TEKA Kom. Mot. Energ. Roln. – OL PAN, 2007, 7, 250–260

THEORIES OF FRICTION AND THEIR APPLICABILITY TO SOIL

Arvids Vilde, Guntis Sevostjanovs*, Janusz Nowak**

*Research Institute of Agricultural Machinery, Latvia University of Agriculture,
e-mail: vilde@delfi.lv
**Management and Production Engineering, Agricultural University of Lublin,
Głęboka Str. 28, 20-612 Lublin, Poland

Summary. A brief review of the friction theories is given. The mechanical, molecular, molecular-kinetic and molecular-mechanical theories of friction are analysed and discussed. For the investigations of soil the most acceptable is Deryagin's opinion and the two-part formulae that allow determination of soil friction as a specific adhesion on the working surfaces.

Key words: theories of friction, the mechanical friction theory, the molecular friction theory, the molecular-kinetic friction theory, the molecular-mechanical friction theory

INTRODUCTION

In most cases soil, when it is cultivated, moves by sliding along the steel surfaces of the operating tools of the soil tillage machines. The sliding resistance of soil affects significantly the draft resistance of the machines. For example in ploughing, friction caused by sliding of soil along a steel surface may constitute 35–50% and more of the total draft resistance [Vilde *et al.* 2004]. Therefore great importance is always attached to the issues how to reduce the sliding resistance of soil moving along the operating tools both when new structures are designed and when the existing machines are used. In order to tackle these issues skilfully, it is necessary to know the regularities which determine the value of the sliding friction of soil. Insufficient knowledge of the impact of various factors upon the frictional properties of soil made us start the survey of the existing general theories of friction and their applicability to soil.

GENERAL THEORIES OF FRICTION

The sliding of soil along a steel surface is a particular case in the totality of every kind and sort of cases when one body slides in relation to another. That is why, before we start discussing the issues connected with sliding of soil along a steel surface, it is expedient to make a brief analysis of the existing most general scientific developments in the theory of friction. Besides, particular attention was paid to those assumptions which could contribute to the recognition of the regularities of the issue to be studied.

The mechanical friction theory. Amontons' Law

The study of friction of solid bodies has a five centuries' old history. According to references [Kragelski 1956, 1968, Deryagin 1963], the first man who formulated the basic regularities of dry friction was the great Italian scientist Leonardo da Vinci at the end of the 15th century. In 1699, on the basis of the experimental data the French physicist G. Amontons arrived at the same conclusions made by Leonardo da Vinci and formulated the law of friction in its classical form: resistance to the relative movement of solid bodies is proportional to the normal load and does not depend on the contact area between the bodies; i.e.

$$F = fN, \tag{1}$$

where:

F – resistance to the relative movement of solid bodies;
N – the normal load on the surface of friction;
f – the friction coefficient (according to Leonardo da Vinci f = 0,25; according to Amontons f = 0,3 for all bodies).

In its further development the study of friction took the path of verification and specification of individual assumptions of this law [Krotova 1960, Semenov *et al*. 1961, Kragelski 1968]. Among the scientists whose works are devoted to these problems was the Swiss mathematician L. Euler. In 1765 he deduced an equation, which allowed determination of the tractive effort for the movement of flexible bodies with friction along a cylindrical surface, as well as their friction force against the cylinder:

$$T_1 = T_2 e^{f\alpha}, \tag{2}$$

$$F = T_2(e^{f\alpha} - 1), \tag{3}$$

where:

T_1 and T_2 – the tension of the guiding and the guided branches of a flexible body (thread);
e = 2,718..... – the base of natural logarithms:
f – the friction coefficient;
α – the angle of contact of the cylinder (block) with the flexible body (thread).

Euler's formula (2, 3) finds wide application in our times, too, in calculations of underground devices, belt transmissions, band brakes, etc. The works by American scientists R.D. Doner and M.L. Nicols [Doner *et al*.1934], as well as by the authors of this article [Vilde 1967] show that similar phenomena take place also when soil slides along the curved surface of the operating tools of soil tillage machines.

The works by many researchers basically confirmed Amontons' Law, it was also found out that the values of the friction coefficient for various materials are different and peculiar to particular pairs of rubbing materials. At the same time there was an opinion that the force of friction depends on the contact surface and the friction coefficient depends on the normal pressure.

Coulomb's binomial formula. The molecular theory of friction

In 1779 the French physicist C.-A. Coulomb expressed the force of friction as a binomial [Kragelski and Shchedrov 1956, Deryagin1963], the contemporary appearance of which is:

$$F = C + f\,N \tag{4}$$

The first term in this formula of the expression is dependent on the adhesion degree of the surfaces but the second – on the value of pressure upon them.

It is evident from the expressions (1) and (4) that Amontons' Law is a particular case of a more general regularity revealed by Coulomb. Coulomb already distinguished static friction and dynamic (kinetic, sliding) friction. He believed that friction at the starting moment of sliding of one surface along the other is dependent on four reasons: 1) the nature of the materials; 2) the length of the surfaces; 3) the pressure between the surfaces; 4) the length of time which has passed since the moment of the contact of the two surfaces, but in the case of relative sliding, the friction between the surfaces depends only on the first three factors.

In spite of its great completeness Coulomb's formula was forgotten for almost entirely a century, and Amontons' formula was applied in technical calculations. It should be noted that the latter formula was often used in our days, too, in technical calculations where application of the binomial formula of friction resistance is required.

Further investigations after Coulomb were generally directed to the establishment of the dependence of the force of friction on the nature of the rubbing bodies, as well as on the application of lubricants. It was discovered that the resistance force to sliding depends on the nature of the bodies, area, the length of time of the contact and specific pressure. With heterogeneous bodies the force of friction is determined by the abrasion of the softer body, and this force is greater for softer bodies and lesser for solid bodies; reduction in the force of friction, when a lubricant is used, depends on the nature of the lubricant and is not dependent on the nature of the mutually sliding bodies. Tables of the mean values of coefficients were drawn up for different materials.

Besides the mechanical theories of friction in which it was asserted that friction is resistance to the lifting of a body over a great number of small surface irregularities (Amontons), or resistance to the elimination of surface irregularities (Leslie), or it is determined by the adhesion of irregularities, which are deformed or onto which the body should be raised (De la Guir, Coulomb), or else it is connected with the scratching process of the rubbing surfaces at the places where overload occurs (Gümbel), hypotheses of molecular friction were advanced [Semenov *et al.* 1961, Kragelski 1968]. They attracted particular interest just during the last fifty years. Various authors proposed again the forgotten Coulomb's formula with different physical substantion (Sachs, Morrow, Deryagin, Prishin). A connection is shown in several articles between the friction coefficient and the hardness of the rubbing bodies (Rebinder, Ernst, Merchant) [Kragelski and Shchedrov 1956, Kragelski 1960, Krotova 1960, Semienov and Chaykovskiy 1961, Dieryagin 1963, Vilde and Rucins 2006].

According to the purely molecular theory by Tomlinson (1929) [Kragelski and Shchedrov 1956, Semienov and Chaykovskiy 1961] friction is a process of successive disengagement of contacting molecules and appearance of molecular contacts. The friction work is equal to the energy consumed for the disengagement of molecules. However, the subsequent investigations carried out by a number of researchers, such as Deryagin, Kragelski [Kragelski 1956, 1968, Deryagin 1963], produced data which significantly differed from Tomlinson's theory and put it in doubt.

Deryagin's molecular-kinetic theory

More perfect is the molecular-kinetic theory of friction worked out and experimentally founded Deryagin (1934) [Kragelski 1956, Deryagin 1963]. Deryagin asserts that the external friction of bodies is determined by the molecular surface roughness, which is an inevitable consequence of the atomic structure of the bodies. He points to the fact that the actual coefficient of friction depends on the molecular-atomic roughness of the contacting surfaces, but the force of friction, besides the molecular roughness, is also dependent on the forces of molecular interaction [Deryagin 1963]. Deryagin arrived at the conclusion that Amontons' Law, which deals with the proportionality of the force of friction to normal load, results from the molecular-atomic surface roughness of the bodies but deviations from this law are determined by the action of the molecular adhesion forces (attraction) that are proportional to the contact area of the bodies.

Deryagin explains the friction mechanism due to the molecular roughness of surfaces by means of a visual model. The unshaded circles in the friction model (Fig. 1) represent atoms of one body, the shaded circles – of the other body. It is supposed that the form and the size of the contacting atoms does not change. In addition to it, since this model should explain the occurring forces of friction, no forces of friction are allowed at the contact points of adjacent atoms. Besides, under its proper weight *P*, the upper body assumes a position in which the centre of gravity of the body turns out to be in the lowest position 0 (Fig. 1a and b). Under the impact of the horizontal force the upper body is displaced from the lowest position to a new one, and the centre of gravity describes a certain curve 00_1 (Fig. 1c).

A formula is deduced from the equilibrium condition of the body located on the inclined surface:

$$\frac{F}{P} = \mathrm{tg}\alpha \tag{5}$$

If *F/P* exceeds some limit, equal to $\mathrm{tg}\alpha_\mathrm{m}$, where α_m – the maximum inclination of the trajectory of the centre of gravity, then equilibrium will be violated, and continued sliding of the upper body along the lower one will begin. Hence the expression of the coefficient of static friction μ follows:

$$\mu = \mathrm{tg}\alpha_m \tag{6}$$

In such a way the appearance of the force of friction can be explained by the molecular roughness, which is the result of the atomic structure of the bodies.

A phenomenon, similar to the one described by the model occurs with any sliding hard surfaces being in mutual contact. The path of the movement of the centre of gravity of the upper body in relation to the lower one always has the form of an undulatory curve (Fig. 1c), the height of the humps of which depends on size of the atoms and molecules located on the contact surfaces. Such undulatory movements result not from the roughness of the surface but on the molecular roughness which cannot be eliminated by any polishing because it is connected with the atomic-molecular structure of solid bodies.

In addition to it, as mentioned above, the force of friction depends also on the molecular interaction (attraction) of both bodies.

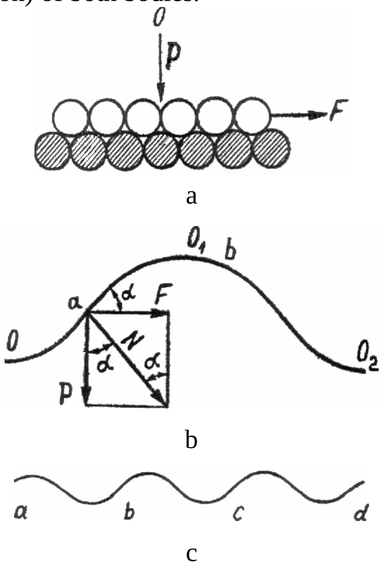


Fig. 1. A model of the molecular friction mechanism according to Deryagin [1963]:
a – the model of rubbing surfaces; b – the scheme of equilibrium of the upper body;
c – the trajectory of the centre of gravity of the body sliding along the „surface" area

Deryagin's binomial law of friction is a generalisation of the laws proposed by Amontons and Coulomb expressed by the following formula [Deryagin 1963, Kragelski 1956, 1968]:

$$F = f_0(N + N_0) = f_0(N + S_k p_0) = f_0 S_k(p + p_0)\,, \tag{7}$$

where:

N_0 – the force of molecular interaction;
S_k – the genuine contact area;
p – the specific pressure;
p_0 – the resultant of forces of molecular attraction per unit of the genuine contact area.

In this formula f_0 is the actual friction coefficient of sliding whose constant character results from the discussed theory. In contrast to it, as evident from the formula, the common design coefficient of friction (more precisely, the coefficient of sliding resistance), which is equal to the relation of the entire force of friction to the external load

$$F = FN^{-1} = f_0(1 + N_0 N^{-1})\,, \tag{8}$$

is not a constant value: it increases, as confirmed by the experiments, when the load decreases.

The genuine contact area S_k is many times lesser than the area of the apparent (nominal) contact of the rubbing surfaces, and it may vary with the value of the nominal load. It follows from the binomial law of friction (7) that the force of friction $F \neq 0$ when $N = 0$ but $F = f_0 N_0 = f_0 S_k p_0$.

Deryagin points out that in reality the apparent compliance with the monomial Amontons' Law may be an application of the binomial law of friction. This may take place in those cases when one of the rubbing bodies is plastic (paraffin, Wood's metal, soap, etc.) and the genuine contact area varies in proportion to load N:

$$S_k = Np_0^{/}, \tag{9}$$

where:

$p_0^{/}$ – the specific pressure at which plastic deformation of the body (metal) takes place, the fluidity limit.

In this case the second term, like the first one in the binomial law of friction (7), becomes proportional to load N, which explains the applicability of Amontons' Law. Then the binomial formula of friction has the appearance:

$$F = f_0 (1 + p_0 \, p_0^{/-1}), \tag{10}$$

i.e. the force of friction will be proportional to the external load.

By changing the genuine contact area depending on the load different regularities the force of friction for the variable can be obtained (the sliding resistance, to be more precise) when the rubbing surfaces are loaded and unloaded. When friction is measured at loads gradually increasing from zero, then the force of friction will increase from zero as well, in proportion to load (Fig. 2). When the force of friction is measured at loads gradually decreasing, then it will also decrease in proportion to load, yet much more slowly than it increased as loads were increased (Fig. 2). This phenomenon can be explained by the fact that the genuine contact area increased in proportion to load when the load was increased but in case the load was decreased it remained almost constant and at maximum load it was equal to the genuine contact area.

In Bowden's theory [Kragelski *et al.* 1956, Semenov 1961], which is widely applied abroad, it is indicated that due to the roughness of surfaces, the genuine contact area $S_к$ constitutes a very small part of the contour area of contact. Therefore Bowden considers that the contact is always of plastic character, and the tension on it is equal to hardness H of the material. He assumes that the equality is as follows:

$$S_к = N H^{-1} \tag{11}$$

This leads to a conclusion that friction coefficient depends neither on the area, nor the roughness of the rubbing surfaces, nor pressure.

Bowden's theory is not supported by Russian scientists [Kragelski 1956, 1968, Deryagin 1963], who regard it correct and only due to the plastic contact when contacting takes place at constant tension (see Formula 10). For elastic materials, such as steel, the genuine contact area varies not in proportion to load but according to a more complicated dependency.

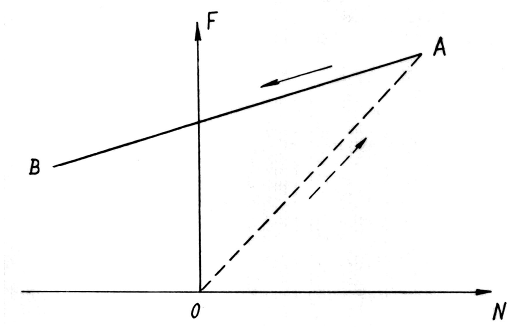


Fig. 2. Regularity of the variations of the force of friction at a plastic contact (according to Deryagin [1963]): OA – variation of the force of friction at an increasing load; AB – the same at a decreasing load

Deryagin discusses the mechanism of static and kinetic friction, as well as the instances of the fluid, dry and borderline friction (with borderline lubrication). He points out that external friction is the most surface friction among all the surface phenomena of two bodies, and it is less dependent on the peculiarities of the material of the rubbing bodies than on the condition of the rubbing surfaces. Therefore, external friction is extraordinarily sensitive to the condition of the surface. Tiny dirty stains, which little affect the other surface phenomena, can change the value of the friction coefficient several times. However, the surfaces of bodies are not, as a rule, absolutely clean in nature. They are covered with an adsorption monomolecular layer of gases and organic compounds, oxides, etc. affecting the lubricant effect on the rubbing surfaces (Fig. 3 and 4). Particularly strong influence upon the decrease in the friction coefficient is exerted by the monolayers of fatty acid. Moreover, the greater is the number the carbon atoms within the molecule, the greater is its length and the more rapid is the decrease of the friction coefficient. For example, if the coefficient of friction between clean steel surfaces is equal to 0.7–0.8, its value diminishes to 0.2 when they are covered with an adsorption monomolecular layer of valeric or acetic acid, to 0.1 when they are covered with an adsorption monomolecular layer of stearic acid, and even more with the fatty acids of a still greater molecular weight [Deryagin 1963]. When the molecular weight increases, the adhesion of surface-active substances (adhesives) increases but it decreases when temperature is raised.

When borderline lubrication is present, the atmospheric pressure does not affect the force of friction [Deryagin1963].

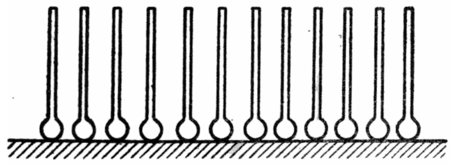


Fig. 3. A molecular pile (an oriented monolayer of molecules) on the surface of a body (according to Deryagin [1963])

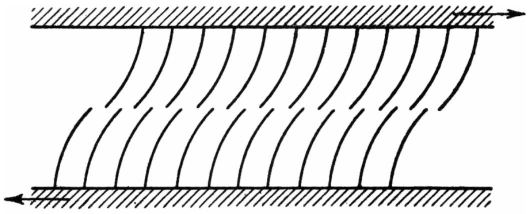


Fig. 4. A model of the lubricant action of the monolayers of organic molecules [Deryagin1963]

Deryagin and Kragelski note that Amontons' Law can be applied to surface friction at the presence of adsorption layers with still greater exactness than when they are absent [Krahielskiy and Shchedrov 1956, Krahielskiy 1960, Krotova 1960]. Borderline lubrication on even surfaces ensures their smooth sliding. If the boundary film is interrupted because of the roughness of the surfaces or insufficient strength of the film the smoothness of sliding is violated, friction takes place at a mixed contact, and the friction coefficient passes through its maximum value when the load increases. This is the reason for a stepwise change of the friction resistance. A reason of jumps in the sliding friction may be alternation of the stages of sliding and standstill in the case of elastic engagement of one of the sliding bodies. It is the difference in the values of the friction coefficients at standstill and in motion that causes these jumps. A reason for such jumps may also be a change of the force of friction as a result of electrostatic attraction or relaxation oscillations [Kragelski and Shchedrov 1956, Deryagin 1963].

In a number of cases quite a strong impact on friction is exerted by such a surface phenomenon as adhesion. Adhesion is determined by the forces resisting to the separation of two solid bodies being in a mutual contact. Deryagin and Krotova [Krotova 1960, Deryagin 1963] point out that the respective forces of adhesion formally differ from the forces of friction only by the fact that they represent resistance to mutual movement of two contacting bodies in a direction which is normal to the contact surface whereas the forces of friction exert resistance to the tangential movement, i.e. sliding. Because of this, they consider it possible and expedient to make a classification of the adhesion phenomena, similar to the phenomena of friction. There are distinguished: static adhesion – resistance to the start of the tear-off, kinetic adhesion – resistance to the tear-off at various stages of the process of increasing the gap between the two bodies, resistance which may depend on the speed of this process. In addition to it, as with the variety of the friction modes depending on the thickness and nature of the lubricant layer, there are: dry adhesion, liquid adhesion and boundary adhesion. At the same time essential difference is underlined between the phenomena of kinetic friction and adhesion. In the first instance a stationary state is possible corresponding to the movement with a constant gap. In the second instance the stationary state is impossible when the two solid bodies are separated; only a quasi-stationary tear-off process is possible when the flexible films lift off.

Despite the harmony of Deryagin's molecular-kinetic theory of friction it still cannot explain many things in the processes of friction. It is regarded as being more correct when the processes of friction are treated at comparatively small pressures of the bodies with relatively smooth surfaces of sliding and the bodies which are subject to plastic deformations [Kragelski *et al.* 1956].

The molecular mechanic theory of friction

On the basis of his theoretical and experimental works aimed at the study of the value of the friction coefficient depending on various factors Kragelski proposed a theory of elastic-viscous contact, or the so-called molecular mechanic (adhesion-deformation) theory of friction [Kragelski, Shchedrov 1956, Krotova 1960]. He arrived at a conclusion that the presence of roughness on the rubbing surfaces creates a non-uniformly tense state of the friction surface, which leads to the appearance of local elastic and plastic deformations during the process of friction (see Fig. 5).

Kragelski explains the phenomena of friction by interaction (connection) at the place of contact of the two bodies of elastic-viscous character. During the mutual movement of the bodies simultaneous destruction and formation of this connection takes place. The force of friction is the force wasted in order to destruct this connection but the coefficient of friction is the relation of the force of friction to normal pressure.

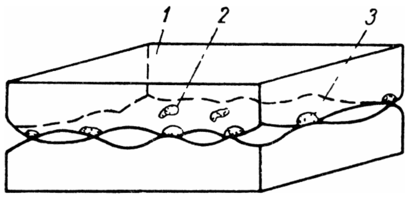


Fig. 5. The scheme of two rough contacting surfaces [Kragelski 1968].
Areas: 1 – nominal, 2 – contour, 3 – genuine

In studying the phenomena of friction Kragelski divides the surface of the mutually rubbing bodies into the areas of three types: 1) the nominal (geometrical) contact area, 2) the contour contact area, and 3) the genuine (physical) contact area (Fig. 5). The nominal contact area is the plane locus of all the possible genuine contact areas, and it is confined by the dimensions of the contacting bodies. The contour contact area is an area created by the voluminal contortion of the bodies determined by the undulation of their surfaces. The value of the contour area depends both on the geometrical outline of the surfaces and load. The genuine contact points are located on the contour contact area, whose summary area is the function of the geometrical outlines of individual surface irregularities and their respective loads. An essential characteristic of the genuine area is the density of the contact, which is the number of spots per 1 cm^2 of the area of an immobile contact.

Under the impact of the compressing load upon the surface of the bodies an increasing number of points approach and, as they approach, they come into contact with each other. Besides, the interacting elements are elastically deformed at the beginning, after that, as pressure increases, the elastic deformation passes into the plastic one. Due to the discreet contact individual contact spots appear during the sliding of the surfaces. High temperature arises on the contact spots and the mechanical properties of the rubbing materials change correspondingly.

The interaction of the rubbing surfaces manifests itself in the formation of contact spots. Not only projections with the films covering them take part in this process, but also the material touching closely these projections, which is deformed correspondingly.

The contact spots, arising as a result of mutual sliding of two bodies who exist and disappear under joint impact of the normal and tangential forces, are called frictional connections.

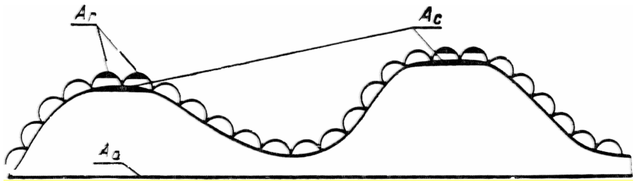


Fig. 6. The model of an undulated rough surface (according to Kragelski [1960])

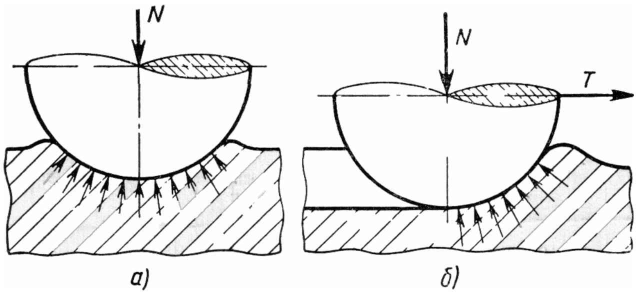
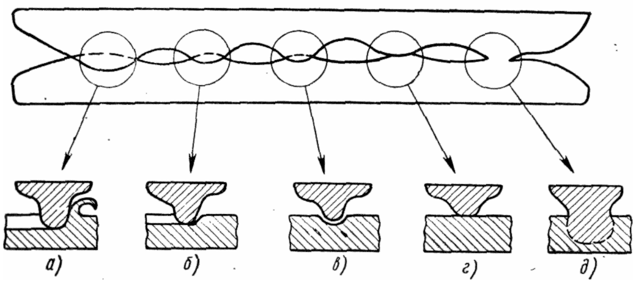


Fig. 7. An interaction scheme of a rigid spherical surface irregularity with a plastic deformed semispace [Kragelski 1960]): a – in a immobile state, b – when displaced



a b c d e

Fig. 8. Five types of violations of the functional links (according to Kragelski [1960]):

a – a section (chip, scratch, comb) of the material, b – plastic pushback of the material, c – elastic deformation of the material, d – surface destruction of the weak molecular bridge (overcoming adhesion), e – deep destruction of the strong molecular bridge (destruction of the material)

When discussing the general picture of the friction of solid bodies, Kragelski points out that „an essential thing for the understanding of the friction process is the circumstance that, due to the roughness and undulation of surfaces (Fig. 6), contacting is always discreet, i.e. it takes place at individual spots. Interaction of the surfaces at these spots has dual molecular mechanic nature. The mechanic interaction is determined by reciprocal penetration of separate contact points (Fig. 7). One surface penetrates into the other not only at the expense of the existing roughness but also the mechanical properties arising under load due to anisotropy. Even entirely smooth surfaces may turn out to be rough under the impact of the compressive force. The molecular interaction is determined by mutual attraction of the surfaces of two bodies. Since it varies in inverse proportion to the 4th degree of distance, then the molecular attraction is practically very weak or such that it leads to the appearance of a strong connection which may be destroyed only in the newly-developed volume (in the case of the external friction in the intermediate layer between the solid bodies).

The configuration of the penetrating elements and the depth of penetration are different for different contact points. The correlation between the forces of adhesion and cohesion is different, too. In this connection there are distinguished five types of violation of the functional links (Fig. 8): 1) chip or section of the material, 2) plastic pushback of the material, 3) elastic deformation of the material, 4) overcoming adhesion, 5) destruction of the material. In a general case they may be all the five types but in individual particular cases only some of these types predominate.

In order to ensure external friction without heavy destruction of the rubbing bodies, it is purposeful to have a weaker intermediate layer between the contacting surfaces."

Kragelski points out that damage always arises to the contacting surfaces during the tangential displacement as a result of their mutual penetration – tear-off of the finest particles, scratches. Friction without scratches and without surface destruction is almost impossible.

CONCLUSIONS

1. It follows from the discussed theories of friction that most completely the essence of friction, more precisely, resistance to sliding of one body along the other, can be expressed by binomial formulae (Deryagin's, Kragelski's formulae) in which the value of one term depends on the roughness of surfaces but of the other – on the reciprocal molecular attraction of the sliding materials. The monomial formula of friction (Amontons' Law), applied basically in technical calculations in our days, too, is a particular case of the general regularity. It is true for rigid materials and high loads.
2. Soil is a material which can be comparatively easily deformed; under certain conditions it is plastic with pronounced adhesion. In order to reveal and explain the character of the sliding resistance of the material in soil (also of soil sliding along the operating tools of the soil tillage machines) it is expedient to use Deryagin's binomial formula as the most appropriate.

REFERENCES

Deryagin B.V. 1963: Shto takoye trienye. Mashinostroyeniye, p. 230.

Doner R. D., Nichols M. L., 1934: The Dynamic Properties of Soil. V. Dynamics of Soil on Plow Moldboard Surfaces Related to Scouring. – Agricult. Eng., vol. 15, 9–12.

Kragelski I.V., Shchedrov V.S. 1956: Pasvitye nauki o trinii. Isd. AN SSSR, p.235.

Kragelski I.V. 1960: Trieniye i isnos. Mashinostroyeniye, p.480.

Krotova N. A. 1960: O skleivanii i prilipanii. Isd. AN SSSR, p. 168.

Semienov A. N., Chaykovskiy B.I. 1961: Is istorii ucheniya o teorii tvierdykh tiel. Trudy Kishinievskoho Instytuta, t. XXIII, 249–268.

Vilde A. 1976: Tyahovoye soprotivleniye klina pri podyomie pochviennoho plasta. Trudy Latv. NIIMESHa, t. 1., 262–313.

Vilde A., Rucins A. 2004: The impact of soil physical and mechanical properties on draft resistance of ploughs. Polish Academy of Sciences Branch in Lublin. TEKA Comm. Mot. Power Ind. Agricult. Polish Academy of Sciences Branch in Lublin, IV, 243–248.