========================= **PHYSICS** =========================

Diffusion in Systems with Slowly Relaxing Jahn–Teller Centers

M. A. Ivanov*, Corresponding Member of the RAS É. A. Pastukhov**, and A. Ya. Fishman**

Crystalline materials with Jahn–Teller centers are ideal for simulating diffusion processes in systems with a multiwell potential. The Jahn–Teller mechanism of interparticle interaction allows the description of various phase transformations, including the transition to Jahn–Teller glass [1, 2], structure transitions with ferromagnetic-distortion and antiferromagnetic-distortion ordering [3], spin-reorienting transitions [4, 5], phase decay [6], etc. In previous studies [7, 8], it was shown that the presence of multiwell-potential centers in a crystal lattice can substantially affect the magnitude and shape of diffusion potential barriers. In this case, the migration of diffusing atoms was considered as for nondegenerate systems under the assumption of a statistically average (i.e., thermodynamically equilibrated) configuration of the crystalline surroundings. At the same time, the properties of degenerate systems depend strongly on the characteristic time of relaxation of degenerate centers and on the distribution function for this time. For this reason, the diffusion coefficients for anions in crystals are theoretically analyzed here under the assumption that there is no time to establish the equilibrium population of degenerate levels in a Jahn–Teller subsystem during a diffusion migration. In this case, the symmetry of the anion surroundings of a diffusing atom at a saddle point is determined by a random configuration of vibron states of the nearest Jahn–Teller centers (see Figs. 1–3). The displacements (corresponding to these states) of anions from the symmetric positions give rise to the change in the shape of the potential barrier and, therefore, in activation energy for the migration frequencies. Taking into account the multiwell shape of the potential energy of anions and the characteristic scale of Jahn–Teller strains (10^{-2}–10^{-1} for $3d$ ions), one may expect that the above effects can give rise to both qualitative and quantitative changes in the diffusion coefficients.

The investigation was carried out for a spinel lattice, where anions form a close-packed fcc structure and cations occupy half of the octahedral sites and one-eighth of the tetrahedral sites [9]. In this case, each anion has three nearest octahedral cations (the bonds are directed along the coordinate axes) and one tetrahedral cation (the bond is directed along the trigonal axis). We assume that Jahn–Teller ions occupy only octahedral sites.

Let a Jahn–Teller center (Jahn–Teller cation plus anion octahedron, see Fig. 1) be characterized by three lowest vibron states each corresponding to the extension of the octahedron along one of the coordinate axes and compression along the other two axes [10]. Then, the position (displacement) of each anion is determined by the set of vibron wave functions specified at the three nearest Jahn–Teller octahedral cations (see Fig. 2). There are nine such states of the anion. In this case, if Jahn–Teller deformations are accompanied by a small change in volume, the anion displacement u_{JT} along the bond to a Jahn–Teller cation in two of three possible vibron states of this cation is half of the displacement when the anion moves away from the cation in the third state.

It is obvious that, if Jahn–Teller states are random in diffusion migration, the corresponding distribution of activation energies for the frequencies of these migrations must also be random. In this case, there arise con-
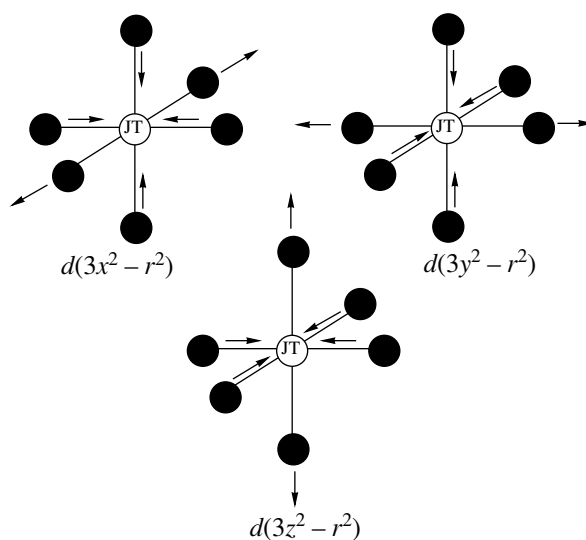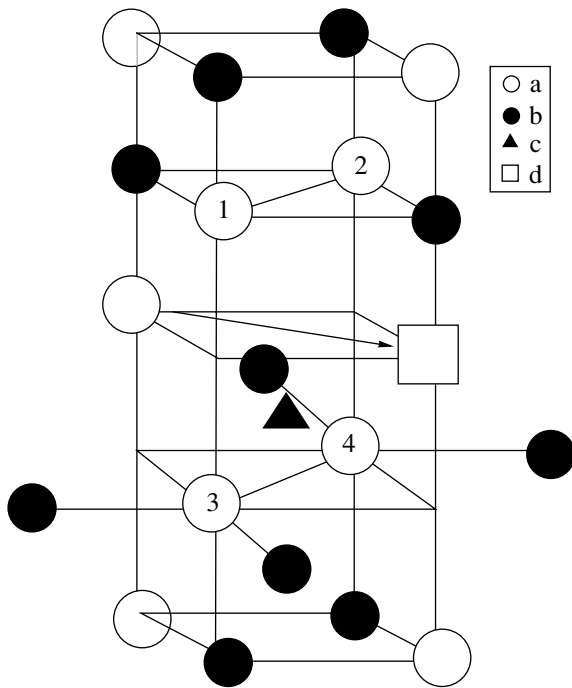

Fig. 1. Displacements of anions in the three lowest vibron states with the indicated symmetries of wave functions.

* *Institute of Metal Physics,*
  National Academy of Sciences of Ukraine,
  pr. Vernadskogo 36, Kiev, 252680 Ukraine
** *Institute of Metallurgy, Ural Division,*
  Russian Academy of Sciences, Yekaterinburg, Russia

416   IVANOV *et al.*

Fig. 2. Configurations of anions and cations near a saddle point: (open circle) anion, (closed circle) octahedral cation, (triangle) tetrahedral cation, and (square) anion vacation. Anions nearest to the saddle point are numbered.
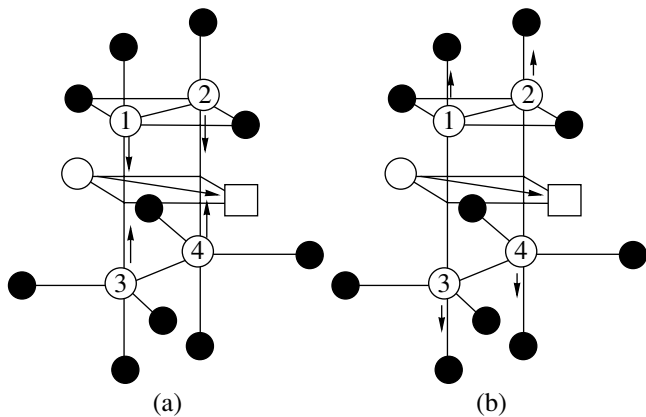
Fig. 3. Displacements of anions in two vibron configurations.

figurations that give rise to both a decrease and an increase in activation energies compared to the case where the Jahn–Teller displacements of anions in a crystal lattice are absent.

The above behavior is pronounced in the elastic model of diffusion potential barriers. In this model, activation energy is attributed to the work expended on deforming a lattice when a diffusing particle moves through a channel (passage) whose center is at a saddle point and whose size is, as a rule, smaller than the diffusing atom. The difference between these two sizes

determines the desired activation energy. This energy is proportional to the square of this size difference and must noticeably vary under rather small variations in the channel diameter that are induced by the Jahn–Teller displacements of anions. Figure 3 shows two possible configurations of anions for which changes in the sizes of the passage channel must be of different signs. For simplicity, only vibron states with cation-anion bonds oriented along the Z-axis are shown. We analyze the effect of these Jahn–Teller displacements caused by the multiwell shape of the potential energy on the diffusion coefficient of anions.

In the high-symmetric cubic phase of cooperative Jahn–Teller systems (i.e., for temperatures T above the structure-transition temperature T_D), the diffusion-coefficient tensor D_{ij} is isotropic, i.e., $D_{ij} = D^0\delta_{ij}$. In the low-symmetric phase, the diffusion coefficients can be anisotropic because of Jahn–Teller strains e_{JT}. The random walk method provides the following expressions for the corresponding D_{ii} [11]:

$$D_{xx} = \frac{1}{2}\sum_k \Gamma_k x_k^2, \quad D_{yy} = \frac{1}{2}\sum_k \Gamma_k y_k^2,$$

$$D_{zz} = \frac{1}{2}\sum_k \Gamma_k z_k^2, \tag{1}$$

where x_k, y_k, and z_k are the projections of the displacement of a diffusing atom onto the coordinate axes upon k-type migration and Γ_k is the frequency of these migrations. In this case, one can assume that migrations of each type are characterized by the Arrhenius temperature dependence

$$\Gamma_k = \Gamma_k^0 \exp\{-E_k/(k_B T)\}, \tag{2}$$

where Γ_k^0 is the pre-exponential factor and E_k is the potential-barrier height.

In the lattice under consideration, there are three basic types of migrations between the nearest positions. If Jahn–Teller strains are ignored, these types have the same configuration of anions and differ only in octa- and tetrahedral cations surrounding the saddle point. For brevity, we consider only one type of diffusion migration under the assumption that the migrations shown in Figs. 2 and 3 are determining. In this case, the diffusion coefficient has the form

$$D = D^0 \sum_i W_i \exp\{-\Delta E_i/k_B T\}, \tag{3}$$

where W_i is the probability of the ith configuration of vibron states of Jahn–Teller centers nearest to the saddle point, ΔE_i is the change in the potential-barrier height in the ith configuration caused by the displacements of anions, and D^0 is the diffusion coefficient in

the absence of Jahn–Teller interactions. The case where the relaxation of Jahn–Teller levels is slow compared to the characteristic time of the above-barrier migration of an anion is analyzed. For this reason, the anisotropy of the diffusion coefficients of anions is not considered.

For simplicity, we analyze the variation of activation energy ΔE_i in the elastic model of diffusion potential barriers [12, 13]. In this case,

$$
\Delta E_i = K_{\mathrm{eff}} v_{\mathrm{eff}}
$$
$$
\times \left\lfloor (R_{\mathrm{pol}}^i - R_{\mathrm{anion}})^2 - (R_{\mathrm{pol}}^0 - R_{\mathrm{anion}})^2 \right\rfloor / R_{\mathrm{anion}}^2, \qquad (4)
$$

where R_{pol}^i is the passage radius in the ith configuration; R_{pol}^0 is this radius in the configuration undisturbed by Jahn–Teller configurations; R_{anion} is the radius of a diffusing anion; and K_{eff} and v_{eff} are the effective values of the elastic modulus and volume, respectively, which describe the work expended on deforming the lattice upon diffusion migration ($v_{\mathrm{eff}} \sim 4\pi R_{\mathrm{anion}}^3/3$). In this case, it is evident that the changes in the potential barrier heights alternate in sign (see Fig. 3). In the stretched rectangle involving atoms 1–4, the passage diameter increases, and therefore the activation energy of migration decreases ($\Delta E_i < 0$). In contrast, activation energy increases ($\Delta E_i > 0$) in the contracted rectangle.

In systems with a small number of Jahn–Teller bonds of anions nearest to a saddle point, the entire spectrum of possible changes in energy ΔE_i can be easily exhausted. This case can be simulated by taking into account only those displacements of anions 1–4 indicated in Fig. 3 which are determined by Jahn–Teller bonds oriented along the Z-axis. The most substantial increase in the channel diameter and decrease in activation energy must occur in the configuration shown in Fig. 3b. The probability of this configuration is quite high and is equal to 16/81. The dominant role of one such contribution to Eq. (3) must give rise to the simultaneous decrease in both activation energy and pre-exponential factor in the diffusion coefficient. In this case, the effect of decreasing activation energy must undoubtedly prevail.

In general, when the displacements of four oxygen atoms nearest to the saddle point are determined by electron states of ten Jahn–Teller cations in octahedral sites, 3^{10} configurations of atoms 1–4 in Fig. 2 should be considered (the number of Jahn–Teller ions in the cluster in question is 10). In view of this circumstance, the summation with respect to configurations i in Eq. (3) can be replaced by integration by introducing the distribution function $f(\Delta E)$ for the activation ener-

gies of migration frequencies. Then,

$$
D = D^0 \int_{-\infty}^{\infty} d\varepsilon\, f(\varepsilon) \exp\{-\varepsilon/k_{\mathrm{B}} T\},
$$

$$
D = D^0 \exp\left\{ \left(\frac{\Delta}{2 k_{\mathrm{B}} T} \right)^2 \right\}, \qquad (5)
$$

$$
\text{if} \quad f(\varepsilon) = \frac{1}{\Delta \sqrt{\pi}} \exp\left\{ -\left(\frac{\varepsilon}{\Delta} \right)^2 \right\}.
$$

For simplicity, the function $f(\Delta E)$ is taken in the Gaussian form. In this case, it is actually seen that the diffusion coefficient increases due to the multiwell shape of the potential relief ($D > D^0$). This effect is attributed to the fact that a diffusing particle under conditions where $\tau \Gamma^0 \gg 1$ (τ is the relaxation time of a Jahn–Teller center) migrates at the instant when the configuration with the lowest potential barrier is realized.

Temperature dependences of the diffusion coefficients similar to Eq. (5) can also occur in nondegenerate systems with random deformation fields caused by defects [8]. However, the conditions determining the domain of existence of this dependence differ fundamentally from the above conditions. In nondegenerate systems, in order to obtain this dependence, the interaction between a diffusing particle and defects must be long-range sign-alternating and lower than $k_{\mathrm{B}} T$ for any allowable distances in the lattice. In the presence of short-range interactions such as blocking or capture of a diffusing atom by defects, the above effect was absent.

By using Eq. (4), the dispersion Δ in Eq. (5) is easily estimated as

$$
\Delta \cong E^0 \left| \frac{u_{\mathrm{JT}}}{R_{\mathrm{pol}}^0 - R_{\mathrm{anion}}} \right|. \qquad (6)
$$

If the difference $R_{\mathrm{anion}} - R_{\mathrm{pol}}^0$ between the anion size and the size of the passage at the saddle point is comparable with Jahn–Teller displacements $u_{\mathrm{JT}} \sim (10^{-1}\text{--}10^{-2}) R_{\mathrm{anion}}$, D must increase considerably compared to D^0 in the temperature range $k_{\mathrm{B}} T < \Delta$ ($\Delta \geq 10^{-1}$ eV).

We now discuss the possibility of observing the above effect. The diffusion of anions is usually observed at temperatures considerably above room temperature. At these temperatures, the condition $\tau \Gamma^0 > 1$ can be satisfied in both low-symmetric and high-symmetric phases of cooperative Jahn–Teller systems with the transition temperature $T_D \sim (10^2\text{--}10^3)$ K.[1] The presence of the cooperative Jahn–Teller effect ensures diffusion in a crystal through fast channels of

---

[1] The relaxation time τ of a multilevel system decreases with increasing temperature and splitting energies of degenerate levels [10].

anion migration, and the relatively high temperature provides quite large local Jahn–Teller displacements.

The results can be verified experimentally by replacing Jahn–Teller ions in the systems under consideration with orbitally nondegenerate ions and with other types of replacements. These replacements must result in a noticeable decrease in the diffusion coefficients. This decrease is associated with the decrease in the total number of degenerate centers or in the concentration of degenerate centers with relatively long relaxation times. As a result, the dispersion of activation energies in Eq. (5), which is related to various configurations of the Jahn–Teller displacements of anions, decreases, and therefore the diffusion coefficient D also decreases.

Note that the anisotropic deformation of the lattice in the low-symmetric phase of the Jahn–Teller system for $\tau\Gamma^0 < 1$ can also give rise to a noticeable variation of the diffusion coefficients D_{ij} at least in certain directions. For replacements resulting in the suppression of Jahn–Teller ordering, this effect gradually vanishes. In contrast to this case, noticeable anisotropy of the diffusion coefficients is absent in the case where $\tau\Gamma^0 > 1$ and the effect remains in the high-temperature phase.

ACKNOWLEDGMENTS

REFERENCES

1. M. A. Ivanov, V. Ya. Mitrofanov, and A. Ya. Fishman, Fiz. Tverd. Tela (Leningrad) **20** (10), 3023 (1978) [Sov. Phys. Solid State **20**, 1744 (1978)].

2. F. Mehran and K. W. H. Stevens, Phys. Rev. B **27** (5), 2899 (1983).

3. J. Kanamori, J. Appl. Phys. **31** (5), 14S (1960).

4. M. Kataoka and J. Kanamori, J. Phys. Soc. Jpn. **32** (1), 113 (1972).

5. K. I. Kugel' and D. I. Khomskiĭ, Usp. Fiz. Nauk **136** (4), 621 (1982) [Sov. Phys. Usp. **25**, 231 (1982)].

6. V. F. Balakirev, V. P. Barkhatov, Yu. V. Golikov, and S. G. Maĭzel', *Manganites: Equilibrium and Unstable States* (Yekaterinburg, 2000).

7. A. Ya. Fishman, M. A. Ivanov, V. Ya. Mitrofanov, and V. B. Vykhodets, Z. Phys. Chem. **201**, 285 (1997).

8. V. B. Vykhodets, M. A. Ivanov, V. Ya. Mitrofanov, *et al.*, Fiz. Met. Metalloved. **77** (3), 112 (1994).

9. S. Krupička, *Physik der Ferrite und der verwandten magnetischen Oxide* (Academia, Prague, 1973; Mir, Moscow, 1976), Vol. 1.

10. F. S. Ham, in *Electron Paramagnetic Resonance* (Plenum, New York, 1972), pp. 1–119.

11. J. R. Manning, *Diffusion Kinetics for Atoms in Crystals* (Van Nostrand, New York, 1968; Mir, Moscow, 1971).

12. Y. A. Bertin, J. Parisot, and J. L. Gacougnolle, J. Less-Common Met. **69**, 121 (1980).

13. J. D. Eshelby, Solid State Phys. **3**, 79 (1956).

Translated by R. Tyapaev

$=$ **PHYSICS** $=$

Conductance Anisotropy of δ-Si Doped GaAs Layers Grown by Molecular Beam Epitaxy on (111)A GaAs Substrates and Misoriented in the [2̄1̄1̄] Direction

G. B. Galiev, Corresponding Member of the RAS V. G. Mokerov, V. A. Kul'bachinskiĭ,
V. G. Kytin, R. A. Lunin, A. V. Derkach, and I. S. Vasil'evskiĭ

Epitaxial films with quasi-one-dimensional and one-dimensional structures have recently been actively investigated. A promising method for preparing them is the growth of epitaxial structures (ESs) by molecular-beam epitaxy (MBE) on the vicinal surface of GaAs with δ-doped layers [1–3]. Previously, for this purpose, tin was used as an electrically active impurity for δ-doping. Due to the significant difference in the atomic radius of tin compared to that of gallium and to high diffusion mobility, tin segregates [4] and is preferentially accumulated at the edges of steps [1, 5]. For this case, epitaxial structures δ-doped with tin (so-called δ-Sn doping) were synthesized and investigated. The results of relevant studies are presented in [6, 7].

In this paper, the conductance anisotropy of δ-Sn-doped ESs was studied for ESs grown by the MBE method on (111)A GaAs substrates misoriented by the angles 0.5°, 1.5°, and 3° in the [2̄1̄1̄] direction. As a dopant impurity in GaAs, silicon possesses amphoteric properties, manifesting them most strongly in the case of the (111)A substrate [8]. It is well known that in GaAs with the (100) orientation, silicon predominantly behaves as a donor. At the same time, in ESs grown on (111)A GaAs substrates, both strongly compensated semi-insulating layers and either n- or p-layers can be obtained [8–12] depending on the growth temperature T_g and ratio γ of arsenic and gallium fluxes. (Here, $\gamma = P_{As}/P_{Ga}$, where P_{As} and P_{Ga} are the partial pressures of As and Ga for the ES growth region in the MBE setup.) In addition, it is well known [8] that for (111)A GaAs substrates misoriented in the [2̄1̄1̄] direction by a small angle α ranging from 0.5° to 3°, a vicinal surface arises with terraces of (111)A orientation and steps of (100) orientation. The arrangement of Ga and As atoms on the vicinal surface is shown in Fig. 1 for the indicated misorientation direc-

tions. When growing a δ-Si-doped layer on this surface, we can expect the predominant doping of steps with (100) orientation by donor silicon. The terraces with (111)A orientation are doped differently depending on the growth conditions, so that we can expect the formation of Si atomic chains on the steps of such a vicinal surface. These structures are promising from the standpoint of the development of so-called quantum wires on their basis.

To check the nonuniformity of silicon doping of the vicinal (111)A GaAs surface, the conductance anisotropy of such structures was studied. The samples were grown using a TsNA-24 MBE setup (made in Russia) on semi-insulating (111)A GaAs substrates misori-
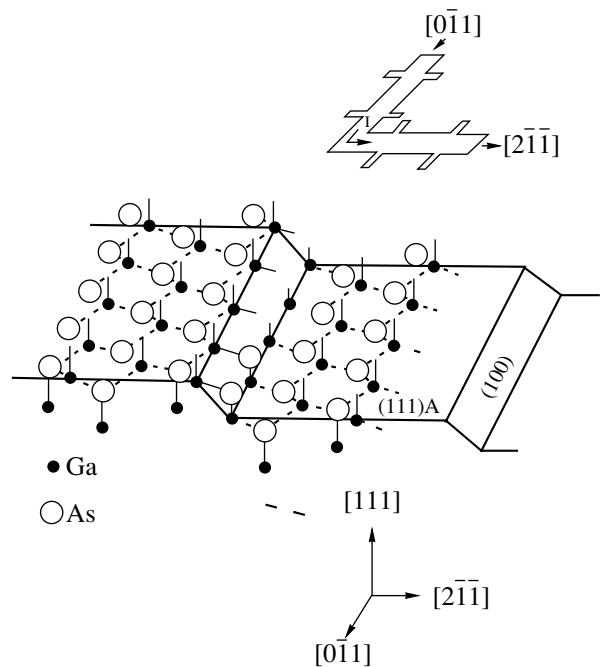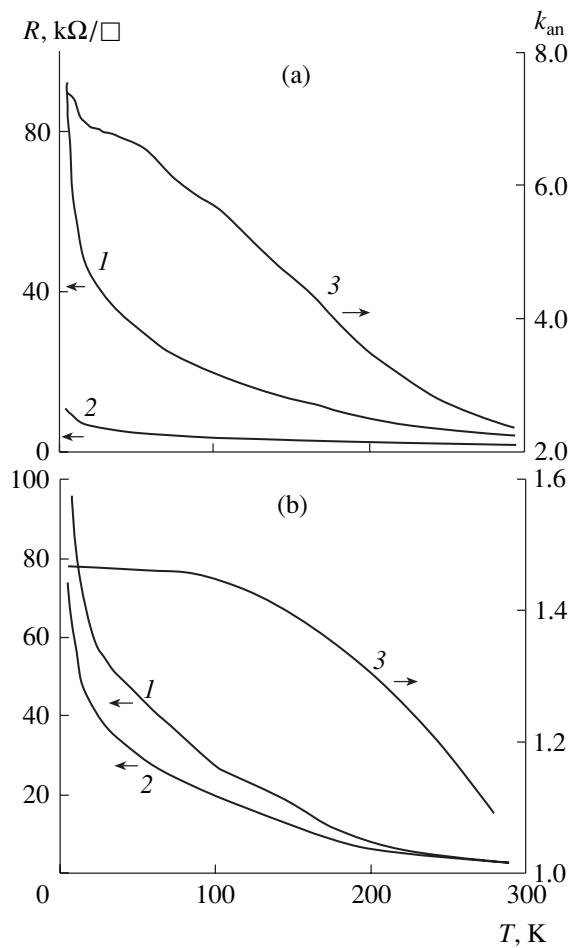

Fig. 1. Schematic arrangement of Ga and As atoms on the (111)A surface for misorientation of substrates in the [2̄1̄1̄] direction and schematic sketch of the Hall bridge for measuring the resistance anisotropy.

Institute of Radio Engineering and Electronics,
Russian Academy of Sciences,
ul. Mokhovaya 11, Moscow, 103907 Russia

Fig. 2. Temperature dependences of the resistance for samples (a) **2** and (b) **3** in the (*1*) $[2\bar{1}\bar{1}]$ and (*2*) $[0\bar{1}1]$ directions, respectively, and (*3*) the anisotropy coefficient $k_{an} = R_{pe}/R_{pa}$.

ented from the (111)A plane in the $[2\bar{1}\bar{1}]$ direction by the angles $\alpha = 0.5°$, $1.5°$, and $3°$. All of these samples, as well as a test sample with (100) orientation, were grown simultaneously with each other. For this purpose, the indicated substrates were glued with indium onto a common molybdenum holder. The structures that were grown included an undoped buffer GaAs layer 0.42 μm thick, a δ-Si layer, an undoped GaAs layer 50 nm thick, and a homogeneously doped GaAs layer 30 nm thick with a Si concentration of ~10^{18} cm^{-3} for filling surface states. The epitaxial growth was performed at a temperature $T_g = 610°C$, and the value of γ was 14. After the growth had been completed, the samples were prepared by the photolithography method in the form of an *L*-type Hall bridge for further measurements. A view of this bridge is shown in Fig. 1.

The resistance of the structures was simultaneously measured by four probes parallel and perpendicular to the step faces of the vicinal surface. Temperature dependences of the resistance were measured within the range 4.2–300 K. The Hall effect and magnetoresistance were investigated in magnetic fields up to 0.5 T. The measurement results are listed in the table.

Nonuniform distribution of the impurities on the vicinal surface results in anisotropy of the resistance measured along and across the steps. As an example, temperature dependences of the resistance for samples **2** and **3** in the $[0\bar{1}1]$ direction (*pa*-direction) and the $[2\bar{1}\bar{1}]$ direction (*pe*-direction), as well as the coefficient of the resistance anisotropy $k_{an} = R_{pe}/R_{pa}$, are shown in Figs. 2a and 2b. For all vicinal samples, the resistance R_{pe} perpendicular to step edges is much larger than the resistance R_{pa} parallel to these edges. The coefficient k_{an} of the resistance anisotropy increases as the temperature decreases similar to vicinal GaAs-structures with Sn-δ doping studied previously [6, 7]. Conductance anisotropy was not found in test sample **1** with (100) orientation.

Resistance anisotropy is associated with different effects of the nonuniform impurity distribution on the carrier mobility for various current directions. The carrier mobility for motion along the steps of the vicinal surface is mainly determined by scattering on ionized impurities randomly located along the terraces. Since silicon is predominantly accumulated at the edges of steps, if the current direction is along the steps, the dispersion of step widths [1] and the random formation of steps with a height of several monolayers [13] significantly affect the carrier mobility. This results in the appearance of a periodic potential, which additionally scatters electrons [14, 15].

The values of R_{pa}, k_{an}, and Hall concentrations n_H for the samples at temperatures of 300, 77, and 4.2 K are listed in the table. The resistance of the structures

Table

Sample	300 K			77 K			4.2 K		
	R_{pa}, ω/η	k_{an}	n_H, 10^{12} cm^{-2}	R_{pa}, ω/η	k_{an}	n_H, 10^{12} cm^{-2}	R_{pa}, ω/η	k_{an}	n_H, 10^{12} cm^{-2}
1, (100)	305	1.0	−12	280	1.0	−11.6	293	1.0	−10.4
2, $\alpha = 0.5°$	1750	2.3	28	3910	6.0	3.3	10700	7.4	−
3, $\alpha = 1.5°$	2600	1.1	21	23100	1.5	3.6	73800	1.5	−
4, $\alpha = 3.0°$	3540	1.0	21	23700	1.0	3.6	139000	−	−

$\ln R$
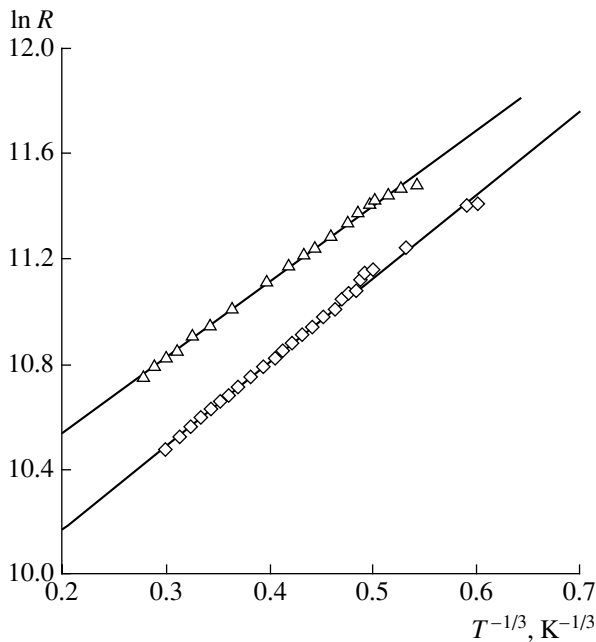
Fig. 3. Resistance logarithm for samples **2** (squares) and **3** (triangles) as a function of $T^{-1/3}$ according to measurements along the $[2\bar{1}\bar{1}]$ direction.

increases and the resistance anisotropy decreases with an increase in the misorientation angle α from 0.5° to 3°. This indicates enhancement of the contribution independent of the carrier scattering direction at large misorientation angles.

Studies of the Hall effect have shown that sample **1** possesses n-type conductance with an electron mobility of 2000 cm^2/V s and that all the vicinal samples have p-type conductance with a hole mobility of about 80 cm^2/V s that weakly decreases with temperature. Thus, for the chosen regimes of epitaxial growth on vicinal surfaces with (111)A orientation, silicon behaves predominantly as an acceptor. The Hall concentration in samples **2–4** decreases by almost a factor of 6 from room to nitrogen temperatures (see table). At helium temperature, we were unable to measure the Hall coefficient.

In the low-temperature region (below 50 K), the resistance of the samples obeys the Mott law for hopping conductivity in the two-dimensional case: $\rho = \rho_0 \exp\left(\dfrac{T_0}{T}\right)^{1/3}$ (Fig. 3), i.e., the charge carriers are localized on the nonuniformities of the potential relief. The parameter T_0 is equal to 31 and 23 K for a current direction across the steps and 22 and 18 K for a current direction along the steps for samples **2** and **3**, respectively. This parameter is associated with both the density of states at the Fermi level and the localization radius $T_0 = C(N_{E_{\mathrm{F}}} a^2)^{-1}$, where $C = 13.8$ is a numerical coefficient. The localization radius a for sample **2** obtained in

this manner is approximately 60 and 72 for current directions across and along the steps, respectively. In other words, in the pe-direction, the holes are more strongly localized, which is again testifies to relief anisotropy.

Thus, the experiments carried out have shown that the nonuniform distribution of silicon in δ-doped GaAs layers grown by the MBE method on (111)A GaAs substrates misoriented in the $[2\bar{1}\bar{1}]$ direction results in anisotropy of the resistance both along and across the steps of the vicinal surface.

ACKNOWLEDGMENTS

REFERENCES

1. B. Etienne, F. Lelarge, Z. Z. Wang, and F. Laruelle, Appl. Surf. Sci. **113/114**, 66 (1997).
2. L. Daweritz, K.-J. Friedland, J. Behrend, and P. Schutzendube, Phys. Status Solidi A **146** (1), 277 (1994).
3. Y. Tokura, T. Saku, and Y. Horikoshi, Phys. Rev. B **53** (16), R10528 (1996).
4. J. J. Harris, D. E. Ashenford, C. T. Foxon, *et al.*, Appl. Phys. A **33**, 87 (1984).
5. G. S. Petrich, A. M. Dabiran, J. E. Macdonald, and P. I. Cohen, J. Vac. Sci. Technol. B **9** (4), 2150 (1991).
6. A. D. Visser, V. I. Kadushkin, V. A. Kul'bachinskiĭ, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **59**, 339 (1994) [JETP Lett. **59**, 363 (1994)].
7. V. A. Kul'bachinskiĭ, V. G. Kytin, V. I. Kadushkin, and A. P. Senichkin, Fiz. Tverd. Tela (St. Petersburg) **37**, 2693 (1995) [Phys. Solid State **37**, 1481 (1995)].
8. Y. Okano, M. Shigeta, H. Seto, *et al.*, Jpn. J. Appl. Phys. **29**, L1357 (1990).
9. F. Piazza, L. Pavesi, M. Henin, and D. Johnston, Semicond. Sci. Technol. **7**, 1504 (1992).
10. G. B. Galiev, M. V. Karachevtseva, V. G. Mokerov, *et al.*, Dokl. Akad. Nauk **367**, 613 (1999) [Dokl. Phys. **44**, 510 (1999)].
11. G. B. Galiev, V. G. Mokerov, and Yu. V. Khabarov, Dokl. Akad. Nauk **376**, 749 (2001) [Dokl. Phys. **46**, 88 (2001)].
12. G. B. Galiev, V. G. Mokerov, É. R. Lyapin, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **35** (4), 421 (2001) [Semiconductors **35**, 409 (2001)].
13. F. Lelarge, Z. Z. Wang, A. Cavanna, *et al.*, Europhys. Lett. **39** (1), 97 (1997).
14. Y. Nakamura, S. Koshiba, and H. Sakaki, J. Cryst. Growth **175/176**, 1092 (1997).
15. Y. Tokura, Phys. Rev. B **58** (11), 7151 (1998).

Translated by T. Galkina

===== **PHYSICS** =====

Effect of Dispersion Strengthening of a Catalyst on the Properties of Polycrystalline Carbonado Diamonds

Academician A. V. Elyutin, A. A. Ermolaev, A. I. Laptev, and A. V. Manukhin

Received December 10, 2001

Polycrystalline carbonado diamonds are synthesized from a carbon material in the presence of a nickel alloy catalyst. A polycrystal is a composite material that consists of a diamond phase and a binder (up to 20%) whose composition is close to the catalyst. The properties of the binder significantly affect the properties of the composite material as a whole. The strengthening properties of nickel alloys can be improved by doping them with disperse particles.

In this study, we determined the conditions for forming the dispersion-strengthened binder in a polycrystalline diamond. Carbonado samples were synthesized in a high-pressure toroid-type chamber at an initial pressure of 8.0 GPa and a temperature of 1800–2000 K in 12 s. The mass of the resulting polycrystal was 0.8 carat. We used MGOSCh graphite as the carbon material.

In this study, we also used ultrafine powders (UDPs) that were obtained by the plasmochemical method and had the following characteristics:

nickel and molybdenum (MIKhM) with an average particle size of 100 nm and an oxygen content of 3 and 5%, respectively;

titanium nitride (INKhP, Chernogolovka) with an average particle size of 70 nm.

Ultrafine powders were preliminarily mixed in a barrel mixer. To distribute the components more uniformly, the ultrafine-powder mixture was additionally treated with high-frequency (22 kHz) ultrasound in hexane for 1 min. The mixture was further compacted under a pressure of 800 MPa into rods of the desired sizes.

The compacts were used to synthesize polycrystalline carbonado diamond. However, polycrystals were not formed because the catalyst contained a large quantity of gaseous impurities, primarily oxygen, hydrogen, nitrogen, carbon dioxide, and carbon monoxide [1]. Annealing the compacts in hydrogen at a temperature of 1070 K for 1 h leads to the removal a great deal of

the gaseous impurities and to sintering of the ultrafine powders with an increase in the strength of the samples. The ability of the compacts to form polycrystalline diamond is totally restored and is comparable with the reaction ability of a compact catalyst.

Strengthening disperse particles can be distributed more uniformly and their size can be reduced by using the internal nitriding method [2] to produce dispersion-strengthened catalysts. For nitriding, we used sintered (Ni–15Mo)–Ti samples obtained by mouthpiece compaction. Nitriding was performed in a nitrogen-filled gasostat under a pressure of 150 MPa and at a temperature of 1250 K for 1 h. In the process of nitriding, titanium, which was previously in the nickel solid solution, interacts with nitrogen and, as a result, forms disperse particles of TiN. In this case, due to the low thermodynamic stability of molybdenum nitride [2], molybdenum remains in the γ-solid solution and conserves the nickel strengthened alloyed solid solution.

The porous structure of the sintered samples (a porosity of ~15%) allows the easy propagation of nitrogen over the entire volume in the process of nitriding. As a result, uniformly distributed disperse TiN precipitate is formed in the γ-solid solution matrix. To determine the TiN lattice constant, we used the reflection from the (220) plane and obtained the value $(4.243 \pm 0.003) \times 10^{-10}$ m corresponding to the stoichiometric composition of TiN [3]. The average size of the disperse TiN particles was determined by the X-ray method from widening of the (220) diffraction line and was equal to 30–40 nm. The method of determining the dispersivity of particles in such dispersion-strengthened materials is well known [4].

The quantitative content of TiN in the nitrided samples was determined by various methods.

For the quantitative X-ray phase analysis, we used the calibration curve method [5]. The standard samples were mixtures of powders of nickel and ultrafine TiN with nearly stoichiometric composition, a lattice constant of $(4.240 \pm 0.003) \times 10^{-10}$ m, and an average particle size of approximately 70 nm. The ratio of intensities of the (220) diffraction lines of TiN and nickel in the samples was determined at least three times for each composition. The results of the quantitative X-ray phase analysis of nitrided samples are as follows: the

Moscow State Institute of Steel and Alloys
(Technical University),
*Leninskiĭ pr. 4, Moscow, 117936 Russia*

content of titanium in the initial powder mixture was equal to 2, 3, 4, 6, and 10%, whereas the content of titanium nitride in the nitrided samples was equal to 1.6 ± 0.5, 2.4 ± 0.6, 2.9 ± 0.6, 3.9 ± 0.7, and $6.1 \pm 0.8\%$.

The nitrogen content in the nitrided samples was determined by the meltback method on a LECO TC-136 installation. The nitrogen content was determined with an accuracy of 2%. The amount of TiN was calculated from the stoichiometry of its composition under the assumption of the absence of dissolved nitrogen in the nickel solid solution. The last assumption does not introduce any substantial error into the calculation, because nitrogen solvability in solid nickel is less than 0.07% [2]. The results are listed in the table. A comparison between the data described above and listed in the table shows that the results obtained for the titanium nitride content in the nitrided catalysts by two different methods coincide within the errors.

After synthesizing polycrystalline diamonds, we compared their strength properties. The polycrystals were crushed, a 400/315 fraction was sifted, and the crushing strength was determined according to the GOST 9206-80 method using the PA-4E installation. The results of the tests were processed by mathematical statistics methods. For a confidence level of 95% and a number of measurements of at least 100, the confidence interval for strength values did not exceed 6 N (the relative confidence deviation was no more than 8%).

The results of the strength tests of the diamond polycrystals obtained by using dispersion-strengthened catalysts are shown in Figs. 1–3. It should be noted that the effect of strengthening of the diamond polycrystals is observed if the disperse phase is distributed reasonably uniformly in the initial catalyst. In particular, if compacted powder mixtures are used as the catalysts, the synthesized polycrystals are strengthened only in the case of additional treatment of an ultrafine-powder mixture by ultrasound (Fig. 1, curve 4). If coarse-grained powders of nickel and molybdenum are used, the additions of both ultrafine and coarse-grained TiN powders only slightly affect the strength of the synthesized diamond polycrystals. The highest strengthening of carbonado polycrystals is achieved by using the dispersion-strengthened catalysts obtained by the internal nitriding method (Fig. 2, curve 2). In such catalysts, the TiN strengthening phase is more dispersed and uniformly distributed.

The form of experimental strengths of diamond polycrystals as functions of disperse TiN admixtures in the initial catalyst is likely caused by the effect of disperse particles on the strength of the nondiamond component of the polycrystals. Strengthening of the metal phase by noncoherent particles takes place, as a rule, according to the Orowan mechanism [4]. In this case, the fraction of the strengthening phase generally must not exceed several percent, whereas the distances between particles of the phase and their sizes must differ by an order of magnitude or more. It should be noted
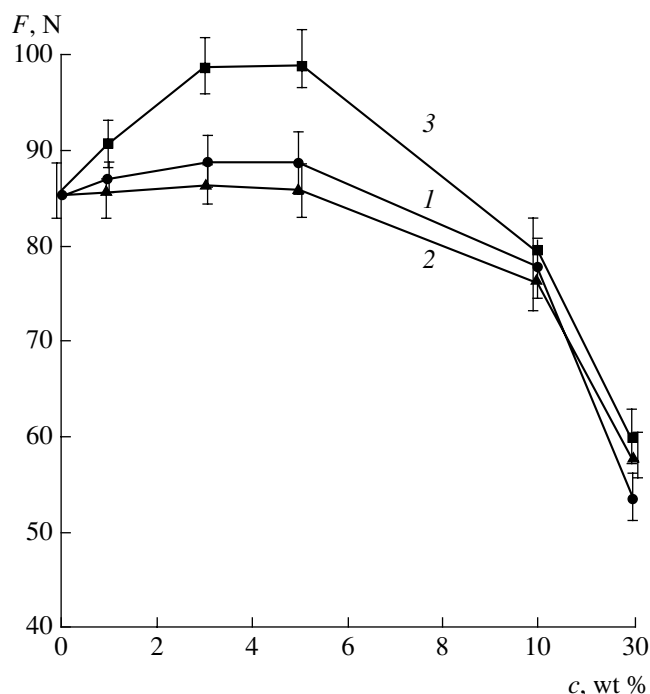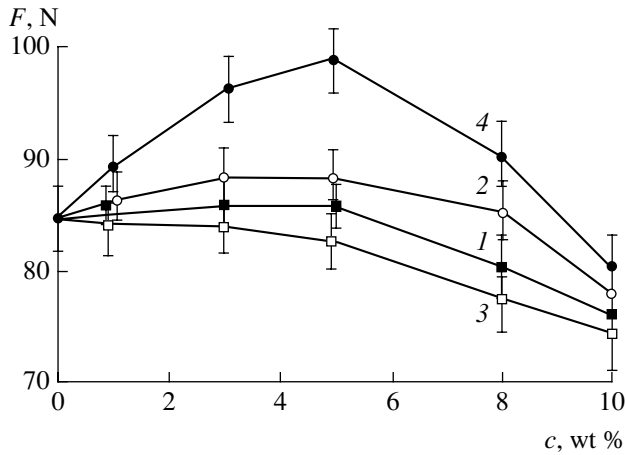
Fig. 1. The strength F of ARK4 400/315 diamond powders vs. the Al_2O_3 content c in the Ni–15Mo catalyst for (1) coarse-grained Ni, Mo, and Al_2O_3 (32/20) powders; (2) ultrafine Ni, Mo, and Al_2O_3 powders mixed in a mixer; and (3) the same as 2, but the mixture is additionally treated by ultrasound.

that the nonuniformity of the distribution of disperse particles significantly reduces the degree of strengthening [4].

The above results make it possible to explain both the effect of dispersivity and the uniformity of the strengthening-phase distribution in the initial catalyst on the degree of strengthening of the synthesized diamond polycrystals and the character of concentration strength dependences with a maximum. In particular, the large fraction of disperse phase violates the optimal

Results of quantitative analysis of nitrided samples by the meltback method

Content of titanium in the initial powder mixture, %	Content of nitrogen in the nitrided samples, %	Calculated TiN content in the nitrided samples, %
0.5	0.14	0.62
1	0.25	1.1
2	0.42	1.9
3	0.54	2.4
4	0.68	3.0
6	0.99	4.4
10	1.5	6.6

ELYUTIN *et al.*

Fig. 2. The same as in Fig. 1, but vs. the titanium content in (*1*) and (*2*) a catalyst consisting of coarse-grained Ni–15Mo and ultrafine TiN powders, (*3*) and (*4*) a catalyst consisting of ultrafine Ni–15Mo and TiN powders, (*2*) and (*4*) mixtures additionally treated by ultrasound.
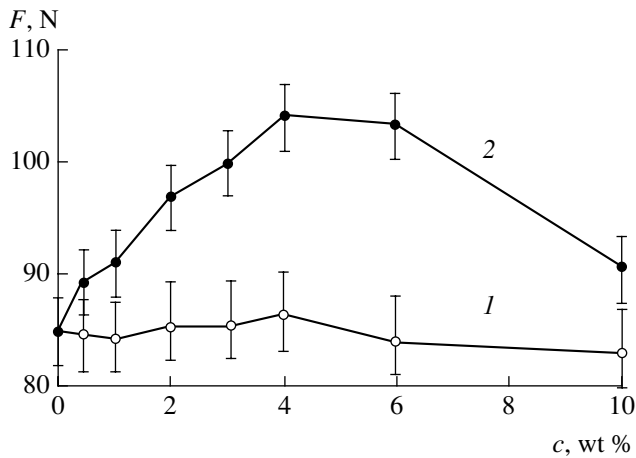
Fig. 3. The same as in Fig. 1, but vs. the titanium content in (*1*) the synthesized catalyst and (*2*) the nitrided catalyst.

geometric parameters for dispersion strengthening, embrittles the nondiamond component, and therefore worsens the strength properties of the polycrystal.

CONCLUSIONS

(i) A method for growing polycrystalline diamonds with a dispersion-strengthened intercrystalline binder was developed.

(ii) On the basis of the methods of powder metallurgy and internal nitriding, a method of producing a dispersion-strengthened catalyst with a certain porosity and composition was developed for synthesizing polycrystalline diamonds.

REFERENCES

1. R. U. Kalamazov, Yu. V. Tsvetkov, and A. A. Kal'kov, *Ultrafine Powders of Tungsten and Molybdenum* (Metallurgiya, Moscow, 1988).

2. S. S. Kiparisov and Yu. V. Levinskiĭ, *Internal Oxidation and Nitriding of Alloys* (Metallurgiya, Moscow, 1979).

3. B. V. Mitrofanov, Yu. G. Zaĭnulin, S. I. Alyamovskiĭ, *et al.*, Izv. Akad. Nauk SSSR, Neorg. Mater. **10**, 745 (1974).

4. K. I. Portnoĭ and B. N. Babich, *Dispersion-Strengthened Materials* (Metallurgiya, Moscow, 1974).

5. *Crystallography, Radiography, and Electron Microscopy*, Ed. by Ya. S. Umanskiĭ, Yu. A. Skakov, A. N. Ivanov, and L. N. Rastorguev (Metallurgiya, Moscow, 1982).

Translated by V. Bukhanov

=================== **PHYSICS** ===================

Nonmonotonic Concentration Dependence of T_c in the Van Hove Scenario

A. N. Kozlov and Corresponding Member of the RAS L. A. Maksimov

Received January 30, 2002

The theory of high-T_c superconductivity extensively uses a scenario considering competitive normal-state instabilities: spin-density wave (SDW), charge-density wave, and Cooper instability (primarily against d pairing). Many numerical and other nonperturbative calculations were performed in the two-dimensional Hubbard model with migrations over the nearest neighbors [1–4]. In this model, the dominant instability is apparently the SDW instability caused by nesting, which is substantial at doping degrees for which the Fermi surface passes near Van Hove singularities of the electron spectrum. In this scenario, an alternative mechanism of superconductivity is d pairing induced by the closeness of a system to the point of the quantum transition to the state with a spin-density wave [5–13]. Continuing study [8], we here consider T_c as a function of doping in the Hubbard model with weak repulsion in the carrier-concentration range where the Cooper amplitude depends on the closeness of a system to the transition to the state with a spin-density wave.

The following excitation spectrum of a two-dimensional electron system for a square lattice with migrations over the nearest neighbors,

$$\gamma_{\mathbf{p}} = 4t \cos \frac{p_x a}{2} \cos \frac{p_y a}{2}, \qquad (1)$$

has saddle points $\left(\pm \dfrac{\pi}{a}, \pm \dfrac{\pi}{a} \right)$ at the vertices of the so-called magnetic Brillouin zone. Using the jump integral t and lattice constant a as units of measurement and measuring quasimomenta from saddle points, e.g., $p = (-\pi + x, -\pi + y)$, we have $\gamma_{\mathbf{p}} = xy$ near such a point. If the Fermi surface passes near saddle points, which corresponds to low hole concentrations and $\mu \ll t$, and interaction is weak, it is customary to speak of the Van Hove scenario. The relation $\gamma_{\mathbf{p}} + \gamma_{\mathbf{p} + \mathbf{Q}} = 0$, where vector $\mathbf{Q}$ is any side of the magnetic zone, is referred to as nesting, which is one of the aspects of the Van Hove scenario.

For scattering on the Fermi surface by vector $\mathbf{Q} + \mathbf{q}$, where $\mathbf{q} \ll \mathbf{Q}$, the polarization operator

$$\Pi_0(\mathbf{q}, 0) = \frac{1}{2\pi^2} \iint\limits_{xy < \mu} \frac{2 dx dy}{xy + (x + q_x)(y + q_y)} \qquad (2)$$

has a double logarithmic singularity, because the chemical potential measured in units of t is $\mu \ll 1$. In fact, the denominator of the integrand has a minimum at the point $x = -\dfrac{q_y}{2}, y = -\dfrac{q_x}{2}$ and, with logarithmic accuracy, we obtain the expression

$$\Pi_0(\mathbf{q}, 0) = \frac{1}{2\pi^2} \int\limits_{\frac{q_y}{2}}^{1} \frac{dx}{x} \int\limits_{\frac{q_x}{2}}^{1} \frac{dy}{y} = \frac{1}{8\pi^2} \ln \frac{1}{q_x^2} \ln \frac{1}{q_y^2}. \qquad (3)$$

In this case, the minima of q_x^2 and q_y^2 are equal to μ.

For weak Hubbard repulsion U (measured in units of t), the ladder amplitude in the SDW channel,

$$\Gamma = \frac{U}{1 - \eta_x \eta_y}, \quad \eta_{x, y} = \sqrt{g} \ln \frac{1}{q_{x, y}^2}, \quad g = \frac{U}{8\pi^2} \qquad (4)$$
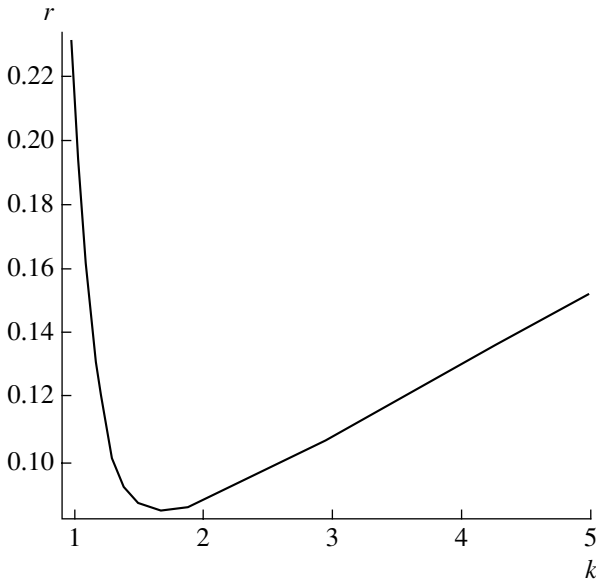
has a singularity at $k = \sqrt{g} \ln \dfrac{1}{\mu} = 1$.

In the logarithmic approximation, it is possible to write a closed ladder equation for the mass operator on the mass shell $\Big[$ displacement from the mass shell gives rise to relatively small corrections of the order of $\left(\ln \dfrac{1}{\mu} \right)^{-1} \Big]$, more exactly, for the factor

$$Z_{\mathbf{p}} = 1 - \frac{\partial}{\partial \varepsilon} \sum (\mathbf{p}, \varepsilon = \gamma_{\mathbf{p}} - \mu) = 1 + \Lambda_{\mathbf{p}}. \qquad (5)$$

Let us consider the following Π operator with the exact Green's functions $G = (\varepsilon Z_{\mathbf{p}} - \gamma_{\mathbf{p}} + \mu)^{-1}$:

$$\Pi(\mathbf{q}, 0) = \frac{1}{2\pi^2} \iint\limits_{xy < \mu} \frac{2 dx dy}{\gamma_{\mathbf{p}} Z_{\mathbf{p} + \mathbf{q}} + \gamma_{\mathbf{p} + \mathbf{q}} Z_{\mathbf{p}}}. \qquad (6)$$

Russian Research Centre Kurchatov Institute,
pl. Akademika Kurchatova 1, Moscow, 123182 Russia

Minimum of the denominator in Eq. (10) vs. concentration for weak coupling $2\sqrt{g} = 0.2$.

Quantities $Z_{\mathbf{p}}$ depending on slowly varying logarithmic variables $\eta_{x,y}$ can be taken at the point of a minimum of the expression $\gamma_{\mathbf{p}} + \gamma_{\mathbf{p}+\mathbf{q}}$; i.e., the replacement $Z_{\mathbf{p}+\mathbf{q}} \longrightarrow Z_{\mathbf{p}} \longrightarrow Z_{\mathbf{q}/2} \longrightarrow Z_{\mathbf{q}}$ can be made. Thus,

$$\Pi(\mathbf{q}, 0) = Z_{\mathbf{p}}^{-1}\Pi_0(\mathbf{q}, 0). \qquad (7)$$

In the Van Hove scenario, the kernel of integral equations for quantity Λ and superconducting gap Δ with the d symmetry is the ladder amplitude

$$\frac{U\Pi}{1 - U\Pi} = \frac{U\Pi_0}{Z - U\Pi_0} = \frac{\eta_x\eta_y}{Z_{\mathbf{q}} - \eta_x\eta_y} = \frac{\eta_x\eta_y}{1 + \Lambda_{\mathbf{q}} - \eta_x\eta_y}. \qquad (8)$$

In the logarithmic approximation, scattering on one side of a square corresponding to the Fermi surface is described by the kernel $\dfrac{k\eta}{1 + \Lambda_\eta - k\eta}$ if the initial logarithmic momentum η is lower than the final momentum η'; otherwise, the kernel has the same form but with the replacement of η by η'. With the same logarithmic accuracy, scattering from one side of the square to another side makes the same contribution as the first type of scattering. The equations for Λ and Δ have the form

$$\Lambda_\eta = 2\sqrt{g}\int_0^\eta \frac{k\eta'd\eta'}{1 + \Lambda_{\eta'} - k\eta'} + 2\sqrt{g}\int_\eta^k \frac{k\eta d\eta'}{1 + \Lambda_\eta - k\eta}, \qquad (9)$$

$$\lambda\Delta_\eta = 2\sqrt{g}\int_0^\eta \frac{k\eta'd\eta'}{1 + \Lambda_{\eta'} - k\eta'}\frac{\Delta_{\eta'}}{1 + \Lambda_{\eta'}}$$

$$+ 2\sqrt{g}\int_\eta^k \frac{k\eta d\eta'}{1 + \Lambda_\eta - k\eta}\frac{\Delta_{\eta'}}{1 + \Lambda_{\eta'}}, \quad \frac{1}{\lambda} = \ln\frac{\mu}{T_c}. \qquad (10)$$

In order to solve the self-consistent equation for the parameter Λ_η, it is convenient to introduce the function $z(x) = 1 + \Lambda(x) - k^2x$, which is the inverse scattering amplitude in the equation for a gap. Differentiating the integral equation for Λ_η, we arrive at the first-order differential equation

$$z' = \frac{(1-x)z - (2\sqrt{g}k)^{-1}z^2}{(1-x)x + (2\sqrt{g}k^3)^{-1}z^2}. \qquad (11)$$

The numerical solution of this equation indicates that the function $z(x)$ attains its minimum at the point $x = 1$, which is a boundary of the interval $(0, 1)$. For weak interaction, quantity $r(k) = z(1, k)$ is a positive function of the doping parameter $k = \sqrt{g}\ln\dfrac{1}{\mu}$ and vanishes at the point $k = 1$ in the ladder approximation without mass corrections. This behavior corresponds to the critical point of the quantum transition to the state with a spin-density wave. Although there is a small parameter $\sqrt{g}$, this equation cannot be analyzed analytically, because the point $x = 1$ is an essential singularity of its solution $z(x)$ in the limit $r(k) \longrightarrow 0$.

The figure shows the function $r(k) = z(1, k)$ for weak interaction with $2\sqrt{g} = 0.2$. The quantity $r(k)$ characterizes closeness to the critical point. With allowance for mass corrections, the figure indicates that the critical point is shifted to the region $k > 1$, i.e., to the region of lower gap concentrations, and is not reached in the self-consistent description being considered. A state described by the self-consistent equation can be treated as a metastable state from which the system transits to the superconducting state and is therefore is stabilized. If interaction is weak and $r \ll 1$, the dominant contribution to integrals in the equation for the gap is made by the vicinity of the point $x = 1$, where $z(x, k) = r(k) + k^2\tau$, $\tau = 1 - x$. In this case, the equation for the gap takes the form

$$\frac{\lambda}{2\sqrt{g}}\Delta_\tau = \int_\tau^1 \frac{k^3(1-\tau')d\tau'}{r + k^2\tau'}\frac{\Delta_{\tau'}}{r + k^2}$$

$$+ \int_0^\tau \frac{k^3(1-t)d\tau'}{r + k^2\tau}\frac{\Delta_{\tau'}}{r + k^2}. \qquad (12)$$

Assuming that $\ln\dfrac{1}{r} \gg 1$, we formally obtain the

BCS limit. Assuming that the gap is a constant, we find

$$\lambda = \frac{2\sqrt{g}}{k} \ln \frac{1}{r(k)} \qquad (13)$$

in the logarithmic approximation.

The quantity $\ln \dfrac{1}{r(k)}$ near the minimum of the function $r(k)$ varies rapidly compared to k. For this reason, the coupling constant λ and $T_c = \mu^{-1/\lambda}$ as functions of k have maxima near the critical doping parameter k.

Thus, T_c depends nonmonotonically on carrier concentration near optimal doping in the Van Hove scenario.

ACKNOWLEDGMENTS

REFERENCES

1. Ye Jinwu and A. Millis, cond-mat/0101032.
2. R. J. Birgeneau, Phys. Rev. B **59**, 3759 (1999).
3. N. Furukawa, N. M. Rice, and M. Salmhofer, Phys. Rev. Lett. **81**, 3195 (1998).
4. S. Daul, D. J. Scalapino, and S. R. White, Phys. Rev. B **61**, 15 526 (2000).
5. V. S. Babichenko and A. N. Kozlov, Pis'ma Zh. Éksp. Teor. Fiz. **47**, 454 (1988) [JETP Lett. **47**, 536 (1988)].
6. P. Monthoux and D. Pines, Phys. Rev. B **50**, 16015 (1994).
7. A. N. Kozlov, Supercond., Phys. Chem. Technol. **2**, 64 (1988).
8. A. N. Kozlov, L. A. Maksimov, and A. F. Barabanov, Physica C (Amsterdam) **200**, 183 (1992).
9. Z. Fisk and D. Pines, Nature **394**, 22 (1998).
10. I. I. Mazin and D. J. Singh, Phys. Rev. Lett. **79**, 733 (1997).
11. P. Monthoux and G. G. Lonzarich, Phys. Rev. B **59**, 14 598 (1997).
12. A. N. Kozlov and L. A. Maksimov, Phys. Lett. A **253**, 223 (1999).
13. K. B. Blagoev, Phys. Rev. Lett. **82**, 133 (1999).

Translated by R. Tyapaev

=========================== **PHYSICS** ===========================

Raster Patterns
of Multifocus Holographic Phase Microelements

Academician A. L. Mikaelian, S. A. Prokopenko, and A. N. Palagushkin

Received February 13, 2002

We consider a method of producing two-dimensional phase structures that provide the desired intensity of electromagnetic-field distribution in the near-field diffraction zone. In a particular case, such a phase element plays the role of a lens having many foci with different intensities (Fig. 1), which is impossible to realize by traditional methods [1]. In a more general case, we deal with images formed in the focal plane. Therefore, such phase structures are artificial holograms simulated by computer-aided synthesis [2], which are called computer-generated holograms (CGHs). We show the possibility of combining multifocus phase CGHs in a continuous raster pattern, which is important not only for parallel data-processing systems, but also for forming complex images in the focal plane. When synthesizing such raster patterns, the problem of matching image fragments produced by individual CGHs and eliminating mutual interference is solved.

We perform the calculation of multifocus phase CGHs in the Fresnel approximation. In the region of Fresnel diffraction, the electromagnetic field in an observation plane is written as

$$U(x_0, y_0) = \frac{\exp(jkz)}{j\lambda z} \exp\left[j\frac{k}{2z}(x_0^2 + y_0^2)\right]$$

$$\times \int\int_{-\infty}^{\infty} \left\{ U(x_1, y_1) \exp\left[j\frac{k}{2z}(x_1^2 + y_1^2)\right]\right\}$$

$$\times \exp\left[-j\frac{2\pi}{\lambda z}(x_0 x_1 + y_0 y_1)\right] dx_1 dy_1,$$

where $U(x_0, y_0)$ is the electromagnetic-field amplitude in the observation plane, $U(x_1, y_1)$ is the amplitude in the plane directly behind a CGH, z is the distance between these planes, and $k = \frac{2\pi}{\lambda}$ is the wave number.

We now solve the inverse problem, i.e., construction of an artificial phase hologram that produces the given distribution of the field amplitude in the focal plane.

The function $U(x_0, y_0)$ can be found as the Fourier transform of the function

$$U(x_1, y_1) \exp\left[j\frac{k}{2z}(x_1^2 + y_1^2)\right].$$

It is worth noting that the expression $\exp\left[j\frac{k}{2z}(x_1^2 + y_1^2)\right]$ corresponds to the transmission function of a thin lens with the focal length $F = z$.

In the discrete representation, the phase structure of a CGH and the complex amplitude-phase distribution in the focal plane are set on a two-dimensional $M \times M$ computational net. The iteration procedure for CGH synthesis starts from the specification of a random phase distribution in the focal plane for the desired amplitude distribution. The inverse discrete Fresnel transformation of the resulting distribution yields the corresponding amplitude-phase distribution in the
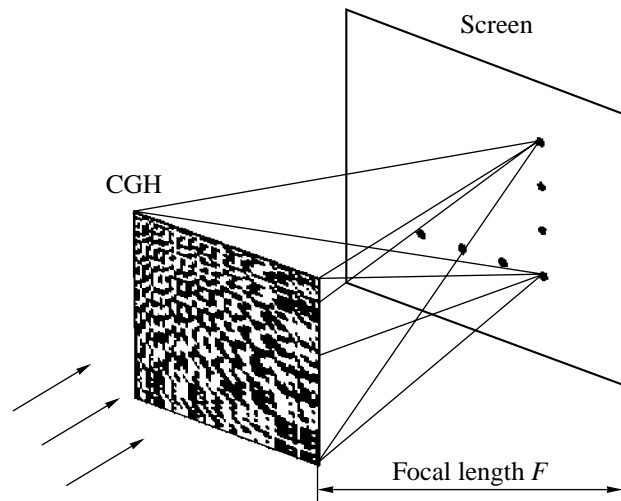
Institute of Optoneural Technologies,
Russian Academy of Sciences,
ul. Vavilova 44/2, Moscow, 117333 Russia

Fig. 1. Scheme of a multifocus CGH.

CGH plane. This distribution is then replaced by a uniform distribution, whereas the phase function is retained (or is quantized to the assigned number of levels). The direct discrete Fresnel transformation provides a new complex distribution in the focal plane, which, generally speaking, differs from the desired distribution. After replacing the amplitude distribution in the focal plane by the desired one (with the retained phase distribution), we pass to the next iteration cycle.

At the first stage of synthesis, we perform a fixed number of iterations, retaining the best intermediate result corresponding to the maximum energy efficiency for the minimum deviation from the desired final distribution. In the process of synthesis, the optimum quantization of the CGH continuous phase function is achieved by tuning phase discrimination levels. In this case, we use the optimizing algorithm of coordinate-wise descent, with the number of degrees of freedom being equal to the number of quantization-phase levels.

At the second stage of synthesis, optimization of the phase distribution obtained in the focal plane at the first stage occurs. The purpose of this operation is to find the best (continuous or quantized) CGH phase function. In this case, we use the optimizing algorithm of coordinate-wise descent for the number of degrees of freedom that is equal to the number of output diffraction orders with nonzero intensity.

We should bear in mind that, in the discrete representation of a phase structure with a $D \times D$ quadratic aperture, the discretization step $\dfrac{D}{M}$ must be smaller than the minimum size of a phase-structure element. In the simplest case of an axial focusing CGH forming one focus (i.e., in the case of a phase Fresnel lens), the local period of the phase structure is minimal at the aperture edge and equals

$$\Delta_{min} = \frac{\lambda F}{R},$$

where R is the radius of the last Fresnel zone $\left(R = \dfrac{D}{\sqrt{2}} \right)$, and F is the focal length.

Thus, for N levels of a phase quantization, the following condition must be fulfilled:

$$\frac{D}{M} < \frac{\sqrt{2}\lambda F}{DN}. \tag{1}$$

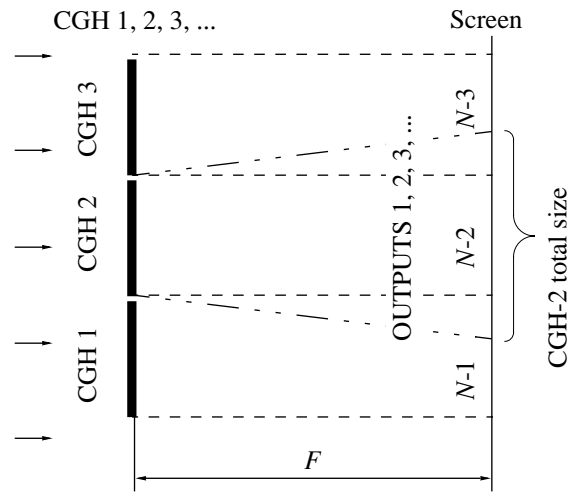Hence, we have the following constraints on the maximum possible size of the CGH quadratic aperture

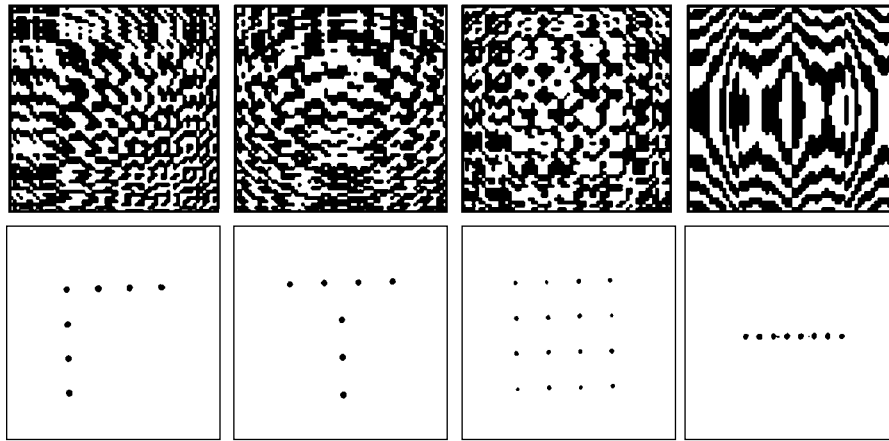**Fig. 2.** Raster pattern of focusing CGHs.

or the minimum possible focal length:

$$D < D_{max} = \sqrt{\frac{\sqrt{2}FM\lambda}{N}},$$

$$F > F_{min} = \frac{D^2 N}{\sqrt{2}M\lambda}. \tag{2}$$

It is clear that the minimum size of an element of the CGH phase structure also depends on the complexity of the desired light distribution in the focal plane. In the simplest case of two-level phase quantization ($N = 2$) (within the visible range $\lambda = 0.65$ μm), it follows from inequalities (2) that, for the focal length $F = 2.0$ mm, the constraint imposed on the aperture is $D < 0.24$ mm for $M = 64$, which corresponds to the angular aperture $\dfrac{D}{F} < 0.12$. The aperture D can be slightly increased when using a larger step of the discretization net. For example, for $M = 128$, we obtain $D < 0.34$, which corresponds to the angular aperture $\dfrac{D}{F} < 0.17$. However, we should keep in mind that, in this case, both the calculation time and the amount of topological information increase, which complicates the preparation of phase structures.

Thus, it follows from inequalities (2) that the realization of large apertures with small focal lengths is impossible. This restriction can be overcome by producing raster patterns, i.e., by combining several multi-focus structures into a unified phase matrix. This makes it possible to realize complex light distributions over a large area at relatively small focal lengths. When synthesizing a raster pattern, the given electromagnetic-field distribution in the focal plane is initially divided into several regions. In each of them, the field is formed by a corresponding phase element. As is seen from

Fig. 3. Calculated phase structures of individual multifocus CGHs and corresponding experimental light distributions of prepared samples.

Fig. 2, for the raster pattern composed of focusing CGHs located close to each other, the computational nets for light distributions of individual phase structures overlap in the focal plane, which causes mutual interference.

To avoid this interference when specifying light distribution formed by each individual CGH in the focal plane, we should use only the central part of the discretization net, which is smaller than the value of M by a factor of α. The coefficient α defines the magnification of the linear size D_F in the focal plane with respect to the aperture D of an individual phase element. Taking into account the fact that the linear size between neighboring readings in the plane of a phase element is $\dfrac{D}{M}$, and in the focal plane, it is $\dfrac{\lambda F}{D}$, we find

$$\alpha = \frac{D_F}{D} = \frac{\lambda F M}{D^2}. \tag{3}$$

In order to avoid the superposition of fields formed by individual axial focusing CGHs, it is necessary (for specifying these fields in the focal plane) to use a computational net with dimensions $C \times C$ (with area $D \times D$). With allowance for Eq. (3), we arrive at

$$C = \frac{M}{\alpha} = \frac{D^2}{\lambda F}, \tag{4}$$

and this dimension is independent of the value M of the discretization net.

Thus, in the case of synthesis of a continuous raster pattern, the general final distribution should be divided into regions of $C \times C$ readings, in each of which the distribution is formed by an individual focusing CGH.

For the above example ($M = 64$, $N = 2$, $\lambda = 0.65$ μm, and $F = 2.0$ mm) at $D = 0.192$ mm, we obtain from relationship (4) that $C = 28$. In this case, $\alpha = 2.26$. There-

fore, in order to specify a light distribution in the focal plane from the entire discretization net consisting of 64×64 readings, we need to use its central part with the dimension of 28×28 readings.

In the synthesis of individual focusing CGHs, we used a Pentium III/733 MHz/256 MB personal computer. The adjustment of the computer-synthesis methods of CGH-phase focusing was performed for a discretization net of size $M = 64$. Furthermore, in order to decrease calculation errors associated with discretization of the phase structure, the estimate of the parameters of the CGH synthesized for the discretization net $M = 64$ was performed for a larger net with $M = 1024$.

The preparation of samples of focusing CGHs for the visible spectral range requires high precision in forming the surface phase relief (errors in the depth and over the plane are ~0.01 and ~0.1 μm, respectively) [3]. Therefore, we applied an *ad hoc* method of direct electron-beam recording of the phase relief in an electron-resist layer, which was then subjected to chemical and plasmochemical treatments [4]. The method is based on dosed electron irradiation of phase regions and makes it possible (in one technological cycle) to form 2…8-level relief quantized in depth within the electron-resist layer over an area of ~6×6 mm^2 with the desired accuracy. The electron-beam setup based on a ZRM-20 (Carl Zeiss) scanning electron microscope enables us to achieve the accuracy above-mentioned for phase-relief formation.

As an illustration, we consider an example of the synthesis and preparation of the raster pattern of focusing CGHs. Figure 3 shows the calculated phase structures of individual multifocus CGHs and the corresponding light distributions formed by these multifocus elements in the focal plane. The parameters of the elements are $F = 2.0$ mm, $\lambda = 0.65$ μm, $D = 0.192$ mm, $M = 64$, and $N = 2$.
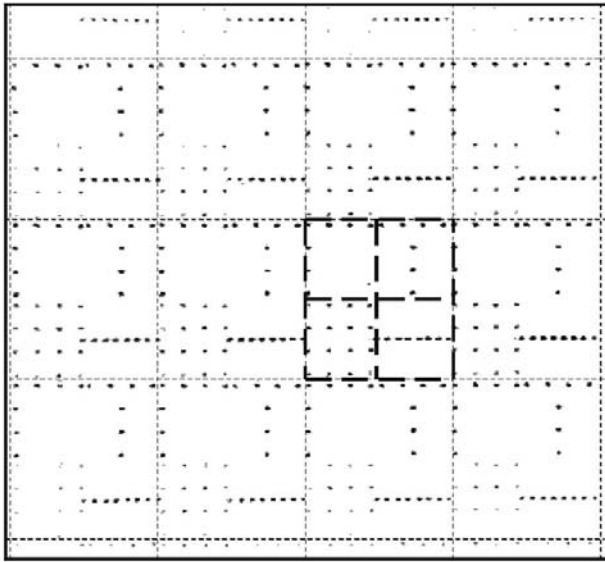
Fig. 4. Experimental output distribution formed by a raster pattern with a general aperture of 1.536×1.536 mm^2. The raster is based on 8×8 multifocus elements with the focal length of 2.0 mm.

Figure 4 shows the experimental output light distribution formed by a raster pattern with $8 \times 8 = 64$ multifocus CGHs with a total aperture of 1.536×1.536 mm^2. Note that according to formula (3), the magnification factor for a linear size of the computational net in the focal plane with respect to that of the CGH aperture is $\alpha = \dfrac{\lambda F M}{D^2} = 2.26$. In order to avoid overlap of the fields formed by the individual focusing CGHs in the focal plane, the operating fields were written into the $C \times C$ reading matrices of discretization-net nodes, where, according to relationship (4), $C = \dfrac{M}{\alpha} = 28$.

In Fig. 4, the distributions corresponding to the individual focusing CGHs are outlined.

The experimental results obtained correspond to the calculated ones with an accuracy sufficient for practical applications.

The developed technique for synthesizing multifocus CGHs and the method of direct electron-beam recording made it possible to produce a new class of elements for laser optics. Combining multifocus CGHs into raster patterns enables us to obtain complex light distributions with a large total aperture for small focal lengths, which is impossible using conventional methods.

The multifocus CGH raster patterns considered in this study can be applied in fiber-optics communication, in image processing systems, and for making optical neural networks. In particular, when using these raster patterns, the problem of splitting light from one fiber into several is solved rather simply and more efficiently (see, e.g., [5]).

ACKNOWLEDGMENTS

REFERENCES

1. A. L. Mikaelian, S. A. Prokopenko, and A. N. Palagushkin, Proc. SPIE **4296**, 189 (2001).

2. S. A. Prokopenko, A. N. Palagushkin, A. L. Mikaelian, *et al.*, Proc. SPIE **3348**, 22 (1998).

3. S. A. Prokopenko, A. N. Palagushkin, A. L. Mikaelian, *et al.*, Proc. SPIE **3348**, 83 (1998).

4. A. L. Mikaelian, S. A. Prokopenko, and A. N. Palagushkin, Opt. Mem. Neural Netw. **6** (3), 221 (1997).

5. M. R. Gordova, J. Linares, A. A. Lipovskii, *et al.*, Opt. Eng. **40** (8), 1507 (2001).

Translated by Yu. Vishnyakov

═══════════════════ **PHYSICS** ═══════════════════

Electron-Impact Excitation of Resonant Transitions in Tantalum Atoms

Yu. M. Smirnov

Presented by Academician A.I. Leont'ev February 12, 2002

Received December 21, 2001

1. Inelastic collisions of electrons with metal atoms have been studied for about 70 years. Originally, these studies were based on employing a vapor-filled cell-method, which had rather limited potentialities. Somewhat later, a significantly more efficient method of crossed beams was developed. However, the operating parameters associated with practical realizations of the method by several experimental groups allow them to study only low-temperature metals. The highest operating temperature (almost 1900 K) was attained in [1] in experiments with gadolinium. In these experiments, the authors noted that "the study of electron-impact excitation… of gadolinium atoms involves considerable experimental difficulties."

The limitations associated with the vaporization temperature of the samples being analyzed were overcome in 1969 after the method of extended crossed beams had been proposed. A universal setup based on this method was developed in 1974 [2]. The setup makes it possible to study any chemical element, e.g., boron (2400 K) [3], zirconium (2700 K) [4], and niobium (2900 K) [5].

2. In this study, the method of extended crossed beams was employed to investigate inelastic collisions of slow electrons with tantalum atoms. Both the excitation of resonant transitions and the branching accompanying it were studied.

An extended beam of tantalum atoms with a cross section of 200×26 mm^2 in the region of intersection with an electron beam was formed by surface evaporation from a graphite substrate and by subsequent shaping of the atomic beam with the use of three cooled diaphragms. Evaporation occurred as a result of the electron beam heating the tantalum surface to a temperature of 3400 K. The tantalum-atom concentration in the beam intersection region reached 2.1×10^9 cm^{-3}. The molten-zone diameter was about 9 mm; its further growth was restrained by extremely large radiation losses from the tantalum surface. Since the melt was

bounded by the autocrucible and did not come in contact with the substrate, its dissolution in the molten metal and contamination of the beam by carbides were completely excluded.

Like other refractory metals, tantalum atoms have a series of low-lying levels belonging to the fundamental $5d^36s^2a^4F$ term. Although splitting is fairly pronounced for this term, these levels can be populated as a result of thermal excitation during evaporation of tantalum. Assuming that the Boltzmann distribution is applicable, we obtained the following estimate for the level populations under our experimental conditions (atomic concentrations in the beam, percent): $5d^36s^2a^4F_{3/2}$ (0)—40.6; $5d^36s^2a^4F_{5/2}$ (2010)—26.1; $5d^36s^2a^4F_{7/2}$ (3963)—15.3; and $5d^36s^2a^4F_{9/2}$ (5621)—9.5, where the numbers in the parentheses are the level energies expressed in cm^{-1}. Since the energy differences for these levels are very large, their populations decrease fairly sharply with an increase in the excitation energy, in spite of increasing statistical weight and very high temperatures. Therefore, in this experiment, the summed atomic concentrations at the levels of the principal term, at three levels of the $5d^36s^2a^4P$ term (for $E > 6000$ cm^{-1}), and at the levels of the sextet $5d^46sa^6D$ term (for $E > 9700$ cm^{-1}) are equal to 91.5, 6.0, and 2.5%, respectively. These values should be taken into account when comparing experimental data with theoretical results. However, when solving practical problems, this distribution is unlikely to significantly affect the accuracy of the results, because, in fact, a similar low-energy distribution always occurs in an actual plasma.

The presence of very intense background radiation is one of the most important factors in experiments with refractory elements. At the above-mentioned temperature of molten metal, it is very difficult to shield the photodetector from the radiation scattered by structural elements of the electron gun. Under the conditions of this experiment, the presence of intense background light from the surface of the molten metal allowed us to carry out the measurements only in the ultraviolet region for $\lambda < 370$ nm. In the longer wave range, the noise level at the photodetector input increased intolerably. This problem could be partially solved if, instead

Moscow Power Institute (Technical University),
Krasnokazarmennaya ul. 14, Moscow, 111250 Russia

of thermal evaporation, sputtering by ion bombardment were used. However, this method has many of its own intrinsic difficulties.

The spectral resolution of our optical system reaches about 0.1 nm. Although this value is close to the best one used in beam experiments, it did not allow us to avoid blending in certain experiments with tantalum, because the spectrum of TaI has a lot of lines. The electron-beam current density did not exceed 1.0 mA/cm^2, and the width of the electron energy distribution was no greater than 1.0 eV (for 90% of electrons). Other experimental conditions, the technique applied, and the measurement procedure are thoroughly described in [6].

The measurement errors of relative values of the cross sections vary from 10 to 22%. The absolute values of the cross sections were found to an accuracy of ±25 to 37%, because the atomic concentration of tantalum in the beam was quite small.

3. We have investigated the optical emission spectrum of tantalum atoms under bombardment by a monoenergetic electron beam with an energy of 50 eV. The excitation cross sections for more than 100 spectral lines of TaI were measured within the spectral range from 230 to 370 nm. About one-third of these lines correspond to resonance lines and to transitions to higher levels, which compete with them in branching. The optical excitation functions (OEFs) were measured for a series of lines within the electron energy range from 0 to 250 eV. No lines that could reliably be assigned to the spectrum of singly charged tantalum ions were found. The same occurs in the case of niobium atoms. At the same time, as was found in [7], the excitation efficiency for singly charged vanadium ions is not much lower than that for atomic lines.

The results obtained for the resonance lines and the transitions competing with them are presented in Table 1. The wavelength λ, the transition type, the internal quantum number J, and the lower and upper energy levels E_{low} and E_{up} are presented according to the data of [8, 9]. The table also contains the position $E(Q_{max})$ of the maximum, as well as the excitation cross sections Q_{50} and Q_{max} at an electron energy of 50 eV and at the OEF maximum, respectively. The numbers in the OEF column correspond to those for the curves in the figure. The optical excitation functions for the lines with a common upper level are naturally identical. Unfortunately, only the parity and the quantum number J are known for almost all upper levels under investigation; their configurations and terms are unknown. All the energy levels lying lower than 33 500 cm^{-1} and only a few higher lying levels of TaI were identified in [9].

As is well known, methods based on optical recording of atoms excited by electron impact make it possible to measure the excitation cross section Q_{ki} for a spectral line (k and i stand for the upper and lower levels of the transition being studied, respectively). At the same time, the excitation cross section q_k for a k level is
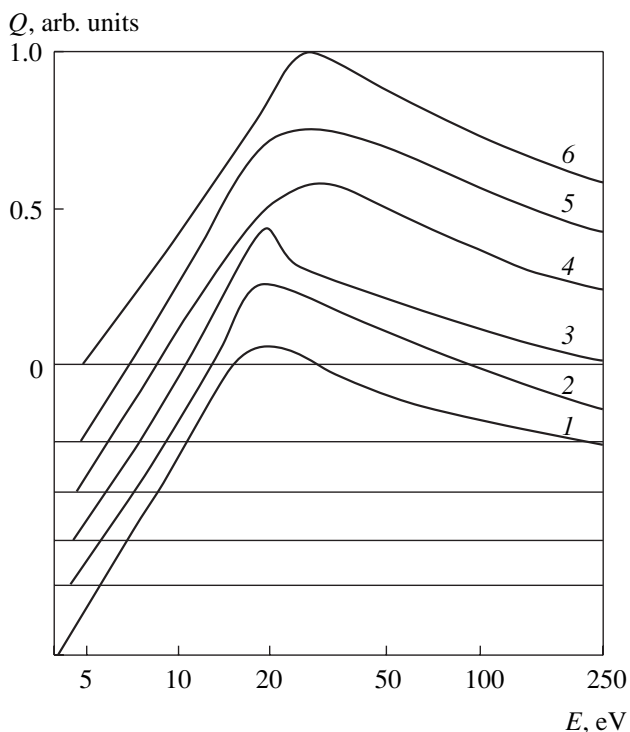
Optical excitation functions for tantalum atoms.

a basic collision characteristic for both theoretical considerations and most practical problems. These quantities are related to each other by a simple relationship that is a particular form of the balance equation under conditions of steady excitation:

$$q_k = \sum_i Q_{ki} - \sum_l Q_{lk} \equiv \Sigma Q_{50} - \Sigma Q'. \qquad (1)$$

Here, the first sum allows for branching, i.e., competition between spontaneous emitting transitions from the k level. The second sum is the contribution of cascade transitions to the k level from all higher lying l levels. The simplified notations introduced in our previous publications are used on the right-hand side of Eq. (1). It follows from Eq. (1) that the excitation cross section for an energy level can be completely determined by measuring Q_{ki} and requires no additional information.

However, the accuracy of the cross section q_k found in this manner is determined by both the measurement accuracy of Q_{ki} (and Q_{lk}) and the completeness of allowance for possible transitions associated with the k level. In actual experiments, such completeness is always limited for many reasons. The following factors are the basic ones: (i) constraints imposed on the working spectral range, (ii) insufficient sensitivity of the recording system, and (iii) the possibility of blending.

At present, electron-atomic collisions are usually studied in the ultraviolet and visible regions. Far fewer studies are carried out in the vacuum-ultraviolet and

Table 1. Excitation cross sections for spectral lines of tantalum atoms

λ, nm	Transition	J	E_{low}, cm^{-1}	E_{up}, cm^{-1}	Q_{50}, 10^{-18} cm^2	Q_{max}, 10^{-18} cm^2	$E(Q_{\text{max}})$, eV	OEF
242.764	$5d^36s^2\ a^4F-$	3/2–3/2	0	41179	46.5			
244.717	$5d^36s^2\ a^4F-$	3/2–5/2	0	40851	34.9			
254.680	$5d^36s^2\ a^4F-$	3/2–3/2	0	39253	37.1			
255.943	$5d^36s^2\ a^4F-$	3/2–5/2	0	39059	83.3	95.8	26	6
256.370	$5d^36s^2\ a^4F-$	3/2–3/2	0	38994	30.4			
257.379	$5d^36s^2\ a^4F-$	5/2–5/2	2010	40851	106			
264.747	$5d^36s^2\ a^4F-$	3/2–3/2	0	37760	654	695	26	5
265.661	$5d^36s^2\ a^4F-$	3/2–5/2	0	37630	199	219	30	4
266.807	$5d^36s^2\ a^4F-$	9/2–9/2	5621	43090	90.0			
266.862	$5d^36s^2\ a^4F-$	3/2–1/2	0	37461				
268.428	$5d^36s^2\ a^4F-$	5/2–3/2	2010	39253	94.0			
269.131	$5d^36s^2\ a^4F-$	3/2–5/2	0	37145	38.3			
269.830	$5d^36s^2\ a^4F-$	5/2–5/2	2010	39059	119	137	26	6
270.306	$5d^36s^2\ a^4F-$	5/2–3/2	2010	38994	24.5			
271.013	$5d^36s^2\ a^4F-$	7/2–5/2	3963	40851	89.5			
271.467	$5d^36s^2\ a^4F-$	3/2–5/2	0	36825	460	597	19	3
273.292	$5d^36s^2\ a^4F-$	3/2–3/2	0	36580	20.3			
277.588	$5d^36s^2\ a^4F-$	3/2–5/2	0	36014	91.2	111	19	2
279.634	$5d^36s^2\ a^4F-$	5/2–3/2	2010	37760	109	116	26	5
280.658	$5d^36s^2\ a^4F-$	5/2–5/2	2010	37630	77.7	85.4	30	4
284.852	$5d^36s^2\ a^4F-$	7/2–5/2	3963	39059	64.8	74.5	26	6
287.142	$5d^36s^2\ a^4F-$	5/2–5/2	2010	36825	85.6	112	19	3
287.336	$5d^36s^2\ a^4F-$	3/2–5/2	0	34792	67.0			
287.356	$5d^36s^2\ a^4F-$	7/2–5/2	3963	38753				
289.184	$5d^36s^2\ a^4F-$	5/2–3/2	2010	36580	86.2			
293.355	$5d^36s^2\ a^4F-$	3/2–3/2	0	34078	181	218	20	1
296.947	$5d^36s^2\ a^4F-$	7/2–5/2	3963	37630	37.2	40.9	30	4
304.956	$5d^36s^2\ a^4F-$	5/2–5/2	2010	34792	72.1			
317.359	$5d^36s^2\ a^4F-5d^26s^26p\ z^2P°$	3/2–1/2	0	31500	40.5			
318.095	$5d^36s^2\ a^4F - 5d^26s^26p\ z^2F°$	3/2–5/2	0	31428	67.6			
333.241	$5d^36s^2\ a^4P-$	5/2–3/2	9253	39253	22.6			
339.833	$5d^36s^2\ a^4F-5d^26s^26p\ z^2F°$	5/2–5/2	2010	31428	21.4			
340.694	$5d^36s^2\ a^4F - 5d^26s^26p\ y^4G°$	3/2–5/2	0	29343	35.8			
344.515	$5d^46s\ a^6D-$	3/2–3/2	9975	38994	14.6			
348.052	$5d^36s^2\ a^4P-$	3/2–5/2	6068	34792	26.2			
348.462	$5d^36s^2\ a^4F - 5d^36s6p\ x^4D°$	3/2–3/2	0	28689	10.7			
356.479	$5d^36s^2\ a^2D-$	3/2–3/2	10950	38994	21.4			
356.672	$5d^36s^2\ a^4P-$	1/2–3/2	6049	34078	23.9	28.8	20	1

infrared (especially for $\lambda > 1$ μm) regions. The spectral range of tantalum atoms is basically limited by the boundary of the vacuum ultraviolet region, whereas many spectral lines of TaI, which greatly contribute to the cascade population of the levels, are related to the red and infrared regions. Under the conditions of our experiment, these lines cannot be recorded. The limitation is a result of the intense background light mentioned above rather than insufficient sensitivity of the recording system (this background causes intense noise at the input of the photodetector and can render it inoperative).

Since no data on the cascade transitions of TaI are available, Eq. (1) cannot be used to completely determine q_k. However, the total excitation cross section ΣQ_{50} for a level, allowing for both direct electron impact excitation of the level from the initial states and the contribution of cascade filling, can be found for a series of levels. For atoms similar to tantalum, a typical contribution of cascade filling amounts 10 to 50% of the total excitation cross section.

The competing transitions from a k level can be taken into account more accurately by using available data on the branching ratios (BR). As a rule, these data are obtained with the use of gas–discharge radiation sources that provide a radiation intensity several decimal orders of magnitude larger than that in the beam crossing region. Since $\dfrac{Q_{ki}}{Q_{km}} = \dfrac{A_{ki}}{A_{km}}$ (where A_{ki} is the radiative-transition probability), when determining the total excitation cross sections, the measurement data for A_{ki} can be used to find the branching ratios.

Many branching ratios for 253 spectral lines of TaI within the range from 312 to 825 nm were found in [10]. These transitions belong to 31 upper levels within the spectral range from 23300 to 35500 cm^{-1}. A gas–discharge hollow-cathode lamp was used as a radiation source with tantalum sputtered by buffer-argon ion bombardment of the cathode. The measurements were performed for all known branching lines of the levels mentioned above, in particular, for the lines whose contribution was less than 1%. Although the long-wave region recorded in [10] was bounded by the beginning of the near-ultraviolet region, the analysis showed that no important infrared branches were omitted.

The total excitation cross sections calculated for 13 energy levels of TaI are presented in Table 2. Unfortunately, of all levels investigated in the present study, only two were considered in [10]. For other levels, the branching was taken into account according to the data of Table 1. As is seen from Table 2, the total excitation cross sections for the 8 levels being studied exceed 10^{-16} cm^2. At present, the theoretical evaluation of the excitation cross sections for these levels of tantalum atoms is impossible, because it is necessary to assign

Table 2. Total excitation cross sections for energy levels of tantalum atoms

λ, nm	E_{up}, cm^{-1}	Q_{50}, 10^{-18} cm^2	BR [10]	ΣQ_{50}, 10^{-18} cm^2
348.462	28689	10.7	0.116	92.3
340.694	29343	35.8	0.674	53.1
339.833	31428	21.4		89.0
318.095		67.6		
356.672	34078	23.9		204.9
293.355		181		
348.052	34792	26.2		98.3
304.956		72.1		
289.184	36580	86.2		106.5
273.292		20.3		
287.142	36825	85.6		545.6
271.467		460		
296.947	37630	37.2		313.9
280.658		77.7		
265.661		199		
279.634	37760	109		763
264.747		654		
356.479	38994	21.4		90.9
344.515		14.6		
270.306		24.5		
256.370		30.4		
284.852	39059	64.8		267.1
269.830		119		
255.943		83.3		
333.241	39253	22.6		153.7
268.428		94.0		
254.680		37.1		
271.013	40851	89.5		230.4
257.379		106		
244.717		34.9		

the levels to configurations and terms and to determine their wave functions.

4. Thus, electron-impact excitation of chemical elements with an evaporation temperature higher than 3000 K has been studied. The data obtained for the total cross sections can be used in astrophysics, plasma chemistry, and plasma physics. At present, it is impossible to interpret the results in detail because no data on the configurations and terms are available for the levels studied.

REFERENCES

1. L. L. Shimon, I. V. Kurta, and I. I. Garga, Opt. Spektrosk. **56** (4), 601 (1984) [Opt. Spectrosc. **56**, 368 (1984)].

2. Yu. M. Smirnov, *Physics of Electron and Atomic Collisions* (FTI Akad. Nauk SSSR, Leningrad, 1985), pp. 183–193.

3. A. N. Kuchenev and Yu. M. Smirnov, Opt. Spektrosk. **51** (2), 210 (1981) [Opt. Spectrosc. **51**, 116 (1981)].

4. A. N. Kuchenev and Yu. M. Smirnov, Zh. Prikl. Spektrosk. **39** (1), 21 (1983).

5. A. N. Kuchenev and Yu. M. Smirnov, Zh. Prikl. Spektrosk. **36** (2), 188 (1982).

6. Yu. M. Smirnov, J. Phys. II **4** (1), 23 (1994).

7. Yu. M. Smirnov, Teplofiz. Vys. Temp. **30** (6), 1049 (1992).

8. P. F. A. Klinkenberg, G. J. van den Berg, and J. C. van den Bosch, Physica (Amsterdam) **16** (11/12), 861 (1950).

9. G. J. van den Berg, P. F. A. Klinkenberg, and J. C. van den Bosch, Physica (Amsterdam) **18** (4), 221 (1952).

10. E. A. Den Hartog, D. W. Duquette, and J. E. Lawler, J. Opt. Soc. Am. B **4** (1), 48 (1987).

Translated by V. Chechin

═══════════════════════════ **PHYSICS** ═══════════════════════════

Fundamental Properties of the Coulomb Collision Operator in the Averaged Master Equation

Corresponding Member of the RAS D. P. Kostomarov and F. S. Zaĭtsev*

Received February 1, 2002

1. INTRODUCTION

One of the basic mathematical models of plasma is the set of master equations for the distribution functions of ions and electrons with the Coulomb collision operators [1]. The master equation is written in the six-dimensional phase space uniting the coordinate space and velocity space. This model in the general form is intricate, because it is multidimensional and involves effects with substantially different characteristic times. However, particular cases allow a change from Cartesian coordinates (X, Y, Z, v_1, v_2, v_3) to specially chosen phase variables separated into slow and fast variables. Then, averaging with respect to fast variables can be performed. This procedure reduces the dimension of the problem, excludes fast oscillating processes, and separates slow-evolution variations of the distribution function, which are of principal interest.

We deal with a high-temperature toroidal plasma used in experiments on controlled thermonuclear fusion in tokamaks. In these systems, the degrees of freedom can be separated into three fast and three slow variables, and one can use the symmetry in the azimuth angle. After averaging with respect to the fast motions, the master equation with three phase variables is obtained.

Various modifications of the averaging method for master equations were analyzed in many works. Their results were generalized and developed in [2]. However, the fundamental properties of the averaged Coulomb collision operator were not appropriately studied.

In the initial master equation, the Coulomb collision operator (see, e.g., [3]) (i) vanishes for Maxwell distributions of particles with the same temperature and average transport velocity, (ii) conserves the number of particles, (iii) conserves the total momentum of particles, (iv) conserves the total energy of particles, (v) does not increase the H function (analog of the Boltzmann H theorem), and (vi) is elliptic. These are the six fundamental properties of this operator.

_____

Moscow State University, Vorob'evy gory,
Moscow, 119899 Russia
* e-mail: zaitsev@cs.msu.su

The averaging of the master equation is not an identical mathematical transformation. Therefore, it is reasonable to determine the extent to which the averaged equations conserve the above properties. This problem will be discussed in this paper.

2. AVERAGED EQUATION

A detailed derivation of the averaged master equation was presented in [2]. Here, we use the results of that study as input data for analyzing the properties of the averaged Coulomb collision operator.

We consider a plasma column of toroidal configuration with noncircular section and introduce two special coordinate systems in the six-dimensional phase space. One system, the so-called local-coordinate system, unites the toroidal coordinates (γ, ξ, η) in the coordinate space and spherical coordinates (v, θ, φ) in the velocity space with the polar axis directed along the magnetic field: $(\gamma, \xi, \eta, v, \theta, \varphi) \equiv (x^1, x^2, x^3, x^4, x^5, x^6) \equiv x$. Here, γ is the label of a magnetic surface, e.g., half the surface width in the equatorial plane; and η and ξ are the toroidal and generalized poloidal angles, respectively. The other coordinate system, the so-called system with the constant of motion, involves the constants of the collisionless motion of particles (see, e.g., [4]) $(\gamma_0, v_0, \theta_0) \equiv \bar{X}$ and fast angular variables (ξ, η, φ): $(\gamma_0, \xi, \eta, v_0, \theta_0, \varphi) \equiv (\bar{x}^1, \bar{x}^2, \bar{x}^3, \bar{x}^4, \bar{x}^5, \bar{x}^6) \equiv \bar{x}$. Quantities $(\gamma_0, v_0, \theta_0)$ are any three independent constants of motion. A particular choice can be determined, e.g., by convenience of physical interpretation or numerical solution of the master equation [2, 5, 6].

It is assumed that the six-dimensional phase space can be divided into domains, each allowing the change from the local coordinates x to the coordinates $\bar{x}$.

The averaged three-dimensional nonlinear integrodifferential master equation has the form [2]

$$\frac{\partial f_\alpha^0}{\partial t} = \sum_\beta \bar{C}_{\alpha, \beta}(f_\alpha^0, f_\beta^0), \tag{1}$$

where $f_\alpha^0 \equiv f_\alpha^0(t, \gamma_0, v_0, \theta_0)$ is the desired distribution function of α particles. The summation with respect to

β is preformed over all types of plasma particles, including α particles:

$$\bar{C}_{\alpha,\beta}(f_\alpha^0, f_\beta^0) = \frac{Z_\alpha^2 Z_\beta^2 e^4 \ln \Lambda_{\alpha\beta}}{8\pi\varepsilon_0 m_\alpha}$$

$$\times \sum_{n=1,4,5} \frac{1}{\langle 1 \rangle} \frac{\partial}{\partial \bar{x}^n} \left\langle \frac{\partial \bar{x}^n}{\partial v_i} \int dv_1' dv_2' dv_3' \right. \tag{2}$$

$$\times \left. \left[\frac{f_\beta^0(t, \bar{X}')}{m_\alpha} \frac{\partial f_\alpha^0(t, \bar{X})}{\partial v_k} - \frac{f_\alpha^0(t, X)}{m_\beta} \frac{\partial f_\beta^0(t, \bar{X}')}{\partial v_k'} \right] U_{ik} \right\rangle.$$

Here, $U_{ik} = \dfrac{\delta_{ik}}{u} - \dfrac{u_i u_k}{u^3}$, where $\mathbf{u} = \mathbf{v} - \mathbf{v}'$ is the relative velocity; $u = |\mathbf{u}|$; δ_{ik} is the Kronecker delta, e is the electron charge; $\ln \Lambda_{\alpha\beta}$ is the Coulomb logarithm; e_α, e_β and m_α, m_β are the charges and masses of particles, respectively; $Z_\alpha = \dfrac{e_\alpha}{|e|}$; and $\bar{X}' \equiv \bar{X}(X, Y, Z, v_1', v_2', v_3')$. We use the International System of Units and imply summation with respect to repeated subscripts i and k from 1 to 3. Angular brackets denote the component of a certain integrable function $a(\bar{x})$ that is independent of fast variables:

$$\langle a \rangle \equiv \int_{\Delta\xi} d\xi \int_{\Delta\eta} d\eta \int_{\Delta\varphi} d\varphi\, a(\bar{x}) \sqrt{|\bar{g}|},$$

where the integrals are calculated over the domains of varying ξ, η, and φ, which are covered by the particle trajectory with specified constants of motion $\bar{X} \equiv (\gamma_0, v_0, \theta_0)$, and $\sqrt{|\bar{g}|}$ is the Jacobian of the transformation to coordinates $\bar{x}$.

Equation (1) is written for each kind of plasma particle; i.e., the set of equations is considered. This set with corresponding boundary and initial conditions allows computer calculations of various kinetic effects in a toroidal plasma. Numerical methods of solving Eq. (1) were presented and justified in [5].

When the other three angular variables vary, a fixed set of the constants of motion $(\gamma_0, v_0, \theta_0)$ specifies a certain hypersurface containing the particle trajectory. Equation (1) describes the evolution induced in the distribution function of these surfaces by Coulomb collisions. Coulomb interaction changes the constants of collisionless motion. It is interesting that the collision operator in the nonaveraged equation acts only in the velocity space, whereas in the averaged equation, this operator acts in the space of three phase variables, which depend not only on velocities, but also on coordinates.

3. PROPERTIES OF THE AVERAGED COULOMB COLLISION OPERATOR

For brevity, we present the properties without proofs.

Property 1. Averaged Coulomb collision operator (2) vanishes for Maxwell distributions $f_{M,\alpha}^0(t, \bar{X})$ and $f_{M,\beta}^0(t, \bar{X})$ of α and β particles with the same uniform (in coordinate space) temperature $T = $ const and uniform densities $n_\alpha = $ const and $n_\beta = $ const, where

$$f_{M,\alpha}^0(t, \bar{X}) = n_\alpha \left(\frac{m_\alpha}{2\pi T} \right)^{3/2} \exp\left(-\frac{m_\alpha v_0^2}{2T} \right).$$

Here, T is the temperature measured in energy units, v_0 corresponds to the total particle energy conserved in stationary magnetic and electric fields,

$$\frac{m_\alpha v_0^2}{2} = m_\alpha \frac{\mathbf{v} - \mathbf{v}_p}{2} + e_\alpha \Phi(\gamma, \xi),$$

$\Phi(\gamma, \xi)$ is the electrostatic potential, and $\mathbf{v}_p = $ const is the uniform transport velocity.

If n_α, n_β, T, or $\mathbf{v}_p$ depends on γ_0 or γ, the operator $\bar{C}_{\alpha,\beta}(f_{M,\alpha}^0, f_{M,\beta}^0)$ is generally nonzero. The radial flux generally does not vanish either. Operator $\bar{C}_{\alpha,\beta}(f_{M,\alpha}^0, f_{M,\beta}^0)$ differs in this property from the nonaveraged operator, where density, temperature, and transport velocity can depend on coordinates [3]. This situation is physically attributed to the fact that the Coulomb collision operator in the averaged equation carries information about the motion of particles and acts in both velocity space and coordinate space, whereas the operator in the nonaveraged equation acts only in velocity space.

Property 2. Averaged Coulomb collision operator (2) conserves the number of α particles in the system.

Property 3. The averaged Coulomb collision operator $\bar{C}_{\alpha,\beta}(f_\alpha^0, f_\beta^0) + \bar{C}_{\beta,\alpha}(f_\beta^0, f_\alpha^0)$ conserves the generalized total toroidal moment $P_\alpha + P_\beta$ of α and β particles if $P_\alpha = P_\alpha(\gamma_0, v_0, \theta_0)$ and $P_\beta = P_\beta(\gamma_0, v_0, \theta_0)$.

For proof, it is necessary to consider the generalized toroidal moment

$$P_\alpha = m_\alpha \sqrt{g_{33}^{\text{real}}}\, v_3 \cos\chi - e_\alpha \psi(\gamma),$$

$$P_\beta = m_\beta \sqrt{g_{33}^{\text{real}}}\, v_3 \cos\chi - e_\beta \psi(\gamma),$$

averaged with respect to gyroangle φ. Here, χ is the angle of the magnetic field with respect to the toroidal direction, g_{33}^{real} is the component of the metric tensor of the transformation from (X, Y, Z) to (γ, ξ, η), and $\psi(\gamma)$

is the poloidal flux of the magnetic field. Moments P_α and P_β are constants of the drift motion in an axisymmetric plasma [4] and are expressed in terms of $(\gamma_0, v_0, \theta_0)$ [2].

In contrast to the nonaveraged case, moments in other directions are not generally conserved, because averaging provides the transition to the description of extended objects that change their shape upon collisions.

Property 4. The averaged Coulomb collision operator $\overline{C}_{\alpha,\beta}(f_\alpha^0, f_\beta^0) + \overline{C}_{\beta,\alpha}(f_\beta^0, f_\alpha^0)$ conserves the total energy of α and β particles.

Property 5 (analog of the Boltzmann H theorem). The averaged Coulomb collision operator $\sum_\alpha \sum_\beta \overline{C}_{\alpha,\beta}(f_\alpha^0, f_\beta^0)$ does not increase the Boltzmann H function: $\frac{dH}{dt} \leq 0$. If the distribution functions of all kinds of particles are Maxwell distributions with uniform densities and the same uniform temperature, we have $\frac{dH}{dt} = 0$.

Property 6. When $f_\beta^0 > 0$, the averaged Coulomb collision operator is elliptic and elliptic-parabolic or elliptic-ultraparabolic at any point $\overline{X}$ of the space of the constants of motion if matrix $\left(\frac{\partial \overline{x}^n}{\partial v_i}\right)$, $n = 1, 4, 5$ and $i = 1, 2, 3$, is nondegenerate and degenerate, respectively.

Note that properties 1 and 4 were proven by using the specific definition of the constant of motion v_0. In property 3, it is assumed that the generalized toroidal moment is expressed in terms of $(\gamma_0, v_0, \theta_0)$. Thus, properties 1, 3, and 4 impose some specific (although quite general) requirements on the coordinate system with the constants of motion. The domain of applicability of properties 2, 5, and 6 is wider, because they require only a particular form of the averaged Coulomb collision operator.

4. CONCLUSION

The physical and mathematical properties of the Coulomb collision operator in the averaged master equation have been investigated. This operator was shown to have the basic properties of the original operator in a certain weakened form.

Properties 1–5 of the averaged Coulomb collision operator were proven only integrally over the entire coordinate space. Moreover, property 3 was proven in a truncated form. This situation is not surprising, because the Coulomb operator in the averaged equation carries information about the particle trajectory and acts not only in the velocity space, but also in the coordinate space. In addition, the averaged operator describes interaction between extended objects whose shape can change upon collisions.

The above study leads to the conclusion that the basic properties of the general averaged master equation correspond to the physical knowledge of processes occurring in a toroidal plasma and that this equation can be used for simulations. Therefore, the applicability of various particular averaged kinetic models to an adequate description of a tokamak plasma is corroborated.

REFERENCES

1. L. D. Landau, Zh. Éksp. Teor. Fiz. **7**, 203 (1937).
2. F. S. Zaitsev, M. R. O'Brien, and M. Cox, Phys. Fluids B **5** (2), 509 (1993).
3. Yu. N. Dnestrovskiĭ and D. P. Kostomarov, *Mathematical Modeling of Plasma* (Nauka, Moscow, 1993).
4. A. I. Morozov and L. S. Solov'ev, in *Reviews of Plasma Physics*, Ed. by M. A. Leontovich (Gosatomizdat, Moscow, 1963; Consultants Bureau, New York, 1966), Vol. 2.
5. F. S. Zaitsev, V. V. Longinov, M. R. O'Brien, and R. Tanner, J. Comput. Phys. **147**, 239 (1998).
6. T. P. Kiviniemi and J. A. Heikkinen, Comput. Phys. Commun., 149 (1997).

Translated by R. Tyapaev

=========================== **PHYSICS** ===========================

Use of Effective Relaxation Time as a Characteristic of Relaxation Spectra of Polymers

L. I. Kuzub and V. I. Irzhak

Presented by Academician V.B. Timofeev December 18, 2001

The spectrum of relaxation times is a convenient characteristic of the viscoelastic behavior of linear polymers and allows the prediction of the relaxation properties of polymers under almost any experimental conditions [1]. The most popular models based on molecular dynamics of the Rouse and Doi–Edwards polymer chains [2] provide this characteristic in an explicit form. Agreement between experimental data and a theoretical model can be estimated by comparing the corresponding relaxation spectra. For this reason, methods for calculating a relaxation spectrum are developed on the basis of analyzing experimental data, primarily, the dependences of the elastic modulus on time and frequency. However, in the mathematical sense, this problem belongs to a class of ill-posed problems that are extensively solved by the Tikhonov method of nonlinear regularization [3–5]. A number of software packages have been developed to calculate the relaxation spectrum under certain assumptions [4–6].

However, the question arises as to whether it is always necessary to represent the relaxation-time distribution function in an explicit form.

In recent papers [7, 8], the concept of effective relaxation time τ_{eff} was introduced and its time dependence was proposed as a qualitative characteristic of the relaxation spectrum. Similarly, one can use the frequency dependence of the mean relaxation time. In this case, there are two experimental relations: the real $G'(\omega)$ and imaginary $G''(\omega)$ parts of the complex elastic modulus. Therefore, there are two parameters with different degrees of averaging over relaxation times.

In the terminal zone ($\omega \longrightarrow 0$), the effective relaxation time defined as $\tau_{\text{eff, w}}(\omega) = \dfrac{G'(\omega)}{\omega G''(\omega)}$ takes a limiting value equal to the ratio of the second and first moments of the relaxation-time distribution, i.e., is the

Institute of Problems of Chemical Physics,
Russian Academy of Sciences,
Chernogolovka, Moscow oblast, 142432 Russia

weighted-mean relaxation time [9, 10]

$$\frac{\sum_i h_i(\tau)\tau_i^2}{\sum_i h_i(\tau)\tau_i} = \langle\tau\rangle_{\text{w}}. \tag{1}$$

At $\omega \longrightarrow 0$, another effective relaxation time defined as $\dfrac{G''(\omega)}{\omega G_{\text{N}}^0} = \tau_{\text{eff, n}}(\omega)$ is the mean relaxation time

$$\frac{\sum_i h_i(\tau)\tau_i}{\sum_i h_i(\tau)} = \langle\tau\rangle_{\text{n}}. \tag{2}$$

Thus, the ratio $\gamma = \dfrac{\langle\tau\rangle_{\text{w}}}{\langle\tau\rangle_{\text{n}}}$ characterizes the width of the relaxation-time distribution.

Figure 1 presents the frequency dependence of the mean relaxation times (curves *1*, *2*), whereas Fig. 2 shows their ratio. The calculations were performed by the corresponding formulas from [2] for the chain lengths $n = 11–501$. As is seen, these parameters are constant in the terminal zone (I) but vary with frequency in the region of the high-elasticity plateau (II), where the models give different behaviors of these parameters: $\tau_{\text{eff, w}} \approx \omega^{-1}$ and $\tau_{\text{eff, n}} \approx \omega^{-1/2}$ according to the Rouse model, and $\tau_{\text{eff, w}} \approx \omega^{-1/2}$ and $\tau_{\text{eff, n}} \approx \omega^{-3/2}$ according to the Doi–Edwards model. Consequently, the ratio of the mean times γ decreases as the square root of frequency in the first model and increases proportionally to frequency in the second model. In frequency zone III, $\tau_{\text{eff, w}}$ takes a constant value, whereas $\tau_{\text{eff, n}}$ decreases with increasing frequency due to the effect of high-frequency modes. This behavior gives rise to an increase in the distribution width.

Thus, different frequency dependences obtained for the mean relaxation times and their ratio in these two models are attributed to substantial differences in the relaxation spectra in these models.

The mean relaxation times and their ratios are very sensitive to the structure of polymer chains. This sensitivity can be illustrated by the example of modified Rouse chains, whose relaxation properties were described recently in [11].

Let us consider a chain whose structure is expressed in terms of the friction coefficients of the links as

$$100–(1–1–1–10–)_m–1–1–1–10,$$

where $m = 0, 2, 4, 6, 8$ and the numbers are the ratios of the friction coefficient of a given link to the value for a certain chosen link.

The mean times and their ratios are plotted versus frequency in Figs. 1 and 2 (curves *3*), respectively. In the terminal region, these quantities are constant and the mean times $\langle\tau\rangle_n$ and $\langle\tau\rangle_w$ are proportional to the chain length $n = 5m + 1$ to the power of 1.66 and 1.72, respectively, while their ratio is $\gamma \sim n^{0.063}$.

In the region of the high-elasticity plateau, the ratio of the mean times varies as the square of the frequency. Thus, the relaxation behavior of the modified Rouse model is close to that of the Doi–Edwards model. This result shows that the above relaxation characteristics are very sensitive to both the type of motion (Rouse and Doi–Edwards models) and the chain structure.

The analysis of experimental relaxation curves obtained in [9, 12–15] for certain linear polymers with narrow mass distribution of molecules shows that the effective relaxation times in the terminal zone are independent of frequency and, according to Eqs. (3) and (4), can be determined as the mean relaxation times $\langle\tau\rangle_w$ and $\langle\tau\rangle_n$. As an example, Figs. 1 and 2 (curves *4*) show data for polystyrene [12]. Similar results are obtained as well for poly(methyl methacrylate) [9] and 1,4-polybutadiene [13].

The average relaxation times as functions of molecular mass (M) are shown in Fig. 3. One can see that $\langle\tau\rangle_w$ and $\langle\tau\rangle_n$ are power functions, where the second time varies more slowly than the first one. The dependence $\langle\tau\rangle_n \sim M^{3,4}$ (see table) obtained for all polymers being investigated agrees with the familiar dependence of viscosity on molecular mass [1]. For both the limiting values and the high-elasticity plateau (zone II), the exponent of M depends on the method of averaging the relaxation time. For all polymers being considered, which have different molecular masses, $\tau_{\text{eff, n}}$ decreases almost linearly with increasing frequency (the exponent is close to 1.2), whereas $\tau_{\text{eff, w}} \approx \omega^{-0.7}$. Consequently, the ratio of these times increases as the square root of frequency. These dependences apparently show that the relaxation spectra of polymers with narrow mass distribution of molecules are intermediate between wide and narrow spectra characteristic of the Rouse and Doi–Edwards models, respectively.
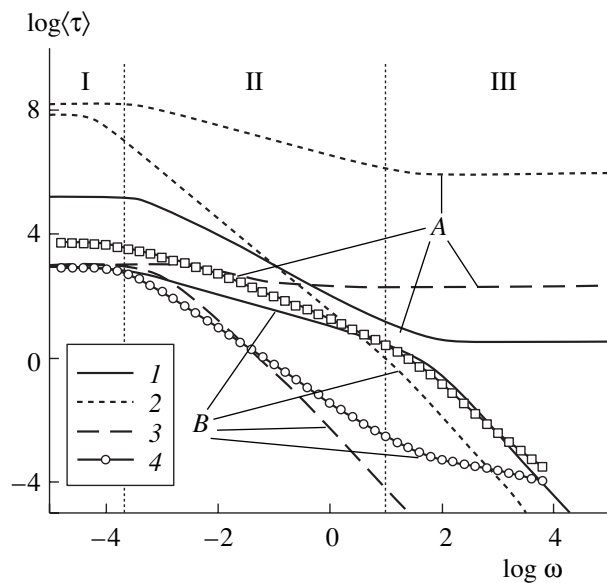
Fig. 1. Frequency dependence of the mean relaxation times for (*1*) the Rouse model at $n = 501$, (*2*) the Doi–Edwards model at $n = 501$, (*3*) the modified Rouse model at $m = 8$, and (*4*) polystyrene (experimental data from [12], molecular mass is equal to 3×10^6); *A* and *B* mark $\tau_{\text{eff, w}}$ and $\tau_{\text{eff, n}}$, respectively; I, II, and III show the terminal zone, the zone of high-elasticity plateau, and the zone of high-frequency modes (vitrification), respectively.
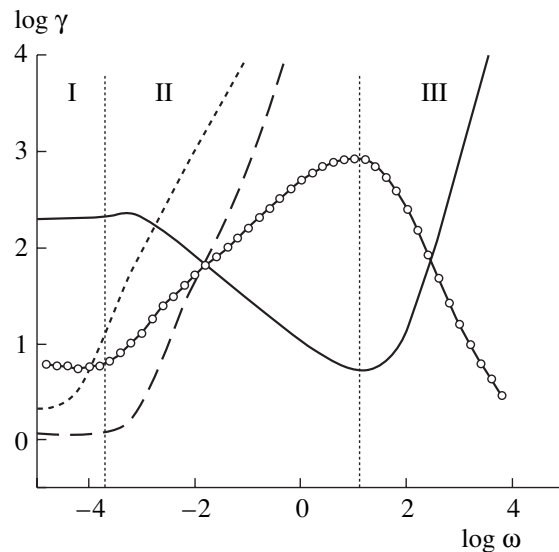
Fig. 2. The same as in Fig. 1, but for the ratio of the average relaxation times γ.

In zone III, the relaxation behavior of the polymer systems differs considerably from that of the models. The spectral width of model chains increases due to the contribution of high-frequency modes but decreases for actual polymers. This is associated with vitrification: all low-frequency modes are frozen out, so that $\tau_{\text{eff, w}}$ decreases even more rapidly than $\tau_{\text{eff, n}}$.

In contrast to the results of the Rouse and Doi–Edwards models, γ is proportional to a relatively low
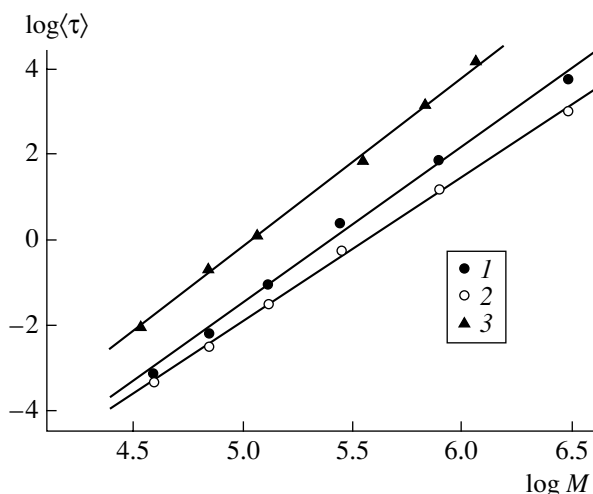
Fig. 3. The mean relaxation times as functions of molecular mass for polystyrene: data from [12] for (*1*) $\langle\tau\rangle_w$, (*2*) $\langle\tau\rangle_n$, and (*3*) τ_{max} (according to [14]).

power of M (see table). Nevertheless, its distribution widens with increasing molecular mass.

Being free of calculation errors, the method of characterizing the spectrum by the mean relaxation times can be a reliable basis for analyzing the relaxation properties of polymers. In a certain sense, the approach proposed here is similar to the method that estimates the mass distribution of polymer molecules by using the

Dependence of the relaxation times on molecular mass of a polymer for different averaging methods

Polymer	Dependence form			
	$\langle\tau\rangle_n$	$\langle\tau\rangle_w$	τ_{max}	$\gamma = \dfrac{\langle\tau\rangle_w}{\langle\tau\rangle_n}$
Rouse model	M^1	M^2	M^2	M^1
Doi–Edwards model	M^3	M^3	M^3	M^0
Modified Rouse model	$M^{1.66}$	$M^{1.72}$	$M^{1.73}$	$M^{0.063}$
Polystyrene	$M^{3.4a}$	$M^{3.7a}$	M^{4b}	$M^{0.28}$
Poly(methyl methacrylate)[c]	$M^{3.4}$	$M^{3.65}$	–	$M^{0.26}$
1,4-polybutadiene[d]	$M^{3.2}$	$M^{3.4}$	–	$M^{0.13}$

[a] [12], [b] [8, 15], [c] [9], [d] [13].

ratio of mean values $\overline{M}_w/\overline{M}_n$ instead of the distribution curve. Of course, the average molecular masses and distribution function are structural characteristics of a polymer, whereas the relaxation spectrum reflects its dynamic behavior. However, the relaxation spectrum is uniquely related to the polymer structure and therefore is physical like the mass distribution of molecules.

REFERENCES

1. G. V. Vinogradov and A. Ya. Malkin, *Rheology of Polymers* (Khimiya, Moscow, 1977).
2. M. Doi and S. Edwards, *The Theory of Polymer Dynamics* (Clarendon, Oxford, 1986; Mir, Moscow, 1998).
3. A. N. Tikhonov and V. Ya. Arsenin, *Solutions of Ill-Posed Problems* (Nauka, Moscow, 1974; Halsted Press, New York, 1977).
4. T. Roths, D. Maier, Ch. Friedrich, *et al.*, Rheol. Acta **39** (2), 163 (2000).
5. T. Roths, M. Marth, J. Weese, and J. Honerkamp, Comput. Phys. Commun. **139** (3), 279 (2001).
6. M. Baumgaertel, A. Schausberger, and H. H. Winter, Rheol. Acta **29** (5), 400 (1990).
7. V. I. Irzhak, Usp. Khim. **69** (3), 283 (2000).
8. T. F. Irzhak, L. I. Kuzub, O. V. Nikitina, and V. I. Irzhak, Vysokomol. Soedin., Ser. A **42**, 1404 (2000).
9. K. Fuchs, Ch. Friedrich, and J. Weese, Macromolecules **29**, 5893 (1996).
10. H. Watanabe, Prog. Polym. Sci. **24**, 1253 (1999).
11. V. I. Irzhak, Dokl. Akad. Nauk **379** (5), 613 (2001) [Dokl. Phys. **46**, 544 (2001)]; Dokl. Akad. Nauk **380**, 31 (2001) [Dokl. Phys. **46**, 609 (2001)].
12. A. Schausberger, G. Schindlauer, and H. Janeschitz-Kriegl, Rheol. Acta **24** (3), 220 (1985).
13. J. K. Jackson, M. E. De Rosa, and H. H. Winter, Macromolecules **27**, 2426 (1994).
14. W. J. McGrory and W. H. Tuminello, J. Rheol. **34**, 867 (1990).
15. A. V. Tobolsky, J. J. Aklonis, and G. Akovali, J. Chem. Phys. **42**, 723 (1965).

Translated by Yu. Verevochkin

==================== **PHYSICS** ====================

Equation of Nonlinear Waves in a Scattering Medium

Corresponding Member of the RAS O. V. Rudenko* and V. A. Robsman**

One of the best known equations describing the propagation of nonlinear waves in a dissipative medium is the Bürgers equation (see, e.g., [1])

$$\frac{\partial p}{\partial x} - \frac{\varepsilon}{c^3 \rho} p \frac{\partial p}{\partial \tau} = \frac{b}{2c^3 \rho} \frac{\partial^2 p}{\partial \tau^2}. \qquad (1)$$

For definiteness, we hereafter use the notation accepted in the theory of intense acoustic perturbations [1], namely, p is the pressure increment; x is the coordinate along the propagation direction for a plane wave; $\tau = t - \frac{x}{c}$ is the time in a coordinate system moving at the speed of sound c; ρ is the density of the medium; and ε and b are the coefficients of nonlinearity and of effective viscosity, respectively.

Dissipative effects are described by the term with the second derivative on the right-hand side of Eq. (1). Evolution equation (1) serves as a universal model for traveling waves of various physical natures provided that the dispersion equation for the corresponding infinitesimal perturbations takes the form [2]

$$k = \frac{\omega}{c} + i \frac{b}{2c^3 \rho} \omega^2, \qquad (2)$$

i.e., contains a small imaginary correction depending on the frequency squared. According to Eqs. (1) and (2), the damping coefficient for a weak harmonic signal is proportional to ω^2. For elastic waves, this dependence is determined by the processes of viscosity and thermal conduction in the continuum.

Equation (1) has the following solution with the steady-state (independent of the coordinate) profile:

$$p = p_0 \tanh\left(\frac{\varepsilon}{b} p_0 \tau\right). \qquad (3)$$

This solution is a solitary shock, i.e., a shock wave with a finite front thickness. Profile (3) forms as a result of competition between nonlinear steeping and dissipative

smoothing of the wave front. Solutions to Eq. (1) and the relevant nonlinear wave processes have been well studied [1].

If the additional term in dispersion equation (2) is real-valued and proportional to the frequency cubed, the right-hand side of the corresponding differential equation, i.e., the Korteweg–de Vries equation, involves the third derivative with respect to τ, in contrast to the second derivative in Eq. (1). Solutions to this equation have also been studied in detail (see, e.g., [3–5]). For the Korteweg–de Vries equation, the steady-state profile has the form of a solitary undamped pulse (soliton) resulting from the competition between nonlinear wave front steeping and dispersion spread of harmonics forming the pulse.

Signal attenuation proportional to the frequency to the fourth power was observed in numerous experiments. Such a dependence occurs in media containing small inhomogeneities, for example, in rocks [6], in spongy cranial bones [7], and in any media with small-scale parameter fluctuations for which the Rayleigh scattering law is valid [8]. Strong monopole-type scattering is observed in liquids containing gas bubbles. In this case, the coefficient of sound damping has the form [9]

$$\alpha = \beta \omega^4, \quad \beta = \frac{4\pi n a^2}{\omega_*^4}, \qquad (4)$$

where a is the bubble radius, n is the volume bubble concentration, and ω_* is the resonance frequency. In all the cases mentioned above, dispersion equation (2) should take the form

$$k = \frac{\omega}{c} + i\beta \omega^4. \qquad (5)$$

Hence, evolution equation (1) is reduced to the form

$$\frac{\partial p}{\partial x} - \frac{\varepsilon}{c^3 \rho} p \frac{\partial p}{\partial \tau} = -\beta \frac{\partial^4 p}{\partial \tau^4}. \qquad (6)$$

As far as we know, the processes described by Eq. (6), in contrast to the Bürgers equations and Korteweg–de Vries equations, have not been analyzed.

_____

* *Moscow State University, Vorob'evy gory,*
  Moscow, 119899 Russia
** *Research Institute of Transport Building,*
  Kol'skaya ul. 1, Moscow, 129329 Russia

To do this, it is convenient to rewrite Eq. (6) in dimensionless variables:

$$\frac{\partial V}{\partial z} - V\frac{V}{\partial \theta} = -R\frac{\partial^4 V}{\partial \theta^4}, \qquad (7)$$

where

$$V = \frac{p}{p_0}, \quad \theta = \omega_0 \tau, \quad z = \frac{\varepsilon}{c^3 \rho}\omega_0 p_0 x = \frac{x}{x_s},$$

$$R = \alpha x_s = \frac{\beta \omega_0^3 c^3 \rho}{\varepsilon p_0}. \qquad (8)$$

Here, x_s is the characteristic nonlinear length (discontinuity formation distance [1]), ω_0 is the characteristic frequency, and the dimensionless number R is the unique similarity criterion analogous to the inverse Reynolds–Goldberger acoustic number in Eq. (1) [1].

If we ignore the nonlinear term in Eq. (7), then the solution corresponding to the initial perturbation

$$V(z = 0, \theta) = V_0(\theta) \qquad (9)$$

takes the form

$$V(z, \theta) = \int_{-\infty}^{\infty} G(z, \theta - \theta')V_0(\theta')d\theta'$$

$$= \exp\left(-Rz\frac{\partial^4}{\partial \theta^4}\right)V_0(\theta), \qquad (10)$$

where

$$G(z, \theta) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(-R\omega^4 z - i\omega\theta)d\omega \qquad (11)$$

is Green's function. In particular, stepwise initial perturbation (9) $V_0 = \text{sgn}(\theta)$ evolves as

$$V = \frac{1}{\pi}\int_0^{\infty} \frac{\sin(\omega\theta)}{\omega}\exp(-R\omega^4 z)d\omega = V_0(z^{-1/4}\theta). \quad (12)$$

Solution (12) is self-similar: the thickness of the infinitely steep (at $z = 0$) wave front increases with the distance covered by the shock wave as $z^{1/4}$.

If we ignore the higher derivative in (7), which describes the scattering, nonlinear effects cause a narrowing of the front thickness as $(z_0 - z)^{-1}$ until the point $z = z_0$ is attained, at which a discontinuity forms. It is evident that the two contradictory tendencies might lead to the formation of a steady-state wave similar to (3). The dynamics of front formation can be described by the solution to Eq. (7), which has the form $V = \tanh\dfrac{\theta}{\theta_f}$, where $\theta_f = \theta_f(z)$ is an unknown front thickness. Substituting this expression into Eq. (7) and

allowing for only the terms linear in θ, which are responsible for the front slope in the vicinity of the front center at $\theta = 0$, we arrive at

$$\frac{d\theta_f}{dz} = -1 + 16R\theta_f^{-3}. \qquad (13)$$

Equation (13) has the solution

$$\frac{z - C}{(16R)^{1/3}} = \frac{1}{6}\ln\frac{S^2 + S + 1}{(S - 1)^2} + \frac{1}{\sqrt{3}}\arctan\frac{2S + 1}{\sqrt{3}} - S,$$

$$S = \frac{\theta_f}{(16R)^{1/3}}, \qquad (14)$$

where the constant C is determined by the given initial front thickness $\theta_f(z = 0)$. The solution describes the approach of the front thickness $\theta_f(z)$ to its steady-state value $(16\Gamma)^{1/3} = \lim \theta_f$ as $z \longrightarrow \infty$.

It is worth noting that the steady-state profile

$$V = \tanh\frac{\theta}{(16R)^{1/3}} \qquad (15)$$

is not an exact solution to Eq. (7) but satisfies a different equation

$$R\frac{d^4 V}{d\theta^4} = \left(V - \frac{3}{4}V^3\right)\frac{dV}{d\theta}. \qquad (16)$$

The latter is a steady-state variant of the generalized evolution equation

$$\frac{\partial V}{\partial z} - \frac{\partial f(V)}{\partial \theta} = -R\frac{\partial^4 V}{\partial \theta^4}, \qquad (17)$$

where the nonlinearity is defined as $f(V) = \dfrac{V^2}{2} - \dfrac{3V^4}{16}$.
The ordinary differential equation similar to (16)

$$R\frac{d^4 V}{\partial \theta^4} = V\frac{dV}{d\theta} \qquad (18)$$

corresponds to Eq. (7) and describes a steady-state wave in a medium with quadratic nonlinearity. Equation (18) has a monotonically increasing particular solution in quadratures,

$$\theta = \left(\frac{40}{9}R\right)^{1/3}\int_0^V \frac{dy}{(1 - y^2)^{2/3}}, \qquad (19)$$

which attains the value $V = 1$ in a finite time $\theta = \theta_*$, with

$$\theta_* = \left(\frac{20}{9}R\right)^{1/3}\frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{3}\right)}{2\Gamma\left(\frac{5}{6}\right)} \approx 3.4R^{1/3}, \qquad (20)$$

where Γ is the gamma function. Solution (19) is shown in Fig. 1. Curves *1–5* correspond to the numbers R equal to 2.2×10^{-4}, 6.1×10^{-3}, 2.8×10^{-2}, 7.7×10^{-2}, and 2.2×10^{-1}, respectively. The scattering process intensifies with increasing R, and the front becomes thicker.

It is worth noting that both the first and second derivatives of the function V vanish at the point $\theta = \theta_*$, while the third derivative is positive. Hence, the function V keeps growing in the region $\theta > \theta_*$. Thus, the steady-state front described by solution (19) is formed for a perturbation indefinitely increasing at infinity. The difference from the behavior of steady-state solution (3) to Bürgers equation (1) (for which the function V tends to unity as $\theta \longrightarrow \infty$) is due to the fact that the stronger frequency dependence of the energy loss for scattering (ω^4 instead of ω^2) can be compensated by a higher energy influx to the front region. The influx is provided by the function V, which increases indefinitely as $\theta \longrightarrow \infty$.

Another steady-state solution obtained by the numerical integration of Eq. (18) for $R = 0.5$ is shown in Fig. 2, where curves *1–3* correspond to the function $V(\theta)$ and its first and second derivatives, respectively. As is seen, the solution is bounded (as $\theta \longrightarrow \infty$) and its amplitude increases nonmonotonically. The function $V(\theta)$ approaches unity asymptotically, accomplishing damped oscillations with a period equal to the settling time.

It should be noted that for a scattering medium with quadratic nonlinearity, evolution equation (7) has a self-similar solution of the form

$$ V = \frac{1}{z^{3/4}} \Phi\left(\xi = \frac{\theta}{z^{1/4}} \right). \tag{21} $$

Substituting solution (21) into Eq. (7), we arrive at an ordinary fourth-order differential equation:

$$ R\Phi^{(4)} = \Phi\Phi' + \frac{1}{4}\xi\Phi' + \frac{3}{4}\Phi. \tag{22} $$

If we ignore the nonlinear term, Eq. (22) has the solution

$$ \Phi(\xi) = C\int_0^\infty \omega^2 \sin(\omega\xi)\exp(-R\omega^4)d\omega. \tag{23} $$

Function (23) behaves like the third derivative with respect to ξ of a stepwise perturbation similar to (12); i.e., it describes a solitary tripolar pulse. The form of this pulse can be found with allowance for nonlinearity only by the numerical integration of Eq. (22).

We now consider the decrease in the energy density $E = \langle V^2 \rangle$ of a wave. Here, the angle brackets stand for either the average value over a period (for signals periodic in θ) or the integral taken over the time interval for which $V \neq 0$ (for perturbations bounded in θ). Multiply-
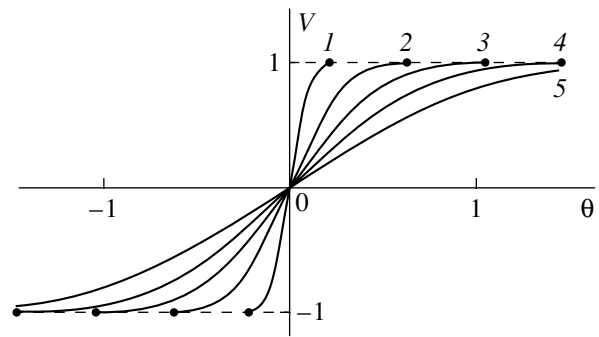

Fig. 1. Shock-front profile described by exact solution (19) to Eq. (7) for various values of the dimensionless parameter R.
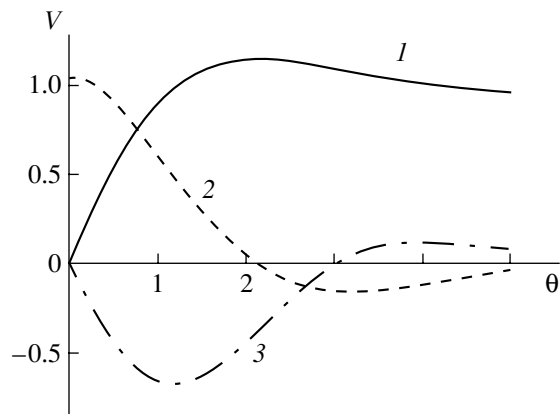

Fig. 2. (*1*) Shape of a steady-state bounded wave and its (*2*) first and (*3*) second derivatives for $R = 0.5$.

ing Eq. (7) by $2V$ and integrating the result over θ, we arrive at the relationship

$$ \frac{dE}{dz} = -2R\left\langle \left(\frac{\partial^2 V}{\partial\theta^2} \right)^2 \right\rangle. \tag{24} $$

In the linear approximation, using Eqs. (10) and (11), we obtain

$$ \frac{dE}{dz} = -4\pi R \int_{-\infty}^\infty \omega^4 |S_0(\omega)|^2 \exp(-2Rz\omega^4)d\omega, \tag{25} $$

where S_0 is the initial spectrum. It follows from Eq. (25) that the rate of a decrease in the energy E of a traveling wave is proportional to the number R [see (8)]. Since the linear solution is given by explicit expressions (10) and (11), the left-hand side of Eq. (25) can be found immediately and expression (24) may be proved to be valid identically.

We failed to find the general solution to Eq. (7) with allowance for the nonlinear term. However, we were able to evaluate the energy decrease for a steady-state wave. After integrating by parts and choosing a new

variable (V instead of θ), we reduced the average on the right-hand side of Eq. (24) to the expression

$$\left\langle \left(\frac{\partial^2 V}{\partial \theta^2}\right)^2 \right\rangle = -\int_{-1}^{1} \frac{d^3 V}{d\theta^3} dV = \frac{2}{3R}. \qquad (26)$$

When calculating the integral, we used steady-state equation (18). Substituting Eq. (26) into Eq. (24), we obtain

$$\frac{dE}{dz} = -\frac{4}{3}. \qquad (27)$$

Therefore, in the vicinity of the shock-wave front, the rate of energy-density decrease is independent of the linear energy-loss coefficient R as in the case of a steady-state wave described by the Bürgers equation.

In certain scattering media, the frequency dependence of the wave damping coefficient differs from the ω^4 power law. This situation occurs, e.g., in biological tissues [7] and in randomly inhomogeneous media [8] (in which the ω^4 dependence turns into a slower ω^2 one for sufficiently high frequencies). For example, for media with refractive-index fluctuations, dispersion equation (5) should be rewritten in the form

$$k = \frac{\omega}{c} + i \frac{\beta \omega^4}{1 + \gamma \omega^2}$$
$$\beta = \frac{8\langle \mu^2 \rangle a^3}{c^4}, \quad \gamma = \frac{4a^2}{c^2}, \qquad (28)$$

where $\langle \mu^2 \rangle$ is the mean square of refractive-index fluctuations and a is the correlation radius. The corresponding generalization of evolution equation (7) takes the form

$$\left(A^2 \frac{\partial^2}{\partial \theta^2} - 1 \right)\left(\frac{\partial V}{\partial z} - V \frac{\partial V}{\partial \theta} \right) = R \frac{\partial^4 V}{\partial \theta^4}. \qquad (29)$$

In Eq. (29), we use dimensionless notation (8) with $A^2 = \gamma \omega_0^2$. Equation (29) can be rewritten in several integrodifferential forms. One of them is

$$\frac{\partial V}{\partial z} - V \frac{\partial V}{\partial \theta} = \frac{R}{A^2} \frac{\partial^2}{\partial \theta^2} \int_0^{\infty} K(\xi) V(\theta - \xi) d\xi, \qquad (30)$$

where the kernel is given by the expression

$$K(\xi) = \delta(\xi) + \frac{1}{A} \sinh\left(\frac{\xi}{A}\right). \qquad (31)$$

It is easy to see that in the extreme cases of large and small values of the number A, the model described by Eqs. (29)–(31) is reduced to conventional Bürgers equation (1) and to Eq. (7), respectively.

The behavior of the damped nonlinear waves described by the evolution equations with dissipative terms of the fourth and higher orders (and by their combinations) is of importance for a series of applications in geophysics, in nonlinear diagnostics, and in medicine.

It is natural that numerical methods should be employed for solving particular problems. Nevertheless, since these equations are universal, it is advisable that their general properties be analyzed and their analytical solutions be found.

REFERENCES

1. O. V. Rudenko and S. I. Soluyan, *Theoretical Foundations of Nonlinear Acoustics* (Nauka, Moscow, 1975; Consultants Bureau, New York, 1977).

2. M. B. Vinogradova, O. V. Rudenko, and A. P. Sukhorukov, *Theory of Waves* (Nauka, Moscow, 1990).

3. *Solitons in Action*, Ed. by K. Lonngren and A. Scott (Academic, New York, 1978; Mir, Moscow, 1981).

4. *Nonlinear Waves*, Ed. by S. Leibovich and A. R. Seebass (Cornell Univ. Press, London, 1974; Mir, Moscow, 1977).

5. M. J. Ablowitz and H. Segur, *Solitons and the Inverse Scattering Transform* (SIAM, Philadelphia, 1981; Mir, Moscow, 1984).

6. V. S. Yamshchikov, *Wave Processes in a Massif of Rocks* (Nedra, Moscow, 1984), pp. 25–26.

7. *Physical Principles of Medical Ultrasound*, Ed. by C. R. Hill (Ellis Horwood, New York, 1986; Mir, Moscow, 1989).

8. L. A. Chernov, *Waves in Randomly Inhomogeneous Media* (Nauka, Moscow, 1975), p. 42.

9. M. A. Isakovich, *General Acoustics* (Nauka, Moscow, 1973), p. 363.

Translated by V. Chechin

═══════════════════════ **PHYSICS** ═══════════════════════

Quasicrystals with the Infinite Point Group as a Symmetry Base of Diamond-Like Structures

M. I. Samoĭlovich, A. L. Talis, and M. I. Mironov

Presented by Academician E.M. Dianov October 4, 2001

A solid-state structure described by the Fedorov group has a unified and simply connected electronic subsystem. Overlap integrals, which determine both the structure of the energy bands and the corresponding potentials, depend on the average position of nuclei and are calculated within Brillouin zones, on which the entire space of the wave vectors of single-electron functions is divided. Thus, there is no necessity for the strict requirement of translational equivalence of the structure, since the band structure is primarily determined by the polyhedral partition of the reciprocal space and by the requirements of Fermi–Dirac statistics. For small rms deviations of the atomic positions (from their positions in the corresponding crystal structures), only slight variations in the band structure and density of the tails of the localized states occur [1, 2]. The variant is ideal when the total rms deviation in the atomic positions within each generating polyhedron is close or equal to zero in a structure that is not translationally invariant. Polyhedral partition allows us to use the standard methods of algebraic geometry in the local approach and in the linear approximation within the framework of equi-affine transformations.

In 1985, a 14-atom cluster [3], which is a nonconvex diamond parallelohedron [4] and is a generating diamond cluster (GC$_D$) consisting of two seven-vertex single-cap octahedrons, was separated from the diamond structure. The graph of this cluster is the graph of incidency of the common subconfiguration of the finite projective plane $PG(2, 2) = 7_3$ and the Desargues configuration 10_3, and its group of collineations and correlations is the group $PGL_2'(7)$, which is isomorphic to $PGL_2(7)$ and has the orthogonal subgroup $\bar{3}\,m$ [4]. A single-cap octahedron is the joining of an octahedron and a tetrahedron (with a common face) from the fcc lattice generated by vectors $\left(\frac{1}{2}, \frac{1}{2}, 0\right)$, $\left(\frac{1}{2}, 0, \frac{1}{2}\right)$, and

All-Russia Research Institute for the Synthesis of Minerals,
Aleksandrov, Vladimir oblast,
Institutskaya ul. 1, 601650 Russia

$\left(0, \frac{1}{2}, \frac{1}{2}\right)$. The connection of the single-cap octahedron vertex common to the tetrahedron and octahedron with the six remaining vertices determines the subsystem Δ of six fcc lattice vectors:

$$\Delta = \{\alpha_1, \alpha_2, \alpha_1 + \alpha_2, \alpha_2 + \alpha_3, \alpha_1 + \alpha_2 + \alpha_3; \alpha_1 + 2\alpha_2 + \alpha_3\},$$

where $\alpha_1 = \left(\frac{1}{2}, \frac{1}{2}, 0\right)$, $\alpha_2 = \left(0, -\frac{1}{2}, \frac{1}{2}\right)$, $\alpha_3 = \left(-\frac{1}{2}, \frac{1}{2}, 0\right)$.

The subsystem Δ is a subsystem of the root system C_3 generating the root lattice D_3 coinciding with the fcc lattice. The first five vectors of Δ belong to the first coordination sphere of the fcc lattice (and to the root system D_3), and the sixth vector belongs to the second coordination sphere. The symplectic group $Sp(2L, K)$ above the field K [5] corresponds to the C_3 algebra (this is important when the antisymmetry of the total wave function of the electronic subsystem is taken into account).

The diamond structure D_3^+ is the self-glueing of two lattices D_3:

$$D_3^+ = D_3 \cup [1]\, D_3. \tag{1}$$

Therefore, 14 vertices of GC$_D$ determine 13 nonzero vectors:

$$GC_D = \{\Delta \cup (\alpha_1 + 2\alpha_2 + \alpha_3)\,[1]\,\Delta_-\}, \tag{2}$$

where $[1] = \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$ is the self-glueing vector of the lattice D_3 [6], $\alpha_1 + 2\alpha_2 + \alpha_3 = (0, 0, 1) \in D_3$, $\Delta_- = -\Delta = \{-\alpha_1, -\alpha_2, -(\alpha_1 + \alpha_2), -(\alpha_2 + \alpha_3), -(\alpha_1 + \alpha_2 + \alpha_3); -(\alpha_1 + 2\alpha_2 + \alpha_3)\}$.

The finite projective plane $PG(2, 2)$ is determined above the Galois field $GF(2)$. Subsequently, group $PGL_2'(7)$ is obtained from the symplectic group $Sp(2L, K)$ by transition from K field to $GF(2)$ field [5]. In this case, invariants of the integer quadratic form determining the conic section on the projective plane correspond to $GC_D(2)$, because it is a subconfiguration of $PG(2, 2)$. Group $PGL_2(7)$ is the homomorphic image of the (2, 3, 7) group, which maps a hyperbolic plane divided into hep-

tagons [7] onto the same plane. The Fedorov group of diamond is the $Fd\bar{3}m$ group. Therefore, the diamond structure divided into the 14-vertex parallelohedrons GC_D should be invariant with respect to the G_A group containing subgroups (2, 3, 7) and $R\bar{3}m$, $R\bar{3}m \subset Fd\bar{3}m$:

$$PSL_2(7) \longleftarrow (2,3,7) \subset G_A = \cup gPGL_2'(7) \supset R\bar{3}m \supset \bar{3}m, \quad (3)$$

where $G_A \supset g \notin PGL_2'(7)$. The minimum group that can contain G_A is the symplectic group $Sp(2, C)$ isomorphic to $SU(2)$ group and to a three-dimensional sphere S^3; $Sp(2, C) = Sp(1, H)$, where C and H are the fields of complex numbers and quaternions, respectively [8].

In 1986, S.P. Novikov [8, 9] suggested the most general definition of quasicrystallographic groups: subgroup G of the $E(n)$ group involving all motions of n-dimensional Euclidean space E^n is called an n-dimensional quasicrystallographic group if its intersection with subgroup $T(n) \subset E(n)$ of all translations is a certain quasilattice T, which is a finitely generated subgroup (i.e., it has a rank and is discrete) generating E^n as a linear space. In the standard terminology, T is a quasilattice of G and the orthogonal part of the factor group $G/T \longleftrightarrow R$ is the point group of the quasicrystal. If R is infinite, the group of linear parts of this n-dimensional quasicrystallographic group $G_\infty(n)$ belongs to the group of linear parts of a certain finitely generated quasicrystallographic group and is an algebraic subgroup of the general linear group $GL(n, Z)$ set above the ring Z [9]. Therefore, the construction of $G_\infty(3)$ reduces to the construction of an infinitely generated subgroup $R \subset O(n)$, $n \geq 3$ that is not quasiresolvable (i.e., it does not contain a resolvable subgroup of a final index) and maps a certain quasilattice T into itself.

By virtue of Eq. (1), quasicrystallographic diamond-like structures cannot be obtained by the standard method of the truncation of a certain multidimensional crystal [10], which allows us to obtain only "physical" quasicrystallographic groups characterized by finite point groups R and quadratic quasilattices T. This means that only groups $G_\infty(3)$ whose linear approximations represent the symmetry of the approximants of a quasicrystal with an infinite R can be the quasicrystallographic groups for diamond-like structures. The finite approximant is a 240-vertex structure on a three-dimensional sphere, i.e., diamond-like polytope {240} [11, 12]. The infinite approximants are diamond in E^3 and {6, 3, 3} honeycombs in a three-dimensional hyperbolic space, which involve a {6, 3}$_{2,1}$ map (seven hexagons on the torus) determining the graphs of 14-vertex generating clusters [4, 7, 11].

Since the form of the physical equations must be independent of the choice of coordinate system and gauge invariant, it is necessary to use the method of fiber spaces and equivariant isomorphisms of fibering,

which induce the identity mappings of the base space. The corresponding equations and their solutions for gauge fields are not considered in this paper. However, the gauge in which the linear approximation for each section is specified or the chiral field (in the general case of equations for gauge potentials) determines the finite or infinite approximants of a quasicrystal with $G_\infty(3)$. The fibering corresponding to $G_\infty(3)$ is nontrivial and does not admit a continuous global section. Therefore, gauge transformations are described by a pair of functions (dependent on the section) taking values in the holonomy groups (i.e., of a given connectedness). For any vector fibering above the compact base, the $GL(n, R)$ group reduces to the $O(n)$ subgroup and sewing functions are defined as matrix functions of dimensionality n. If the transformations are constant, the gauge is substantially local and belongs to the center of holonomy, and connectedness sections form an Abelian group. In this case, such local gauge transformations do not change any connectedness. Thus, gauge potentials are specified not as connectednesses in the principal bundle, but rather as vector functions taking values in the corresponding Lie algebra [5, 8].

The root lattice E_8 (defined finally by a four-dimensional icosahedron, the {3, 3, 5} polytope [11]) contains sublattices D_3, D_6, and H_4, which are basic for diamond-like structures [5–7]. This allows us to represent $G_\infty(3)$ in terms of the Bruhat decomposition [5]:

$$G_\infty(3) = \cup BwB, \quad w \in W, \quad (4)$$

where B is the Borel subgroup assigned by the choice of the corresponding Weyl chamber [5] that is in the E_8 lattice and, in turn, is determined by the GC vectors; and W is the Weyl group of the second-rank Tits system of an infinite dihedral group in the case under consideration [5]. On going over from the quasicrystal with $G_\infty(3)$ to diamond defined by Eqs. (3), the subgroup of the (2, 3, 7) group, which is specified by $GC_D(2)$, corresponds to group B.

Group $G_\infty(3)$ defined by the root lattice can also be considered as a reductive group [5]. Therefore, the following relations are valid:

$$G_\infty(3) \subset Sp(2L, K) \cdot V = \cup g(Sp(2L, K) \cap V). \quad (5)$$

Here, $(G, G) \supseteq PGL(2, Q) \supseteq Sp(2L, K)$, where (G, G) is the derived subgroup of $G_\infty(3)$ and the fields Q and K are different; group V is the subgroup of the centralizer of a singular torus; $g \in G_\infty(3)$; $Sp(2L, K) \cap V$ is the finite group that is obtained from $Sp(2L, K)$ by the transition from K to $GF(q)$ [5] and represents the symmetry of the initial generating cluster. Relations (5) allow us to adequately describe a diamond-like structure formed of the clusters elementarily similar [13] to the initial generating cluster.

The {240} polytope contains a 14-vertex $GC_{30/11}$ (part of a Petrie polygon [11] of the {240} polytope),
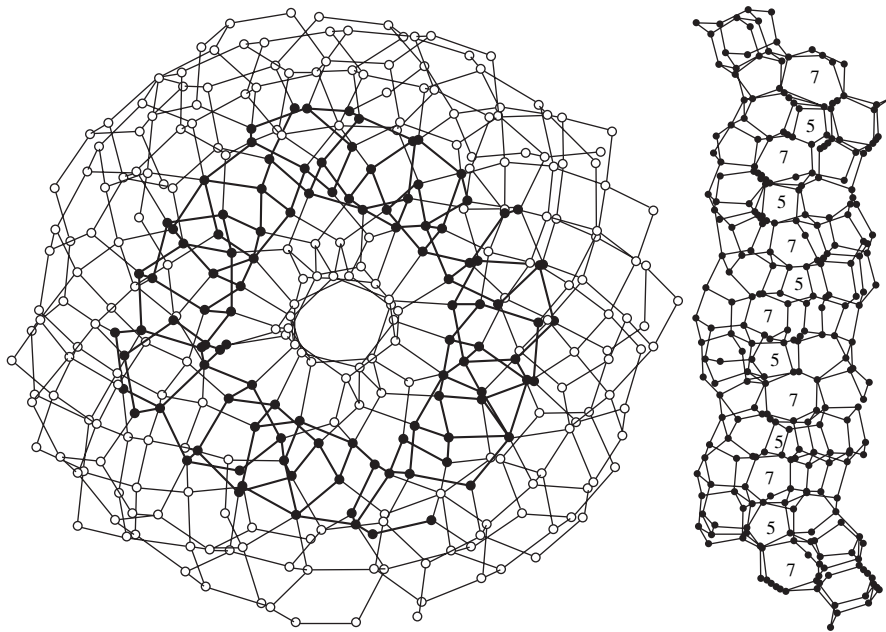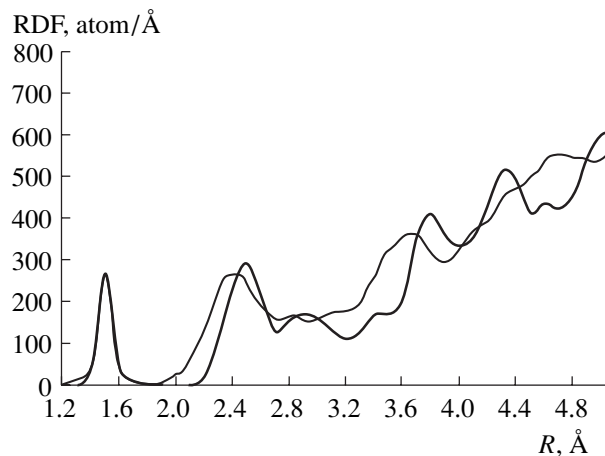
Fig. 1. Computer model of the maximum determined noncrystalline diamond-like structure (left part) across (along the twisted diamond channel) and (right part) along the rod axis. The twisted diamond channel is surrounded by (dark circles and links) four 30/11 channels, which are marked in black. The axis of the rod is formed by the alternating penta- and heptacycles marked by 5 and 7.

which is elementarily similar to GC_D and contains only hexacycles of the twisted-boat type. The whole maximum determined noncrystalline diamond-like structure can be assembled from GC_D and $GC_{30/11}$ containing only hexacycles by means of the algorithm defined finally by Eqs. (5). In this case, channels with transverse penta- and heptacycles appear between GC_D and $GC_{30/11}$ [2, 4].

The maximum determined noncrystalline diamond-like structure was simulated by computer model [2] realized in E^3 as a rodlike structure assembled from linear substructures (channels) generated by GC_D and $GC_{30/11}$ (Fig. 1). The model allowed us to calculate its structural parameters, in particular, the radial distribution function. Figure 2 shows this function along with the corresponding function for the Polk model—a continuous random network of tetrahedrally connected atoms, which involves only penta-, hexa-, and heptacycles [1]. The radial distribution function in the Polk model, which is in fact a model of amorphous silicon, has no more than four peaks. In contrast to the Polk model, the radial distribution function for the maximum determined noncrystalline diamond-like structure (Fig. 2) has eight to ten peaks. Since no more peaks are observed, this structure differs from the crystalline structure. On average, this maximum determined noncrystalline diamond-like structure has diamond bond lengths and diamond angles between them and is a maximum determined diamond analog of semiconducting amorphous silicon. The calculations show [14] that diamond-like materials of this type will not have the drawbacks inherent in modern polycrystalline dia-

mond-like films. The characteristics (mobility, breakdown voltage, etc.) of diamond classified among wide-gap semiconductors allow us to expect the development of semiconductor structures of the next generation. The maximum determined noncrystalline diamond-like structure is expected to provide a basis for a nanocomposite diamond-like substance with unique properties, for which the description of the band structure can be reduced to just the description of its linear part. For a quasicrystal with $G_\infty(3)$, the soliton-type solution will be inherent in the electron subsystem [8].

Fig. 2. Radial distribution function (RDF) for (thick line) the computer model of the maximum determined noncrystalline diamond-like structure (340 atoms) and (thin line) the Polk model (588 atoms). Smoothing of discrete functions was carried out by the standard procedure.

REFERENCES

1. A. Madan and M. Shaw, *The Physics and Applications of Amorphous Semiconductors* (Academic, Boston, 1988; Mir, Moscow, 1991).

2. M. I. Samoĭlovich, A. L. Talis, and M. I. Mironov, in *Proceedings of XII International Symposium on Diamond Films and Films of Related Materials* (Kharkov, 2001), p. 177.

3. N. A. Bul'enkov, Dokl. Akad. Nauk SSSR **284**, 1392 (1985).

4. A. L. Talis, in *Synthesis of Minerals* (Aleksandrov, 2000), Vol. 3, pp. 321–405.

5. J. E. Humphreys, *Linear Algebraic Groups* (Springer-Verlag, New York, 1975; Nauka, Moscow, 1980).

6. J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices, and Groups* (Springer-Verlag, New York, 1988; Mir, Moscow, 1990), Vols. 1, 2.

7. H. S. M. Coxeter and W. O. J. Moser, *Generations and Relations for Discrete Groups* (Springer-Verlag, New York, 1972; Nauka, Moscow, 1980).

8. B. A. Dubrovin, A. T. Fomenko, and S. P. Novikov, *Modern Geometry: Methods and Applications* (Nauka, Moscow, 1979; Springer-Verlag, New York, 1984, 1985, 1990), Parts 1–3.

9. Le Ty Kuok Thang, S. A. Piunikhin, and V. A. Sadov, Usp. Mat. Nauk **48** (1), 41 (1993).

10. Z. Olami and S. Alexander, Phys. Rev. B **39**, 1478 (1989).

11. H. S. M. Coxeter, *Regular Polytopes* (Dower, New York, 1973).

12. R. Mosseri, D. P. Di Vincenzo, T. F. Sadoc, and M. H. Brodsky, Phys. Rev. B **32**, 3974 (1985).

13. A. L. Talis, in *Proceedings of IV International Conference "Crystals: Growth, Properties, Real Structure, Applications," Aleksandrov, 1999*, p. 219.

14. M. I. Samoĭlovich, A. L. Talis, and M. I. Mironov, in *Proceedings of II All-Russia Scientific Conference on Molecular Physics of Nonequilibrium Systems, Ivanovo, 2000*, p. 248.

Translated by T. Galkina

═══════════ **PHYSICS** ═══════════

Crystallographic Analysis of Defects in Cementite and Evolution of the Lamellar-Perlite Structure in Carbon Steel under Annealing

Corresponding Member of the RAS V. M. Schastlivtsev, I. L. Yakovleva, L. E. Kar'kina, Yu. V. Khlebnikova, and T. I. Tabatchikova

The isothermal decomposition of austenite in eutectoid carbon steel is known to lead to the formation of perlite, i.e., a structure composed of alternating plates of ferrite and cementite, the crystallogeometric matching between them being specified by the orientation relationships (ORs). The evolution of the lamellar carbide phase under annealing involves processes of crushing the plates, spheroidization of the obtained fragments, and their coalescence. The processes of spheroidization and coalescence are explained well by the tendency of a system (in our case, the two-phase mixture of ferrite and cementite) to decrease the area of the interface surface of carbide particles [1, 2]. However, the kinematics of the initial stages of carbide-phase coagulation is insufficiently studied.
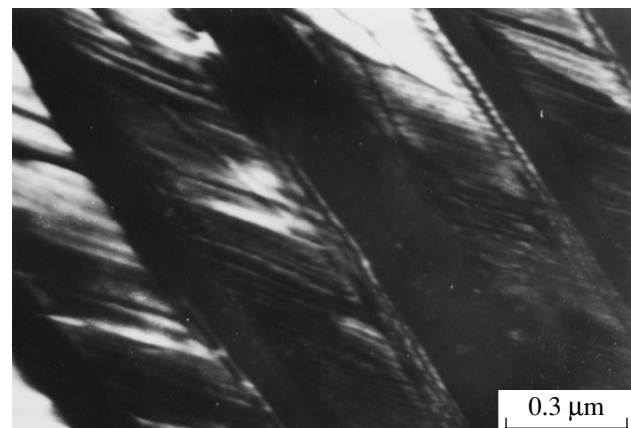
Analysis of thermodynamic features of the process of modifying the shape of cementite plates under annealing and of the mechanism of redistribution of carbon atoms shows that the character of the progress of spheroidization essentially depends on cementite-structure defects possessing additional energy [1, 2].

In this paper, we analyze the defect structure of cementite, clarify the role of defects in the crystalline structure of cementite in the processes of dissolution and coagulation of the carbon phase, and also determine the crystallographic characteristics of defect formations.

The investigations were carried out with samples of U8 carbon eutectoid steel containing 0.8% C, 0.18% Mn, 0.22% Si, 0.17% Cr, 0.12% Ni, and 0.10% Cu. The perlite structure in U8 steel was obtained by the isothermal aging of blanks previously heated to 1050°C for 5 and 30 min at 500, 600, 650, and 700°C. The fine structure was studied by electron microscopy both immediately after the isothermal perlite decomposition and after the subsequent heating of the perlite structure to 600, 650, or 700°C with holding for 5, 30, 180 min, and 30 h. The crystallographic characteristics of the planar defects in perlite cementite were determined by $g \cdot b$ analysis [3].

Electron microscopy analysis of the orientation relationships between structure components of perlite obtained at various transformation temperatures showed that both Pitch–Patch ORs and Bagryatskiĭ ORs known from the literature were observed. For both types of ORs in the lamellar perlite obtained immediately after completing the perlite transformation, defects were found in the cementite plates in the form of a system of parallel planes. The planar defects in cementite were observed for all investigated temperatures of the perlite transformation within the range 500–700°C. The dark-field image of a perlite-structure fragment obtained at 650°C for 30-min holding is shown in Fig. 1. As is seen, the cementite plate has a nonuniform contrast and contains multiple defects in

Institute of Metal Physics, Ural Division,
Russian Academy of Sciences,
ul. S. Kovalevskoĭ 18, Yekaterinburg, 620219 Russia
E-mail: schastliv@imp.uran.ru; phym@imp.uran.ru

Fig. 1. Dark-field image of the perlite microstructure for U8 carbon steel. The image is obtained at 650°C, with a 30-min exposure in the $g = 12\bar{1}_c = 110_\alpha$ reflection common to ferrite and cementite.
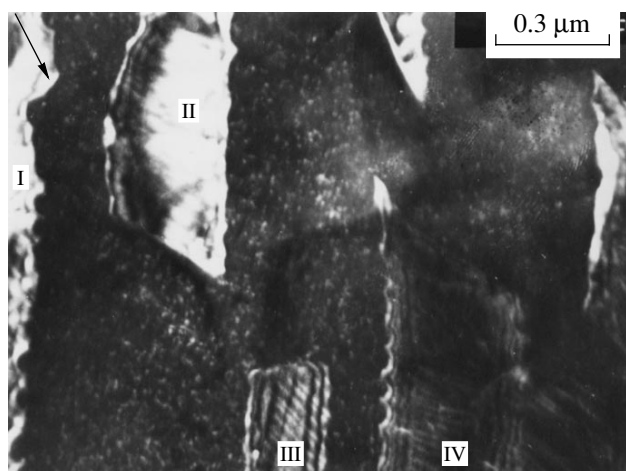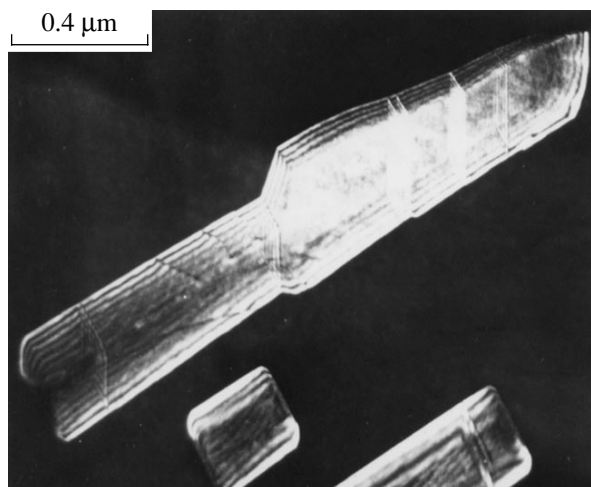
Fig. 2. Dark-field image of the perlite microstructure in the $g = 103_c = 1\bar{1}0_\alpha$ common reflection. The image is obtained at 600°C after additional annealing at 650°C for 30 h.

Fig. 3. Dark-field image of the perlite microstructure in the $g = 2\bar{1}\,\bar{1}_c$ reflection after isothermal decomposition at 700°C and additional annealing at 700°C for 30 h.

the form of thin parallel bands arranged at a certain angle to the ferrite–cementite interface. The distances between different bands are 15–25 nm.

On the basis of crystallographic analysis of diffraction patterns obtained from cementite plates containing defect planes and using the statistical treatment of the results from 20 perlite colonies, we determined the crystallographic types of the defect planes, namely, $\{101\}_c$ and $\{103\}_c$. Specific features of the arrangement of atoms in these planes enable us to assume the possibility of forming planar defects in them. On the basis of an analysis of the crystallographic relationship between phases participating in the isothermal transformation of austenite with the formation of lamellar perlite, we considered the mechanism of the appearance of planar defects in cementite as a possible way of compensating a mismatch between the crystal lattices of austenite, ferrite, and cementite [4].

Upon an additional annealing of lamellar perlite, we observed a modification in the morphology of the cementite plates. Figure 2 shows a microfragment of the perlite structure after decomposition at 600°C and an additional annealing at 650°C for 30 h. The arrow in Fig. 2 shows an intermediate stage of the formation of a neck on cementite plate I. It is seen that swelling of the cementite plate II located in front of the neck occurs during annealing. This experimentally corroborates the mechanism found in [1, 2] of redistributing carbon atoms, which leads to the dissolution of certain plates and the growth of other plates. Furthermore, acute-angled segments of cementite appear (see Fig. 2, carbide fragment IV), which cannot be explained by existing models of the dissolution of lamellar structures.

The transport of carbon atoms from defect sites of the cementite plates can lead to nucleation in the ferrite

matrix of fine globular cementite particles with a more perfect structure. The results of this study have experimentally corroborated this assumption. After additional annealing, fine-disperse precipitates of globular carbides were found in the α state in the ferrite component of perlite (see Fig. 2). The average size of these globular carbides is ~7–8 nm, which is much smaller than the width of the cementite plates.

From Fig. 2, we can also see that the planar interface of cementite plate IV becomes irregular and jagged. A connection between the line of the planar defect exit to the interface and ripples on the boundary can be traced. (The average spacing between the ripples is about 25 nm.) Thus, an accelerated outflow of carbon takes place in the sites of intersection of the planar defects with the ferrite–cementite habit plane.

The most important mechanism of modifying the cementite-plate shape under annealing is fragmentation. In this study, along with the previously considered mechanisms of neck formation, thinning, and plate separation, we experimentally found another mechanism of fragmenting cementite plates. It is associated with the formation of planar defects in cementite during annealing, which were called fracture planes in [7]. We observed sequential stages of the cementite-plate division that begins with the formation of a planar defect, continues as the formation of a thin ferrite interlayer (Fig. 3), and is finally finished by division of the plate into blocks (Fig. 4). In this case, the thickness of the separated plate fragment does not vary substantially compared to the rest of the fragment; i.e., active dissolution of the plate takes place near the fracture plane. It should be noted that in each individual perlite colony, we observed only one type of fracture plane for all cementite plates, which also coincides with the plane that limits the ends of the cementite fragments (see

Fig. 3). Furthermore, in the process of coagulation and spheroidization of carbides, the formed plate fragments bounded by the fracture planes serve as nuclei for generating globular particles (see Fig. 4). A unified spatial orientation of carbide particles was quite often observed in perlite colonies containing coagulated carbides. On the basis of crystallographic analysis of diffraction patterns obtained for cementite plates containing fracture planes, we determined the fracture-plane indices: $\{110\}_c$, $\{101\}_c$, and $\{103\}_c$. For a cementite plate divided into blocks, an increase in the total surface of cementite particles takes place that at first glance contradicts the coalescence principle. However, this seeming contradiction is eliminated if surfaces with a lower specific energy appear.

During the dissolution of cementite plates, the structure of the ferrite–cementite interface also changes. The arising system of grooves and ripples at the exits of the defect cementite planes to the habit plane creates local elastic stresses [5, 6]. On the other hand, this system becomes a source generating dislocations in the ferrite matrix. Thus, when perlite is annealed, the conditions for the generation and motion of dislocations in ferrite arise. The dislocations arising in the ferrite matrix are often decorated by a chain of fine-disperse globular carbide particles.

The investigation of the spatial distribution of dislocations showed that only a minor fraction of them are chaotically distributed in the ferrite-matrix bulk. The major fraction of the dislocations are spatially redistributed in the stress fields generated by the cementite-plate fragments and forms groups of dislocations, subboundaries, nets, or walls whose spatial disposition correlates with that of carbide fragments.

Thus, we pioneered in discovering planar defects of lamellar perlite in cementite, which are arranged in the $\{101\}_c$ and $\{103\}_c$ planes. The exits of the defect planes to the habit plane are sites of accelerated carbon outflow, whereas the terraces arising on the phase interface are sources of local elastic stresses and dislocations in the ferrite matrix. Upon additional annealing of lamellar perlite, we experimentally found that the fragmentation of cementite plates proceeds by division into blocks bounded by a single set of planes. The preferen-

Fig. 4. Dark-field image of the perlite microstructure in the $g = 1\bar{1}1_c$ reflection after isothermal decomposition at 600°C and additional annealing at 700°C for 30 h.

tial outflow of carbon takes place along these planes with the subsequent formation of ferrite bridges.

REFERENCES

1. V. M. Schastlivtsev, I. L. Yakovleva, and D. A. Mirzaev, Fiz. Met. Metalloved. **78** (3), 94 (1994).
2. V. M. Schastlivtsev, I. L. Yakovleva, and D. A. Mirzaev, Fiz. Met. Metalloved. **78** (3), 104 (1994).
3. L. M. Utevskiĭ, *Diffraction Electronic Microscopy in the Metallurgical Science* (Metallurgiya, Moscow, 1973).
4. I. L. Yakovleva, L. E. Kar'kina, Yu. V. Khlebnikova, and V. M. Schastlivtsev, Fiz. Met. Metalloved. **92** (3), 77 (2001).
5. G. J. Shiflet and J. H. van der Marwe, Metall. Mater. Trans. A **25** (9), 1885 (1994).
6. J. P. Hirth, Metall. Mater. Trans. A **25** (9), 1885 (1994).
7. I. L. Yakovleva, L. E. Kar'kina, Yu. V. Khlebnikova, *et al.*, Fiz. Met. Metalloved. **92** (6), 81 (2001).

Translated by V. Bukhanov

==================== TECHNICAL ====================
TECHNICAL
PHYSICS

On the Study of Deformation and Failure of Solids

A. M. Ivanov, E. S. Lukin, and Academician V. P. Larionov

Received January 9, 2002

The behavior of an elastoplastic material under loading (in particular, at the supercritical stage) is rather intricate. The material passes through different deformation stages from the development of different kinds of damage to gross failure. In this case, the complete description of changes in the physicomechanical properties of material is complicated by the fact that, due to rearrangement of the structure, the initially uniform stressed state (because of the development of significant plastic deformations) becomes nonuniform. Moreover, deformation is accompanied by the formation and development of defects in the form of pores and microcracks at the supercritical stage and the intensity of the processes changes. Hence, determination of the material properties at the different stages of deformation is complicated.

The occurrence of limiting states (the loss of stability of one or another process) is accompanied by changes in the mechanisms of deformation, structural modifications (attainment of the critical values), development of discontinuities, and changes in the dynamics of the processes occurring in a material.

It is of interest to investigate the development of deformation and damage accompanying deformation in an integrated approach uniting the method for constructing the true deformation curve and the methods for recording the variations of different physical parameters.

A study of the development of elastoplastic deformations and damage will provide physically justified estimates for the mechanical properties of a material.

In this paper, the deformation and failure of a strengthening elastoplastic material are studied by analyzing the true deformation curve and variation in temperature of a deformed sample.

It is known that the transition to the plastic stage of deformation is accompanied by the appearance of irreversible plastic deformations and dissipation of energy into heat [1–11]. The engineering of advanced measuring systems with quite high IR sensitivity makes it possible to analyze IR images in real time and to determine

Institute of Physicotechnical Problems of the North,
Siberian Division, Russian Academy of Sciences,
ul. Oktyabr'skaya 1, Yakutsk, 677891 Russia

the time variation in temperature of the sample under investigation.

We investigated plane samples that were made of the low-alloyed 18G2S steel and had operative parts of $65 \times 15 \times 2.4$ mm dimensions. The mechanical characteristics determined from testing by the standard technique are as follows: yield point $\sigma_t = 406.5$ MPa, breaking point $\sigma_v = 563$ MPa, residual elongation $\delta_k = 25.85\%$, and narrowing $\psi_k = 55\%$. The samples were subjected to uniaxial tension on an Instron-1195 testing machine with a rate of 8.3×10^{-5} m/s.

The kinetics of development, localization of deformation, and variation in temperature of the samples were investigated by the IR radiation method. The technique simultaneously records a deformation curve and the pattern of IR radiation and variation in temperature of a sample right up to its failure. The kinetics of the IR-field development and variation in temperature of a sample were investigated with the aid of a Thermovision 550 (Agema) IR imager with an IR sensitivity of ± 0.1 K [10]. The sample temperature in the process of deformation was also measured by using copper–constantan thermocouples and a Hewlett-Packard 3497A Data Acquisition/Control Unit with an accuracy of ± 0.05 K.

Figure 1 shows typical IR-radiation patterns that characterize the kinetics of IR-field development along the sample and correspond to the different stages of deformation. The dynamics of IR-field development represent the processes of plastic deformation and their localization. In order to determine an increase in temperature, as well as the agreement between thermal-front propagation, thermal-effect localization, and temperature variation, the deformation rate must exceed the rate of temperature variation. It is worth noting that this boundary depends on the initial structure of the sample material.

The appearance of a local zone of IR radiation and the propagation of the IR-field front along the sample corresponds to the yield plateau where the motion of Chernov–Luders bands is observed. Previously, this process was observed in holographic interferograms [12]. Toward the end of the yield plateau, the thermal front reaches the opposite end of the sample. The velocity of the IR-field front provides information on the dynamics of the motion of the deformation center in the

(a)   (b)   (c)
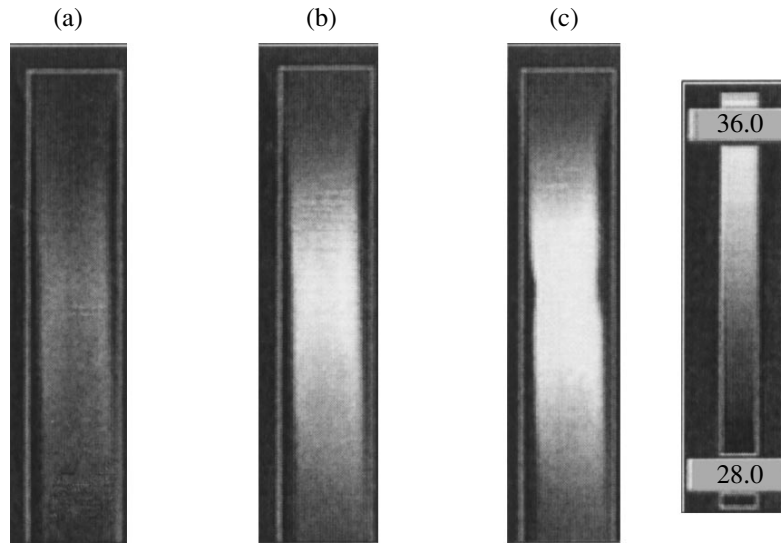
Fig. 1. Thermograms of a standard smooth sample made of 18G2S steel: (a) $\sigma = 669.1$ MPa, $\varepsilon = 14.8\%$; (b) $\sigma = 791.0$ MPa, $\varepsilon = 20.3\%$; and (c) $\sigma = 1134.8$ MPa, $\varepsilon = 23.8\%$.

process of plastic deformation. The appearance of a local zone of intense IR radiation corresponds to the onset of narrowing formation in the sample. When the sample fails, a bright flash of IR radiation is observed.

Joint analysis of the deformation curve and the curve of maximum temperature makes it possible to judge the character of the deformation process, structural variations, and the development of damage in the material (Fig. 2). Analysis of the maximum temperature as a function of strain testifies to the existence of several sections characterized by different rates of material heating. Therefore, the ΔT–ε dependence can be used to study the stages of elastoplastic deformation and the development of the defect structure in the material. The basic stages of deformation strengthening can be described in accordance with [13].

At the elastic stage of material deformation, a decrease in temperature is observed (section OA' in curve 2). The yield plateau is characterized by a monotonic rise in temperature (section $A'B'$). At the initial stage of deformation strengthening, the sample temperature is somewhat stabilized (section $B'E'$), which may be caused by a decrease in the number of movable dislocations and the onset of deformation strengthening.

A further rise in strain is accompanied by an increase in temperature. After loss of the stability of plastic deformations, which is accompanied by narrowing formation in the sample, the rise in temperature becomes substantial (section $C'D'$).

At the prefailure stage, temperature decreases. A short-term decrease in temperature in the localization zone of plastic deformations is observed at the supercritical stage of material deformation (point F' in curve 2). A decrease in temperature and the corresponding equivalent critical strain $\overline{\varepsilon}_f = 0.96$ precede

the complete failure of the sample. Here, $\overline{\varepsilon}_f = \varepsilon_f/\varepsilon_k$, where ε_f is the critical strain corresponding to the point F' and ε_k is the final strain. This phenomenon is characteristic of ductile failure of an elastoplastic material and can indicate the onset of failure, i.e., the formation of an incipient crack. At the instant of sample failure, a sharp jump in temperature occurs (not shown in Fig. 2).

It should be noted that an increase in the slope of the ΔT vs. ε dependence characterizes a rise in the deformation rate. The propagation of the IR-field front and the increase in temperature on it indicate that the deformation is more intense at the front of the Chernov–Luders bands. The presence of characteristic sections in the maximum temperature plot is associated with the evolution of a dislocation structure and damage, as well as with the existence of stages in the development of elas-
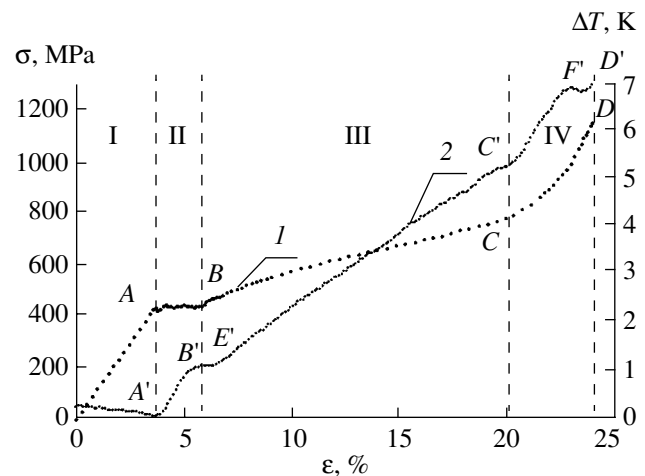
Fig. 2. (*1*) Deformation and (*2*) temperature curves for a smooth sample made of 18G2S steel.
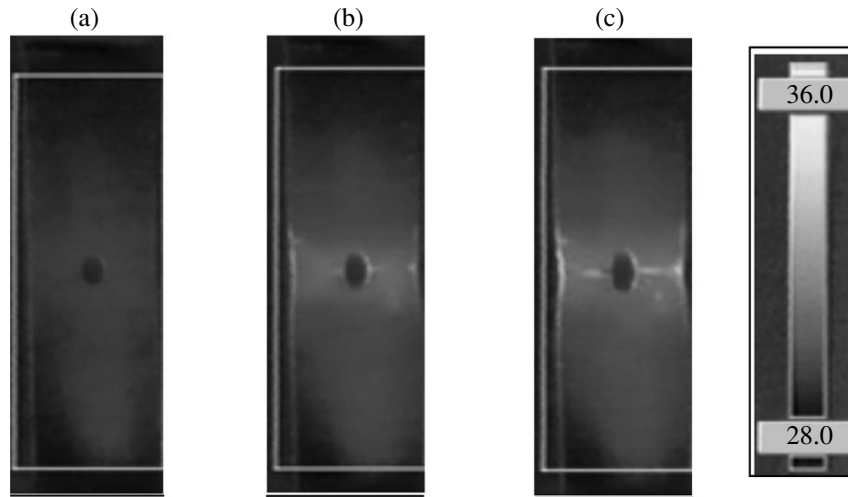
Fig. 3. Dynamics of the pattern of IR radiation for a sample made of 18G2S steel with a central circular hole: P = (a) 32.2, (b) 32.75, and (c) 20 kN.

toplastic deformations. The transition points at the boundaries between the stages of the strain dependence of temperature are related to the instability of the processes (occurrence of limiting states): attainment of the yield point, loss of the stability of plastic deformations, and the formation of an incipient crack.

We also investigated samples made of 18G2S steel annealed at a temperature of 1073 K, and strengthened under preliminary loading to a residual strain of 4.8%, as well as standard samples made of D16 aluminum alloy, poly(methyl methacrylate), and solidified PN-12 polyester resin.

For samples made of strengthened 18G2S steel and D16 alloy, the deformation curves do not have yield plateaus, but the onset of yield is unambiguously determined by the onset of an increase in the sample temperature. The temperature curve for the sample of strengthened 18G2S steel has no section II corresponding to the yield plateau. After the onset of yield, the temperature rises monotonically to the point C' and then as for the original material. Annealing leads to an increase in the length of the stable-temperature section at the initial stage of material strengthening. The character of the temperature change for a polymeric material differs significantly from that for metallic alloys. The temperature diagram consists of three sections: temperature decreases on the section of elastic deformation, is stable on the second section, and increases abruptly on the third section in a time shorter than the corresponding time for the metal alloys.

Thus, a complete description of the properties of a material (particularly when nonuniform deformation is developed) requires not only the mechanical characteristics, but also additional information about the processes accompanying deformation, damage development, and failure.

To investigate the effect of a stress concentrator on the variation in the temperature regime, we tested plane samples made of 18G2S steel that had operative part dimensions of $95 \times 30 \times 2.25$ mm and a central circular hole 6 mm in diameter. Figure 3 shows the typical IR-radiation patterns obtained by the IR imager for uniaxial tension with the same loading parameters. The presence of the geometric stress concentrator results in the localization of deformations and, accordingly, the IR field in bounded regions where zones of plastic deformations are developed. Analysis of the tensile stress–strain diagram and the maximum temperature curve shows that the onset of the increase in temperature at the point A' corresponds to the yield of the material over the minimum section, and a further significant increase in temperature at the point B' corresponds to the onset of ductile failure (Fig. 4). A technique for determining the zone of plastic deformation by measuring the size of
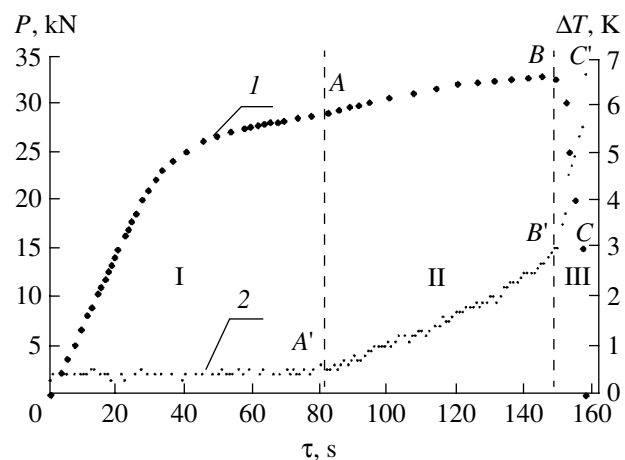
Fig. 4. (*1*) Tensile stress–strain diagram and (*2*) temperature curve for a sample made of 18G2S steel with a central circular hole.

an IR field was described in [14]. The possibility of experimentally estimating the shape and size of a plastic zone around the geometric stress concentrator on the basis of holographic interferograms was shown in [12].

Thus, the IR radiation method can be used to investigate the process of deformation and failure of materials that differ in chemical composition, structure, and manufacturing methods. It has been shown that, by analyzing the kinetics of IR-field development on the IR-radiation pattern, one can trace the process of deformation and its localization. The following features of a deformed elastoplastic material have been established: temperature stabilization on heating at the initial stage of material strengthening and a decrease in temperature at the prefailure stage. The occurrence of limiting states on the macroscopic level can be determined from the characteristic variations in the temperature of a sample, regardless of differences in the deformation mechanisms and the development of damage.

REFERENCES

1. Y. Nakada, Philos. Mag. **11**, 251 (1965).

2. J. C. Erdmann and J. A. Jahoda, Appl. Phys. Lett. **4**, 204 (1964).

3. Ya. B. Fridman, *Mechanical Properties of Metals: Deformation and Failure* (Mashinostroenie, Moscow, 1974), Part 1.

4. A. V. Nikiforov and O. V. Klyavin, Fiz. Tverd. Tela (Leningrad) **18**, 1167 (1976) [Sov. Phys. Solid State **18**, 672 (1976)].

5. G. A. Malygin, Fiz. Tverd. Tela (Leningrad) **19**, 3152 (1977) [Sov. Phys. Solid State **19**, 1845 (1977)].

6. A. A. Tupik, N. P. Valuev, and B. V. Manegin, Dokl. Akad. Nauk SSSR **272** (4), 858 (1983) [Sov. Phys. Dokl. **28**, 890 (1983)].

7. W. Dahl and P. Belche, in *Werkstoffkunde Eisen und Stahl*, Ed. by W. Dahl and W. Anton (Verlag Stahleisen, Düsseldorf, 1983; Metallurgiya, Moscow, 1986).

8. K. B. Abramova, B. P. Peregud, and I. P. Shcherbakov, Zh. Tekh. Fiz. **60** (8), 80 (1990) [Sov. Phys. Tech. Phys. **35**, 919 (1990)].

9. V. E. Remorov, Zavod. Lab., No. 5, 27 (1992).

10. A. M. Ivanov and E. S. Lukin, in *Proceedings of the XXXVI International Workshop "Topical Problems of Strength," Vitebsk, 2000*, Part 2, p. 586.

11. O. P. Maksimkin and M. N. Gusev, Pis'ma Zh. Tekh. Fiz. **27** (24), 85 (2001) [Tech. Phys. Lett. **27**, 1065 (2001)].

12. M. D. Novopashin, S. V. Suknev, and A. M. Ivanov, *Elastoplastic Deformation and Limiting State of Construction Units with Stress Concentrators* (Nauka, Novosibirsk, 1995).

13. V. I. Trefilov, *Strain Hardening and Destruction of Polycrystalline Metals* (Metallurgiya, Moscow, 1989).

14. A. M. Ivanov, Ya. S. Semenov, and E. S. Lukin, Zavod. Lab. **67** (10), 55 (2001).

Translated by Yu. Vishnyakov

Types of Discrete Symmetries of Convection in a Plane Fluid Layer

A. V. Kistovich and Yu. D. Chashechkin
Presented by Academician D.M. Klimov January 25, 2002

Received January 28, 2002

The Rayleigh–Benard convection in a plane fluid layer is actively studied both theoretically and experimentally. Diverse types of structures observed under natural and laboratory conditions [1, 2] should be classified, and the conditions of their existence should be determined. In this paper, in order to determine permissible types of convection structures in a fluid layer heated uniformly from below, the method of calculating discrete symmetries is first applied. It unites the methods of continuous-group theory, the formalism of differential forms, and the method of imbedding in higher dimension space [3, 4].

We consider the steady motion of a single fluid whose density depends only on temperature as $\rho = \rho_0(1 - T)$. The corresponding system of equations in the Boussinesq approximation has the form

$$n_u(\mathbf{u}\nabla)\mathbf{u} = -\nabla P + \nu\Delta\mathbf{u}$$

$$+ \nu'_T(2(\nabla T\nabla)\mathbf{u} + \nabla T \times (\nabla \times \mathbf{u})) - T\mathbf{g},$$

$$n_T\mathbf{u} \cdot \nabla T = \chi\Delta T + \chi'_T\nabla T \cdot \nabla T + H\delta(z), \tag{1}$$

$$n_\rho\nabla \cdot (\rho\mathbf{u}) = \nabla \cdot (D\nabla\rho).$$

Here, $\mathbf{u}$ is the velocity; P is the pressure free of the hydrostatic component; T is the dimensionless temperature obtained by the change of variables $\alpha T \longrightarrow T$, where α is the constant thermal expansion coefficient; H is the intensity of the heat source; $\delta(z)$ is the Dirac delta function; and $\mathbf{g}$ is the acceleration of gravity. In addition, the kinematic viscosity ν, the thermal diffusivity χ, and the diffusion coefficient D are assumed to be functions of medium temperature and a prime denotes differentiation with respect to temperature.

The parameters n_u, n_T, and n_ρ are equal to unity in the general case and to zero in regions where $(\mathbf{u}\nabla)\mathbf{u} \approx 0$ or $\mathbf{u} \cdot \nabla T \approx 0$. They are introduced into Eqs. (1) for more compact representation of the results and to minimize calculations of symmetries.

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

The symmetries of the structures are analyzed by the method of seeking the discrete symmetries of the differential equations [3, 4]. For this purpose, system (1) is represented in the differential form

$$(p_P^k + n_u p_{u_k}^j u_j - \nu'_T p_T^j(p_{u_k}^j + p_{u_j}^k) + Tg\delta_{kz})dV$$

$$= \nu dp_{u_k}^j \wedge (\partial_{x_j}|dV),$$

$$(n_T p_T^j u_j - \chi'_T p_T^j p_T^j)dV = \chi dp_T^j \wedge (\partial_{x_j}|dV), \tag{2}$$

$$(n_\rho p_T^j u_j - D'_T p_T^j p_T^j - p_{u_j}^j)dV = Ddp_T^j \wedge (\partial_{x_j}|dV).$$

Here, δ_{ki} is the Kronecker delta, p_i^j is the partial derivative of the ith field with respect to the jth coordinate, $dV = dx \wedge dy \wedge dz$, $\wedge$ and $|$ are the symbols of exterior and interior multiplication, and summation is implied over the repeated indices.

Considering only translations of spatial variables and following the procedure discussed in [3, 4], we introduce new variables, which are denoted by a tilde, as

$$\tilde{x} = a_x x + b_x y + x_0, \quad \tilde{y} = a_y x + b_y y + y_0, \quad \tilde{z} = a_z z,$$

$$\tilde{T} = A_T T + A_u u + A_\nu v + A_w w + A(x, y, z),$$

$$\tilde{P} = B_P P + B_T T + B_u u + B_\nu v + B_w w$$

$$+ B_x(x) + B_y(y) + B_z(z), \tag{3}$$

$$\tilde{u} = \Phi_T T + \Phi_u u + \Phi_\nu v + \Phi_w w + \Phi(x, y, z),$$

$$\tilde{v} = \Psi_T T + \Psi_u u + \Psi_\nu v + \Psi_w w + \Psi(x, y, z),$$

$$\tilde{w} = \Omega_T T + \Omega_u u + \Omega_\nu v + \Omega_w w + \Omega(x, y, z),$$

where $a_x, \ldots, A_T, A_u, \ldots$ are constants, while u, v, and w represent components of the velocity $\mathbf{u}$.

The introduction of common notation G_k^i for both coefficients and functions in Eq. (3) so that $\Psi_T \equiv G_T^\psi$ or

$\Omega(x, y, z) \equiv G_R^\omega$ leads to the system of constitutive equations

$$(a_x^2 + b_x^2)\alpha_i + (a_y^2 + b_y^2)\beta_i + a_z^2\gamma_i$$

$$= \tilde{\chi}\left(\Delta G_R^i + \frac{1}{\chi}G_T^i S_T + \frac{1}{\nu}G_{u_k}^i S_P^k\right), \qquad (4)$$

$$\alpha_i + \beta_i + \gamma_i = \tilde{S}_T, \quad a_x b_x \alpha_i + a_y b_y \beta_i = \tilde{\chi}\frac{\partial^2}{\partial x \partial y}G_R^i.$$

Here,

$$S_P^k = p_P^k + n_u p_{u_k}^j u_j - \nu'_T p_T^j(p_{u_k}^j + p_{u_j}^k) + Tg\delta_{kz},$$

$$S_T = n_T p_T^j u_j - \chi'_T p_T^j p_T^j,$$

and α_i, β_i, and γ_i are unknown functions.

If system (4) is degenerate, i.e., the principal determinant of its left-hand sides vanishes, Eqs. (3) describe regular flow symmetries and locally are the following rotations in the horizontal plane and extensions with the factor c:

$$a_x = c\cos\theta, \quad b_x = x\sin\theta, \quad a_y = -c\sin\theta,$$
$$b_y = c\cos\theta, \quad a_z = c. \qquad (5)$$

The degeneration of system (4) can be of the first kind, when $\theta \neq \frac{\pi n}{2}$, and of the second kind, when $\theta = \frac{\pi n}{2}$. Both the solution of system (4) and the form of Eqs. (3) depend only on the parameters n_T and n_u, which determine the physical flow pattern.

If convective temperature transport can be neglected ($n_T = 0$), we have

$$\tilde{T} = A_T T + T_0, \quad T_0 = \text{const}; \qquad (6)$$

where A_T is a nonzero constant, whereas transformations of the other field variables are of general form (3).

At $n_T \neq 0$, the flow pattern depends strongly on convective heat transport and

$$\tilde{T} = A_T T + T_0, \quad T_0 = \text{const}, \quad c = \pm 1, \quad \frac{\tilde{\chi}'_T}{\tilde{\chi}}A_T = \frac{\chi'_T}{\chi},$$

$$\tilde{u} = \frac{\tilde{\chi}}{c\chi}(u\cos\theta + v\sin\theta), \qquad (7)$$

$$\tilde{v} = \frac{\tilde{\chi}}{c\chi}(-u\sin\theta + v\cos\theta), \quad \tilde{w} = \frac{\tilde{\chi}}{c\chi}w,$$

whereas the pressure transformation has form (3).

If advection is weak ($n_u = 0$), we have

$$\tilde{T} = \frac{\tilde{\nu}}{\nu}T + T_0, \quad T_0 = \text{const},$$

$$\tilde{u} = u\cos\theta + v\sin\theta + \varphi_1 z + \varphi_2 x + \varphi_3 y + \varphi_0,$$

$$\tilde{v} = -u\sin\theta + v\cos\theta + \psi_1 z + \varphi_2 y - \varphi_3 x + \psi_0,$$

$$\tilde{w} = w + (\varphi_2\cos\theta + \varphi_3\sin\theta)z \qquad (8)$$
$$+ (\psi_1\sin\theta - \varphi_1\cos\theta)x - (\varphi_1\sin\theta + \psi_1\cos\theta)y + \omega_0,$$

$$\tilde{P} = \frac{\tilde{\nu}}{c\nu}P + 2\frac{\tilde{\nu}\tilde{\nu}'_T}{c^3\nu}(\varphi_2\cos\theta + \varphi_3\sin\theta)T - cT_0g,$$

where φ_1, φ_2, φ_3, and ψ_1 are arbitrary constants.

If advection is substantial ($n_u \neq 0$),

$$\tilde{T} = A_T T + T_0, \quad T_0 = \text{const}, \quad c = \pm 1,$$

$$\frac{\tilde{\nu}'_T}{\tilde{\nu}}A_T = \frac{\nu'_T}{\nu}, \quad \tilde{P} = \frac{\tilde{\nu}^2}{\nu^2}P - cT_0g,$$
$$\qquad (9)$$

$$\tilde{u} = \frac{\tilde{\nu}}{c\nu}(u\cos\theta + v\sin\theta),$$

$$\tilde{v} = \frac{\tilde{\nu}}{c\nu}(-u\sin\theta + v\cos\theta), \quad \tilde{w} = \frac{\tilde{\nu}}{c\nu}w,$$

and the condition $\tilde{\nu}\tilde{\nu}'_T = c\nu\nu'_T$ must be satisfied.

Transformations (6)–(9) involve geometric symmetries and specific symmetries corresponding to the temperature dependence of the kinetic coefficients of the medium.

To analyze the geometric symmetries with allowance for the effect of a heat source, it is sufficient to consider the most general transformations (9), which, together with equations relating the tangent spaces of the original and new variables, generate transformations of the form

$$\tilde{p}_i^x\cos\theta - \tilde{p}_i^y\sin\theta = \frac{\tilde{\nu}^2}{c\nu^2}p_i^x,$$

$$\tilde{p}_i^x\sin\theta + \tilde{p}_i^y\cos\theta = \frac{\tilde{\nu}^2}{c\nu^2}p_i^y, \qquad (10)$$

$$\tilde{p}_i^z = \frac{\tilde{\nu}^2}{c\nu^2}p_i^z, \quad c = \pm 1.$$

To separate geometric symmetries, we consider the kinetic coefficients as constants ($\tilde{\nu} = \nu$, $\tilde{\chi} = \chi$). The cyclicity condition ($\tilde{f}^{(n)} = f$) for the n-fold application of Eqs. (9) and (10) to an input quantity f determines the
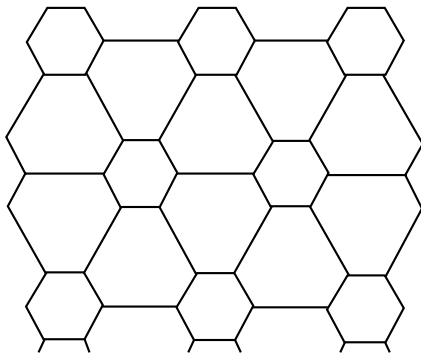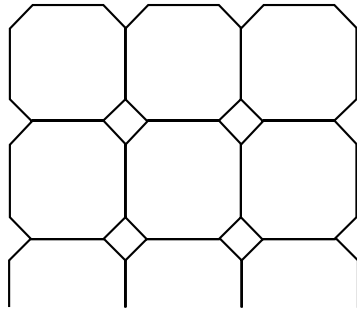
Fig. 1.

Fig. 2.

quantum of the rotation angle as

$$\theta = \frac{\pi}{2n}(3 - (-1)^n) \text{ at } c = 1,$$

$$\theta = \frac{2\pi}{n} \text{ at } c = -1. \tag{11}$$

In the first case ($c = 1$) $\tilde{T} = T$; i.e., the temperature field does not vary. At $n = 1$, (11) the transformation is identical because $\theta = 2\pi$. At $n > 1$, the velocity field is symmetric about the rotation group $SO_{2n}(2)$ or $SO_n(2)$ when n is 2 to an integer power or otherwise, respectively.

Although the method of seeking discrete symmetries deals with locally defined quantities (derivatives and differentials), the elements of the above groups act globally, i.e., on the entire fluid volume. All convective cells belong to the same symmetry group only at $n = 2$, 3, 4, 6. This restricts the possible types of regular structures, which can consist of one regular or several possibly irregular elements.

One-element structures are formed by squares ($n = 2$, symmetry degeneration of the second kind), triangles ($n = 3$, first-kind degeneration), and hexagons ($n = 6$, first-kind degeneration).

At $n = 6$, the group $SO_6(2)$ has the subgroup $SO_3(2)$, and, as is shown in Fig. 1, the multielement structures are formed of regular [symmetry $SO_6(2)$] and irregular [symmetry $SO_3(2)$] hexagons. In this case, the global flow pattern is symmetric under rotations about the centers of reg-

ular [by the angle $\frac{\pi}{3}$, which is a discrete element of the group $SO_6(2)$] and irregular [by the angle $\frac{2\pi}{3}$, which is a discrete element of the group $SO_3(2)$] hexagons. Symmetry degeneration of all elements is of the first kind.

At $n = 4$, the $SO_8(2)$ group has the $SO_4(2)$ subgroup and, as is shown in Fig. 2, multielement structures consist of squares and octagons (regular or irregular). If these octagons are regular, degeneration is of the first and second kinds for octagonal and square cells, respectively. The flow is symmetric under rotations about the centers of the regular octagons by the angle $\frac{\pi}{4}$ [group $SO_8(2)$] and about the centers of squares by the angle $\frac{\pi}{2}$ [group $SO_4(2)$]. If the octagons are irregular [subgroup $SO_4(2)$], degeneration is of the second kind for all cells. All structures described here were observed in the convection patterns [2].

The symmetry $SO_2(2)$ with the discrete rotation angle π is absent in media with constant kinetic coefficients ($c = 1$). For so-called roll convection, when cells are horizontal cylinders, this means that vorticity cannot reverse sign in neighboring elements.

When kinetic coefficients depend on temperature ($c = -1$), flow circulation in the cells changes sign when $\tilde{\nu}\tilde{\nu}'_T = -\nu\nu'_T$, i.e., when the derivative of the kinetic coefficient changes sign. This phenomenon was observed in [5]. According to Eqs. (7), a similar effect occurs at $\tilde{\chi}\tilde{\chi}'_T = -\chi\chi'_T$ for slightly varying kinematic viscosity and strong temperature dependence of thermal diffusivity. In this case, as for constant kinetic coefficients, the above set of geometric symmetries is supplemented only by the symmetry group $SO_2(2)$. In the case of roll convection, this symmetry makes alternation of vorticity sign in the cells possible, and this effect is observed in media with strong temperature dependence of kinematic viscosity [2]. In this case, convection rolls are very unstable in the longitudinal direction and tend to turn into square cells.

REFERENCES

1. D. J. Tritton, *Physical Fluid Dynamics* (Oxford Univ. Press, Oxford, 1988).
2. D. B. White, J. Fluid Mech. **191**, 247 (1988).
3. A. V. Kistovich and Yu. D. Chashechkin, Dokl. Akad. Nauk **380** (6), 757 (2001) [Dokl. Phys. **46**, 718 (2001)].
4. A. V. Kistovich and Yu. D. Chashechkin, Regul. Chaot. Dyn. **6** (3), 327 (2001).
5. E. Palm, T. Ellingsen, and B. Gjevik, J. Fluid Mech. **30**, 651 (1967).

Translated by Yu. Verevochkin

THEORETICAL PHYSICS

On the Consistency of Experiments Supporting General Relativity

Academician L. B. Okun' and K. G. Selivanov

Here we present our reaction to the paper by V.V. Okorokov "On the Discrepancy of Experiments Supporting Certain Conclusions of General Relativity," which was recently published in *Doklady Physics* [1].

As is stated in [1], experimental data related to the so-called red shift of photons contradict experiments in which the acceleration of the atomic-clock rate with increasing distance from the Earth was verified. This is an erroneous statement by Okorokov that initiated a thorough analysis of the problem in question in 1998 [2–5]. It was clearly demonstrated that the results of both types of experiments are in excellent agreement with each other.

The seeming contradiction is of a purely terminological nature. It is caused by the fact that the interpretation of experiments on the photon red shift in a number of papers is insufficiently consistent. In actual conditions, the change in the photon frequency is of a relative, rather than an absolute character. In the Earth's static gravitational field, a photon emitted by a source located in the basement of a building cannot cause the reverse nuclear transition in a detector placed in the building's attic. However, this is not due to a decrease in the photon energy but rather to the fact that the distance between nuclear levels in the attic is larger than in the basement. The same argument is also true for atomic layers in an atomic clock. Thus, the photon frequency decreases only with respect to the nuclear transition frequency or with respect to the atomic-clock frequency (cf. [6, Section 88]).

In the scientific and popular-science literature, this argument is insufficiently clarified. This explains Okorokov's error and the necessity of publishing papers [2–5]. It should be emphasized that absolute conservation of the photon frequency and absolute acceleration of the clock rate take place only in a static gravitational field. In the picture described above, the same time is used along the entire photon path, e.g., the time counted off by the clock at infinity. In [6], this time is called universal.

In the general case of a nonstatic field, there is no global time, and global clock synchronization is impossible (see, e.g., [6, Section 85]). In this case, a logical possibility is to supply each point in space with a standard clock measuring the local time. In this general covariant picture, a comparison of the clock rate at different points is possible only through an exchange of (light) signals. Clearly, in this case, the effects of the photon red-shift and the acceleration of the clock rate are indistinguishable. This general covariant definition of the red-shift effect was given by H. Weyl in [7].

REFERENCES

1. V. V. Okorokov, Dokl. Akad. Nauk **378** (5), 617 (2001) [Dokl. Phys. **46**, 400 (2001)].
2. L. B. Okun', K. G. Selivanov, and V. L. Telegdi, Usp. Fiz. Nauk **169**, 1141 (1999).
3. L. B. Okun, K. G. Selivanov, and V. L. Telegdi, Am. J. Phys. **68**, 115 (2000).
4. L. B. Okun, Mod. Phys. Lett. A **15**, 1941 (2000).
5. L. B. Okun, Mod. Phys. Lett. A **15**, 2007 (2000).
6. L. D. Landau and E. M. Lifshitz, *Classical Field Theory* (Nauka, Moscow, 1988; Pergamon, Oxford, 1975).
7. H. Weyl, *Raum, Zeit, Materie* (J. Springer, Berlin, 1923), p. 322.

Institute of Theoretical and Experimental Physics,
Bol'shaya Cheremushkinskaya ul. 25,
Moscow, 117218 Russia
E-mail: okun@heron.itep.ru; selivano@heron.itep.ru

Translated by G. Merzon

=============== **MECHANICS** ===============

On the Possibility of Free Convection in the Earth's Crystalline Mantle

Corresponding Member of the RAS V. N. Anfilogov

Received December 17, 2001

In the classic monograph "Mechanics of Continua" [1], L. Landau and E. Lifshitz stated that the nonuniform heating of a solid phase does not result in the appearance of convection as usually takes place in liquids. The cause of this difference between solids and liquids was regarded by the authors to be evident, and therefore they did not analyze it. However, in geological models describing deformation processes in Earth's rocks down to depths of 3000 km, there are widely used concepts that these processes may result in the generation of circular convective fluxes in the Earth's mantle, with the sizes of convection cells ranging from 700 to 2500 km at temperature gradients of 1–2°C per km [2–4]. The authors of these models assume that even though the viscosity of the Earth's mantle is as great as 10^{23}–10^{25} Pa s, the large sizes of the convection cells make it possible to attain values of the Rayleigh number for which free convection arises in liquids [2, 3]. The authors believe that this is the main argument in favor of free convection becoming possible in the conditions under consideration not only in liquids, but in solids as well. We feel that the problem of free convection in the Earth's mantle is not so simple or evident. Therefore, the possibility of occurring such a process in a nonuniformly heated solid should be analyzed in more detail.

In rheology, two ultimate kinds of deformation processes are indicated, namely, perfectly elastic and perfectly fluid regimes [5]. The deformation of a perfectly elastic body is described by Hooke's law, which states that the stress F is proportional to the strain S:

$$F = ES. \tag{1}$$

Here, E is the elastic modulus. In the case of a perfectly elastic body, its deformation is a reversible process so that, when unloaded, the body is completely restored to its original size and shape. In the case of a perfectly fluid body, even the slightest loads result in irreversible strains that are conserved in the absence of loads [5].

Institute of Mineralogy, Ural Division,
Russian Academy of Sciences, Miass,
Chelyabinskaya oblast, 456317 Russia
E-mail: iminchf@ilmeny.ac.ru

Properties of actual bodies depend on both strain and strain rate. By varying the deformation conditions, we can make an arbitrary material fluid, and rheology can describe its deformation from the general standpoint independently of the aggregate state of the material. Nevertheless, there exist deformation regimes that can take place only in liquids and which are unrealizable, in principle, for bodies with a crystalline structure. One such regime is free convection in a liquid, which is caused by the presence of a vertical temperature gradient in it.

In this study, we consider certain features of the mechanism of displacing structural elements in both crystalline substances and liquids during a deformation process. In a crystalline substance, irreversible strains in it are caused by either plastic flow or creep. The plastic deformation of a solid results in the directed motion of a system of dislocations, with the atoms and ions translating in the lattice sites (Fig. 1). In this case, the strains take the form a stepped shift whose magnitude is a multiple of the site spacing. The viscosity of an isotropic solid subjected to plastic deformation is described by the equation [6]

$$\sigma_{ik} = 2\eta \left(u_{ik} - \frac{1}{3} \delta_{ik} u_{//} \right) + \xi u_{//} \delta_{ik}. \tag{2}$$

Here, σ_{ik} is the dissipation tensor, η and ξ are the viscosities, u_{ik} and $u_{//}$ are the derivatives of the strain tensor with respect to time, and δ_{ik} is the unit tensor. It is important to emphasize that Eq. (2) only formally coincides with the expression of [6] for the viscous strain tensor in a fluid.

Creep is a type of plastic deformation that takes place in crystalline bodies at high temperatures and at low strain rates [7]. This deformation process is realized by transverse dislocation glide when the dislocations creep from one lattice plane to another. This leads to bulk saturation of the crystal by dislocations. As a result, this deformation process occurs under fairly small stresses. In this case, the point defects accumulated at grain boundaries play an important role, because their migration promotes the dislocation
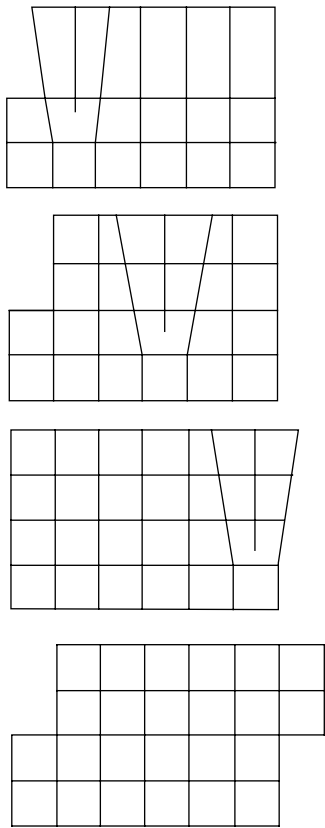
Fig. 1. Scheme of dislocation motion under plastic deformation of crystals.

motion. The viscosity of a solid body subjected to creep deformation is described by the equation [7]

$$\eta = \frac{2kTRL}{3BD\Omega}.$$

Here, L is the grain spacing, K is Boltzmann constant, T is temperature, R is the grain radius, B is a constant, D is the diffusivity, and Ω is the atomic volume.

In the course of plastic deformation or creep of a crystalline body, bulk displacement of its material occurs. In both cases, this proceeds by skipping atoms and ions between lattice sites or vacancies. Such jumps

are possible provided that the load applied to the solid is sufficiently large. It is worth emphasizing that in both cases, these displacements are unidirectional and are directed either parallel or perpendicular to the applied loads. In no case can they form a closed circular motion in the bulk of the solid. If a displacement caused by creep deformation is directed along the temperature gradient, it can be accompanied by heat transfer. By analogy with liquids, this process can be treated as forced convection. In contrast to free convection, forced convection is caused by external forces and ceases after these forces have stopped acting. In order to understand the key differences between deformation in liquids and that in solids, we now consider certain basic features of liquids. There are many models that have been proposed for describing the structure of molecular and nonassociated liquids (such as liquid argon), associated liquids (such as water and silicate melts), ionic liquids (e.g., salt melts), etc. These models treat a liquid as either a disordered crystal lattice (polyhedron hole model), a system of structured fragments separated by vacancies (quasi-crystal model), or as islets with an ordered structure which are separated by cavities filled with gaseous particles (discernible structure model) [8, 9]. A distinctive feature of all these models of fluid bodies is the existence of a certain free volume in them, which is a key difference compared to the crystalline state. The size of this free volume is sufficient for both the deformation of the body and the variation of its shape to occur without a noticeable load.

We now consider a fluid flow under conditions of free convection. The simplest and most obvious example of such a flow is a layer of a liquid bounded by two infinite parallel plates whose temperatures are kept constant, with the temperature of the lower plate being higher than that of the upper plate (Fig. 2). When the liquid is heated, a layer in contact with the lower-plate surface is formed. In this layer, the free volume is larger and therefore the density is lower than those in higher lying layers. As a result, the heated layer begins to float up, forming ascending flows of the heated liquid. A key difference between these ascending flows and arbitrary directed displacements of material in a solid consists in the fact that, in this case, the fluid moves as finite vol-
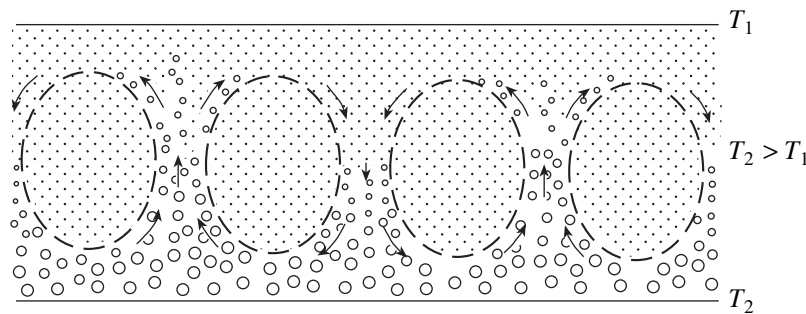
Fig. 2. Vertical cross section of convection cells between two infinite parallel plates. The dimensions of the open circles correspond to the sizes of free volumes in the liquid. The dashed lines bound the domains of a stable structure in the liquid.

umes forming ascending and descending jets between the parallel plates rather than as translations of particles from an equilibrium position to another one. In this case, no specific convective particles originate in the liquid. A system of convection cells with a cross-sectional shape similar to a right hexagon [1] is formed in the space between the plates. The liquid rises to the surface of the upper plate at the center of the hexagon, while the cooled liquid descends at the periphery (Fig. 2) The cell size is determined by the distance between the plates, the temperature difference, and the viscosity of the fluid.

When estimating the possibility of matter moving from the heated part of a body to its cooled part, it is very important to determine the relative values of the forces causing the free convection. This can be performed in the following manner. In contrast to a solid, the shape of a liquid is gravitationally unstable. Therefore, under the action of gravity, the liquid is deformed, taking the form of the vessel in which it is contained. Forces acting in the process of free convection are determined by the difference in the densities of heated and cooled layers of the liquid. These forces are much weaker than the gravitational forces that affect the shape of the liquid. Hence, in this process, the behavior of the liquid to a great extent corresponds to a perfectly fluid body.

We now imagine that instead of a liquid, a giant crystal of a mantle mineral, for example, $MgSiO_3$, is placed between the plates. When the lower part of the crystal is heated, the interatomic distances in this domain (to be more precise, the amplitudes of atomic oscillations with respect to equilibrium positions) slightly increase. However, no free volume or directed pyroxene flow in pyroxene can arise. The strength of solids is much greater than that of liquids. Therefore, in the Earth's gravitational field, a solid conserves its shape for an infinitely long time. Moreover, the difference in densities of heated and cooled parts of the crystal cannot violate interatomic bonds in the crystal and form ascending and descending flows of the crystalline substance. Therefore, the assumption of free convection in the polycrystalline mantle is nonsense from the standpoint of the physics of solids. The reasoning based on the large size of the convection cells is not conclusive because, regardless of the cell sizes, the displacement of the substance in the convection process occurs by the motion of elementary particles that form the structure of the substance.

The groundlessness of the concept of free convection in the mantle is also evident for the following rea-

sons. Let us assume the mantle to be a polycrystalline aggregate with a grain size of 0.5 cm. For a temperature gradient of 1–2°C per km (this value is considered by a number of authors to be sufficient for convection to originate in the mantle), the temperature difference between the lower and upper boundaries of a grain is no higher than 10^{-5}°C. For such a temperature gradient in the grain, internal stresses sufficient for dislocations or vacancies to move in the direction of the gradient cannot, in principle, originate.

The impossibility of free convection in the solid-phase mantle implies that the present-day concept of continental plate tectonics is groundless. Indeed, both the mechanism of motion of the plates and the possibility of their immersion in so-called subduction zones (whose existence has long been questioned) turn out to be extremely doubtful [10]. Nevertheless, this argument does not suggest that large-scale horizontal displacements of crustal blocks are impossible, but it forces us to seek other mechanisms to explain the hidden nature of such motions.

REFERENCES

1. L. D. Landau and E. M. Lifshitz, *Mechanics of Continua* (Gostekhizdat, Moscow, 1953).
2. A. S. Monin, *Early Geological History of the Earth* (Nedra, Moscow, 1987).
3. V. N. Zharkov, *Internal Structure of the Earth and Planets* (Nauka, Moscow, 1983).
4. D. Tercotte and G. Schubert, *Geodynamics* (Wiley, New York, 1982; Mir, Moscow, 1985).
5. F. R. Eirich, in *Rheology: Theory and Applications,* Ed. by R. F. Eirich (Academic, New York, 1956; Inostrannaya Literatura, Moscow, 1956).
6. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 7: *Theory of Elasticity* (Nauka, Moscow, 1965; Pergamon, New York, 1986).
7. G. J. Dyne, in *Rheology: Theory and Applications,* Ed. by R. F. Eirich (Academic, New York, 1956; Inostrannaya Literatura, Moscow, 1956).
8. H. Bloom and J. O. M. Bockris, in *Fused Salts*, Ed. by B. R. Sundheim (McGraw-Hill, New York, 1964; Mir, Moscow, 1966).
9. V. N. Anfilogov, I. B. Bobylev, G. I. Anfilogova, and N. A. Zyuzeva, *Structure and Properties of Silicate-Halogen Melts* (Nauka, Moscow, 1990).
10. S. W. Carey, *Theories of the Earth and Universe: A History of Dogma in the Earth's Science* (Stanford Univ. Press, Stanford, 1988; Mir, Moscow, 1991).

Translated by V. Chechin

<center>═══════════ **MECHANICS** ═══════════</center>

On Nonlinear Stability of Steady Rotation of a Regular Vortex Polygon

L. G. Kurakin and V. I. Yudovich

Presented by Academician G.G. Chernyĭ February 11, 2002

Received February 28, 2002

1. INTRODUCTION

W. Thomson (Lord Kelvin) [1] formulated the stability problem for the permanent rotation of a system of n point vortices situated at the corners of a regular n-gon. He pointed to the profound analogy between this problem and the problem of equilibrium stability for a system of n magnets floating in an external magnetic field. The latter problem was investigated experimentally by Mayer in [2, 3]. Thomson emphasized the importance of determining the largest n value for which the given steady-state regime is stable.

Many authors have investigated the problem. According to [4, 5], the regime is exponentially unstable at $n \geq 8$ and stable in the linear approximation at $n \leq 6$. The case $n = 7$ is questionable, because the zero eigenvalue of the linear problem has a multiplicity of 4, while the normal value is equal to 2.

In this paper, we prove that a regular vortex heptagon is stable. This proof requires a special investigation of nonlinearity; at $n \leq 6$, the linear approximation is sufficient [6]. Thus, the complete answer to Thomson's question is that stability occurs only at $n \leq 7$.

We note that Khazin [7, 8] erroneously stated that nonlinear analysis is necessary at $n \leq 6$ and the given regime is unstable at $n = 7$.

The right conclusion concerning the stability at $n = 7$ was drawn in [9, 10]. However, the analysis carried out in [9] is insufficient, and as the authors of [10] wrote, their result "was accurate to extensive numerical computer calculations."

Excellent reviews of results on the stability of both regular vortex polygons and more complex configurations of point vortices were presented in [11].

In this section based on [12], we present certain general results concerning steady motions in symmetric systems and their stability. In our opinion, this general theory will be useful, in particular, in investigating

Faculty of Mechanics and Mathematics,
Rostov State University, ul. Zorge 5,
Rostov-on-Don, 344090 Russia

many other vortex problems. It is noteworthy that the rigorous definition of the stability of steady motion was obscured in many papers and this fact often led to misunderstanding.

2. STEADY MOTIONS OF SYMMETRIC DYNAMICAL SYSTEMS

In the Banach space V, we consider the autonomous differential equation

$$\dot{u} = F(u), \qquad (2.1)$$

for which the Cauchy problem is globally and uniquely solvable. This means that, for any $v \in V$, there is a single solution $u(t) = N_t v$ defined for all $t \geq 0$, where N_t: $V \longrightarrow V$ is the evolution operator.

Let G be a Lie group and L: $g \mapsto L_g$ be its isomorphism onto a certain subgroup $L(G)$ of the diffeomorphism group $\mathrm{Diff}\, V$ of the space V. We call (N_t, V) a symmetric dynamical system with the symmetry group G if, for any $t \geq 0$ and $g \in G$, the operators N_t and L_g commute with each other:

$$L_g N_t = N_t L_g. \qquad (2.2)$$

According to this equality, the transformations L_g convert motions into motions. Differentiating Eq. (2.2) with respect to t at $t = 0$, we arrive at the symmetry condition in the differential form

$$L_g'(v) F(v) = F(L_g v) \qquad (2.3)$$

for all $g \in G$ and $v \in V$. Here, $L_g'(v)$ is the Frechét derivative of the operator L_g at the point v.

Let $\mathbb{A} = \mathbb{A}_G$ be the Lie algebra of the group G. Each element $a \in \mathbb{A}$ uniquely defines the one-parameter subgroup $g(\tau) = \exp \tau a$, where $\exp$: $\mathbb{A} \longrightarrow G$ is the exponential mapping and a is the generator of the subgroup $g(\tau)$.

We define the generator $T(a)$: $V \longrightarrow V$ (infinitesimal operator) of the one-parameter subgroup $L_{g(\tau)}$ of the group $L(G)$ as

$$T(a) = \left.\frac{d}{d\tau}\right|_{\tau=0} L_{\exp \tau a}. \qquad (2.4)$$

The operator $T(a)$ is generally nonlinear but depends linearly on a.

We call the solution $u(t)$ of Eq. (2.1) steady-state if it can be presented in the form

$$u(t) = L_{\exp ta} v \qquad (2.5)$$

for certain $a \in \mathbb{A}$ and $v \in V$. In other words, steady motion is realized by the transformations of a certain one-parameter subgroup $L_{g(t)}$, $g(t) = \exp ta$ of the group $L(G)$.

Any one-parameter subgroup $g(t) = \exp ta$ and any $v \in V$ satisfy the formula

$$\frac{d}{dt} L_{g(t)} v = L'_{g(t)}(v) T(a) v. \qquad (2.6)$$

Substituting Eq. (2.5) into Eq. (2.1), using Eqs. (2.3) and (2.6), and reducing the reversible operator $L'_{g(t)}(v)$, we obtain the equation of steady motions

$$F(v) = T(a) v, \qquad (2.7)$$

which is closed, at least formally. In a finite-dimensional case, this means that the number of equations coincides with the number of unknowns. The equation is determined, because the presence of the unknown vector a is compensated by the symmetry of Eq. (2.1). If $v \in V$ and $a \in \mathbb{A}$ satisfy Eq. (2.7), steady motion is described by Eq. (2.5)

The simplification of the problem due to the transition from differential equation (2.1) to Eq. (2.7) resembles the separation of variables in the linear case. The equation for a vector function with values in the space V reduces to an equation immediately in the space V. As a result, we obtain a peculiar eigenvalue problem whose parameter is the element a of the Lie algebra of the symmetry group G.

If v is the initial point of a certain steady motion, all points of the orbit $O(v) = \{L_h v: h \in G\}$, which is obtained by the action of the group G and passes through the point v, belong to trajectories of steady motions. In particular, each point $v_\tau = L_{\exp \tau a} v$ belonging to the trajectory of steady-state regime (2.5) satisfies Eq. (2.7) with the same a.

The change of variables $u(t) = L_{g(t)} w(t)$ with $g(t) = \exp ta$ in Eq. (2.1) leads to the equation

$$\dot{w} + T(a) w = F(w), \qquad (2.8)$$

which describes relative motion in the frame of reference given by the subgroup $L_{g(t)}$ of the group $L(G)$. According to Eq. (2.7), the initial point v of steady motion (2.5) is an equilibrium of Eq. (2.8). It is also called the relative equilibrium of Eq. (2.1). Each point $L_{g(\tau)} v$ of trajectory (2.5) is also a relative equilibrium.

Equation (2.8) can be derived for any subgroup $g(t) = \exp ta$ ($a \in \mathbb{A}$). This equation is autonomous precisely because of symmetry. If perhaps only the vector a is changed, the transformations of the group $L(G)$ conserve the form of Eq. (2.8).

In contrast to equilibria, a steady-state regime is never asymptotically stable, because there are other steady-state regimes close to it. In general, it can be stable in the Lyapunov sense for systems with dissipation, when its trajectory is asymptotically stable. For conservative systems, when trajectories of steady-state regimes are not isolated (fill entire submanifolds in the phase space), Lyapunov stability is possible, but only as an infrequent exception. Indeed, if an initial perturbation leads to an adjacent steady-state regime with another trajectory, both steady motions must satisfy a certain isochronism condition. For example, if they are periodic, the perturbation should not change the period. Otherwise, the departure of the perturbed motions from each other in a finite time can become of the order of trajectory diameter. Therefore, other definitions of stability are reasonable in the general theory.

A steady-state solution $u(t) = L_{\exp ta} v$ of Eq. (2.1) is stable in the Routh sense if its trajectory $\mathcal{T} = \{w: w = L_{\exp \tau a} v, \tau \in \mathbb{R}\}$ is a stable family of equilibria of relative-motion equation (2.8). In more detail, for any $\varepsilon > 0$, there exists $\delta > 0$ such that the inequality $\rho(w(t), \mathcal{T}) < \varepsilon$ (ρ is the distance) is satisfied at all $t \geq 0$ for each solution $w(t)$ of Eq. (2.8) when $\rho(w(0), \mathcal{T}) < \delta$.

The steady motion $u_s(t) = L_{\exp ta} v$ is called G-stable if, for each $\delta > 0$, there exists $\varepsilon > 0$ such that the inequality $\|w_0 - v\| < \delta$ at initial time $t = 0$ leads to the inequality $\rho(w(t), O(v)) < \varepsilon$ at all $t > 0$. Here, $w(t)$ is the solution of the Cauchy problem for Eq. (2.8) with the initial condition $w(0) = w_0$.

This steady motion is asymptotically G-stable if it is stable and the invariant set $O(v)$ is attracting; i.e., $\rho(w(t), O(v)) \longrightarrow 0$ at $t \longrightarrow +\infty$ if the initial point w_0 is so close to $O(v)$ that $\rho(w_0, O(v)) < \delta_0$, where the number $\delta_0 > 0$ is determined by the given steady-state regime.

It was assumed above that, for any initial condition $u(0) = u_0 \in V$, the Cauchy problem for Eq. (2.1) has a solution $u(t)$ defined at all $t > 0$. In general, the definition of stability must include the continuability of any solution of Eq. (2.1) with the initial point u_0 in a certain vicinity of the point v onto the semi-infinite interval $[0, \infty)$.

Lyapunov and his followers studied a more general concept of stability with respect to certain variables (see, e.g., [13]).

We note that the concepts of steady motion and its G stability depend on the considered symmetry group G. In order for steady motion to be stable, this group should be maximal in a certain sense.

CERTAIN CRITERIA FOR STABILITY OF STEADY MOTIONS

Hereafter, all given functions are assumed to be C^∞-smooth.

Proposition 2.1. *A steady-state regime* $u^0(t) = L_{\exp ta} v$ *is stable in the sense of Routh and G-stable under the following conditions:* (i) *The function* $u^0(t) = L_{\exp ta} v$ *is a steady-state solution of Eq.* (2.1) *with the symmetry group G, and its orbit* $O(v)$ *is compact;* (ii) *The function V is given in the vicinity* $O(v)$, *invariant with respect to the transformation group* $L_{\exp ta}$ *generating the regime* u^0, *and reaches a transversally strict minimum* (*or maximum*) *in this orbit* (*locally strict minimum or maximum in the space of orbits of the group* $L_{\exp ta}$); (iii) *The derivative* $\dot{V}$ *is negative* (*positive*) *according to Eq.* (2.1) *and, in particular, can be an integral of this equation.*

We define this proposition concretely for the following Hamiltonian system in $\mathbb{R}^{2n}$:

$$\dot{u} = J \operatorname{grad} H(u). \qquad (2.9)$$

We conventionally assume that the symmetry group G acts in the phase space $\mathbb{R}^{2n}$ of Eq. (2.9) by means of symplectic diffeomorphisms. In this case, the Hamilton equation

$$\dot{u} = J \operatorname{grad} M(u) \qquad (2.10)$$

with the Hamiltonian M corresponds to each one-parameter subgroup $\exp ta$ of the group G. In addition, the operator $T(a)$ is the right-hand side of Eq. (2.10):

$$T(a)u = J \operatorname{grad} M(u). \qquad (2.11)$$

The algebra of functions with the Lie multiplication given by the Poisson brackets $\{M_1(u), M_2(u)\} = (J \operatorname{grad} M_1(u), M_2(u))$ corresponds to the Lie algebra $\mathbb{A}$.

The Noether theorem for Eq. (2.9) states that, if $J \operatorname{grad} M(u)$ is a symmetry, $M(u)$ is an integral of Eq. (2.9).

Equation (2.7), which describes steady motion generated by the symmetry group with the Hamiltonian $\lambda M(u)$ ($\lambda \in \mathbb{R}$), is written as

$$\operatorname{grad}(H(v) - \lambda M(v)) = 0. \qquad (2.12)$$

The corresponding equation of relative motion (2.8) takes the form

$$\dot{w} = J \operatorname{grad}(H(w) - \lambda M(w)). \qquad (2.13)$$

The steady-state regime $u^0(t)$ can be found as a solution of the Cauchy problem

$$\dot{u}^0 = J \operatorname{grad} M(u^0), \quad u^0|_{t=0} = v. \qquad (2.14)$$

Equation (2.12) allows the following variational interpretations.

(1) At a given λ, v is the critical point for the function $H - \lambda M$.

(2) The point v is critical for the contraction of the function H to the set of the M: $M(u) = M(v)$ function level.

(3) At $\lambda \neq 0$, v is the critical point for the contraction of the function M to the set of the $H(u) = H(v)$ level.

Here, in items (2) and (3), the quantity λ is the Lagrange multiplier for these problems of a conditional extremum.

We call the function $H - \lambda M$, where λ is determined by the given steady motion, the relative (reduced) Hamiltonian. This is the most natural Lyapunov function for the problem of the Routh stability. The relative Hamiltonian $H - \lambda M$ is an integral of the equation of relative motion (2.13) and is invariant under the transformations $L_{\exp \tau a}$ at all $\tau \in \mathbb{R}$. Indeed, the function H is invariant according to the original symmetry condition, and the function M is the Hamiltonian corresponding to the transformation group $L_{\exp \tau a}$. Therefore, Proposition 2.1 leads directly to the following statement.

Proposition 2.2. *Let the relative Hamiltonian* $H - \lambda M$ *reach a transversally strict minimum or maximum in the trajectory of the steady-state regime* u^0 *with a compact orbit. Then, the steady-state regime* u^0 *is stable in the Routh sense and G-stable.*

The instability of a steady motion in the Routh sense does not exclude the stability of an invariant set of steady motions (the group orbit of the given steady motion). If this stability occurred, perturbations would grow only along the invariant set. In reality, exponential instability in the linear approximation generally causes the growth of perturbations in the transverse direction as well. For families of periodic motions, this fact was established in [14, 15].

3. STEADY ROTATION OF A SYSTEM OF POINT VORTICES

Plane motion of a system of n point vortices is described by the following Kirchhoff equations in the Hamiltonian formalism:

$$\kappa_k \frac{dx_k}{dt} = \frac{\partial H}{\partial y_k}, \quad \kappa_k \frac{dy_k}{dt} = -\frac{\partial H}{\partial x_k}, \quad 1 \leq k \leq n. \quad (3.1)$$

Here,

$$H = -\frac{1}{4\pi} \sum_{1 \leq j < k \leq n} \kappa_j \kappa_k \ln \sqrt{(x_j - x_k)^2 + (y_j - y_k)^2} \quad (3.2)$$

is the Hamiltonian, x_k and y_k are the Cartesian coordinates of the kth vortex, and κ_k is its strength.

Equation (3.1) corresponds to the Poisson brackets for the functions U and V:

$$\{U, V\} = \sum_{k=1}^{n} \frac{1}{\kappa_k} \left(\frac{\partial U}{\partial x_k} \frac{\partial V}{\partial y_k} - \frac{\partial U}{\partial y_k} \frac{\partial V}{\partial x_k} \right). \qquad (3.3)$$

Kirchhoff found the following four integrals of system (3.1): the function H, the momentum components

$$I_1 = \sum_{k=1}^{n} \kappa_k x_k \text{ and } I_2 = \sum_{k=1}^{n} \kappa_k y_k, \text{ and the total moment of}$$

inertia $M = \sum_{k=1}^{n} \kappa_k (x_k^2 + y_k^2)$.

Introduction of the complex variables $z_k = x_k + i y_k$ and $z_k^* = x_k - i y_k$ reduces the Kirchhoff equations to the system of n complex differential equations

$$\kappa_k \dot{z}_k = -2 i H_{z_k^*}, \quad k = 1, 2, \ldots, n. \quad (3.4)$$

Now,

$$H = -\frac{1}{4\pi} \sum_{1 \le j < k \le n} \kappa_j \kappa_k \ln[(z_j - z_k)(z_j^* - z_k^*)] \quad (3.5)$$

is a function of the variables z_k and z_k^* and the equations of motion take the form

$$\dot{z}_k^* = \frac{1}{2\pi i} \sum_{j=1}^{n} {}' \frac{\kappa_j}{z_k - z_j}, \quad k = 1, 2, \ldots, n, \quad (3.6)$$

where the prime denotes omission of the term with $j = k$. The phase space Z of system (3.6) is $\mathbb{C}^n$ with cuts along all hyperplanes $z_j = z_k$ $(j \ne k)$.

It is noteworthy that, if the strengths have identical signs, the Cauchy problem for system (3.6) is globally solvable and its trajectories are compact in Z.

Below, we consider the symmetry group. System (3.6) is invariant with respect to the group G of the Euclidean motions of the plane $\mathbb{R}^2$. For any point $z = (z_1, z_2, \ldots, z_n) \in Z$ and any motion $g \in G$, the action $g \mapsto L_g$ of this group in the phase space Z is determined by the equality $L_g z = (g z_1, g z_2, \ldots, g z_n)$. The group G includes the specular reflection $j: z \mapsto z^*$, translations $g^{tr}: z \mapsto z + h$ $(h \in \mathbb{C})$, and rotations $g^{rot}: z \mapsto e^{i\alpha} z$ $(\alpha \in \mathbb{R})$. We note that, according to the Noether theorem, the momentum integrals and the integral of moment of inertia arise due to the translational and rotational invariances of the Hamiltonian H, respectively.

The equation of steady motions (2.7) and (2.11) that corresponds to the translation group g^{tr} takes the form

$$v^* = \frac{1}{2\pi i} \sum_{j=1}^{n} {}' \frac{\kappa_j}{z_k - z_j}, \quad k = 1, 2, \ldots, n \quad (3.7)$$

with the unknowns $v, z_1, z_2, \ldots, z_n \in \mathbb{C}$. After analyzing it, we conclude that a nonequilibrium steady-state regime corresponding to the translation group can exist only when $\sum_{k=1}^{n} \kappa_n = 0$.

According to Eq. (2.5), the steady motion corresponding to the rotation subgroup g^{rot} is sought in the form $z_k = e^{i\omega t} u_k$. Simultaneously, the equation of steady motions (2.7) and (2.11) is written as

$$i \omega u_k^* = \frac{1}{2\pi i} \sum_{j=1}^{n} {}' \frac{\kappa_j}{u_k - u_j}, \quad k = 1, 2, \ldots, n \quad (3.8)$$

for the unknowns $u_1, u_2, \ldots, u_n \in \mathbb{C}$ and $\omega \in \mathbb{R}$.

For identical strengths $\kappa_1 = \ldots = \kappa_n = \kappa$, system (3.8) has the familiar exact solution ($R > 0$ is arbitrary)

$$u_k = R e^{\frac{2\pi}{n} i(k-1)}, \quad k = 1, 2, \ldots, n,$$
$$\omega = \kappa \frac{n-1}{4\pi R^2}. \quad (3.9)$$

The moment integral is related to R by the equality

$$M = \kappa \sum_{k=1}^{n} |z_k|^2 = \kappa n R^2.$$

The corresponding steady-state regime is described by the equations

$$z_k(t) = R e^{i\omega t} u_k, \quad k = 1, 2, \ldots, n. \quad (3.10)$$

The existence of steady-state rotation regimes is also proven in the general case of arbitrary strengths κ_k with identical signs.

4. STABILITY OF A REGULAR VORTEX n-GON

It is assumed that all vortices have the identical strength κ. Then, a system of vortices that are situated along a circle of radius R at the corners of a regular n-gon rotates with a constant angular velocity. This motion corresponds to steady-state solution (3.10) and relative-motion equation (2.13) with the relative Hamiltonian

$$E(w) = H(w) - \lambda M(w),$$
$$M = \kappa \sum_{k=1}^{n} (v_k^2 + \mu_k^2), \quad \lambda = -\frac{\omega}{2}. \quad (4.1)$$

Here, $w = (v_1, v_2, \ldots, v_n, \mu_1, \mu_2, \ldots, \mu_n) \in R^{2n}$ and λ is determined from the equation of steady motion (2.12).

In each plane of the variables (v_k, μ_k), we introduce polar coordinates and represent $v_k + i\mu_k$ in the form

$$v_k + i\mu_k = R(1 + r_k) e^{i\left(\frac{2\pi}{n}(k-1) + \theta_k\right)}, \quad R \ge 0. \quad (4.2)$$

In terms of the variables $r = (r_1, r_2, \ldots, r_n)$ and $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$, steady motion (3.10) corresponds to the continuous family that involves equilibria of sys-

tem (2.13) and is situated on the straight line $\Gamma = \{(r, \theta) \in R^{2n}: r = 0, \theta_1 = \theta_2 = \ldots = \theta_n\}$.

The Taylor expansion of the function $E(w(\rho))$, $\rho \overset{\text{def}}{=} (r, \theta)$ has the following identical form in the vicinity of each equilibrium of the family Γ:

$$E(w(\rho)) = \frac{\kappa^2}{4\pi}(E_0 + E_2(w(\rho))$$
$$+ E_3(w(\rho)) + E_4(w(\rho)) + \ldots). \qquad (4.3)$$

Here, the ellipsis means the terms of powers exceeding 4,

$$E_2(w(\rho)) = (B_1 r, r) + (B_2 \theta, \theta),$$

$$E_3(w(\rho)) = (B_3 r, r^2) + (B_4 \theta, r^2)$$
$$+ 2(B_4 r, r\theta) - (B_4 \theta, \theta^2),$$

$$E_4(w(r, 0)) = (B_5 r^2, r^2) + (B_6 r, r^3),$$

where (for $k = 2, 3$)

$$r^k \overset{\text{def}}{=} (r_1^k, r_2^k, \ldots, r_n^k), \quad r\theta \overset{\text{def}}{=} (r_1 \theta_1, r_2 \theta_2, \ldots, r_n \theta_n),$$
$$\theta^2 \overset{\text{def}}{=} (\theta_1^2, \theta_2^2, \ldots, \theta_n^2).$$

All matrices B_j ($j = 1$–6) are circulant matrices, i.e., polynomials of the cyclic matrix $\mathbb{C} = (c_{ij})_{ij=1}^n$, whose nonzero elements are $c_{1,2} = \ldots = c_{n-1,n} = c_{n,1} = 1$. As a result,

$$B_1 = b_1 I + B_0, \quad B_2 = b_2 I - B_0, \quad B_3 = b_3 I - B_0,$$

$$B_0 \overset{\text{def}}{=} \sum_{m=1}^{n-1} a_m \mathbb{C}^m, \quad B_4 = 2 \sum_{m=1}^{n-1} \sin\frac{2\pi m}{n} a_m^2 \mathbb{C}^m,$$

$$B_5 = b_5 I + \frac{3}{2} \sum_{m=1}^{n-1} a_m^2 \mathbb{C}^m, \quad B_6 = -2 \sum_{m=1}^{n-1} \cos\frac{2\pi m}{n} a_m^2 \mathbb{C}^m,$$

where

$$a_m = \frac{1}{2}\frac{1}{1 - \cos\dfrac{2\pi m}{n}}, \quad b_1 = \frac{1}{12}(n-1)(11-n),$$

$$b_2 = \frac{1}{12}(n^2 - 1),$$

$$b_3 = \frac{1}{12}(n-1)(n-3),$$

$$b_5 = \frac{1}{1440}(n-1)(n^3 + n^2 - 109n + 251).$$

The eigenvalues λ_{1k} and λ_{2k} of the matrices B_1 and B_2 are, respectively,

$$\lambda_{1k} = n - 1 - \frac{1}{2}k(n-k),$$

$$\lambda_{2k} = \frac{1}{2}k(n-k), \quad 1 \le k \le n.$$

The following theorem justifies the application of the linearization method to the problem of stability of a vortex n-gon at $n \ne 7$. It is noteworthy that this theorem treats instability in the strongest sense; i.e., the invariant set of steady rotations is transversally unstable.

Theorem 3.1. *Steady rotation* (3.10) *of a regular vortex n-gon is stable in the Routh sense* (*and G-stable*) *at $n \le 6$ and unstable at $n \ge 8$.*

Proof. At $n \le 6$, all eigenvalues of the matrices B_1 and B_2 except $\lambda_{2n} = 0$, which corresponds to the proper subspace Γ, are positive. Therefore, the quadratic form E_2 is positive-definite in the subspace $\Gamma^\perp = R^{2n} \ominus \Gamma$. Expansion (4.3) indicates that the relative Hamiltonian $E(w(\rho))$ reaches a transversally strict minimum in the family of equilibria Γ. Thus, the conditions of Statement 2.2 are satisfied.

Proof of instability for $n \ge 8$ is based on the fact that the transverse component of the system under consideration is separated, i.e., is independent of the tangential component. Therefore, the instability of the system follows from the general Lyapunov theorems concerning instability.

At $n = 7$, the matrix B_1 has a double zero eigenvalue so that, in order to solve the problem of the transverse minimum, one should use the next terms of expansion (4.3).

In the R^7 space, the circulant matrices B_1 and B_2 have the common proper basis

$$h_m = (1, \cos(m\alpha), \ldots, \cos(6m\alpha)),$$

$$h_{7-m} = (0, \sin(m\alpha), \ldots, \sin(6m\alpha)),$$

$$h_7 = (1, 1, 1, 1, 1, 1, 1), \quad m = 1, 2, 3, \quad \alpha = \frac{2\pi}{7}.$$

As a result, $B_j h_k = \lambda_{jk} h_k$, where $j = 1, 2$ and $k = 1, \ldots, 7$.

The vectors r and θ can be represented in this basis as

$$r = \hat{r} + \tilde{r}, \quad \hat{r} = \xi_3 h_3 + \xi_4 h_4 \in \ker B_1,$$

$$\tilde{r} = \xi_1 h_1 + \xi_2 h_2 + \xi_5 h_5 + \xi_6 h_6 + \xi_7 h_7, \qquad (4.4)$$

$$\theta = \theta_\perp + \zeta_7 h_7, \quad \theta_\perp = \sum_{k=1}^{6} \zeta_k h_k.$$

The Hamiltonian $E(w(\rho))$ is independent of the variable ζ_7 and, in the vicinity of the family Γ, has the

asymptotic form

$$E(w(\rho)) = \frac{\kappa^2}{4\pi}(E_0 + \mathbb{S}(\hat{r}, \tilde{r}, \theta_\perp)$$

$$+ o(|\hat{r}|^4 + |\tilde{r}|^2 + |\theta_\perp|^2)), \quad (\hat{r}, \tilde{r}, \theta_\perp) \longrightarrow 0, \tag{4.5}$$

where the function $\mathbb{S}$ has the form

$$\frac{2}{7}\mathbb{S} = 3\xi_1^2 + \xi_2^2 + \xi_5^2 + 3\xi_6^2 + 12\xi_7^2$$

$$+ 3\zeta_1^2 + 5\zeta_2^2 + 6\zeta_3^2 + 6\zeta_4^2 + 5\zeta_5^2 + 3\zeta_6^2$$

$$+ \frac{9}{2}\xi_3^2\xi_1 + 6\xi_3^2\xi_7 - 9\xi_3\xi_4\xi_6 + 6\xi_4^2\xi_7 - \frac{9}{2}\xi_4^2\xi_1$$

$$+ \frac{9}{2}\xi_3^2\zeta_6 + 9\xi_4\xi_3\zeta_1 - \frac{9}{2}\xi_4^2\zeta_6 + 6\frac{3}{8}(\xi_3^2 + \xi_4^2)^2. \tag{4.6}$$

For arbitrary $x \in R$, we present the polynomial $\mathbb{S}$ as the following sum:

$$\frac{2}{7}\mathbb{S} = \mathbb{S}_1 + \mathbb{S}_2 + \xi_2^2 + \xi_5^2 + 5\zeta_2^2 + 6\zeta_3^2 + 6\zeta_4^2 + 5\zeta_5^2, \tag{4.7}$$

$$\mathbb{S}_1 = 3\xi_1^2 + 12\xi_7^2 + 3\zeta_6^2 + \frac{9}{2}\alpha_3\xi_1 + 6\alpha_3\xi_7 + 6\alpha_4\xi_7$$

$$- \frac{9}{2}\alpha_4\xi_1 + \frac{9}{2}\alpha_3\zeta_6 - \frac{9}{2}\alpha_4\zeta_6 + 6\frac{3}{8}\alpha_3^2 + 6\frac{3}{8}\alpha_4^2 + 2x\alpha_3\alpha_4,$$

$$\mathbb{S}_2 = 3\xi_6^2 + 3\zeta_1^2 - 9\beta\xi_6 + 9\beta\zeta_1 + \left(12\frac{3}{4} - 2x\right)\beta^2,$$

where $\alpha_3 = \xi_3^2$, $\alpha_4 = \xi_4^2$, and $\beta = \xi_3\xi_4$. At $x \in \left(-4\frac{7}{8}, -\frac{3}{8}\right)$, the quadratic forms $\mathbb{S}_1$ and $\mathbb{S}_2$ are simultaneously positively definite, and therefore the Sylvester criterion is satisfied.

Thus, the function $\mathbb{S}$ has a strict local minimum at zero, and the Hamiltonian $E(w(\rho))$ satisfies the conditions of Proposition 2.2. Consequently, the following theorem is valid.

Theorem 3.2. *The steady rotation* (3.10) *of a regular vortex heptagon is stable in the Routh sense and G-stable.*

REFERENCES

1. W. Thomson, Nature **18**, 13 (1878).
2. A. M. Mayer, Nature **17**, 487 (1878).
3. A. M. Mayer, Nature **18**, 258 (1878).
4. J. J. Thomson, *A Treatise on the Motion of Vortex Rings* (Macmillan, London, 1883).
5. T. H. Havelock, Philos. Mag., Ser. 7 **11** (70), 617 (1931).
6. L. G. Kurakin, Dokl. Akad. Nauk **335** (6), 729 (1994) [Phys. Dokl. **39**, 284 (1994)].
7. L. G. Khazin, Dokl. Akad. Nauk SSSR **230** (4), 799 (1976) [Sov. Phys. Dokl. **21**, 567 (1976)].
8. L. G. Khazin, in *N. E. Kochin and Development of Mechanics* (Nauka, Moscow, 1984), pp. 174–185.
9. G. J. Mertz, Phys. Fluids **21** (7), 1092 (1978).
10. H. E. Cabral and D. S. Schmidt, SIAM J. Math. Anal. **31** (2), 231 (1999).
11. H. Aref, J. Fluid Mech. **290**, 167 (1995).
12. V. I. Yudovich, Doctoral Dissertation in Physics and Mathematics (Inst. Prikl. Mekh. Akad. Nauk SSSR, Moscow, 1972).
13. V. V. Rumyantsev, Vestn. Mosk. Univ., Ser. 1: Mat., Mekh., No. 4, 9 (1957).
14. V. I. Yudovich, Dokl. Akad. Nauk SSSR **195** (3), 574 (1970).
15. N. V. Nikolenko, Izv. Vyssh. Uchebn. Zaved., Mat., No. 9, 54 (1973).

Translated by Yu. Verevochkin

═══════════ **MECHANICS** ═══════════

On the Fundamental Possibility of Using Cable Systems to Carry out Gravitational Maneuvers in the Vicinity of a Planet

I. M. Sidorov

Presented by Academician A.A. Galeev January 30, 2002

Received February 7, 2002

In [1], we gave preliminary estimates of the possibility of designing transport routes in space on the basis of cable systems. We propose to displace space objects by arranging two-way transport of equal-mass loads by their exchange at contact centers.

In this study, we extend the problem of applying the cable systems and consider the possibility of directly converting the mechanical energy of a space-object's motion into electric energy and vice versa. In order to implement this project, it is necessary to solve the following problems.

(i) The direct conversion of the energy accumulated aboard a space object, namely, in a cable system, into the mechanical energy of orbital motion. This implies a variation of the space-object's orbit by means of direct energy conversion without employing jet propulsion. We propose to use as an energy source an electric accumulator or a generator that produces electric energy aboard a spacecraft. The same scheme also makes it possible to realize the reverse process, i.e., the conversion of mechanical energy into electric energy.

(ii) The transfer of mechanical energy from one object to another.

(iii) The use of the mechanical energy of motion of natural celestial bodies, in particular, of the moon.

When a space object performs a gravitational maneuver near the moon, the mechanical energy of this object can increase. This energy can be used to change a spacecraft's orbit or be converted into electric energy aboard the spacecraft.

The purpose of this study is to show that the problems formulated above can be solved on the basis of existing systems of space technology and modern materials.

For practical realization of this project, it is necessary to preliminarily deploy a group of cable systems in circumterrestrial orbits. The design of each object of this group must be specialized for the optimal realization of a particular problem. Each object is assumed to occupy a preassigned orbit. For the optimal choice of parameters of this group, it is necessary to solve a number of scientific and technological problems. In addition to the problems listed in this paper, the construction of this space system will make it possible to carry out transport operations and to launch spacecrafts into their orbits at lower costs than those inherent in conventional means of rocket engineering. In the framework of this study, we assume that a group of cable systems exists, and we develop principles of its operation.

The group consists of cable systems with a unified structural scheme. Henceforth, we call this cable system a sling. The sling shown in Fig. 1 is made up of two blocks *B* and *C*, which are linked by a long cable. The
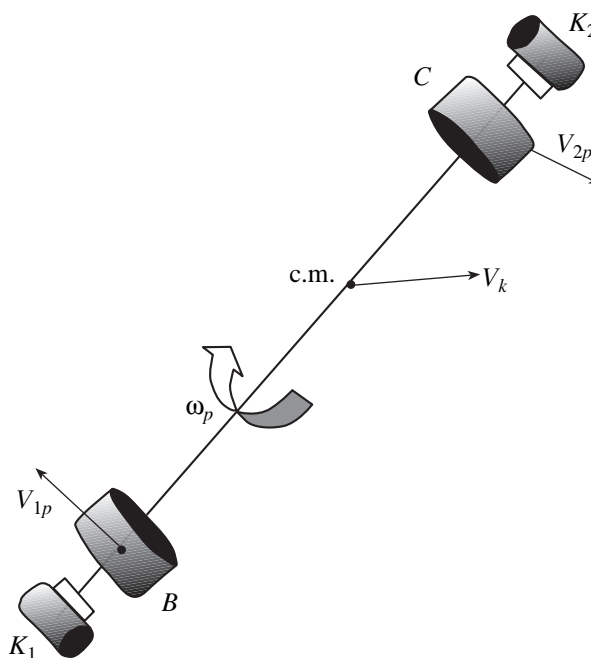
Institute for Space Research,
Russian Academy of Sciences, ul. Profsoyuznaya 84/32,
GSP-7, Moscow, 117810 Russia

Fig. 1. Sling-motion scheme.

blocks are mated with detachable containers K_1 and K_2. The cable system is launched into a satellite orbit. The mean angular velocity of system rotation with respect to its center of mass is ω_p, the rotation plane being coincident with the orbit plane.

1. ENERGY CONVERSION

We denote by $v_{1p}(B)$ and $v_{2p}(C)$ the tangential velocities of the blocks B and C with respect to the center of mass of the cable system. Modern high-strength materials make it possible to design a cable system several hundred kilometers long with tangential velocities on the order of 1500 m/s.

The dynamics of the cable-system motion was investigated on the basis of integrating equations of motion [1] in the gravitational field of two loads linked by an imponderable stretchable cable. The motion proceeds in the orbit plane. We consider a variant of an equatorial orbit; therefore, the orbital plane undergoes no perturbations due to the Earth's nonsphericity.

The cable is assumed to be in tension during rotation, the tension force being proportional to the cable elongation according to Hooke's law:

$$
\begin{aligned}
m_j x_j'' &= -\frac{m_j \mu_1 x_j}{R_j^3} + \gamma_j F(dL)\frac{x_2 - x_1}{L_u}, \\
m_j y_j'' &= -\frac{m_j \mu_1 y_j}{R_j^3} + \gamma_j F(dL)\frac{y_2 - y_1}{L_u}, \quad j = 1, 2.
\end{aligned}
\tag{1}
$$

Here, x_j and y_j are the coordinates of two bodies; $\gamma_j = 1$ for $j = 1$ and $\gamma_j = -1$ for $j = 2$; R is the Earth's radius; $\mu_1 = 398606$ km^3/s^2; m_j is the mass of the end blocks; $R_j = \sqrt{x_j^2 + y_j^2}$; and $L_u = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$.

The cable tension is determined by the expression

$$
F(dL) = \frac{dL E_m S_h}{L_0}.
\tag{2}
$$

Here, E_m is the elastic modulus of the cable material, S_h is the cable cross-sectional area, dL is the cable elongation, $dL = L_u - L_0$, and L_0 is the cable length in the nonstretched state.

Along with a simplified model of an imponderable cable, we investigated the model of spatial motion for a sling with a heavy stretchable cable. The qualitative results for this model coincide with the solution to Eqs. (1).

In Eqs. (1), we multiply each of the first two equations by x_1' and y_1', respectively, and the other two equations—by x_2' and y_2'. Furthermore, we sum all the equations obtained and take the integral of both the right-hand and left-hand sides. As a result, we obtain the expression for the total mechanical energy Q_s of the sling:

$$
Q_s = Q_1 + Q_2 + Q_p,
$$

$$
Q_1 = m_1\left(\frac{\mu_1}{R} - \frac{\mu_1}{R_1} + \frac{v_1^2}{2}\right), \quad Q_2 = m_2\left(\frac{\mu_1}{R} - \frac{\mu_1}{R_2} + \frac{v_2^2}{2}\right),
\tag{3}
$$

$$
Q_p = \frac{E_m S_h (L_0 \ln L_u - L_u)}{m_1}, \quad v_j^2 = x_j'^2 + y_j'^2, \quad j = 1, 2.
$$

In (3), Q_1 and Q_2 determine the mechanical energy of the end blocks of the sling. The quantity Q_p characterizes the energy of longitudinal vibrations of the stretchable cable.

Certain results in the analysis of the sling-motion dynamics were obtained using the methods of multiple-frequency vibrations in nonlinear systems [2]. However, numerical integration of Eqs. (1) was used as a basic method of investigation. To realize this approach, it is necessary to select variants of particular numerical values for cable-system parameters and to set the initial conditions for Eqs. (1) in the corresponding manner.

When integrating Eqs. (1), we additionally determined the coordinates of the motion of the sling's center of mass:

$$
x_c = \frac{k_m x_2 + x_1}{1 + k_m}, \quad y_c = \frac{k_m y_2 + y_1}{1 + k_m}, \quad k_m = \frac{m_2}{m_1}.
$$

The current altitude of the sling's center of mass in its orbital motion is $H = \sqrt{x_c^2 + y_c^2} - R$.

The tangential velocities v_{1p} and v_{2p} of blocks B and C with respect to the cable-system's center of mass determine the rotation regime of the sling around its center of mass. Upon decreasing these velocities, a resonance regime is possible for which the period of rotation around the center of mass equals the period of orbital rotation of the center of mass. Since this regime is unstable, the tangential velocities are chosen from the condition for the orbital-rotation period T_0 to exceed the period T_p of the sling rotation. In this case, the rotation regime is steady. Depending on the initial conditions, the sling rotates either clockwise or counterclockwise.

We also calculated the components (L_x, L_y) of the vector directed along the stretched cable from block B to block C with respect to the orbital system of coordinates, i.e., the components along both the radius vector and the perpendicular to it. The components $L_x(t)$ and $L_y(t)$ are periodic functions that can be represented in

the form of the sum of harmonics with frequencies that are multiples of $\omega_p = \dfrac{2\pi}{T_p}$. The cable tension $F(dL)$ oscillates about a mean value with the double frequency $2\omega_p$. $F(dL)$ attains its maximum or minimum values if the cable is directed along the radius vector or the perpendicular to it, respectively.

We now sequentially consider the methods of solving the formulated problems.

The idea that orbit parameters can vary with the size of a space object was advanced in the well-known monograph [4]. In the present study, we consider a sling to be such an object and show that, by properly varying the geometric size of the cable system, it is possible to transform the orbital parameters and to convert the energy accumulated aboard a space object into mechanical energy and vice versa. We assume that reversible electric motors are installed in the end blocks of the sling operating according to the recuperation principle, as well as accumulators and electric-current sources.

The sling dimensions vary provided that an electric motor situated in one or both end-block motors begins at once to pull or release the cable in the time interval $(0 < t < T)$ so that the cable's free-state length $L_0(t)$ corresponds to the formula

$$L_0(t) = L_0 - l(t). \tag{4}$$

To carry out this operation, the energy

$$\Delta_1 Q = \int_0^T \frac{dl(t)}{dt} F(t) dt \tag{5}$$

is required, where $F(t)$ is the cable tension and $\dfrac{dl(t)}{dt}$ is the rate of cable-length variation provided by the electric motors. The recuperation loss is ignored in formula (5). According to formula (3), at the initial time $t = 0$, the total mechanical energy of the sling is $Q_s(0)$. At $t > T$, when the operation of the motors is completed, the sling-motion parameters are changed, and, according to (3), the total mechanical energy of the sling becomes equal to $Q_s(T)$. The results of integrating equations of motion (1) within the interval $(0 < t < T)$ with a cable length varying according to law (4) show that

$$Q_s(T) - Q_s(0) = \Delta_1 Q; \tag{6}$$

where $\Delta_1 Q$ is determined as a result of calculating integral (5). If $\Delta_1 Q > 0$, this implies that the electric energy was converted into the mechanical energy of the sling. If $\Delta_1 Q < 0$, the inverse conversion takes place.

Furthermore, we consider certain options for controlling the cable length.

In option 1, the cable is pulled with a constant velocity, $\Delta l > 0$, $l(t) = t\Delta l$. When the tangential velocities increase, the cable tension and the mechanical energy

of the sling also grow. According to relation (5), we determine the energy expenditures for the work of the electric motors. When the cable is released, $\Delta l < 0$, the processes have the opposite sign and the corresponding quantity of energy arrives at the accumulators of the end blocks.

In option 2, we propose a law (4) of cable-length variation such that after completing the control cycle, the electric energy is converted into mechanical energy, while the sling dimensions remain the same. We use the fact that the tension force $F(t)$ of the sling vibrates about an average position with a frequency $2\omega_p$. We propose to vary the cable length within the interval $0 < t < T$ according to the law

$$l(t) = p_1 f_\omega(t), \quad f_\omega(t) = \frac{L_x L_y}{L_u^2}. \tag{7}$$

The quantity $f_\omega(t)$ represents a sinusoid with the frequency $2\omega_p$, and p_1 is the parameter. When the control is set in this form, synchronization is provided between the excitation-force frequency and the sling-rotation frequency that varies in the control process. The quantities L_x and L_y can be measured by autonomous means for determining the angular orientation aboard the space object. In option 2, an increment in the mechanical energy of the sling is distributed differently than in option 1: the sling-rotation velocity varies simultaneously with the orbital velocity of the sling's center of mass. If $p_1 > 0$ in control law (7), and the sling rotates clockwise, the angular velocity ω_p of the sling rotation decreases, the long semiaxis a_p of the orbit increases, and the energy supply $\Delta_1 Q$ in the accumulators grows. For $p_1 < 0$, the variations of the parameters have the opposite sign. If the sling rotates counterclockwise and $p_1 > 0$, the quantities ω_p and a_p increase, while $\Delta_1 Q$ decreases; i.e., the energy is spent.

In option 3, we simultaneously realize control according to options 1 and 2. This makes it possible to solve the problem of winding the sling cable; i.e., by pulling the cable, we are able to join blocks B and C. We also solve the problem of unrolling the cable to a given length and imparting a given angular rotation velocity to the sling. The law of the cable-length variation in option 3 is

$$l(t) = t\Delta l + p_1 f_\omega(t).$$

For the chosen parameters Δl and p_1, when the cable is being pulled, the quantity Δl reduces the cable length, while the quantity p_1 decreases the angular velocity ω_p. After completing the process of winding the sling, the long semiaxis a_p of the center-of-mass orbit increases, and additional energy arrives at the accumulator. When the sling is unwound, energy expenditure is required.

In options 1 and 2, the angular velocity of the sling rotation varies, which limits the possible values of the converted energy $\Delta_1 Q$. In option 4, the entire energy is

spent on varying the orbital parameters and the angular velocity is virtually invariable. This enables us to convert a considerably higher energy $\Delta_1 Q$. In option 4, the orbit of the motion of the sling's center of mass must be elliptic. In this case, it turns out that the cable tension varies depending on the altitude H of the orbit. The law of cable-length variation is

$$L_0(t) = L_0 + p_1 f_\omega(t) + h_p H f_\omega(t).$$

When choosing the corresponding values of p_1 and h_p, the following evolution of motion parameters is observed. The consumption of electric energy is associated with the decrease in the orbit perigee r_p, but the orbit apogee r_a increases much more strongly. Thus, the long semiaxis a_p of the orbit, the mechanical energy, and the eccentricity e_p also increase. If the values p_1 and h_p are chosen with the opposite sign, the evolution of parameters is opposite and the energy supply in the accumulators increases. The quantity ω_p remains invariable. A sling operating according to option 4 must be specialized for solving the problems of conversion of electric energy into mechanical energy and vice versa. The sling orbit must have optimal parameters, for example, $r_p = 500–600$ km and $r_a = 2500–3500$ km. At considerable distances from the Earth, the efficiency of the operations decreases. The converted mechanical energy should be transferred to another object of the cable-system group for the purpose of a preset variation of spacecraft orbits.

2. ENERGY TRANSFER FROM ONE SPACE OBJECT TO ANOTHER

We consider the model of a possible interaction between two spacecraft that move in the same plane along elliptic orbits with various orbital parameters as in Fig. 2. We assume that the collision of these two objects has occurred at point D. At the moment of collision, the relative velocity of one body with respect to another can attain 1000–1500 m/s. It is evident that, if we consider conventional spacecraft, mating is excluded and failure of the contacting objects will occur. However, if one of these spacecrafts is a sling, then their contact and mating into a unified object is possible. In Fig. 2, we show a general case of the contact. Near point D, the intersection of orbits of the objects occurs along lines *1* and *2*. Object A_1 is the sling, whose center of mass moves along line *1*. The components of the velocity vector are $\mathbf{u}_1 \{u_{1x}, u_{1y}\}$. A_{11} and A_{12} are the end blocks of the sling with masses m_1 and m_2. The rotation of the sling is clockwise. Object A_2 is the exchange load with a mass Δm moving along line *2* and having the velocity vector $\mathbf{u}_2 \{u_{2x}, u_{2y}\}$. The velocity of object A_2 with respect to object A_1 is $\Delta \mathbf{u}$ with the vector components $(\Delta u_1, \Delta u_2)$. If the circular velocity of sling block A_{12} coincides with $\Delta \mathbf{u}$, contact between object A_2 and end block A_{12} is possible. For this to occur, the sling's center of mass must be at a given point O_1 of the trajectories at the moment of contact, and the phase φ_1 of the sling rotation must also correspond to the calculated one. Launching of the sling to the point of contact, which corresponds to the given conditions, can be realized by controlling the cable length according to options 1 and 2. This control can change the orbit, the sling rotation velocity, and the cable length.

The method of solving the problem of controlling the motion of load A_2 in the zone of contact with the end block is based on the method [3] of controlling the pro-
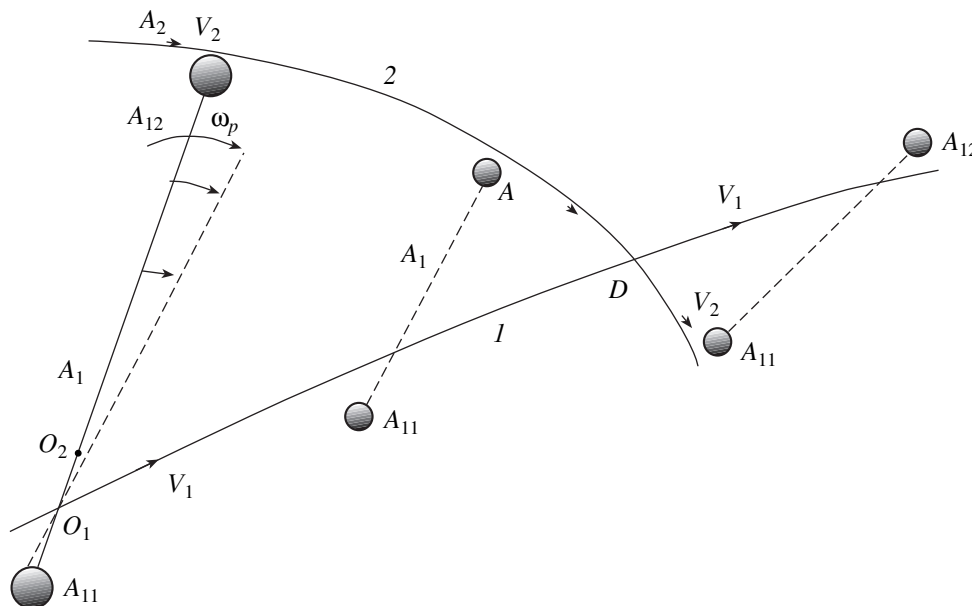

Fig. 2. Scheme of contacting block A_2 with sling A.

cess of landing a free-flying platform onto a landing pad of an orbital station. End block A_{12} moves in the contact zone along a cycloid whose cuspidal point coincides with the contact point.

After completing the contact, the velocity $\mathbf{u}_{1v}$ of the sling's center of mass connected with the object A_2 is

$$\mathbf{u}_{1v} = \mathbf{u}_1 + \frac{\Delta\mathbf{u}\Delta m}{m_1 + m_2 + \Delta m}.$$

The sling's center of mass displaces to the point O_2 in a jumplike manner, as can be seen in Fig. 2. Thus, mating takes place for two bodies, which arrive at the contact point with substantially different velocities. As the contact between object A_2 and end block A_{12} is accomplished with virtually zero velocity, this operation requires no energy. Therefore, the momentum of the unified body is equal to the sum of momenta of A_1 and A_2, excluding the expenditure for the excitation of longitudinal cable vibrations. When object A_2 separates from the sling, the energy carried away by this object depends on the phase of sling rotation at the moment of separation.

The principal scheme of the energy exchange is the following. As an example, we consider a system of two objects A_1 and A_2 which move along their orbits near the Earth. A_1 is a heavy object, such as an orbital station. This object is tied by a cable to end block A_{12}, which rotates around the common center of mass in the orbital plane. Thus, A_1 is a sling. A_2 is a sling energy converter, which moves along an elliptic orbit. We formulate a problem of the transfer of a fraction of the energy from object A_2 to object A_1 such that the masses of objects remain invariable after completing the operation. At a given time, a load G breaks away from end block A_{12}, and then mating of this load with the end block of sling A_2 takes place. Later, in the corresponding rotation phase, the load breaks away from sling A_2 and joins end block A_{12} of sling A_1. The problem parameters can be chosen in such a way that after separation from the sling, the load G carries away a higher energy than that transferred by it at the moment of mating with sling A_2. The energy excess obtained is transferred to sling A_1. When performing this maneuver, the long semiaxis of sling A_2's orbit decreases.

Thus, if sling A_2 has a power source, its energy can be converted into the mechanical energy of the motion of orbital station A_1. In order to attach practical significance to this mechanism, it suffices to have a reasonably powerful energy source. As a prospect, it is possible to consider the fundamental possibility of using the mechanical energy of natural celestial bodies, namely, the moon.

3. TRANSFERRING THE MOON'S MECHANICAL ENERGY TO CABLE SYSTEMS

The principles of solving this problem are the following. The cable group is supplemented by sling A_3 moving along an extended elliptic orbit such that its perigee occurs near the terrestrial surface, and the orbit apogee is approximately equal to the radius of the moon's orbit. The rotation period of A_3 is on the order of 7–8 days. In Fig. 3, we show an example of the calculation for the gravitational maneuver of object A_3. The trajectory of the moon's motion is line L, and that of sling A_3 is line A_3. The trajectories were obtained by directly calculating the set of equations for a plane restricted three-body problem with the corresponding choice of initial conditions. Of course, this example does not reflect the actual complexity of the problem under consideration. At the moment $t = 0$, object A_3 is in orbit perigee [position $A_3(0)$]. The moon's coordinate was denoted by $L(0)$ at $t = 0$. Furthermore, in the process of the motion, the straight lines in Fig. 3 connect the coordinates of the objects at equal moments of time. The time is indicated in hours. The first turn is performed by the sling when the moon moves along a remote portion of the orbit and the sling orbit is close to ellipsis. At the next turn, the sling approaches the moon in the orbit apogee. The sling orbit undergoes a substantial perturbation, and its long semiaxis increases, which leads to an increment in the sling energy. The energy increment is the result of the moon–sling energy
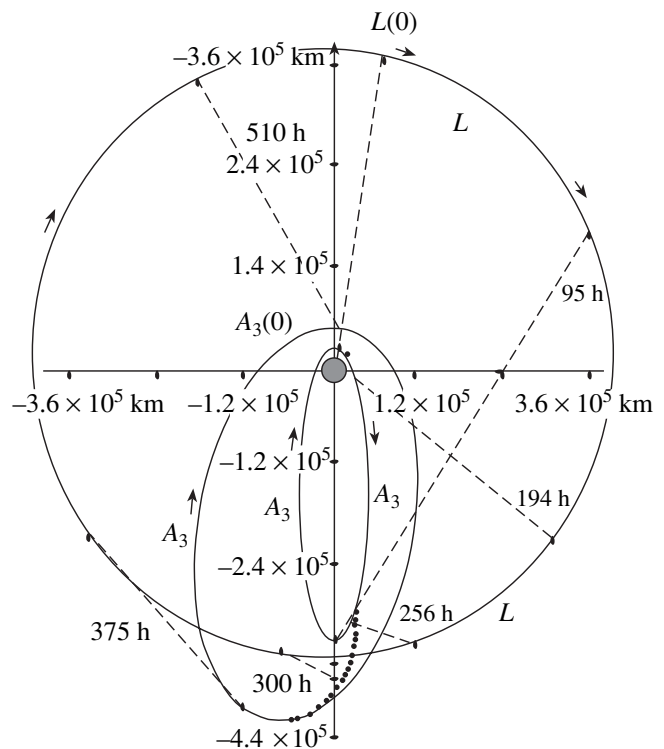

Fig. 3. Example of the gravitational maneuver of sling A_3.

exchange. The energy increment should be transferred to a sling of another group whose orbit is closer to the Earth. It is assumed that after the energy exchange, the orbital parameters of sling A_3 will return to approximately the initial position and the described gravitational maneuver can be repeated. The cycle takes nearly a month. The energy increment is proportional to sling A_3's mass. The energy obtained can be used in the form of mechanical energy to change the parameters of the orbit of space objects. This energy can be also accumulated in the form of electric energy.

REFERENCES

1. I. M. Sidorov, Polet, No. 8, 34 (2000).

2. I. M. Sidorov and V. V. Timofeev, *Multiple-Frequency Oscillations* (Nauka, Moscow, 1984).

3. I. M. Sidorov and V. A. Frolov, in *Proceedings of the ASCE International Conference, Albuquerque, 2000*, p. 1340.

4. V. V. Beletskiĭ, *Essays on Motion of Space Bodies* (Nauka, Moscow, 1977), p. 432.

Translated by V. Bukhanov

============================== **MECHANICS** ==============================

Free Motion of Solid Particles in a Fluid

É. G. Shifrin

In this paper, it is proved that the limiting trajectory of a solid body as small as is wished which is placed in the flow of a viscous incompressible fluid does not coincide with the trajectory of a fluid particle (except for the case of a quasi-solid flow). Therefore, visualizing flows by solid particles (or by their suspensions) is not quite correct.

One method of hydrodynamic visualization consists in marking flows by small solid particles whose density is equal to the fluid density. The method is based on a hypothesis asserting the identity of a fluid-particle trajectory $r(t)$ [i.e., a vector line of the velocity field $V(r, t)$] and the limiting trajectory $\lim_{\delta \to 0} R_\delta(t)$ of a freely moving solid particle of radius δ. For example, visualization by 45-μm solid balls was employed to study the statistical properties of turbulent flows [1]. The method was substantiated by the convergence of the sequence of measured trajectories. However, this implies, at best, only the existence of a limiting particle trajectory but does not prove the hypothesis. Moreover, the following analysis performed in the framework of the mechanics of a viscous incompressible fluid shows that the corresponding identity

$$\forall r_0 \in D_\delta \subset\subset D, \quad R_\delta(0) = r_0 = r(0),$$
$$\lim_{\delta \to 0} R_\delta(t) \equiv r(t) \tag{1}$$

occurs only in the case of a quasi-solid flow.

Let, as in [1], a solid particle be a 3D-ball $\overline{B_\delta(t)}$ of radius δ with the boundary $\Gamma_\delta(t)$. Let the densities of the ball and fluid be equal to unity. We consider a flow in the cylinder $\overline{D_T} = \overline{D \times T}$, where D is a bounded 3D-domain with the boundary $S \in C^{2+2h}, 2h < 1$, and T is the time interval.

The flow corresponds to the solution of the initial boundary value problem A described by the Navier-

_____

Institute for Computer-Aided Design,
Russian Academy of Sciences,
Vtoraya Brestskaya ul. 19/18, Moscow, 123056 Russia

Stokes equations for the 3D-velocity vector $V(r, t)$ and the pressure $p(r, t)$:

$$\operatorname{div} V = 0, \quad V_t + (V \cdot \nabla)V + \nabla p - \mu \Delta V = 0. \tag{2}$$

Let the initial and boundary conditions be given by smooth functions

$$V(r, 0) \in C^{2+2h}(\overline{D}), \quad V(r|_S, t) \in C^{2+2h, 1+h}(\overline{S \times T}).$$

Let the traces of these functions at $t = 0$, $r \in S$ be concordant. Then, there exists the unique solution

$$V(r, t) \in C^{2+2h, 1+h}(\overline{D_T}), \quad \nabla p(r, t) \in C^{2h, h}(\overline{D_T})$$

to the problem A either at $T = T(\mathrm{Re})$ or at any T if the Reynolds number Re defined by the conditions of the problem being discussed is sufficiently small [2, 3]. Below, we assume that the time interval T is defined by one of these conditions.

For simplicity, let the flow motion begin from a quiescent state, i.e.,

$$V(r, 0) = 0, \quad V(r|_S, 0) = \frac{\partial V(r|_S, 0)}{\partial t} = 0.$$

We now place a ball $B_\delta(0)$ of a sufficiently small radius $\delta > 0$ at an arbitrary point of the domain D at the time $t = 0$. Afterwards, the ball begins its free motion under the action of only hydrodynamic forces. We denote by $R_\delta(t)$ the trajectory of the ball's center of inertia. In accordance with the laws of motion, the ball's movement is characterized by both the velocity vector $U_\delta(t) = \dfrac{dR_\delta(t)}{dt}$ of the center of inertia and the angular velocity vector $\Omega_\delta(t)$. In accordance with the accepted simplification, $U_\delta(0) = \Omega_\delta(0) = 0$.

We formulate the initial boundary value problem A_δ for system (2), which describes the flow in the subdomain external with respect to the ball. Apart from the velocity and pressure fields $\{V_\delta^-(r, t), p_\delta^-(r, t)\}$ determined in the bent cylinder $\overline{(D \backslash B_\delta(t)) \times T}$, the solution involves the vector functions $U_\delta(t)$ and $\Omega_\delta(t)$, $t \in \overline{T}$.

The velocity field $V_\delta^+(r, t)$ inside the ball is given by the formulas

$$\forall t \in T, \, r \in B_\delta(t),$$

$$V_\delta^+(r, t) = U_\delta(t) + \Omega_\delta(t) \times (r - R_\delta(t)),$$

$$R_\delta(t) = r_0 + \int_0^t U_\delta(t) dt. \qquad (3)$$

The field $V_\delta^+(r, t)$ for $\Omega = \text{const}$ corresponds to the exact solution to the system (2) for the so-called quasi-solid motion. By virtue of the nonslip condition, we obtain $V_\delta^+(r|_{\Gamma_\delta(t)}, t) = V_\delta^-(r|_{\Gamma_\delta(t)}, t)$, the boundary condition on S remaining the same: $V_\delta^-(r|_S, t) = V(r|_S, t)$.

The functions $R_\delta(t)$, $\Omega_\delta(t)$, $U_\delta(t) = \dfrac{dR_\delta(t)}{dt}$ together with $V_\delta^-(r, t)$, $p_\delta^-(r, t)$ should be determined when solving the problem A_δ from Newton's law and expressions for the hydrodynamic force and moment acting on the ball $B_\delta(t)$:

$$\frac{dU_\delta(t)}{dt} = \frac{F_\delta(t)}{w_\delta}, \quad \frac{d\Omega_\delta(t)}{dt} = \frac{K_\delta(t)}{I_\delta}, \qquad (4)$$

$$F_\delta(t) = \oiint_{\Gamma_\delta(t)} \Pi_\delta^-(r, t) \cdot n ds,$$

$$K_\delta(t) = \oiint_{\Gamma_\delta(t)} [(r - R_\delta(t)) \times \Pi_\delta^-(r, t)] \cdot n ds.$$

The initial conditions are $R_\delta(0) = r_0 \in D_\delta$, $\Omega_\delta(0) = 0$, $U_\delta(0) = 0$, where the distance between the subdomain $D_\delta \subset\subset D$ and S is δ (for sufficiently small δ).

In formulas (4), $\Pi_\delta^-(r, t) = -p_\delta^-(r, t)E + \mu(\nabla V_\delta^-(r, t) + V_\delta^-(r, t)\nabla)$ is the stress tensor, E is the unit tensor, $\mu > 0$ is the viscosity, w_δ is the volume, and I_δ is the moment of inertia of the ball B_δ. The tensor $\Pi_\delta^-(r, t)$ functionally depends on $R_\delta(t)$, $\Omega_\delta(t)$, $U_\delta(t)$.

The problem A_δ differs from the initial boundary value problem with the mobile boundary by the fact that the functions $R_\delta(t)$, $\Omega_\delta(t)$, and $U_\delta(t)$ are given. As was noted in [3], the solvability of the above problem in the class of smooth functions is the same as in the case of the initial boundary value problem with the fixed boundary. The solution to the latter problem continuously depends on initial and boundary conditions [3]. If this behavior is conserved for the boundary value problem with the mobile boundary—i.e., if the solution to this problem, which is determined in $\overline{(D\backslash B_\delta(t)) \times T}$,

continuously depends on $R_\delta(t)$, $\Omega_\delta(t)$, and $U_\delta(t)$—then the problem A_δ has a solution in the same class of smooth functions. Indeed, in this case, the force and the moment acting on the ball $B_\delta(t)$ from the fluid are expressed by continuous functional dependences on the functions $R_\delta(t)$, $\Omega_\delta(t)$, and $U_\delta(t)$. In accordance with the Peano theorem, by virtue of the continuity of the right-hand part of system (4) with respect to these arguments, there exists a solution to Cauchy problem (4), which determines the trajectory $R_\delta(t)$ in the time segment $\overline{T}$. This reasoning allows us to make the following assumption.

Assumption 1. For each reasonably small positive δ, there exists a solution to the problem A_δ

$$V_\delta^-(r, t) \in C^{2+2h, \, 1+h}(\overline{D_T\backslash(B_\delta(t) \times T)}),$$

$$\nabla p_\delta^-(r, t) \in C^{2h, \, h}(\overline{D_T\backslash(B_\delta(t) \times T)}),$$

$$0 < h = h(\delta) < \frac{1}{2},$$

$$R_\delta(t) \in C^2(T), \quad U_\delta(t) = \frac{dR_\delta(t)}{dt} \in C^1(T),$$

$$\Omega_\delta(t) \in C^1(T).$$

We consider that the ball $B_\delta(t)$ follows the flow if it does not disturb the original flow on the entire trajectory; i.e.,

$$\forall t \in T, \quad r \in \overline{D}\backslash B_\delta(t) \quad V_\delta^-(r, t) \equiv V(r, t).$$

It is evident that if the original flow $V(r, t)$ in $\overline{D \times T}$ is quasi-solid, then the ball B_δ of an arbitrary radius follows the flow irrespective of its initial position in D_δ. [In this case, identity (1) is also valid.]

Experimental data on the self convergence of the sequence of trajectories $R_\delta(t)$ [1] make it possible to accept the following assumption.

Assumption 2. We make an assumption that

$$\exists \lim_{\delta \to 0} R_\delta(t) = R_0(t) \in C^2(T).$$

By virtue of this assumption, $\exists V_0^-(r, t)$, $(r, t) \in (\overline{D}\backslash R_0(t)) \times \overline{T}$, so that

$$\forall t \in T, \quad r \in \overline{D}\backslash B_\delta(t), \quad 0 < \forall \delta \ll 1,$$

$$V_\delta^-(r, t) = V_0^-(r, t) + o(\delta).$$

Identity (1) expressing the disappearance of the disturbance produced by the ball as $\delta \longrightarrow 0$ corresponds to the equality

$$V_0^-(r, t) \equiv V(r, t), \quad (r, t) \in (\overline{D}\backslash R_0(t)) \times \overline{T}.$$

We now show that if identity (1) occurs, then the initial flow is quasi-solid. Let

$$\forall t \in T, \quad r \in \overline{D} \backslash B_\delta(t),$$

$$V_\delta^-(r, t) = V(r, t) + o(1), \quad \delta \longrightarrow 0.$$

Then

$$\forall t \in T, \quad r \in \Gamma_\delta(t),$$

$$V(r, t) = V_\delta^-(r, t) + o(1) = V_\delta^+(r, t) + o(1), \quad \delta \longrightarrow 0.$$

By this construction, $V(r, 0) \equiv V_\delta^+(r, 0) \equiv 0$. By virtue of the continuous dependence of the solution to the initial boundary value problem with a mobile boundary on the boundary condition and on $\Gamma_\delta(t)$, the velocity field $V(r, t)$ of the original flow in $\overline{B_\delta(t) \times T}$ for $0 < \delta \ll 1$ is approximately quasi-solid. In other words, within the accuracy to the term $o(1)$, it coincides there with the field $V_\delta^+(r, t)$. By virtue of the arbitrariness of both the ball radius δ and its initial position in B_δ, the approximately quasi-solid character of the initial flow

exists everywhere in $D_\delta \times \overline{T}$. Passing to the limit as $\delta \longrightarrow 0$, we obtain that the original flow should be quasi-solid everywhere in $D \times \overline{T}$.

Thus, in the framework of the assumptions made above, we have proved the following theorem.

Theorem. *In order for the trajectory of a solid ball B_δ freely driven in a flow of a viscous incompressible fluid (of the same mass density) to coincide, as $\delta \longrightarrow 0$, with the trajectory of a fluid particle passing through the same initial point, it is necessary and sufficient for the flow to be quasi-solid.*

REFERENCES

1. A. La Porta, G. A. Voth, A. M. Crawford, *et al.*, Nature **409**, 1017 (2001).

2. V. A. Solonnikov, Proc. Steklov Math. Inst. **73**, 221 (1964).

3. O. A. Ladyzhenskaya, *Mathematical Aspects of Dynamics of Viscous Incompressible Fluid* (Nauka, Moscow, 1970).

Translated by É. Shifrin

===== **MECHANICS** =====

Mesomechanics of Failure
in Cold-Rolled Metal Polycrystals Subjected to Tension

Academician V. E. Panin, V. S. Pleshanov, and S. P. Burkova

Received February 27, 2002

Investigation of the kinetics of crack nucleation and development in a loaded solid is among the most important problems in the mechanics and physics of failure [1, 2]. Traditionally, when interpreting the nature of failure in a loaded material, it is assumed that the onset of failure is associated with a random fluctuation in a sample cross-section region in which the maximum value of the applied stress is attained [3]. From the standpoint of physical mesomechanics, which combines the approaches of continuum mechanics and the physics of plasticity [4–7], this is an erroneous assumption, since the process of failure in solids is strictly deterministic. In mesomechanics, plastic deformation is considered to be a local loss in shear stability of a material at different scale levels, and failure in itself is accounted for by a global loss of shear stability at the macroscopic level. Failure occurs in a localization region of a stress macroconcentrator and is determined by the mechanics of the development of macrobands of localized deformation, where the relaxation processes at the microscale and macroscale levels have an accommodation character. This study is devoted to the experimental substantiation of this fundamentally important aspect.

Deformation and failure of a loaded solid are always associated with the action of maximum shear stresses on a sample from a base stress concentrator (test-machine grip). In plastic materials, this is manifested in the onset of loading and is expressed as propagating surface-defect flows in conjugate directions of the maximum tangent stresses τ_{max} [7, 8]. There is a possibility of increasing the scale level of this effect and of determining the role of deformation accommodation mesobands in failure of polycrystals. To do this, it is necessary to block surface-defect flows using special treatment of the materials. In this case, the defect flows must propagate by developing meso- and macrobands of a localized deformation in the conjugate directions along the vector of τ_{max}. These bands determine the mechanism of deformation and failure of the material. In this

study, the treatment consisted in the preliminary cold rolling of polycrystals.

In our investigations, we used polycrystals of armco iron, high-nitrogen Kh17AG18 steel, and commercial BT-1 titanium as typical materials with body-centered cubic, face-centered cubic, and face-centered close-packed crystal lattices, respectively. In a state cold-hardened by cold rolling, these materials exhibit different degrees of shear stability at the mesoscale level. The initial average grain sizes were 30, 15, and 200 μm for armco iron, high-nitrogen steel, and titanium, respectively. The cold rolling was performed using a laboratory mill at room temperature 293 K, and different compaction degrees were obtained as a result of multiple runs. Samples with dimensions of the working part of $30 \times 4 \times 1$ mm^3 were cut by the electric-erosion method. The experiments were performed in the conditions of static sample tension at a rate of 3×10^{-5} m s^{-1} using an IMASh-20-78 testing machine with an attached TOMSC measuring optical and TV system [4] for investigating the evolution of deformation mesostructures immediately in the loading process.

The curves of plastic flow in polycrystals of armco iron, which were obtained under tension in the initial state and with different degrees of preliminary cold rolling, are shown in Fig. 1. (The curves for high-nitrogen steel and titanium have a similar qualitative character.) Figure 1 illustrates the well-known fact of the instability of the cold-shaped state of metals, which is obtained by rolling followed by subsequent tension [9, 10]. It is this instability under conditions of a high level of external applied stresses that causes the formation of deformation mesobands and development of plastic flow at the mesoscale level. In cold-rolled materials, the high degree of imperfection of the crystal lattice and the well-pronounced texture prevent the occurrence of crystallographic deformation mechanisms at the microscale level. Under these conditions, the local loss in lattice shear stability within the regions of stress microconcentrators does not occur. The level of deforming stresses progressively increases and attains values where the shear stability in the extended regions of a loaded polycrystal is lost, and the stress mesoconcentrators are formed at the onset of plastic flow. They relax by means of noncrystallographic propagation of

Institute of Strength Physics and Materials Science,
Siberian Division, Russian Academy of Sciences,
Akademicheskiĭ pr. 2/1, Tomsk, 634021 Russia

σ, MPa
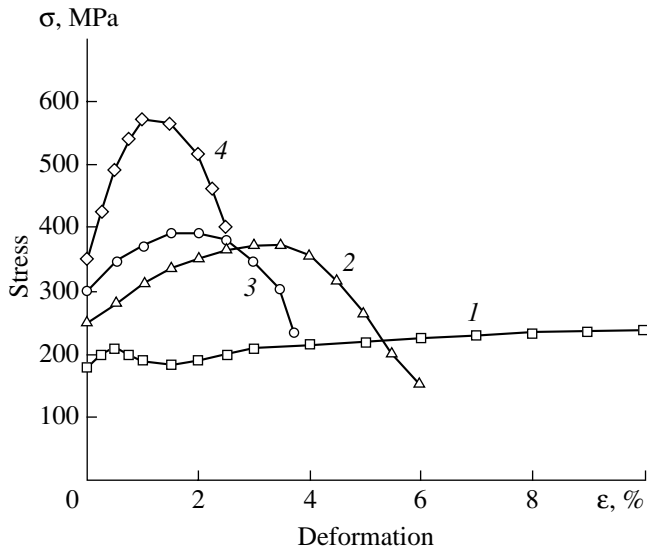
Fig. 1. Curves for flows of armco-iron polycrystals with a different degree of preliminary cold rolling: (*1*) 0, (*2*) 30, (*3*) 50, and (*4*) 80%.

deformation-defect flows through all internal interfaces independently of the crystal structure of the material. In this case, in a polycrystal at the stage of parabolic strengthening, a well-pronounced fragmented mesoband structure forms. Elements of this structure (bands of localized plastic deformation) are oriented in conjugate directions with respect to τ_{max} (Fig. 2). The nature of these bands is related to the displacement of different parts of the crystal with respect to each other. A displacement along one band is inevitably accompanied by a constrained rotation in the sample and by bending of its axis. Since the position of the sample axis is fixed by the testing machine, a counter (induced) stress mesoconcentrator arises at the opposite end of the sample, which generates the displacement and formation of a mesoband in the conjugate direction. The vector sum of these displacements shifts the sample points in the

direction of the acting force, which obeys the laws of continuum mechanics. The dynamics of the mesoband structure in the tension process determines the features of the stress–deformation curves shown in Fig. 1, and the stage nature of plastic deformation at the mesolevel.

Failure of polycrystals represents the final stage in the evolution of the mesostructure. Ultimately, the deformation is concentrated in one of the most developed bands, the so-called superlocalization band. The descending part of the curves in Fig. 1 corresponds to this stage. The stress macroconcentrator formed in a given region of the material generates two adjacent subbands (macroband dipole) within the superlocalization band along which the counterdisplacement of two polycrystal parts occurs. In each half of the sample, rotational deformation modes with opposite sign are developed and, as a result, the sample fails. In this case, the main crack accommodates the difference in material rotations in the two adjacent bands when the possibilities of fragmentation in the material as an accommodation process of the rotational type at the mesoscale level become exhausted. Failure of the material occurs in a quasibrittle manner without the neck formation.

In general, the character of failure is determined by the condition of self-consistency in the failure region for all deformation scale levels. The crack trajectory can be determined by either the superlocalization band (direction of τ_{max}) or the action of normal stresses (the failure occurs perpendicular to the axis of tension). This depends on the degree of development of accommodation processes of plastic flow in adjoining material at a lower scale level. These processes lead to relaxation of the stress macroconcentrator in the superlocalization band. Figure 3 illustrates failure of the materials under investigation.

Failure of the armco-iron polycrystals occurs strictly along the macroband dipole of the localized deformation (Fig. 3a). This is associated with the fact that the development of accommodation processes at lower scale levels is manifested very weakly.
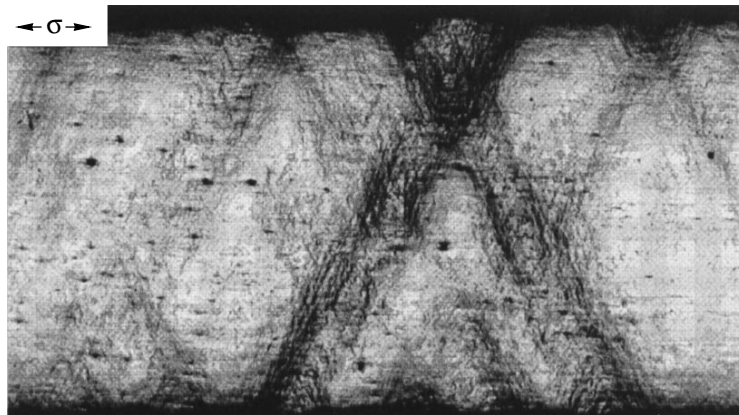
Fig. 2. Formation of the mesoband structure in the process of tension of cold-rolled high-nitrogen steel (the rolling degree is 80%); ε = 0.4%. Magnification is ×10.

(a)

0.5 mm

(b)

0.5 mm

(c)

0.5 mm
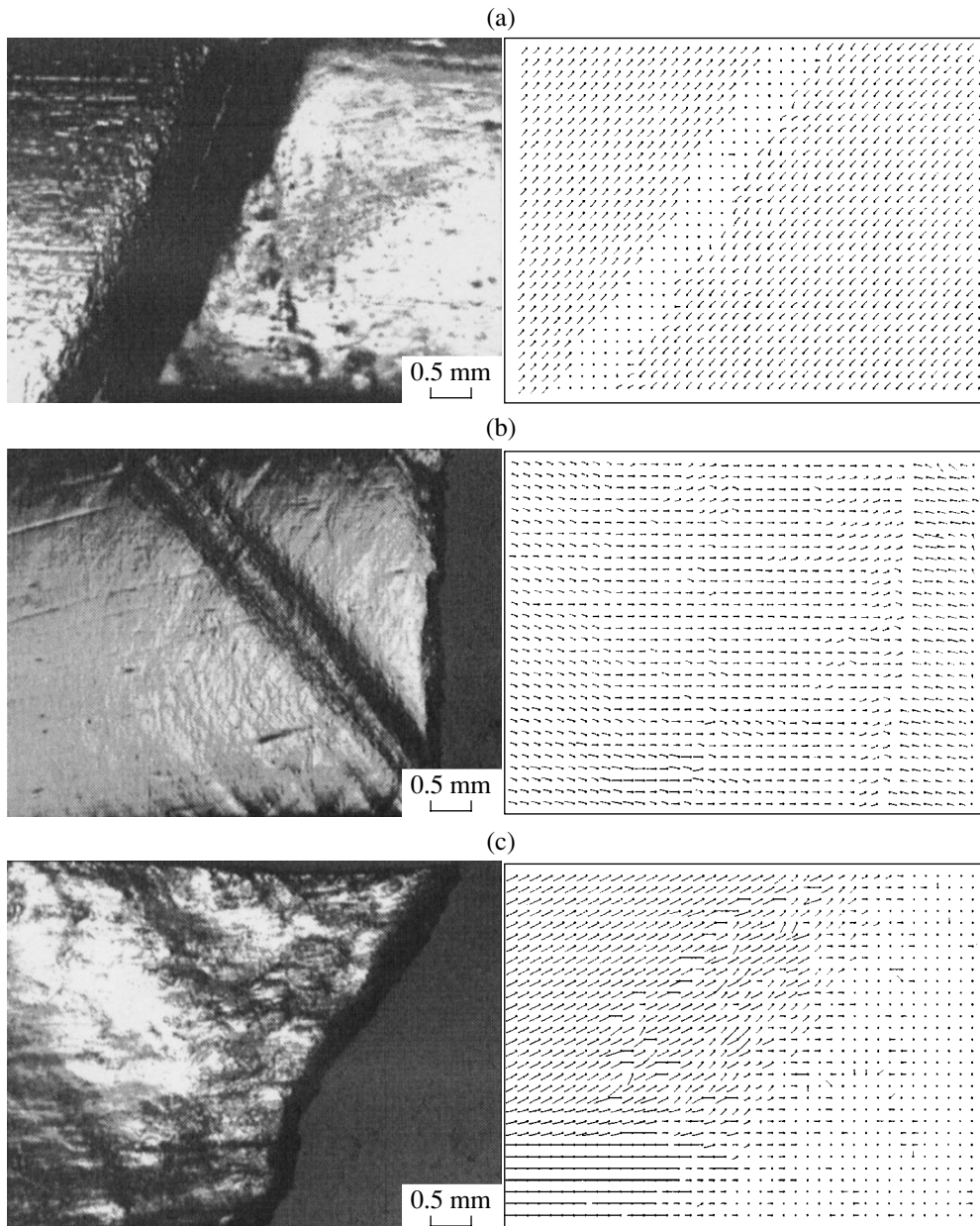
Fig. 3. Character of failure in polycrystals and corresponding displacement-vector fields preceding failure: (a) armco iron, (b) high-nitrogen steel, and (c) titanium.

The character of failure of high-nitrogen steel corresponds to normal separation (Fig. 3b). Although this failure was preceded by the strong development of a localized deformation within the macroband dipole, the main crack propagated across the sample rather than along the macroband. This is caused by the intense development of a system of conjugate bands of accommodation plastic flow at the mesoscale level in the region of failure. The superposition of such conjugate displacements at the mesoscale level determines the development of failure of the normal-separation type.

Failure of Ti polycrystals an the intermediate character. Initially, the main crack propagates along the

macroband dipole and later develops according to the normal-separation scheme. In this case, the accommodation shears at the mesoscale level are developed relatively weakly and are intensified only at the final stage of failure.

In conclusion, it is worth noting that the degree of activity of the plastic-deformation accommodation processes at the mesoscale level in the region of the super-localization macroband determines not only the character of failure but, in many respects, also the macrome-chanical properties of polycrystals. For example, the values of the ultimate strengths obtained in this study for armco iron, titanium, and high-nitrogen steel are

570, 820, and 1500 MPa, respectively, which correlates well with the above failure mechanisms.

Thus, the results obtained convincingly testify to the fundamental importance of the multilevel approach to solving the problems of deformation and failure of solids.

REFERENCES

1. D. Broek, *Principles of Fracture Mechanics* (Noordhoff, Groningen, 1974; Vysshaya Shkola, Moscow, 1980).

2. K. Hallan, *Introduction to Fracture Mechanics* (McGraw-Hill, New York, 1984; Mir, Moscow, 1988).

3. V. I. Vladimirov, *Physical Nature of Metal Failure* (Metallurgiya, Moscow, 1984).

4. V. E. Panin, V. E. Egorushkin, P. V. Makarov, *et al.*, *Physical Mesomechanics and Computer-Aided Design of Materials* (Nauka, Novosibirsk, 1995), Vol. 1.

5. V. E. Panin, Fiz. Mezomekh. **1** (1), 5 (1998).

6. V. E. Panin, Izv. Vyssh. Uchebn. Zaved., Fiz. **41** (1), 7 (1998).

7. V. E. Panin, Fiz. Mezomekh. **4** (3), 5 (2001).

8. A. V. Panin, V. A. Klimenov, N. L. Abramovskaya, and A. A. Son, Fiz. Mezomekh. **3** (1), 83 (2000).

9. A. Korbel and H. Dybiec, Acta Metal. **29**, 89 (1981).

10. A. Korbel and P. Martin, Acta Metal. **36**, 2575 (1988).

Translated by Yu. Vishnyakov