

# Chapter 13

## Contact and Friction of One- and Two-Dimensional Nanostructures

Yin Zhang and Ya-pu Zhao

**Abstract** Because their thickness dimension is very small compared with other dimensions, the one-dimensional (1D) nanostructures (such as nanowire, nanotube, and nanobelt) and two-dimensional (2D) nanostructures (such as graphene) are highly prone to bend. Because of their large bending flexurality, the 1D and 2D nanostructures exhibit different contact behavior from those chunky ones. Without considering the flexurality effect, the analysis on the experimental data of 1D and 2D nanostructures can lead to different and even contradicting results/conclusions on their mechanical properties. One focus of this chapter is on what can go wrong in the indentation and three-point bending tests of 1D nanostructures if the flexurality effect is not accounted. At the same time, the 1D and 2D nanostructures also exhibit abnormal friction behavior. The assumptions of the classical contact are reviewed, and their possible deficiencies and difficulties of being used to analyze the contact and friction of 1D/2D nanostructures are also discussed.

### 13.1 Introduction

The so-called one-dimensional (1D) nanostructures [1] or wirelike nanoentities [2], such as nanowire (NW) [3–10], nanotube (NT) [11–15], and nanobelt (NB) [16–22], have attracted much interest in scientific community because of their remarkable mechanical, electrical, and thermal properties and potential applications in wide variety of devices. The mechanical failure of 1D nanostructures can lead to the malfunction or even failure of an entire device, and 1D nanostructures may also have size-dependent properties. Therefore, an accurate measurement of their mechanical properties is of critical importance when integrating them into nanodevices. Unfortunately, the experimental measurements of 1D nanostructures mechanical properties are often different and even contradicting to one another [1, 9]. For example, in the Young's modulus measurement of Ag NWs using the three-point bending test, the following different and contradicting experimental observations were reported: (1) Cuenot et al. [3] observed that the Young's moduli of

Ag NWs increase monotonously (76–140 GPa) with the decrease in their diameters (diameter range of 30–250 nm), and they are all larger than the bulk modulus of 76 GPa; (2) Wu et al. [4] observed that the Young's moduli of Ag NWs (diameter range of 20–35 nm) scatter randomly in the range of 60–140 GPa; and (3) Chen et al. [5] observed that Young's moduli of Ag NWs (diameter range of 65–140 nm) have no size effect at all; they are all around the bulk value. The same scenario occurs in the Young's modulus measurements of ZnO NB: (1) Ni and Li [18] measured the Young's modulus of ZnO NBs with the  $2b/t$  range of 2–10 ( $2b$ ,  $t$ : NB width and thickness, respectively) as  $31.1 \pm 1.3$  GPa (nanoindentation test) and  $38.2 \pm 1.8$  GPa (three-point bending test); (2) Bai et al. [19] measured the ZnO NB Young's modulus using mechanical resonance test as 52 GPa ( $2b/t = 1.1$ – $1.7$ ); (3) Mai and Wang [20] measured the ZnO NB Young's modulus using the three-point bending test as  $118 \pm 14$  GPa ( $2b/t = 1.09$ ),  $105 \pm 10$  GPa ( $2b/t = 1.21$ ), and  $162 \pm 12$  GPa ( $2b/t = 1.29$ ), respectively; (4) Lucas et al. [21] measured the ZnO NB Young's modulus using nanoindentation test as  $62 \pm 5$  GPa ( $2b/t = 1.2$ ),  $38 \pm 5$  GPa ( $2b/t = 1.6$ ), and  $17 \pm 5$  GPa ( $2b/t = 3.3$ ), respectively; and (5) another nanoindentation test by Lucas et al. [16] showed the ZnO NB Young's modulus decreasing from 100 to 10 GPa with increase of  $2b/t$  from 1.2 to 10.3. Besides the quantity difference, it is also noticed that there is no (obvious) dependence of ZnO Young's modulus on  $2b/t$  in references [18, 19], but strong  $2b/t$  dependence in references [16, 20, 21]. We are fully aware that the 1D nanostructures from different groups are fabricated/synthesized and processed differently, which can cause different experimental observations. Physically, surface tension [3], microstructure [4], surface layer [8], etc., can all be the mechanisms responsible for the size-dependent properties of 1D nanostructures. The core-shell model is proposed to explain the increase in the ZnO NW Young's modulus with the decrease in diameter [8]. The shell is formed by the surface layer with larger Young's modulus and the proportion of shell increases with the decrease in NW diameter [8]. For this core-shell model to have noticeable surface effects on the 1D ZnO nanostructures, the surface-to-volume ratio is required to be  $0.08 \text{ nm}^{-1}$  or larger [16]. However, the surface-to-volume ratio of the NBs showing strong  $2b/t$  dependence of Young's modulus is more than one order of magnitude smaller than the required value [16]. Furthermore, Lucas et al. [16, 21] showed that the Young's moduli of ZnO NBs have no (clear) correlation with the surface-to-volume ratio, and the width-to-thickness ratio of  $2b/t$  is the only key parameter. Unlike the surface-to-volume ratio (larger value indicates smaller physical dimensions), the width-to-thickness ratio is purely a geometric characterization of a NB. Although a growth-direction-dependent aspect ratio, stacking faults in NBs growing along particular directions [16] and point defects [21] are proposed as the (possible) mechanisms, a clear physical picture on the ZnO NB Young's modulus dependence on the width-to-thickness ratio is still unavailable [16]. Because of their extremely small dimensions, the manipulation of 1D nanostructures is of great difficulty and imposes tremendous challenges to many existing testing techniques, which can also be responsible for the experimental error. However, the variance of the measured data (e.g., those given in reference [3]) is beyond the range of the instrument

system error. As there are very few models for extracting the mechanical properties of 1D/2D nanostructures [7], the models used in those experiments can also be responsible for the different and even contradicting experimental observation. For 1D and 2D nanostructures, their thickness/diameter is much smaller than other dimensions, which makes them highly flexural. The flexurality of 1D and 2D nanostructures is not incorporated in many models used to extract their mechanical properties from experiments. For example, in the three-point bending test of NW, the clamped–clamped boundary conditions are assumed because of the NW small size (thus large surface-to-volume ratio) and strong adhesion effect [3, 4]. Chen et al. [5] found in their experiments that the NW boundary conditions may not be a clamped–clamped one, depending on the external load and the NW diameter. Zhang and Zhao [23] presented an adhesive contact model for the NW three-point bending test and found that in general the NW boundary conditions are neither clamped–clamped nor hinged–hinged, but an intermediate one due to the lift-off caused by the flexurality. The clamped–clamped and hinged–hinged boundary conditions will cause four times difference in the Young’s modulus in the three-point bending test [23]; the intermediate boundary conditions can offer some insight into the experimental observations that the Young’s modulus of Ag NW is 2–3 times larger than its bulk one [3]. In a nanoindentation test, the Oliver–Pharr method [24] and Sneddon method [25] are used to extract the 1D nanostructures’ hardness [6, 18, 22] and Young’s modulus [6, 14]. However, the standard Oliver–Pharr method [24] assumes the indented sample as a monolithic, semi-infinite elastic half-space [26], so does the Sneddon method [25]. Recently, lift-off induced by the structural flexurality has been shown to have significant impact on the interpretation of the experimental data obtained in the NW [27] and NB [28] indentation tests. Because of its small size of thickness/diameter, the substrate effect may also influence the test results [28]. When the indentation depth is more than 10% of the 1D/2D nanostructure thickness/diameter, the elastic field under indenter is not confined to the nanostructure itself; it extends into substrate [29, 30], and therefore, the substrate stiffness has impact on the measured contact stiffness.

The 1D and 2D nanostructures also exhibit abnormal friction behavior because of the structural flexurality. For example, Bhushan et al. [31] found that the intertube shear strength of carbon NT is much higher than its interlayer shear strength; Lee et al. [32] found that friction monotonically increases as the number of the layers of four atomically thin 2D nanostructures decreases. Bhushan et al. [31] have proposed that the current contact theory (such as the JKR model) may not accurately calculate the NT contact area [31]. With the decrease of the layer number, the 2D film bending stiffness also decreases, which makes the film easier to bend and forms “puckering” [32]. Lee et al. [32] compared the friction tests of atomically thin sheets on substrate and in suspension and concluded that the increased contact area due to the flexurality/bending of 2D nanostructures is the (only) reason responsible for the increase of friction when the layer number decreases. The possible deficiencies and difficulties of using the classical contact models to analyze the contact and friction of 1D/2D nanostructures are reviewed.

This chapter is organized into three parts: (1) review of the classical contact theories such as Hertz, Bradley, Johnson–Kendall–Roberts (JKR), Derjaguin–Muller–Toporov (DMT), and Maugis–Dugdale (MD) models which are for the sphere contact where the structural flexurality is not considered; (2) the flexural contact and its application in the three-point bending test and nanoindentation test; and (3) friction of the flexural 1D and 2D structures.

## 13.2 Classical Contact Theories

The transition and relation between those classical contact models have been given by Maugis [33], Johnson and Greenwood [34]. The parameter they use to show the transition between the classical contact models is the elasticity parameter [33, 34] and the external compressive load [34]. Here, we reveal the transition and relation from another different angle: the surface interaction outside contact area and the pressure profile in the contact area [35, 36]. Here, only a brief review is given, the detailed analysis and other issues such as curved surface effect and instability jump are presented in references [35, 36].

### 13.2.1 Surface Interaction Force and Derjaguin Approximation

If  $\sigma(h)$  is the surface interaction force per unit area, the total surface force outside the contact area,  $P_{\text{vdw}}$ , can be summed up as follows [36]:

$$P_{\text{vdw}} = \int \sigma(h) dA = \int_a^\infty \sigma(h) d(\pi r^2) = 2\pi R \int_z^\infty \sigma(h) dh, \quad (13.1)$$

where  $a$  is the contact radius and  $h$  is the separation distance. In the above equation, the parabolic approximation is used for the spheres surface profiles, i.e.,  $h = r^2/2R$ . For the Lennard-Jones (LJ) force law,  $\sigma(h)$  is given as follows [36, 37]:

$$\sigma(h) = \frac{16\gamma}{3z_0} \left[ \left( \frac{z_0}{h} \right)^3 - \left( \frac{z_0}{h} \right)^9 \right]. \quad (13.2)$$

Here,  $2\gamma$  is the work of adhesion, and  $z_0$  is the equilibrium separation of two half-spaces in the LJ surface force law [36, 38]. The positive value of  $\sigma$  indicates attraction and the negative repulsion. The 3–9 force law of  $\sigma(h)$  in Eq. 13.2 is the surface force per unit area of two flat half-spaces separated by a distance  $h$ ,

which is derived from the LJ 6–12 potential law for two isolated molecules [36,38]. Substituting Eq. 13.2 into Eq. 13.1, we have

$$P_{\text{vdw}} = 4\pi R\gamma \left[ \frac{4}{3} \left( \frac{z_0}{z} \right)^2 - \frac{1}{3} \left( \frac{z_0}{z} \right)^8 \right]. \quad (13.3)$$

During the integration of Eq. 13.1,  $z$  is treated as a constant [4]. Therefore, Eq. 13.1 physically gives the interaction forces of two flat surfaces. The Derjaguin approximation states that the interaction energy/force between small areas of (slightly) curved solids can be (accurately) approximated by the interaction energy/force of two flat surfaces [37,39]. In many self-consistent computations, Eq. 13.2 is used as the local pressure of a curved contact interface. So, in that sense, besides Eq. 13.3 which expresses the total surface force, Eq. 13.2 is also referred to as the Derjaguin approximation [36]. From the point of view of elasticity,  $\sigma(h)$  is a normal surface traction [38]. Instead of using the Derjaguin approximation of Eq. 13.2, Argento et al. [40] derived the surface traction by directly summing up the body forces of molecules as given by the LJ 6–12 potential through a double volume integration. Argento et al. [40] also showed that when (at least) one contacting body is a half-space, the Derjaguin approximation which ignores the surface profile and is usually regarded as an approximation, in fact, gives the exact total surface force. The maximum attractive surface interaction force obtained from Eq. 13.3 is  $P_{\text{vdw}} = 4\pi R\gamma$  at  $z = z_0$ . The tensile external load required to balance the attractive surface interaction force is thus  $P = -P_{\text{vdw}}$ , and the maximum tensile external load is  $-4\pi R\gamma$ , which is the DMT pull-off force. Here, it needs to be emphasized that  $P_{\text{vdw}}$  is the surface force outside the contact area and the equilibrium equation of  $P = -P_{\text{vdw}}$  can only be valid for the following two cases: (1) the Bradley model (there is no contact region for two rigid spheres, i.e.,  $a = 0$ ;  $z$  is “the nearest distance between the centers of surface molecules of the spheres” [41], and therefore,  $P_{\text{vdw}}$  of Eq. 13.3 is the only force for an external load to balance); (2) for the DMT model at the point of separation (for the DMT model, the spheres separate at  $a = 0$ , and therefore, there is no force contribution from the contact region at that critical point), the equation of  $P = -P_{\text{vdw}}$  does not deal with any elastic deformation, which is also often referred to as the Bradley model [37].

### 13.2.2 Contact Pressure and Equilibrium

The contact pressure profile in the JKR model is assumed to have the following form [42]:

$$p(r) = p_0 (1 - r^2/a^2)^{1/2} + p'_0 (1 - r^2/a^2)^{-1/2}. \quad (13.4)$$

The first term in Eq. 13.4 is the compressive Hertz pressure; the second term is a tensile one, which is induced by adhesion and often referred to as the Boussinesq pressure [36, 42].  $p_o$  and  $p'_o$  are given as follows [35, 36, 42]:

$$p_o = \frac{2Ea}{\pi R}, \quad p'_o = -\sqrt{\frac{4E\gamma}{\pi a}}. \quad (13.5)$$

During the derivation of  $p'_o$ , the contact area is assumed to be flat, which can be proved by an elasticity approach [35, 36]. However, Attard and Parker's self-consistent computation shows that the contact area is not exactly flat [43]. In general, when the two contacting bodies have similar Young's moduli, the flat contact area is a good approximation [36], and the effects of the curved contact area are discussed in reference [36].

The following amount of external load  $P$  is required to balance the force due to the surface interaction force outside the contact area and elastic deformation:

$$P + P_{vdw} = \int_0^a 2\pi r p(r) dr = \left( \frac{2}{3} p_o + 2p'_o \right) \pi a^2. \quad (13.6)$$

Equation 13.6 offers a framework to show how different classical hard contact models are related:

*Bradley Model:* Because there is no elastic deformation for rigid spheres,  $p_o = p'_o = 0$ , Eq. 13.6 recovers the Bradley model.

*Hertz Model:* Because it involves a nonadhesive contact,  $P_{vdw}$  of the surface interaction force outside the contact region and  $p'_o$  induced by the adhesion are both zero. When  $P_{vdw} = p'_o = 0$ , Eq. 13.6 recovers the Hertz model with  $p_o = 2Ea/\pi R$ .

*DMT Model:*  $p'_o = 0$  because in the DMT model, there is no tensile pressure inside the contact region. Because of its short "neck," the surface interaction forces outside the contact region are accounted for. So when  $p'_o = 0$  and  $P_{vdw}$  is taken with the maximum value of  $P_{vdw} = 4\pi R\gamma$ , Eq. 13.6 recovers the DMT model with  $p_o = 2Ea/\pi R$ .

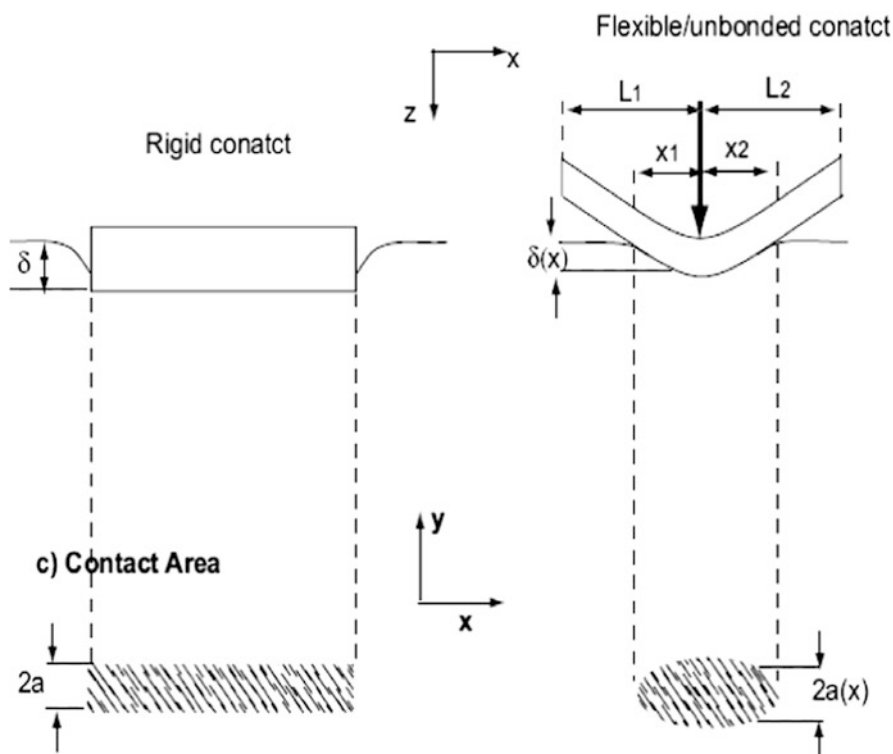
*JKR Model:* Because a flat contact area is formed,  $p'_o = -\sqrt{4E\gamma/\pi a}$ . Also because of its large "neck" height, the surface interaction force outside contact region can be ignored. So when  $p'_o = -\sqrt{4E\gamma/\pi a}$  and  $P_{vdw} = 0$ , Eq. 13.6 recovers the JKR model with  $p_o = 2Ea/\pi R$ .

The MD model is expressed by two coupled equations [33], which cannot be recovered by Eq. 13.6. The MD pull-off force is between the JKR one and DMT one [33, 36]. The contact models are usually presented by the  $P - a$  curves, which can be found by substituting Eq. 13.5 into Eq. 13.6. The detailed discussion on the transition and relation of the  $P - a$  curves of different contact models is presented in [35, 36], which is essentially the same as the above discussion on  $p_o$ ,  $p'_o$ , and  $P_{vdw}$ .

### 13.3 Flexural Contact

The above contact models are for spheres, and for the above models to hold the dimensions of the contact must be small compared (a) with the dimensions of each contacting bodies and (b) with the relative radii of curvature of the surfaces [42], which is to say that deformation only occurs around the small contact area. Similar approach is also used to derive the contact model for cylinder [42, 44]. In those models [42, 44], the elastic deformation of a cylinder is also around the contact interface, and the cylinder does not bend; therefore, one cross-section deformation can stand for the whole cylinder. Here, we call this kind of contact “rigid contact” because the cylinder flexural rigidity is so large that the cylinder does not bend during contact, as shown in Fig. 13.1. In general, because of the small thickness/diameter dimension, the 1D/2D nanostructures bend during loading and some parts of nanostructures separate/lift-off from the supporting substrate. As seen in Fig. 13.2, the pileup due to plastic deformation only forms around the loading site and the tilting angle as large as  $50^\circ$  is also observed [12]. Without considering this bending effect, the interpretation on the experimental data of the 1D nanostructures tests can be different and even contradicting [23, 35, 36].

Before we further discuss the 1D nanostructures tests, it is helpful for us to have a brief review and comment on the previous contact models which have obfuscating names. The contact problem of flexural structure is called differently as the receding contact [45], unbonded contact [46–48], tensionless contact [49–52], and unilateral contact [53]. Despite their different names, they all emphasize one essential thing: bending and lift-off of flexural structure. The name of receding contact [45] emphasizes that the contact area under loading is smaller than the unloaded one because of lift-off. The name of unbonded contact [46–48] emphasizes that the flexural structure is allowed to lift-off/separate from the contacting medium. The name of tensionless contact [49–52] emphasizes that in nonadhesive contact, tensile stress cannot be transmitted to the lifting-off parts of structures. The names of tensionless contact and unilateral contact [53] both emphasize that only compressive stress exists in contact area. All these contact mechanics models of flexural structure [45–53] reach the same conclusion that the contact length is independent on the load magnitude, which is responsible for the localized effect. Also in all of the above models, the adhesion effect is not considered. Because of the large surface-to-volume ratio of 1D nanostructures, adhesion can be important in some scenarios. As indicated in Eqs. 13.4 and 13.5, adhesion induces the tensile Boussinesq pressure, which, from the viewpoint of force balance, will lead to a larger contact area [27, 28, 35, 36]. Generally speaking, adhesion energy is rather weak compared with elastic energy [28]; only when the dimensions of the contacting bodies and elastic deformation are small enough can adhesion have significant influence. The difference between the adhesive contact model (such as JKR) and nonadhesive contact (such as Hertz) only stand out when the external load is a tensile one or a compressive one with small magnitude [28, 35, 36]. Usually, in indentation test, the compressive external load is very large and the adhesion effect can be

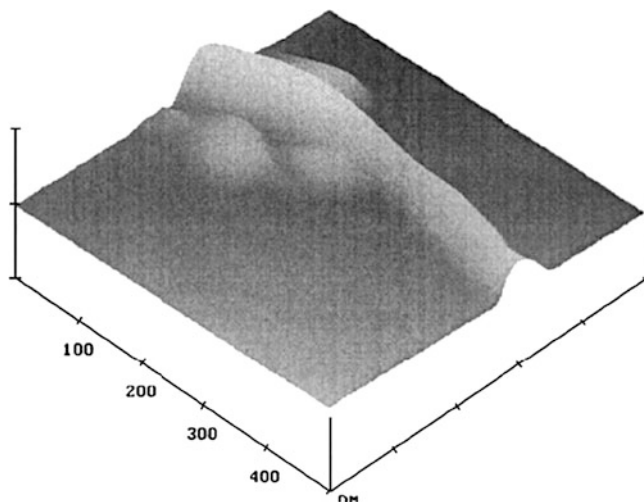


**Fig. 13.1** For the rigid contact scenario, the whole cylinder sinks into the elastic medium with a constant indentation depth. For the flexible contact scenario, the cylinder lifts-off and the indentation depth varies at different locations. Also for the rigid contact, the contact area is a rectangle; for the flexible contact scenario, the contact area is an ellipse-like zone (After [27])

ignored. However, with the proper choice of loading, size, and adhesion parameters, the adhesion can significantly reduce or even suppress the lift-off of a NW [27].

There are two types of problem formulations on this flexural contact problem: integral and differential formulations. When the elastic medium is modeled as an elastic half-space [46–48, 50], the formulation is always the integral one. The integral equation of the elastic half-space model, which indicates the displacement, strain, and stress at a point, is determined by the elastic deformation all over the area. In contrast to this, the elastic foundation model, for example, the Winkler foundation model, assumes the elastic medium consists of a series of independent springs; therefore, in an elastic foundation, the displacement, strain, and stress at a point are locally determined. The elastic foundation model leads to the differential formulation of this contact problem [49, 51, 52]. In the viewpoint of continuum approach, the elastic foundation model may lead to some physically unrealistic results, especially on the stress analysis. But Kerr [54] pointed out that the elastic foundation model is introduced to study the foundation surface response, not



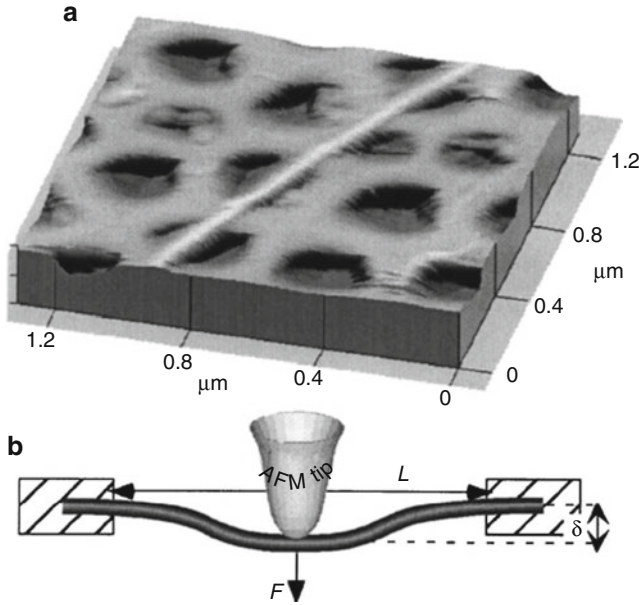


**Fig. 13.2** A restored NT with pileups on the surface of silicon wafer. As noticed, only the portion around the loading/indenter forms pileups and other parts keep intact, which is due to the lift-off of NT (After [12])

the stress caused inside the foundation. The mission of our modeling on NW and NB [27, 28] is stated as follows: finding the indentation depth (i.e., surface deformation of substrate) for a given indentation load. The elastic foundation model is introduced to study the foundation surface response, not the stress caused inside the foundation, the deficiency of elastic foundation model on the stress analysis in general should not cause serious problem. Elastic foundation model mathematically is much simpler than the elasticity approach of the elastic half-space model [54]. The integral formulation in general is much more complex and lengthy than the differential one. Furthermore, the relationship of line load displacement for the 2D cylinder plane-strain contact is indeterminate in an elastic half-space model [55] due to some uncertainties of its 2D elasticity features [56]. The Winkler elastic foundation model is derived for the cylinder-shaped NW [27] and the prism-shaped NB [28]. From the modeling aspect, the Winkler foundation only consists of the spring layer [54]. The Reissner foundation, which consists of the spring, shear, and bending layers [54], is a much better model to approximate the elastic half-space. The flexural contact of a beam with the Reissner foundation model can be found in references [49, 52].

### 13.3.1 Three-Point Bending Test

The 1D micro-/nanomaterials in test are first suspended over porous material, or a strip, or an etched hole, or over a trench as shown in Fig. 13.3 [23, 57]. The atomic force microscopy (AFM) tip then exerts a concentrated force on the suspended



**Fig. 13.3** (a) AFM image of NT adhered to the porous alumina ultrafiltration membrane. (b) Schematic of the measurement: the AFM applies a concentrated load to the NT which is modeled as an Euler–Bernoulli beam. Notice that the clamped–clamped boundary conditions are assumed (After [57])

section to form a typical three-point bending test [23, 57]. The AFM tip is often placed at the center of the suspended section, which is thus referred to as the midpoint test [23]. By assuming the Euler–Bernoulli beam model, the midpoint displacement has the following relationship with the testing material properties [23]:

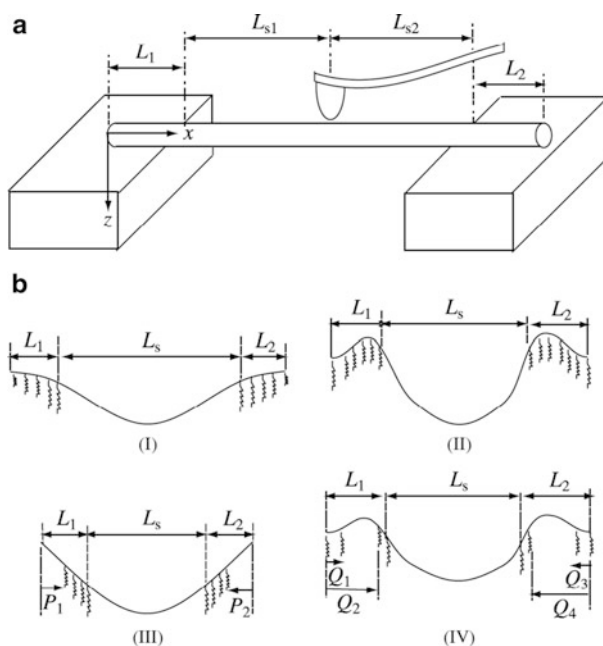
$$k = \frac{F}{z_{\text{mp}}} = K \frac{E_1 I}{L_s^3}, \quad (13.7)$$

where  $k$  is the effective spring stiffness of the beam and the AFM measures the data for the force–displacement ( $F - z_{\text{mp}}$ ) curves [5, 7–11]. The term  $F$  is the concentrated load exerted at the center of the beam;  $z_{\text{mp}}$  is the beam midpoint displacement which is also the maximum displacement for a clamped–clamped (C–C) beam and a hinged–hinged (H–H) beam with the load at its center.  $K$  is a constant depending on the testing materials boundary conditions;  $K = 192$  for the C–C boundary conditions, and  $K = 48$  for the H–H boundary conditions.  $L_s$  is the beam suspension length,  $E_1$  is the beam (bending) Young’s modulus,  $I$  is the beam cross-sectional moment of inertia, and  $I = \pi R^4/4$  for a solid circular beam ( $R$  is the radius). Therefore, by measuring the midpoint displacement, the concentrated load, the suspension length, and the radius and choosing the boundary conditions accordingly, the Young’s modulus of the testing material can be found from the

above equation. Keep in mind that the spring stiffness,  $k$ , of a C–C beam (with the same  $E_1$ ,  $R$ , and  $L_s$ ) is four times larger than that of a H–H beam because of the difference in the  $K$  values. It should be noticed that for the three-point bending test of a NW which is bonded to the support by adhesion, the C–C and H–H boundary conditions can both be the choice [5, 23].

Because of the small size of NW/NT (which results in the large surface-to-volume ratio), and because of the large adhesion effect [3], many researchers assume C–C boundary conditions for a NW/NB/NT suspended on a trench/pore [23]. From Eq. 13.7, the Young's modulus can be calculated as  $E_1 = kL_s^3/(KI)$  with  $k$  supplied by the  $F - z_{\text{mp}}$  data, measured by the AFM. With  $K$  fixed at 192 for a C–C beam,  $I(R)$ ,  $L_s$  can be measured with relatively high accuracy. By analyzing the experimental  $F - z_{\text{mp}}$  data [2, 5–7], Cuenot et al. [3] found that the Young's modulus of Ag NW increases significantly with a decrease of its diameter—usually 2–3 times that of the bulk value. Cuenot et al. [3] offered a surface stress theory to explain the increase in the Young's modulus with decreasing Ag NW diameter. According to their theory, the increase of the Ag NW Young's modulus is proportional to  $R^{-3}$ , which is also supported by their experimental observations [3]. However, the experiment by Wu et al. [4] on Ag NWs does not agree with Cuenot's experiments and theories [3]. Wu's experiment shows that most of the measured Ag NWs' Young's moduli are higher than the bulk one, but their Young's moduli are not sensitive to the change of NW radius at all, as shown in their Fig. 13.3 [4]; a novel fivefold twin microstructure mechanism is proposed to explain the increase of the Ag NW Young's modulus [4]. Unlike the midpoint test which measures the force–displacement data only at the suspension center [3, 4], Chen et al. conducted the three-point bending test by measuring force–displacement data for the whole profile of the suspended portions of the testing NWs [5]. The reason for conducting the multiple-point measurements rather than the single point measurement is that unlike the model used in those experiments [3, 4] which assumes the C–C boundary conditions for the testing NWs, Chen et al. [5] found that the boundary conditions of the suspended NW may change with the change of the diameter/load. This was verified by their experimental data. With multiple-point force–displacement data and curve fitting, the boundary conditions can be specified [5]. By taking into account that boundary conditions may change, Chen et al. [5] conclude that in their Ag NW test the Young's modulus of Ag NW was not significantly different from the bulk property. Now for the Young's modulus three-point bending test of the Ag NWs, three different and contradicting trends were observed: (1) Cuenot et al. [3] observed that the Young's moduli of Ag NWs (diameter range of 30–250 nm) increase monotonously with the decrease of the diameter, and they were all larger than the bulk property; (2) Wu et al. [4] observed that the Young's moduli of Ag NWs (diameter range of 13–35 nm) can be either larger or smaller than the bulk property, and Young's moduli were found not to be sensitive to the NWs diameters at all; (3) Chen et al. [5] concluded that there is no size effect for Ag NWs (diameter range of 65–140 nm).

Chen's findings [5, 58] in their experiments with both Ag and GaN NWs are summarized as follows [23]: (1) The NW boundary conditions were found to be



**Fig. 13.4** (a) Dimension of a NW under an AFM loading and its coordination system. The schematic diagram of a beam laid on a trench. The beam dimensions and the coordinate system are also shown. (b) Four typical boundary conditions of the beam under loading: (I) and (II) full contact, (III) and (IV) partial contact (After [23])

C–C for those with small diameters and H–H for those with larger diameters (in their experiments, the suspended length  $L_s$  is fixed); (2) the magnitude of the applied concentrated force influenced on the boundary conditions transition; (3) a NW can have asymmetric boundary conditions of the clamped–hinged (C–H) type; and (4) as shown in Fig. 3d and 3e of reference [5], the deflection of a NW with a relatively large diameter is between a C–C deflection curve and a H–H curve. Chen et al. also noticed intermediate boundary conditions [58]. All these experimental observations can be explained by the model presented in reference [23] as shown in Fig. 13.4. Chen et al. [58] believe that adhesion and its competition with the applied load is the key to understanding the boundary conditions transition, and the model proposed in reference [23] confirmed it. In the three-point bending test for NWs bonded with support by adhesion, the boundary conditions are a key issue. As indicated in Eq. 13.7, the boundary conditions, being either C–C or H–H, can contribute to a Young’s modulus evaluation of four times difference. In our model [23], we find that the NW boundary conditions in general are the intermediate ones, which is to say that the NW end is neither clamped nor hinged. The hinged end cannot take any bending moment and a rotation will occur; the clamped end can take a bending moment and there is no rotation. For the end with the intermediate boundary conditions, it can take some bending

moment, but there is also a rotation. The intermediate boundary conditions are common in many as-grown or as-deposited film/substrate structures [59]. The NW boundary conditions can be asymmetric, of the clamped–hinged (C–H) type [5, 58], with  $K \approx 107$  in Eq. 13.7 for the C–H beam in the midpoint test. The effective structure stiffness of a beam with the intermediate boundary conditions or with the asymmetric C–H boundary conditions is between that with H–H boundary conditions and that with C–C boundary conditions. The fact is that a NW with intermediate boundary conditions or with C–H type during three-point bending test may offer an insight into the experimental observations that the Ag NW Young’s modulus is 2–3 times larger than its bulk one [3].

Even in those experiments which assume the C–C boundary conditions, the researchers all show their concerns for the boundary conditions of NWs under test and realize the importance of boundary conditions on determining the NW Young’s modulus [3, 57]. Sliding is one of the mechanisms which may cause the change of boundary conditions [5, 58]. Although the AFM exerts a vertical force, the AFM tip–NW contact [23] and the midplane stretching during deflection [60] can both induce a horizontal tensile force which can cause the NW in test to slide. Here, the vertical direction is the  $z$ -axis direction, and horizontal direction is the  $x$ -axis direction as shown in Fig. 13.4a. However, the experimental observations show very little or no sliding [23]. Lift-off was proposed as another possible mechanism causing the boundary conditions to change [5, 23]. Unlike sliding which occurs in horizontal direction and can be observed relatively easily by SEM imaging [61], lift-off occurs in a vertical direction, which is experimentally difficult to be observed in the three-point bending test of NWs. Cuenot et al. [62] assumed that strong adhesion should prevent lift-off, and thus, the boundary conditions should be the C–C type. However, Chen et al. [58] suspected that under a relatively large applied force, the adhesion may not be sufficient enough to hold the NW ends clamped. Paulo et al. pointed out that “the nonideal anchoring using adhesion forces” will introduce “interfacial mechanical instabilities,” which means the nanobeam may not be “solidly connected to prefabricated microstructures” [63]. To achieve “a mechanically rigid anchor” (which is the clamped boundary condition), Paulo et al. [63] used the vapor–liquid–solid (VLS) method to grow the NWs from small catalyst particles deposited on a substrate. Similarly, in order to make sure the NWs two ends are clamped after they are dispersed on a trench, Zhu et al. [1] and Wu et al. [4] enhanced the NWs bonding with the trench by electron-beam-induced deposition (EBID) welding. However, the clamps formed by EBID may change the composition and structure of 1D nanostructure and affect measurement accuracy [17, 28]. The adhesion force is relatively weak, and its effect stands out only when the surface-to-volume ratio is large. When the NW diameter is small (the surface-to-volume ratio is large) and applied force is relatively small, the adhesion effect can be strong enough to hold the NW two ends clamped. However, when the NW diameter is large or the applied force is large, the adhesion may not be strong enough to hold the two ends clamped, which leads to the boundary condition transition from the C–C type to the H–H one, which has been in Chen’s experiments [5, 58] and Zhang and Zhao’s computation [23]. In reference [23], a dimensionless parameter is defined, which is

the key parameter used to show how the boundary conditions changes and given as follows:

$$\alpha = 4\sqrt[4]{2} \sqrt{\frac{\gamma}{\pi E^* R}} \left( \frac{E^*}{2E_1} \right)^{3/16}. \quad (13.8)$$

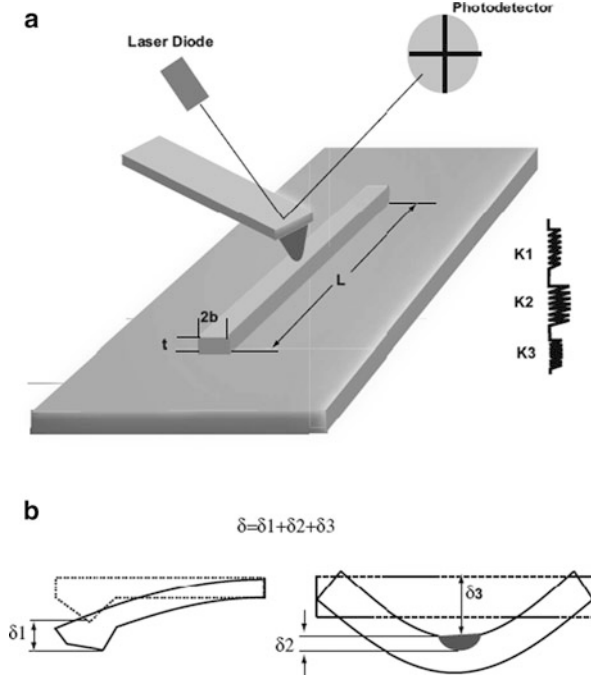
Here,  $E^*$  is the reduced modulus. The adhesion effect ( $\gamma$ ) and cylinder size effect ( $R$ ) are incorporated in the parameter  $\alpha$ .  $\alpha$ , in essence, indicates the (order of) adhesion contribution to the line load, as compared with that due to the Hertzian contact [23].  $\alpha$  together with the suspension length and external load determines how the boundary conditions are formed in the three-point bending test [23].

The model presented in reference [23] does not consider the friction effect in the horizontal direction. When the suspended NW is pushed down, its midplane is stretched, and a tension force is thus generated. This tension force is balanced by the NW–trench interfacial friction. When the tension reaches the critical value of static friction, the NW slides. However, the interfacial shear strength ( $\tau$ ) is relatively large for a NW, for example,  $\tau$  lies between 134 and 139 MPa for Ag NW on Au substrate [64], and the horizontal force required to cause the NW to slide is so large that the fracture of NW is observed in the experiment [65]. In general, the deflection of the suspended NW in a three-point bending test is small, and there is little or no sliding [61].

### 13.3.2 Indentation Test

The basic idea on modeling the flexural contact of NB and NW is the same, and the difference only lies in the different contact pressure due to their different geometries [27, 28]. The indentation of a flexural NW based on the adhesive cylinder contact model given by Chaudhury et al. [44] is given in reference [27]. Here, only the NB indentation is discussed.

Figure 13.5a shows an NB with Young's modulus  $E_B$ , Poisson ratio  $\nu_B$ , width  $2b$ , thickness  $t$ , and length  $L$  indented by an atomic force microscope (AFM). The supporting substrate is with Young's modulus  $E_S$  and Poisson ratio  $\nu_S$ . Here, a three-spring-in-series system is formed. Spring  $K_1$  is due to the bending stiffness of AFM cantilever; spring  $K_2$  is due to the contact between the AFM tip indenter and NB; spring  $K_3$  is due to the contact between NB and substrate. Because there are two contacting interfaces, the indenter/NB interface and NB/substrate interface, this kind of contact is also referred to as double contact [7]. In experiment, AFM measures the travel distance along the indentation direction [14], which is the total displacement of  $\delta$ . As shown in Fig. 13.5b,  $\delta$  consists of three parts: (1) AFM tip normal displacement due to cantilever bending ( $\delta_1$ ), (2) indenter indentation depth into NB ( $\delta_2$ ), and (3) NB indentation depth into substrate ( $\delta_3$ ). Because  $\delta = \delta_1 + \delta_2 + \delta_3$  and the three springs are in series (i.e.,  $P/K = P/K_1 + P/K_2 + P/K_3$ ), the system compliance is additive, which is  $1/K = 1/K_1 + 1/K_2 + 1/K_3$ .  $K$  is the effective stiffness of the system.  $K_1$ ,  $K_2$ , and  $K_3$  are given as follows [28]:



**Fig. 13.5** (a) AFM indenting a NB on a substrate: a three-spring-in-series system.  $K_1$  is due to the bending stiffness of AFM cantilever;  $K_2$  is due to the contact between indenter and NB;  $K_3$  is due to the contact between NB and substrate.  $L$ ,  $2b$ ,  $t$ ,  $E_B$ , and  $\nu_B$  are the length, width, thickness, Young's modulus, and Poisson's ratio of NB, respectively. (b) The total AFM normal displacement  $\delta = \delta_1 + \delta_2 + \delta_3$ .  $\delta_1$  is the AFM tip displacement due to the cantilever bending deflection;  $\delta_2$  is the indenter indentation depth into the NB; and  $\delta_3$  is the NB indentation depth into the substrate (After [28])

$$K_1 = \frac{E_c t_c^3 b_c}{2L_c^3}, \quad K_2 = \frac{4E^* a^2 / R - 3\sqrt{2\pi a \gamma_1}}{2a/R - \sqrt{\pi \gamma_1 / (2E^* a)}}, \quad K_3 = 2.01 b E_s, \quad (13.9)$$

where  $E_c$ ,  $t_c$ ,  $2b_c$ , and  $L_c$  are the Young's modulus, thickness, width, and length of AFM cantilever, respectively.  $a$  is the contact radius,  $R$  is the AFM tip radius, and  $\gamma_1$  is the work of adhesion for the indenter/NB contact. When  $\gamma_1 = 0$ , which is the nonadhesive Hertz contact,  $K_2$  recovers the previous result obtained by Lucas et al. [16] and Carpick et al. [66] as follows:

$$K_2 = 2E^* a = 2E^* \sqrt{R\delta_2}. \quad (13.10)$$

Clearly,  $K_2$  is a nonlinear function of  $\delta_2$ .  $E^*$  can be found from Eq. 13.10 as the following:

$$E^* = K_2 / (2\sqrt{R\delta_2}). \quad (13.11)$$



In standard indentation test,  $K_2$  is extracted from the initial unloading curve [24, 26] to exclude the plasticity influence. The derivation of  $K_2$  does not consider the influence of tangential friction force [28]. If the indenter and NB have different elastic constants, the interaction between normal and tangential forces will have impact on the force-indentation depth curve and make the mechanical property extraction of NB from the indentation test extremely difficult [67]. In the real-world application of AFM indentation, the AFM tip cannot be perpendicular to the sample surface, which causes friction force in tangential direction [6, 66, 68]. Therefore, using a nanoindenter (for vertical loading) in conjunction with an AFM (for visualization purpose to find 1D nanostructures to indent) is proposed [6, 17, 18]. One of the remarkable things from the modeling analysis of the flexural NW/NB indentation is that the indentation depth is independent of the loading location [27, 28], which is corroborated by Sohn's experimental observation of the identical force-indentation depth curves at five different indentation locations of a silicon NW [10].

Let us have a discussion on how the two-spring-in-series model [16, 66] can induce the measurement error of NB Young's modulus and the size effect in a nanoindentation test. Because the indentation test measures force-indentation depth, for Eq. 13.10 to be used to extract nanobelt Young's modulus, modification on Eq. 13.10 is needed. For Hertz contact,  $a = \sqrt[3]{3RP/(4E^*)}$  ( $P$  is the concentrated load exerted by an indenter), and substituting it into Eq. 13.10, we have [28]

$$K_2 = \frac{dP}{d\delta_2} = 2E^*a = (6E^*RP)^{1/3}, \quad (13.12)$$

where  $K_2$  is the slope extracted from the initial unloading curve of  $P - \delta_2$  experimental data and  $R$  is a given value of indenter radius. Therefore, the following equation is used to extract the NB Young's modulus from indentation test:

$$E^* = \sqrt{\frac{K_2^3}{6RP}}. \quad (13.13)$$

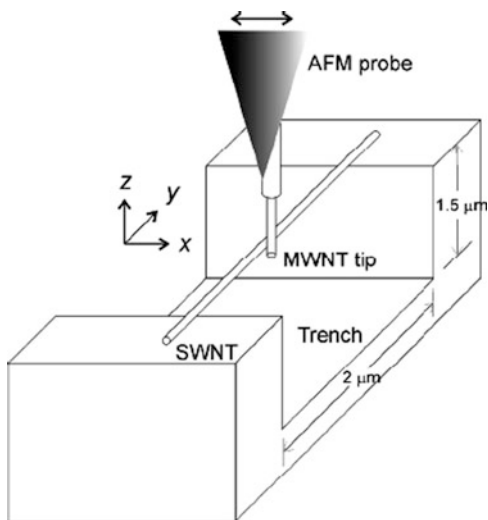
For the two-spring-in-series model,  $\delta = \delta_1^* + \delta_2^*$ ; for the three-spring-in-series model,  $\delta = \delta_1 + \delta_2 + \delta_3$ . Suppose that the AFM cantilever stiffness is accurately calibrated and  $\delta_1^* = \delta_1$ . Therefore,  $\delta_2^* = \delta_2 + \delta_3$ , which is to say that physically the indentation depth of two-spring-in-series model is the indentation depth summation of the indenter/NB contact and NB/substrate contact. So the extraction of  $K_2 = dP/d\delta_2^*$  from experimental data will differ from the real one of  $K_2 = dP/d\delta_2$ , and consequently, Eq. 13.13 can yield erroneous values. Because  $\delta_3 = P/K_3 = P/(2.01bE_s)$  is NB width (size) dependent, the error of  $dP/d\delta_2^*$  is dependent on the NB width, which may be responsible for the ZnO NB Young's modulus dependence on the width-to-thickness ratio observed in the nanoindentation test [16, 21]. Shen et al. [12] have realized that large normal indentation load can cause their NT bending and deformation of substrate, which can influence the measurement accuracy. So they carefully controlled the normal indentation force



$P$  up to  $16.7\ \mu\text{N}$ .  $K_2$  varies nonlinearly with  $P$  or  $\delta_2$ . If  $K_2$  is not accurately extracted (e.g.,  $K_2$  extracted from the initial unloading curve of  $P - \delta_2^*$ ), Eq. 13.13 can easily lead to a (wrong) prediction of  $P$ -dependent Young's modulus of indented sample as shown in Shen's Fig. 3 in reference [12]. The similar error on the hardness measurement of NW/NB may also arise [28].

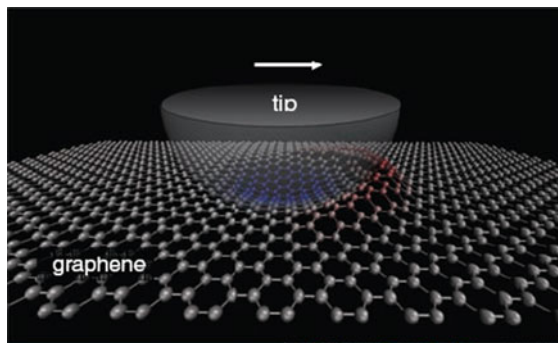
### 13.4 Friction of the Flexural 1D and 2D Structures

As shown in Fig. 13.6, a multiwalled carbon nanotube (MWNT) tip attached to a conventional atomic force microscopy (AFM) probe scans across a single-walled carbon nanotube (SWNT) suspended over a trench [31]. The coefficient of friction of  $0.006 \pm 0.003$  is obtained for the sliding between NTs, which is comparable to the reported value of graphite on nanoscale. The shear strength between NTs is derived to be  $4 \pm 1\ \text{MPa}$  by using a continuum model, which is nearly 2 orders of magnitude larger than the interlayer shear strength of  $0.05\ \text{MPa}$  reported for MWNT in vacuum. Bhushan et al. [31] attributed the huge difference between the intertube and the interlayer shear strength to the following two reasons: (1) the presence of water at the nanotube–nanotube interface in ambient; (2) that the continuum model (the JKR model) might not be accurate enough for the calculation of the contact between NTs. Lee et al. [32] conducted the friction tests on four different atomically thin materials and revealed a universal trend that friction monotonically increased as the number of layers decreased. Because the four materials in test have widely varying electronic and vibrational properties (one metallic, one insulating and two semiconductor materials), as well as the strong substrate effects, Lee et al. [32]



**Fig. 13.6** Schematic of the experimental setup for tribological measurement. A MWNT attached to an AFM scans across a SWNT suspended over a trench (After [31])

**Fig. 13.7** A schematic showing the proposed puckering effect, where adhesion to the sliding AFM tip creates out-of-plane deformation of a graphene sheet, leading to increased contact area and friction; the color scale of the atoms indicates their out-of-plane positions (After [32])

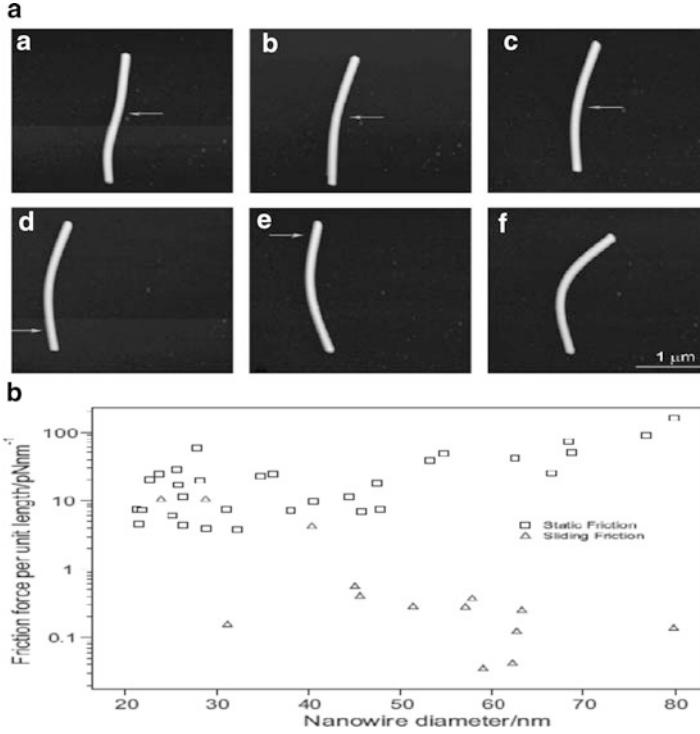


excluded the electronic and phononic dissipations as the cause. Also because they use an amorphous silicon contact tip, the dislocation mechanism is excluded [32]. The out-of-plane bending as shown in Fig. 13.7, which is referred to as puckering, is considered by Lee et al. [32] as the only mechanism. Less number of layers means less bending stiffness. Therefore, under the same normal load, the sheet with less number of layers bends more, which increases both the contact area and friction [32]. Conache et al. [65] measured the friction of InAs NW on  $Si_3N_4$  substrate by using an AFM to horizontally push the NW as shown in Fig. 13.8a. The following abnormal friction property as shown in Fig. 13.8b is found [65]: For large wires, there is a difference between the coefficients of the sliding and static friction of two to three orders of magnitude. At smaller diameters (around 40 nm), this difference disappears and the sliding friction values rise to become equal with those for static friction. Conache et al. [65] interpret this as a transition to stick-slip motion. Figure 13.8b also shows an increase in static friction with NW diameter, which Conache et al. [65] ascribed to the consequence of a change in the NW Young's modulus due to the oxide shell on NW. It is clearly seen in Fig. 13.8a that during the friction test, there is a severe in-plane bending of NW. With a fixed length, the bending stiffness increases with the NW diameter. Therefore, the in-plane bending should play some role in this friction test.

Despite the heavy volume of work performed, so far there is still no fundamental understanding of tribological process [69]. Friction, as quoted from Carpick and Salmeron [69], is “one of the most common, yet least understood, physical phenomenon.” It may be fair to say that except that friction is ultimately due to the electromagnetic forces between the electrons and nuclear particles [70], any other characterization of friction is an open question. The Amontons–Coulomb law states the following:

$$F = \mu P, \quad (13.14)$$

where  $F$  is the friction force and  $P$  is the normal external compression load.  $\mu$  is the friction coefficient, which is a constant and independent of nominal contact area and sliding velocity. There are two types of the frictions: the static and kinetic



**Fig. 13.8** (a) Sliding and static friction experiments performed on a single InAs NW on  $Si_3N_4$ . This wire was 49 nm in diameter and about 2.5 mm in length. First, the whole wire is translated uniformly by pushing at its mid-point. Then it is bent into a tighter curve by pushing on one end at a time. (b) The friction force per unit length versus NW diameter for InAs NWs deposited on a  $Si_3N_4$  substrate. The friction is shown on a logarithmic scale (After [65])

frictions [71]. The static friction is the force needed to start sliding, and the kinetic friction is the force required to maintain sliding [71]. Usually, the static friction is larger than the kinetic one [71]. If adhesion effect is considered, Derjaguin proposed the following law [72]:

$$F = \mu(P_o + P). \quad (13.15)$$

For an adhesive contact, there is still a finite contact area and friction when the external load  $P$  becomes zero [73].  $P_o$  is the “internal” load added to external load to account for the adhesive force [74], and more specifically,  $P_o$  has the same magnitude and opposite sign of the pull-off force. However, Eq. 13.15 encounters some inconsistency of defining the friction coefficient [74]: (1) If the friction coefficient is defined as a ratio, from Eq. 13.15, we have  $F/P = \mu(1 + P_o/P)$ . As  $P_o$  is fixed for given contacting bodies geometries and adhesion,  $\mu$  is not a constant with varying external load, and  $\mu = \infty$  when  $P = 0$ . (2) If the friction

coefficient is defined as a slope, then  $dF/dP = \mu$  is constant. The above friction laws are just a phenomenological characterization. Despite the success of the Amontons–Coulomb law, there is no microscopic theory that explains its origins and wide-ranging applicability [75]. Some analytical theories [76], experiments [77, 78], and computations [79, 80] show that the static friction (almost) vanishes between a pair of clean surfaces that deform elastically. Tomlinson presented an ingenious molecular theory of friction in 1929 [81], which was called by himself as “molecular plucking action” theory. The basics of Tomlinson’s theory are as follows: When two molecules come into contact and then separate, they enter the repulsion and attraction zone of their intermolecular force field. Because of the nonlinear feature of the intermolecular force, the molecules will experience an instability jump during the contact-separate process, which loses energy and is manifest as friction. Muser [82, 83] thinks that this instability jump is an essential thing for friction, and without this instability, friction should vanish or become extremely small. However, one challenge is that in most cases the interbulk interactions are too weak to cause instabilities at the atomic scale [83]. Therefore, the theory that “third bodies” adsorbed on the contacting surfaces arrange to lock the contacting surfaces together was proposed [75, 83], which can also be used as the molecular mechanism to explain the Amontons–Coulomb law. From the fracture mechanics viewpoint, the static friction is the critical tangential force needed to start a mode II crack under a normal load [84, 85]. Johnson suggests that the principal contribution to sliding friction comes from propagation dislocation through the whole area of the interface, i.e., the effective Peierls stress, rather than from nucleation at the periphery [85]. However, as mentioned above, Lee’s experiment results [32] suggest that the Peierls stress may have nothing to do with friction. Savkoor and Briggs [84] derived the following friction–load relation (their Eq. 13.19):  $F \propto (4\gamma\pi RP + 12\gamma^2\pi^2 R^2)^{1/2}$  (notice that their  $\gamma$  has the difference of a factor of 2 with ours defined in Eq. 13.2 and the static friction is called by them as the critical tangential force), which resembles neither the Amontons–Coulomb law nor the Bowden–Tabor law given as follows. Further confusion on the static friction arises according to the recent experiment by Ternes et al. [86]. Ternes et al. [86] used an AFM to measure the force to move a cobalt atom on Pt(111) and Cu(111) surfaces, and they found that the lateral force to move the atom remained constant, whereas the vertical force varied by a large factor. In contrast, the Amontons–Coulomb law, “third bodies” theory [75, 83] and fracture mechanics theory [84, 85] all correlate friction with the vertical/normal force.

As for the kinetic friction, the argument was raised as early as 1804 by Leslie who argued that the energy expended on dragging an asperity to the top of another is simply recovered when it falls down on the other side [74]. Therefore, no energy is ever lost; the two surfaces should simply continue to move once they are set in motion without the need of a constant driving force (to overcome the nonexistent friction force). Some “energy-dissipating” mechanism was therefore called for. Ultimately, energy is dissipated by conversion of kinetic energy of the moving bodies into lattice vibrations (heat) [87]. More specifically, the vibrations of the

surface atoms are damped by energy transfer to bulk phonon modes [88] and in metals also by electronic excitations [87].

The Amontons–Coulomb law is a phenomenological and macroscopic characterization of friction. At the microscopic level, friction is characterized by the following Bowden–Tabor law [89]:

$$F = \tau A. \quad (13.16)$$

$\tau$  is the shear strength and  $A$  is the real contact area. According to Bhushan et al. [90], friction consist of two parts: (1) the force required to shear the junctions formed at the regions of real contact, which is described by the above Eq. 13.16, and (2) the force required to plow the surface of the softer material by the asperities of the harder. It is also important to keep in mind that during the friction test, the instruments actually measure the lateral forces which arise not only from friction but also from the local surface slope [69, 90]. Local variations in the microscale friction of scratched surfaces can be significant and depend on the local surface slope rather than the surface height distributions [69, 90]. In the friction tests of the flexural 1D [31] and 2D [32] nanostructures, the bending-induced out-of-plane deformation should thus contribute more or less the (abnormal) friction observation.

Friction and contact area are measured/calculated in experiment, thus the shear strength can be obtained as  $\tau = F/A$  [73]. If the interaction force of two contacting surfaces is van der Waals described by Eq. 13.2, Xu et al. [91] derived  $\tau = 0.43\sigma_o$  ( $\sigma_o$  is the maximum value given by Eq. 13.2, which is also referred to as the cohesive strength [36]); if friction is thought to result from the movement of dislocation and shear strength, it is regarded as the Peierls stress [92],  $\tau \approx G/30$  [67, 93]. ( $1/G = (2 - \nu_1)/G_1 + (2 - \nu_2)/G_2$ .  $G_1$ ,  $G_2$  and  $\nu_1$ ,  $\nu_2$  are the shear moduli and Poisson's ratios of the two contacting bodies, respectively.) Because several different contact models are presented in Sect. 13.2, which give different contact areas, a question that arises for the real contact is, which model is the appropriate one to describe it? Carpick's experiment [73] shows that the JKR model almost perfectly characterize this real contact even in the range in which the DMT is supposed to apply. There are still some arguments about this real contact area. At the microscale level, the real contact area is an ill-defined concept [35, 36]. For example, at the contact edge, the contact pressure is zero and infinite for the Hertz and the JKR model [35]. At the molecular level, the contact area between two atoms or molecules is an undefined and unnecessary quantity [74]. According to Gao et al. [74], the real contact area is a non-fundamental parameter and the fundamental parameter is the number or density of atoms, molecules, or bonds.

The Hertz model indicates that the real contact area  $A \propto P^{2/3}$  and thus  $F \propto P^{2/3}$  according to the Bowden–Tabor law, which is inconsistent with  $F \propto P$  of the Amontons–Coulomb law. Therefore a natural question arises: How does the Bowden–Tabor law relate with the Amontons–Coulomb law? The simple answer is the surface roughness because the Bowden–Tabor law characterizes the microscopic friction and the Amontons–Coulomb law characterizes the macroscopic friction. Bowden and Tabor [89] gave a straightforward explanation: Because of the surface,

only small portions of the two surfaces are actually in contact, and all these contact portions deform plastically, which gives the result that the contact area is proportional to the external normal load, i.e.,  $P = p_m A$  ( $p_m$  is the “flow pressure” and is approximately equal to the hardness of surface layers [94]). However, not all of the contacting portions deform plastically, and elastic deformation also plays an important role in contact [95]. In Sect. 13.2, there are two implicit assumptions: (1) The two contacting surfaces are “molecularly smooth,” and (2) linear elasticity is used in the derivation of the Hertz, JKR, DMT, and MD models [35, 36]. For those adhesive contact models such as JKR and DMT models, the contact area is not proportional to the normal load, either. Fuller and Tabor [96] showed that adhesion is very sensitive to the surface roughness, which can dramatically reduce the adhesion effect. Therefore, in the study of the rough surface, the Hertz model is sufficient. Greenwood and Williamson [95] developed a model considering the surface roughness and showed how the overall contact area becomes proportional to the external normal load, whereas each contacting asperity is governed by the Hertz model. There are three assumptions used in Greenwood and Williamson’s model [95]: (1) Asperities have a Gaussian height distribution (Persson [70] later showed that it does not matter whether the distribution is exactly Gaussian or not; the only important feature of the height distribution is that the number of asperities of a given height falls off rapidly with increasing height). (2) Asperities can be treated as having spherical caps (for which the Hertz model applies) and as being identical except for their heights. (3) “Peaks” on a surface profile, which are the points higher than their immediate neighbors at the sampling interval used, correspond to asperities. The third assumption is quite wrong, which was realized by Greenwood himself quite some time ago [94]. The second assumption also encounters the “embarrassment” as pointed out by Barber and Ciavarella [97]: The existence of an apparently inexhaustible sequence of smaller and smaller length scales is an embarrassment to asperity model theories, because the definition of an asperity is scale dependent. This scale-dependent model suggests that instruments with different solutions and scan length yield different values of the statistical parameters for the same surface [98, 99]. Therefore, the predictions of the scale-dependent model based on these parameters may not be unique. For example, Greenwood and Williamson defined a dimensionless parameter called “plasticity index,” which measures the extent of plastic deformation to be anticipated in elastic–plastic asperity contact and is given as  $\psi = (E'/H)\sqrt{\sigma/\beta}$  ( $E'$ , reduced modulus, and  $H$ , hardness;  $\sigma$ , the standard deviation of the surface height distribution, and  $\beta$ , the radius of asperity which is treated as a spherical cap by Greenwood and Williamson’s 2nd assumption). Clearly, this plasticity index depends on how the surface is measured ( $\sigma$  and  $\beta$ ). Furthermore, this plasticity index appears to increase without limit as the sampling length is reduced. Another example is that Persson et al. [100, 101] defined a dimensionless parameter called magnification,  $\zeta = L/\lambda$  ( $L$  is the arbitrarily chosen reference scale of the lateral size of the nominal contact area;  $\lambda$  is the shortest wavelength roughness which can be resolved at magnification  $\zeta$ ), and they found that the contact area decreases with the increasing magnification! Furthermore, the fact that surface has hierarchical (or say, fractal)

structure also challenges Greenwood and Williamson's second assumption. Archard in 1957 presented a truly prescient "protuberance on protuberance on protuberance" model [102], which in essence is to say that the surface is fractal. Again, Archard's model [102] demonstrates how the overall contact area becomes proportional to the external normal, whereas each protuberance/asperity is governed by the Hertz model. The basic idea of the Archard's model is like this [70, 102]: Of the primary result of increasing the load is to cause existing contact areas to grow (as described by the Hertz model), then the real contact area will not be proportional to the load. But if the primary result is to form new contact areas of contact, then the area and load will be proportional (as described by those surface roughness models [96, 102]). To avoid the dilemma of the above scale-dependent asperity models, Majumdar and Bhushan [98, 99] used the Weierstrass–Mandelbrot function to characterize the fractal surface as the following:

$$z(x) = G^{(D-1)} \sum_{n=n_1}^{\infty} \frac{\cos(2\pi\gamma^n x)}{\gamma^{(2-D)n}}, \quad (13.17)$$

where  $z(x)$  is the surface profile,  $G$  is a scaling constant, and  $D$  is the fractal dimension; the frequency modes  $\gamma^n$  correspond to the reciprocal of the wavelength of roughness as  $\gamma^n = 1/\lambda_n$ . The above Weierstrass–Mandelbrot function has the following self-affine property:

$$z(\gamma x) = \gamma^{(2-D)} z(x). \quad (13.18)$$

As  $1 < D < 2$  and  $\gamma > 1$ , Eq. 13.16 converges and the advantage of this fractal characterization stands out:  $z(x)$  is composed of a superposition of infinite frequency modes, it is a multiscale function, and there is no need/worry on the truncation problem at small scales as encountered by the above asperity models. Choosing  $\gamma$  is essential as Eq. 13.17 takes its power to form a geometric series.  $\gamma = 1.5$  was taken by Majumdar and Bhushan [98, 99]. However, the choosing of  $\gamma$  and the whole idea of the multiscale characterization of surface as given by Eq. 13.17 are challenged by Greenwood and Wu [94]. The main argument of Greenwood and Wu [94] is based on the following statement given by Greenwood and Williamson [95]: "When the asperity has been compressed to the stage shown, the micro-asperities may possibly have some influence on the actual area of contact, but the nature will depend on the large scale geometry of the asperity."

It should be kept in mind that the above Amontons–Coulomb law and Bowden–Tabor law only characterize the shearing part of friction. As stated before, friction consists of two parts [90]: shearing and plowing. Plowing can play more important when the load is large. Bhushan and Kulkarni's experiment [103] shows that the Amontons–Coulomb law breaks down, which is that coefficient of friction increases with the normal load due to plowing. The Amontons–Coulomb law should only apply for the chunky structure such as a sphere or a block, in which the deformation only occurs around the contacting interface as discussed at the



beginning of Sect. 13.3. Increasing normal external load brings more contact area either by enlarging the existing contact area as described by those contact models given in Sect. 13.2 or by bringing more new asperity into contact as for a rough surface, which is true for a chunky structure but not necessarily true for a flexural structure. If the main function of the external normal load is not to increase the contact area, the Amontons–Coulomb law breaks down as shown by Yang et al. [104] in the contact of NT rafts. Because the NT's Young's modulus in the radial direction is relatively small compared with its Young's modulus in the axial direction, which leads to relatively large cross-section deformation, the work done by the normal pressure is transformed more into the elastic energy rather than to increase the contact area, which leads to the breakdown of the Amontons–Coulomb law [104]. Besides the cross-section flexibility of the shell structure, the bending flexibility may play more important role in determining the contact area as discussed in Sect. 13.3 (for solid NW and NB, there is no worry on the cross-section flexibility). To use the Bowden–Tabor law to characterize the friction problem of the flexural 1D and 2D, the good news is that the tangential friction force will not change the contact area [85]. However, we still face the following additional difficulties or challenges. Firstly, it is to describe the contact area between indenter and flexural nanostructure. The flexural contact models in Sect. 13.3 are about the contact between the flexural nanostructure and substrate. The contact models of Sect. 13.2 used to predict the contact area are for the nonconforming contact [42], which is that two contacting surfaces have dissimilar profiles. As seen in Fig. 13.7, due to the bending of the graphene sheet, the sheet closely wraps the indenter, which may form the conforming contact or say, the classical contact model may break down as pointed out by Bhushan et al. [31]. Because the dimensions of the contact is very small compared with the relative radii of curvature of the surfaces, the relative radii of curvature ( $R$ ) is not influenced by the contact deformation and assumed to be a constant [42]. Here  $1/R = 1/R_1 + 1/R_2$  ( $R_1$ ,  $R_2$  are the radii of curvature of the indenter and nanostructure, respectively). As seen in Fig. 13.7,  $R_1$  of indenter is safe to be assumed constant, but  $R_2$  of the graphene sheet is not.  $R_2$  should vary with the normal load and location. Because the constant  $R$  is an important parameter used in the kinematic assumption to derive the contact model [42], the variation of  $R_2$  may make the contact model to deviate more or less. Also, the elastic half-space model/assumption is also applied in the contact model derivation [42]. The tiny dimensions of the flexural 1D and 2D nanostructures for sure will break the elastic half-space assumption, the substrate effect [28, 29] must be considered. Secondly, it is to evaluate the surface slope effect induced by the bending. Unlike the local surface slope due to surface roughness which has a fixed titling angle [69, 90], the local surface slope induced by the bending of the flexural structure varies with the normal load and mechanical properties of the flexural structure and substrate. It is expected to be rather difficult to characterize. In general, the indenter–flexural structure–substrate forms a rather complex coupled system, which is the essential challenge for modeling the friction of such system. Thirdly, in atomic scale, the classical contact mechanics breakdown physically due to the atomic-scale roughness [105]. Equation 13.2 of surface interaction due to LJ 6–12 potential works fine when



the surface separation is less 10 nm [106]. Due to surface roughness, the LJ 6–12 cannot accurately describe the surface interaction when the surface separation is more than 50 nm [106]. The Casimir force needs to be brought in to model the contact [106, 107]. Furthermore, this review is limited to the dry contact. Once there is liquid trapped between the two contacting surfaces, the capillary effect may play a more important role in contact [106]. Fourthly, at nanometer scale, friction is also influenced by the way how the two contacting crystalline surfaces align, i.e., the degree of commensurability [70, 75, 108]. For 1D nanostructure, the combination of the sliding and rolling motions can arise due to the change of alignment/orientation of 1D nanostructure [65, 109].

**Acknowledgements** This work is supported by the National Natural Science Foundation of China (NSFC nos. 10721202 and 11023001) and the Chinese Academy of Sciences (Grant no. KJCX2-EW-L03). We are also very thankful to Prof. Xinghua Shi of LNM and Dr. Weixue Tian of Caterpillar Champaign Simulation Center for proofreading the manuscript and their revision advices.

## References

1. Y. Zhu, C. Ke, H.D. Espinosa, *Exp. Mech.* **47**, 7 (2007)
2. Z. Wang, *Adv. Mater.* **12**, 1295 (2000)
3. S. Cuenot, C. Frégnigny, S. Demoustier-Champagne, B. Nysten, *Phys. Rev. B* **69**, 165410 (2004)
4. B. Wu, A. Heidelberg, J. Boland, J.E. Sader, X. Sun, Y. Li, *Nano Lett.* **6**, 468 (2006)
5. Y. Chen, B. Dorgan, D. McIlroy, D. Aston, *J. Appl. Phys.* **100**, 104301 (2006)
6. X. Li, H. Gao, C.J. Murphy, K.K. Caswell, *Nano Lett.* **3**, 1495 (2003)
7. G. Feng, W.D. Nix, Y. Yoon, C. Lee, *J. Appl. Phys.* **99**, 074304 (2006)
8. C. Chen, Y. Shi, Y. Zhang, J. Zhu, Y. Yan, *Phys. Rev. Lett.* **96**, 075505 (2006)
9. Y. Zhu, F. Xu, Q. Qin, W. Fung, W. Lu, *Nano Lett.* **9**, 3934 (2009)
10. Y. Sohn, J. Park, G. Yoon, J. Song, S. Jee, J. Lee, S. Na, T. Kwon, K. Eom, *Nanoscale Res. Lett.* **5**, 211 (2010)
11. M. Yu, T. Kowalewski, R.S. Ruoff, *Phys. Rev. Lett.* **85**, 1456 (2000)
12. W. Shen, B. Jiang, B. Han, S. Xie, *Phys. Rev. Lett.* **84**, 3634 (2000)
13. I. Palaci, S. Fedrigo, H. Brune, C. Klinke, M. Chen, E. Riedo, *Phys. Rev. Lett.* **94**, 175502 (2005)
14. M. Minary-Jolandan M. Yu, *J. Appl. Phys.* **103**, 073516 (2008)
15. G. Stan, C. Ciobanu, T. Thayer, G. Wang, J. Creighton, K. Purushotham, L. Bendersky, R. Cook, *Nanotechnology* **20**, 035706 (2009)
16. M. Lucas, W. Mai, R. Yang, Z. Wang, E. Riedo, *Nano Lett.* **7**, 1314 (2007)
17. X. Li, X. Wang, Q. Xiong, P. Eklund, *Nano Lett.* **5**, 1982 (2005)
18. H. Ni, X. Li, *Nanotechnology* **17**, 3591 (2006)
19. X. Bai, P. Gao, Z. Wang, E. Wang, *Appl. Phys. Lett.* **82**, 4806 (2003)
20. W. Mai, Z. Wang, *Appl. Phys. Lett.* **89**, 073112 (2006)
21. M. Lucas, Z. Wang, E. Riedo, *Appl. Phys. Lett.* **95**, 051904 (2009)
22. S. Mao, M. Zhao, Z. Wang, *Appl. Phys. Lett.* **83**, 993 (2003)
23. Y. Zhang, Y. Zhao, J. Adhes. Sci. Tech. **25**, 1107 (2011)
24. W.C. Oliver, G.M. Pharr, *J. Mater. Res.* **7**, 1564 (1992)
25. I.N. Sneddon, *Int. J. Eng. Sci.* **3**, 47 (1965)

26. W.C. Oliver, G.M. Pharr, *J. Mater. Res.* **19**, 3 (2004)
27. Y. Zhang, Y. Zhao, *J. Appl. Mech.* **78**, 011007 (2011)
28. Y. Zhang, *J. Appl. Phys.* **107**, 123518 (2010)
29. R. Saha, W.D. Nix, *Acta Mater.* **50**, 23 (2002)
30. K. Geng, F. Yang, E. Grulke, *Mater. Sci. Eng. A* **479**, 157 (2008)
31. B. Bhushan, X. Ling, A. Jungen, C. Hierold, *Phys. Rev. B* **77**, 265428 (2008)
32. C. Lee, Q. Li, W. Kalb, Z. Liu, H. Berger, R.W. Carpick, J. Hone, *Science* **328**, 76 (2010)
33. D. Maugis, *J. Colloid Interface Sci.* **150**, 243 (1992)
34. K.L. Johnson, J.A. Greenwood, *J. Colloid Interface Sci.* **192**, 326 (1997)
35. Y. Zhang, *J. Adhes. Sci. Tech.* **22**, 699 (2008)
36. Y. Zhang, *J. Adhes. Sci. Tech.* **25**, 1435 (2011)
37. J.A. Greenwood, *Proc. Roy. Soc. London Ser. A* **453**, 1277 (1997)
38. J.A. Greenwood, *Phil. Mag.* **89**, 945 (2009)
39. J.N. Israelachvili, *Intermolecular and Surface Forces*. (Academic, London, 1985)
40. C. Argento, A. Jagota, W.C. Carter, *J. Mech. Phys. Solids* **45**, 1161 (1997)
41. R.S. Bradley, *Phil. Mag.* **13**, 853 (1932)
42. K.L. Johnson, *Contact Mechanics*. (Cambridge University Press, Cambridge, 1985)
43. P. Attard, J.L. Parker, *Phys. Rev. A* **46**, 7959 (1992)
44. M.K. Chaudhury, T. Weaver, C.Y. Hui, E.J. Kramer, *J. Appl. Phys.*, **80**, 30 (1996)
45. L.M. Keer, J. Dundurs, K. Tsai, *J. Appl. Mech.* **39**, 1115 (1972)
46. S.L. Pu, M. Hussain, *J. Appl. Mech.* **37**, 859 (1970)
47. G.M. Gladwell, *J. Appl. Mech.* **43**, 263 (1976)
48. Y. Weitsman, *J. Appl. Mech.* **36**, 505 (1969)
49. Y. Weitsman, *J. Appl. Mech.* **37**, 1019 (1970)
50. Y. Weitsman, *Int. J. Eng. Sci.* **10**, 73 (1972)
51. Y. Zhang, K.D. Murphy, *Int. J. Solids Struct.* **41**, 6745 (2004)
52. Y. Zhang, *Int. J. Mech. Sci.* **50**, 1035 (2008)
53. L. Ascione, A. Grimaldi, *Meccanica* **19**, 223 (1984)
54. A.D. Kerr, *J. Appl. Mech.* **31**, 491 (1964)
55. A.E. Giannakopoulos, *J. Mech. Phys. Solids* **54**, 1305 (2006)
56. F. Yang, Y.T. Cheng, *J. Mater. Res.* **24**, 1976 (2009)
57. J. Salvétat, G.A.D. Briggs, J. Bonard, R. Bacsá, A.J. Kulik, T. Stöckli, N.A. Burnham, L. Forró, *Phys. Rev. Lett.* **82**, 944 (1999)
58. Y. Chen, I. Stevenson, R. Pouy, L. Wang, D.N. McIlroy, T. Pounds, M.G. Norton, D.E. Aston, *Nanotechnology* **18**, 135708 (2007)
59. W. Fang, J.A. Wickert, *J. Micromech. Microeng.* **6**, 301 (1996)
60. Y. Zhang, Y. Zhao, *Microsyst. Technol.* **12**, 357 (2006)
61. G. Jing, H. Duan, X. Sun, Z. Zhang, J. Xu, Y. Li, J. Wang, D. Yu, *Phys. Rev. B* **73**, 235409 (2006)
62. S. Cuenot, S. Demoustier-Champagne, B. Nysten, *Phys. Rev. Lett.* **85**, 1690 (2000)
63. A. San Paulo, J. Bokor, R.T. Howe, R. He, P. Yang, D. Gao, C. Carraro, R. Maboudian, *Appl. Phys. Lett.* **87**, 053111 (2005)
64. Y. Zhu, Q. Qin, Y. Gu, Z.L. Wang, *Nanoscale Res. Lett.* **5**, 291 (2010)
65. G. Conache, S.M. Gray, A. Ribayrol, L.E. Fröberg, L. Samuelson, H. Pettersson, L. Montelius, *Small* **5**, 203 (2009)
66. R.W. Carpick, D. Ogletree, M. Salmeron, *Appl. Phys. Lett.* **70**, 1548 (1997)
67. M.A. Lantz, S.J. O'Shea, M.E. Welland, K.L. Johnson, *Phys. Rev. B* **55**, 10776 (1997)
68. Y. Zhang, K.D. Murphy, in *Scanning Probe Microscopy in Nanoscience and Nanotechnology*, ed. by B. Bhushan (Springer, Berlin, 2010), Chapter 8
69. R.W. Carpick, M. Salmeron, *Chem. Rev.* **97**, 1163–1194 (1997)
70. B.N.J. Persson, *Sliding Friction: Physical Principles and Applications*, 2nd edn. (Springer, Berlin, 2000)
71. E. Rabinowicz, *Friction and Wear of Materials*. (Wiley, New York, 1965)
72. B.V. Derjaguin, *Z. Phys.* **88**, 661 (1934)

73. R.W. Carpick, N. Agrait, D.F. Ogletree, M. Salmeron, Langmuir, **12**, 3334 (1996)
74. J. Gao, W. Luedtke, D. Gourdon, M. Ruth, J.N. Israelachvili, U. Landman, J. Phys. Chem. B, **108**, 3410 (2004)
75. G. He, M.H. Muser, M.O. Robbins, Science, **284**, 1650 (1999)
76. M.O. Robbins, E.D. Smith, Langmuir **12**, 4543 (1996)
77. J. Krim, D.H. Solina, R. Chiarello, Phys. Rev. Lett. **66**, 181 (1991)
78. C. Mak, J. Krim, Phys. Rev. B **58**, 5157 (1998)
79. M. Cieplak, E.D. Smith, M.O. Robbins, Science **265** 1209 (1994)
80. M.S. Tomassone, J.B. Sokoloff, A. Widom, J. Krim, Phys. Rev. Lett. **79**, 4798 (1997);
81. G.A. Tomlinson, Philos. Mag. Ser. 7, 905 (1929)
82. M.H. Muser, D. Shakhvorostov, Science, **328**, 52 (2010)
83. M.H. Muser, Phys. Rev. Lett. **89**, 224301 (2002)
84. A.R. Savkoor, G.A.D. Briggs, Proc. R. Soc. Lond. A. **356**, 103 (1977)
85. K.L. Johnson, Proc. R. Soc. Lond. A. **453**, 163 (1997)
86. M. Ternes, C.P. Lutz, C.F. Hirjibehedin, F.J. Giessibl, A.J. Heinrich, Science, **319**, 1066 (2008)
87. J.Y. Park, D.F. Ogletree, P.A. Thiel, M. Salmeron, Science, **313**, 186 (2006)
88. R.J. Cannara, M.J. Brukman, K. Cimat, A.V. Sumant, S. Baldelli, R.W. Carpick, Science, **318**, 780 (2007)
89. F.P. Bowden, D. Tabor, *The Friction and Lubrication of Solids*. (Clarendon Press, Oxford, 1954)
90. B. Bhushan, J.N. Israelachvili, U. Landman, Nature, **374**, 607 (1995)
91. D. Xu, K. Ravi-Chandar, K.M. Liechti, J. Colloid Interface Sci. **318**, 507 (1992)
92. A. Cottrell, *Introduction to the Modern Theory of Metals*. (Institute of Metals, London, 1988)
93. R. Szoszkiewicz, B. Bhushan, B.D. Huey, A.J. Kulik, G. Gremaud, J. Appl. Phys. **122**, 144708 (2005)
94. J.A. Greenwood, J.J. Wu, Meccanica **36**, 617 (2001)
95. J.A. Greenwood, J.B.P. Williamson, Proc. R. Soc. Lond. A. **295**, 300 (1966)
96. K.N. Fuller, D. Tabor, Proc. R. Soc. Lond. A. **345**, 327 (1975)
97. J.R. Barber, M. Ciavarella, Int. J. Solids Struct. **37**, 29 (2000)
98. A. Majumdar, B. Bhushan, ASME J. Tribol. **112**, 205 (1990)
99. A. Majumdar, B. Bhushan, ASME J. Tribol. **113**, 1 (1991)
100. B.N.J. Persson, Eur. Phys. J. E **8**, 385 (2002)
101. C. Yang, U. Tartaglino, B.N.J. Persson, Eur. Phys. J. E **19**, 47 (2006)
102. J.F. Archard, Proc. R. Soc. Lond. A. **243**, 190 (1957)
103. B. Bhushan, A.V. Kulkarni, Thin Solid Films **278**, 49 (1996)
104. W. Yang, H. Wang, Y. Huang, ASME J. Eng. Mater. Technol. **127**, 383 (2005)
105. B. Luan, M.O. Robbins, Nature **435**, 929 (2005)
106. F. Delrio, M.P. de Boer, J.A. Knapps, E. Reedy Jr., P. Clews, M. Dunn, Nature Mat. **4**, 629 (2005)
107. Y.P. Zhao, L. Wang, Y.T.X. Yu, J. Adhes. Sci. Tech. **17**, 519 (2003)
108. M.R. Falvo, J. Steele, R.M. Taylor II, R. Superfine, Tribol. Lett. **9**, 73 (2000)
109. M.R. Falvo, R.M. Taylor II, A. Helser, V. Chi, F.P. Brooks Jr., S. Washburn, R. Superfine, Nature **397**, 236 (1999)