Doklady Physics, Vol. 45, No. 5, 2000, pp. 187–190. Translated from Doklady Akademii Nauk, Vol. 372, No. 1, 2000, pp. 36–39.

PHYSICS

Separation of Random Signals

Yu. V. Andreev, A. S. Dmitriev, E. V. Efremova, and Corresponding Member of the RAS V. I. Pustovoĭt

Received August 31, 1999

The discovery of the phenomena of random synchronization [1, 2] and random synchronous response [3] has called a widespread interest to applications of random signals in communication systems. This interest is caused, first of all, by the possibility of the implementation of receivers (guided dynamic systems) self-synchronized with transmitters (guiding random systems). However, these are not the only potential applications of the random synchronization and the random synchronous response in communication systems. Based on these phenomena, it is possible, in principle, to decompose a sum of random signals into initial random components [4–6]. Such an approach suggests new concepts of signal multiplexing and demultiplexing in the communication systems.

Conventional methods of signal multiplexing are the frequency separation, time separation, and code separation. The separation based on the random synchronization is not similar to any of these methods. The principal idea of such a method is the following. We consider m pairs of random receivers and transmitters. For transmitting random signals $x_j(k)$, $j = 1, 2, \ldots, m$ from the receivers to the transmitters, we use a single communication channel in which signals are summed up. In general, the sum of random signals is supplemented by noise $\eta(k)$.

At the receiving terminal, each receiver separates out its specific signal from the total sum implementing the dynamic effect of random synchronization of processes in the transmitters and the receiver.

We now consider this problem for the specific case of two receiver–transmitter pairs. The dynamics of transmitters is assumed to be governed by one-dimensional mappings f of the same type, but with different values of parameters μ_1 and μ_2

$$\begin{aligned} x_1(k+1) &= f_1(x_1(k), \mu_1), \\ x_2(k+1) &= f_2(x_2(k), \mu_2). \end{aligned} \tag{1}$$

Institute of Radio Engineering and Electronics, Russian Academy of Sciences, ul. Mokhovaya 11, Moscow, 102907 Russia

The signal in the channel has the form

$$u(k) = x_1(k) + x_2(k) + \eta(k). \tag{2}$$

In [4], the following scheme of signal separation was suggested. The receivers from each pair are mutually coupled, and their dynamics is described by the following equations

$$\begin{aligned} y_1(k+1) &= f_1(y_1(k)) \\ &+ \alpha[u(k) - f_1(y_1(k)) - f_2(y_2(k))], \\ y_2(k+1) &= f_2(y_2(k)) \\ &+ \alpha[u(k) - f_1(y_1(k)) - f_2(y_2(k))], \end{aligned} \tag{3}$$

where α is the coupling constant. In the case of synchronization, the terms in square brackets become negligible [at the noise level of $\eta(k)$], and we have $y_1(k) \approx x_1(k)$ and $y_2(k) \approx x_2(k)$, i.e., the "proper" signals are separated out by the receiver. The analysis of the signal-separation scheme demonstrates that (i) the separation is possible in the case of $\mu_1 \neq \mu_2$, and (ii) the scheme is efficient only for very low noise level.

In paper [7], information aspects of the random synchronization were analyzed. The random synchronization of two systems was defined there as the situation when the output signals of the systems coincide or asymptotically tend to each other. In this case, it is not necessary for the signals of one system to directly affect another system. Instead, we can only transmit to one system a minimum amount of information concerning the state of another system. It was also shown that the implementation of such a random synchronization provides an opportunity to improve significantly the stability of synchronization with respect to noise.

In this paper, we introduce a new separation scheme for random signals. The scheme is based on the aforementioned type of random synchronization. We show that this scheme is more stable to the noise than the scheme proposed in [4].

For clarity, both sources of the randomness are described below by the logistic-parabola map

$$\begin{aligned} x_1(k+1) &= \mu_1 x_1(k)(1 - x_1(k)), \\ x_2(k+1) &= \mu_2 x_2(k)(1 - x_2(k)). \end{aligned} \tag{4}$$

1028-3358/00/4505-0187$20.00

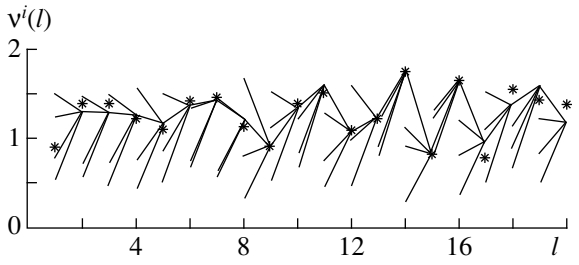


Fig. 1. Branching of the solution while iterating inverse maps. l is the iteration number, $\nu^i(l) = y_1(l) + y_2(l)$, $i = 1, ..., 4$.

The essence of the proposed scheme of signal separation is the following. The receiver, being the separator of random signals, involves copies of the maps generating random signals of the transmitters. The signals $u(k)$ are fed to the input of the receiver. Assume that at a time moment k, the receiver contains not only an estimate of the sum of random signals in the form $u(k)$, but also separate estimates for magnitudes of random signals $X_1(k)$ and $X_2(k)$. We iterate each of the maps

$$\begin{aligned} y_1(k+1) &= \mu_1 y_1(k)(1 - y_1(k)), \\ y_2(k+1) &= \mu_2 y_2(k)(1 - y_2(k)) \end{aligned} \tag{5}$$

by one backward step with the initial conditions $X_1(k)$ and $X_2(k)$, respectively. This operation is equivalent to iterating the maps f^{-1} (inverse with respect to the maps f) by one step forward. Some properties of these maps were discussed in [8]. Maps (5) iterated forward are stretching on the average, hence, they are compressing at the backward iteration. Therefore, the estimates for X_1 and X_2 corresponding to the $(k-1)$th time moment and obtained from the estimates at the kth time point have a higher accuracy on the average than the latter estimates themselves. However, it should be taken into account that maps (5) are two-valued in the case of the backward iteration in time. Iterating the maps for one step, we obtain two values rather than one for each of the map. Hence, it is necessary to choose the "correct" branch after the iteration. The net signal at the kth time point and the estimates for each individual signal do not contain the information necessary for the correct choice of branches at the $(k-1)$th step and cannot be used for this purpose. However, at $(k-1)$th time point, the signal $u(k-1)$ being the estimate of the sum of two random signals at this moment of time is fed to the receiver. Signal $u(k-1)$ can be used for the correct choice of the branch. Actually, we have the following combination of branches, which minimizes the root-mean-square deviation of the estimates for the sum of two random signals from the net signal fed to the receiver (under condition of low noise)

$$\delta = \left\| u(k-1) - y_1^j(k-1) - y_2^i(k-1) \right\|. \tag{6}$$

Here, $y_1^j(k-1)$ and $y_2^i(k-1)$ are values of the variables y_1 and y_2 for various branches. This relationship allows us to choose the correct branch. The procedure providing the choice of the "correct" combination of branches is illustrated in Fig. 1. Asterisks denote magnitudes of the signal $u(l)$ corresponding to $l < k$. The figure demonstrates that at the lth step, we choose only one from four possible values, which is the closest to $u(l)$. In turn, the chosen value gives rise to four branches at the $(l-1)$th step. Repeating this procedure provides an opportunity to obtain the separation of signals at an arbitrary time point $l < k$.

We have discussed the scheme of signal separation under the condition that the estimates of signals $X_1(k)$ and $X_2(k)$ are known. However, in the general case, the separate estimate for the state of guiding systems at the time instant k is absent. The calculations show that, as such an estimate, we can take a pair of arbitrary points $y_1(k)$ and $y_2(k)$ belonging to the attractors of maps (5). Starting from these initial conditions, we obtain the calculated trajectories converging with time to the transmitter trajectories. However, the time of convergence strongly depends on the choice of particular conditions specified for the attractor. To ensure the acceleration of the convergence process and improve the initial estimates for the signals $X_1(k)$ and $X_2(k)$, we proceed in the following manner. We specify a series of initial values for the variable $y_1(k)$ at the time moment k: $y_1^j(k-1)$, $j = 1, 2, \ldots, m$. Each initial value must belong to the attractor, and the whole series should be arranged more or less uniformly on it. The similar series of initial values is specified for the variable $y_2(k)$. Let ε be an admissible accuracy of the estimate for the variables $y_1(k)$ and $y_2(k)$ being separated. From these series of initial values, we form pairs $(y_1^j(k), y_2^i(k))$, $i, j = 1, 2, \ldots, m$, and choose those meeting the condition

$$\left\| u(k) - y_1^j(k) - y_2^i(k) \right\| < \varepsilon. \tag{7}$$

We iterate each of chosen pairs by one step backward. As a result, we obtain the doubled number of pairs. Among them, we choose a pair minimizing expression (6). Then, we repeat the procedure described above.

The efficiency of the proposed separation scheme for random signals was tested using the guiding systems described by logistic maps with $\mu_1 = 3.7$ and $\mu_2 = 3.8$. The signal-to-noise ratio for each of the isolated signals was calculated as a function of the signal-to-noise ratio for the signal $u(k)$. As a criterion of the efficiency for separation of signals, we can consider the relationship between the noise levels in the isolated signals and in the signal $u(k)$. The range of noise levels in the channel, in which the noise level of the isolated signal turns out to be below than the noise level of a signal in the channel, can be referred to as a range of efficient separation of signals. The separation is inefficient if the noise level in the isolated signal far exceeds the noise level in the channel. The results of calculations are

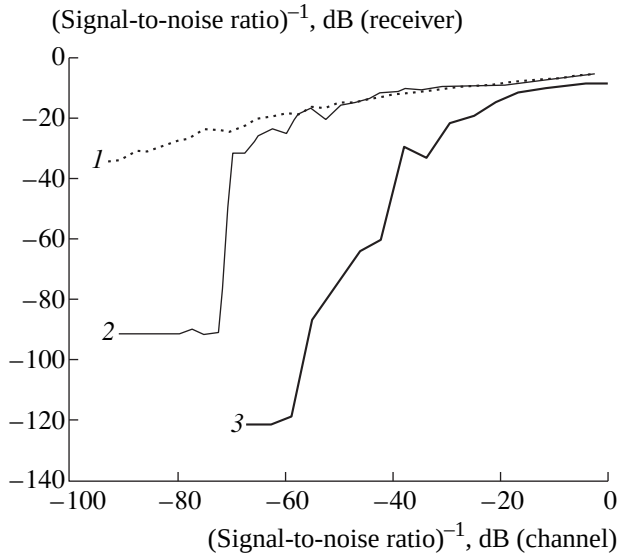


Fig. 2. Signal-to-noise ratio at the receiver output as a function of the signal-to-noise ratio in the channel: (*1*) method reported in [4]; (*2*) algorithm with the single sequence; (*3*) algorithm with M sequences ($M = 16$).

illustrated by Fig. 2 (curve *2*). In the process of signal separation, failures can sporadically arise (Fig. 3). The failures lead to a pronounced scatter in the estimates based on the criterion of the signal-to-noise ratio. Therefore, we used additionally the second criterion for the signal-separation efficiency, namely, relative time τ needed for the efficient separation of signals (Fig. 4, curve *2*). It is clear that the separation is efficient only if τ is close to unity. The results presented in Figs. 2 and 4 show that efficient signal separation is possible for signal-to-noise ratio exceeding 70 dB. For comparison, in Figs. 2 and 4, we present the corresponding results (curve *1*) for the scheme discussed in [4]. From the comparison of the curves, it follows that the proposed scheme for separating random signals remains efficient in the case of the noise level exceeding by 20–30 dB the corresponding level for the scheme proposed in [4].

Analysis of situations reconstructed by the above method shows that its efficiency decreases with increasing external noise due to failures in the process of separation. The failures stem by the incorrect choice of branches for the inverse maps, which is caused by the effect of external noise. This leads to an idea to improve the algorithm by monitoring several other branches in addition to the "optimum" one and to choose between them by averaging the desynchronization signal over long time interval.

This idea was realized in the form of the following algorithm. In contrast to the original version of the algorithm, we choose for the $(k-1)$th time point M pairs of estimates (branches) instead of one. At all subsequent iteration steps for the f^{-1} map, each branch splits into four prototypes, see Fig. 1. The number of branches increases by the factor of four. Some of these branches nearly coinciding with each other can form "bundles." In the course of further iterations, the branches belonging to the same bundle approach one another exponentially and form essentially the same solution. Therefore, we take only one branch from each bundle. After that, we choose M branches among the remaining ones. The chosen branches correspond to the minimum values of the root-mean-square deviation of the estimate for the sum of two random signals from the net signal fed to the receiver. Moving in such a manner from the final element of sequence $u(k)$ to its beginning, we obtain M versions for the pairs of sequences (y_1^m, y_2^m), $m = 1, 2, \ldots, M$. Then, we apply the criterion of the minimum distance squared between the sum of reconstructed sequences and the observed sequence $u(i)$

$$\min_{m = 1, 2, \ldots, M} \sum_{i=1}^{k} \left(u(i) - y_1^m(i) - y_2^m(i)\right)^2. \qquad (8)$$

According to this criterion, we chose the best pair (y_1^j, y_2^j), which is assumed to be the actual pair of sequences obtained as a result of demultiplexing.

The results of numerical simulations of the random signal separation based on the algorithm with a nonunique sequence are presented in Figs. 2 and 4 (curves *3*). These results show that such an algorithm allows us to

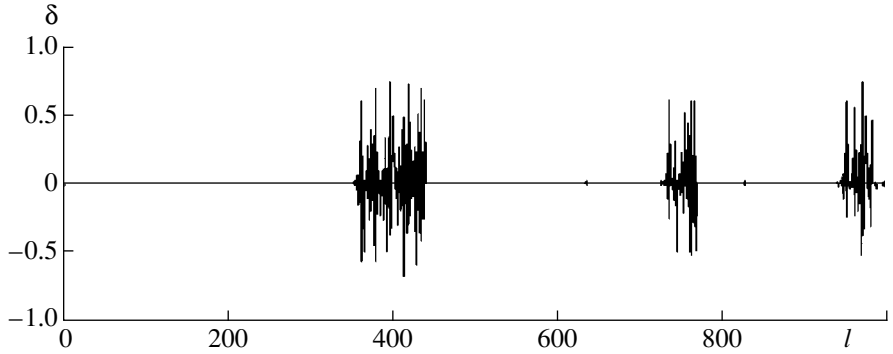


Fig. 3. Failures in the process of signal separation (δ is the difference between the transmitted signal and the signal at the receiver output).

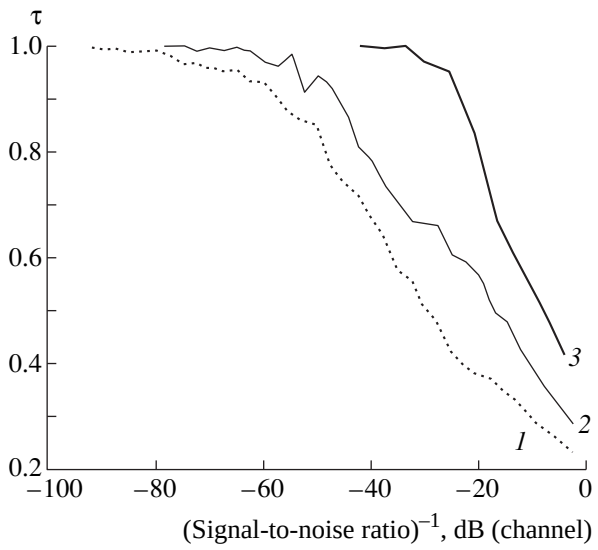


Fig. 4. Relative mean time of synchronization τ as a function of the signal-to-noise ratio in the channel.

separate random signals with the noise level approximately by 20 dB higher than those processed using the algorithm with the single sequence.

Thus, we have proposed and analyzed an algorithm for separation of the sum of random signals, which is based on their dynamic properties. The algorithm is efficient not only in the absence of noise, but at nonzero noise as well. The results presented in this and previous papers suggest the possibility and viability of a fundamentally new approach to multiplexing the signals in communication systems. The approach is based on the dynamic properties of systems with chaos inherent in them.

ACKNOWLEDGMENTS

This work was supported in part by the Russian Foundation for Basic Research, project nos. 97-01-00800 and 99-02-18315.

REFERENCES

1. H. Fujisaka and T. Yamada, Progr. Theor. Phys. **69**, 32 (1983).
2. V. S. Afraĭmovich, N. N. Verichev, and M. I. Rabinovich, Izv. Vyssh. Uchebn. Zaved., Ser. Radiofiz. **29**, 1050 (1986).
3. L. M. Pecora and T. L. Caroll, Phys. Rev. Lett. **64**, 821 (1990).
4. L. S. Tsimring and M. M. Sushchik, Phys. Lett. A **213**, 155 (1996).
5. A. S. Dmitriev and S. O. Starkov, Usp. Sovrem. Radioélektron. (Zarubezh. Radioélektron.), No. 11, 4 (1998).
6. Yu. G. Tratas, Usp. Sovrem. Radioélektron. (Zarubezh. Radioélektron.), No. 11, 57 (1998).
7. A. S. Dmitriev, Izv. Vyssh. Uchebn. Zaved., Ser. Radiofiz. **41**, 1497 (1998).
8. V. I. Pustovoĭt, Élektromag. Volny Élektron. Sistemy **2**, 12 (1997).

Translated by V. Bukhanov

Doklady Physics, Vol. 45, No. 5, 2000, pp. 191–193. Translated from Doklady Akademii Nauk, Vol. 372, No. 2, 2000, pp. 174–176.

PHYSICS

Octahedral Cr^{3+} Sites in $LiGaO_2$

A. V. Gaĭster*, E. V. Zharikov, G. M. Kuz'micheva***, Yu. M. Papin**, V. B. Rybakov****, and V. A. Smirnov****

Presented by Academician V.V. Osiko May 25, 1999

Received June 15, 1999

In laser materials activated by transition metal ions, activator ions often possess different formal charges. The relative content of each type of ion in a crystal depends on the crystal-growth conditions, in particular, on the oxidizing potential of the growth ambient. The ionic radius of cations and specific features of their electron structure, which correspond to the given value of the formal charge, determine the specific crystallographic positions (one ore more) occupied by the activators under study in the structure typical of the given chemical composition and morphology.

The above considerations are undoubtedly applicable to chromium ions, which can have various values of the formal charge ranging from +2 to +6. According to the present-day knowledge, Cr^{4+} and Cr^{6+} ions occupy, as a rule, tetrahedral sites in crystal structures, whereas Cr^{2+} and Cr^{3+} ions occupy octahedral sites. Thus, two ions, Cr^{3+} and Cr^{4+}, the most important for the laser physics, occupy crystallographic positions with different coordination numbers. Therefore, in the case of crystals activated by Cr^{3+} and Cr^{4+} ions, octahedral and tetrahedral positions should be provided, respectively. If a structure involves simultaneously positions of two or more types, having both the octahedral and tetrahedral coordination with respect to cations, chromium enters into the crystalline structure predominantly in the form of Cr^{3+} ions occupying the octahedral sites [1]. The content of Cr^{4+} in such structures is usually very low (as a rule, it does not exceed few percent of the total chromium content in the crystal [2]).

The increase in Cr^{4+} content of crystals can be attained for such compounds in which the host crystal contains only tetrahedral sites. In this case, the concentration of Cr^{4+} ions occupying the tetrahedral sites depends on the formal charge of the substituted cation: the Cr^{4+} content for the isovalent substitution is usually higher than for the heterovalent one. However, in practice, it is difficult to ensure a significant increase in the Cr^{4+} content of crystals even when all aforementioned conditions are met. This is illustrated by the results presented below.

We have studied lithium-gallate crystals doped with chromium, which were grown with the help of the Kristall-2 facility by the Czochralski method in iridium crucibles at a rate of 3 mm/h in a weakly oxidizing ambient (N_2 + 2 vol % of gas mixture). The grown crystals were studied by methods of X-ray diffraction analysis and fluorescence spectroscopy.

The structure of $LiGaO_2$ is a derivative of the wurtzite structure. Within this structure, Li^+ and Ga^+ ions are ordered occupying tetrahedral sites. In addition, the $LiGaO_2$ structure involves a large number of octahedral voids usually unoccupied [3]. Cr^{4+} ions occupy tetrahedral positions substituting Ga^{3+} ions. The concentration

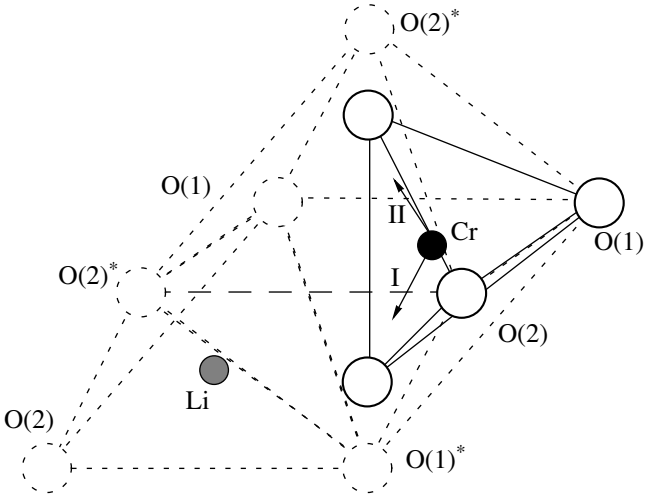


Fig. 1. Scheme for the transformation of GaO_4 tetrahedrons to octahedrons in $LiGaO_2$ crystals.

* *Mendeleev University of Chemical Technology, Miusskaya pl. 9, Moscow, 125190 Russia*

** *Institute of General Physics, Russian Academy of Sciences, ul. Vavilova 38, Moscow, 117942 Russia*

*** *Lomonosov State Academy of Fine Chemical Technology, pr. Vernadskogo 86, Moscow, 117571 Russia*

**** *Moscow State University, Vorob'evy gory, Moscow, 117234 Russia*

1028-3358/00/4505-0191$20.00 © 2000 MAIK "Nauka/Interperiodica"

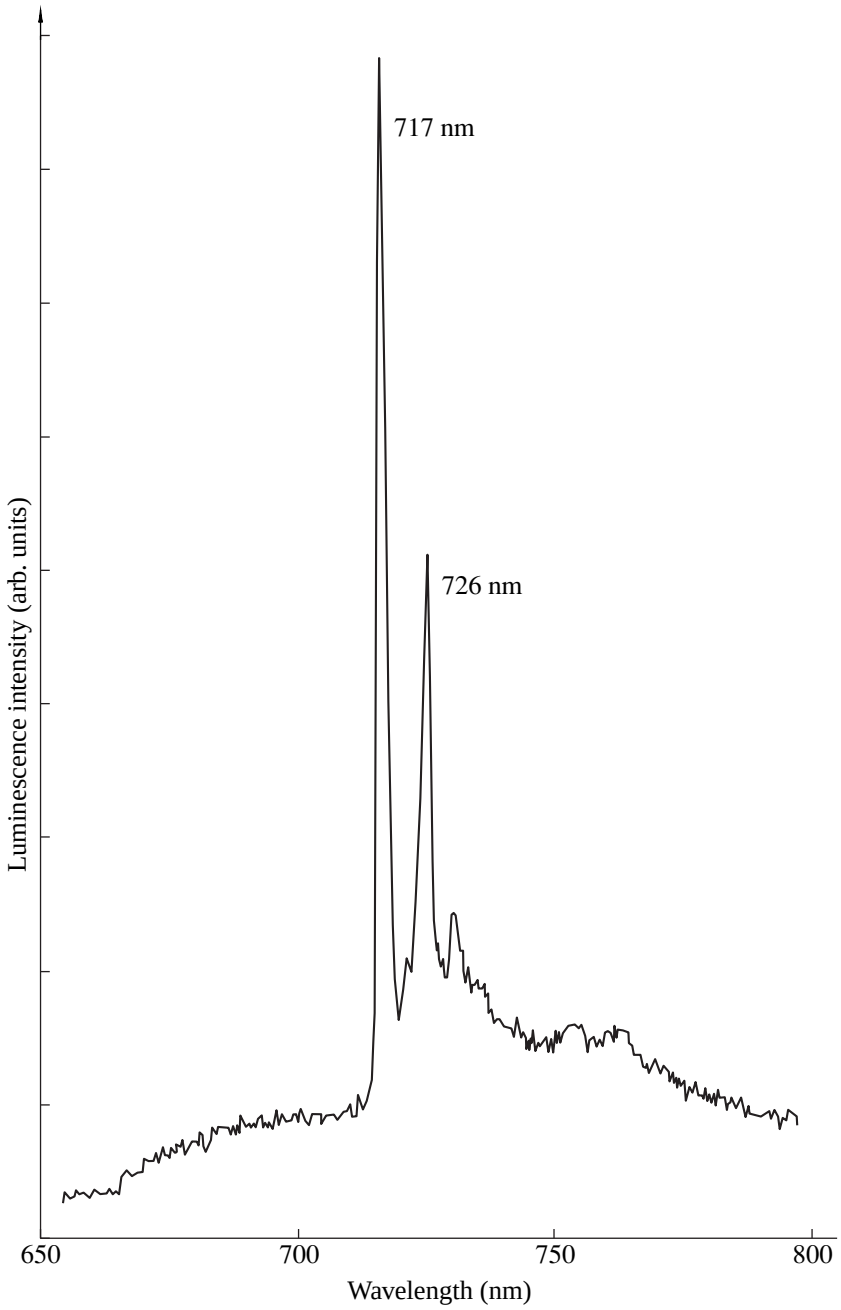


Fig. 2. Luminescence spectrum for Cr^{3+} ions in $LiGaO_2$ at 77 K.

of chromium ions grows with the concentration of vacancies in lithium positions (V'_{Li}) [4].

At the same time, the specific features of electron structure determine for Cr^{3+} ions a tendency toward the occupation of octahedral positions. This tendency is so pronounced that it seems to be quite probable for chromium ions to enter the vacant octahedral voids in the $LiGaO_2$ structure and form interstitial $Cr_i^{\cdots}$ atoms. This is possible only in the case of simultaneous formation of cation vacancies, obeying the relationships $2V'_{Li} + V'''_{Ga}$ or $2V'''_{Ga} + V'_{Li}$ (we use here the Kröger–Wink notation for point defects with maximum values of formal charges). The relationships stem from the small spacings between interstitial cations in the octahedron and adjacent cations in the tetrahedron. Such a process is the most probable in the central part of $LiGaO_2$:Cr crystals, because it is the most defect-rich region in a sample (the bulk defect), see [4].

On the other hand, it is unreasonable to exclude a possibility for rearrangement of some tetrahedrons in the structure of $LiGaO_2$ with forming octahedral groups. This phenomenon was reported previously for

melilite crystals [5]. The X-ray diffraction studies of our $LiGaO_2$:Cr samples reveal for the tetrahedrons (Ga, Cr)O_4 the decrease either in the (Ga, Cr)–O(1)* distance (from 1.821 to 1.785 Å) or (Ga, Cr)–O(2)* distance (from 1.816 to 1.742 Å) compared to the corresponding distances in the undoped $LiGaO_2$ crystal [4]. The observed changes in interatomic distances of (Ga, Cr)O_4 tetrahedrons suggests the tendency to rearrangement of (Ga, Cr)O_4 tetrahedron in $LiGaO_2$:Cr^{3+} structures to the octahedron with Cr^{3+} ions displaced toward O(1)* (Fig. 1, direction I) or O(2)* (Fig. 1, direction II).

Our studies of the luminescence spectra also present an indication for the simultaneous existence of both Cr^{4+} and Cr^{3+} ions in $LiGaO_2$:Cr samples. In the near-infrared range, we observed the broadband luminescence with $\lambda_{max} \sim 1.25$ μm. The analysis of the kinetics of the luminescence relaxation reveals the nonexponential nature of the decay for excited states of chromium ions. The lifetime of the excited state increases from 9 μs at initial stages of the process to 22 μs at the final stages of the luminescence relaxation. Taking into account the spectral range of the luminescence band and characteristic microsecond lifetimes, we relate (similar to [6, 7]) the band near $\lambda_{max} \sim 1.25$ μm to the $^3T_2$–$^3A_2$ transition occurring in Cr^{4+} ions. In [7], this decay was reported to be exponential with the lifetime $\tau = 25$ μs, whereas the decay curve in [6] was presented as a superposition of two exponential ones with $\tau_1 = 4$ μs and $\tau_2 = 18$ μs. The significant difference in the data concerning the decay kinetics for the excited state of Cr^{4+} ions can be explained by uncontrolled impurities in the initial mixture prepared for the crystal growth. Another possibility might also be uncontrolled deviations from stoichiometry in the melt due to the intense evaporation of the components (both lithium oxide and gallium oxide). In the latter case, not only the structural perfection and optical quality but also the luminescence spectra of $LiGaO_2$ samples turn out to be sensitive to specific features of the crystal growth technology (such as temperature gradients in the system, total pressure of the gas ambient, the existence of gas flow in the growth zone, etc.) and can vary even within the same growth batch.

While exciting $LiGaO_2$:Cr crystals by the second harmonics of the neodymium laser ($\lambda = 532$ nm), we have observed, for the first time, the narrow luminescence bands in the vicinity of the wavelength of 720 nm. The lifetime of the excited state at the final stages of luminescence relaxation was about 1 μs at 295 K and 2.8 μs at 77 K. The characteristic features (line widths, spectral range, and the lifetime of the excited state) of the observed bands make it possible to identify them as *R* line with the accompanying vibrational structure of the optical centers based on Cr^{3+} ions (^2E–$^4A_2$ transition) in the octahedral surrounding. The typical luminescence spectrum of Cr^{3+} ions in $LiGaO_2$ crystals (77 K) is shown in Fig. 2. The luminescence decay curves of Cr^{3+} are nonexponential. This can be related to both Cr^{3+} centers of different types and extinguishing their luminescence at lattice defects. Unfortunately, the available spectral data do not provide an opportunity to put forward any justified assumptions concerning the number of Cr^{3+}-based optical centers and their nature.

Thus, $LiGaO_2$:Cr crystals grown in the weakly oxidizing ambient contain chromium ions in two valence states, namely, Cr^{4+} and Cr^{3+}. The results of spectroscopic studies imply the formation of Cr^{3+} centers with the octahedral coordination in $LiGaO_2$ structure characterized by purely tetrahedral coordination for all cations. In accordance with X-ray diffraction data, Cr^{3+} ions can be located within distorted (close to octahedral) gallium tetrahedrons or within the octahedral voids accompanied by the vacancies in the adjacent cation positions.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, projects nos. 98-03-32755 and 98-02-17340.

REFERENCES

1. E. V. Zharikov, V. V. Laptev, A. A. Maĭer, and V. V. Osiko, Izv. Akad. Nauk SSSR, Neorg. Mater. **20**, 984 (1984).
2. E. V. Zharikov, Tr. Inst. Obshch. Fiz. Akad. Nauk SSSR **26**, 50 (1990).
3. M. Marezio, Acta Crystallogr. **18**, 481 (1965).
4. G. M. Kuz'micheva, A. V. Gaĭster, V. B. Rybakov, *et al.*, Perspekt. Mater., No. 1, 48 (1999).
5. G. M. Kuz'micheva, B. V. Mukhin, V. B. Rybakov, *et al.*, Zh. Neorg. Khim. **40**, 569 (1995).
6. S. Kueck, S. Hartung, K. Petermann, and G. Huber, in *Proceedings of the Conference on Advanced Solid-State Lasers* (Opt. Soc. Am., Memphis, 1995), Vol. 24, pp. 486–490.
7. B. H. T. Chai and X. X. Zhang, in *Proceedings of the Conference on Advanced Solid-State Lasers Tech. Digest* (Opt. Soc. Am., Memphis, 1995), pp. 289–291.

Translated by K. Kugel'

Doklady Physics, Vol. 45, No. 5, 2000, pp. 194–197. Translated from Doklady Akademii Nauk, Vol. 372, No. 2, 2000, pp. 177–180.

PHYSICS

Kinetic Equations for Large Systems with Fractal Structures

Ya. L. Kobelev*, L. Ya. Kobelev*, and Corresponding Member of the RAS E. P. Romanov**

Received December 2, 1999

In recent years, a large number of studies and monographs [1, 2] have been devoted to the effect of fractal characteristics of crystallite surfaces on many properties of solids (including their strength, brittleness, etc.). Fractal structures in solids are spontaneously formed in processes of crystallization from melts. These structures are characterized by the self-similarity of heterogeneities (crystallochemical, mechanical, crystallographic ones, etc.) in various scales of consideration. A number of properties of solutions to the diffusion equation and Klimontovich kinetic equation in the fractal space [3, 4] and the self-induced-wave processes described by nonlinear equations of fractal diffusion [5] were also studied. In [6], we considered fractal dimensions of time and space, which determine potentials of physical fields as functions of coordinates and time. We proposed also to describe the dynamics of the processes proceeding in fractal manifolds by the mathematical formalism of generalized fractional derivatives. In this case, the fractal dimensions can be considered, in particular, as the order parameters (in the Ginzburg–Landau sense [7]). In [6, 8], the fractal dimension of a crystallite surface was considered (on the basis of results of [6]) as a phenomenological order parameter, and the dependence of the fractal dimensions on potentials determining the crystal-lattice structure was presented. A question arises: whether exist the fractal structures of dynamic type stemming the time (possibly, also space) memory of particles on their motion in large systems of interacting objects, which are described by kinetic equations? (We imply systems of interacting particles, electron–ion plasma, biological, chemical, and other large systems.) Whether these structures arise during the short-term formation of the self-similar (both time and space) density perturbances? What role do they play in the behavior of such systems? In particular, how do these structures affect the stability conditions for system states (e.g., in plasma)? If the fractal structures exist, for their description (i.e., for including the fractal characteristics into kinetic equations), it is necessary to consider the system of a large number of objects as a multifractal set. The properties of this set must depend on the fractal dimension corresponding to each time and space point. To allow for the effect of fractal structures arisen in a multifractal system on its properties in the kinetic equations describing the behavior of the system, it is necessary, according to [6], to introduce the following modifications:

—To replace time derivatives by generalized fractional derivatives (making it possible to describe the dynamics of a multifractal system and the existence of the time memory in this system) with the fractional index $\nu = 1 + d(\mathbf{r}, t)$, where ν is the fractional dimension (local fractal dimension) of the multifractal system under consideration;

—To write out equations determining d (a fractional addition to the system fractal dimension) using the system characteristics (interaction potentials and distribution functions).

In this case, the dependence of d on physical characteristics (for example, electric potentials of electrons and ions, while analyzing properties of the electron–ion plasma) of the system can be estimated considering d as an order parameter [6, 8] and determining $d(t)$ from the equations for order parameters and distribution functions corresponding to the kinetic equations in the statistical theory of open systems [9].

The goal of this study is to obtain the kinetic equations for large systems, which take into account the presence of fractal structures in these systems (see also [10]). In this case, the appearance of fractal structures is considered as a consequence of self-organization processes proceeding in the system. The answer to the question on the existence of fractal structures and conditions under which they arise can be obtained from the combined solving equations for the fractal dimensions considered as the order parameters and from the kinetic equations for the distribution functions of system objects. In this study, we formulated the general equations of such a type. As an example, we present the

* *Ural State University, pr. Lenina 51, Yekaterinburg, 620083 Russia*

** *Institute of Metal Physics, Ural Division, Russian Academy of Sciences, ul. S. Kovalevskoĭ 18, Yekaterinburg, 620219 Russia*

1028-3358/00/4505-0194$20.00

equations for electron–ion plasma (without taking into account the pumping fields and the magnetic self-fields). Electron–ion plasma with fractal structures is described in the framework of the Vlasov kinetic equation [11] with the Klimontovich collision integrals [12] for an electrically neutral system of interacting particles.

1. APPEARANCE OF A FRACTAL DIMENSION IN LARGE SYSTEMS

When considering a large system of objects as a multifractal set, we can choose the set R^1 (at $\nu < 1$) or R^2 (at $\nu > 1$) for the time memory and R^3 (or R^4) for the space memory as a measure carrier. Each point of these sets is characterized by a local fractal dimension $\nu_\alpha(\mathbf{r}, t)$ ($\alpha = t, \mathbf{r}, \mathbf{p}$). This consideration assumes certain trajectories to be inaccessible for particles of the system, and the dimension of the set of accessible trajectories to be fractional similar to strange attractors. From a mathematical standpoint, the presence of the time memory for the past (future) is described by introducting generalized fractal derivatives with respect to time (in terms of the equation formalism of generalized fractional derivatives) instead of integer derivatives. We consider a system consisting of a large number of identical objects of arbitrary nature (physical, chemical, biological, economical, etc.) and interacting with each other. We also leave room for describing the behavior of this system by methods of statistical physics for open systems (kinetic equations, Langevin equations, reaction-diffusion equations, etc.). We denote the probability density for localizing an object with the momentum $\mathbf{p}$ at the point $\mathbf{r}$ and at the moment of time t as $n_i(\mathbf{r}, \mathbf{p}, t)$. We write out the corresponding collision integrals as $I_{i,\,\mathrm{col}}$. Such a system is described by the nonlinear equations of statistical physics for open systems

$$L_i(\partial/\partial t, \nabla_i, I_{i,\,\mathrm{col}})n_i(\mathbf{r}, t) = 0. \qquad (1)$$

(Here, the dependence of L_i on operators is determined by the choice of a particular problem.) In this system, under certain conditions, processes of self-organization, in particular, phase transitions leading to the appearance of ordered structures of various types, which also include the fractal structures, can arise. In this study, we consider the behavior of a system affected by only fractal structures, i.e., structures characterized by the fractal dimension (determined, for example, as the box dimension or the Hausdorff dimension). These structures are assumed to be order parameters (similar to the order parameters from [7]). The appearance of fractal dimensions in a system of this type can be described by the methods used in [6]. To explain the appearance of the fractal dimension in the system, we use the notation introduced previously for the fractional addition to the local fractal dimension $d(\mathbf{r}, t)$ and the fractal dimension (at the points $\mathbf{r}, t$) of the system: $\nu = 1 + d(\mathbf{r}, t)$. In the absence of fractal structures, the dimension of the system under consideration is integer-valued. Following [6], we consider the equation for $d(\mathbf{r}, t)$ as the order parameter (determined by varying the Ginzburg–Landau-type functional of the fractal dimension for the system free energy):

$$\frac{\partial d}{\partial S_i} = \frac{\partial}{\partial P_i}\left[A(P_i)\frac{\partial d}{\partial P_i}\right] + (\varphi - \varphi_0)d + k_1 d^2. \qquad (2)$$

Here, we choose the time as the variable S_i determining the dependence of the fractal dimension on physical variables. In this case, equation (2) takes the following form (when using the integer derivatives $\partial/\partial t$ instead of the generalized fractional derivatives in the equation determining d):

$$\frac{\partial d}{\partial t} = \frac{\partial}{\partial P_i}\left[D(\mathbf{r}, t)\frac{\partial d}{\partial P_i}\right] + (\varphi - \varphi_0)d + k_1 d^2. \qquad (3)$$

According to [6], $\varphi' - \varphi'_0$ can be presented in the form of the functional for the physical variables P_i, which depends on the distribution functions

$$\varphi' - \varphi'_0 = \int d\mathbf{r}'\frac{\partial P_i(\mathbf{r}, t)}{\partial \mathbf{r}'}d\mathbf{p}\{\beta_j A(P_i, \mathbf{r} - \mathbf{r}') \times n_j(\mathbf{r}', \mathbf{p}, t) - \beta_i A(P_j, \mathbf{r} - \mathbf{r}')n_i(\mathbf{r}', \mathbf{p}, t)\}. \qquad (4)$$

In [4], $n_j(\mathbf{r}, \mathbf{p}, t)$ and $n_i(\mathbf{r}, \mathbf{p}, t)$ are the densities of the one-particle distribution functions for objects of various type (for example, electrons, ions, etc.), β'_a are the dimensional constants providing the dimensionless ($a = j, i$) fractional dimension d_i, and A are the functions determining the dependence on P_i. The choice of parameters P_i depends also on the choice of a model. For P_i, we can choose time [13], pressure [14], coordinates, temperature, etc.

2. FRACTAL DIMENSION AND FRACTIONAL DERIVATIVES

To describe the dynamics of functions given on fractal sets, it is necessary to use the mathematical formalism of fractional derivatives, because the ordinary differentiation and integration of such functions lead to the loss of essential properties characteristic of fractal systems. For systems conserving the time memory, it is convenient to use the generalized fractional derivatives (first proposed in [6]), which are the generalization of the fractional derivatives and the Riemann–Liouville integrals:

$$D^{\nu}_{+,\,t}f(t) = \left(\frac{d}{dt}\right)^{\alpha}\int_a^t dt'\frac{f(t')}{\Gamma(\alpha - \nu(t'))(t - t')^{\nu(t') - \alpha + 1}}, \qquad (5)$$

$$D^{\nu}_{-,\,t}f(t) = (-1)^{\alpha}\left(\frac{d}{dt}\right)^{\alpha}\int_t^b dt'\frac{f(t')}{\Gamma(\alpha - \nu(t'))(t - t')^{\nu(t') - n + 1}}. \qquad (6)$$

Here, $\nu = 1 + d$, Γ is the gamma-function, a and b are constants (taking values from 0 to ∞), $\alpha - 1 < \nu \le \alpha$, $\alpha = \{\nu\} + 1$, $\{\nu\}$ is the integer part of $\nu \ge 0$, and $\alpha = 0$ for $\nu < 0$. For $\nu = \text{const}$, the generalized-fractional derivatives (5), (6) coincide with the Riemann–Liouville fractional derivatives. For $\nu = n + d(t)$ and $d(t) \longrightarrow 0$, the generalized-fractional derivatives can be presented through the integer derivatives and integrals. For $\nu = 1$, generalized-fractional derivatives (5) and (6) coincide with the first derivative with respect to time. It is possible to show (see [6]) that, for the fractal dimensions differing only slightly from the integer values ($\nu \sim 1$, $|d| \ll 1$), the generalized-fractional derivative $D^{\nu}_{+,t}$ from (5) takes the form

$$D^{\nu}_{+,t} n \approx D^{\nu}_{-,t} n \approx \frac{\partial}{\partial t} n + a \frac{\partial}{\partial t} (\Gamma^{-1}(1-\nu) n d). \qquad (7)$$

In (7), a is the numerical coefficient depending on the choice of the integral regularization in the fractional derivative and the coefficients of the gamma-function expansion into series (for $\nu < 1$ and $a = 0.5$). The functions under the integrals in (5) and (6) are considered as the generalized functions given in the set of finite functions [15]. In the presence of the system fractal structures, the integer derivatives with respect to time should be replaced by integral operators (5) or (6). As was noted above, the dependence $d(\mathbf{r}, t)$ on time and coordinates is determined by the distribution of potential fields generated by objects, constituting the system, and by their distribution functions.

3. KINETIC EQUATIONS WITH FRACTAL STRUCTURES

Let a large system or several systems each composed of i ($i = 1, 2, \ldots$) subsystems of N_i objects be described by the kinetic equation for the distribution functions $f_i(t)$. In the presence of the system under description of fractal structures, in addition to the equations determining the functions $f_i(t)$, it is necessary to also write out the equations for the fractal dimensions $\nu_i(\mathbf{r}, t)$ of these systems. The set of equations describing the behavior of these objects takes the form

$$L_i(D^{d_i}_{+,t}, D^{d_i}_{-,t}, D^{d_i}_{+,\mathbf{r}}, D^{d_i}_{-,\mathbf{r}}, \ldots, d_i, d_j, I_{i,\text{col}}) f_i(\mathbf{r}, t) = 0, \qquad (8)$$

$$\tilde{L}_i(D^{d_i}_{+,t}, \ldots, D^{d_i}_{+,\mathbf{r}}, \ldots, f_i, f_j) d_i(\mathbf{r}, t) = 0. \qquad (9)$$

In (8) and (9), L_i and $\tilde{L}_i$ are the operator functions determining a particular form of kinetic equations (depending on the chosen statistical model for description of the system of the objects under consideration) and equations for the fractal dimensions (obtained according to the rules of the previous section). The quantities $I_{i,\text{col}}$ are the collision integrals for the corresponding systems and $D^{d_i}_{+,\alpha}$ and $D^{d_i}_{-,\alpha}$ are the generalized fractional derivatives replacing the ordinary derivatives with respect to time, coordinates, and momenta. Equations (8) and (9) describe the behavior of large systems with allowance for appearing in them the fractal structures. In the case of the absence of fractal structures, equations (9) yield zero solutions for the fractional addition to the dimensions of time and space. In this case, the generalized fractional derivatives coincide with ordinary derivatives, and equation (8) coincide with the known kinetic equations (including the Langevin equations, equations of the reaction-diffusion type, etc.).

4. KINETIC EQUATIONS FOR PLASMA WITH FRACTAL STRUCTURES

We consider, as an example, the kinetic equations for the Coulomb plasma with allowance for fractal structures. We introduce the following notation: let $a = e, i$ correspond to the electron and ion plasma components, e_a and m_a are masses of the particles, N_a is the total number of particles of the type a,

$$n_a = \frac{N_a}{V} = \int f(\mathbf{r}, \mathbf{p}, t) d\mathbf{r} d\mathbf{p},$$

V is the volume of the system, and $\sum_a e_a N_a = 0$. The kinetic equations for the function f_a of such a system (the Vlasov equations with the Klimontovich dissipative collision integrals) with allowance for the fractal structures have the following form (for simplification in (8) and (9), the generalized fractional derivatives are conserved only to take into account the fractal structures with the time memory):

$$D^{1+d_a(\mathbf{r},t)}_{+,t} f_a + \mathbf{v}\frac{\partial f_a}{\partial \mathbf{r}} + e_a \mathbf{E}(\mathbf{r}, t)\frac{\partial f_a}{\partial \mathbf{p}} = I^{(\nu)}_a(\mathbf{r}, \mathbf{p}, t) + I^{(\mathbf{r})}_a(\mathbf{r}, \mathbf{p}, t), \qquad (10)$$

$$\frac{\partial d_a}{\partial t} = \frac{\partial}{\partial P_i}\left[D_0(\mathbf{r}, t)\frac{\partial d_a}{\partial P_i}\right] + (\varphi_a - \varphi_{a0}) d_a + k_1 d_a^2, \qquad (11)$$

where $I^{(\nu)}_a$ is the Landau dissipation, $I^{(\mathbf{r})}_a$ is the Klimontovich collision integral describing spatial diffusion, and the equations for E have the form:

$$\nabla \times \mathbf{E} = 0, \quad \nabla \mathbf{E} = 4\pi \sum e_a n_a \int f_a(\mathbf{r}, \mathbf{p}, t) d\mathbf{p}.$$

We consider a plasma state for which the fractal structures arisen have the fractal dimension not containing explicitly the dependence on time and on the parameter P_i. In this case, we find for $d_a(\mathbf{r}, t)$ ($a = i, e$, $a \ne a'$, and V_a is the Coulomb interaction):

$$d_a(\mathbf{r}, t) = \beta_0 \int d\mathbf{r} d\mathbf{p} [(V_a(\mathbf{r} - \mathbf{r}') f_a(\mathbf{r}, \mathbf{p}, t) - V_{a'}(\mathbf{r} - \mathbf{r}') f_{a'}(\mathbf{r}, \mathbf{p}, t)] - \beta_0 \langle V_0 \rangle. \qquad (12)$$

Instead of a set of equations for f_a and d_a, we obtain after the substitution of (12) into (10) (e.g., for $\nu < 1$, $\nu = 1 + d$, $|d| \ll 1$)

$$\frac{\partial}{\partial t} f_a + \mathbf{v}\frac{\partial f_a}{\partial \mathbf{r}} + e_a \mathbf{E}(\mathbf{r}, t)\frac{\partial f_a}{\partial \mathbf{p}} + \alpha e_a(\langle V_a\rangle - \langle V_0\rangle) \times \frac{\partial}{\partial t} f_a + \alpha e_a \frac{\partial}{\partial t}\langle V\rangle f_a = I_a^{(\nu)}(\mathbf{r}, \mathbf{p}, t) + I_a^{(\mathbf{r})}(\mathbf{r}, \mathbf{p}, t), \quad (13)$$

$$\langle V_a(\mathbf{r}, t)\rangle = \beta_0 \int d\mathbf{r} d\mathbf{p}[V_a(\mathbf{r} - \mathbf{r}') f_a(\mathbf{r}, \mathbf{p}, t) - V_{a'}(\mathbf{r} - \mathbf{r}') f_{a'}(\mathbf{r}, \mathbf{p}, t)]. \quad (14)$$

The values $\langle V_a\rangle - \langle V_0\rangle$ can be replaced by $\langle T\rangle - \langle T_0\rangle$ etc. depending on the choice of a model for describing the fractal structures arising in plasma. For changes in the spectrum of elementary excitations in plasma nearby the state with a uniform density $f_0(\mathbf{p})$, we find using the known calculation methods (for the spectrum of longitudinal oscillations):

$$\omega' = \omega\left\{1 + \beta_0\left(\langle V\rangle - \langle V_0\rangle + \frac{\partial}{\omega \partial t}\langle V\rangle\right\}. \quad (15)$$

For cold electron plasma, relationship (15) leads to a minor renormalization of the frequency. For certain temperatures at $\langle V\rangle - \langle V_0\rangle < 0$ and $d \sim 1$, equation (10) is a nonlinear integro-differential equation with the generalized-fractional derivative. In this case, the spectrum of elementary excitations can be more complicated than (15) and reflects a substantial effect of fractal structures on the plasma behavior.

The method proposed of allowance for fractal structures in the systems described by kinetic equations leads to additional equations determining the dependence of the system fractal dimension on the distribution functions of the system objects. The equations for the distribution functions, written out by means of the generalized-fractional derivatives and involving the fractal dimensions, are supplemented by equations for determining the fractal dimensions themselves, which depend on the distribution functions. The consistent set of equations formed involving the generalized fractional derivatives is more complicated than the initial set of kinetic equations (into which it transforms when the fractional additions to the fractal dimension vanish) even for a small value of these additions compared to the integer dimensions. The appearance of new nonlinear terms in equations of the system can lead to the generation of new singular points of the bifurcation type, phase transitions, etc. This enables us to hope that, with allowance for external fields, we can obtain more convenient methods for controlling the behavior of complicated large statistical systems.

REFERENCES

1. B. B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, New York, 1982).
2. V. S. Ivanova, A. S. Balankin, I. Zh. Bunin, *et al.*, *Sinergetics and Fractals in Materials Science* (Nauka, Moscow, 1994).
3. V. L. Kobelev, E. P. Romanov, L. Ya. Kobelev, *et al.*, Dokl. Akad. Nauk **361**, 755 (1998) [Dokl. Phys. **43**, 537 (1998)].
4. L. Ya. Kobelev, V. L. Kobelev, and Ya. L. Kobelev, Available from VINITI, No. 967-B98 (Yekaterinburg, 1998).
5. L. Ya. Kobelev, Ya. L. Kobelev, and E. P. Romanov, Dokl. Akad. Nauk **370**, 757 (1999) [Dokl. Phys. **44**, 752 (1999)].
6. L. Ya. Kobelev, *Fractal Theory of Time and Space* (Konross, Yekaterinburg, 1999).
7. V. L. Ginzburg and L. D. Landau, Zh. Éksp. Teor. Fiz. **20**, 1064 (1950).
8. L. Ya. Kobelev, Ya. L. Kobelev, V. L. Kobelev, and E. P. Omanov, Available from VINITI, No. 1668-B99 (Yekaterinburg, 1999).
9. Yu. L. Klimontovich, *Statistical Theory of Open Systems* (Kluwer, Dordrecht, 1995).
10. V. L. Kobelev, E. P. Romanov, L. Ya. Kobelev, and Ya. L. Kobelev, Available from VINITI, No. 1667-B99 (Yekaterinburg, 1999).
11. A. A. Vlasov, Zh. Éksp. Teor. Fiz. **8**, 291 (1938).
12. Yu. L. Klimontovich, J. Plasma Phys. **59**, 647 (1998).
13. V. L. Kobelev and L. Ya. Kobelev, in *Abstracts of the IX International Meeting on Ferroelectricity, Seoul, August, 4–29, 1997*. Seoul, 1997, p. 10.
14. V. L. Kobelev, E. P. Romanov, L. Ya. Kobelev, and Ya. L. Kobelev, Fiz. Met. Metalloved. **85** (4), 54 (1988).
15. I. M. Gel'fand and G. E. Shilov, *Generalized Functions* (Fizmatgiz, Moscow, 1958; Academic Press, New York, 1964), Vol. 1.

Translated by V. Bukhanov

Doklady Physics, Vol. 45, No. 5, 2000, pp. 198–200. Translated from Doklady Akademii Nauk, Vol. 372, No. 3, 2000, pp. 319–321.

PHYSICS

Effect of Submicron Crystalline Structure on Field Emission of Nickel

L. R. Zubairov*, E. A. Litvinov, R. R. Mulyukov**, R. Sh. Musalimov**, and Yu. M. Yumaguzin*****

Presented by Academician G.A. Mesyats June 23, 1999

Received June 23, 1999

Materials consisting of ultrafine grains attracted widespread attention among specialists in physics and materials science about ten years ago [1–4]. This class of materials involves nanocrystalline ones with the mean grain size of about 10 nm and materials with the grains of submicron (about 100 nm) size. The interest in these materials stems from their significant difference in physical properties compared to conventional coarse-grained materials. This fact opens new possibilities in producing advanced materials with specified and even ultimate properties. However, physical mechanisms underlying the specific features of such materials are not yet quite clear [5]. To solve this problem, it seems reasonable to study their electronic structure.

In this paper, we present the results of our experimental studies of the electronic structure, which were performed by the field-emission electron spectroscopy.

As a material for our studies, we have chosen nickel (99.98% of purity). Its submicron-grained structure was formed by significant plastic deformations produced by the torsional technique under quasi-hydrostatic pressure of about 8 GPa on the basis of the Bridgman anvil-type facility [3, 4, 6]. The field-emissive cathodes were made of submicron-grained nickel by the method of electrochemical etching. The cathodes had a shape of pointed tips with the curvature radius of about 100 nm. The experiments were performed under conditions of ultrahigh vacuum ($<10^{-8}$ Pa) and employed a setup including an autoelectronic projector for continuous monitoring emission patterns and a dispersive electrostatic energy analyzer with the resolution not worse than 30 meV [7]. The cathode surface was subjected to cleaning in order to attain the stability of both the emission current and the emission pattern. The cleaning was performed by the method of evaporation in the reverse electric field. The cathode microstructure was studied by the JEM-2000EX electron microscope. For comparison, we also studied a sample of the nickel single crystal produced by annealing of the submicron-grained pointed tip in the experimental setup at about 800°C.

As a result of nickel processing by deformation, we obtained submicron-grained samples with a homogeneous grain structure having the mean grain size of 0.1 μm and nonequilibrium grain boundaries. This nonequilibrium state manifests itself by the diffuse contrast of grains in the electron-microscopy images and in curvy extinction contours within the grains. These results testify to the long-range internal stresses, whose sources are evidently induced by grain boundaries [3, 4]. Such a microstructure was also conserved in submicron-grained tips of autoelectronic emissive cathodes manufactured by electrochemical etching.

In the course of electron field-emission studies of the submicron-grained samples, the emission pattern included two regions with parallel alternating light and dark stripes. The emission pattern for the single-crystalline sample looked like a fragment of the conventional emission pattern for nickel [8].

The energy distributions of emitted electrons were recorded from both the light regions and the intermediate ones located between light and dark stripes. For a submicron-grained sample, we have obtained two types of electron distributions depending on the choice of the emission region at the cathode surface. While scanning the light region, we obtained spectra having only a single peak (Fig. 1a), whereas scanning intermediate regions between light and dark stripes has revealed an additional peak in the low-energy part of the distribution (Fig. 1b). For annealed nickel, the electron total-energy distribution had the conventional form corresponding to the free-electron model [9].

* *Institute of Problems of Metal Superplasticity, Russian Academy of Sciences, ul. Khalturina 39, Ufa, Bashkortostan, 450001 Russia*

** *Institute of Electrophysics, Ural Division, Russian Academy of Sciences, ul. Komsomol'skaya 34, Yekaterinburg, 620219 Russia*

*** *Bashkortostan State University, ul. Frunze 32, Ufa, Bashkortostan, 450074 Russia*

1028-3358/00/4505-0198$20.00

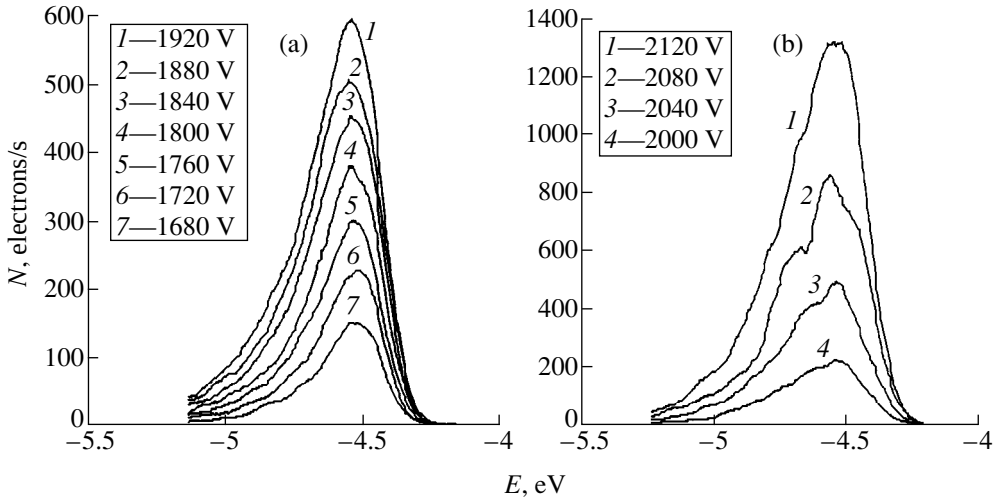


Fig. 1. Total-energy distributions for electrons emitted from submicron-grained nickel at different autoelectronic-emission voltages.

For submicron-grained samples, the electron distribution differed from the conventional one. In contrast to annealed nickel, we observed significant broadening the peaks and their shift toward low-energies. For the energy distribution corresponding to a submicron-grained sample, the full peak width at the half-maximum was 0.35 eV or larger. For nickel single crystals, this value equaled 0.24 eV (Fig. 2). The average shift of the peak in the energy distribution for the submicron-grained sample was 0.16 eV compared to single-crystal samples.

For interpretation of the experimental data obtained, we present the electron total-energy distributions in the following form [10, 11]:

$$N = \frac{j}{e} S f(x, y),$$

where j is the emission-current density, e is the electron charge, S is the effective emission area for the spectra recorded. The function $f(x, y)$ is defined as

$$f(x, y) = \frac{(\exp(x))^y}{1 + \exp(x)},$$

where $x = (\varepsilon - \varepsilon_F)/k\,T$, $y = T/2T_i$, ε is the energy of emitted electrons; ε_F is the Fermi energy; k is the Boltzmann constant; T is the absolute temperature; T_i is the inversion temperature for which the calorimetric Nottingham effect vanishes [12], i.e.,

$$T_i = \frac{eh}{8\pi\sqrt{2m\varphi}\,\eta\left(\frac{\sqrt{e^3E}}{\varphi}\right)}E; \qquad (1)$$

h is the Planck constant; m is the effective electron mass; φ is the work function; E is the electric-field intensity; $\eta(x)$ is a slightly varying function of x

$$\eta(x) = \theta(x) - \frac{2}{3}x\frac{d\theta(x)}{dx};$$

and $\theta(x)$ is the Nordheim function.

The electron total-energy distribution N expressed in terms of the function $f(x, y)$ has the maximum

$$N_m = \frac{j}{e} S(1-y)\left(\frac{y}{1-y}\right)^y$$

at an energy

$$\varepsilon_m = \varepsilon_F - kT\ln\left(\frac{1-y}{y}\right). \qquad (2)$$

Relationship (2) shows that the position of the maximum in the horizontal axis shifts to the left with

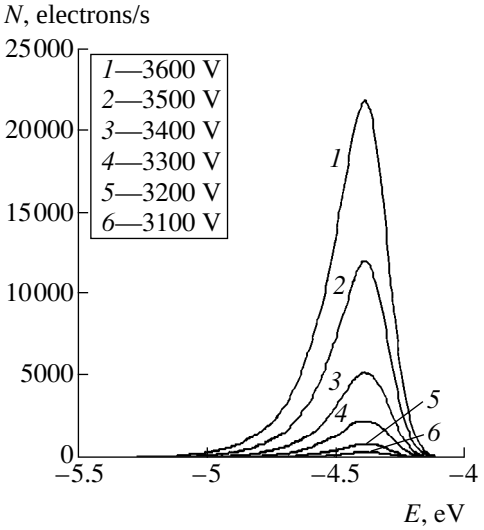


Fig. 2. Energy distributions for electrons emitted from annealed nickel at different emission voltages.

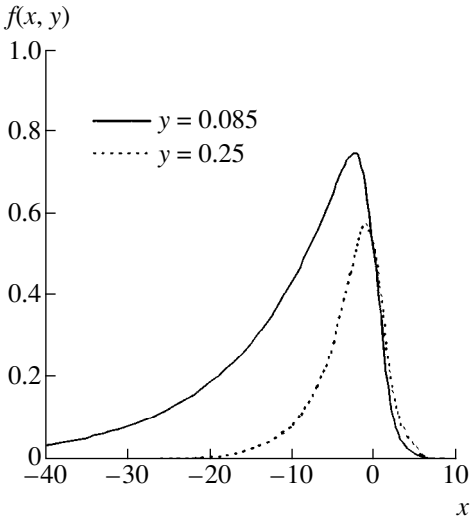


Fig. 3. Plots for the function $f(x, y)$.

decreasing y. As an illustration, we present in Fig. 3 plots for the function $f(x, y)$. As is seen, the curve becomes wider, and its maximum shifts to the left with decreasing y. The decrease in y can be caused by elevating the inversion temperature T_i. It is clear from (1) that the inversion temperature can increase due to the growth of the electric field E or/and decrease in the work function φ. Figures 1 and 2 demonstrate that the appreciable emission from a submicron-grained sample occurs at lower voltages. This implies the decrease in the work function. Therefore, the broadening of the peak in the energy distribution for electrons emitted from submicron-grained samples and the leftward shift of this peak can be attributed to the lowering of the effective work function. Additional peaks in the distribution curves recorded from the boundary regions between the light and dark domains can be related to the competition between the emission processes of regions with low and high work functions.

ACKNOWLEDGMENTS

The work was supported by the Russian Foundation for Basic Research, project no. 97-02-17236 and by the Academy of Sciences of Bashkortostan.

REFERENCES

1. I. D. Morokhov, L. I. Trusov, and V. I. Lapovok, *Physical Phenomena in Ultradisperse Media* (Nauka, Moscow, 1984).
2. R. Birringer and H. Gleiter, in *Encyclopedia of Materials. Sci. and Eng. Suppl. 1* (Pergamon Press, Oxford, 1988), pp. 339–351.
3. R. Z. Valiev, A. V. Korznikov, and R. R. Mulyukov, Mater. Sci. Eng., A **168**, 141 (1993).
4. R. R. Mulyukov, *Structure and Properties of Submicron-Crystalline Metals Obtained in the Process of Intense Plastic Deformation*, Doctoral Dissertation in Mathematical Physics (Moscow Institute of Steel and Alloys, Moscow, 1997).
5. *Proceedings of the International Conference on Nanostructured Materials, Stockholm, June 14–19, 1998.*
6. N. A. Smirnova, V. I. Levit, V. P. Pilyugin, *et al.*, Fiz. Met. Metalloved. **61**, 1170 (1986).
7. R. Z. Bakhtizin, V. M. Lobanov, and Yu. M. Yumaguzin, Prib. Tekh. Éksp., No. 4, 247 (1987).
8. N. Rihon, Phys. Status Solidi A **49**, 697 (1978).
9. A. Modinos, *Field, Thermoionic and Secondary Electron-Emission Spectroscopy* (Plenum Press, New York, 1984; Nauka, Moscow, 1990).
10. E. L. Murphy and R. H. Good, Phys. Rev. **102** (6) (1956).
11. P. H. Levine, J. Appl. Phys. **33** (2) (1962).
12. W. B. Nottingham, Phys. Rev. **59** (11) (1941).

Translated by K. Kugel'

Doklady Physics, Vol. 45, No. 5, 2000, pp. 201–204. Translated from Doklady Akademii Nauk, Vol. 372, No. 3, 2000, pp. 322–325.

PHYSICS

On Application of the Morse Potential for Evaluating Anharmonic Vibrations of Polyatomic Molecules with XH_n-Complexes

G. V. Yukhnevich

Presented by Academician V.V. Osiko September 17, 1999

Received September 18, 1999

Molecular vibrations are known not to be strictly harmonic. In the case of XH-bonds, their anharmonicity is sufficiently large to affect the frequencies of the first-order transitions. For example, while deuteration of a certain compound, the decrease in the valence-vibration frequency squared of its XH-bond is smaller than its reduced-mass rise. For compensating this effect, the so-called "spectroscopic" masses of the protium and deuterium atoms were introduced into the theory of harmonic molecular vibrations [1].

Anharmonic molecular vibrations are usually calculated on the basis of the perturbation theory. This method involves large computational difficulties and, therefore, allows the evaluation of only second-order or third-order transitions for 3-atomic to 5-atomic molecules [2, 3]. It was believed that calculations for higher-lying levels and larger-size molecules could be performed by using modern computers. However, this hope has turned out to be not quite justified.

The application of the Ritz variational procedure for evaluating anharmonic vibrations of polyatomic molecules makes it possible to significantly simplify the calculations and to find all energy levels lying in the lower quarter of a potential well [4–7]. Moreover, a code was recently developed, which has made it possible to calculate all vibrational energy levels for molecules with a practically arbitrary number of atoms [8]. In spite of these arguments, calculation of the anharmonic vibrations has not yet become an ordinary procedure in spectrochemical practice. Such a calculation is impossible because potentials of bonds for specific molecules are unknown. In contrast to a harmonic potential function, arbitrary anharmonic potentials can be defined by not less than two parameters. In the case of the Morse potential

$$E = D_e[1 - \exp\alpha(r_e - r)]^2,$$

which is used by the code described in [8], these parameters are the depth D_e of the potential well and the factor α, which is inversely proportional to the well width. The former parameter can usually be found from independent measurements, while the latter, as a rule, remains unknown. The absolute value of the bond equilibrium length r_e, easily measurable by electron-diffraction methods and according to rotational molecular spectra, does not explicitly influence the form of the Morse potential.

At the same time, from general physical reasons, the width of a potential well could be assumed to increase with the bond length. If such an assumption is valid, and a one-to-one correspondence between the length of a bond and the width of its potential well can be ascertained, the calculation of anharmonic vibrations for a large variety of molecules could be easily performed with the help of the method described in [8]. This paper is devoted to searching for a correlation between the length r_{XY} of an X–Y-bond and the coefficient α_{XY} entering into the expression for the Morse potential energy of this bond.

To solve this problem, we have analyzed all known neutral and charged diatomic hydrides (50 protium molecules and 38 their deuterium analogues). For ground states of these compounds, the anharmonicity coefficients $\omega_e x_e$ (standing ahead the vibrational quantum number squared in the expression for the molecular energy), as well as the interatomic distances r_e have been accurately measured [9]. This made it possible to calculate the desired quantities α_{XY}. To compare with the hydrides, we analyzed also 28 homopolar molecules formed by atoms of the first and seventh groups of the periodic system of elements, as well as those of the first period.

Using the expression for the anharmonicity coefficient $\omega_e x_e$ in the Morse function (see, for example, [8]), we obtain the relationship

$$\alpha = 0.2436(M\omega_e x_e)^{1/2}.$$

Here, the coefficient α, reduced mass M, and $\omega_e x_e$ are expressed in $Å^{-1}$, atomic units, and cm^{-1}, respectively.

Kurnakov Institute of General and Inorganic Chemistry, Russian Academy of Sciences, Leninskiĭ pr. 31, Moscow, 117907 Russia

1028-3358/00/4505-0201$20.00

Reduced masses M (atomic units), anharmonicity coefficients $\omega_e x_e$ (cm^{--1}), exponential factors α^{--1} (Å) of the Morse function, and equilibrium bond lengths r_e (Å) for diatomic molecules and ions

Molecule	M	$\omega_e x_e$	α^{-1}	r_e	Molecule	M	$\omega_e x_e$	α^{-1}	r_e
HeH^+	0.8051	157.7	0.364	0.770	IH	0.9999	39.64	0.652	1.61
LiH	0.8812	23.20	0.908	1.595	CsH	1.000	12.93	1.142	2.49
BeH	0.9064	36.31	0.716	1.34	BaH	1.000	14.50	1.078	2.23
BH	0.9233	49.40	0.608	1.23	YbH	1.002	21.06	0.894	2.05
CH	0.9297	63.02	0.536	1.12	LuH	1.002	22.0	0.874	1.91
NH	0.9402	78.35	0.478	1.04	PtH	1.002	46.0	0.605	1.53
OH	0.9481	84.88	0.458	0.970	AuH	1.003	43.12	0.624	1.52
OH^+	0.9480	78.52	0.476	1.029	HgH^+	1.003	40.9	0.641	1.59
FH	0.9571	89.88	0.443	0.917	TlH	1.003	22.7	0.860	1.87
FH^+	0.9571	89.00	0.445	1.00	PbH	1.003	29.75	0.751	1.84
NaH	0.9655	19.72	0.941	1.89	BiH	1.003	31.6	0.729	1.81
MgH	0.9672	31.89	0.739	1.73	H_2	0.504	121.3	0.525	0.741
MgH^+	0.9672	31.94	0.738	1.65	Li_2	3.508	2.61	1.357	2.67
AlH	0.9715	29.09	0.772	1.65	LiNa	5.376	1.61	1.394	2.81
SiH	0.9728	35.51	0.698	1.52	B_2	5.504	9.35	0.572	1.59
SiH^+	0.9728	34.24	0.711	1.504	C_2	6.000	13.34	0.459	1.242
PH	0.9761	44.5	0.623	1.42	N_2	7.002	14.32	0.410	1.097
SH	0.9770	59.9	0.537	1.34	O_2	7.997	11.98	0.419	1.21
ClH	0.9796	52.82	0.571	1.275	F_2	9.49	11.24	0.398	1.41
ClH^+	0.9796	52.54	0.572	1.315	Na_2	11.49	0.725	1.422	3.08
KH	0.9824	14.3	1.095	2.24	FCl	12.31	6.16	0.471	1.63
CaH	0.9830	21.8	0.887	2.00	NaK	14.46	0.511	1.510	3.59
CrH	0.989	32.0	0.730	1.66	FBr	15.31	4.05	0.521	1.76
MnH	0.990	28.8	0.769	1.73	FI	16.52	3.12	0.572	1.91
CoD	1.948	17.59	0.701	1.52	Cl_2	17.48	2.68	0.600	1.99
NiH	0.9906	38	0.669	1.47	K_2	19.48	0.283	1.749	3.91
CuH	0.9919	37.51	0.676	1.46	ClBr	24.23	1.84	0.614	2.14
ZnH	0.992	55.14	0.555	1.59	AgCl	26.35	1.17	0.739	2.28
ZnH^+	0.992	39.0	0.660	1.51	ClI	27.41	1.50	0.640	2.32
GaH	0.9933	28.77	0.768	1.66	Cu_2	31.46	1.03	0.723	2.22
GeH	0.9939	37.0	0.677	1.59	As_2	37.46	1.12	0.635	2.10
BrH	0.9954	45.22	0.612	1.414	Br_2	39.46	1.08	0.630	2.28
BrH^+	0.9954	47.4	0.598	1.45	Kr_2	41.96	1.34	0.547	4.03
RbH	0.9960	14.21	1.091	2.37	Rb_2	42.46	0.105	1.944	3.94
SrH	0.9964	17.0	0.997	2.146	BrI	48.66	0.814	0.652	2.47
PdD	1.977	19.59	0.660	1.53	I_2	63.45	0.615	0.657	2.67
AgH	0.9984	34.06	0.704	1.52	Xe_2	65.19	0.65	0.63	4.36
CdH^+	0.999	35.4	0.690	1.67	Cs_2	66.45	0.082	1.755	4.47
InH	0.999	25.61	0.812	1.84	Au_2	98.48	0.42	0.638	2.47

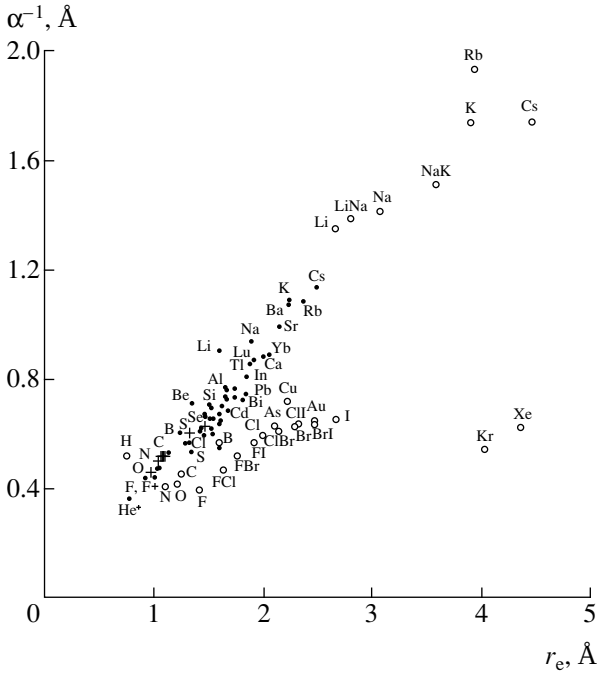
Using this formula for protium and deuterium forms of the hydrides, we have evaluated the values of α_H and α_D. In the table, to reduce its size, we list the calculated results only for protium forms of the hydrides (except for CoD and PdD).

It follows from the results obtained that, for the molecules and ions under consideration, the width parameter α of the potential function varies more than by the factor of 5 (from 0.514 to 2.745 $Å^{-1}$). However, for such pairs of unlike molecules as CH and SH, CrH and BiH,

and ZnH^+ and PdD, the values of α coincide with the accuracy better than 0.1%. This situation can be explained by both the incomplete adequacy of the Morse function for actual potentials of existing bonds [1] and the complicated dependence of potentials under consideration on different bond parameters. To find the lower estimate for the difference between the Morse function and the actual form of a bond potential, we can compare the values of the exponential factors α_H and α_D, which were calculated for the same potential. Indeed, it follows from the results presented in the table, that the ratio α_H/α_D averaged over the 38 molecules is equal to 1.001. In this case, the root-mean-square deviation is 0.011. Therefore, the value of α for each molecule is found within the error of not less than 1%. Such an error, fairly significant when estimating the parameter of a specific molecule, becomes unessential when analyzing the totality of the molecules for which this parameter varies by the factor of 5.

To examine the influence of the bond length on the parameter α_{XY}, which is inversely proportional to the width of the potential well of an XY-molecule, we plot the values of α_{XY}^{-1} as a function of the corresponding bond lengths (see figure). As is seen, there exists no unified dependence between the parameters α_{XY}^{-1} and r_{XY} for the entire totality of the molecules under consideration. However, for certain more narrow molecular groups, we can observe very clear correlations between these parameters. Indeed, almost all the homopolar molecules (shown in figure by circles) can be divided into three sets: (F_2, FCl, FBr, FI, Cl_2, ClBr, As_2, ClI, Br_2, Au_2, BrI, and I_2), (N_2, O_2, C_2, and B_2), and (Li_2, LiNa, Na_2, NaK, K_2, Rb_2, and Cs_2), which obey three different dependences. It is curious that the first of these sets contains Au_2 and As_2 molecules in addition to those formed by atoms of seventh group of the periodic system of elements. For the hydrogen molecule, none of these dependences holds. Thus, the presence of different dependences for various groups of the homopolar molecules implies that, in general, the parameter α_{XY} is determined not only by the bond length.

Moreover, the results presented indicate that in the case of diatomic hydrides quite a different situation takes place. Indeed, for 48 of 50 hydrides (except LiH and BeH), for which we succeeded in evaluating the parameters α_{XH}^{-1} and r_{XH}, the proportionality factor between them varies from 0.4 to 0.5 (dots in the figure). If this spread is considered to be a statistical straggling, the mean value of the factor will be approximately equal to 0.45. However, this spread is unlikely random. More probably, it is caused by a specific electronic structure of various atoms. Therefore, for the bonds such as OH, NH, FH, CH, and BH, which are the most popular in spectroscopic practice, the relation $\alpha_{XH}^{-1} = (0.475 \pm 0.015) r_{XH}$, based on the corresponding values shown in figure, is more correct. It is worth noting that the values of α_{XH}^{-1}, which were found by solving inverse spectroscopic problems [8] for larger molecules (H_2O, H_2S, H_2Se, HCN, C_2H_2, C_2H_4, COH_2, and NH_3) shown by crosses in figure satisfy this relation as well.



Dependence between the bond length and width of the corresponding potential well given in the form of the Morse function: circles, dots, and crosses correspond to diatomic homopolar molecules, diatomic hydrides, and XH-bonds of polyatomic molecules, respectively.

Thus, it follows from the results obtained that, in the case of approximating the potential of XH-bonds by the Morse function, the coefficients α_{XY} turn out to be inversely proportional to the lengths of these bonds. Since the proportionality factor remains essentially constant in passage from one atom X to another, we can expect that this factor is unaffected by varying the length of the XH-bond, which is caused by intramolecular or intermolecular interactions. Therefore, for an arbitrary molecule with the known length of its XH-bond, the dependence found makes it possible to evaluate the parameter α_{XH}. As a result, anharmonic vibrations of XH-bonds and their reaction capabilities can be successfully calculated for a wide class of molecules and chemical complexes.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project no. 99-03-32004.

REFERENCES

1. A. M. El'yashevich, *Atomic and Molecular Spectroscopy* (Fizmatgiz, Moscow, 1962).
2. J. Pliva, in *Critical Evaluation of Chemical and Physical Structural Information*, Ed. by D. R. Lide, Jr. and M. A. Paul (National Academy of Science, Washington, 1974), Chap. 5, pp. 283–311.
3. V. V. Rossikhin and V. P. Morozov, *Potential Constants and Electron-Optical Parameters of Molecules* (Énergoizdat, Moscow, 1983).
4. L. A. Gribov, A. I. Pavlyuchko, and G. F. Lozenko, Zh. Prikl. Spektrosk. **36**, 87 (1982).
5. A. I. Pavlyuchko, G. F. Lozenko, G. V. Khovrin, *et al.*, Opt. Spektrosk. **52**, 64 (1982) [Opt. Spectrosc. **52**, 37 (1982)].
6. L. A. Gribov, A. I. Pavlyuchko, and G. F. Lozenko, Zh. Prikl. Spektrosk. **36**, 274 (1982).
7. A. I. Pavlyuchko, Zh. Prikl. Spektrosk. **43**, 119 (1985).
8. L. A. Gribov and A. I. Pavlyuchko, *Variational Methods for Solving Anharmonic Problems in Theory of Molecule Vibrational Spectra* (Nauka, Moscow, 1998).
9. K. P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure. IV. Constants of Diatomic Molecules* (Van Nostrand Reinhold, New York, 1979; Mir, Moscow, 1984).

Translated by V. Chechin

Doklady Physics, Vol. 45, No. 5, 2000, pp. 205–207. Translated from Doklady Akademii Nauk, Vol. 372, No. 1, 2000, pp. 40–42.

TECHNICAL PHYSICS

A Regime of the Solid-State Reaction in the Self-Propagating High-Temperature Synthesis

M. A. Korchagin, T. F. Grigor'eva, A. P. Barinova, and Corresponding Member of the RAS N. Z. Lyakhov

Received November 9, 1999

1. A concept of the "true solid-state combustion" (TSC) has long been in use for processes of self-propagating high-temperature synthesis (SHS) [1]. It was put forward based on the comparison of adiabatic combustion temperatures (T_{ad}) and characteristic temperatures of the most low-melting eutectics in the equilibrium phase diagrams. For example, Ni + 2Si mixture is characterized by T_{ad} = 1869 K, whereas the lowest eutectic temperature in the phase diagram for Ni–Si system is 1560 K. Therefore, it is natural to expect that it is sufficient to introduce about 20 wt % of the final product into the mixture (and thereby to lower T_{ad}) to attain a solid-state combustion regime [1]. However, the TSC was still not observed in experiments despite much efforts made in this direction. Moreover, the available data on nonequilibrium eutectics cast a certain doubt on the possibility of predicting the TSC based only on analysis of equilibrium phase diagrams.

Estimates obtained within the framework of the model for the reaction-induced diffusion demonstrate [2, 3] that ultrafine particles (of ≤0.1 μm) of SHS reagents are needed to realize the true solid-state combustion. Currently, it is impossible to obtain the powders with such small particles of SHS components. In [1], several possible ways to attain the TSC were suggested, but none of them has been implemented yet.

In this paper, it is shown that the true solid-state combustion in SHS systems can actually be implemented using preliminary mechanical activation of the reacting mixtures by intense grinding in the planetary mill.

2. The preliminary mechanical activation of reacting mixtures was performed in the argon ambient using a water-cooled AGO-2 planetary ball mill [4].

The combustion of the samples having bulk density was also performed in the argon ambient at the atmospheric pressure. The SHS was initiated using the fuse ignited by the current-heated nichrome spiral.

Institute of Solid-State Chemistry and Mineral Processing, Siberian Division, Russian Academy of Sciences, ul. Kutateladze 18, Novosibirsk, 630128 Russia

For measuring the combustion rate U_c and temperature T_c, the tungsten–rhenium thermocouples 100 μm in diameter were used. They were placed at certain distances between them. The signals from thermocouples were recorded by an N-117 loop oscilloscope.

Reacting mixtures after mechanical activation and SHS products were analyzed by the X-ray diffraction technique and by means of scanning electron microscopy. X-ray diffraction patterns were recorded by DRON-3M and URL-63 diffractometers (using CuK_α radiation). Electron microscopy studies were performed using JSM-T20 and JEM-2000FX-II microscopes supplied by JEOL.

The experiments were performed using the well-known and rather thoroughly studied Ni–Ti SHS system, as well as certain compositions for Ni–Si, Ni–Al, and Fe–Fe systems. To prepare reacting mixtures, we used such powders as carbonyl nickel PNK-1L5, ultrapure carbonyl iron (grade A), aluminum PA-4, titanium PTOM, and high-purity semiconducting silicon.

3. It is well known that the SHS in Ni–Ti systems cannot be initiated without preliminarily heating the reacting mixture [3]. Our experiments demonstrated that after a short-term preliminary mechanical activation, Ni + 45 wt % Ti samples began to burn at the room temperature. For this composition, the plots of U_c and T_c as a function of mechanical activation time are presented in Fig. 1. We call attention to the anomalously low values of T_c. These very low values of T_c after the preliminary mechanical activation turned out to be characteristic of all systems under study. The cause of this behavior is yet not quite clear.

The electron microscopy studies of SHS products demonstrated that the solid-state combustion regime appears at an arbitrary duration of the mechanical activation. None of the samples exhibited any traces of melting. The morphology and microstructure of initial samples remained unchanged in the SHS products.

It was also found that the products of mechanical activation had very complicated structures, and their morphology was changing in the course of mechanical processing. Even after 30 s of mechanical activation, it was practically impossible to find initial nickel and tita-

1028-3358/00/4505-0205$20.00

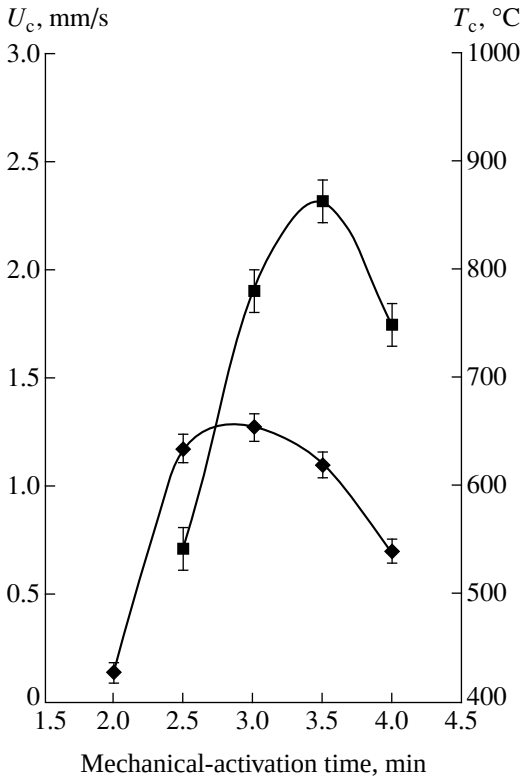


Fig. 1. Combustion rate U_c and temperature T_c as a function of duration of preliminary mechanical activation for Ni + Ti (45 wt %) composition.

nium particles in the sample. The products contained agglomerates of various shapes and sizes. As a rule, large agglomerates consisted of smaller rounded particles that, in turn, also appeared to be agglomerates of initial nickel and titanium.

When the duration of mechanical activation grows, the fraction of large agglomerates increases. With the growth of their density, these agglomerates acquire a plateletlike shape. The formation of so-called layered composites begins [5]. An external view of these composites is shown in Fig. 2. Even after 2.5 min of mechanical activation, the composites of this system have a rather dense surface with 0.5–0.1-μm grain sizes. In some cases, the surface of composites bears traces of pronounced plastic distortions resembling those produced by drawing.

The further increase in the time of mechanical activation leads to increasing density of composites and to decreasing grain size down to that characteristic of ultrafine grains (≤0.1 μm), see Fig. 3.

The reduction of combustion rate after 3.5 min of mechanical activation is related to the "poisoning" of reacting mixture by intermediate intermetallic compounds arising as a result of mechanically induced fusion [6], and by the consequent destruction of the large-scale layered structure in the composites.

Thus, the obtained results suggest that the implementation of the SHS at the room temperature in a Ni–Ti system and the enhancement of U_c with the duration of the mechanical-activation time are caused by the fine grinding of reagents and the increase in their contact areas within the arising layered composites. The structure of these composites, the clear traces of plastic deformation observed at their surface, as well as the corresponding models of solid-state chemistry clearly indicate the decrease in the characteristic size of inhomogeneities and formation of a new structure with high density of defects [7].

The fine grinding of reagents down to ultrafine grain size apparently gives rise to the TSC.

For all durations of the mechanical-activation time, the X-ray diffraction patterns exhibit only the NiTi lines (the basic phase) with a small admixture of reflections corresponding to the Ti_2Ni phase. The lines corre-

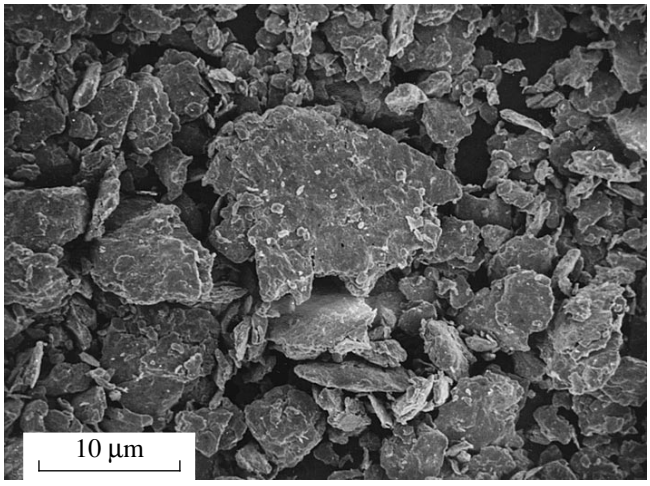


Fig. 2. Composites formed after 2.5-min mechanical activation of Ni + Ti (45 wt %).

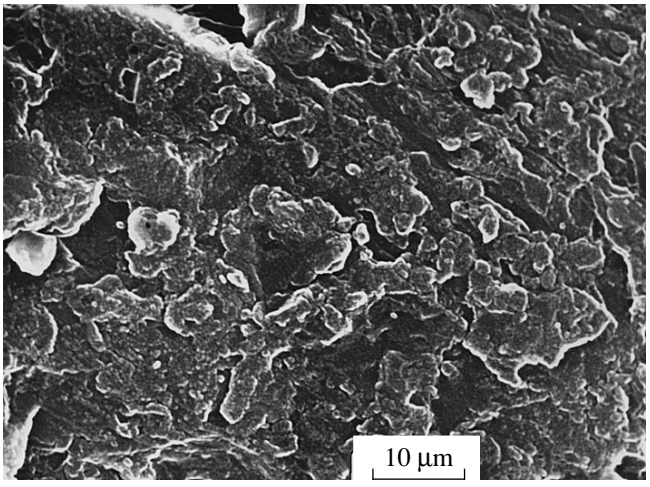


Fig. 3. Surface of the layered composite [Ni + Ti (45 wt %), 3.5-min mechanical activation].

sponding to titanium and nickel are not observed. Just the same composition of products is also observed for the SHS occurring in the case of preliminary heating the mixture of conventional powders [3].

We have succeeded in attaining the solid-state combustion regime for some other systems, in particular, for Ni + Si (10 wt %) mixtures and, strange as it may seem, for Ni + Al (13 wt %). The latter fact is surprising, because this system involves such a low-melting reagent as aluminum. Note that it is impossible to attain the SHS in these compositions without their preliminary mechanical activation. Similarly to the previous case, the increase in mechanical activation time for these compositions results in the formation of layered composites, which are characterized by the ultrafine grinding of reagents causing the enhanced contact area. For these mixtures, the plots of the combustion rates as functions of preliminary mechanical activation time have a pronounced peaked shape.

The electron-microscopy studies of these two systems did not reveal any melting of SHS products in all the samples except those corresponding to the maximum values of U_c and T_c (mechanical activation for 2 min).

After the SHS, has occurred the samples consist of the same layer composites, even the traces of drawing remain at their surface. The grain size remains nearly unchanged. The only difference between them and the initial samples is, maybe, the lower density of certain composites after the SHS. The microscopic image of the Ni + Al (13 wt %) composites in the SHS products with the flaked upper platelets is presented in Fig. 4 (after 1.5 min of mechanical activation).

The SHS products for these compositions corresponding to maximum values of U_c and T_c contain, although quite rarely, spherical partially-melted particles. The surface of these particles clearly exhibits the pattern typical of the crystallization from the liquid phase.

For the Fe–Si system, we succeeded in attaining the solid-state combustion for Fe + Si, 5Fe + Si, and 3Fe + Si compositions (after 3 min of mechanical activation). It was impossible to attain the SHS for the Fe + 2Si composition (after the same mechanical activation during 3 min). According to the data reported in [8], the SHS for this composition can be performed by heating a sample up to about 400°C after mechanical activation for 4 h. However, the mill of distinctly different design was used in this case.

In our case, the electron microscopy revealed that the Fe + 2Si composition differs from other ones of this system: products of the mechanical activation contain ultrafine particles of the reagents, but they do not form dense composites. Ultrafine particles form only loose agglomerates. Obviously, this fact related to the volumetric fraction of low-plasticity silicon in the mixture.

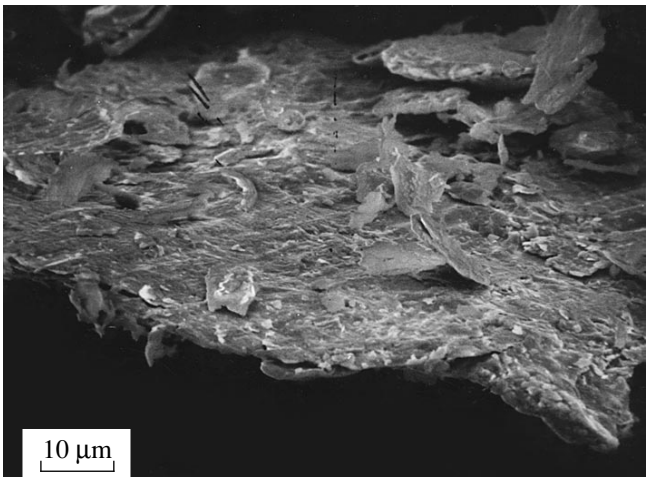


Fig. 4. Layered composite in SHS products for Ni + Al (13 wt %) composition.

(The value of the latent heat of $FeSi_2$ formation even exceeds slightly the ΔH value for FeSi).

This result seems to be quite interesting, since it unambiguously demonstrates that not only fine grinding is important for the SHS process after mechanical activation but the formation of dense composites as well, in which the drastic increase in the contact surface area occurs.

REFERENCES

1. A. G. Merzhanov, *Combustion Process and Synthesis of Materials* (Chernogolovka, 1998).
2. A. G. Merzhanov, *Regularities and Mechanism of Combustion for Explosive Mixtures of Titanium and Boron* (Chernogolovka, 1978).
3. V. I. Itin and Yu. S. Naĭborodenko, *High-Temperature Synthesis of Intermetallic Compounds* (Tomsk. Univ., Tomsk, 1989).
4. E. G. Avvakumov, A. R. Potkin, and O. I. Samarin, *Planetary Mill*, USSR Inventor's Certificate No. 975068, Byull. Izobret., No. 43 (1982).
5. P. S. Gilman and J. S. Benjamin, Ann. Rev. Mater. Sci. **13**, 270 (1983).
6. T. F. Grigorieva, A. P. Barinova, V. V. Boldyrev, *et al.*, in *Proceedings of the International Symposium on Metastable, Mechanical Alloyed and Nanocrystalline Materials, ISMANAM'96, Rome, 1996*, pp. 577–582.
7. T. F. Grigor'eva, *Mechanochemical Synthesis of Metastable Intermetallic Phases and Their Reacting Ability*, Candidate's Dissertation in Chemistry (Institute of Solid-State Chemistry and Mineral Processing, Siberian Division, Russian Academy of Sciences, Novosibirsk, 1988).
8. C. Gras, E. Gaffet, F. Bernard, *et al.*, Mater. Sci. Eng., A **264**, 94 (1999).

Translated by T. Galkina

Doklady Physics, Vol. 45, No. 5, 2000, pp. 208–212. Translated from Doklady Akademii Nauk, Vol. 372, No. 3, 2000, pp. 326–329.

TECHNICAL PHYSICS

The General Mechanism for Generation of the $1/f$ Noise

A. G. Budarin

Presented by Academician B.V. Bunkin June 23, 1999

Received June 25, 1999

The flicker noise ($1/f$ noise) is a universal phenomenon observed while measuring transport processes in various systems. In studies [1–3], the mathematical theory of the processes generating the $1/f$ noise was proposed. This theory explains its nature from a unified standpoint. The basic assumption consists in the fact that the $1/f$ noise, in itself, is not a physical process but results from the conversion, in the course of measurements, of a certain actual physical process (with an essential dependence of the spatial coordinate) into a purely temporal random function with the $(1/f)$-type spectrum. According to the second assumption of the theory, the mechanism responsible for this process is a sequential combination of two mechanisms: the formation of fluctuations and their essentially slower relaxation described, respectively, by pulsed Poisson processes $S(x, t)$ and $R(x, t)$ with the pulses $a_i s(x - x_i)\delta(t - t_i)$ and $a_i r(x - x_i, t - t_i)$, where $i = 1, 2, \ldots, K$. Here, K is the number of pulses for the observation time t_0, a_i is their amplitude, x_i and t_i are the parameters characterizing the location of fluctuations and moments for their initiation, which are distributed uniformly within the interval $(0, L)$, and $(0, t_0)$; $r(x - x_i)$, $(t - t_i)$ is the solution to the diffusion equation $r'_t = Ur'_x + \kappa_D r''$ with the initial condition $r(x - x_i, 0) = s(x - x_i)$ and the boundary condition $r(0, t) = r(L, t) = 0$. According to the third assumption of the theory, the observed value is a relaxation flux $J(x, t)$ formed by a sum of contributions

$$j(x - x_i, t - t_i) = -\kappa_D a_i r'_x(x - x_i, t - t_i).$$

In [1–3], the measured spectrum $F(\omega)$ of such a process was proven to have the desired shape $F(\omega) \sim \omega^{-\alpha}$ only if $s(x) \sim x^{-\beta}$, $\beta = 3/2 - \alpha$, $\beta \in (0, 1)$, and $\alpha \in (0.5, 1.5)$. In the same studies, the general reason is indicated for the initiation of the process $S(x, t)$, which is the basic one in the generation of the $1/f$ noise.

The goal of this study is mathematical description of the general physical mechanism for the formation of the process with power-type fluctuations and estimation on this basis of the amplitude of the $1/f$ noise spectrum. According to the general theory of nonlinear waves [4, 5], such density distributions for a physical parameter of the medium (e.g., mass, temperature, and momentum) arise near discontinuities for the solutions to evolution equations of the shock-wave type. The appearance of these distributions is a characteristic result of developing smooth perturbations propagating in continuous media. The principal conditions of such a development is a nonlinear relation between the density of a medium parameter and its flux, as well as a sufficiently large amplitude of perturbations required to overcome the smoothing action of the relaxation processes. The second condition implies that the initial fluctuation pulses inherently constitute rare events and, thereby, give rise to a nearly Poisson process. Thus, the process $S(x, t)$ is simply a random sequence of shock waves arising due to fluctuations of the field of a continuous physical parameter. Since this physical situation has an exceptionally general nature [4], the universality of the $1/f$ noise as the observable phenomenon becomes understandable.

The more detailed presentation of the proposed physical picture is based on the general evolution equation for a medium parameter [4, 5]:

$$\rho_t + q'_x = I(\rho), \tag{1}$$

where $\rho = \rho(x, t)$ is the density of the medium parameter, $q = q(x, t) = Q(\rho) - \xi\rho'_x$ is its flux, and $I(\rho)$ is the operator determining the variation in the number of particles in the flux due to collisions (collision integral).

We now make standard simplifying assumptions inherent in both the theory of nonlinear simple waves [4] and relaxation approximation for $I(\rho)$ [6]. We consider equation (1) to be one-dimensional and the relation between q and ρ to be pure functional ($\xi = 0$) and quadratic:

$$Q(\rho) = c_0[(\rho - \bar{\rho}) + B(\rho - \bar{\rho})^2/2].$$

We also take into account that the state of the medium is close to the equilibrium one, i.e., $I(\rho) = (\rho - \bar{\rho})/\tau_r$ (τ_r is the relaxation time), and $Z = B(\rho - \bar{\rho}) \ll 1$. In this case, equation (1) and the boundary condition for the quantity ρ can be written in the reduced form as

$$Z'_t + c_0(1 + Z)Z'_x + Z/\tau_r = 0, \quad Z(0, t) = BR(t). \tag{2}$$

Here, the smooth function $R(t)$ is the time dependence of the initial perturbance in a small physical neighborhood of the point $x = 0$ in which all the physical parameters used are considered to be determinate.

1028-3358/00/4505-0208$20.00

Changing variables $\zeta = x$, $\tau = t - x/c_0$, $Z = U(\zeta, \tau)$ in (2), we obtain in the first approximation for $Z \ll 1$ the equation $U'_\zeta - UU'_\tau/c_0 + U/\delta = 0$, where $\delta = c_0\tau_r$. Solving this equation by the method of characteristics, we find the implicit expression for the form of the evolving perturbance among the initial variables, which is valid until the moment t_{cr} of the shock-front formation:

$$t - x/c_0 - \tau_r Z[1 - \exp(x/\delta)] = R^{-1}(B^{-1}Z\exp(x/\delta)). \quad (3)$$

For $t > t_{cr}$, the localized field perturbance of an arbitrary initial shape degenerates gradually into a moving pulse of the triangular shape with the parameter density distribution

$$Z \sim [(x_F(t) - x)\mathrm{sign}c_0]^{-1/2},$$

where $x_F(t)$ is the pulse-front coordinate [4], i.e., represents the distribution of the desired type, which yields $F(\omega) \sim \omega^{-1}$. The general condition $(Z'_x)_{t = t_{cr}} = \infty$ for the onset of the formation of the shock front is specified for (3) in the form of the implicit limitation from below for the amplitude of initial fluctuations, which depends essentially on their shape:

$$(R')_{\max} > R' = [B\tau_r(1 - \exp(-x/\delta))]^{-1} > (B\tau_r)^{-1}. \quad (4)$$

We determine the explicit form of this condition on the basis of a standard approximation for $R(t)$ as a sum of exponential curves (de Prony approximation) and assumptions that their exponents are real and the initial fluctuations have the Markovian behavior. The first condition implies that only aperiodic motions take part in the formation of these fluctuations. The second condition implies that the medium is close to the equilibrium state, which leads to the fact that the function $R(t)$ is even.

The condition of smoothness for $R(t)$ requires the presence of at least two exponentials in its approximation. Thus, we arrive at the simplest model shape of initial fluctuations corresponding to all the conditions:

$$R(t) = R(0)(1 - p)^{-1}[\exp(-|t|/\tau_u) - p\exp(-|t|/\tau_s)], \quad p = \tau_s/\tau_u \quad (5)$$

(for the physical sense of τ_s and τ_r see below).

The maximum value $(R'(t))_{\max}$ is attained for $t_m = \tau_s \ln p$ and is $(R'(t))_{\max} = R(0)(1 - p\ln p)/\tau_u$. Since $p\ln p \longrightarrow 0$ for $p \longrightarrow 0$, the value of $R'(t_m)$ is independent of τ_s for reasonably small values of p and is equal to $R(0)/\tau_u$. From here, we find the critical amplitude $A = R(0) = Z(0)/B$ for fluctuations of the field $\rho(x, t)$:

$$A = \tau_u/B\tau_r. \quad (6)$$

Thus, this is the presence of the limitations of type (4) and (6) which leads to the fact, that the fluctuations capable of forming the shock compressions, are related to the most significant and, therefore, rare ones. Hence, they form the time sequence close to the pulsed-Poisson process [1]. This fact confirms the physical explanation for the form of the process $S(x, t)$ and, together with it, the entire picture of generating the 1/*f* noise, which was presented above.

We note that the most general consideration of the evolution of initial perturbances with the conservation of the gradient term in the definition of the function $q(x, t)$ (i.e., for $\xi \neq 0$) leads to the same qualitative conclusions. In the important particular case of $I(\rho) = 0$, equation (1) transforms into the Bürgers equation [4]: $c'_t + cc'_x = \xi c''_{xx}$, where $c = Q'(\rho)$. The difference from case (2) consists here only in the method of determining the critical amplitude. As far as the dynamics of acoustic processes and transport fluxes is successfully simulated by the Bürgers equation [4], the above arguments offer a physical explanation for the 1/*f* noise generation in these processes.

We now extend the theory under discussion to conducting systems. According to this theory, the fluxes inducing the 1/*f* noise are the diffusion fluxes of conduction electrons. Fluctuations of the electron density are efficiently suppressed by the Coulomb screening already at distances on the order of interatomic ones [6]. Therefore, the generation of electron fluxes of the desired type by an electron subsystem as itself is impossible. However, it is possible as a consequence of macroscopic-fluctuation motions of the shock type in the phonon subsystem. Their principal feasibility is provided by the equations of phonon hydrodynamics [7] describing the evolution of phonon wave packets:

$$\frac{\partial P}{\partial t} + S\frac{\partial(\delta T)}{\partial x} + \frac{P}{\tau_r} = 0, \quad C\frac{\partial(\delta T)}{\partial t} + ST\frac{\partial V}{\partial x} = 0, \quad P = \eta V. \quad (7)$$

Here, $\delta T(x, t)$, $P(x, t)$, and $V(x, t)$ are the local quasi-equilibrium and, therefore, functionally related fields of temperature perturbations, averaged phonon quasi-momentum, and averaged drift velocity of phonons, respectively. The quantities T, C, and S are the local magnitudes of temperature, specific heat, and entropy density; the definition of η is given in [7].

We prove this statement reducing (7) to (2) by the iterative method. In the zero approximation with respect to τ_r^{-1}, we ignore the relaxation term P/τ_r and, integrating the simplified set of equations by the method used in [8], while solving the set of gas-dynamic equations, we find its solution in the form of the Riemannian simple wave $\delta T = g(x \pm tc_T(\delta T))$ [g is

the arbitrary smooth function, $c_T = \sqrt{TS^2/\eta C} \simeq c_T^{(0)}(1 + Z)$, $Z = B\delta T$]. The differential bonds

$$\left(\frac{d(\delta T)}{dP}\right)^2 \frac{dP}{dV} = \left(\frac{d(\delta T)}{dV}\right)^2 \frac{dV}{dP} = \frac{T}{C}$$

between the fields δT, V, and P are transformed into the relationships

$$\delta T = V\sqrt{\frac{\eta T}{C}} = P\sqrt{\frac{T}{\eta C}}$$

under assumption of smallness of their amplitudes. Substituting the relations found into the first equation of set (7), we obtain relationship (2).

Next, we describe the mechanism for formation of the large-scale fluctuations $\delta T(t) = \delta T(0, t)$, which define the boundary condition in (2), by the example of a uniform metal steady-state system with the electric current. The basis of this mechanism is assumed to be the equilibrium fluctuations Δn of the phonon density $n = \bar{n} + \Delta n$ stemmed by the heat exchange of a sample with a thermostat and causing the fluctuations ΔI_e and ΔV_e of both the mean free path l_e and the drift velocity $V_e = \bar{V}_e + \Delta V_e$ for the conduction electrons. Since $l_e \sim n^{-1}$ and $V_e \sim l_e$, then $\overline{\Delta V_e^2}/\bar{V}_e^2 = \overline{\Delta l_e^2}/\bar{l}_e^2 = \overline{\Delta n^2}/\bar{n}^2 = E$ and $\bar{V}_e = \bar{J}/(n_e e)$ (where $\bar{J}$ is the mean current density, and n_e is the density of conduction electrons). The process of development of the large-scale fluctuations $\delta V_e(t)$ caused by them can be described by the Langevin equation [10]

$$\delta\dot{V}_e = -\frac{\delta V_e}{\tau_u} + \Gamma(t),$$

where $\Gamma(t)$ is the "Langevin force" with the exponential correlation function

$$K_\Gamma(t) = \left(\frac{\overline{\Delta V_e^2}}{\tau_s}\right)\exp\left(-\frac{|t|}{\tau_s}\right),$$

corresponding to the Markovian behavior of the equilibrium fluctuations. Here, τ_s is the time of the electron mean-free path between collisions with phonons, and τ_u is the relaxation time of the parameter V_e, which is on the order of the relaxation time for the total momentum of the electron flux. Solving this equation under conditions of the Langevin approach $\tau_s/\tau_u \ll 1$ (which is valid provided that the predominant interaction mechanism for electrons and phonons is the electron–phonon scattering [6]), we find the correlation function $K_{V_e}(t)$ for the parameter V_e:

$$K_{V_e}(t) = \bar{V}_e^2 E\left(\frac{\tau_u}{\tau_s}\right)\left[1 - \left(\frac{\tau_s}{\tau_u}\right)^2\right]^{-1}$$
$$\times\left[\exp\left(-\frac{|t|}{\tau_u}\right) - \left(\frac{\tau_s}{\tau_u}\right)\exp\left(-\frac{|t|}{\tau_s}\right)\right]$$

which is of type (5), whose form is repeated by the process $\delta V_e(t)$ [9].

Since in a metal system, $\tau_u \gg \tau_e > \tau_s$ (τ_e is the relaxation time of the electron quasi-momentum in the N-processes [6]) and $|\delta V_e| \ll |\bar{V}_e| \ll c_s \ll c_F$ (c_F is the Fermi velocity), the carrying away of the phonon subsystem by slow perturbances in the electron subsystem proceeds adiabatically and is virtually complete. This fact causes the conversion of fluctuations $\delta V_e(t)$ into perturbances of the phonon drift velocity $V(t)$ with the approximate conservation of their shape and amplitude ($|\delta V_e| \approx |V|$). By these means, the field $V(t)$ of large-scale fluctuations [with $\delta T(t) = \delta T(V)$ and the correlation functions $K_V(t) \simeq K_{V_e}(t)$ and $K_T(t) = (\eta T/C)K_V(t)$] is formed in the phonon subsystem in the neighborhood of the point $x = 0$. The further evolution of these functions leads to formation of abrupt fronts in the field $\delta T(x, t)$. They, in turn, induce the thermal excitation of electron fluxes $\delta J_e(x, t)$, which was observed in experiments.

We now estimate the amplitude $F_{\delta J}(\omega)$ of the $1/f$ spectrum for the fluxes δJ_e. Since they are of the thermoelectric nature [6], then

$$\delta J_e = Q_T\sigma\frac{\partial T}{\partial x} = \left(Q_T\frac{\sigma C}{\kappa}\right)J_T.$$

Here, $J_T = -\kappa_D\left(\frac{\partial T}{\partial x}\right)$, $\kappa_D = \kappa/C$, Q_T is the differential thermal electromotive force, σ is the electrical conductivity, and κ is the thermal conductivity. Substituting the general expression $F(\omega)$ from [1] for fluxes, we obtain

$$F_{\delta J}(\omega) = \pi(Q_T\sigma)^2 C\overline{a^2}\bar{\nu}/(\kappa\omega).$$

Here, the parameter $\bar{\nu}$ is the mean number of maxima for the random process $\delta\dot{T}(t)$, which are located above critical level (4) in a unit volume. This number is equal to

$$L^{-1}\int_{b}^{\infty}\int_{-\infty}^{0}|x_3|w(x_1, 0, x_3)dx_1dx_3,$$

where $w(x_1, x_2, x_3)$ is the common density of the distribution $\delta\dot{T}$, $\delta\ddot{T}$, $\delta\dddot{T}$ [11]. According to (5), $b = (B\tau_r)^{-1}$,

τ_r is the characteristic attenuation time $\delta T(x, t)$ for waves, which is on the order of time for attaining the thermal equilibrium with a thermostat. Owing to the steadiness and the Gaussian behavior of $\delta T(t)$, we have $w(x_1, x_2, x_3) = w(x_1, x_3)w(x_2)$, where $w(x_1, x_3)$, $w(x_2)$ are the two-dimensional and one-dimensional Gaussian densities of the distribution, respectively, with the variances $d_i^2 = (-1)^i K_T^{(2i)}(0)$, $i = 1, 2, 3$, and the correlation coefficient

$$-\gamma = \frac{\langle d\dot{T} d\dddot{T}\rangle}{d_1 d_3} = -\frac{\langle d\ddot{T}^2\rangle}{d_1 d_3} = -\frac{d_2^2}{d_1 d_3}.$$

From here, $w(0) = (2\pi d_2^2)^{-1/2}$, $d_1^2 = d_0^2/(\tau_u\tau_s)$, where $d_0^2 = D[\delta T] = \bar{V}_e^2 E\eta T C^{-1}(\tau_u/\tau_s)$. For a low attenuation of waves determined by the condition

$$b' = \frac{b}{d_1\sqrt{2(1-\gamma^2)}} \ll 1,$$

we obtain $\bar{\nu} = (4\pi)^{-1}(d_3/d_2)(1-\gamma)$. It is impossible to use the value $K_T(t)$ found for estimating d_3/d_2 and γ due to the absence of the multiple differentiability of the process $\delta T(t)$ with such a correlation function [11]. Therefore, we represent $K_T(t)$ in the de Prony general form:

$$K_T(t) = \sum_{i=1}^{m} a_i \exp\left(-\frac{|t|}{\tau_i}\right), \quad m > 4,$$

where $\tau_s = \tau_1 \ll \tau_2 \ll \ldots \ll \tau_{m-1} \ll \tau_m = \tau_u$. Thus, $d_3/d_2 \sim \tau_s^{-1}$ and $\gamma \sim 1$. According to the similarity method $\gamma = \gamma(\tau_s/\tau_u, \tau_2/\tau_u, \ldots, \tau_{m-1}/\tau_u)$, with $\gamma \simeq \gamma(\tau_s/\tau_u)$, as far as just the parameters τ_s and τ_u dominate in the definition of $\delta T(t)$. Since $\tau_s/\tau_u \ll 1$, $\gamma \simeq 1 - a(\tau_s/\tau_u)$ and $\gamma \sim 1$, then this implies also that $a \sim 1$. Thus, for $b' \ll 1$ $\bar{\nu} \sim (4\pi L\tau_u)^{-1}$ and is independent of $\bar{V}_e$ and $\bar{J}$.

The parameter

$$\overline{a^2} = \int_A^\infty (\delta T)^2 \exp\left[-\frac{(\delta T)^2}{2d_0^2}\right] d(\delta T)$$

is the mean power of a pulse in the $S(x, t)$ process. In the case of a large amplitude of fluctuations $\delta T(t)$ $b'' = A/d_0 \ll 1$. Hence, it follows that $\overline{a^2} \sim d_0^2/2$. Taking into account that $\sigma \sim n_e e^2\tau_e/m_e$, $N_e = n_e L$ is the total number of conduction electrons in a sample of the unit cross section, we obtain for $b' \ll 1$ and $b'' \ll 1$ the 1/f noise spectral density equal to $F_{\delta J}(f)$ ($f = \omega/2\pi$) and its desired amplitude α_H in the Hooge form:

$$\frac{F_{\delta J}(f)}{\bar{J}^2} = \frac{\alpha_H}{N_e f}, \quad \alpha_H \sim 10^{-2} TE\eta m_e^{-1}\frac{Q_T^2}{\kappa}\sigma\frac{\tau_e}{\tau_s}. \qquad (8)$$

From the meaning of physical parameters in (8), it follows that the Hooge coefficient α_H is independent of N_e in correspondence with the Hooge empirical law [12].

In an actual conducting system, the values of E, η, and B depend essentially on a particular form of the density for system phonon states [7, 9]. At the same time, the values τ_s and τ_e depend on the parameters of various and mutually interfering scattering mechanisms [6]. This fact makes it extremely difficult to obtain a rigorous estimate of α_H. Nevertheless, at not too low temperatures $T \geq 0.1\theta_D$, the standard estimates lead to

$$\tau_e/\tau_s \sim 1, \quad \sigma T/\kappa \sim e^2, \quad \eta \sim CT/c_s^2,$$

$$E \sim \bar{n}^{-1} \sim (N_i T/\theta_D)^{-1}, \quad C \sim N_i, \quad Q_T \sim T/e\varepsilon_F,$$

$M_i c_s^2 \sim \theta_D$ (N_i is the concentration of ions in the lattice, M_i is the ion mass, c_s is the sonic velocity, θ_D is the Debye temperature, and ε_F is the Fermi energy.) These estimates obtained in the framework of the Debye model for a phonon subsystem [7] and gas model of the electron subsystem [6] yield the value

$$\alpha_H \sim 10^{-2}\frac{M_i}{m_e}\left(\frac{T}{\varepsilon_F}\right)^2,$$

which agree quite well with experimental values $\alpha_H^{(exp)} \sim 10^{-4}$–$10^{-2}$ [12].

The estimates for values b' and b'' [$b' \sim b'' = (d_0 b)^{-1}(\tau_u/\tau_r)$)] is extremely difficult due to the same reasons. However, because of $\tau_u/\tau_r \ll 1$, the conditions $b' \ll 1$ and $b'' \ll 1$ for the derivation of formula (8), which correspond to the experiment [12], are likely not strict in actual conducting systems and can be easily satisfied.

This fact explains the dependence $F_{\delta j}(f) \sim \bar{J}^2$ and its independence of the value of τ_r that experiences an influence of particular experimental conditions and the state of the sample surface. This theory presents also the derivation of both the Hooge law and interval for the variation of the exponent α strictly corresponding to the experiment [12]. As a result, this theory yields the justified theoretical derivation of basic empirical regularities established for the 1/f noise.

Thus, in accordance with the data of [1–3] and the results of this study, there exists a general physical mechanism responsible for occurrence of the 1/f noise in the case of observation of various transport phenomena involving acoustic, transport, and electrical ones.

This mechanism consists in the transformation of the form of large-scale fluctuations of a macroscopic parameter characterizing the transport process into an essentially inhomogeneous form of the shock-wave type with the subsequent generation of diffusion fluxes corresponding to this parameter. The $1/f$ noise, in itself, is not the physical process. It is a random signal formed in the course of statistical measurements for these fluxes.

REFERENCES

1. A. G. Budarin, Dokl. Akad. Nauk **359**, 615 (1998) [Dokl. Phys. **43**, 218 (1998)].
2. A. G. Budarin, Radiotekhnika, No. 11, 8 (1997).
3. A. G. Budarin, in *Proceedings of the I International Conference "Digital Signal Processing and Its Application"* (Moscow, 1998), Vol. 2, pp. 203–205.
4. G. B. Whitham, *Linear and Nonlinear Waves* (Wiley, New York, 1974; Mir, Moscow, 1977).
5. Ya. B. Zel'dovich and A. D. Myshkis, *Principles of Mathematical Physics* (Nauka, Moscow, 1973).
6. A. A. Abrikosov, *Fundamentals of the Theory of Metals* (Nauka, Moscow, 1987; North-Holland, Amsterdam, 1988).
7. V. L. Gurevich, *Kinetics of Phonon Systems* (Nauka, Moscow, 1981).
8. L. D. Landau and E. M. Lifshitz, *Hydrodynamics* (Nauka, Moscow, 1986).
9. L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Nauka, Moscow, 1976; Pergamon Press, Oxford, 1980).
10. S. M. Rytov, *Introduction to Statistical Radiophysics* (Nauka, Moscow, 1976), Chap. 1. Random Processes.
11. B. R. Levin, *Theoretical Grounds of Statistical Radioengineering* (Radio i Svyaz', Moscow, 1989).
12. Sh. M. Kogan, Usp. Fiz. Nauk **145**, 285 (1985) [Sov. Phys. Usp. **28**, 170 (1985)].

Translated by V. Bukhanov

Doklady Physics, Vol. 45, No. 5, 2000, pp. 213–215. Translated from Doklady Akademii Nauk, Vol. 372, No. 1, 2000, pp. 43–45.

MECHANICS

On the Energy Condition for Fracture of Solids

I. M. Dunaev* and V. I. Dunaev**

Presented by Academician V.A. Babeshko on May 26, 1999

Received June 29, 1999

In [1], A. Griffith suggested the following energy condition for determining critical stresses and sizes of defects formed as a result of single loading solids:

$$-dU_\mathrm{p} + \gamma d\Sigma = dA. \tag{1}$$

Here, $U_\mathrm{p} = U_\mathrm{p}^{(1)} - U_\mathrm{p}^{(0)}$ is a change in the potential energy for the defect formation; $U_\mathrm{p}^{(1)}$ and $U_\mathrm{p}^{(0)}$ are the potential energies of a solid with and without the defect, respectively; γ is the specific energy for the formation of a unit of the defect surface Σ; and A is the work of external forces. In problems of uniaxial and biaxial tension (compression) of a plane with a defect, the critical stresses determined by condition (1) have usually the same value for tension and compression. However, this conclusion contradicts to experimental data. In this paper, we suggest another energy condition similar to that of Griffith but differing from it, namely,

$$-dU + \gamma d\Sigma = dA. \tag{2}$$

Here, $U = U_\mathrm{p} + U_\mathrm{k}$ is the change in the internal energy, and $U_\mathrm{k} = U_\mathrm{k}^{(1)} - U_\mathrm{k}^{(0)}$ is the change in the heat component of the internal energy. Condition (2) represents the first law of thermodynamics, while condition (1) follows from (2) in the case of ignoring the energy U_k in it. For isothermal deformation of solids, which obeys Hooke's law, components of the internal energy U have the form [2]

$$\begin{aligned} U_\mathrm{p} &= \frac{1}{2}\int\limits_{V_1} \sigma_{ij}^{(1)}\varepsilon_{ij}^{(1)}dV - \frac{1}{2}\int\limits_{V_0} \sigma_{ij}^{(0)}\varepsilon_{ij}^{(0)}dV, \\ U_\mathrm{k} &= \alpha_0 T_0\left[\int\limits_{V_1} \sigma_{ij}^{(1)}\delta_{ij}dV - \int\limits_{V_0} \sigma_{ij}^{(0)}\delta_{ij}dV\right], \\ & i, j = 1, 2, 3, \end{aligned} \tag{3}$$

where σ_{ij} and ε_{ij} are the stress-tensor and strain-tensor components, respectively; δ_{ij} is the Kronecker delta; V_1 and V_0 are the volumes of the solid with and without the defect, respectively; α_0 is the coefficient of linear thermal expansion; and T_0 is temperature. We now analyze conditions (1) and (2) for the fragile strength in the case of two models of isolated defects. In the first model (A), the stresses on the exterior surface of the solid are given, whereas in the second model (B), the displacements are given, which correspond to stresses applied prior to formation of a defect. In both the models (A) and (B), stresses on the defect surface are zeros. In the first (A) model, the external forces do work on the exterior defect boundary, whereas in the (B) model, the work of external forces equals zero. Consider the integrals comprising U_k in (3) for the models (A) and (B) in the case of the plane problem of elasticity theory:

$$\begin{aligned} U_\mathrm{k} = \alpha_0 T_0 \chi\Bigg[&\iint\limits_{D_1}\left(\sigma_x^{(1)} + \sigma_y^{(1)}\right)dxdy \\ &- \iint\limits_{D_0}\left(\sigma_x^{(0)} + \sigma_y^{(0)}\right)\Bigg]dxdy. \end{aligned} \tag{4}$$

Here, σ_x and σ_y are the stress components in a Cartesian coordinate system; $\chi = 1$ or $(1 + \nu)$ in the cases of a plane stressed state and plane deformation, respectively; ν is the Poisson's ratio; and D_1 and D_0 are the domains occupied by the solid with and without a defect, respectively. Stresses σ_x and σ_y and the boundary condition in the first boundary value problem, which are expressed in terms of the Airy stress function [3], have the form

$$\sigma_x = \frac{\partial^2 F}{\partial y^2}, \quad \sigma_y = \frac{\partial^2 F}{\partial x^2}, \tag{5}$$

$$\begin{aligned} F_x(s) &= \frac{\partial F}{\partial x} = -\int\limits_0^s \sigma_{ny}(s)ds + C, \\ F_y(s) &= \frac{\partial F}{\partial y} = \int\limits_0^s \sigma_{nx}(s)ds + D, \end{aligned} \tag{6}$$

where $(x, y) \in S = \Gamma_0 + \Sigma$ (Γ_0 is the outer contour of the solid); σ_{nx} and σ_{ny} are the stress components on the con-

* *Kuban State Technological University, ul. Moskovskaya 2, Krasnodar, 350072 Russia*

** *Kuban State University, ul. Stavropol'skaya 149, Krasnodar, 350072 Russia*

tour S; n is an outward normal to the contour; and C and D are arbitrary constants. We now make use of the fact of the validity in the model (A) the boundary conditions $\sigma_{nx}^{(1)} = \sigma_{nx}^{(0)}$, $\sigma_{ny}^{(1)} = \sigma_{ny}^{(0)}$ on Γ_0, and $\sigma_{ny}^{(1)} = \sigma_{ny}^{(0)} = 0$ on Σ, and the equality $\oint_\Sigma Cdx + Ddy = 0$. Thus, after substituting (5) into (4) with allowance for relations (6) and using the Green's formulas, we obtain $U_k = 0$. Hence, the model (A) leads to a Griffith condition (1). Taking (3) into account, we can write out the internal-energy increment in the form

$$U = \frac{1}{2}\oint_{\Gamma_0+\Sigma}\left(\sigma_{ij}^{(1)} + \sigma_{ij}^{(0)}\right)\left(u_i^{(1)} - u_i^{(0)}\right)n_j ds + \alpha_0 T_0 k_1 \oint_{\Gamma_0+\Sigma}\left(u_i^{(1)} - u_i^{(0)}\right)\delta_{ij} n_j ds - \frac{1}{2}\int_V \left[\sigma_{ij}^{(0)}\varepsilon_{ij}^{(0)} + 2\alpha_0 T_0 \chi \sigma_{ij}^{(0)}\delta_{ij}\right]dV, \quad V = V_0 - V_1, \quad i, j = 1, 2, \tag{7}$$

where $k_1 = E/(1-\nu)$ or $E/(1-2\nu)$ for the plane-stressed state and the plane deformation, respectively; E is the Young's modulus; u_i are the displacement vector components; n_i are direction cosines. Expression (7) was derived with the use of Betti's reciprocal theorem, Hooke's law, Green's formula, and the equation of equilibrium. Since in the (B) model, $u_i^{(1)} = u_i^{(0)}$ on Γ_0 and $\sigma_{ij}^{(1)} n_j = 0$ on Σ, expression (7) is reduced to

$$U = -\frac{1}{2}\oint_\Sigma \sigma_{ij}^{(0)} u_i^{(1)} n_j ds - \alpha_0 T_0 k_1 \oint_\Sigma u_i^{(1)}\delta_{ij} n_j ds. \tag{8}$$

Here, $U_k \neq 0$ and, hence, the model (B) leads to condition (2).

We now consider a problem of the plate fracture due to formation of a defect with an elliptic cross section under the action of principal tensile and/or compressing stresses P_1, and P_2. Let the P_1 stress direction form an angle α with the xO-axis, and let the ellipse principal semiaxes be denoted by a and b, $a \geq b$. Then, integral (8) can be represented in the following complex form:

$$U = \frac{1}{4}\mathrm{Re}\left\{\frac{1}{i}\oint_\Sigma\left[\left(u^{(1)} + iv^{(1)}\right)\left(\sigma_y^{(0)} - \sigma_x^{(0)} + 2i\sigma_{xy}^{(0)}\right) - \left(\overline{u^{(1)} + iv^{(1)}}\right)\left(\sigma_x^{(0)} + \sigma_y^{(0)}\right)\right]dz\right\} + \alpha_0 T_0 k_1 \mathrm{Re}\left[i\oint_\Sigma\left(\overline{u^{(1)} + iv^{(1)}}\right)dz\right], \tag{9}$$

where stresses σ_x, σ_y, and σ_{xy} and displacements u and v are calculated according to the formulas of [3]

$$2\mu(u + iv) = k\varphi(z) - z\overline{\varphi'(z)} - \overline{\psi(z)}, \quad \sigma_x + \sigma_y = 2[\varphi'(z) + \overline{\varphi'(z)}], \quad \sigma_y - \sigma_x + 2i\sigma_{xy} = 2[\bar{z}\varphi''(z) + \psi'(z)] \tag{10}$$

with relevant indices; $k = 3 - 4\nu$ and $(3-\nu)/(1+\nu)$ for the plane deformation and the plane stressed state, respectively; and $\mu = E/2(1+\nu)$. In the plane case and with allowance for (10), we substitute into (9) the functions from [3]

$$\varphi_0(z) = \frac{P_1 + P_2}{4}z, \quad \psi_0(z) = \frac{P_1 - P_2}{2}ze^{-2i\alpha},$$

the boundary condition $\varphi_1(z) + z\overline{\varphi_1'(z)} + \overline{\psi_1(z)} = 0$ on the ellipse contour; and the function [3]

$$z = \omega(\xi) = c\left(\xi + \frac{m}{\xi}\right), \quad c = \frac{a+b}{2}, \quad m = \frac{a-b}{a+b},$$

that maps the plane with an elliptic hole onto the plane $|\xi| > 1$ with a round hole. We also use the function $\varphi_1(\xi)$, which, in this case, is the same for problems being solved within models (A) and (B) [3],

$$\varphi_1(\xi) = c\frac{P_1 + P_2}{4}\left(\xi - \frac{m}{\xi}\right) + \frac{P_1 - P_2}{2}\frac{c}{\xi}e^{2i\alpha}.$$

Then, taking into account that

$$\gamma\Sigma = 4\gamma\alpha E(q), \quad E(q) = \int_0^{\pi/2}(1 - q^2\sin^2\beta)^{1/2}d\beta, \quad q^2 = 1 - \frac{b^2}{a^2},$$

we integrate (9) along the circle $|\xi| = 1$ and, as a result, obtain the expression for the increment of the total energy

$$W(a, b) = -U + \gamma\Sigma = -\frac{(k+1)\pi c^2}{4\mu}\left[\frac{(P_1 + P_2)^2}{4}(1 + m^2) - (P_1^2 - P_2^2)m\cos 2\alpha + \frac{(P_1 - P_2)^2}{2}\right] - \frac{\alpha_0 T_0 k_1 (k+1)\pi c^2}{2\mu}\left[\frac{P_1 + P_2}{2}(1 + m^2) - (P_1 - P_2)m\cos 2\alpha\right] + 4\gamma a E(q). \tag{11}$$

Substituting (11) in (2) and considering that $dA = 0$ in the model (B), we can obtain the macroscopic criterion for the fragile fracture as a function of the physicome-

chanical parameters γ, α_0, μ, ν of the material and the temperature T_0 for an arbitrary combination of critical stresses P_1 and P_2, and sizes a, b, as well as the location (α) of the defect. For example, in the case of $m = 0$ ($a = b = a_0$) or $m = 1$ ($b = 0$), we obtain from (11) and (2) the criterion of the fragile fracture for a plate having a round defect with the radius a_0 or a cut with a half-length a, respectively,

$$2(P_1^2 + P_2^2) + (P_1 - P_2)^2 + 4\alpha_0 T_0 k_1 (P_1 + P_2) - \frac{16\mu}{k+1}\frac{\gamma}{a_0} = 0,$$

$$P_1^2 + P_2^2 - (P_1^2 - P_2^2)\cos 2\alpha + 2\alpha_0 T_0 k_1 [(P_1 + P_2) - (P_1 - P_2)\cos 2\alpha] - \frac{32\mu}{\pi(k+1)}\frac{\gamma}{a} = 0.$$

In the particular case of $P_1 = P_2 = P$, these expressions make it possible to obtain the critical stresses for biaxial tension P^+ and biaxial compression P^-,

$$P^{\pm} = -\alpha_0 T_0 k_1 \pm \sqrt{(\alpha_0 T_0 k_1)^2 + \frac{4\mu}{k+1}\frac{\gamma}{a_0}},$$

$$P^{\pm} = -\alpha_0 T_0 k_1 \pm \sqrt{(\alpha_0 T_0 k_1)^2 + \frac{16\mu}{\pi(k+1)}\frac{\gamma}{a}}.$$

We obtain also the formula $P^+ + P^- = -2\alpha_0 T_0 k_1$, which is valid for arbitrary m and α.

REFERENCES

1. A. Griffith, in *Proceedings of the I International Congress for Applied Mechanics, Brussels, 1924*, Vol. 8, pp. 51–58.
2. A. A. Il'yushin, *Continuum Mechanics* (Mosk. Gos. Univ., Moscow, 1990).
3. N. I. Muskhelishvili, *Certain Basic Problems of Mathematical Elasticity Theory* (Nauka, Moscow, 1966).

Translated by V. Troitskiĭ

Doklady Physics, Vol. 45, No. 5, 2000, pp. 216–218. Translated from Doklady Akademii Nauk, Vol. 372, No. 1, 2000, pp. 46–49.

MECHANICS

Geometry of Helical Vortices in Homogeneous Ideal Fluid

A. V. Kistovich and Yu. D. Chashechkin

Presented by Academician A.Yu. Ishlinskiĭ July 14, 1999

Received July 15, 1999

The goal of this study is to investigate steady vortex flows of homogeneous ideal fluid, whose helicity, locally defined as $h = \mathbf{u} \cdot \nabla \times \mathbf{u}$ [1], is bounded. In this case, a set of constitutive hydrodynamic equations contains both Euler equations and continuity equations:

$$(\mathbf{u}\nabla)\mathbf{u} = -\nabla p, \quad \nabla \cdot \mathbf{u} = 0. \tag{1}$$

Here, $\mathbf{u}$ is the flow velocity and pressure p is normalized to the density.

For steady flows of ideal fluid, trajectories of marked particles coincide with streamlines [2] and form integral lines. The proposed geometric description is based on the assumption that the integral flow lines are geodesics at the second-order surfaces being parametrized and fill in the space occupied by the fluid. At the same time, the velocity field is formed by geodesic flows determined by these surfaces. In this case, an individual integral second-order surface is characterized by the vector function $\mathbf{f}$ and is given in a parametric form. The parameters t and τ are responsible for motion along a specific geodesic transition from one geodesic to another on the same surface. The vectors $\mathbf{f}'_t$ and $\mathbf{f}'_\tau$ are tangents and vector $[\mathbf{f}'_t \times \mathbf{f}'_\tau]$ is the normal to the integral surface. Primes denote differentiation, and indices signify the differentiation variables. For describing a transition from one integral surface to another, the variable n is used defining the vector

$$\mathbf{f}'_n = \alpha\mathbf{f}'_t + \beta\mathbf{f}'_\tau + \gamma[\mathbf{f}'_t \times \mathbf{f}'_\tau], \tag{2}$$

where the functions α, β, and γ depend on t, τ, and n.

The direction vectors $\mathbf{f}'_t$, $\mathbf{f}'_\tau$, and $\mathbf{f}'_n$ specify a curvilinear coordinate system used furthermore for describing vortex flows.

A tangential geodesic flux $\mathbf{J}$ satisfying the condition $D_\mathbf{J}\mathbf{J} = 0$ is one of the characteristics for integral surfaces formed by geodesic lines [3]. The covariant derivative D of an arbitrary vector field $\mathbf{X}$ with respect to the tangential field $\mathbf{Y}$ is calculated according to the following rules:

$$D_\mathbf{Y}\mathbf{X} = \nabla_\mathbf{Y}\mathbf{X} - \frac{\mathbf{N}(\nabla_\mathbf{Y}\mathbf{X} \cdot \mathbf{N})}{||\mathbf{N}||^2}, \quad \nabla_\mathbf{Y}\mathbf{X} \equiv (\mathbf{Y}\nabla)\mathbf{X}.$$

Here, $\mathbf{N} = \nabla F|_{\mathbf{r} \in F}$ is a field of normals to the second-order surface $F = F_0$.

Since the geodesic-flux vector $\mathbf{J}$ coincides with the vector $\mathbf{f}'_t$ tangential to a geodesic line at each point of a surface, the equation

$$(\mathbf{f}'_t\nabla)\mathbf{f}'_t - \frac{\nabla F(\nabla F \cdot (\mathbf{f}'_t\nabla)\mathbf{f}'_t)}{||\nabla F||^2} = 0 \tag{3}$$

follows from the above definition.

Using the geometric vector field $\mathbf{f}'_t$, we can construct an infinite set of equally directed physical vector fields $\mathbf{u}$,

$$\mathbf{u} = u(t, \tau, n)\mathbf{f}'_t. \tag{4}$$

These fields differ from one another by the amplitude function $u(t, \tau, n)$. Velocity-vector field (4) determines a hydrodynamic flow in the case when this field satisfies the set of equations (1).

Substituting (4) into (1) with allowance for (3), we can reduce initial system (1) to the form

$$u^2(\nabla \cdot \mathbf{f}'_t)\mathbf{f}'_t = \nabla p - \frac{\nabla F(\nabla F \cdot \nabla p)}{|\nabla F|^2}, \quad (uG)'_t = 0. \tag{5}$$

From the second equation of set (5), it follows that

$$u(t, \tau, n) = \frac{U(\tau, n)}{G}, \quad G = \mathbf{f}'_n \cdot [\mathbf{f}'_t \times \mathbf{f}'_\tau].$$

Using definition of the vorticity $\boldsymbol{\omega} = \nabla \times \mathbf{u}$ and excluding the pressure from the first equation of system (5), we obtain the equation

$$\nabla_\mathbf{u}\boldsymbol{\omega} = \nabla_\boldsymbol{\omega}\mathbf{u}. \tag{6}$$

Employing the representation $\boldsymbol{\omega} = p\mathbf{f}'_t + q\mathbf{f}'_\tau + r\mathbf{f}'_n$, where p, q, and r are functions of the variables t, τ, and

Institute of Problems in Mechanics,
Russian Academy of Sciences,
pr. Vernadskogo 101, Moscow, 117526 Russia

1028-3358/00/4505-0216$20.00

n, we can transform equation (6) to the form

$$up_t' = pu_t' + qu_\tau' + ru_n', \quad q_t' = 0, \quad r_t' = 0,$$
$$p = \frac{A_\tau^{tn} - A_n^{t\tau}}{G}, \quad q = \frac{A_n^{tt} - A_t^{tn}}{G}, \quad r = \frac{A_t^{t\tau} - A_\tau^{tt}}{G}, \quad (7)$$
$$A_k^{ij} = \frac{\partial(ug_{ij})}{\partial x_k}, \quad g_{ij} = \mathbf{f}_{x_i}' \cdot \mathbf{f}_{x_j}'.$$

Here, x_k is one of the variables t, τ, n.

Integrating (7), we transform the first equation of system (5) to the form

$$T_t' = \frac{2\varepsilon G^2 R}{g_{tt}}, \quad T = \frac{g_{tt}\xi - g_{t\tau}\varepsilon}{g_{tt}}, \quad (8)$$

where T is a twist of an individual integral line.

The elements ε and ξ of the second quadratic form of a surface are determined by the relations

$$\mathbf{f}_{ab}'' = \Gamma_{ab}^t \mathbf{f}_t' + \Gamma_{ab}^\tau \mathbf{f}_\tau' + \lambda_{ab}\mathbf{n}, \quad \mathbf{n} = \frac{[\mathbf{f}_t' \times \mathbf{f}_\tau']}{(g_{tt}g_{\tau\tau})^{1/2}},$$

where Γ_{ab}^c is the Christoffel symbol, while $\varepsilon = \lambda_{tt}$ and $\xi = \lambda_{t\tau}$. According to (8), the change in the twist for an individual particle trajectory is determined by the normal component R of the vorticity vector $\boldsymbol{\omega}$.

The class of possible solutions narrows in the case of the local coincidence of the tangential spaces of integral surfaces for the velocity field and vorticity field. This makes it possible to reduce (8) and, consequently, system (1) to the equation $T_t' = 0$, which implies invariance of the twist for an individual integral line.

Equations (8) contain only functions describing the internal geometry for flow integral surfaces. Since the incompressibility equation transforms into the form, representing the velocity field, then the initial hydrodynamic problem is reduced to the pure geometric one. We imply the problem of determining characteristics of a Riemannian space, whose internal geometry is constrained by condition (8). In this case, the boundary conditions impose restrictions on the external geometry of the flow under investigation [3].

From the smoothness condition for the operation of the mixed differentiation in the curvilinear coordinate system (t, τ, n), it follows that equation (8) must be supplemented with the Gauss and Peterson–Codazzi equations

$$R_{t,t\tau}^t = g_{t\tau}\mathcal{K}, \quad R_{t,t\tau}^\tau = -g_{tt}\mathcal{K}, \quad R_{\tau,t\tau}^t = g_{\tau\tau}\mathcal{K},$$
$$R_{\tau,t\tau}^\tau = -g_{t\tau}\mathcal{K}, \quad \varepsilon_\tau' + \varepsilon\Gamma_{\tau\tau}^\tau = \xi_t' + 2\xi\Gamma_{t\tau}^\tau, \quad (9)$$
$$\xi_\tau' + \varepsilon\Gamma_{\tau\tau}^t = \eta_t' + 2\xi\Gamma_{t\tau}^t.$$

Here, $\mathcal{K} = \varepsilon\eta - \xi^2$ is the Gaussian curvature of the surface, $R_{j,nm}^i$ are the components of the curvature tensor, and $\eta = \lambda_{\tau\tau}$.

We can pass to an orthogonal coordinate system bound to the surface (as has been done, for example, in [4]). Then, the first and second quadratic forms become diagonal simultaneously, and equations (8) and (9) acquire a more symmetric form, which simplifies solving the problem.

The new parametrization corresponding to such an orthogonal coordinate system is described by two parameters $\{a, b\}$, so that the vectors $\mathbf{f}_a'$ and $\mathbf{f}_b'$ are tangential to the integral surfaces. The orthogonal pair ($\mathbf{f}_a'$, $\mathbf{f}_b'$) for the new system can be obtained by a local rotation of the orthogonal pair of the initial system ($\mathbf{f}_t'$, $\mathbf{f}_\tau' - (g_{t\tau}/g_{tt})\mathbf{f}_t'$) through the angle $\varphi(t, \tau)$ determined by the diagonalization condition. As a result of this transformation, equations (9) take the form

$$\left(\frac{\lambda}{\sqrt{g_{aa}}}\right)_b' - \frac{\nu}{g_{bb}}(\sqrt{g_{aa}})_b' = 0,$$
$$\left(\frac{\nu}{\sqrt{g_{bb}}}\right)_a' - \frac{\lambda}{g_{aa}}(\sqrt{g_{bb}})_a' = 0, \quad (10)$$
$$\left(\frac{1}{\sqrt{g_{aa}}}(\sqrt{g_{bb}})_a'\right)_a' + \left(\frac{1}{\sqrt{g_{bb}}}(\sqrt{g_{aa}})_b'\right)_b' + \frac{\lambda\nu}{\sqrt{g_{aa}g_{bb}}} = 0$$

and (8) reduces to the equation

$$\sin(2\varphi)\left(F_a'\frac{\cos\varphi}{\sqrt{g_{aa}}} + F_b'\frac{\sin\varphi}{\sqrt{g_{bb}}}\right)$$
$$+ 2F\cos(2\varphi)\left(\frac{\varphi_a'\cos\varphi}{\sqrt{g_{aa}}} + \frac{\varphi_b'\sin\varphi}{\sqrt{g_{bb}}}\right) = 0, \quad (11)$$
$$F = \frac{\lambda}{g_{aa}} - \frac{\nu}{g_{bb}}.$$

Here, g_{aa}, g_{bb}, and λ, ν are the coefficients for the first and second quadratic forms in the new parametrization, respectively. Thus, the following relations hold:

$$\varphi_a'\cos\varphi + \varphi_b'\sin\varphi\sqrt{\frac{g_{aa}}{g_{bb}}} + \frac{\sin\varphi}{2g_{bb}}\frac{\partial g_{bb}}{\partial a}$$
$$- \frac{\cos\varphi}{2g_{aa}}\sqrt{\frac{g_{aa}}{g_{bb}}}\frac{\partial g_{aa}}{\partial b} = 0,$$
$$-\varphi_a'\sin\varphi + \varphi_b'\cos\varphi\sqrt{\frac{g_{bb}}{g_{aa}}} - \frac{\cos\varphi}{2g_{aa}}\frac{\partial g_{aa}}{\partial a} \quad (12)$$
$$- \frac{\sin\varphi}{2g_{bb}}\sqrt{\frac{g_{bb}}{g_{aa}}}\frac{\partial g_{bb}}{\partial b} = \sqrt{g_{aa}}(\ln\Delta)_t',$$

where Δ is the transformation Jacobian for passing to the new parametrization.

The right-hand side of the second equation in (12) can be reduced to zero if the additional condition of the conservation of a surface local element (the equireality condition for the performed affine transformation) is imposed.

Straightforward methods are not feasible for finding the angle φ necessary for constructing solutions to symmetrized equations (10), (11). The reason is that system (12) is reduced to the dual Riccati equations for the function Q by introducing a new variable $Q = \tan\varphi$. Although these equations always possess the solution, regular methods for its construction are not available.

In the given physical problem, additional topological flow characteristics, i.e., the density of helicity h and density of the flow parity $\wp$ (the latter depending on conditions for the generation of the helical structure), must be used for determining the local rotation angle φ. In steady-state problems, these topological quantities are related by the equation of the helicity-density conservation for Hamiltonian flows [5]

$$\nabla \cdot \mathbf{T} = \wp, \quad \mathbf{T} = h\mathbf{u} - \frac{u^2}{2}\boldsymbol{\omega},$$

which rearranges in the coordinate system (a, b) to the form

$$\frac{\cos\varphi}{\sqrt{g_{aa}}}G_a' + \frac{\sin\varphi}{\sqrt{g_{bb}}}G_b' = -\frac{G^3}{2TU^3}\wp. \qquad (13)$$

Equations (13) and (12) allow us to find an explicit expression for the angle φ providing the most compact and symmetric form

$$\varphi = \frac{\pi}{2} + \Psi_G - \Psi_\wp. \qquad (14)$$

Here, the following notation is used:

$$\cos\Psi_G = \frac{B}{\sqrt{A^2 + B^2}}, \quad \sin\Psi_G = -\frac{A}{\sqrt{A^2 + B^2}},$$

$$\cos\Psi_\wp = \sigma, \quad \sin\Psi_\wp = \sqrt{1 - \sigma^2},$$

$$A = \frac{1}{\sqrt{g_{aa}}}G_a', \quad B = \frac{1}{\sqrt{g_{bb}}}G_b',$$

$$\sigma = -\frac{G^3}{2TU^3}\frac{\wp}{\sqrt{A^2 + B^2}},$$

and, in this case, $\Psi_\wp = \frac{\pi}{2}$ for $\wp = 0$.

Relation (14) has a simple geometric interpretation. In the case of zero parity, the space of angles for internal geometry Ψ_G (the space definition involves only coefficients of the first and second quadratic forms of the surface) is orthogonal to the space of parity angles $\Psi_\wp$ (which is defined invoking topological flow characteristics). For nonzero parity, this orthogonality is broken, the spaces are projected onto each other, which results in changing the angle φ.

Employing geometric images provides a means for treating the internal flow structure as a set of integral surfaces possessing certain properties that can be specified with a reasonable degree of arbitrariness. In particular, using the conditions that integral surfaces of the velocity field are isoenergetic and integral surfaces of the vorticity field in fluid with a free boundary are isovortical, we can construct the solution in the form of a set of modified Rankine vortices [6]. These vortices, unlike the classic one [7], have bounded integral invariants, namely, the energy, momentum, angular momentum, vorticity, and helicity (here, total quantities are implied).

Reducing hydrodynamic equations to equations of differential geometry makes it possible to analyze properties of steady helical flows in a new way and construct a formalism of searching for their hidden features that do not manifest themselves immediately in studies of physical fields alone. Among these characteristics, there are the helicity density and parity density and their integral analogues, i.e., the total flow helicity and total flow parity, since only these quantities allow us to unambiguously classify different-type helical flows.

The proposed geometric approach is also applicable to the problems of helical-structure evolution in viscous media. However, in this case, the problem of helicity and parity sources arises, since in viscous media, in contrast to Hamiltonian ones, both dissipation and generation of topological flow characteristics occur.

ACKNOWLEDGMENTS

This study was supported in part by the Russian Foundation for Basic Research, project no. 99-05-64980, and by the Ministry of Education of the Russian Federation (Federal Purposeful Program “Integratsiya,” project no. 2.1-304).

REFERENCES

1. H. K. Moffat, J. Fluid Mech. **35** (1), 117 (1969).
2. G. K. Batchelor, *An Introduction to Fluid Dynamics* (Cambridge Univ. Press, Cambridge, 1967; Mir, Moscow, 1973).
3. S. P. Novikov and A. T. Fomenko, *Principles of Differential Geometry and Topology* (Nauka, Moscow, 1987).
4. B. G. Konopelchenko, J. Phys. A **30**, 437 (1997).
5. R. M. Kiehn, in *Topological Fluid Mechanics*, Ed. by H. K. Moffat and T. S. Tsinober (Cambridge Univ. Press, Cambridge, 1990), pp. 449–458.
6. A. V. Kistovich and Yu. D. Chashechkin, *Vortex and Helical Structures in Homogeneous Ideal Fluid*, Preprint No. 627, IPM RAN (Institute of Applied Mechanics, Russian Academy of Sciences, Moscow, 1998).
7. H. Lamb, *Hydrodynamics* (Cambridge Univ. Press, Cambridge, 1932, 6th ed.; Gostekhizdat, Moscow, 1947).

Translated by V. Tsarev

Doklady Physics, Vol. 45, No. 5, 2000, pp. 219–221. Translated from Doklady Akademii Nauk, Vol. 372, No. 1, 2000, pp. 50–52.

MECHANICS

On Compressing an Elastoplastic-Material Layer by Rigid Anisotropically Rough Plates

L. A. Maksimova

Presented by Academician A.Yu. Ishlinskiĭ February 15, 1999

Received June 15, 1999

The ideal-plasticity condition suggested in [1] has the form

$$\sigma_1 = \sigma_2, \quad \sigma_3 = \sigma_1 + 2\kappa, \quad \kappa = \text{const}, \tag{1}$$

where σ_i are the principal-stress components and κ is the yield stress.

According to [2], plasticity condition (1) can be written as

$$\begin{aligned} \sigma_x &= \sigma - \frac{2\kappa}{3} + 2\kappa n_1^2, \quad \tau_{xy} = 2\kappa n_1 n_2, \\ \sigma_y &= \sigma - \frac{2\kappa}{3} + 2\kappa n_2^2, \quad \tau_{yz} = 2\kappa n_2 n_3, \\ \sigma_z &= \sigma - \frac{2\kappa}{3} + 2\kappa n_3^2, \quad \tau_{xz} = 2\kappa n_1 n_3, \\ n_1^2 &+ n_2^2 + n_3^2 = 1, \end{aligned} \tag{2}$$

where σ_x, τ_{xy}, ... are the stress components in the Cartesian xyz coordinate system.

Relations (2) can be rewritten in the form

$$\begin{aligned} \sigma_x &= \sigma - \frac{2\kappa}{3} + \frac{\tau_{xy}\tau_{xz}}{\tau_{yz}}, \quad \sigma_y = \sigma - \frac{2\kappa}{3} + \frac{\tau_{xy}\tau_{yz}}{\tau_{xz}}, \\ \sigma_z &= \sigma - \frac{2\kappa}{3} + \frac{\tau_{xz}\tau_{yz}}{\tau_{xy}}, \end{aligned} \tag{3}$$

$$\frac{\tau_{xy}\tau_{xz}}{\tau_{yz}} + \frac{\tau_{xy}\tau_{yz}}{\tau_{xz}} + \frac{\tau_{xz}\tau_{yz}}{\tau_{xy}} = 2\kappa. \tag{4}$$

Following ideas of Prandtl, who has treated the plane problem on compressing a layer of an ideal plastic material by rough plates [3], we assume that

$$\begin{aligned} \tau_{xz} &= az + c_1, \quad \tau_{yz} = bz + c_2, \\ a, b, c_1, c_2 &= \text{const}. \end{aligned} \tag{5}$$

It follows from (4) and (5) that

$$\begin{aligned} \tau_{xy} &= \frac{(az + c_1)(bz + c_2)}{(az + c_1)^2 + (bz + c_2)^2} \\ &\times \left[\kappa \pm \sqrt{\kappa^2 - (az + c_1)^2 - (bz + c_2)^2}\right]. \end{aligned} \tag{6}$$

With regard to (3) and (6), we have

$$\begin{aligned} \sigma_x &= \sigma - \frac{2\kappa}{3} + \tau_{xy}\frac{az + c_1}{bz + c_2}, \\ \sigma_y &= \sigma - \frac{2\kappa}{3} + \tau_{xy}\frac{bz + c_2}{az + c_1}, \\ \sigma_z &= \sigma - \frac{2\kappa}{3} + \frac{1}{\tau_{xy}}(az + c_1)(bz + c_2). \end{aligned} \tag{7}$$

From the equations of equilibrium $\sigma_{ij,j} = 0$ and relations (5)–(7), we obtain

$$\sigma = -ax - by + C + \frac{2\kappa}{3} - \frac{\tau_{xz}\tau_{yz}}{\tau_{xy}}, \quad C = \text{const}. \tag{8}$$

Combining (7) and (8) yields

$$\begin{aligned} \sigma_x &= -ax - by + C + \frac{\tau_{xy}\tau_{xz}}{\tau_{yz}} - \frac{\tau_{xz}\tau_{yz}}{\tau_{xy}}, \\ \sigma_y &= -ax - by + C + \frac{\tau_{xy}\tau_{yz}}{\tau_{xz}} - \frac{\tau_{xz}\tau_{yz}}{\tau_{xy}}, \\ \sigma_z &= -ax - by + C, \end{aligned} \tag{9}$$

where τ_{xy}, τ_{xz}, and τ_{yz} are determined from (5) and (6).

We now consider the case of a plastic-material layer with the thickness $2h$ and assume that the x- and y-axes lie in the middle plane $z = 0$. In what follows, we pass to dimensionless variables. Namely, we normalize the stresses to the yield stress κ and all linear dimensions—to the quantity h.

Chuvash State Pedagogical University, ul. Karla Marksa 38, Cheboksary, 428000 Russia

1028-3358/00/4505-0219$20.00

According to (5), for the upper and bottom sides of the layer, we obtain, respectively,

$$\begin{aligned} \tau_{xz}^{+} &= a + c_1, \quad \tau_{yz}^{+} = b + c_2, \quad z = 1, \\ \tau_{xz}^{-} &= -a + c_1, \quad \tau_{yz}^{-} = -b + c_2, \quad z = -1. \end{aligned} \tag{10}$$

At these sides, the tangential-stress vectors take the form

$$\mathbf{T}_1 = \tau_{xz}^{+}\mathbf{i} + \tau_{yz}^{+}\mathbf{j}, \quad \mathbf{T}_2 = \tau_{xz}^{-}\mathbf{i} + \tau_{yz}^{-}\mathbf{j}, \tag{11}$$

where $\mathbf{i}$ and $\mathbf{j}$ are the unit vectors along the x- and y-axes, respectively.

According to (10) and (11), the total tangential stresses on the layer sides are given by

$$\begin{aligned} T_1 &= \sqrt{\tau_{xz}^{+2} + \tau_{yz}^{+2}} = \sqrt{(a + c_1)^2 + (b + c_2)^2} = \kappa_1, \\ T_2 &= \sqrt{\tau_{xz}^{-2} + \tau_{yz}^{-2}} = \sqrt{(-a + c_1)^2 + (-b + c_2)^2} = \kappa_2, \end{aligned} \tag{12}$$

where κ_1 and κ_2 are dimensionless values of the total tangential stresses, with $\kappa_1 \le 1$ and $\kappa_2 \le 1$.

With regard to (10) and (11), the angle between the vectors $\mathbf{T}_1$ and $\mathbf{T}_2$, as well as the directions of these vectors are given by the relations

$$\cos\varphi = \frac{\mathbf{T}_1 \cdot \mathbf{T}_2}{T_1 T_2} = \frac{c_1^2 + c_2^2 - (a^2 + b^2)}{\kappa_1 \kappa_2}, \tag{13}$$

$$\begin{aligned} \tan\mu_1 &= \frac{\tau_{yz}^{+}}{\tau_{xz}^{+}} = \frac{b + c_2}{a + c_1}, \quad \tan\mu_2 = \frac{\tau_{yz}^{-}}{\tau_{xz}^{-}} = \frac{b - c_2}{a - c_1}, \\ \varphi &= \mu_2 - \mu_1. \end{aligned} \tag{14}$$

The values of a, b, c_1, and c_2 are being found according to the given values of κ_1, κ_2, μ_1, and μ_2.

The case $c_1 = c_2 = 0$ and $\mu_1 = \mu_2$ was considered in [4].

We now assume that, on the upper and lower plate sides, the tangential stresses attain the yield stress (i.e., $\kappa_1 = \kappa_2 = 1$) and the vector $\mathbf{T}_1$ is directed along the x-axis:

$$\mathbf{T}_1 = (a + c_1)\mathbf{i}, \quad a + c_1 = 1, \quad b + c_2 = 0. \tag{15}$$

The vector $\mathbf{T}_2$ takes the form

$$\mathbf{T}_2 = (-a + c_1)\mathbf{i} + (-b + c_2)\mathbf{j}, \quad T_2 = 1. \tag{16}$$

It follows from (15) and (16) that

$$b^2 = a(1 - a). \tag{17}$$

In the case under consideration, according to (14)–(17), we have

$$\begin{aligned} a &= \cos^2\frac{\varphi}{2}, \quad b = \sin\frac{\varphi}{2}\cos\frac{\varphi}{2}, \\ \mu_1 &= 0, \quad \varphi = -\mu_2 \end{aligned} \tag{18}$$

and, with allowance for (18) and (9),

$$\sigma_z = -\cos\frac{\varphi}{2}\left(x\cos\frac{\varphi}{2} + y\sin\frac{\varphi}{2}\right) + C. \tag{19}$$

Introducing the variables

$$\xi = x\cos\frac{\varphi}{2} + y\sin\frac{\varphi}{2}, \quad \eta = -x\sin\frac{\varphi}{2} + y\cos\frac{\varphi}{2}, \tag{20}$$

we obtain, in accordance with (19) and (20),

$$\sigma_z = -\cos\frac{\varphi}{2}\xi + C. \tag{21}$$

Thus, the pressure increases linearly along the bisectrix of the angle between the vectors $\mathbf{T}_1$ and $\mathbf{T}_2$. In this case, the following relations take also place:

$$\begin{aligned} \sigma_z &= -\xi + C, \text{ for } \varphi = 0, \\ \sigma_z &= -\frac{\sqrt{2}}{2}\xi + C, \text{ for } \varphi = \frac{\pi}{2}. \end{aligned} \tag{22}$$

To determine the constant C entering into relation (9), we assume that the plate edge $\xi = 0$ is stress-free [5], and the normal stress averaged over the layer depth, σ_ξ, vanishes:

$$\int_{-1}^{1} \sigma_\xi dz = 0. \tag{23}$$

It follows from (23), (9), and (19) that

$$2C = \int_{-1}^{1} \left[\frac{\tau_{xy}\tau_{xz}}{\tau_{yz}}\cos^2\frac{\varphi}{2} + \frac{\tau_{xy}\tau_{yz}}{\tau_{xz}}\sin^2\frac{\varphi}{2} + \tau_{xy}\sin\varphi\right] dz. \tag{24}$$

The equations defining the kinematics of plastic flow can be written as

$$\begin{aligned} \varepsilon_x + \varepsilon_{xy}\frac{n_2}{n_1} + \varepsilon_{xz}\frac{n_3}{n_1} &= \varepsilon_{xy}\frac{n_1}{n_2} + \varepsilon_y + \varepsilon_{yz}\frac{n_3}{n_2} \\ &= \varepsilon_{xz}\frac{n_1}{n_2} + \varepsilon_{yz}\frac{n_2}{n_3} + \varepsilon_z, \end{aligned} \tag{25}$$

$$\varepsilon_x + \varepsilon_y + \varepsilon_z = 0, \tag{26}$$

where ε_x, ε_{xy}, … are the components of the plastic-strain rate [2].

Using conditions (2), we can rewrite equations (25) in the form

$$\begin{aligned} \tau_{yz}\left(\frac{\varepsilon_x}{\tau_{yz}} + \frac{\varepsilon_{xy}}{\tau_{xz}} + \frac{\varepsilon_{xz}}{\tau_{xy}}\right) &= \tau_{xz}\left(\frac{\varepsilon_{xy}}{\tau_{yz}} + \frac{\varepsilon_y}{\tau_{xz}} + \frac{\varepsilon_{yz}}{\tau_{xy}}\right) \\ &= \tau_{xy}\left(\frac{\varepsilon_{xz}}{\tau_{yz}} + \frac{\varepsilon_{yz}}{\tau_{xz}} + \frac{\varepsilon_z}{\tau_{xy}}\right). \end{aligned} \tag{27}$$

Finally, we write out the components U, V, and W of

the strain rate as

$$U = p_1 x + q_1 y + u(z), \quad V = p_2 x + q_2 y + v(z), \quad W = pz. \tag{28}$$

With regard to (28), relations (27) take the form

$$\tau_{yz}\left(\frac{2p_1}{\tau_{yz}} + \frac{q_1 + p_2}{\tau_{xz}} + \frac{u'}{\tau_{xy}}\right) = \tau_{xz}\left(\frac{q_1 + p_2}{\tau_{yz}} + \frac{2q_2}{\tau_{xz}} + \frac{v'}{\tau_{xy}}\right) = \tau_{xy}\left(\frac{u'}{\tau_{yz}} + \frac{v'}{\tau_{xz}} + \frac{2p}{\tau_{xy}}\right). \tag{29}$$

In (29), the components of tangential stresses τ_{ij} are given by (5) and (6) and depend on the coordinate z. Two equations (29) define two continuous functions $u(z)$ and $v(z)$.

Thus, the components of both the stresses and strain rates can be determined from relations (5)–(7), (28), and (29).

REFERENCES

1. A. Haar and T. Karman, in *Plasticity Theory* (Inostrannaya Literatura, Moscow, 1948), pp. 41–56.
2. D. D. Ivlev, *Theory of Ideal Plasticity* (Nauka, Moscow, 1966).
3. L. O. Prandtl, in *Plasticity Theory* (Inostrannaya Literatura, Moscow, 1948).
4. D. D. Ivlev, Izv. Ross. Akad. Nauk, Mekh. Tverd. Tela, No. 1 (1998).
5. L. M. Kachanov, *Principles of Plasticity Theory* (Nauka, Moscow, 1969).

Translated by V. Chechin

Doklady Physics, Vol. 45, No. 5, 2000, pp. 222–224. Translated from Doklady Akademii Nauk, Vol. 372, No. 1, 2000, pp. 53–55.

MECHANICS

Propagation of Surface Waves in Vertically Nonuniform Elastic Media

S. G. Saakyan

Presented by Academician G.G. Chernyĭ October 19, 1998

Received November 12, 1998

Propagation of surface waves in a uniform elastic half-space was first treated by Rayleigh [1]. The paper of Rayleigh initiated a large body of theoretical and experimental studies concerning the surface-wave propagation along free boundaries of spherical layers, cylinders, plates, shells, and other elastic bodies [2–4].

In this paper, we consider free vibrations driven by the surface waves in vertically nonuniform elastic half-spaces. We derive dispersion equations for the surface waves and find their phase velocities.

1. We consider a plane wave $\{u(x, z, t), 0, w(x, z, t)\}$ propagating in a nonuniform elastic medium occupying domain $z \geq h$ in the Cartesian xyz coordinate system. If external bulk forces are absent, the equations of motion for this nonuniform elastic medium take the form

$$\frac{\partial}{\partial x}\left[\lambda\left(\frac{\partial u}{\partial x}+\frac{\partial w}{\partial z}\right)+2\mu\frac{\partial u}{\partial x}\right] + \frac{\partial}{\partial z}\left[\mu\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}\right)\right] = \rho\frac{\partial^2 u}{\partial t^2}, \tag{1.1}$$

$$\frac{\partial}{\partial x}\left[\mu\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}\right)\right] + \frac{\partial}{\partial z}\left[\lambda\left(\frac{\partial u}{\partial x}+\frac{\partial w}{\partial z}\right)+2\mu\frac{\partial w}{\partial z}\right] = \rho\frac{\partial^2 w}{\partial t^2}. \tag{1.2}$$

We assume that the stresses are zero at the boundary $z = h$ of the nonuniform elastic half-space. Using the Hooke's law, we can write the boundary conditions at $z = h$ as

$$\tau_{xz} = \mu\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}\right) = 0, \tag{1.3}$$

$$\sigma_{zz} = \lambda\left(\frac{\partial u}{\partial x}+\frac{\partial w}{\partial t}\right)+2\mu\frac{\partial w}{\partial z} = 0. \tag{1.4}$$

In general, the Lamé coefficients λ and μ and density ρ of the medium are functions of spatial coordinates. In the our case of a vertically nonuniform medium, the Lamé coefficients and the density depend only on the z-coordinate, and the Poisson's ratio σ equals 0.5.

We seek a solution for free vibrations of the vertically nonuniform elastic half-space in the form

$$u(x, z, t) = \frac{f(z)}{\sqrt{\mu(z)}}\exp[i(kx-\omega t)], \tag{1.5}$$

$$w(x, z, t) = \frac{\varphi(z)}{\sqrt{\mu(z)}}\exp[i(kx-\omega t)]. \tag{1.6}$$

This form corresponds to the sine wave with the frequency ω and wave number k propagating along the x-axis.

Substituting (1.5) and (1.6) into equations (1.1) and (1.2), we obtain the following system of differential equations for unknown functions $f(z)$ and $\varphi(z)$

$$\frac{d^2 f}{dz^2}-3n_1^2 f+2ik\frac{d\varphi}{dz} = 0, \tag{1.7}$$

$$\frac{d^2\varphi}{dz^2}-\frac{1}{3}n_2^2\varphi+\frac{2ik}{3}\frac{df}{dz} = 0. \tag{1.8}$$

Here,

$$n_1 = k^2-k_1^2+\frac{1}{3}g, \quad n_2^2 = k^2-k_2^2+3g, \tag{1.9}$$

$$k_1 = \frac{\omega}{c_d}, \quad k_2 = \frac{\omega}{c_s}, \quad c_d = \sqrt{3}c_s, \quad c_s = \frac{\mu}{\rho}, \tag{1.10}$$

$$g\left(\frac{\mu'}{\mu}\right) \equiv \left(\frac{\mu'}{2\mu}\right)' + \left(\frac{\mu'}{2\mu}\right)^2. \tag{1.11}$$

We now consider the vertically nonuniform media assuming that the following conditions are met

$$c_s = \text{const}, \quad g\left(\frac{\mu'}{\mu}\right) = \text{const}. \tag{1.12}$$

The condition $c_s = \text{const}$ implies that the functions μ/μ_0

1028-3358/00/4505-0222$20.00

and ρ/ρ_0 are identical, i.e.,

$$\frac{\mu}{\mu_0} = \frac{\rho}{\rho_0} \equiv \upsilon(z). \tag{1.13}$$

The second condition in (1.12) is a differential equation for an unknown function $\upsilon(z)$

$$g\left(\frac{\mu'}{\mu}\right) = g\left(\frac{\upsilon'}{\upsilon}\right) \equiv \left(\frac{\upsilon'}{2\upsilon}\right)' + \left(\frac{\upsilon'}{2\upsilon}\right)^2 = a. \tag{1.14}$$

Here, a is an arbitrary constant.

Solving equations (1.14) for the initial conditions $\mu(h) = \mu_0$ and $\mu'(h) = \mu_0'$, we obtain the relationships for the following vertically nonuniform media

$$\text{(A)}\quad \frac{\mu}{\mu_0} = \frac{\rho}{\rho_0} = \left[1 + \frac{\mu_0'}{2\mu_0}(z-h)\right]^2, \quad \text{for } a = 0, \tag{1.15}$$

$$\text{(B)}\quad \frac{\mu}{\mu_0} = \frac{\rho}{\rho_0} = \left[\cos\frac{z-h}{p} + \frac{\mu_0' p}{2\mu_0}\sin\frac{z-h}{p}\right]^2, \tag{1.16}$$

$$\text{for } a = -\frac{1}{p^2},$$

$$\text{(C)}\quad \frac{\mu}{\mu_0} = \frac{\rho}{\rho_0} = \left[\cosh\frac{z-h}{p} + \frac{\mu_0' p}{2\mu_0}\sinh\frac{z-h}{p}\right]^2, \tag{1.17}$$

$$\text{for } a = \frac{1}{p^2}.$$

2. According to (1.5) and (1.6), the solution to equations (1.7) and (1.8), which decreases exponentially with the distance from the free surface of the vertically nonuniform medium described by (1.15), determines the components of the displacement vector

$$u = \frac{1}{\sqrt{\mu(z)}} \times [A\exp(-zn_1) + B\exp(-zn_2)]\exp[i(kx - \omega t)], \tag{2.1}$$

$$w = \frac{1}{\sqrt{\mu(z)}} \times \left[\frac{n_2}{k}A\exp(-zn_1) + \frac{k}{n_2}B\exp(-zn_2)\right]\exp[i(kx - \omega t)], \tag{2.2}$$

where A and B are arbitrary constants. The branches n_1 and n_2 of the radicals are specified by the existence conditions for a surface wave

$$\arg n_i = 0, \text{ for } k^2 - k_1^2 > 0 \quad (i = 1, 2). \tag{2.3}$$

Substituting expressions for u and w into boundary conditions (1.3) and (1.4), we arrive at the system of linear homogeneous equations with respect to the constants A and B:

$$n_2(2n_1 + \varepsilon)A + (n^2 + \varepsilon n_2)B = 0, \tag{2.4}$$

$$n_2(n^2 + 3\varepsilon n_1)A + k^2(2n_2 + 3\varepsilon)B = 0, \tag{2.5}$$

where,

$$n^2 = 2k^2 - k_2^2, \quad \varepsilon = \frac{\mu_0'}{2\mu_0}. \tag{2.6}$$

If the nontrivial solution to equations (2.4) and (2.5) exists, the dispersion equation takes the form

$$n^4 - 4k^2 n_1 - \varepsilon(n^2 - zk^2)(3n_1 + n_2) + 3\varepsilon^2(n_1 n_2 - k^2) = 0. \tag{2.7}$$

Factorization of dispersion equation (2.7) yields

$$(N_1 N_2 - K^2)(3N_1 N_2 - K^2) = 0, \text{ for } \varepsilon = 0, \tag{2.8}$$

$$(N_1 N_2 - K^2)[(3 + 2\sqrt{3})N_1 + N_2 + (3 + \sqrt{3})H][(3 - 2\sqrt{3})N_1 + N_2 + (3 - \sqrt{3})H] = 0. \tag{2.9}$$

It follows from (2.8) and (2.9) that equation (2.7) has the unique positive solution at an arbitrary value of the parameter $\varepsilon \neq 0$.

Henceforth, we introduce the following notation

$$n_1 = k_2 N_1, \quad n_2 = k_2 N_2, \quad K = \frac{k}{k_2}, \tag{2.10}$$

$$N_1 = \sqrt{K^2 - \frac{1}{3}}, \quad N_2 = \sqrt{K^2 - 1}, \quad H = \frac{\varepsilon}{k_2}. \tag{2.11}$$

If $\varepsilon = 0$ ($\mu_0' = 0$), then the elastic medium is uniform. It follows from (2.8) that

$$c_R^\rho = c_s\sqrt{\frac{2}{3}(3 - \sqrt{3})} \approx 0.9194 c_s.$$

If $\varepsilon \neq 0$, then the elastic medium is nonuniform. In this case, the surface-wave phase velocity is determined from (2.9) and can be presented by the expression

$$\frac{c_R}{c_R^0} = \frac{\Omega\sqrt{3 + \sqrt{3}}}{2\sqrt{\frac{1}{3}\Omega^2 + \left[\frac{2 + \sqrt{3}}{4}\sqrt{\frac{4\sqrt{3} - 1}{3}\Omega^2 + 3\varepsilon^2} - \frac{3}{4}\varepsilon\right]^2}}, \tag{2.12}$$

where $\Omega = \dfrac{\omega}{c_s}$.

Expression (2.12) demonstrates that the surface wave, propagating in a vertically nonuniform medium, is characterized by the pronounced dispersion, in contrast to the case of a uniform medium. When $\varepsilon > 0$, such

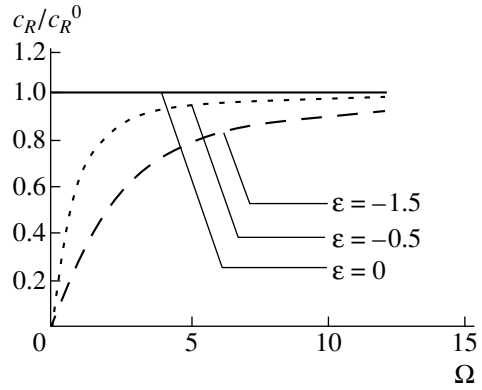


Fig 1.

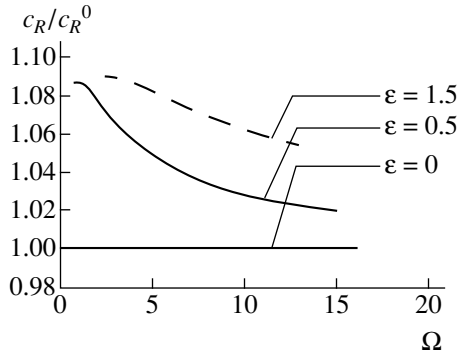


Fig. 2.

a wave can arise at frequencies $\Omega \geq \Omega_0$, where $\Omega_0 = \varepsilon\sqrt{3(2+\sqrt{3})}$. The phase velocity attains its maximum value c_s when $\Omega = \Omega_0$. Then, this velocity gradually decreases with increasing Ω and tends to c_R^0 as $\Omega \longrightarrow \infty$.

At $\varepsilon > 0$, the surface wave can arise at an arbitrary frequency. In this case, the phase velocity increases monotonically within the range $(0, +\infty)$, and c_R tends to c_R^0 at $\Omega \longrightarrow \infty$. However, in both cases, $c_R \leq c_s$.

Thus, the nonuniformity of the medium affects the phase velocity at low frequencies. At high frequencies, the surface-wave phase velocity is close to the Rayleigh-wave velocity in the uniform elastic medium.

Plots of phase velocity as a function of frequency Ω are presented in Figs. 1 and 2 for various values of ε.

REFERENCES

1. J. W. Rayleigh, Proc. London Math. Soc. **17** (253), 4 (1885).
2. P. Pfeiffer, in *Hanbbuch der Physik. Band IV. Mechanik der elastischen Körper*, Ed. by R. Grammel (Springer, Berlin, 1928; ONTI, Moscow, 1934).
3. I. A. Viktorov, *Sonic Surface Waves in Solids* (Nauka, Moscow, 1981).
4. G. A. Maugin, Adv. Appl. Mech. **23**, 373 (1983).

Translated by V. Chechin

Doklady Physics, Vol. 45, No. 5, 2000, pp. 225–228. Translated from Doklady Akademii Nauk, Vol. 372, No. 2, 2000, pp. 181–184.

MECHANICS

On Interaction of Strong Shock Waves with Weak Low-Intensity Shocks

N. A. Ostapenko

Presented by Academician G.G. Chernyĭ January 28, 2000

Received February 7, 2000

At present, the theory of interaction of shock waves propagating through perfect gas with a constant heat capacity appears to be a completely developed field of supersonic gas dynamics. For problems of interference and diffraction of shock waves, this theory provides a way of interpreting results for numerous experiments. It is performed under the assumption that in the small vicinity of branch points, flows separated by surfaces of parameter discontinuities are uniform.

However, it should be emphasized that, to a large extent, the concept of completeness of the theory has an *a posteriori* nature. This is associated with the fact that in theoretical (calculated) studies of shock-wave interaction problems, especially in the case of three-dimensional gas flows, it is frequently taken that the structure of branch points results from solving the boundary value problem as a whole. This is a possible reason why certain properties of interaction of strong shock waves with weak low-intensity shocks, which are the subject of the present study, remain uninvestigated.

Perturbances of the slope of the shock-wave and parameters behind it in the vicinity of a branch point are shown to be finite in the general case when an oblique shock of an arbitrarily low intensity impinges on a strong shock wave. If the shock-wave slope with respect to a uniform-flow velocity is equal to a certain value depending on the Mach number and the adiabatic exponent, then the perturbance of this quantity will also be arbitrarily small.

Let u, p, and ρ be the velocity, pressure, and density of a uniform flow with a Mach number M. We assume that there exists in this flow a strong shock wave (the causes for its occurrence are inessential), which is characterized by the slope θ with respect to the direction of the unperturbed velocity, and σ is the flow turning angle in the shock wave. The intensity of a weak oblique shock incident upon the strong shock wave is assumed to be small:

$$\frac{p_k}{p} - 1 = \Delta p \ll 1. \tag{1}$$

Using familiar relations and retaining only the terms linear in Δp, we can write out expressions for parameters characterizing the oblique shock and flow behind it:

$$\begin{gathered}\theta_k = \arcsin\frac{1}{\mathrm{M}} + \frac{\gamma+1}{4\gamma\sqrt{\mathrm{M}^2-1}}\Delta p + O(\Delta p^2),\\ \sigma_k = \frac{\sqrt{\mathrm{M}^2-1}}{\gamma\mathrm{M}^2}\Delta p + O(\Delta p^2),\\ \rho_k = 1 + \frac{\Delta p}{\gamma} + O(\Delta p^2), \quad u_k = 1 - \frac{\Delta p}{\gamma\mathrm{M}^2} + O(\Delta p^2),\\ \mathrm{M}_k = \mathrm{M}\left[1 - \frac{1}{\gamma\mathrm{M}^2}\left(1 + \frac{\gamma-1}{2}\mathrm{M}^2\right)\Delta p + O(\Delta p^2)\right].\end{gathered} \tag{2}$$

Here, γ is the ratio of the specific heat capacities.

We now assume that polars (heart-shaped curves), describing possible shock transitions from states in the unperturbed flow to those behind the oblique shock, intersect each other. In this case, the condition for the velocities of locally uniform gas flows to be parallel on each side of the contact discontinuity emerging from the triple point of the λ-configuration for the shock waves must hold:

$$\begin{gathered}\theta_1 - \arctan\left[\tan\theta_1\left(\varepsilon + \frac{2}{(\gamma+1)\mathrm{M}^2\sin^2\theta_1}\right)\right]\\ = \sigma_k + \theta_2 - \arctan\left[\tan\theta_2\left(\varepsilon + \frac{2}{(\gamma+1)\mathrm{M}_k^2\sin^2\theta_2}\right)\right],\\ \varepsilon = \frac{\gamma-1}{\gamma+1}.\end{gathered} \tag{3}$$

Here, θ_1 and θ_2 are slopes for the shock waves emerging from the branch point with respect to the directions of velocities of the unperturbed flow and the uniform flow

Institute of Mechanics, Moscow State University, Michurinskiĭ pr. 1, Moscow, 119899 Russia

1028-3358/00/4505-0225$20.00

behind the oblique shock, respectively. These slopes can be found from the following formulas:

$$\sin\theta_1 = \pm\frac{1}{M}\sqrt{\frac{\gamma+1}{2\gamma}\left(\frac{p_1}{p}+\varepsilon\right)},$$
$$\sin\theta_2 = \pm\frac{1}{M_k}\sqrt{\frac{\gamma+1}{2\gamma}\left(\frac{p_1}{p_k}+\varepsilon\right)}. \quad (4)$$

The "plus" and "minus" signs correspond to the position of the point of intersection for polars in the plane of the variables σ, p at $\sigma > 0$ and $\sigma < 0$, respectively. The presence of the pressure p_1 under the radical sign in equation (4) testifies to the fact that one more condition holds at the contact discontinuity: gas pressures on each side of the contact discontinuity are equal.

Substituting p_k from (1) and M_k from (2) to the second relation of (4), we find the relationship between the angles θ_2 and θ_1:

$$\sin\theta_2 = \sin\theta_1[1 + A\Delta p + O(\Delta p^2)],$$
$$A = \frac{1}{\gamma M^2}\left[1 + \frac{(\gamma-1)}{2}M^2\right] - \frac{p_1}{2(p_1+\varepsilon p)}. \quad (5)$$

Equation (5) can be written out in the form

$$\theta_2 = \theta_1 + A\tan\theta_1\Delta p + O(\Delta p^2), \quad (6)$$

which holds true at an arbitrary finite value of the coefficient $A\tan\theta_1$. Note that according to (4) and (5), we have for $\theta_1 = \pi/2$

$$A = -\frac{1}{2\gamma M^2}(M^2 - M_0^2), \quad M_0 = \sqrt{\frac{\gamma+3}{2}}. \quad (7)$$

By employing relations (2) and (6), we can recognize that terms linear in the oblique-shock intensity Δp [see (1)] are the principal ones in equation (3). Ignoring for simplicity terms on the order of ε^2 compared to unity, we can show that at arbitrarily small values of Δp, for satisfying equation (3), the angle θ_1 must be determined by the formulas

$$\cos\theta_1 = B(\sqrt{D}-1), \quad M > M_0, \quad \sigma_1 < 0,$$
$$\cos\theta_1 = B(1-\sqrt{D}), \quad M_D \le M < M_0, \quad \sigma_1 > 0,$$
$$B = a_1^{3/2}\frac{c_1}{c_2}\sqrt{M^2-1},$$
$$D = 1 - \frac{\varepsilon(\gamma+1)}{a_1^3}\left(1-\frac{2\varepsilon}{M^2}\right)\frac{c_2}{c_1^2}\frac{AM^2}{M^2-1}, \quad (8)$$
$$a_1 = \frac{2\gamma}{\gamma+1}, \quad c_1 = \frac{a_1^2}{\gamma^2 M^4}(1+\gamma a_2 M^2)^2 + a_2,$$
$$a_2 = \frac{\varepsilon}{a_1^2}, \quad c_2 = 2\left[a_1^2\left(1-\frac{\varepsilon}{M^2}\right) - \gamma A M^2\left(1-\frac{2\varepsilon}{M^2}\right)\right].$$

Here, M_D is the Mach number for an unperturbed flow for which $D = 0$. According to the calculation for air ($\gamma = 1.4$), $M_D = 1.064$. From (8), it follows that for $M < M_D$, polars do not intersect each other. In these cases, the structure of the branch point must be supplemented with a centered rarefaction wave. The relevant isentrope must be drawn from the sound point of the inner polar until it intersects the external polar ($\sigma > 0$).

As $M \longrightarrow M_0$ [see (8)] $\theta_1 \longrightarrow \pi/2$, so that the coefficient $A\tan\theta_1$ in (6) remains finite. As $M \longrightarrow \infty$, solution (8) to equation (3) indicates the existence of the asymptote: $\cot\theta_1 = -\sqrt{\varepsilon}$. The sound-flow regime is realized behind the shock wave having the indicated slope with respect to the direction of the unperturbed flow.

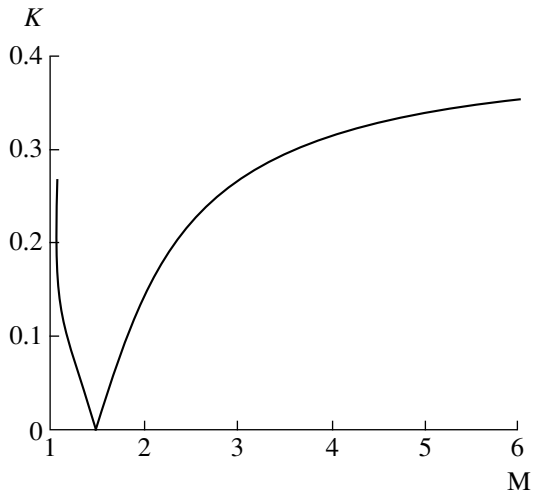


Fig. 1.

The result obtained testifies to the fact that, at an arbitrarily small Δp, perturbances of a strong shock wave (i.e., of its slope) and parameters behind it in a certain vicinity of the branch point (whose relative size can only be determined as a result of solving the boundary value problem) are finite in the general case. Indeed, if the slope of a shock wave subjected to a weak perturbance is equal to $\theta = \theta_1(M, \gamma)$ [see (8)], then, according to (6), the wave, in itself, and parameters behind it acquire perturbances on the order of $O(\Delta p)$. However, if the slope of the unperturbed shock wave is $\theta \neq \theta_1(M, \gamma)$, the slopes of the shock waves emerging from the branch point acquire finite changes for an arbitrarily small intensity of the oblique shock.

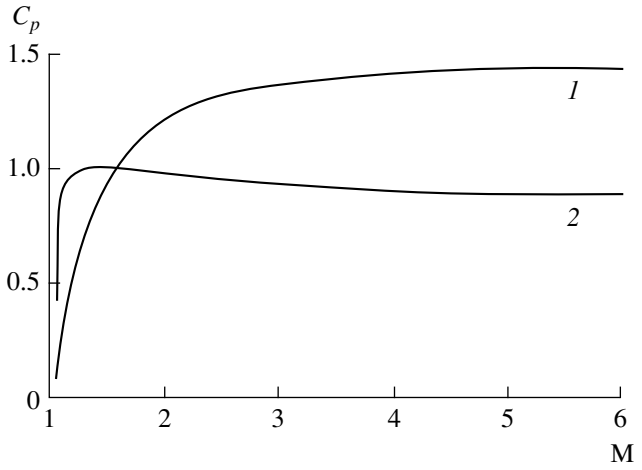


Fig. 2.

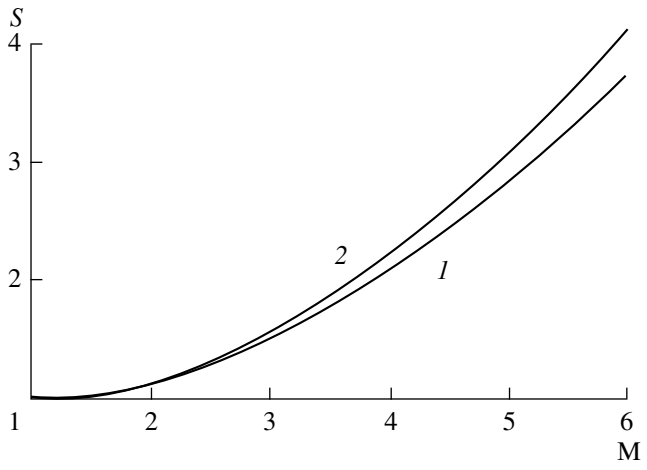


Fig. 3.

The dependence $K(M) \equiv |\cot\theta_1|$, which was calculated in accordance with (8) for $M \in [M_D, 6]$ at $\gamma = 1.4$, is shown in Fig. 1. This dependence characterizes the slope of the perturbed strong shock wave at the branch point. As is seen, the normal shock wave ($\theta = \pi/2$) in air is weakly perturbed when the Mach number of the unperturbed flow is $M = M_0$. As $M \longrightarrow \infty$, $K \longrightarrow \sqrt{\varepsilon}$.

As an illustration, for the same range of variations of the Mach-number (see Fig. 2), the pressure coefficient C_p in the region behind the shock wave with the slope θ_1 (curve *1*) is shown. The ratio of C_p to the magnitude of the pressure coefficient behind the normal shock wave for the same values of M (curve *2*) is also demonstrated. This plot allows us to gain an impression of the pressure variations behind the perturbed normal shock wave under the action of a weak shock with an arbitrarily small intensity.

We introduce the entropy function $S = (p_1/p)/(\rho_1/\rho)^\gamma$ and the coefficient of the total pressure restoration, which is equal to the ratio of the isentropically stagnant gas pressure behind the shock wave with the slope θ_1 to that ahead of the wave $K_v = p_{1t}/p_t$. These quantities are related to each other by the formula $S = K_v^{1-\gamma}$.

Figure 3 shows the function $S(M)$ for $M \in [M_D, 6]$ (curve *1*) and, for comparison, variation of the entropy function behind the shock wave (curve *2*). The function $K_v(M)$ is shown in Fig. 4 (curve *1*). Curve *2* in the same figure represents the ratio of the total pressure restoration behind the normal shock wave to that behind the wave with the slope θ_1. According to these data, at high supersonic velocities, after perturbance of the normal shock wave, the total pressure increases by more than 20%.

The function $M_1(M)$ and variation of the flow Mach number behind the normal shock wave are also shown in Fig. 4 by curves *3* and *4*, respectively. Fast divergence of the indicated curves from each other for $M > 2$ and (more than twofold) increase of the M_1 behind the perturbed shock wave compared to the value of the same parameter behind the normal shock wave for $M > 5$ provide the indicated increase by more than 20% of the restoration coefficient for the total pressure. This result may provide an explanation for the "anomalous" phenomena of the force and heat nature on external and internal (air intakes) elements of an aircraft under the conditions of a hypersonic flow, which are consequences of "weak interactions."

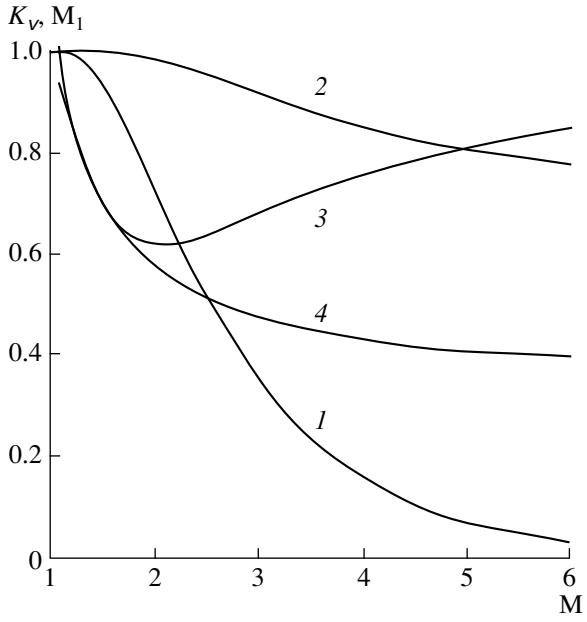


Fig. 4.

It follows from the calculations (see Fig. 4, curve *3*) that, in a small vicinity of the number $M = M_D$, the flow velocity behind the shock wave with the slope θ_1 becomes supersonic. This is associated with the fact that the solutions to equation (3) are approximated by relations (8). For $M = M_D$, the Mach number M_1 must be equal to unity. Indeed, for all $1 < M < M_D$, when the polars do not intersect each other, the isentrope emerging from the sound point of the internal polar at $\Delta p \longrightarrow 0$ will intersect the external polar at a point arbitrarily close to its sound point. Therefore, in the case of the weak interaction within the indicated range of the numbers M, the limiting slope θ_1 for the shock wave will be that for which the sound regime of the flow is realized behind the shock wave. Therefore, in the interval $[1, M_D]$, the function $K(M)$ (see Fig. 1) will increase monotonically from zero to $K(M_D)$, and $M_1 \equiv 1$.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (projects no. 97-01-00099 and 00-01-00234).

Translated by V. Tsarev

Doklady Physics, Vol. 45, No. 5, 2000, pp. 229–232. Translated from Doklady Akademii Nauk, Vol. 372, No. 2, 2000, pp. 185–188.

MECHANICS

On a Certain Feature of Non-Self-Similar Interaction for Gas-Dynamic Discontinuities

Academician G. G. Chernyĭ

Received February 7, 2000

An important problem of gas dynamics is studying flows with discontinuity surfaces intersecting each other. We imply, e.g., the intersection of a shock wave with a tangential discontinuity, in other words, with a contact discontinuity (and its limiting cases, such as a rigid wall and a free surface) or with another shock wave. For one-dimensional unsteady flows, a relatively simple local problem of intersecting discontinuities is always solvable and has been studied comprehensively in [1]. It should be noted that since initial boundary value problems are of the hyperbolic type, their known local solutions make possible continuation of the solution to be obtained into the domain of its determinacy.

In the general case of steady-state three-dimensional flows, the problem can be reduced locally to intersection of rectilinear discontinuities with a constant intensity in a certain plane. In this statement, there are no length scales inherent in the problem. Therefore, its solutions are self-similar. These solutions are also well studied throughout the entire domain of their existence [2, 3].

However, in contrast to one-dimensional unsteady flows, these locally self-similar solutions are not necessarily determined by the given local conditions. This is explained by the fact that the corresponding boundary value problems are nonhyperbolic. For example, regions of subsonic flow, which occur either as a result of interaction of discontinuities or due to the existence (even while formulating the problem) of subsonic-flow regions. By virtue of these reasons, a possibility arises for conditions given downstream to affect the flow near a point of intersecting discontinuities.

The progress in analytically investigating non-self-similar flows with intersecting discontinuities is rather insignificant, because of complexity and diversity of these flows. It is sufficient to mention the semi-centennial history of yet unfinished investigation of the problem for irregular (Mach) intersection of counter-propagating shock waves and a specific case of the problem presented by irregular reflection of a shock wave from a rigid wall.

In this paper, we consider non-self-similar interaction of a shock wave with discontinuities of both possible types with subsonic flows behind them, i.e., with a contact discontinuity surface or another shock wave.

We begin our consideration with the problem concerning reflection of a shock wave from a contact surface presented by a boundary between supersonic and subsonic flows. This problem has self-similar solutions only in two limiting cases. They are the reflection of a shock wave with a supersonic (or sonic) flow velocity behind it from a free surface (i.e., from a boundary with a quiescent gas) and the regular reflection of a shock from a rigid wall. In all other cases, the self-similar solutions do not exist.

This fact made the authors of [2] to declare that "the intersection of a shock wave with a tangential discontinuity having the nonzero (but subsonic) fluid velocity behind it is impossible at all" [2, p. 582]. The excessive rigorism of this statement will be shown below.

Seemingly, the authors of [4] were the first, who investigated the non-self-similar interaction of a local perturbance propagating from a supersonic-flow region to its boundary with a subsonic flow. They considered in the linear approximation unidirectional uniform flows (supersonic in one half-plane and subsonic in the other). These flows were perturbed by a general-form pressure wave propagating from the supersonic region to the boundary between the flows. The analytic solution obtained was applied, in particular, to calculating the flow distorted by a localized continuous triangular wave of increased pressure. Since, being unbounded in its width, the subsonic region requires pressure to be constant at infinity, the author has not considered a stepwise pressure wave, describing approximately a weak-shock wave.

In the same linear approximation, this has been done in [5]. In this paper, a supersonic flow occupies a half-plane as before. At the same time, a subsonic-flow region represents a finite-width layer separating the supersonic region and flowing along a plane-rigid wall. The authors have considered the problem of reflecting a weak shock wave from the boundary between the flows. The other problem solved by the authors con-

Institute of Mechanics, Moscow State University, Michurinskiĭ pr. 1, Moscow, 117234 Russia

1028-3358/00/4505-0229$20.00

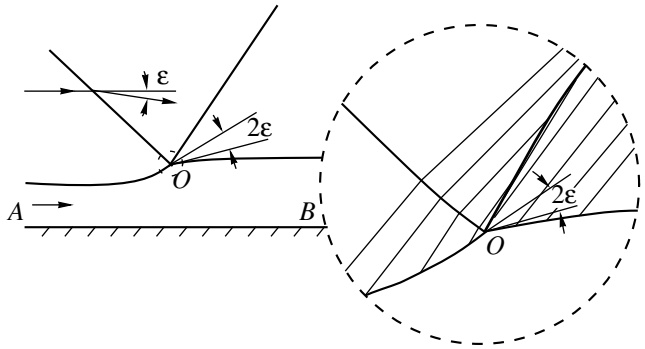


Fig. 1. Intersection of a weak pressure shock with a contact discontinuity surface.

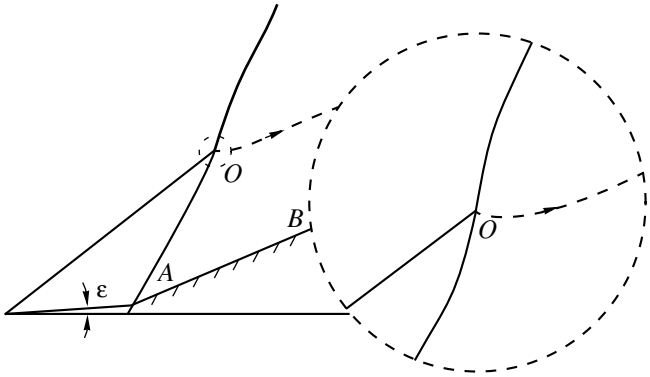


Fig. 2. Flow region in the hodograph (V, θ) plane.

cerns the flow perturbed by the small deviation of a part of the wall toward the side occupied by gas (flow around a concave angle) in the absence of perturbances coming from the supersonic region.

In accordance with the solution to the first problem, as flow approaches the point of meeting the shock, the pressure at the contact surface increases starting from its value in the incoming flow. Behind this point, the pressure drops and tends to its value behind the shock reflected from the rigid wall. At the very point of intersection, the pressure has a logarithmic singularity tending to infinity. Being concave toward the supersonic flow before the point of meeting with the shock, the boundary line is convex behind this point and has a vertical tangent (with respect to the wall confining the flow).

In a rather inaccessible paper [6] published many years ago by the author of the present study, the same problems as posed in [5] had been solved with nonlinear effects taken into account. The consideration was restricted by the assumption of low-intense shock waves, which made it possible to ignore arising vorticity of the supersonic flow. The allowance for the nonlinearity in the problems under consideration enabled us to find a structure of the non-self-similar flow, which occurred in the case when a sufficiently weak shock wave had intersected a contact surface bounding a region of subsonic flow. This flow structure is shown in Fig. 1. At the point O of the shock incidence, the separating contact line has a break forming an angle with a concavity towards the subsonic region. Thus, for the subsonic gas flow, the point O is a braking point with the zero velocity and maximum pressure at this point. A simple compression wave forms in the supersonic flow ahead of the incident shock wave due to the pressure increase transferred forward through the subsonic region. The wave is refracted while passing by the shock and gives rise to a reflected shock. Near the point O, this shock interacts with a centered rarefaction wave outgoing from the same point (at this point, the incoming shock is reflected from the boundary as from a free surface, at which the pressure is equal to the Pitot pressure in the subsonic flow). As a result of this interaction, the infinitely weak reflected shock arises already at the point O. The shock intensity grows gradually, and, at infinity, takes the value that corresponds to the reflection from the rigid wall having none of adjacent subsonic layer.

Thus, the flow asymptotic behavior is observed at a small distance from the point of shock intersection with the contact surface, as well as far from the rigid wall. This asymptotic behavior corresponds to the two above-mentioned exceptional cases of the existence of self-similar flows.

Figure 2 shows the region of the subsonic flow in the motion-hodograph plane (in the polar V, θ-coordinates, where V and θ represent the magnitude of the velocity and the angle of inclination of the velocity vector to the direction of the unperturbed flows, respectively). The lines AO and BO correspond to the contact discontinuity. To obtain them, the relation between pressure p and the angle θ, which corresponds to simple waves, both ahead and behind the incident shock and the Bernoulli integral in the subsonic flow were used. The rectilinear segment AB corresponds to the wall bounding the flow. In this segment, the stream function ψ is equal to zero, while, in the contour AOB, it is equal to a gas-flow rate in the subsonic layer ($\psi = Q$). An assigned value of Q determines the characteristic dimension of the problem, i.e., a width of the layer in the unperturbed state.

We pay attention to the facts indicated in [6] but not emphasized there. The point O shown in Fig. 2 remains fixed no matter of the intensity of the incident shock. Hence, even for an arbitrary weak shock, the flow perturbance is finite in the vicinity of this point. In particular, pressure at the point O differs significantly from that in the incoming flow: it is equal to the Pitot pressure of this flow. The angle of the separating-line approaching the incident shock does not depend as much on the shock intensity. However, the angle of the separating-line break at the intersection point decreases in proportion with a decrease in the shock intensity.

For invariable initial parameters of the problem and an unlimited increase in the width of the subsonic layer,

the flow in a fixed vicinity of the point O tends to the self-similar one. This corresponds to the reflection of the shock from a free surface, with the pressure on it equal to the Pitot pressure of the subsonic flow. Thus, as the intensity of a shock impinging onto the boundary between supersonic and subsonic flows tends to zero, the maximum perturbance remains finite and invariable. This nonlinear character of the interaction holds in progressively decreasing vicinity of the point at which the discontinuities interact with each other.

Recently, N.A. Ostapenko has informed the author on a similar property of a solution to another problem. It concerns interaction of an initial plane shock wave, behind which flow is subsonic, with a low-intensity pressure shock approaching it from the front and having the same or opposite direction. The self-similar solution to this problem exists only in particular cases when parameters of both shocks and the incident flow are linked by a certain relation. This restriction is caused by subsonic gas velocity behind the shock. Here, the self-similar solution loses additional arbitrariness occurring in the case of supersonic flow and is associated with a possibility of appearing either a centered rarefaction wave or a pressure shock, both outgoing from the interaction point. In all other cases, no self-similar solutions exist.

On the basis of these facts, Ostapenko has drawn a conclusion that, for the above-mentioned relation between the defining parameters, a low-intensity incident shock causes similar small perturbances of both the main shock and the flow behind it. In the general case, this perturbance remains finite at arbitrary low intensity of the incident shock. The behavior of the flow is quite analogous to that described above, while analyzing the first problem.

In the case of interaction of discontinuities, which was considered by Ostapenko, mathematical statement of the problem is almost completely equivalent to the first case.

We now analyze in detail both the correspondence of and difference in the two problems. In the first case, the simple-wave relation between the pressure p and the angle θ of the velocity-vector inclination is used from both sides of the interaction point at the boundary separating the subsonic-flow region. The position of this boundary in the physical plane is unknown beforehand. In the case considered by Ostapenko, a form of this relation corresponds to the conditions existing in the shock wave. In addition, a tangential discontinuity penetrates from the point of interaction of the shocks into the subsonic region. In the case of the existence of the tangential discontinuity, it is preferable to map a subsonic-flow region not onto the hodograph plane, as is done in Fig. 2, but onto the (p, θ) plane. The triangular region AOB shown in Fig. 3 represents an example of such mapping. Figure 4 presents the configuration of the discontinuities in the flow plane. In Fig. 3, identical letters are used to mark points and regions of the flow

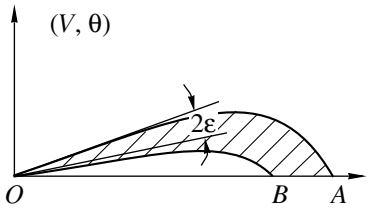


Fig. 3. Flow region in the (p, θ) plane for the second problem.

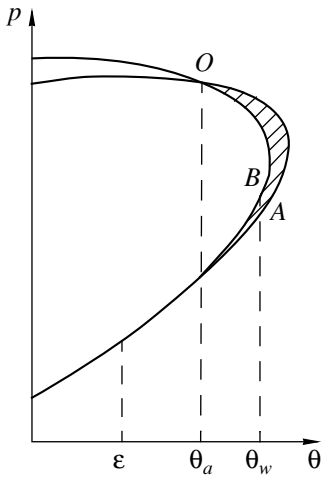


Fig. 4. Intersecting shocks in the case of the subsonic gas velocity behind them.

plane and the corresponding states. The definiteness of mapping is provided by the condition of boundedness for the subsonic-flow region by the wall AB inclined to the approaching flow at an angle θ_w, which differs from the angle θ_a corresponding to the self-similar flow. The same condition provides the existence of a linear scale that is necessary in the non-self-similar case and can be taken, for example, as the length of the segment OC. Similarly to the previous problem, the shock bounding the subsonic region has a break at the point O, which forms a concave angle with a concavity towards this region.

At the point O behind the shock, both pressure and the angle of inclination for the velocity vector are continuous, while the value of the velocity has a jump. This leads to forming the tangential discontinuity OB. If the intensity of the interacting shocks is sufficiently high, it is necessary to take into account the vorticity in a wake adjacent to the tangential discontinuity OB behind the most curved part of the shock near the point O.

By virtue of complicated boundary conditions for the stream function in the region AOB of the (p, θ) plane, it is more convenient to carry out actual solving of the problem directly in the flow plane. Certainly, this makes solving the problem more complicated compared to the first problem.

Similarly to [6], for constant parameters of the problem and unlimited distance between the points *C* and *O*, the flow in a fixed vicinity of the point *O* becomes progressively close to the self-similar and corresponds to the above-mentioned special relation between the initial parameters.

In contrast to Fig. 2, the position of the point *O* in Fig. 3 depends on the intensity of the first shock, namely, the angle ε. However, as was shown by Ostapenko, the position of the point *O* becomes stabilized as $\varepsilon \to 0$. Consequently, in this limiting case, nonlinear effects remain significant at arbitrarily small ε in a progressively decreasing vicinity of the point *O*. At the same time, the intensity of the tangential discontinuity vanishes together with ε. Obviously, it should be kept in mind that, when the region with essentially nonlinear effects becomes sufficiently small, large gradients of gas-dynamic quantities in it require taking into account the viscosity and heat conduction of gas.

It is noteworthy that, in the linear approximation and with allowance for the flow vorticity, pressure behind the shock has the logarithmic singularity at the point *O* as in the problem discussed in [6]. Generally speaking, in the linear approximation (with respect to ε), mathematical problems related to interaction of a weak pressure wave, propagating from a supersonic region, with a pressure shock or a contact discontinuity having a subsonic flow velocity behind them turn out to be almost completely identical.

ACKNOWLEDGMENTS

The author is grateful to N.A. Ostapenko for fruitful discussions that have stimulated appearance of this paper.

The work was supported by the Federal Program "Integratsiya" (project no. AO112) and the Russian Foundation for Basic Research (project no. 99–01–00002).

REFERENCES

1. B. L. Rozhdestvenskiĭ and N. N. Yanenko, *Systems of Quasi-Nonlinear Equations and Their Applications in Gas Dynamics* (Nauka, Moscow, 1968).
2. L. D. Landau and E. M. Lifshitz, *Hydrodynamics* (Nauka, Moscow, 1988).
3. G. G. Chernyĭ, *Gas Dynamics* (Nauka, Moscow, 1988).
4. L. Howarth, Proc. Cambridge Philos. Soc. **44**, 380 (1948).
5. H. S. Tsien and M. Finstone, J. Aeronaut. Sci. **16** (1949).
6. G. G. Chernyĭ, in *Theoretical Hydromechanics* (Oborongiz, Moscow, 1952), Vol. 7, No. 9.

Translated by Yu. Verevochkin

Doklady Physics, Vol. 45, No. 5, 2000, pp. 233–235. Translated from Doklady Akademii Nauk, Vol. 372, No. 3, 2000, pp. 330–332.

MECHANICS

Model of Unsteady Flow Distribution in Multiloop Hydraulic Circuits

O. A. Balyshev

Presented by Academician O.N. Favorskiĭ July 2, 1999

Received July 8, 1999.

1. The problem of calculating unsteady flow distributions is very urgent in studying multiloop hydraulic systems (urban heat supply and water supply, water heating and building ventilating, cooling technological facilities, adverse-waste removing, etc.). The necessity for solving this problem arises, for example, in evaluating consequences of hydraulic shock and developing measures for its protection, in analyzing controlling processes in the case of normal and emergency conditions, in designing regimes of putting into operation and shutting-down long pipelines, and in many other cases.

2. Unsteady motion of compressible viscous continuous fluid in an isolated pipe is described by the classical system of hydrodynamic equations [1, 2]. Based on them, N.E. Joukowski has derived formulas for calculating propagation of shock waves [3].

Models for fluid motion in a pipe were extended to unsteady flow distributions in multiloop systems. For this purpose, in certain cases, the original system of hydrodynamic equations was used directly (see, for example [4, 5]), while in other cases, the solution was based on the Joukowski equation [6, 7]. Significant computational difficulties were revealed in both the cases. These difficulties are related to a necessity of specifying two separate boundary conditions (for pressures and rates) for each section of the calculation scheme. Such a necessity arises in the course of solving the system of partial differential equations (being reduced under certain conditions to equations of the hyperbolic type). The boundary conditions involve time-dependent variables being unknown functions of the flow distribution. The simplification of computational difficulties by transforming computational schemes, for example through passing from close graphs to trees, restricts possibilities in the formulation of the flow-distribution problems. Provision for requirements of the approximation, stability, and convergence of computational procedures, as well as finding the numerical solution with acceptable accuracy turn out to be rather complicated.

In order to overcome difficulties arising in simulation of unsteady processes in complex hydraulic systems, we may employ a technique developed in the theory of hydraulic circuits, which allows us to represent laws of one-dimensional flow in simpler form compared to the traditional approach, i.e., the "network" form.

3. Based on analogy between electrical and hydraulic circuits, V. Ya. Khasilev and A. P. Merenkov laid, in the 1960s, foundations for theory of hydraulic circuits [8–10]. Its subject is direct and inverse problems of flow distributions and problems of optimum synthesis (i.e., choice of schemes and parameters) for hydraulic circuits.

As applied to problems of flow distributions, models in the theory of hydraulic circuits are based on the representation of an actual circuit network by an oriented marked graph (circuit) specified by the number of nodes m, branches (arcs, sections) n, linearly independent loops c, as well as by matrix $\mathbf{A}$ with the dimension $(m-1)\times n$ for joints of linearly independent nodes and branches and by the matrix $\mathbf{B}$ with the dimension $c \times n$ for coincidence of linearly independent directed loops and branches, etc.

Each branch is described by integral parameters: the desired flow rate x_i, loss in head y_i, constant resistance s_i, and the total head H_i.

In the framework of the theory of hydraulic circuits, the following methods were developed for solving direct problems of the steady flow distribution in circuits with lumped parameters: the loop flow-rate method (the analogue of the loop-current method in electric circuits), node pressure method (the analogue of the node-potential one), and the method of searching for the extremum of a function, whose form is determined on the basis of principles of mechanics or the second law of thermodynamics [11].

In the loop flow-rate method, the initial system of equations has the form

$$\mathbf{Ax} = \mathbf{Q}, \tag{1}$$

Melent'ev Institute of Power Systems,
Siberian Division, Russian Academy of Sciences,
ul. Lermontova 130, Irkutsk, 664033 Russia

1028-3358/00/4505-0233$20.00

$$\mathbf{By} = 0, \quad (2)$$

$$y_i + H_i = f_i(x_i), \quad i = 1, \ldots, n, \quad (3)$$

where $\mathbf{Q}$ is the vector of inflows and discharges in nodes of a hydraulic circuit.

4. In order to use the network approach in the theory of hydraulic circuits, while solving problems of unsteady flow distribution, we can, in particular, pass from a system of algebraic equations (1)–(3) to the system of algebraic-differential equations

$$\mathbf{A}(t)\mathbf{x}(t) = \mathbf{Q}(t), \quad (4)$$

$$\mathbf{B}(t)\mathbf{y}(t) = 0, \quad (5)$$

$$y_i(t) + H_i(t) = f_i\left(x_i(t), \frac{dx_i}{dt}, t\right), \quad i = 1, \ldots, n, \quad (6)$$

where $\mathbf{x}(t)$, $\mathbf{y}(t)$, $\mathbf{H}(t)$, $\mathbf{Q}(t)$ are, respectively, vectorial functions of flow rates, pressure loss, total head in branches and inflows (discharges) at nodes of a hydraulic system.

5. To determine the form of the function $f_i\left(x_i(t), \frac{dx_i}{dt}, t\right)$, we take into account the fact that circuits with lumped parameters are considered (the density ρ and temperature T being constant). We assume, in addition, that the dependence of the pressure loss on the flow rate in a branch is quadratic. Then, the system of one-dimensional hydrodynamic equations takes the form

$$\rho\frac{\partial w}{\partial \zeta} = 0, \quad (7)$$

$$\rho\frac{\partial w}{\partial t} = \rho Z - \frac{\partial P}{\partial \zeta} + \frac{\lambda\rho}{2D}w^2 + c_0 w, \quad (8)$$

where w and P are the rate and pressure of a continuous medium; ζ and t are geometric and time coordinates of the pipeline section; Z, λ, and c_0 characterize gravity, viscosity, and roughness, respectively.

By changing the variables w and P, that are the functions of geometric and time coordinates, and of the integral parameters (the flow rate x and the loss in the head y within a branch of the circuit), we obtain a closing equation in the differential form

$$\begin{aligned} y_i(t) + H_i(t) &= R_i(t)\frac{dx_i(t)}{dt} + S_i(t)x_i^2(t) \\ &+ C_i(t)x_i(t) \pm H_{gi}(t) = f_i\left(x_i(t), \frac{dx_i}{dt}, t\right). \end{aligned} \quad (9)$$

Here, $y_i(t)$ and $x_i(t)$ are the desired functions of time in problems of flow distributions, which determine the state of the ith section; $R_i(t)$, $S_i(t)$, and $C_i(t)$ are known parameters of the ith section; $H_{gi}(t)$ and $H_i(t)$ are, respectively, the gravity head and total head in the ith section of the circuit.

Based on closing relationship (9) and exploiting foundations and the matrix technique of the theory of hydraulic circuits, we can formulate the mathematical model of an unsteady flow-rate distribution for a hydraulic circuit with lumped parameters. This model naturally fits the network model of the first and second Kirchhoff laws written out in form (4), (5) and closing relation (9) for all branches of the circuit ($i = 1, \ldots, n$).

6. Furthermore, we followed the method of simplifying the realization of computational algorithms, which was used by V.Ya. Khasilev and A.P. Merenkov in the case of a steady flow distribution [8–10]. To do this, we employ transformations of the original multicircuit digraph into two subgraphs: i.e., a tree (graph without directed circuits) and chords (branches complementing the tree to a multicircuit graph). This made it possible to pass from a space with the dimension $2n$ to those of dimensions c (in the method of circuit flow-rates) or $m - 1$ (in the method of node pressures).

As a result of the manipulations performed with respect to the desired vectorial function for the flow rates on chords [$\mathbf{x}_c(t)$], we obtain the canonical system of nonlinear differential equations

$$\begin{aligned} \frac{d\mathbf{x}_c(t)}{dt} &= \mathbf{H}_Q(t) - \sigma(t)\mathbf{x}_c(t) - \mathbf{S}_c(t)\mathbf{Z}_c\mathbf{x}_c(t) \\ &- \mathbf{B}_d(t)\mathbf{S}_d(t)\mathbf{Z}_d(\mathbf{B}_d^T(t)\mathbf{x}_c(t)), \end{aligned} \quad (10)$$

where $\mathbf{Z}_c$ and $\mathbf{Z}_d$ are square-diagonal matrices with components equal to the absolute value of flow rates in sections:

$$\Omega(t) = \mathbf{A}_d^{-1}(t)\mathbf{Q}(t); \quad \mathbf{B}_d(t) = -[\mathbf{A}_d^{-1}(t)\mathbf{A}_c(t)]^T;$$

$$\frac{d\Omega(t)}{dt} = \mathbf{Q}(t)\frac{d}{dt}[\mathbf{A}_d^{-1}(t)] + \mathbf{A}_d^{-1}(t)\frac{d\mathbf{Q}(t)}{dt};$$

$$\frac{d}{dt}[\mathbf{B}_d^T(t)] = -\mathbf{A}_c(t)\frac{d}{dt}[\mathbf{A}_d^{-1}(t)] - \mathbf{A}_d^{-1}(t)\frac{d}{dt}[\mathbf{A}_c(t)];$$

$$\mathbf{H}_Q(t) = [\mathbf{R}_c(t) + \mathbf{B}_d(t)\mathbf{R}_d(t)\mathbf{B}_d^T(t)]^{-1}$$

$$\times\left\{\mathbf{B}(t)[\mathbf{H}(t) \pm \mathbf{H}_g(t)]\right.$$

$$\left. - \mathbf{B}_d(t)\left[\mathbf{R}_d(t)\frac{d\Omega(t)}{dt} + \mathbf{C}_d(t)\Omega(t) + \mathbf{S}_d(t)\Omega^2(t)\right]\right\};$$

$$\sigma(t) = [\mathbf{R}_c(t) + \mathbf{B}_d(t)\mathbf{R}_d(t)\mathbf{B}_d^T(t)]^{-1}$$

$$\left\{\mathbf{C}_c(t) + \mathbf{B}_d(t)\left[\mathbf{R}_d(t)\frac{d}{dt}[\mathbf{B}_d^T(t)] + \mathbf{C}_d(t)\mathbf{B}_d^T(t)\right.\right.$$

$$+ 2\mathbf{S}_d(t)\Omega(t)\mathbf{B}_d^T(t)\Big]\Big\}.$$

System of equations (10) complemented by initial conditions has a unique solution (the Cauchy problem).

7. On the basis of the extremum approach [10, 11], we can formulate, in the variational form, the problem of analysis of unsteady regimes in hydraulic circuits with lumped parameters.

We write out the equation of virtual displacements for a branch of a hydraulic circuit in the form

$$h_i(t)dx_i = \Big[R_i(t)\frac{dx_i(t)}{dt} + s_i(t)x_i^2(t) + c_i(t)x_i(t) \pm H_{gi} - H_i(t)\Big]dx_i.$$

Then the Lagrange function for the hydraulic circuit as a whole can be written as

$$W = \sum_{i=1}^{n}\Big[R_i(t)x_i(t)\frac{dx_i}{dt} + s_i(t)\frac{x_i^3(t)}{3} + c_i(t)\frac{x_i^2(t)}{2} \pm H_{gi}x_i(t) - H_i(t)x_i(t)\Big]. \quad (11)$$

This implies the following formulation of the extremum problem: It is necessary to determine those components of the vectorial function $\mathbf{x}(t)$ in a given hydraulic circuit with n branches, which satisfy the conditions of the material balance at circuit nodes of (4) and minimize the Lagrange function (11).

8. Steady states of a hydraulic circuit, which corresponded to limiting points of models (10) and (4), (11) (as $d/dt[\mathbf{x}(t)] \longrightarrow 0$), were determined on the basis of solutions to the systems of nonlinear algebraic equations [8–10] related to systems appearing in the loop flow-rate method. These states are, in particular, useful for determination of the initial flow distributions, while solving unsteady problems.

9. In the framework of the theory for hydraulic circuits, preliminary investigations of inverse problems for the unsteady flow distribution and stability of hydraulic systems were also performed. Studies associated with the experimental verification of the models proposed now are being conducted, whose first results have shown a qualitative agreement between the model solutions and actual unsteady processes. As a whole, the approach proposed to the description of unsteady flows in multiloop circuits must, seemingly, be useful in both developing theory of hydraulic circuits and improving the efficiency of applications of this theory in practice.

ACKNOWLEDGMENTS

The author must give credits for memories of professor V.Ya. Khasilev and Corresponding Member of the RAS A.P. Merenkov, who supported the development of this scientific approach in the framework of the theory of hydraulic circuits.

The author is also grateful to B.M. Kaganovich for permanent attention to the present work.

This study was supported by the Russian Foundation for Basic Research, project no. 96-15-98185.

REFERENCES

1. L. I. Sedov, *Continuum Mechanics* (Fizmatgiz, Moscow, 1973), Vol. 1.
2. L. D. Landau and E. M. Lifshitz, *Hydrodynamics* (Nauka, Moscow, 1988).
3. N. E. Joukowski, *On a Hydraulic Shock in Water Pipelines* (Gostekhizdat, Moscow, 1949).
4. O. F. Vasil'ev, É. A. Bondarev, A. F. Voevodin, and M. A. Kanibolotskiĭ, *Nonisothermic Gas Flow in Pipelines* (Nauka, Novosibirsk, 1978).
5. A. F. Voevodin and S. M. Shugrin, *Methods of Solving Problems for One-Dimensional Evolution Systems* (Nauka, Novosibirsk, 1993).
6. B. N. Gromov, L. P. Kanina, V. G. Sidler, and L. E Sidler, in *New Information Technologies for Controlling Development and Functioning Pipeline Systems* (Irkutsk, 1993), pp. 69–78.
7. K. P. Vishnevskiĭ, *Transitional Processes in Head Water Supply Systems* (Agpromizdat, Moscow, 1986).
8. V. Ya. Khasilev, Izv. Akad. Nauk SSSR, Energ. Transp., No. 1, 69 (1964).
9. V. Ya. Khasilev, Izv. Akad. Nauk SSSR, Energ. Transp., No. 2, 231 (1964).
10. A. P. Merenkov and V. Ya. Khasilev, *Theory of Hydraulic Circuits* (Nauka, Moscow, 1985).
11. B. M. Kaganovich, A. P. Merenkov, and O. A. Balyshev, *Principles in the Theory of Heterogeneous Hydraulic Circuits* (Nauka, Novosibirsk, 1997).
12. O. A. Balyshev, B. M. Kaganovich, and A. P. Merenkov, Izv. Ross. Akad. Nauk, Energetics, No. 2, 96 (1996).

Translated by V. Devitsyn

Doklady Physics, Vol. 45, No. 5, 2000, pp. 236–240. Translated from Doklady Akademii Nauk, Vol. 372, No. 3, 2000, pp. 333–336.

MECHANICS

Three-Dimensional Contact Problem for an Elastic Wedge with Allowance for Friction in an Unknown Contact Region

D. A. Pozharskiĭ

Presented by Academician I.I. Vorovich June 25, 1999

Received July 2, 1999

In [1], the solution to the problem concerning action on a face of a three-dimensional elastic wedge of normal and tangential (but perpendicular to the wedge edge) loads was obtained. In solving the problem, the complex-valued Fourier–Kontorovich–Lebedev integral was used, and the problem of the elasticity theory was reduced to the Hilbert boundary value problem generalized by I.N. Vekua (see [2]). In the special case when a wedge angle corresponds to a half-space, the formulas obtained for displacements and stresses in the wedge, whose one face is stress-free, coincide with the known solutions to the Boussinesq and Cerruti problems [3].

In this paper, a contact quasi-static problem based on the solution obtained in [1] is considered. We analyze a die moving along a face of an elastic wedge in the direction orthogonal to the wedge edge. Since the die represents an elliptic paraboloid strongly elongated along the edge, friction forces can be approximately considered as collinear with and directed oppositely to the die motion. The statement of the problem generalizes the well-known case of die motion in a half-space [4]. The integral equation of the contact problem with an unknown contact region is solved by the method of nonlinear boundary integral equations [5, 6]. The effect of the Coulomb friction factor on the dependence between an impressing force and die upsetting is investigated for different wedge angles. After solving the contact problem, the effective stress [7] is calculated on the symmetry axis of the contact region. Being important in applications to the Novikov gear transmissions [8, 9], this effective stress is calculated as a function of the wedge angle, distance between the die and the wedge edge, and direction and magnitude of the friction forces.

Other methods for solving three-dimensional contact problems in the presence of friction were used in [10, 11].

Institute of Mechanics and Applied Mathematics, Rostov State University, pr. Stachki 200/1, Rostov-on-Don, 344090 Russia

In the quasi-static contact problem under consideration, a rigid die, initially imbedded into a face of an elastic wedge with the opening angle α, begins its sufficiently slow skewness-free motion along this face in the direction perpendicular to the wedge edge. Below, we use the cylindrical coordinates r, φ, z with the z-axis directed along the wedge edge. The problem is assumed to be symmetric with respect to the z-coordinate. The other wedge face ($\varphi = 0$) is stress-free. The boundary condition describing contact of the bodies $u_\varphi(r, \alpha, z) = -[\delta - f(r, z)]$, $(r, z) \in \Omega$, where δ corresponds to the die upsetting and the function $f(r, z) = (r - a)^2/(2R_1) + z^2/(2R_2)$ presents the shape of the wedge base ($R_1 \ll R_2$), is satisfied due to application of formula (4) from [1]. As a result, we obtain an integral equation with respect to the unknown normal contact pressure $\sigma_\varphi(r, \alpha, z) = -q(r, z)$ in the unknown contact region $(r, z) \in \Omega$:

$$\iint_\Omega q(x, y)\left[\frac{1}{R} - \mu\frac{1-2\nu}{2(1-\nu)}\frac{r-x}{R^2} + K(x, y, r, z)\right]dxdy = 2\pi\theta[\delta - f(r, z)], \quad (1)$$

$$K(x, y, r, z) = \frac{4}{\pi^2}\int_0^\infty\int_0^\infty \sinh\frac{\pi u}{2} W(u, \beta x)K_{iu}(\beta r)\cos\beta(y - z)d\beta du,$$

$$W(u, \beta x) = W_1(u, \alpha)\Phi_1(u, \beta x) - W_2(u, \alpha)\Phi_2(u, \beta x)$$

$$+ 2\cosh\frac{\pi u}{2}K_{iu}(\beta x)[W_0(u, \alpha) - \coth\pi u + \mu f_0(u, \alpha)]$$

$$+ \frac{\mu}{2(1-\nu)}\cosh\frac{\pi u}{2}\int_0^\infty\left\{W_1(u, \alpha)h_1(t, \alpha)\right. \quad (2)$$

$$\left. - W_2(u, \alpha)h_2(t, \alpha) - (1 - 2\nu)\left[\coth\frac{\pi u}{2}\tanh\frac{\pi t}{2}\right.\right.$$

$$\left.\left. + \tanh\frac{\pi u}{2}\coth\frac{\pi t}{2}\right]\right\}\frac{K_{it}(\beta x)\sinh\pi t dt}{\cosh\pi t - \cosh\pi u},$$

1028-3358/00/4505-0236$20.00

$$W_0(u, \alpha) = \frac{\sinh 2\alpha u + u \sin 2\alpha}{\cosh 2\alpha u - 2u^2 \sin^2 \alpha - 1},$$

$$f_0(u, \alpha) = \frac{2u \sin^2 \alpha}{\cosh 2\alpha u - 2u^2 \sin^2 \alpha - 1}.$$

Here, $\theta = G/(1 - \nu)$, ν is the Poisson's ratio, G is the shear modulus, μ is the Coulomb friction factor, $R = \sqrt{(r-x)^2 + (z-y)^2}$, $K_{i\tau}(x)$ is the MacDonald function, and the functions $\Phi_n(u, \beta)$ ($n = 1, 2$) are calculated by using the integral second-kind Fredholm equations ($0 \le u < \infty$)

$$\Phi_n(u, \beta x) = (1 - 2\nu) \int_0^\infty L_n(u, t)[\Phi_n(t, \beta x) + F_n(t, \beta x)]\, dt, \quad (3)$$

$$F_n(u, \beta x) = \left[1 - \frac{\mu f_n(u, \alpha)}{2(1-\nu)(1-2\nu)}\right] \cosh \frac{\pi u}{2} K_{iu}(\beta x) + \frac{\mu}{2(1-\nu)} \cosh \frac{\pi u}{2} \int_0^\infty h_n(t, \alpha) K_{it}(\beta x) \frac{\sinh \pi t\, dt}{\cosh \pi t - \cosh \pi u}. \quad (4)$$

The functions $W_n(u, \alpha)$, $f_n(u, \alpha)$, $h_n(u, \alpha)$, and $L_n(u, t)$ ($n = 1, 2$) are determined according to formulas (2) from [1]. For $\mu > 0$ and $\mu < 0$, the die begins its motion toward the wedge edge and in the opposite direction, respectively. To improve convergence of the integrals, the singular part is explicitly isolated in the kernel of the integral equation (1). At $\alpha = \pi$, equation (1) coincides with equation (3) from [4] (with allowance for the fact that, at $\mu > 0$, the motion occurs oppositely to the positive direction of the r-axis).

The formal solution to equations (3) can be written as functional series in powers of $(1 - 2\nu)$, which are uniformly convergent in the space of the functions that are continuous and bounded at the semiaxis [1]. The method of mechanical quadratures and the Gaussian quadrature could be used for practical solving equations (3). Singular integrals appearing in expressions (2) and (4) are calculated by the standard-regularization method. It is applied after the infinite range of integration has been partitioned to localize the singularities in a finite interval.

To solve the integral equation (1) under the condition $q(r, z) = 0$, $(r, z) \in \partial\Omega$, the method of nonlinear boundary integral equations [5, 6] is applied. It allows simultaneous determination of normal contact pressure and the contact region. The main features characterizing the integral operator generated by the kernel of equation (1) at $\mu = 0$ (i.e., its strict positiveness and complete continuity [6]) remain valid at $\mu \neq 0$ as well (for $\alpha = \pi$, this was noted in [5]). Therefore, data obtained in [6], at $\mu = 0$, concerning the existence and uniqueness of the solution to equation (1), as well as the method of constructing the solution, completely correspond to the case under consideration when friction forces are taken into account. Furthermore, omitting primes, we use the dimensionless notation (2.1) from [6] (see also the paragraph following formulas (2.1) in [6]).

Analysis of the results obtained shows that, at $\alpha = 180°$, the quantity $P(\delta)$ (see [6]) is practically independent of μ. An explanation of this fact is presented below. At $\alpha = 180°$, the solution to equation (1) can be sought as a series in powers of the small parameter $\varepsilon_* = \mu(1 - 2\nu)/(2 - 2\nu)$ [4]. Rejecting the terms on the order of ε_*^2 (at $\mu = 0.2$ and $\nu = 0.3$, $\varepsilon_*^2 \approx 0.003$) yields that in the case of die motions in positive and negative directions of the r-axis

$$q(r, z) = q_0(r, z) - \varepsilon_* q_1(r, z) + O(\varepsilon_*^2),$$

and

$$q(r, z) = q_0(r, z) + \varepsilon_* q_1(r, z) + O(\varepsilon_*^2),$$

respectively.

Here, the function $q(r, z)$ satisfies the integral equation for the contact problem in the case of the friction-free half-space. The function $q_1(r, z)$ is expressed in terms of $q_0(r, z)$ regardless of ε_*. The normal force P [integral of the function $q(r, z)$ taken over the region Ω] is, obviously, independent of the direction of the die motion in the half-space. Therefore, the integral of the function $q_1(r, z)$ taken over the contact region must be equal to zero. Consequently, the terms describing the effect of friction forces on the function $P(\delta)$ are on the order of $o(\varepsilon_*)$. The friction essentially affects the eccentricity (and the momentum) of the normal force, which provides the skewness-free die motion. As is well known, in the axisymmetric contact problem with allowance for friction [12], friction forces do not affect the dependence $P(\delta)$ at all. (Two independent systems of stresses and strains arise there: one of them determines the normal force regardless of μ, while the other, the force momentum dependent on friction.) The smaller the wedge angle, the stronger the friction and the effect of the motion direction on the function $P(\delta)$. Table 1 (for a quarter of space, $\alpha = 90°$), presents the normal force $P \times 10^3$ as a function of the upsetting $\delta \times 10^3$ at different values of the friction factor μ. Here, the values $A = 0.1$, $B = 0.005$, $\varepsilon = 0.15$, $\lambda = 1$, and $\nu = 0.3$ ($\gamma = 0$) are taken in the notation of [6].

In the case of the angle $\alpha < \pi$, the closer the die to the wedge edge, i.e., the smaller λ [6], the smaller the value of $P(\delta)$ due to the increase in the flexibility of the elastic material. Being known at zero friction [6], this conclusion is valid at fixed $\mu \neq 0$ as well. If the die moves to the wedge edge ($\mu > 0$ and λ is fixed), then $P(\delta)$ is smaller (it is easier to impress the die) compared to its value at $\mu = 0$ and the same value of λ. If the die

Table 1. Force $P \times 10^3$ at $\alpha = 90°$

μ	$\delta \times 10^3 = 4$	4.5	5	5.5	6	6.5
–0.2	0.531	0.632	0.740	0.849	0.965	1.09
–0.1	0.510	0.605	0.708	0.812	0.920	1.03
0	0.491	0.581	0.678	0.777	0.879	0.986
0.1	0.474	0.559	0.650	0.746	0.841	0.942
0.2	0.457	0.539	0.625	0.716	0.808	0.903

moves off the edge ($\mu < 0$ and λ is fixed), $P(\delta)$ exceeds its value corresponding to $\mu = 0$ and to the same value of λ. This increase in the force (at $\mu < 0$) occurs because of higher maximum normal contact pressure. At the same time, the area of the contact region occurring at $\mu < 0$ can be smaller than that at $\mu = 0$, while the latter, in turn, is smaller compared to area of the region Ω at $\mu > 0$. (For example, at $\alpha = 70°$, $\lambda \approx 0$, and $\delta \times 10^3 = 6.5$, values of the rest parameters are the same as in Table 1). At $\delta = \text{const}$, the motion of the die toward the wedge edge ($\mu > 0$) counteracts the loss of the contact in the vicinity of the edge (i.e., to the departure of the edge from the die), which is observed at $\mu = 0$, $\lambda \approx 0$, and sufficiently acute angles α [6].

After the contact problem has been solved, in other words, when the function $q(r, z)$ and the contact area Ω are already known, it is possible to determine the dimensionless effective stress $\sigma_e' = \sigma_e/(2\pi\theta)$, which is important in applications [8, 9]. The method of optical analogy made it possible to determine (see [12, p. 67]) that, for nonzero friction under the smooth die, the zone of the maximum tangential stresses is shifted close to the boundary of the elastic half-plane. Furthermore, based on the concept of the surface strength, we calculate σ_e' at the axis of symmetry of the contact region ($z = 0$). To do this, we employ the first formula (3.1) ([6]) and the following formulas (primes are omitted):

$$\sigma_e = 2^{-1/2}\sqrt{(\sigma_r - \sigma_\varphi)^2 + (\sigma_\varphi - \sigma_z)^2 + (\sigma_z - \sigma_r)^2 + 6\tau_{r\varphi}^2},$$

$$\sigma_r = \frac{\nu}{1-\nu}\left(\frac{\partial u_r}{\partial r} + \frac{\partial u_z}{\partial z}\right) + \frac{\partial u_r}{\partial r} - \frac{\nu}{1-\nu}q(r, 0),$$

$$\sigma_\varphi = -q(r, 0),$$

$$\sigma_z = \frac{\nu}{1-\nu}\left(\frac{\partial u_r}{\partial r} + \frac{\partial u_z}{\partial z}\right) + \frac{\partial u_z}{\partial z} - \frac{\nu}{1-\nu}q(r, 0),$$

$$\tau_{r\varphi} = -\mu q(r, 0),$$

$$\frac{\partial u_r}{\partial r} = -\frac{1-2\nu}{\pi^3}\int_0^\infty\int_0^\infty\int_0^\infty E_1(\beta, t, u)K_*(\beta, t, r_0)\,d\beta\,dt\,du$$

$$-\frac{1}{\pi^3}\int_0^\infty\int_0^\infty E_2(\beta, t)\left[\frac{r_0}{2}K_*(\beta, t, r_0)\right.$$

$$\left. + (1-2\nu)\mathrm{Re}K_{1+it}(\beta r_0)\right]d\beta\,dt - Q(r_0)$$

$$-(1-2\nu)q(r, 0)$$

$$-\frac{\mu(1-\nu)}{\pi}\iint_\Omega \frac{1}{\sqrt{(r_0-x)^2 + y^2}}\frac{\partial}{\partial x}q(x, y)\,dx\,dy,$$

$$\frac{\partial u_z}{\partial z} = \frac{1-2\nu}{\pi^3}\int_0^\infty\int_0^\infty\int_0^\infty E_1(\beta, t, u)\beta K_{it}(\beta r_0)\,d\beta\,dt\,du \qquad (5)$$

$$+\frac{r_0}{2\pi^3}\int_0^\infty\int_0^\infty E_2(\beta, t)\beta K_{it}(\beta r_0)\,d\beta\,dt + Q(r_0),$$

$$Q(r_0) = \frac{1-2\nu}{2\pi}\iint_\Omega \frac{y}{(r_0-x)^2 + y^2}\frac{\partial}{\partial y}q(x, y)\,dx\,dy$$

$$+\frac{\mu\nu}{\pi}\iint_\Omega \frac{(r_0-x)y}{[(r_0-x)^2 + y^2]^{3/2}}\frac{\partial}{\partial y}q(x, y)\,dx\,dy,$$

$$r_0 = \begin{cases} r + \lambda & (\lambda > \varepsilon) \\ r + \varepsilon & (\lambda \le \varepsilon), \end{cases}$$

$$K_*(\beta, t, r_0) = \beta K_{it}(\beta r_0)$$

$$-[t\,\mathrm{Im}K_{1+it}(\beta r_0) - \mathrm{Re}K_{1+it}(\beta r_0)]/r_0,$$

$$E_1(\beta, t, u) = \frac{\sinh\pi t \sinh\pi u}{\cosh\pi t + \cosh\pi u}\left\{W_1(u, \alpha)\coth\frac{\alpha t}{2}\frac{\Psi_1(u, \beta)}{\cosh(\pi u/2)}\right.$$

$$\left. - W_2(u, \alpha)\tanh\frac{\alpha t}{2}\frac{\Psi_2(u, \beta)}{\cosh(\pi u/2)}\right.$$

$$- \coth\frac{\pi u}{2}\coth\frac{\pi t}{2}\left[\left(1 - \frac{\mu u\tanh(\pi u/2)}{2(1-\nu)(1-2\nu)}\right)S(u,\beta)\right.$$

$$\left. + \frac{\mu(1-2\nu)}{2(1-\nu)}\int_0^\infty \tanh\frac{\pi\tau}{2}\frac{S(\tau,\beta)\sinh\pi\tau d\tau}{\cosh\pi\tau - \cosh\pi u}\right]$$

$$- \tanh\frac{\pi u}{2}\tanh\frac{\pi t}{2}\left[\left(1 - \frac{\mu u\coth(\pi u/2)}{2(1-\nu)(1-2\nu)}\right)S(u,\beta)\right.$$

$$\left.\left. + \frac{\mu(1-2\nu)}{2(1-\nu)}\int_0^\infty \coth\frac{\pi\tau}{2}\frac{S(\tau,\beta)\sinh\pi\tau d\tau}{\cosh\pi\tau - \cosh\pi u}\right]\right\},$$

$$E_2(\beta, t) = -2\beta\sinh\pi t\left\{\frac{\sin\alpha}{\cosh(\pi t/2)}\left[\frac{\Psi_1(t,\beta)}{\sinh\alpha t + t\sin\alpha}\right.\right.$$

$$\left.\left. - \frac{\Psi_2(t,\beta)}{\sinh\alpha t - t\sin\alpha}\right]\right.$$

$$\left. - \frac{\mu}{1-\nu}S(t,\beta)\left[\frac{\sinh 2\alpha t - t(1-2\nu)^{-1}\sin 2\alpha}{\cosh 2\alpha t - \cos 2\alpha} - \coth\pi t\right]\right\},$$

$$\Psi_n(u,\beta)$$

$$= -\int_\Omega\int q(r,z)[F_n(u,\beta r_0) + \Phi_n(u,\beta r_0)]\cos\beta z\, dr\, dz,$$

$$n = 1, 2,$$

$$S(t,\beta) = -\int_\Omega\int q(r,z)K_{it}(\beta r_0)\cos\beta z dr dz.$$

In (5), all components of the stress tensor are normalized to $2\pi\theta$. To improve convergence of the integrals occurring in expressions for $\frac{\partial u_r}{\partial r}$ and $\frac{\partial u_z}{\partial z}$ in (5), the terms corresponding to the case $\alpha = \pi$ are explicitly isolated there. For $\alpha = \pi$, we have $E_1(\beta, t, u) \equiv E_2(\beta, t) \equiv 0$. The results of integrating by parts testify to the coincidence of the stresses σ_r and σ_z calculated at the boundary of the half-space by both (5) and the formulas from [3].

If to set in (5) $\mu = 0$ and $\alpha = \pi$ and consider the function $q(r, z)$ as defined by law (3.4) of [6] in the elliptic region Ω, then, at the point of the initial contact (in the center of the ellipse), the expression for σ_e turns into the known formula (3.5) of [6]. For the sufficiently elongated ellipse under consideration, σ_e corresponding to the center of the ellipse, has the maximum value for the contact surface and considerably exceeds the values of σ_e at the ellipse edge [7]. If the shape of the ellipse Ω is close to a circle, the maximum value of σ_e on the contact surface is attained at the edge of the major semiaxis. However, this value only slightly exceeds that at the point of the initial contact [7].

Table 2 presents the values for upsetting $\delta \times 10^3$ and the maximum effective stress $\sigma_e \times 10^3$ occurring at the symmetry axis of the contact region. The data correspond to the constant impressing force $P \times 10^3 = 0.583$ and different friction factors μ (and directions of the die motion), for the half-space, the wedge with the opening angle $\alpha = 110°$, and values of λ, which characterize a degree of the die closeness to the wedge edge [6]. Values of other parameters are the same as in Table 1. If friction forces are taken into account for the half-space (and for the wedge at values of λ and α being not too small), then the maximum of σ_e is attained, as before, at the point of the initial contact, and the value of σ_e grows with increasing $|\mu|$. For the constant impressing force in the vicinity of the wedge edge (and sufficiently small values of λ), a point corresponding to the maximum of σ_e begins to shift from the initial contact point. At $\mu < 0$ and $\mu > 0$, this shift occurs, as a rule, toward the wedge edge (friction forces are directed toward the edge) and in the opposite direction (friction forces are directed from the edge), respectively. For $\alpha = 70°$, $\lambda = \varepsilon$, and values of other parameters corresponding to Table 2, the contact is violated in the vicinity of the edge (due to a significant increase in the upsetting at a constant force).

Table 2. Upsetting and effective stress at a constant force P

μ	$\alpha = 180°$		$\alpha = 110°$, $\lambda = \varepsilon$		$\alpha = 110°$, $\lambda = \varepsilon/2$		$\alpha = 110°$, $\lambda = \varepsilon/4$	
	$\delta \times 10^3$	$\max \sigma_e \times 10^3$	$\delta \times 10^3$	$\max \sigma_e \times 10^3$	$\delta \times 10^3$	$\max \sigma_e \times 10^3$	$\delta \times 10^3$	$\max \sigma_e \times 10^3$
−0.2	4.00	1.61	4.52	1.56	4.74	2.10	5.02	2.40
−0.1	4.00	1.26	4.76	1.17	5.04	1.53	5.32	1.82
0	4.00	1.12	5.00	1.08	5.34	1.49	5.62	1.80
0.1	4.00	1.26	5.23	1.29	5.42	1.56	5.90	2.04
0.2	4.00	1.61	5.46	1.73	5.89	1.99	6.18	2.34

ACKNOWLEDGMENTS

The work was supported by the Russian Foundation for Basic Research, project no. 99-15-96012.

REFERENCES

1. I. A. Lubyagin, D. A. Pozharskiĭ, and M. I. Chebakov, Dokl. Akad. Nauk SSSR **321**, 58 (1991) [Sov. Phys.–Dokl. **36**, 797 (1991)].
2. V. M. Aleksandrov and D. A. Pozharskiĭ, *Nonclassic Spatial Problems in Mechanics of Contact Interactions for Elastic Bodies* (Faktorial, Moscow, 1998).
3. H. Hahn, *Elastizitatstheorie. Grundlagen der Linearen Theorie und Anwendungen* (Teubner, Stuttgart, 1986; Mir, Moscow, 1988).
4. L. A. Galin and I. G. Goryacheva, Prikl. Mat. Mekh. **46**, 1016 (1982).
5. B. A. Galanov, Author's Abstract of Doctoral Dissertation in Mathematical Physics (Moscow, 1989).
6. D. A. Pozharskiĭ, Prikl. Mat. Mekh. **59**, 812 (1995).
7. B. S. Koval'skiĭ, *Calculation of Machinery Components for the Local Compression* (Kharkov, 1967).
8. V. I. Korotkin and D. A. Pozharskiĭ, Probl. Mashinostr. Nadezhnosti Mash., No. 3, 107 (1996).
9. V. I. Korotkin, Vestn. Mashinostr., No. 6, 8 (1997).
10. A. S. Kravchuk, Trenie Iznos **2**, 589 (1981).
11. A. A. Spektor, Prikl. Mat. Mekh. **51**, 76 (1987).
12. L. A. Galin, *Contact Problems in Theory of Elasticity and Viscoelasticity* (Nauka, Moscow, 1980).

Translated by Yu. Verevochkin