviation of slip-plane distance, and fractal roughness of the indenting surface, but is approximately independent of the fractal dimension.

4) The roughness of the indenting surface demonstrates a more dominant effect on the evolution of plasticity than the standard deviation of the slip-plane distance.

5) As the contact load increases, asperity contact interactions evolve at different wavelength scales. Similar to the indenting surface, dislocation interactions below neighboring asperity contacts demonstrate self-similarity.

6) Large-scale plastic flow in the crystalline medium, which cannot be captured by continuum plasticity theories, occurs due to strong long-range dislocation interactions caused by the intensification of the subsurface stress field due to asperity interaction.
Chapter 9

Conclusion

Semi-infinite homogeneous and layered media with smooth, patterned, and rough surfaces subjected to mechanical contact and/or surface forces were examined to elucidate the effects of surface patterning, surface topography, and crystalline microstructure on the resulting deformation behavior and evolution of subsurface stresses. The following main conclusions can be drawn from the results and discussions presented in previous chapters.

Finite element results of plane-strain contact analysis of a rigid flat indenting an elastic-plastic semi-infinite medium with a sinusoidal surface profile showed a uniform rise of the surface material for relatively large interferences, in agreement with experimental observations. The contact pressure increased rapidly when full surface conformity was approached due to the effect of asperity interaction, which constrained subsurface deformation in the vicinity of asperity contacts. Differences between finite element and analytical results were interpreted by accounting for the effect of bulk elastic deformation on surface deformation.

Plane-strain finite element analysis of patterned chromium-quartz microstructures subjected to dynamic pressure loadings representative of those in megasonic cleaning revealed two failure modes: shear band formation in the quartz substrate from stress raiser points (e.g., cavity corners) and cracking in the bulk of the chromium layer (cohesive fracture) or the layer/substrate interface (delamination). Because the collapse pressure of the examined pattern microstructures was found to be higher than the pressure due to bubble cavitation, these mask patterns are not expected to fail instantaneously but undergo fatigue failure during megasonic cleaning. Microstructures with a larger undercut depth, larger zero-aperture distance, thinner chromium layer, and smaller line width showed a higher damage probability. In addition, the likelihood for microstructure failure due to plastic deformation was found to increase significantly with megasonic frequency, especially for chromium layer thickness and line width less than a threshold value, and undercut depth and zero-aperture distance greater than a critical value. A residual stress in the chromium layer was found to enhance microstructure failure and to diminish the structural integrity.

From a contact mechanics analysis of a hard and rough (fractal) surface with its asperities plowing and cutting through a soft and smooth substrate based on the slip-line theory of plasticity, the friction coefficient, energy dissipated in plastic microcontacts, and abrasive wear coefficient (rate) of representative metal-metal,
metallic-ceramic, and ceramic-ceramic contact systems were found to vary significantly, depending on surface topography (roughness), elastic-plastic material properties of the soft surface, global interference (normal load), and interfacial shear strength (adhesion) between the interacting surfaces. The wear coefficients of the examined contact systems were found to be on the order of $10^{-2}$, which is typical of two-body abrasive wear. The friction coefficient, energy dissipated in material removal (wear particle formation), wear rate, and wear coefficient increased with the root-mean-square roughness of the rough surface. Friction was shown to increase with interfacial adhesion, but to depend weakly on global interference (normal load). Both the energy dissipated in the form of material removal and the abrasive wear coefficient decreased with increasing interfacial shear strength (adhesion) and global interference (normal load). For adhesionless interfaces, both friction and energy dissipation due to plastic deformation and material removal were found to be approximately independent of surface roughness and global interference. An inverse proportionality of the friction coefficient to the ratio of the elastic modulus to the hardness of the soft material was observed, while the energy dissipated in material removal and the wear coefficient were found to vary proportionally with the elastic modulus-to-hardness ratio.

Based on an analysis of adhesive wear of rough (fractal) surfaces in normal contact, the plastic deformation at asperity contacts and the adhesive wear rate and wear coefficient were examined in terms of the topography (roughness) of the contacting surfaces, global interference (normal load), elastic-plastic material properties, and work of adhesion which depends on the surface energies of the contacting surfaces and interfacial adhesion affected by the material compatibility and contact environment. Elastic deformation was found to dominate at the asperity level over a wide range of global interference (normal load). The wear coefficients of the analyzed contact systems were found to vary in the range of $\sim 10^{-4} - 10^{-3}$ which is typical of severe adhesive wear. Both wear rate and wear coefficient increased with surface roughness and decreased with work of adhesion (interfacial adhesion). The wear rate increased monotonically with the global interference (normal load), whereas the wear coefficient decreased rapidly to a steady-state value.

A discrete dislocation plasticity analysis of plane-strain indentation of a single-crystal semi-infinite medium by a rigid smooth (cylindrical) asperity, a rigid rough (fractal) asperity, and a rigid rough (fractal) surface was also performed. It was shown that the dislocation density increases with contact load, dislocation source density, standard deviation of slip-plane distance, sharpness of smooth indenter, and fractal roughness of indenting surface, but is approximately independent of the fractal dimension of the indenting surface. The slip-plane direction parallel to the medium surface was found to be the most favorable direction for dislocation multiplication. Dislocation pinning by obstacles was enhanced with the increase in obstacle density, in accord with the basic mechanism of dislocation hardening. Dislocation pile-ups
induced local stress raisers, which could play the role of precursors for microscale failure mechanisms. For a single-crystal medium indented by a rigid smooth (cylindrical) asperity, crystal damage due to initial yielding was found to be equal to ~30% and to increase with dislocation source density and the decrease of the slip-plane distance. First dislocation dipole emission occurred at a 45° slip-plane orientation angle with respect to the surface normal, in agreement with continuum mechanics yield criteria. Plastic deformation at the theoretical strength of the material was found for indenter radius less than the slip-plane distance. A comparison between results obtained for smooth and rough (fractal) asperities revealed higher dislocation densities for rough asperities for the same contact load because of the higher contact stresses generated by the rough asperity. For a single-crystal medium indented by a rigid rough (fractal) surface, asperity interactions evolve at different wavelength scales with the increase of the contact load. Similar to the indenting rough (fractal) surface, dislocation interactions below neighboring asperity microcontacts demonstrated self-similarity. Plastic flow throughout the crystalline medium, which cannot be captured by continuum plasticity theories, occurred as a result of strong long-range dislocation interaction due to intensification of the subsurface shear stress field caused by interacting asperity microcontacts.

In conclusion, the contribution of this dissertation in the field of contact mechanics is the further fundamental insight into elastic-plastic deformation and tribological characteristics of interacting solid surfaces with surfaces exhibiting single- or multi-scale roughness. Numerical simulations and analytical results revealed the role of various important parameters, such as surface patterning, rough surface topography, interfacial adhesion, and crystalline microstructure on the nanoscale mechanical behavior and tribological properties of metal, ceramic, and thin-film materials used in various industry sectors.


Derjaguin, B.V., Muller, V.M., Toporov, Y.P., 1975. Effect of contact deformations on


of Tribology 127, 503–512.


164


Appendix A

Determination of slip-line field angles

The geometric and trigonometric relationships presented here refer to Fig. A.1.

A.1. Determination of $\alpha_3$, $\alpha_1$, and $r_3$

The angle change from point I to point E along IHE is equal to the angle change moving along IFE, i.e.,

$$\alpha_1 + \alpha_3 = \theta + \phi_0$$  \hspace{1cm} (A1)

Definitions:

$$r_1 = (BF) = (DF) = (DI) = (BD)/\sqrt{2}$$  \hspace{1cm} (A2)

$$r_3 = (O_3E) = (O_3H)$$  \hspace{1cm} (A3)

$$r_5 = (O_5M) = (O_5H')$$ \hspace{1cm} (A4)

$$R_1 = (O_2I) = (O_2H) = (O_2H')$$  \hspace{1cm} (A5)

Triangle $(O_3O_2H)$:

$$(O_3O_2)^2 = (O_2H)^2 + (O_3H)^2$$  \hspace{1cm} (A6)

$$\angle O_3O_2H = \tan^{-1}\left[\frac{O_3H}{O_2H}\right]$$  \hspace{1cm} (A7)

$$\angle O_3O_2I = \angle O_3O_2H - \alpha_1$$  \hspace{1cm} (A8)

Triangle $(DO_2I)$:

$$(DO_2)^2 = (O_2I)^2 + (DI)^2$$  \hspace{1cm} (A9)

$$\angle DO_2I = \tan^{-1}\left[\frac{DI}{O_2I}\right]$$  \hspace{1cm} (A10)

$$\angle DO_2O_3 = \angle DO_2I - \angle O_3O_2I$$  \hspace{1cm} (A11)

Triangle $(O_3O_2D)$:

$$(O_3D)^2 = (DO_2)^2 + (O_3O_2)^2 - 2(DO_2)(O_3O_2)\cos \angle DO_2O_3$$  \hspace{1cm} (A12)

Triangle $(OBD)$:

$$(OB) = R$$  \hspace{1cm} (A13)

$$\frac{1}{\cos\left(\pi-\frac{\pi}{4}-\eta\right)}$$  \hspace{1cm} (A14)

$$\angle OBD = \frac{3}{4}\pi + \eta$$  \hspace{1cm} (A15)

$$\angle BDO = \pi - \angle BOD - \angle OBD$$  \hspace{1cm} (A16)
\[(OD)^2 = (OB)^2 + (BD)^2 - 2(OB)(BD)\cos \angle OBD \quad (A17)\]
\[(DB)/\sin \angle BOD = (OB)/\sin \angle BDO \quad (A18)\]

Triangle \((O_3OE)\):
\[(OE) = R \quad (A19)\]
\[\angle O_3EO = \pi - \eta \quad (A20)\]
\[\angle OO_3E = \pi - \angle O_3EO - \angle EOO_3 \quad (A21)\]
\[(OO_3)^2 = (OE)^2 + (O_3E)^2 - 2(OE)(O_3E)\cos \angle O_3EO \quad (A22)\]
\[(O_3E)/\sin \angle EOO_3 = (OE)/\sin \angle OO_3E \quad (A23)\]

Triangle \((O_3OD)\):
\[(O_3D)^2 = (DO)^2 + (O_3O)^2 - 2(DO)(O_3O)\cos \angle DOO_3 \quad (A24)\]

Triangle \((EHI)\):
\[(EH) = 2(O_3H)\sin(\alpha_3/2) \quad (A25)\]
\[(IH) = 2(O_2I)\sin(\alpha_1/2) \quad (A26)\]
\[\angle EHI = \frac{\pi}{2} - \frac{\alpha_3}{2} - \frac{\alpha_1}{2} \quad (A27)\]
\[(EI)^2 = (EH)^2 + (IH)^2 - 2(EH)(IH)\cos \angle EHI \quad (A28)\]

Triangle \((EBF)\):
\[(BE) = 2(OB)\sin(\phi_0/2) \quad (A29)\]
\[\angle EBF = \eta + \frac{\phi_0}{2} \quad (A30)\]
\[(EF)^2 = (BF)^2 + (BE)^2 - 2(BF)(BE)\cos \angle EBF \quad (A31)\]

Triangle \((EFI)\):
\[(IF) = 2(DI)\sin(\theta/2) \quad (A32)\]
\[\angle EFI = \frac{\pi}{2} + \frac{\theta}{2} + \frac{\phi_0}{2} \quad (A33)\]
\[(EI)^2 = (EF)^2 + (IF)^2 - 2(EIF)(IF)\cos \angle EFI \quad (A34)\]

Angle relationship:
\[\angle DOO_3 = \phi_0 - \angle BOD - \angle EOO_3 \quad (A35)\]

Eqs. (A1)–(A3) and (A5)–(A35) are used to determine \(\alpha_3\), \(\alpha_1\), and \(r_3\) of the admissible slip-line field.

A.2. Determination of \(\alpha_2\), \(\alpha_5\), and \(r_5\)

The angle change from point H to point M along HH'M is equal to the angle change along HEM, i.e.,
\[\alpha_2 + \alpha_5 = \alpha_3 + (\phi - \phi_0 - \lambda) \quad (A36)\]
Triangle (OEM):
\[ (EM) = 2(\text{OE})\sin[(\phi - \phi_0 - \lambda)/2] \]  
\[ \angle OEM = \pi - \frac{\phi - \phi_0 - \lambda}{2} \]  
(A37)
(A38)

Triangle (O_3EH):
\[ \angle O_3EH = \frac{\pi - \alpha_3}{2} \]  
(A39)

Triangle (HEM):
\[ (HM)^2 = (EH)^2 + (EM)^2 - 2(EH)(EM)\cos \angle HEM \]  
(A40)

Triangle (HH'O_2):
\[ (H'H) = 2(O_2H)\sin(\alpha_2/2) \]  
(A41)

Triangle (O_5H'M):
\[ (H'M) = 2(O_5M)\sin(\alpha_5/2) \]  
(A42)

Triangle (HH'M):
\[ \angle HH'M = \frac{\pi - \alpha_2 - \alpha_5}{2} \]  
\[ (HM)^2 = (H'H)^2 + (H'M)^2 - 2(H'H)(H'M)\cos \angle HH'M \]  
(A43)
(A44)

Triangle (H'O_2O_5):
\[ (O_5O_2)^2 = (O_2H')^2 + (O_2H)^2 \]  
\[ \angle O_5O_2H' = \tan^{-1}[(O_5H')/(O_2H')] \]  
\[ \angle O_5O_2H = \angle O_5O_2H' - \alpha_2 \]  
(A45)
(A46)
(A47)

Triangle (O_5O_2O_2):
\[ (O_3O_5)^2 = (O_3O_2)^2 + (O_5O_2)^2 - 2(O_3O_2)(O_5O_2)\cos \angle O_3O_2O_5 \]  
(A48)

Triangle (O_5OM):
\[ (OM) = R \]  
\[ \angle O_5MO = \pi - \eta \]  
\[ \angle OO_5M = \pi - \angle O_5MO - \angle MOO_5 \]  
\[ (O_5O)^2 = (O_5M)^2 + (OM)^2 - 2(O_5M)(OM)\cos \angle O_5MO \]  
\[ (O_5M)/\sin \angle MOO_5 = (OM)/\sin \angle OO_5M \]  
(A49)
(A50)
(A51)
(A52)
(A53)

Triangle (O_3OO_5):
\[ (O_3O_5)^2 = (O_3O)^2 + (O_5O)^2 - 2(O_3O)(O_5O)\cos \angle O_3OO_5 \]  
(A54)
Angle relationships:
\[
\angle HEM = 2\pi - \angle O_3E_0 - \angle O_3EH - \angle OEM
\]  
\[
\angle O_3OO_5 = \angle O_3OE + \angle O_5OE
\]  
\[
\angle O_5OE = \angle EOM - \angle MOO_5
\]  
\[
\angle EOM = \phi - \phi_0 - \lambda
\]  
\[
\angle O_3O_2O_5 = \angle O_3O_2H - \angle O_5O_2H
\]

Eqs. (A3)–(A6), (A19)–(A23), (A25), (A36)–(A59) and \(\alpha_3, \alpha_1, \) and \(r_3\) obtained in Sect. A.1 are used to determine \(\alpha_2, \alpha_5, \) and \(r_5\) of the admissible slip-line field.

Fig. A.1 Slip-line geometry used to determine angles \(\alpha_2, \alpha_3, \) and \(\alpha_5\) in Eqs. (4.12) and (4.13) (Komvopoulos, 2011).