Ch. 10: Fundamental of contact between solids

Actual surface is not smooth. At atomic scale, there are always defects at surface, such as vacancies, ledges, kinks, terraces.

In micro or macro scale, roughness always exists. Even at extreme scale such as surface of a planet, the roughness has the same features at macro or micro scale. What are the observable features of surface roughness?

Surface roughness has multi-scale feature. A rough surface contains a smaller scale of roughness at high magnification. Ex. Surface of the earth is rough, as well as that of a mountain, a hill, a rock, etc.
A normal distribution (Gaussian) will show a straight line of cumulative graph using log scale.

\[ \int_{0}^{L} Z \, dx \]

Roughness can be reported as \( R_a \), \( R_q \) or others. They are average value of peak height distribution. Different roughness values use different way to average the peak height.

<table>
<thead>
<tr>
<th>Table 10.1 Commonly used height parameters.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roughness average (CLA or ( R_a ))</td>
</tr>
<tr>
<td>( R_a = \frac{1}{L} \int_{0}^{L}</td>
</tr>
<tr>
<td>Root mean square roughness (RMS or ( R_q ))</td>
</tr>
<tr>
<td>( R_q = \frac{1}{L} \int_{0}^{L} Z^2 , dx )</td>
</tr>
</tbody>
</table>

Sensitivity of different kinds of roughness value are different. Each kind of roughness value has different sensitivity to different surface feature. In the example above, \( R_a \) is not so sensitive to the peak height, but rather to the volume of the material above datum. But \( R_q \) shows large difference of values for the two kinds of surface.
Spatial distribution of peaks can be characterized by graph of $R(\tau)$ and $\tau$.

Or can be seen from the value of $\beta^*$, which show the rate of change of slope of the normalized $R(\tau)$ graph.

| Autocovariance function (ACVF or $R(\tau)$) | $R(\tau) = \lim_{L \to \infty} \frac{1}{L} \int_0^L z(x)z(x+\tau)dx$ | $R(\tau)$ |
| Power spectral density function (PSDF or $G(\omega)$) | $G(\omega) = \frac{1}{2\pi} \int_0^\infty R(\tau)\cos(\omega\tau)d\tau$ | $G(\omega)$ |

Table 10.2: Statistical functions used to describe spatial characteristics of the real surfaces (adapted from [9]).

where:
- $\tau$ is the spatial distance [m];
- $\beta^*$ is the decay constant of the exponential autocorrelation function [m];
- $\omega$ is the radial frequency $[\text{m}^{-1}]$, i.e., $\omega = 2\pi/\lambda$, where $\lambda$ is the wavelength [m].

Generally, a very smooth surface is not desirable, as well as the very smooth surface. Why so?
This also depends on the severity of contact load.
Real contact area is the summation of contact area of each asperities.

Nominal contact area is the macro scale contact area. This includes real contact area, and any void and non-contacting regions between the rough surfaces.

At asperities, actual stress at contact point is much larger than the average stress (nominal contact pressure) calculated by load divided by nominal contact area.

There are several models for calculation of real contact area, for example, Onions and Archard model:

\[ A_r = n \pi A (2.3 \beta^2) \int_0^{\infty} \frac{F(z^c, C)}{NC} dz^c \]

where:
- \( A_r \) is the true area of contact [m²];
- \( n \) is the number of asperities per unit area of apparent contact;
- \( A \) is the apparent contact area [m²];
- \( \beta^2 \) is the correlation distance obtained from the exponential autocorrelation function of a surface profile [m];
- \( z^c \) is the normalized ordinate, i.e. \( z^c = z/\sigma \) (height/RMS surface roughness);
- \( N \) is the ratio of peaks to ordinates. In this model \( N = 1/3 \) [26];
- \( d \) is the normalized separation between the datum planes of either surface, i.e. \( d = h/\sigma \);
- \( C \) is the dimensionless asperity curvature defined as:
  \[ C = Fr/\sigma \]

where:
- \( b \) is the mean plane separation [m];
- \( \sigma \) is the RMS surface roughness [m];
- \( t \) is the sampling interval. In this model \( t = 2.3 \sigma \) [m];
- \( f^* \) is the probability density function of peak heights and curvatures;
- \( r \) is the mean asperity radius [m] defined as:
  \[ r = \frac{2^{5/2}(2.3 \sigma)^3}{9 \sigma} \]
The expression for a total load is given in a form [26]:

$$ W = \frac{4}{3} \pi E' (2.3 \beta^3) \int \left( \eta - d \right) \int \frac{P(z,\eta,\epsilon)}{\sqrt{C}} dC dz $$

where:
- \( W \) is the total load [N];
- \( E' \) is the composite Young’s modulus (eq. 7.35) [Pa].

The ratio of load to real contact area, i.e., the mean contact pressure, is

$$ P_{mean} = \frac{W}{A_r} = \frac{4 \pi E'}{3 \pi n (2.3 \beta^3)} \int \left( \eta - d \right) \int \frac{P(z,\eta,\epsilon)}{\sqrt{C}} dC dz $$

It can clearly be seen from equation (10.5) that the ratio of load to true contact area depends only on the material properties defined by the Young’s modulus and the asperity geometry.

Mean pressure does not depend on the apparent contact area. For a rough surface, there is a specific proportionality between load and real contact area.

In G-W and W-A models for \( \psi \) and \( \psi^* \) < 0.6 elastic deformation dominates and if \( \psi \) and \( \psi^* > 1 \) a large portion of contact will involve plastic deformation. Plastic deformation causes the surface topography to sustain considerable permanent change, i.e., flattening of asperities. Protective films may also fracture and allow severe wear to occur. When \( \psi^* \) or \( \psi''^* \) is in the range of 0.6 - 1 the mode of deformation is in doubt. In B-J model for values of \( \psi^* < 1 \) the wear rate is negligible, and as \( \psi' \) increases from 1.0 to 3.5 the wear coefficients increase by several orders of the magnitude [74].

Larger degree of plastic deformation, wear of contacting surface is likely. This is because plastic deformation causes permanent damage to the asperities.

Effect of sliding on contact between solid surface.

As load increases, plastic deformation becomes the dominating kind of deformation. Real contact area value approached apparent contact area value, as deep grooves and depression become intact. Level of plastic deformation of contacting surface can be estimated by three models.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( \psi = \left( \frac{E}{H} \right)^\psi )</td>
<td>( \psi^* = \left( \frac{E}{H} \right)^\psi )</td>
<td>( \psi_i = \left( \frac{E}{H} \right)^\psi )</td>
</tr>
</tbody>
</table>

where:
- \( E \) is the composite Young’s modulus (eq. 7.35) [Pa];
- \( H \) is the hardness of the deforming surface [Pa];
- \( \sigma_r \) is the standard deviation of the surface peak height distribution (nm);
- \( r \) is the asperity radius, constant in the G-W model [nm];
- \( \sigma \) is the RMS surface roughness, ‘\( \sigma \)’ refers to the harder surface in the B-J model [nm];
- \( \rho \) is the correlation distance [nm];
- \( k \) is the asperity lip curvature of the harder surface [nm];
- \( P_c \) is the zero pressure of the softer surface [Pa]. Note that \( \psi' \) is a function of friction coefficient, i.e., it decreases with the increase of \( \psi' \) [74];
- \( \psi_i \) is the plasticity index for repeated sliding [73,74,76].

During sliding due to an applied tangential force, asperities dig deeper as the surface which resists the force is reduced to one side only. Ramping effect will eventually lift the asperities up, slight increases separation between the two surfaces.

FIGURE 10.15 Schematic illustration of the transition from static contact to sliding contact for a hard asperity on a soft surface.
The high asperities have larger effect to wear of surface, compared to the lower ones. Some low asperities loose contact with the surface during lifting off.

Wear debris generated by plastic deformation can be trapped in grooves or depression between asperities. This in effect reduced the actual contact real of the two surfaces. Lumps of wear debris eventually replace asperities as the site of true contact.

**Friction and Wear**

Friction is the dissipation of energy between two sliding bodies. Four basic empirical laws:

- there is a proportionality between the maximum tangential force before sliding and the normal force when a static body is subjected to increasing tangential load;
- the tangential friction force is proportional to the normal force in sliding;
- friction force is independent of the apparent contact area;
- friction force is independent of the sliding speed.

Are they true?

When the tangential force is less than the friction force (static), there is an elastic displacement, in submicron scale. This is due to elastic deformation of asperities.

**Friction and Wear**

Friction is the dissipation of energy between two sliding bodies. Four basic empirical laws:

- there is a proportionality between the maximum tangential force before sliding and the normal force when a static body is subjected to increasing tangential load;
- the tangential friction force is proportional to the normal force in sliding;
- friction force is independent of the apparent contact area;
- friction force is independent of the sliding speed.

Are they true?

Where do the static friction coefficient and kinetic friction coefficient apply?
Effect of rate of friction force application:
(a) Low rate of 20 N/s  (b) High rate of 20000 N/s

• Low rate agrees well with classical understanding, there is a static and kinetic friction coefficient. The later has a smaller value. Abrupt sliding occurs.
• High rates show a continuous change of friction force (as well as friction coefficient) and displacement before gross sliding occurs.

Stick-Slip phenomena

An oscillation between the static and kinetic levels of friction can also occur and this is known as 'stick-slip'. Stick-slip is a phenomenon where the instantaneous sliding speed of an object does not remain close to the average sliding speed. Instead, the sliding speed continuously varies between almost stationary periods and moments of very high speed. Stick-slip depends on the variation in the friction coefficient at low sliding speeds and on the vibrational characteristics of the system. In many cases the suppression of stick-slip can be an important as reducing the overall coefficient of friction because of the destructive nature of the vibration caused.

Moderate rate of friction force application: rate of displacement of system can follow. No vibration of the system.
Rapid rate of friction force application: only part of the body adjacent to the contact can follow, resonance can occur, causing severe 'stick-slip'. This is particularly true for a low stiffness materials due to the ease of elastic deformation of the body adjacent to the contact area.

Structural difference between Static and Sliding Contacts

Static Contact: Random distribution of contact points.
Sliding Contact: Lesser number of larger contact areas. Sliding movement is mostly affected by the large asperities (lumps) or large trapped wear debris. This may be due to lift-off effect at asperities in contact during sliding.

Loss of wear debris is an irreversible process. Once it leaves the contact surface, it cannot reenter. In addition, a larger stress exists at center of contact area, compared to at edges. It is likely for wear debris to move toward edges of contact surface, and finally leave the contact area, rather than moving inside the contact area indefinitely.
Other contact phenomena in Rolling

**Corrugation**: formation of wave-shaped profile on the rolled surface by repeated rolling contact.

**Traction**: Ability of a roller to sustain a tangential contact force while continuing to roll, with negligible resistance to motion.

Coefficient of rolling friction: The force required to maintain steady rolling / load. A smooth hard metal roller (rail wheel) has low value of 0.01-0.001. A soft material roller (rubber tire) adheres to underlying surface, will generate a high level of resistance force.

Coefficient of adhesion: Maximum tangential force (traction force) that can be sustained at the rolling contact, divided by load. Typical values are 0.1-1.0.

Micro-slip

Wheel cannot generate traction without micro-slip. A wheel always slips during rolling! Micro-slip occurs due to the restricted movement of roll surface by the flattened contact area.

The larger $\Delta r$, the larger magnitude of slip.

Frictional Heat

Large amount of energy is dissipated in form of heat. Heat generation is concentrated at asperities which make contact, causing large temperature rise. The larger sliding speed, the larger heating power (heat generation/unit time).
Additional effect is expansion asperities in contact, forming lumps (thermal mounds) which cause additional contact friction, and create more heating effect. This is especially true for a dry sliding of material with low thermal conductivity.

Characterization of Wear

Wear Volume:

\[ V = K A_i l = K \frac{W}{H} \]

where:
- \( V \) is the wear volume [m³];
- \( K \) is the proportionality constant;
- \( A_i \) is the real area of the contact [m²];
- \( W \) is the load [N];
- \( H \) is the hardness of the softer surface [Pa];
- \( l \) is the sliding distance [m].

The ‘K’ coefficient, also known as the ‘Archard coefficient’, is widely used as an index of wear severity. The coefficient can also be imagined as the proportion of asperity contacts resulting in wear. The value of ‘K’ is never supposed to exceed unity and in practice, ‘K’ has a value of 0.001 or less for all but the most severe forms of wear. The low value of ‘K’ indicates that wear is caused by only a very small proportion of asperity contacts. In almost all cases, asperities slide over each other with little difficulty and only a minute proportion of asperity contacts result in the formation of wear particles.

Specific wear rate (or wear coefficient!)

\[ k = V / (W x L) \]

where
- \( k \) is specific wear rate (m³/N-m)
- \( V \) is wear volume (m³)
- \( W \) is load (N)
- \( L \) is sliding distance (m)