On the question of theoretical justification of the Amontons-Coulomb law for friction of unlubricated surfaces
V A Zhuravlev
DOI: 10.1243/13506501JET176

The online version of this article can be found at:
http://pij.sagepub.com/content/221/8/893

Published by:
SAGE
http://www.sagepublications.com

On behalf of:
Institution of Mechanical Engineers

Additional services and information for Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology can be found at:

Email Alerts: http://pij.sagepub.com/cgi/alerts
Subscriptions: http://pij.sagepub.com/subscriptions
Reprints: http://www.sagepub.com/journalsReprints.nav
Permissions: http://www.sagepub.com/journalsPermissions.nav
Citations: http://pij.sagepub.com/content/221/8/893.refs.html

>> Version of Record - Aug 1, 2007

What is This?
On the question of theoretical justification of the Amontons–Coulomb law for friction of unlubricated surfaces

V A Zhuravlev


INTRODUCTION TO ZHURAVLEV’S HISTORICAL PAPER

The manuscript was received on 7 February 2006 and was accepted after revision for publication on 5 March 2007.

DOI: 10.1243/13506501JET176

The book on history of tribology by Kragelsky and Shchedrov [1] calls attention to pioneering works on tribological problems by da Vinci, Amontons, Euler, Kotelnikov and Coulomb. These historical studies are referred to by Dowson in his book on the history of tribology [2] which is well known amongst English speaking scientists. These books, and papers that have been drawn based on them, describe the development of tribology in various countries. In particular, they gave credit to some tribological studies in Russia, e.g. Euler and his former student Kotelnikov [3] are credited with the use of the Greek mu ($\mu$) to represent the friction coefficient (see, e.g. Blau [4]). It is worthwhile for the tribology community to be made aware of some other important contributions that have not been published in English, and this contribution seeks to achieve this for Zhuravlev’s paper.

The statistical modelling of a nominally flat rough surfaces by independent spherical, elastic protuberances with a random height distribution that are deformed according to Hertz, is now classic. Assuming that the protuberances have the same radii, but various heights (so that the protuberance number at a given height increases with depth), the result that the true contact area is approximately proportional to the external compressing force can be obtained. A number of publications and text books on tribology have associated this result with the Greenwood and Williamson model of 1966 [5]. In fact, this model for purely elastic contact was developed by Zhuravlev and published in 1940 [6]. A similar scheme developed by Greenwood and Williamson [5] modifies the Zhuravlev model by considering elasto-plastic transition of protuberances. Although Zhuravlev’s results were discussed in a number of Russian books on contact mechanics and tribology (e.g. Kragelsky and Shchedrov [1], Demkin [7], Kragelsky et al. [8], Galin [9], Sviridenok et al. [10], Argatov and Dmitriev [11]), his paper is rarely cited in technical literature published in English. To the best of the author’s knowledge, by 2001 the paper had only been cited three times in English language publications, namely by Johnson [12], Greenwood [13] and Adams and Nosonovsky [14]. In 2001 the author translated the paper into English and started to publicise it to the tribology community (see, e.g. [15]).

Zuravlev did not continue his study in the field started in the paper. Unfortunately little is known of his fate. He was a young researcher who worked in the Tajik capital Dushanbe which at that time was called as Stalinabad. His supervisor was Davidenkov, a member of the USSR Academy of Sciences and a staff member of the Ioffe Physico-Technical Institute. Kragelsky reportedly received uncertain information that Zhuravlev had perished during the Second World War (M. N. Dobychin, 2004, personal communication). The Zhuravlev paper was published in Zhurnal Tekhnicheskoi Fiziki (Journal of Technical Physics). The journal was founded by the Ioffe Physico-Technical Institute of the USSR Academy of Sciences in 1931. In 1956 the Journal was started to be translated into English first as Soviet Physics – Technical Physics, and then as Technical Physics. Currently this journal contains practical information on all aspects of applied physics, especially on molecular physics, surface physics, instrumentation and measurement techniques. When Technical Physics was founded, it was the leading Russian journal not only in applied physics but also in theoretical mechanics. However, the centre of gravity of mechanical publications has since moved to Journal of Applied Mathematics and Mechanics (PMM), another journal of the USSR Academy of
Sciences, was founded in 1937. This may well be the reason that the Zhuravlev paper has not been more widely recognized by the Tribology community.

Kragelsky [16] noted that instead of protuberances that are deformed according to Hertz, Zhuravlev could consider flat-ended protuberances. Indeed, the model of multiple contact between flat-ended protuberances and an elastic half-space was introduced by Goryacheva and Dobychin [17] (see also Galin [9] and Goryacheva [18]). This approach was also used by Borodich and Mosolov [19] in application to fractal rough surfaces. The interactions between contact spots for both flat-ended cylindrical protuberances and spherical protuberances were studied by Goryacheva [18].

Kragelsky [16] also noted that the height distribution of protuberances is approximately Gaussian rather than the linear distribution considered by Zhuravlev. However, Kragelsky did not use the Gaussian distribution in his calculations. Greenwood [13] notes that the Greenwood and Williamson model [5] improves on Zhuravlev by assuming a Gaussian height distribution (still of identical spherical caps). However, as can be seen from [5], they considered an exponential distribution of the sphere tops as an approximation to the Gaussian distribution.

The other significant improvement by Greenwood and Williamson is that Zhuravlev did not consider the question of whether the deformation of a protuberance will be elastic or plastic, whereas including this consideration leads to a new plasticity index [5]

$$\Psi = \frac{E^*}{H} \sqrt{\frac{\sigma}{R}}$$

where $E^*$ is the contact elastic modulus of the two bodies, $H$ is the hardness, $\sigma$ is the standard deviation of the summit heights and $R$ is the radius of curvature of the protuberances. The physical sense of $\Psi$ is as follows [13]: if $\Psi = 1$ then about 1 per cent of the surface is plastically deformed, if $\Psi = 0.9$, then there is hardly any plastic deformation, while for $\Psi = 1.1$, the proportion of plastic area is much more than 1 per cent.

So far, only contact aspects of the Zhuravlev paper have been discussed. However, as is clear from its title, the main purpose of the paper was to consider the friction law. Giving a high appraisal of Zhuravlev paper, Kragelsky and Shchedrov [1] also made the following two criticisms: (a) Zhuravlev assumed that the specific frictional force is constant, i.e. $\mu = F_i/P$, while only one term of two in the Amontons–Coulomb law does not depend on the pressure; and (b) the linear distribution of spherical cap heights is not observed in reality. The latter criticism was also made by Demkin [7]. Indeed, it is known [1, 2, 18] that while the Amontons law for the friction force $F_i$ is $F_i = \mu P$, the Coulomb law suggests that the total friction force could be represented as the sum of a constant force ($A$) depending on sticking of surfaces and a force that depends on the load $P$: $F_i = A + \mu P$. Derjaguin [20] gave a molecular meaning to the Coulomb's force $A$. He wrote $F_i = \mu (P + S_p)$ where $S$ is the true area of the interacting surface, and $p_s$ is the specific attractive force. Hence, the term $A = \mu S_p$ represents the tangential component of the force of molecular interactions. Zhuravlev actually considered the Amontons law rather than the Amontons–Coulomb law. However, if adhesive forces can be neglected then the Amontons–Coulomb law reduces to the Amontons law and Zhuravlev was the first to offer an explanation of this phenomenon. Concerning the criticism on linear distribution of the protuberance heights, it is worthwhile to note that Zhuravlev gave a general expression for the true contact area, as can be seen from the paper (see equation (8)), and he considered the linear distribution as an approximation that allows analytical integration of the expression. Zhuravlev also noted that his preliminary calculations showed that other distribution functions may give better results.

In addition to the comments mentioned above, Kragelsky noted that although the limits of integration were not taken accurately by Zhuravlev, this does not affect the final results from a practical point of view. The question of integration was discussed recently by Argatov and Dmitriev [11].

The idea of offering a translation of Zhuravlev’s paper for publication and to write this introduction was influenced by Rossmanith’s introduction [21] to his translation of Wieghardt’s pioneering paper on fracture mechanics [22]. It is interesting to note that before Rossmanith’s translation was published, only German and Russian researchers had referred to the paper (see, e.g. [23]). One can see that the history of science is a very difficult task. However, studies of the history of science lead to better understanding of the development of ideas and help to avoid repetitions of errors. The author hopes that this is only the first historical paper to appear in the Journal of Engineering Tribology and that one day everyone will be able to read in English the famous historical papers by C. Cataneo, B. V. Derjaguin, and other great scientists who worked in the field of tribology.

ACKNOWLEDGEMENT

The author is grateful to Dr M. N. Dobychin (Institute for Problems in Mechanics, Russian Academy of Sciences, Moscow) for sharing his memories about his late teacher Professor I. V. Kragelsky.
REFERENCES

1 Kragelsky, I. V. and Shchedrov, V. S. Development of the science about friction (in Russian), 1956 (USSR Academy of Sciences Press, Moscow).


3 Kotelnikov, S. K. A book containing the doctrine on equilibrium and motion of bodies (in Russian), 1774 (Naval Noble College, Saint-Petersburg).


7 Demkin, N. B. Actual area of contact between solid surfaces (in Russian), 1962 (USSR Academy of Sciences Press, Moscow).


On the question of theoretical justification of the Amontons–Coulomb law for friction of unlubricated surfaces

V A Zhuravlev


Translated from the Russian by F. M. Borodich.

The laws of friction, first formulated by Amontons [1] and later discovered independently by Coulomb [2], have not yet received a full theoretical justification despite the great practical importance of friction phenomena (this statement relates only to dry friction, i.e. friction of unlubricated surfaces). For friction of lubricated surfaces, there is a hydrodynamic theory due mainly to Reynolds). It seems that the most substantiated among the few suggested theories of dry friction is that presented by Tomlinson [3] (there is a review of theory of friction in a paper by Kontorova (journal 'Uspekhi fizicheskikh nauk', 1937, XVIII, issue 3)). Tomlinson sees cohesion between molecules of the rubbing solids as the cause of the frictional force. He considers the process of friction as successive decohesion of contacting molecules and creation of new molecular contacts. In practice, rubbing surfaces are almost never ideally smooth; so that their contact is realized by a number of contacts between separate surface 'protuberances'. Assuming that both the number of contacts between separate
molecules and the cohesion force are functions of the true contact area, and calculating the area of contact between two protuberances using the Hertz theory of elastic contact between two balls, Tomlinson finds an expression for the coefficient of friction up to a factor. He gives a satisfactory explanation for the observed correlation between the coefficients of friction and elastic constants of solids. However, as shown by Kontorova [4], the Tomlinson theory has a serious disadvantage. It follows from the theory that the coefficient of friction must be inversely proportional to the cube root of normal pressure whereas in reality it is a constant independent of the load. In Kontorova’s opinion the cause of this disagreement between the Tomlinson theory and practice is that the contacts between protuberances are elastic only for small pressures, and the coefficient of friction must indeed be inversely proportional to the cube root of normal pressure for these conditions. However, the character of contact changes for greater loads, as it becomes plastic and ceases to depend on the pressure.

The cause of the independence of the coefficient of friction and the pressure at plastic contacts was not studied by Kontorova in detail. However, this explanation is not very convincing. Indeed, if the validity of the Amontons–Coulomb law at large normal pressures were caused by plastic deformations then it would be unclear why Amontons’s law is valid for brittle solids like glass.

An attempt to derive Amontons’s law theoretically, based on somewhat different assumptions concerning the nature of contact between rubbing surfaces is given below.

2

Following Tomlinson, let us assume that rubbing surfaces are covered by a number of protuberances. However, we assume that the protuberances have various heights and the number of the protuberances (or more precisely, the number of the summits) at a specific height increases as the level of consideration goes deeper into the rough surface. Upon pressing two rough surfaces together, the highest protuberances come into contact first, and they are deformed as the compressive load increases, causing new, deeper protuberances to come into contact. Let us consider an arbitrary element of the area of the compressed surfaces Δσ. We will assume the topography and the elastic properties of the surfaces are the same and that the protuberances of the rough surfaces are spherical and the radii are the same. However, the number of protuberances situated at different levels will vary. Let the distribution of the protuberance summits at various levels for the element Δσ of a rough surface be characterized by a function

\[ n = n(\xi) \]

\[ dn = n(\xi) d\xi \]

is the number of summits situated in the layer dξ at depth ξ. The total number of protuberance summits situated at various levels of the element Δσ from the highest summit to the level of depth x, is

\[ N = \int_0^x n(\xi) d\xi \]

Let a compressive force P act for the element Δσ. Let the compressing displacement between both surfaces be equal to x (under the term ‘displacement’ here and later on, we mean the relative approach between two points of either solid situated at such distances between the points that the compressive deformations can be neglected). Since the surfaces are completely the same, the compressing displacements x₁ and x₂ of each surface are

\[ x_1 = x_2 = \frac{x}{2} \]

Let us calculate the force P that has to act on the element Δσ to give the compressing displacement x. We will assume that deformations are entirely elastic. The force is composed of forces arising at separated contacts between protuberances. The largest forces are at contacts between the highest protuberances (ξ = 0) because their displacements are the greatest.

For these we have from Hertz theory of elastic contact between spheres

\[ x = \sqrt{\frac{9\pi^2 p^2 k^2}{2R}} \]  

(1)

where x is the compressing displacement, p is the force acting on a protuberance, R is the radius of the spherical protuberance, \( k = (1 - \nu^2)/\pi E \), \( \nu \) is the Poisson ratio and E is the Young’s modulus. Writing (1) in terms of the force p, we have

\[ p = \frac{1}{3\pi k} \sqrt{2R} x^{3/2} \]  

(2)

We now consider a protuberance of one of the surfaces (surface A), whose summit is situated at depth \( \xi_1 \). In the process of compression, it contacts the protuberance of the other surface (surface B), whose summit is situated at depth \( \xi_2 \). If the compressing displacement is x then the displacement of the contacting protuberances is \( x - (\xi_1 + \xi_2) \) because these protuberances must have an approach displacement of \( \xi_1 + \xi_2 \) before contact ensues. The force arising at the contact is therefore

\[ \frac{\sqrt{2R}}{3\pi k} [x - (\xi_1 + \xi_2)]^{3/2} \]  

(3)

We now consider a layer dξ₁ of surface A situated at depth \( \xi_1 \). It will have \( n(\xi_1) d\xi_1 \) of the protuberance summits. Let us find the force corresponding to all
protuberances whose sums are in this layer. As shown above, at the contact between one of the protuberances of the layer and a protuberance of the layer \( d\xi_2 \) situated at the depth \( \xi_2 \), the force arising is

\[
\frac{\sqrt{2R}}{3\pi k} [x - (\xi_1 + \xi_2)]^{3/2}
\]

The expected number of such contacts is equal to the product of the number of summits in the layer \( d\xi_1 \) and the probability of a summit of the layer \( d\xi_1 \) (surface A) meeting a protuberance summit of the layer \( d\xi_2 \) of surface B. The probability is equal to the ratio of the number of summits in the layer \( d\xi_1 \) to the total number of summits of the element \( \Delta\sigma \) of surface B (FB): There was a misprint in the original paper where the expression was written \((n(\xi_1) d\xi_2)/N\).

\[
n(\xi_1) d\xi_1 \frac{n(\xi_2) d\xi_2}{N}
\]

Thus the expected number of such contacts is

\[
n(\xi_1) d\xi_1 \frac{n(\xi_2) d\xi_2}{N}
\]

The corresponding force is equal to

\[
\frac{\sqrt{2R}}{3\pi k N} [x - (\xi_1 + \xi_2)]^{3/2} n(\xi_1) n(\xi_2) d\xi_1 d\xi_2
\]

(4)

We obtain the force acting on all protuberances of the surface A with summits in the layer \( d\xi_1 \) by adding together expressions (4) for all active layers of surface B. Thus, integrating (4) with respect to \( \xi_2 \) from 0 to \( x/2 \) gives

\[
\int_0^{x/2} \frac{\sqrt{2R}}{3\pi k N} [x - (\xi_1 + \xi_2)]^{3/2} n(\xi_1) n(\xi_2) d\xi_1 d\xi_2 = \frac{\sqrt{2R}}{3\pi k N} \int_0^{x/2} [x - (\xi_1 + \xi_2)]^{3/2} n(\xi_1) n(\xi_2) d\xi_1 d\xi_2
\]

(5)

Finally, to obtain the force acting on the element \( \Delta\sigma \) of surface A, we must sum the forces acting on its various layers, i.e. integrate the above expression with respect to \( \xi_1 \) from 0 to \( x/2 \). Thus

\[
P = \frac{\sqrt{2R}}{3\pi k N} \int_0^{x/2} \int_0^{x/2} [x - (\xi_1 + \xi_2)]^{3/2} n(\xi_1) n(\xi_2) d\xi_1 d\xi_2
\]

(6)

Now we can calculate the true area of contact between an element \( \Delta\sigma \) of surface A and the corresponding element \( \Delta\sigma \) of surface B. This area is the sum of contact areas for separate protuberances. From the theory of elastic contact between two spheres, we have the formula for the radius \( a \) of the contact region between two spheres under compression

\[
a = \sqrt{\frac{3\pi k R p}{4}}
\]

(7)

The area of an elementary contact region \( \Delta S \) is

\[
\Delta S = \pi a^2 = \pi \left( \frac{3\pi k R p}{4} \right)^{2/3}
\]

Let us express \( \Delta S \) in terms of the compressing displacement \( x \) using (1)

\[
\Delta S = \pi \left( \frac{3\pi k R p}{4} \right)^{2/3} p^{2/3}
\]

\[
= \pi \left( \frac{3\pi k R p}{4} \right)^{2/3} \left( \frac{\sqrt{2R}}{3\pi k} \right)^{2/3} x = \frac{\pi R x}{2}
\]

At the contact between a protuberance of surface A whose summit is situated at depth \( \xi_1 \) and a protuberance of surface B whose summit is situated at depth \( \xi_2 \), the contact area is

\[
\frac{\pi R}{2} [x - (\xi_1 + \xi_2)]
\]

For computing the true contact area \( S \) of all protuberances of the element \( \Delta\sigma \), we use arguments completely analogous to those used above for computing the force \( P \). We thus obtain the contact area as

\[
S = \frac{\pi R}{2N} \int_0^{x/2} \int_0^{x/2} [x - (\xi_1 + \xi_2)] n(\xi_1) n(\xi_2) d\xi_1 d\xi_2
\]

(8)

Let us assume that the number of contacts between individual molecules of surfaces A and B is proportional to the true contact area \( S \). The frictional force \( F \) must be proportional to the number of molecular contacts, therefore, it will also be proportional to the true contact area

\[
F = \alpha S
\]

(9)

where \( F \) is the frictional force, \( S \) is the true contact area, and \( \alpha \) is a coefficient depending on cohesive forces.

From (8) and (9) we obtain the following expression

\[
F = \frac{\alpha \pi R}{2N} \int_0^{x/2} \int_0^{x/2} [x - (\xi_1 + \xi_2)] n(\xi_1) n(\xi_2) d\xi_1 d\xi_2
\]

(10)

Now we can obtain the relation between the frictional force \( F \) and the force of normal compression \( P \). For this purpose, we should exclude \( x \) from expressions (6) and (10).

In order to eliminate \( x \) it is first necessary to evaluate the integrals in these equations. However, integrating (6) and (10), requires knowledge of the function \( n = n(\xi) \), i.e. we should know the distribution of depths for protuberances of the rough surfaces.
Determination of the form of the function \( n = n(\xi) \) is a rather difficult procedure. So in the first instance we will approximate it as a linear function, i.e. we put

\[
n(\xi) = \beta \xi
\]

where \( \beta \) is a proportionality coefficient. Then (6) and (10) become

\[
F = \frac{\alpha \pi R \beta^2}{2N} \int_0^{\pi/2} \int_0^{\pi/2} [x - (\xi_1 + \xi_2)] \xi_1 \xi_2 \, d\xi_1 \, d\xi_2
\tag{11}
\]

and

\[
P = \frac{\sqrt{2} \pi R \beta^2}{3 \pi k N} \int_0^{\pi/2} \int_0^{\pi/2} [x - (\xi_1 + \xi_2)]^{3/2} \xi_1 \xi_2 \, d\xi_1 \, d\xi_2
\tag{12}
\]

Integrating (12), we obtain

\[
P = \frac{\sqrt{2} \pi R \beta^2}{3 \pi k N} C_1 x^{11/2}
\tag{13}
\]

Integrating (11), we obtain

\[
F = \frac{\alpha \pi R \beta^2}{2N} C_2 x^3
\tag{14}
\]

where \( C_1 \) and \( C_2 \) are constants.

Eliminating \( x \) from (13) and (14) and combining the constants into a single factor \( C \), we have

\[
F = C \frac{\alpha R^{1/11} \beta^2}{N^{1/11}} k^{10/11} \nu^{10/11}
\tag{15}
\]

or taking into account that \( k = ((1 - \nu^2)/\pi E) \),

\[
F = C \frac{\alpha R^{1/11} \beta^2}{N^{1/11}} \left( \frac{1 - \nu^2}{\pi E} \right)^{10/11} \nu^{10/11}
\]

\[
= C \frac{\alpha R^{1/11} \beta^2}{N^{1/11}} \left( \frac{1 - \nu^2}{\pi E} \right)^{10/11} \nu^{10/11} P
\tag{16}
\]

It follows from the above expression that the relation between the normal compressing force and the frictional force is close to linear \((10/11 = 0.909 \equiv 1)\), i.e. Amontons’s law is JUSTIFIED in a general way.

For the coefficient of friction \( \mu \), we obtain from (16)

\[
\mu = C \frac{\alpha R^{1/11} \beta^2}{N^{1/11}} \left( \frac{1 - \nu^2}{\pi E} \right)^{10/11} \nu^{10/11} P
\]

Hence, the coefficient of friction \( \mu \) is inversely proportional to the eleventh root of the load, i.e. its dependence on the load is very small.

The dependence of the coefficient of friction on the elastic constants is defined by the factor \(((1 - \nu^2)/\pi E)^{10/11} \mu \). If the above arguments are valid then the ratio \(((1 - \nu^2)/\pi E)^{10/11} \mu \) should be approximately constant for various materials.

The numerical values of the ratio \(((1 - \nu^2)/\pi E)^{10/11} \mu \) for some materials are presented in Table 1.

<table>
<thead>
<tr>
<th>Material</th>
<th>( E ) (kgf/mm(^2))</th>
<th>( \nu )</th>
<th>( \mu )</th>
<th>( ((1 - \nu^2)/\pi E)^{10/11} \mu )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steel</td>
<td>21 500</td>
<td>0.28</td>
<td>0.39</td>
<td>3.40</td>
</tr>
<tr>
<td>Cu</td>
<td>12 600</td>
<td>0.34</td>
<td>0.60</td>
<td>3.45</td>
</tr>
<tr>
<td>Al</td>
<td>7200</td>
<td>0.34</td>
<td>0.94</td>
<td>3.60</td>
</tr>
<tr>
<td>Pt</td>
<td>17 000</td>
<td>0.39</td>
<td>0.44</td>
<td>3.42</td>
</tr>
<tr>
<td>Sn</td>
<td>5500</td>
<td>0.39</td>
<td>1.11</td>
<td>4.00</td>
</tr>
</tbody>
</table>

This ratio should not be expected to be exactly constant because the expression for the coefficient of friction contains factors depending on the cohesive forces that are different for different materials.

As preliminary calculations have shown, the expression for the relation between the normal compressive force and the frictional force can be improved (in the sense of closer approximation to the linear relation observed in the reality) by proper choice of the function \( n = n(\xi) \). For this purpose, a device is being constructed at the Physical Laboratory of the T. G. Shevchenko Pedagogical Institute (Stalinabad) to determine the contact area between rubbing surfaces by measurement of electrical resistance of the junction.

ACKNOWLEDGEMENT

The author expresses his deep acknowledgment to academician N. N. Davidenko who suggested the topic of the work presented and gave suggestions as to its fulfillment.

REFERENCES

1 Amontons, G. Histoire de l’Académie Royale des Sciences avec les mémoires de mathématique et de physique, 1699.
2 Coulomb, C. A. Mémoires de mathématique et de physique de l’Académie Royale, 1785.
3 Tomlinson, G. A. Phil. Mag., 1929, 7, 905.
4 Kontorova, T. UFN, XVIII, 3, 1937.

Stalinabad.

Submitted on 19 June 1940.