Mapping of three-dimensional contact problems into one dimension

Thomas Geike and Valentin L. Popov
Institute of Mechanics, Technische Universität Berlin, D-10623 Berlin, Germany
(Received 26 October 2006; revised manuscript received 29 January 2007; published 27 September 2007)

We consider a contact problem between two three-dimensional bodies with randomly rough surfaces and show how this problem can be reduced from three to one dimension without loss of essential contact properties. This means a huge reduction of computation time and allows the simulation of multiscale systems to include essentially all of the scales from nanometer to macroscopic in a single model.

DOI: 10.1103/PhysRevE.76.036710 PACS number(s): 02.70.Ns, 46.55.+d

I. INTRODUCTION

Many physical systems and processes have a multi-scale fractal structure, which essentially determines their properties [1]. Many tribological systems belong to the fractal system class. In friction processes, both the microscopic and macroscopic scales may play an essential role [2]. For example, a monolayer of impurity atoms on a metal surface changes both friction and wear dramatically. On the other hand, the contact mechanics of real surfaces (fractal in most cases) is controlled by space scales differing by many orders of magnitude [3]. The same problems arise in simulation of rubber: the rolling resistance and the wear in this material can be caused by space and time scales differing by about 10 orders of magnitude [4]. A correct simulation of such systems has to take into account all these scales. The multiscale nature of a system often raises the question of if and how it can be simulated numerically [5].

In the present paper, we discuss this general question for contact mechanics of randomly rough surfaces. We show that in contact problems it is possible to map the main properties of a three-dimensional system onto a simplified one-dimensional system. We achieve this in such a way that the relavant properties of the contact problem remain invariant, while the required computation time is drastically reduced.

II. SINGLE CONTACT

The first main idea of the proposed reduced description is the following [6]: Consider the three-dimensional contact problem of two smooth surfaces with relative radius of curvature R_3 and elastic modulus $E^*=E/(1-\nu^2)$ (ν is the Poisson ratio). According to the Hertz' theory, the correlation between normal force F_3 and approach d reads [7]

$$F_3(d)=\frac{4}{3}E^*\sqrt{R_3 d^3}. \tag{1}$$

The relation between normal force F_3 and radius of contact a is

$$F_3(a)=\frac{4E^*}{3R_3}a^3. \tag{2}$$

Now consider the one-dimensional contact problem depicted in Fig. 1. The respective relations are

$$F_1(d)=\frac{4\sqrt{2}c_n}{3}\sqrt{R_1 d^3} \tag{3}$$

and

$$F_1(a)=\frac{2c_n}{3R_1}a^3, \tag{4}$$

where c_n is the stiffness per unit length. The distance Δx between particles is assumed to be small compared to the size of the contact. The macroscopic relations between force and approach, and force and radius of contact will be identical for the three-dimensional and one-dimensional problem if

$$R_1=\frac{1}{2}R_3, \quad c_n=E^*. \tag{5}$$

Hence, the three-dimensional contact problem for one single spherical contact can be reduced to a one-dimensional problem for an arbitrary curvature radius. The physical nature of this coincidence lies in the proportionality of the stiffness of a three-dimensional contact to the contact length and not to the contact area [7]. This property is valid both for normal and tangential contact. It shows immediately that the one dimensional system with constant stiffness density c_n is an appropriate candidate for simulating three-dimensional contact problems. Note not only the force-displacement dependence, but also other dependencies which are of interest for contact mechanics (e.g., contact length-force and contact area determined as πa^2) coincide for the true three-dimensional and the reduced one-dimensional system. Further, mapping Eq. (5) is completely correct for any cylindrical indenter. Indeed, the stiffness of the contact in this case is $2aE^*$ both in three- and one-dimensional cases.

The local force in the one-dimensional problem is

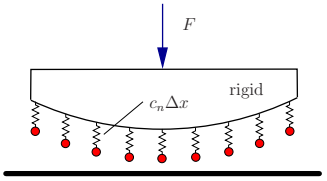

FIG. 1. (Color online) Contact between a rigid plate and a rigid cylinder with an elastic layer.

$$f(x) \propto (a^2 - x^2), \tag{6}$$

which is different from the well-known result for contact pressure in the three-dimensional problem [7]. The desired relation is reached by defining the stress in the one-dimensional problem according to

$$\sigma(x) = \frac{f(x)}{b\sqrt{\delta(x)R_1}}, \tag{7}$$

where $\delta(x)$ is the local deformation and b is the effective width. Then

$$\sigma(x) \propto \sqrt{1 - \frac{x^2}{a^2}}. \tag{8}$$

By choosing the effective width b appropriately, the stress according to Eq. (7) is identical to the three-dimensional result. The definition of stress according to Eq. (7) is not arbitrary, but has a simple physical meaning. First note that the radius of contact is $a=\sqrt{\delta R}$. When taking Eq. (7) as an estimation of the average stress, the correct stress definition for the three-dimensional problem $\sigma = \text{const}\frac{F}{a^2} \approx \text{const}\frac{F/a}{a}\text{const}\frac{f}{\sqrt{R\delta}}$ is obtained. However, if we let the spring force $f(x)$ and δ depend on the coordinate x, we get the exact Hertz stress distribution. That is the "empirical" reason that we use the stress definition (7).

The possibility of reducing three-dimensional asperities to an equivalent one-dimensional problem is also valid in the case of elliptical asperities by using the geometrical average $\sqrt{R^{(1)}R^{(2)}}$ of the principal radii $R^{(1)}$ and $R^{(2)}$ of an asperity instead of R. This is a very good approximation as long as the ratio of both principal radii is not too large. For example, for the ratio $R^{(2)}/R^{(1)}=10$, the difference between the force in a three-dimensional problem and in the one-dimensional problem with equivalent radius $\sqrt{R^{(1)}R^{(2)}}$ will be 2.5% [7]. Even in the extremely unusual case $R^{(2)}/R^{(1)}=100$, the force in a three-dimensional problem will differ from that of the one-dimensional problem by only 18%. Thus, in all typical cases the result of mapping a three-dimensional problem into one-dimension will be excellent and in the worst case still relatively good. The fact that the force-displacement relations (1) and (3) are equivalent independent of the radius of curvature let us hope that this result is very robust and will not depend on the exact form of the asperity. Direct simulations of contact problems with rough surfaces support this (see next section).

III. MULTI ASPERITY CONTACT

To go over from contact of single asperities to randomly rough surfaces, we need a rule of generating the "equivalent" one-dimensional surfaces from two-dimensional ones. As a motivation for the analytical form of the recalculation rule, we use some ideas of Greenwood and Williamson (GW) model [8]. However, we do not claim to support the correctness of GW model and we do not use in our mapping the independence of asperities.

In the Greenwood and Williamson model, the interaction between neighboring asperities is of no importance for the contact problem as long as the size of micro contacts is much smaller than the distance between them. It is rather the statistics of heights and radii of curvature which are important for the contact problem[1] [8] (see also a comment in this section). The statistics of microcontacts determines on one hand the normal forces between bodies. On the other hand, it determines the real area of contact and thus the tangential friction forces. The distributions of normal and tangential forces as well as the contact area of microcontacts are the most important quantities for understanding and qualitative characterization of tribological systems on a microscale. As we have shown that a single three-dimensional asperity can be equivalently substituted by a one-dimensional asperity—independent of the radius of curvature—the next step is to create a one-dimensional surface with the same contact properties. In the GW approximation this would be a surface with the same statistical distributions of height and curvature as the three-dimensional body's two-dimensional surface. The one-dimensional asperity would then have in the GW approximation the same contact properties as the initial three-dimensional body. In the present section we study whether it is possible to create such an "equivalent" one-dimensional surface (line) and if possible, how it is done correctly.

For simplicity, we assume here that the topography of a two dimensional surface (from a three-dimensional body) can be characterized by its surface roughness power spectrum $C_{2D}(q)$ defined by

$$C_{2D}(q) = \frac{1}{(2\pi)^2}\int \langle h(\mathbf{x})h(0)\rangle e^{-i\mathbf{q}\cdot\mathbf{x}}d^2x, \tag{9}$$

where $h(\mathbf{x})$ is the height measured from the average plane, defined so that $\langle h\rangle=0$ and $\langle\cdots\rangle$ stands for ensemble averaging [3]. Since it is assumed that the statistical properties of the surface topography are translationally invariant and isotropic, the surface roughness power spectrum $C_{2D}(q)$ only depends on the magnitude q of the wave vector $\mathbf{q}$.

Similarly, a surface roughness power spectrum $C_{1D}(q)$ can be introduced for a one-dimensional "surface" topography according to

$$C_{1D}(q) = \frac{1}{2\pi}\int \langle h(x)h(0)\rangle e^{-iqx}dx. \tag{10}$$

To generate a one-dimensional surface equivalent to the initial two-dimensional surface, the appropriate surface roughness power spectrum $C_{1D}(q)$ must be defined. The qualitative arguments for the choice of proper one-dimensional spectral density are the following: The height distribution of a fractal surface's asperities generally has the same order of magnitude as the mean square root value of the height $h(x)$ of the

[1]This implies that the interrelation between microcontacts is neglected. As shown in Ref. [9] this is a good approximation at small Hurst exponents of surface roughness, while at larger Hurst exponents the approximation of independent scales (Persson *et al.* [4]) is more adequate. The results for contact area in both approaches have, however, the same analytical form and differ only by a constant factor.

profile. The mean curvature of the asperities has the same order of magnitude as the mean square root value of the curvature $\kappa=\partial^2 h(x)/\partial x^2$. The mean-square values of height for two- and one-dimensional systems

$$\langle h^2\rangle_{2D} = 2\pi \int_0^\infty q C_{2D}(q) dq, \tag{11}$$

$$\langle h^2\rangle_{1D} = 2 \int_0^\infty C_{1D}(q) dq, \tag{12}$$

will be equal, if we take

$$C_{1D}(q) = \pi q C_{2D}(q). \tag{13}$$

It is important to note that the mean square curvatures $\langle \kappa^2 \rangle$ will then be equal as well.[2] The mapping is only possible, if the phases in the Fourier spectrum can be considered as not correlated. Such effects as form difference of asperities and valleys [11] cannot be described with the above procedure. For description of real surfaces, further studies are needed.

We now show that not only the average height and curvature values will be almost equal for these systems, but also their distributions. For this sake, we study how the statistics of asperities of both two- and one-dimensional systems are related to each other. We generated one-dimensional and two-dimensional surface topographies and determined the height and curvature distributions numerically. The surface topography is calculated from the surface roughness power spectrum according to

$$h(\mathbf{x}) = \sum_{\mathbf{q}} B_{2D}(\mathbf{q}) \exp\{i[\mathbf{q}\cdot\mathbf{x} + \phi(\mathbf{q})]\}, \tag{14}$$

where $\phi(\mathbf{q}) = -\phi(-\mathbf{q})$ are randomly distributed in $[0, 2\pi)$ and

$$B_{2D}(\mathbf{q}) = \frac{2\pi}{L}\sqrt{C_{2D}(\mathbf{q})} = \bar{B}_{2D}(-\mathbf{q}). \tag{15}$$

For the one-dimensional case the respective equations are

$$h(x) = \sum_{q} B_{1D}(q) \exp\{i[qx + \phi(q)]\}, \tag{16}$$

$$B_{1D}(q) = \sqrt{\frac{2\pi}{L} C_{1D}(q)} = \bar{B}_{1D}(-q). \tag{17}$$

Numerical generation of surfaces is based on the FFT algorithm rather than directly calculating the sums in Eqs. (14) and (16). The distribution of asperities is calculated for each generated surface topography. We introduce the following ratios ϕ_1, ϕ_2 and ϕ_3, which relate the asperity statistics (index p) to the profile statistics:

$$\phi_1 = \sqrt{\frac{\langle h_p^2\rangle}{\langle h^2\rangle}}, \quad \phi_2 = \frac{\langle \kappa_p\rangle}{\sqrt{\langle \kappa^2\rangle}}, \quad \phi_3 = \sqrt{\frac{\langle \kappa_p^2\rangle}{\langle \kappa^2\rangle}}. \tag{18}$$

Figure 2 shows ϕ_1, ϕ_2, and ϕ_3 for one-dimensional surfaces (top) and for two-dimensional surfaces (bottom), for a constant power spectrum with cutting wave vectors $q_1 = 2q_0$:

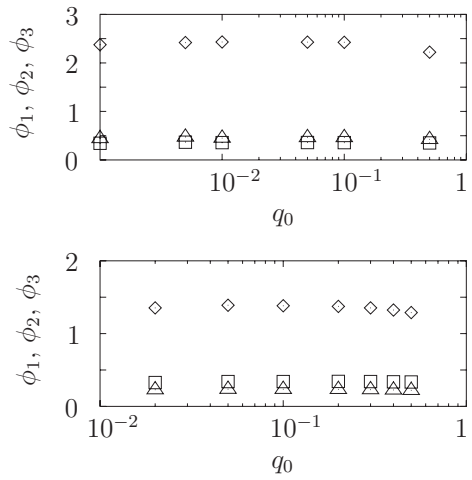

FIG. 2. Ratios ϕ_1, ϕ_2, and ϕ_3 according to Eq. (18) for a 1D surface (top) and for a 2D surface (bottom). $q_1 = 2q_0$, □: ϕ_1, ◇: ϕ_2, △: ϕ_3.

$$C_{2D} = \begin{cases} c & \text{for } q_0 \le q \le q_1, \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

and C_{1D} according to Eq. (13). ϕ_1, ϕ_2, and ϕ_3 hardly depend on the wave number q_0. Further numerical experiments with $2 \le q_1/q_0 \le 10$ show that this characteristic is also present for $q_1 \neq 2q_0$.

The following important conclusion regarding the statistics of asperities (index p) can finally be drawn from numerical studies with generated one- and two-dimensional surface topographies $2 \le q_1/q_0 \le 10$): if the surface roughness power spectrum is transformed according to Eq. (13) the statistics of asperities will transform according to

$$\langle h_p^2\rangle_{1D} \approx \langle h_p^2\rangle_{2D},$$

$$\langle \kappa_p\rangle_{1D} \approx 1.8\langle \kappa_p\rangle_{2D},$$

$$\langle \kappa_p^2\rangle_{1D} \approx 2.0\langle \kappa_p^2\rangle_{2D}.$$

Note that according to Eq. (5) the relation for the average curvature of asperities should preferably be $\langle \kappa_p\rangle_{1D} = 2\langle \kappa_p\rangle_{2D}$. Choosing the stiffness c_n to achieve the correct $F(d)$ relation, the contact radius a will not be exactly equal in the two models. In the case at hand the relation $\langle \kappa_p\rangle_{1D} \approx 1.8\langle \kappa_p\rangle_{2D}$ leads to about 5% error in the radius of contact.

Our argument about the independence of different asperities does not play an important role in this method. In fact, we claim that the method would also work by crossover between different scales of asperities. The physical reason for this is the following. Imagine a roughness with a very large wave length superimposed with very small wave length. If brought into contact, first the small asperities will deform and the analogy will work exactly because all of them are of the same scale and in the first approximation independent. Now consider a very large normal pressure

[2] In accordance with the comment at the end of Sec. II, we define for a two-dimensional surface $\kappa^2 = \kappa^{(1)}\kappa^{(2)}$ with $\kappa^{(1)}$ and $\kappa^{(2)}$ being principal curvatures of the surface.

such that the surfaces are in a complete contact on the of small asperities scale. After this point, the displacement-force dependence will be correct again as now only the larger asperity is valid and the correctness does not depend on the curvature radius. The only question is whether the exact crossover from one scale to the other is correct.

In order to clarify the crossover behavior and support the conclusion regarding the identity of three-dimensional and reduced one-dimensional problems, we have carried out test simulations of a true three-dimensional system and compared it with the one-dimensional counterparts generated according to the rules described in Sec. III. Surface displacement $\bar{u}_z$ of a linear elastic half space under the action of pressure p can be calculated according to

$$\bar{u}_z(x,y) = \frac{1}{\pi E^*} \int \int_{(A)} \frac{p(\hat{x},\hat{y})}{\sqrt{(\hat{x}-x)^2 + (\hat{y}-y)^2}} d\hat{x}d\hat{y}. \quad (20)$$

The total applied normal load is then $\iint_{(A)} p(x,y)dxdy$. The boundary conditions for this equation are the following: Pressure must be positive inside the contact region, and zero outside. Furthermore, the gap is zero inside the contact region and positive outside. Numerical evaluation of displacement and pressure requires Eq. (20) to be discretized and then solved iteratively. This method is a boundary element approach to the normal contact problem, valid under the half space assumption.

Using Eqs. (14) and (19), 450 two-dimensional surface topographies with 64×64 points have been generated. The relation between normal force F and contact area A_c has then been calculated by the boundary element method just described. Figure 3 shows the exemplary contact regions for one surface and one value of normal force. The total contact area A_c is the sum of the areas of all microcontacts. The results of all simulations are shown as error-bars in Fig. 4. Mean values satisfy an approximately linear relation (thin dashed line). Error bars represent the standard deviation of the respective values.

One-dimensional simulations have been performed with spectra according to Eq. (13). Microcontacts are now characterized by their respective lengths a_j. The microcontacts are identified as connected regions in the simulation, and the contact area is then calculated according to

$$A_{c,1D} = \frac{\pi}{4} \sum_j a_j^2. \quad (21)$$

The force-contact area relation for the one-dimensional simulation is also shown in Fig. 4 (thick dotted line). The linear approximation of the three-dimensional result and one-dimensional result match almost exactly, even at relatively large contact areas of about 10% of the apparent area. This is in accordance with results by Campana and Müser [11]. Note that the analytical result considering independent asperities by Bush overestimates the contact area at the given force [9,10], while the analytical results of Persson carried out in the approximation of independent scales underestimate the area [4].

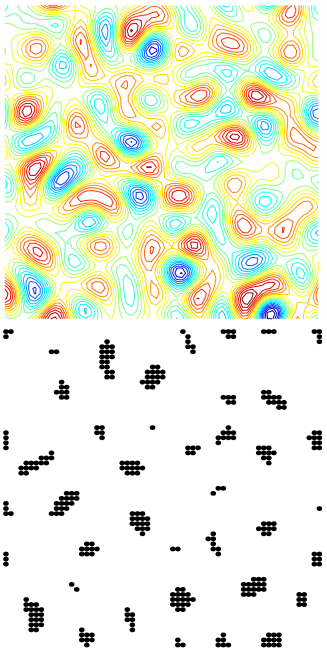

FIG. 3. (Color online) Two-dimensional surface topography (top) and microcontacts due to the applied normal force (bottom).

IV. SUMMARY

For the example of a contact problem between randomly rough surfaces we have studied whether it is possible to reduce the dimension of systems from three to one while leaving the essential contact properties invariant. We have shown that it is indeed possible—as long as the contact area is much smaller then the apparent (macroscopic) contact area. Even at real contact areas equal to 10% of the apparent area, the congruence between the results of the exact three-dimensional and reduced one-dimensional systems is almost exact. We have proved this for a variety of single asperity forms and for random surfaces with the power spectrum localized around some characteristic wave vector. Qualitative

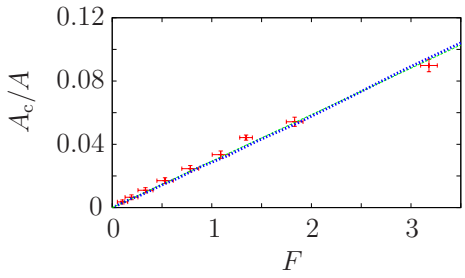

FIG. 4. (Color online) Relation between relative contact area (A_c real contact area, A apparent contact area) and normal force F, three-dimensional results (error bars based on 450 surfaces and thin dashed line for linear approximation of mean values) compared to one-dimensional results (thick dotted line). One-dimensional result and linear approximation of three-dimensional results are difficult to distinguish because they match very closely.

arguments of Sec. III show that the mapping should work also for fractal surfaces, but we have not proved it so far.

This all means a huge reduction of the computation time, which principally allows the simulation of multiscale systems while simultaneously including essentially all scales from nanometer (10^{-9} m) to macroscopic (approximately 10^{-2} m). A model containing about 10^7 points is then necessary for a one-dimensional system, a number still possible for simulations run on fast computers.

[1] B. B. Mandelbrot, Science **279**, 5352 (1998); H. Qi, J. Ma, and P. Z. Wong, Phys. Rev. E **64**, 041601 (2001); H. T. Williams, L. Goodwin, S. G. Desjardins, and F. T. Billings, Phys. Lett. A **250**, 1 (1998); A. E. M. El-Misiery and E. Ahmed, Appl. Math. Comput. **178**, 2 (2006).

[2] B. N. J. Persson, F. Bucher, B. Chiaia, Phys. Rev. B **65**, 184106 (2002); F. Bucher, K. Knothe, and A. Theiler, Wear **253**, 1 (2002); B. N. J. Persson, *Sliding Friction: Physical Principles and Applications*, 2nd ed. (Springer, Berlin, 2000); A. Majumdar and C. L. Tien, Wear **136**, 2 (1990); K. Komvopoulos and N. Ye, J. Tribol. **123**, 7 (2001).

[3] B. N. J. Persson, Phys. Rev. Lett. **87**, 116101 (2001); **89**, 245502 (2002).

[4] B. N. J. Persson, O. Albohr, U. Tartaglino, A. I. Volokitin, and E. Tosatti, J. Phys.: Condens. Matter **17**, 1 (2005).

[5] C. Yang, U. Tartaglino, and B. N. J. Persson, Eur. Phys. J.: Appl. Phys. **19**, 1 (2006).

[6] T. Geike and V. L. Popov, Tribol. Int. **40**, 924 (2007); V. L. Popov and S. G. Psakhie, *ibid.* **40**, 916 (2007).

[7] K. L. Johnson, *Contact Mechanics*, 6th ed. (Cambridge University Press, Cambridge, 2001).

[8] J. A. Greenwood and J. B. P. Williamson, Proc. R. Soc. London, Ser. A **295**, 300 (1966); A. W. Bush, R. D. Gibson, and T. R. Thomas, Wear **35**, 87 (1975).

[9] S. Hyun, L. Pei, J.-F. Molinari, and M. O. Robbins, Phys. Rev. E **70**, 026117 (2004).

[10] A. W. Bush, R. D. Gibson, and T. R. Thomas, Wear **35**, 87 (1975).

[11] C. Campana and M. H. Müser, Europhys. Lett. **77**, 38005 (2007).