Technical Physics, Vol. 47, No. 4, 2002, pp. 375–379. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 1–5.

THEORETICAL AND MATHEMATICAL PHYSICS

The Reliability of Thermal Analysis of Cross-Country Gas Pipelines

B. B. Kudryashov*, V. S. Litvinenko*, and S. G. Serdyukov**

** St. Petersburg State Mining Institute (Technical University),*
Vtoraya liniya 21, St. Petersburg, 199026 Russia
*** Lentransgaz Ltd.*

Received July 26, 2001; in final form, October 10, 2001

Abstract—In the designs of gas pipelines that must be laid on the bottom of north and south seas, the reliability of analytical results for the gas temperature in the pipelines is questionable. The reason is that different authors take into account various factors influencing the temperature conditions in the pipelines. *© 2002 MAIK “Nauka/Interperiodica”.*

Today’s gas pipeline is a very complex and expensive structure that is designed for long-term trouble-free exploitation. The design and construction of any gas pipeline is based on extensive engineering work that takes into account multiple factors, such as preset capacity; mechanical strength; reliability; stability against technogenic and environmental effects, including natural calamities; and environmental safety. In addition, the pipeline material must withstand mechanical wear and corrosion for a long time, the design must be accident-proof, etc.

The validity of the available design methods, specifically, hydraulic (aerodynamic) methods, is beyond question in most cases, since it has been repeatedly checked during the long-term exploitation of the extensive gas-pipe networks. However, the reliability of the analysis of the gas temperature distribution in pipelines is questionable, because different authors variously take into account factors influencing the temperature conditions in gas pipelines.

Over many years, the temperature distribution in gas and oil pipelines has been estimated with well-known Shukhov’s formula

$$t = t_r + (t_0 - t_r)e^{-\frac{k\pi D}{Gc_p}x}, \tag{1}$$

where t and t_0 are the instantaneous and initial gas temperatures, respectively, °C; t_r is the environmental (rock) temperature, °C; k is the coefficient of heat transfer through the pipeline wall, W/(m^2 °C); D is the inner diameter of the pipeline, m; G is the mass flow rate of the gas, kg/s; c_p is the specific heat of the gas at constant pressure, J/(kg °C); and x is the distance from the beginning of the pipeline.

Currently, gas pipelines many thousands of kilometers long pass through permafrost, cross huge water obstacles, and have large level drops. New pipelines are projected to be laid in deep-water layers of the Barents, Black, and Caspian seas. Therefore, Shukhov’s formula can no longer adequately describe the temperature distribution. The authors of handbook [1] suggest the differential equation

$$\begin{aligned} -k\pi D(t - t_r)dx &= Gc_p dt + Gc_p Dh\frac{p_1 - p_2}{l}dx \\ &\quad + Gg\frac{\Delta z}{l}dx + Gd\left(\frac{v^2}{2}\right), \end{aligned} \tag{2}$$

which, as they argue, includes all the factors influencing the gas temperature in the pipeline. Ignoring the gas velocity variation (the last term on the right-hand side) and integrating Eq. (2), they come to an equation for the gas temperature at any point of the pipeline:

$$\begin{aligned} t &= t_r + (t_0 - t_r)e^{-ax} \\ &\quad - Dh\frac{p_1 - p_2}{l}\frac{1 - e^{-ax}}{a} - g\frac{\Delta z}{l}\frac{1 - e^{-ax}}{c_p a}; \end{aligned} \tag{3}$$

$$a = \frac{k\pi D}{Gc_p},$$

where Dh is the Joule–Thomson coefficient for natural gas, °C/10^5 Pa; p_1 and p_2 are the gas pressures at the beginning and at the end of the pipeline, Pa; Δz is a rise in the level, m; and l is the total length of the pipeline, m.

However, both initial equation (2) and Eq. (3), which is used in calculations, lack the term that takes into account the work of friction of the gas flow. Note that the authors of [1] contradict themselves. In the section devoted to the Joule–Thomson effect upon throttling [1, p. 11], they argue that “…the work spent to overcome obstacles… is converted to heat and is received by the substance…adiabatic throttling of gases, vapors, and liquids is an isoenthalpic process,

1063-7842/02/4704-0375$22.00 © 2002 MAIK “Nauka/Interperiodica”

$h = idem$." On the other hand [1, p. 16], it is argued that "in the absence of external heat exchange, the temperature of the gas being transported changes due to only the Joule–Thomson effect and a change in the position of the center of gravity of the flow."

While analyzing the throttling of natural gas, Russian engineers usually take the Joule–Thomson coefficient in the range Dh = 3–5 K/MPa. A particular value depends on the pressure and temperature. The value of Dh can be found by various techniques according to specific conditions [1–3]. We will find it with the formula for adiabatic expansion, taking the formula for gas adiabatic expansion as the basis:

$$\frac{T_2}{T_1} = \left(\frac{p_2}{p_1}\right)^{\frac{k_0-1}{k_0}}. \qquad (4)$$

The parameters of the improved version of the Blue Flow project [4] are the following: rate of gas delivery simultaneously through two pipelines is 16 billion m^3 pear year; the length and outer diameter of the pipelines are 380 km and 24″, respectively; the thickness of the pipeline wall is 31.8 mm; the pipelines must be laid on the Black Sea bottom at a depth of 2150 m; the methane pressure at the beginning and end of the pipeline is 251.5 and 130.8 bar (abs.); the adiabatic exponent for methane (97.5% in natural gas) is $k_0 = 1.314$; the initial gas temperature at the Beregovaya station is $t = 50°\mathrm{C}$ or $T_1 = 273 + 50 = 323$ K (according to the estimates made by Italian engineers [4]); the compressor discharge pressure is $p_1 = 251.5$ bar; and the pressure at the Turkish coast is $p_2 = 130.8$ bar. Then, the temperature at the end of the pipeline should be

$$T_2 = T_1\left(\frac{p_2}{p_1}\right)^{\frac{k_0-1}{k_0}} = 323\left(\frac{130.8}{251.5}\right)^{\frac{1.314-1}{1.314}} = 276.28\ \mathrm{K}.$$

The temperature drop will be $\Delta T = 323–276.28 = 46.72$ K, and the pressure drop will be $\Delta p = 251.5–130.8 = 120.7\ \mathrm{bar} = 120.7 \times 10^5$ Pa. Accordingly, the Joule–Thomson coefficient in this case is

$$Dh = \frac{\Delta T}{\Delta p} = \frac{46.72}{120.7\times 10^5} = 3.87\ \mathrm{K/MPa},$$

which virtually coincides with the midpoint of the range Dh = 3–5 K/MPa.

We do not know the methods of thermal analysis applied by the Italian engineers in the Blue Flow project. It is noteworthy, however, that with the calculated value Dh = 3.87 K/MPa substituted into Eq. (3), which ignores the work of flow friction, we do obtain the gas temperature at the Turkish coast $t_2 = –10°\mathrm{C}$. That is why the Saipem corporation recommended [4] to set up the last 6.4 km of the double pipeline (near the Turkish coast) on supporting piles (the cost of this project is estimated at $ 78.85 million) or to increase the diameter of one of the pipelines from 24 to 26″ (an extra cost of $ 115.02 million). Both recommendations are intended to prevent negative temperatures that are expected at the end of the pipeline.

"Throttling is a method of expanding a gas at constant enthalpy when it passes through a throttle (that is, through a local resistance, such as an aperture, nozzle, pipe connection, valve, cock, tube constriction, etc.). The expansion, in this case, is attended by a temperature change because the energy is spent to overcome the internal molecular forces of mutual attraction" [2, p. 87]. Upon the adiabatic throttling of oil in the near-bottom region of a well, the oil temperature grows, while the gas temperature drops. This makes it possible to evaluate the collecting properties of a pool and determine the structure of strata [5]. In industrial condensed gas equipment, the throttling effect is used for cooling natural gas in order to separate out water and readily condensing hydrocarbons [1]. In all these and similar cases, formula (4) for the Joule–Thomson coefficient is valid and can be applied for calculations under specific conditions, since it holds only for sudden adiabatic gas expansion and this effect can by no means be "diffused" throughout the many-kilometer pipeline.

Actually, when gas transport through a horizontal pipeline lying deep in a rock is steady, the process is similar to the adiabatic gas flow with friction, because the specific heat of rocks is high and their thermal conductivity is low. During the transport, the heat from the gas-heated pipeline wall returns to the gas and the gas in the pipeline cools down by converting the potential energy of pressure to the kinetic energy of the gas (without allowance for external heat exchange). The degree of cooling is defined by the difference between the initial and final velocities squared. In this case, the gas temperature drop can be found, e.g., from the Aleksandrov formula [6]

$$\Delta T = \frac{k_0-1}{k_0}\frac{v_2^2-v_1^2}{2R}, \qquad (5)$$

where v_1 and v_2 are the cross-section-averaged velocities of compressible gas at the beginning and end of the pipeline, m/s; k_0 is the adiabatic exponent (for methane, $k_0 = 1.314$); and R is the gas constant (for methane, R = 500 J/(kg K)).

For the parameters of the Blue Flow version listed above, $v_1 = 4.27$ m/s and $v_1 = 6.94$ m/s. Then, according to formula (5), the gas temperature drop is less than 0.01°C. This value is convincingly confirmed by experience of many years in drilling wells with air (gas) blow and measuring the temperature distribution in them [7–10]. Early in the development of this technology, it was argued in many, mostly American, publications that blow with gas is in many respects a much more promising approach than irrigation but the temperature at the well bottom drops down to –40°C because of the Joule–Thomson effect. Note that American drilling engineers used formula (4) for sudden adiabatic gas expansion.

Consider a typical situation (Fig. 1). Upon rotor drilling with air blow, the ascending air flow velocity at the well bottom must be no less than v_2 = 20 m/s. This value will be taken in subsequent calculations. Let the pressure in the drill pipe string near the cone drill bit be p_1 = 8 bar and in the annular channel, 4 bar abs. Then, if the cross-sectional area of the inner channel of the drill pipe string is equal to that of the annular channel of the well, the air velocity near the drill bit will be v_1 = 10 m/s. Let the compressed air temperature near the drill bit be t_1 = 20°C or T_1 = 293 K. The adiabatic exponent for air is k_0 = 1.41. Under these conditions, formula (4) for the air temperature at the bottom of the annular channel yields

$$T_2 = 293\left(\frac{4}{8}\right)^{\frac{1.41-1}{1.41}} = 239.5\text{ K} = -33.5°\text{C},$$

and the temperature drop is $\Delta T = T_1 - T_2 = 293–239.5$ = 53.5°C.

However, experience of drilling with blow suggests that this technology may encounter some difficulties associated with the insufficient cooling (overheating) of rock-crushing tools (drill bits). For this reason, drilling by diamond bits with blow has been found to be inappropriate. One of the authors showed [8, 10] that the temperature must be calculated with formula (5), which yields a reduction of the temperature at the well bottom by

$$\Delta T = \frac{1.41-1}{1.41}\,\frac{20^2-10^2}{2\times 287} = 0.15°\text{C}$$

at the same initial parameters and the gas constant for air R = 287 J/(kg K).

Subsequently, it was found that the early bench experiments (February, 1951) carried out by specialists from Hughes Tool Co. [11] gave identical results.

In a 2-ft (≈0.6 m)-deep well of diameter 8 3/4″ (≈220 mm), a new three-extension OSC jet drill bit of diameter 6 1/4″ (≈160 mm) designed for drilling with air blow was installed without rotation (the diameter of each of the extensions was 5/16″ (≈8 mm). Air circulated along the drill pipe string of diameter 2″ (≈50 mm) through the extensions at a flow rate of 350 ft^3/min (≈10 m^3/min) and at a temperature of 77°F (25°C). The resulting pressure drop was 80 psi (≈5.5 bar abs). Under these conditions, the calculation by formula (4) for the adiabatic expansion of air yields the temperature drop ΔT = 116.5 K and the temperature at the drill bit end t = –91.5°C. At the same time, the measurements performed after the steady-state circulation had been established (in 17–30 min) showed that the temperature drop is only 9°F (5°C). However, even this minor drop was explained by the absorption of the heat of vaporization when the air is humidified, entering the well.

The general applicability (and, hence, validity) of Aleksandrov formula (5) can be demonstrated with the sudden adiabatic flow of bottled air. Let the initial parameter values remain the same: the pressure and the temperature of the air inside the bottle are p_1 = 8 bar abs and T_1 = 293 K, respectively, and let the external pressure be p_2 = 4 bar abs. It is known that if the bottle pressure exceeds the external one by a factor of more than 1.86, the air velocity equals the velocity of sound; that is, $v_2 \approx 330$ m/s [12]. (At lower pressure drops, the Saint-Venant principle comes into action.) By definition, the bottle air velocity is v_1 = 0 m/s. Formula (5) yields an air temperature reduction upon flowing out of the bottle

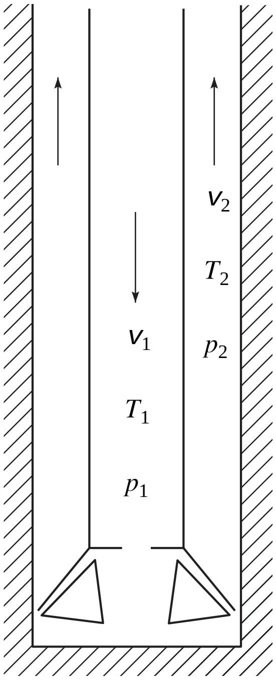


Fig. 1. Schematic representation of air blow of the well bottom.

$$\Delta T = \frac{1.41-1}{1.41}\,\frac{330^2-0^2}{2\times 287} = 55.2°\text{C}.$$

Then, the temperature at the exit from the bottle is $T_2 = T_1 - \Delta T = 293 - 55.2 = 237.8$ K = –35.2°C, which almost equals that predicted from formula (4) for sudden adiabatic expansion.

In our opinion, if starting equation (2) would take into account heat release due to gas flow friction (instead of the unjustified inclusion of the Joule–Thomson effect) and would not ignore the gas velocity variation (the fourth term on the right-hand side), it would yield the same cooling in the pipeline as the Aleksandrov formula (provided that external heat exchange is absent) because cooling may be associated only with the potential-to-kinetic energy conversion under these conditions. At the same time, a change in the kinetic energy can also be neglected, since this change is very

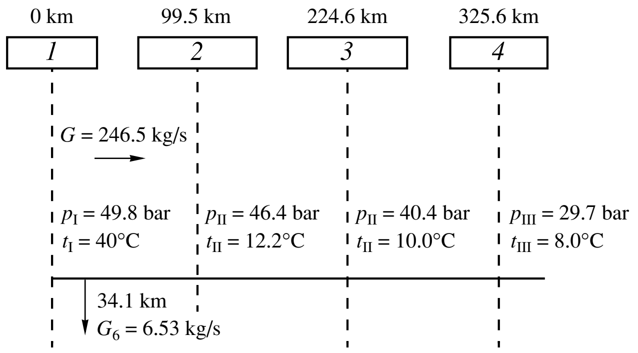


Fig. 2. Segment of the Torzhok–Minsk–Ivatsevichi 3 pipeline (3rd branch) between the Rzhev and Orsha CSs (*1*, Rzhev CS; *2*, Kh. Zhirki CS; *3*, Smolensk CS; and *4*, Orsha CS). The daily average absolute pressure and temperature for the gas being transported are indicated. Subscripts I, II, and III refer to the initial (output), intermediate, and final (input) values, respectively. Note: 1 bar = 10^5 Pa.

small; hence, the gas cooling is also very small (on the order of several hundredths of a degree).

Let us compare the calculated values of the effects considered with the daily average (October 22, 2000) temperature of the gas transported through the Torzhok–Minsk–Ivatsevichi 3 pipeline on the segment between the compressor stations (CSs) in Rzhev and Orsha. This 325.6-km-long segment is schematically depicted in Fig. 2, where the subscripts I, II, and III refer to initial (output), intermediate, and final (input) gas temperatures and pressures, respectively. The data were provided by Lentransgas Ltd.

Ignoring a level drop ($\Delta z = 0$), we supplement the right-hand side of expression (3) by the term allowing for the mechanical work of friction of the gas flow. Then, the temperature distribution along the pipeline is given by

$$t = t_r + \frac{Ggi}{k\pi D} - Dh\frac{\Delta p}{l}\frac{Gc_p}{k\pi D} + \left(t_0 - t_r - \frac{Ggi}{k\pi D} + Dh\frac{\Delta p}{l}\frac{Gc_p}{k\pi D}\right)e^{-\frac{k\pi D}{G v_p}}. \qquad (6)$$

Here, g is the gravitational acceleration ($g = 9.81$ m/s^2) and i is the dimensionless hydraulic gradient, which is found from the expression

$$i = \frac{\Delta p}{g\rho_{\mathrm{av}}l}, \qquad (7)$$

where Δp is the pressure drop in the pipeline ($\Delta p = p_{\mathrm{in}} - p_{\mathrm{f}}$), Pa, and ρ_{av} is the average gas density in the pipeline, kg/m^3.

Comparison of the measured and analytical values of the natural gas temperature along the path Rzhev CS–Orsha CS, °C

Compressor stations	Rzhev	Kh. Zhirki	Smolensk	Orsha
Measurements	40	12.20	10.00	8.00
Calculated data ($i = 0.021$, $Dh = 0$)	40	9.83	8.18	8.14
Calculated data ($Dh = 0.4$°C/bar, $i = 0$)	40	6.06	4.21	4.16
Data calculated by Shukhov's formula	40	6.86	5.05	5.00

In the calculations, we used the actual mass flow rate of the gas $G = 246.5$ kg/s and the inner diameter of the pipeline $D = 1.195$ m. The gas being transported is methane (accounts for 97.5% in natural gas), for which $R = 500$ J/(kg K). The overcompression factor is $z = 0.91$, and the specific heat is $c_p = 2220$ J/(kg °C) (at an average temperature $t = 20$°C). The gas consumption at a rate $G_0 = 6.53$ kg/s between the compressor stations at Rzhev and Kh. Zhirki is neglected. The rock temperature around the pipeline is taken to be constant, $t_r = 5$°C at the end of October, throughout the segment considered. Based on tentative estimates for external heat exchange, the coefficient of heat transfer from the gas to the rock is set equal to $k = 4.3$ W/(m^2 °C).

The calculations were performed by formula (6) for two variants: (1) with regard for the mechanical work of friction of the gas flow ($i = 0.021$) and without taking into account the Joule–Thomson effect ($Dh = 0$) and (2) in view of the Joule–Thomson effect ($Dh = 0.4$ °C/10^5 Pa) and without taking into account the work of friction ($i = 0$). In both cases, external heat exchange was equally taken into consideration [$k = 4.3$ W/(m^2 °C)].

The results of the calculations are summarized in the table, where the data obtained by Shukhov's formula (1) for the same initial conditions are shown for comparison.

Both analytical distributions (with external heat exchange equally taken into account) tend to the constant temperature value faster than the actual (measured) gas temperature distribution along the pipeline. This can be explained, for example, by the variable rock temperature along the 325-km-long pipeline. However, the agreement between the measured and analytical data is of minor concern for us. It is interesting to contrast the calculated gas temperature distributions in the pipeline when the work of friction is taken into account and the Joule–Thomson effect is not, and vice versa. In the former case, the gas temperature at the final point is nearly coincident with its average value on October 22, 2000. In the latter case, the gas temperature calculated for the last two compressor stations (without pumping-over) turned out to be lower than the rock temperature. This can be explained only by incorrectly taking into account the Joule–Thomson effect especially when the work of friction is neglected.

The gas velocities calculated at the beginning and at the end of the segment under consideration were found to be $v_1 = 3.34$ m/s and $v_2 = 5.61$ m/s for the actual temperatures and pressures. The temperature drop due to the potential-to-kinetic energy conversion was as low as $\Delta t = 5 \times 10^{-3}$ °C, as estimated by formula (5).

Note that the associated calculation made by Shukhov's formula (1) for the same initial conditions gave the temperature value lying between those obtained in the above two variants, since it includes neither the mechanical work of friction nor the Joule–Thomson effect. At the end of the 325-km-long pipeline (the Orsha CS), the gas temperature found by Shukhov's formula is equal to the rock temperature. However, the structure of this formula is such that it cannot yield a lower temperature. For many decades, this fact satisfied theorists and engineers engaged in the gas extraction industry. Formula (6) for the first variant ($Dh = 0$) also cannot yield temperatures below the rock temperature, since it takes into account the work of friction of the gas flow. Nearly the same results were obtained for a number of segments in the pipeline network exploited by Lentransgas Ltd.

In the history of exploitation of pipelines passing through the Arctic permafrost and through unfrozen rocks in the European part of Russia and in Western Europe up to Greece and Portugal, engineers from Lentransgas Ltd. have never encountered any problems associated with rock freezing, frost-induced rock heaving, etc., except for a drastic drop in the gas pressure at the distributing stations of gas-transporting networks [13].

In summary, our conclusions, as applied to the calculations within the Blue Flow project, are the following [14]. For the gas pipeline passing on the Black Sea bottom from Russia to Turkey at a minimal water temperature $t = +9$°C, its temperature at the Turkish coast by no means can be negative. Thus, there is no reason for extensive discussion [4] on the danger of mechanical damage to the pipeline at the Turkish coast because of rock freezing and *a fortiori* for the "scientific" justification of extremely questionable proposals on "improving" the initial Blue Flow project, which costs $ 79–115 million.

On November 23, 2000, the coordination meeting on gas transport from the Shtokman condensed gas deposit was held at the St. Petersburg State Mining Institute. Specialists from research institutes and universities, as well as the administration and the leading engineers from the Gasprom joint-stock venture, were present at the meeting (a total of 31 persons). Various and often mutually exclusive ideas concerning whether or not the mechanical work of friction and the Joule–Thomson effect should be taken into account in the analysis of the temperature distribution along the pipeline were put forward. It was even suggested to use the Navier–Stokes equations to solve this relatively simple problem. We greatly appreciate the intimate knowledge of thermodynamics and higher mathematics the participants demonstrated. Yet, some fundamental physical laws should be remembered. One of them states that the giant compressor energy spent on gas transfer over several hundreds of kilometers at a rate of several millions of cubic meters per day is eventually dissipated as heat, and this fact cannot be ignored.

All the above reasoning seems to be valid even without regard for the time factor, which has never been considered in the design and analysis of the pipelines. For the recently built pipelines, however, this factor is of great importance and will be considered in subsequent publications.

The authors admit that the reasoning and calculations reported in this article may make little difference for some of those taking part in the Blue Flow project. Anyway, they make anybody think and try to solve the problem touched upon analytically and experimentally.

REFERENCES

1. B. P. Parshakov, R. N. Bikchentaĭ, and B. A. Romanov, *Thermodynamics and Heat Transfer (in Technological Processes of Oil and Gas Industry): Textbook for Institutes of Higher Education* (Nedra, Moscow, 1987).
2. A. I. Gritsenko, Z. S. Aliev, O. M. Ermilov, *et al.*, *Manual on Study of Wells* (Nauka, Moscow, 1995).
3. *Tables of Physical Quantities: Handbook*, Ed. by I. K. Kikoin (Atomizdat, Moscow, 1976).
4. Saipem. Gazprom-SNAM. Project "Blue Stream": Business and Technical Application, June 30, 1999 (with supplements).
5. Sh. K. Gimatudinov and A. I. Shirkovskiĭ, *Physics of Oil and Gas Seam* (Nedra, Moscow, 1982).
6. V. L. Aleksandrov, *Technical Hydromechanics* (Gostekhizdat, Moscow, 1946).
7. B. B. Kudryashov, Zap. Leningr. Gos. Inst. im. G. V. Plekhanova **62** (2), 43 (1969).
8. B. B. Kudryashov and A. I. Kirsanov, *Boring of Exploratory Wells with Air* (Nedra, Moscow, 1990).
9. B. B. Kudryashov, V. K. Chistyakov, and V. S. Litvinenko, *Well-Boring under Aggregative State Change of Rocks* (Nedra, Leningrad, 1991).
10. B. B. Kudryashov and A. M. Yakovlev, *Drilling in the Permafrost* (Oxonion Press, New Delhi, 1990).
11. Laboratory Report (Hughes Tool Company, 1951).
12. B. B. Kudryashov, *Prom. Energ.*, No. 3, 14 (1962).
13. S. G. Serdyukov, Author's Candidate's Dissertation (St. Petersburg, 1998).
14. B. B. Kudryashov and A. V. Kozlov, *Preliminary Report on Reliability Estimate of Engineering Solutions on Project "Blue Stream"* (St. Petersburg, 2000).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 4, 2002, pp. 380–384. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 6–10.
Original Russian Text Copyright © 2002 by Muratov.

THEORETICAL AND MATHEMATICAL PHYSICS

On the Frenkel Problem of Equivalent Steady Currents in a Sphere

R. Z. Muratov
Moscow State Mining University, Moscow, 117935 Russia

Received July 16, 2001

Abstract—The Frenkel problem of substituting the 3D system of steady currents given in one of two concentric spherical regions by an equivalent system of currents (i.e., by that inducing the same external magnetic field) that is distributed over the surface of the other region is considered. A method of multipole moments providing the direct solution (without calculating the fields) of the problem is described. The case of currents with the density components represented by cubic polynomials of the Cartesian coordinates is considered as an example. *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Dissimilar systems of sources are referred to as equivalent (or adequate) if they, being located inside a finite space, induce the same fields outside the space. These may be equivalent distributions of gravitating masses, electric charges, and, finally, steady electric currents. The characteristic feature of the last-named case is the vector nature of the sources.

It is well known that a charged sphere with a uniform volume charge density and a concentric sphere of arbitrary radius with a uniform surface charge density induce the same external electrostatic fields coinciding with that of a point charge located at the center of the spheres if the total charge is equal in all the three cases. These are examples of scalar equivalent sources. For a magnetostatic field, such a simple example is absent since there are no magnetic charges in the nature.

The problem of finding the source distributions equivalent to a given distribution is not new. In the case of scalar sources (charges), the theoretical statement of the problem and the discussion of general results were given in [1, p. 103; 2, p. 524]. It is worth noting that in the general statement, the problem of equivalent sources has an infinite set of solutions, as also follows from the example given above. The statement of practical interest is that in which the distribution of sources is given in a volume *V* (or on a closed surface *S*) and it is required to find the equivalent distribution of sources on another surface *S*. This provides the uniqueness of the solution. The method for solving such scalar problems in the case of ellipsoidal spatial regions was reported in [3, 4]. As for the vector problem of equivalent steady currents, the author has not encountered an example of its consideration. However, the solution of such a problem is useful for refining the geomagnetic dynamo mechanism, estimating currents induced on superconducting shields, and tackling some other issues of practical magnetostatics.

In this paper, we propose a solution to the simple problem of equivalent steady currents where the "original" currents are specified in the volume of a sphere of radius *a* and the equivalent currents are looked for on a concentric sphere of radius *R*. The problem is considered in Cartesian (rather than in spherical) coordinates in order to subsequently extend the approach used to geometrically more general problems. The solution of the problem, with the Cartesian components of the original current density represented by polynomial functions of the Cartesian coordinates, is obtained by the method of multipole moments. As in the case of scalar sources, this method makes it possible to avoid the cumbersome procedure of calculating the fields.

MULTIPOLE MOMENTS METHOD

As is well known, in a simply connected spatial region that is an outer region with respect to steady electric currents, one can deal with the pseudoscalar potential of the induced magnetic field (hereafter, magnetic potential) and its multipole expansion [5]

$$\Phi(\mathbf{r}) = \sum_{l=1}^{\infty} \frac{(-2)^l}{(2l)!} m_{i_1\ldots i_l} \nabla_{i_1} \ldots \nabla_{i_l} \frac{1}{r}. \qquad (1)$$

The quantities $m_{i_1\ldots i_l}$ appearing in (1) are the components of the lth-rank tensor, which constitute the magnetic multipole moment of the lth order. In what follows, we will distinguish the magnetic multipoles $m^{(\mathbf{j})}_{i_1\ldots i_l}$ of the bulk currents with the density $\mathbf{j}(\mathbf{r})$ and the magnetic multipoles $m^{(\mathbf{i})}_{i_1\ldots i_l}$ of the surface currents with the surface density $\mathbf{i}(\mathbf{r})$. These multipole moments are given by

$$m^{(\mathbf{j})}_{i_1\ldots i_l} = \frac{1}{(l+1)c} \int [\mathbf{r}\mathbf{j}] \nabla_{i_1\ldots i_l} dV, \qquad (2)$$

1063-7842/02/4704-0380$22.00 © 2002 MAIK "Nauka/Interperiodica"

$$m^{(\mathbf{i})}_{i_1\ldots i_l} = \frac{1}{(l+1)c}\oint[\mathbf{ri}]\nabla_{i_1\ldots i_l}dS. \qquad (3)$$

Here, the bracketed vectors mean the vector product.

Expressions (2) and (3) involve the irreducible symmetrical tensor $d_{i_1\ldots i_l}(\mathbf{r})$, which is defined as the product of the components of the vector operator [6]

$$\hat{\mathbf{D}} = 2\mathbf{r}(\mathbf{r}\nabla) - r^2\nabla + \mathbf{r} \qquad (4)$$

(the vectors in parentheses mean the scalar product). The product operates on unity; i.e.,

$$d_{i_1\ldots i_l}(\mathbf{r}) = \hat{D}_{i_i}\ldots\hat{D}_{i_l}\cdot 1. \qquad (5)$$

In particular, $d_x = x$, $d_{xx} = 2x^2 - y^2 - z^2$, $d_{xy} = 3xy$, $d_{xyy} = 3x(4y^2 - x^2 - z^2)$, and $d_{xyz} = 15xyz$, so that

$$m_x = \frac{1}{2c}\int(yj_z - zj_y)dV; \qquad (6)$$

$$m_{xx} = \frac{2}{c}\int x(yj_z - zj_y)dV, \qquad (7)$$

$$m_{xy} = \frac{1}{c}\int\{xzj_x - yzj_y + (y^2 - x^2)j_z\}dV; \qquad (8)$$

$$\begin{aligned} m_{xyy} = \frac{3}{4c}\int\{&10xyzj_x + (x^2 + z^2 - 4y^2)zj_y \\ &+ (4y^2 - 11x^2 - z^2)yj_z\}dV, \end{aligned} \qquad (9)$$

$$\begin{aligned} m_{xyz} = \frac{15}{4c}\int\{&(z^2 - y^2)xj_x + (x^2 - z^2)yj_y \\ &+ (y^2 - x^2)zj_z\}dV.[1] \end{aligned} \qquad (10)$$

Note that, owing to the special “normalization” [5] of the magnetic moments, expansion (1) is identical to the multipole expansion of the electrostatic potential with only one exception: in sum (1), the term corresponding to $l = 0$ (magnetic charge) is absent.

For our purposes, two well-known properties of expansion (1) are essential. First, this expansion is universal, i.e., is applicable to arbitrary distributions of currents if they occupy a finite spatial region. In particular, the expansion is independent on what kind of currents (bulk or surface) induces the magnetic field. Hence, if different systems of currents have the same multipole moments, these systems are adequate. Formally, this implies the coincidence of the infinite number of the magnetic multipole components, for series (1) is infinite.

The second property of the multipole expansion allows one to use the above universality just in the geometric configuration considered. Indeed, if a region filled with currents is represented by either a sphere or its surface and the current density components (bulk or surface) are polynomial functions (of degree $L > 0$) of Cartesian coordinates, sum (1) involves a finite number of terms, since all magnetic multipoles of the $(L + 1)$th or higher orders vanish. Here, we assume that the origin is placed at the center of the sphere.

Now we turn to the general statement of the problem and the method to solve it. Assume that an electric current with the Cartesian density components represented by L-degree polynomials

$$\begin{aligned} j_x &= \sum_{1\le l+m+n\le L} A_{lmn}x^l y^m z^n, \\ j_y &= \sum_{1\le l+m+n\le L} B_{lmn}x^l y^m z^n, \\ j_z &= \sum_{1\le l+m+n\le L} C_{lmn}x^l y^m z^n \end{aligned} \qquad (11)$$

is specified in the volume of a sphere with radius a.[2]

In (11), the coefficients A, B, and C of the polynomials given are related by expressions following from the current steadiness condition

$$\operatorname{div}\mathbf{j} = 0 \qquad (12)$$

and the boundary condition on the sphere surface

$$j_n|_S = 0. \qquad (13)$$

It is required to find the current on the surface of a concentric sphere of the radius R that is adequate to (11). The pattern of calculation is as follows. First, we calculate all the independent components of the magnetic multipoles $m^{(\mathbf{j})}_{i_1\ldots i_l}$ whose rank is in the range $1 \le l \le L$. The total number of such components is

$$N_L^{(M)} = \sum_{l=1}^{L}(2l+1) = L(L+2). \qquad (14)$$

As it was mentioned, any Cartesian component of the equivalent surface current has also to be an L-degree polynomial. Recall that the degree of a polynomial on the surface of a sphere

$$x^2 + y^2 + z^2 = R^2 \qquad (15)$$

is that of the polynomial resulting from the initial one by eliminating [in view of (15)] the even powers of any Cartesian coordinate (e.g., z) and collecting the similar terms. These polynomials may be represented in various equivalent forms (see [3] for details). Hereafter, we will write them in the symmetric (keeping the equiva-

---

[1] Hereafter, in the case of several relations (equations) that differ from each other by either the circular permutation or interchanging of the coordinates, we write only one relation (equation). It is good to bear in mind that, in the case of interchanging, the pseudovariables change their sign.

[2] In (11), we have taken into account that the coefficients before the summands with $x^0y^0z^0$ are equal to zero.

lence of the Cartesian directions) form:

$$i_x = \sum_{l=L}^{L} \sum_{i+j+k=l} \alpha_{ijk} x^i y^j z^k,$$

$$j_y = \sum_{l=L}^{L} \sum_{i+j+k=l} \beta_{ijk} x^i y^j z^k, \tag{16}$$

$$j_z = \sum_{l=L}^{L} \sum_{i+j+k=l} \gamma_{ijk} x^i y^j z^k.$$

Thus, each of the components of the surface current is the sum of homogeneous polynomials of the $(L-1)$th and Lth degrees with the coefficients to be found. Since the number of the coefficients $N_L^{(H)}$ of a homogeneous L-degree polynomial is

$$N_L^{(H)} = \frac{1}{2}(L+1)(L+2),$$

the number of the coefficients in each of the polynomials in (16) is

$$N^{(L)} = N_L^{(H)} + N_{L-1}^{(H)} = (L+1)^2. \tag{17}$$

The number of unknowns in our problem is defined by the total number of coefficients in all three polynomials in (16):

$$N = 3N^{(L)} = 3(L+1)^2. \tag{18}$$

In particular, the number of unknowns is 48 for $L = 3$, whereas, in the similar problem with scalar sources, the number of unknowns is three times smaller.

Let us calculate the number of available linear algebraic equations needed to find the coefficients of the polynomials in (16). First of all, these are the inhomogeneous equations arising as a result of equating the corresponding independent components of the magnetic multipole tensors (of rank from one to L) due to currents (11) and (16). Since any symmetric lth-rank irreducible tensor, such as that of the magnetic multipole of the same rank, has $2l + 1$ independent components, the number of the resulting inhomogeneous equations turns out to be

$$N_L^{(M)} = \sum_{l=1}^{L} (2l+1) = (L+1)^2 - 1. \tag{19}$$

The requirement that the normal component of the surface current density vanish on the surface of sphere (15),

$$i_n|_{\bar{S}} = (\mathbf{ri})|_{\bar{S}} = xi_x + yi_y + zi_z = 0, \tag{20}$$

increases the degree of the polynomials describing the current by unity, as is seen from (20). Consequently, this requirement adds to (19)

$$N_{\bar{S}} = N^{(L=1)} = (L+2)^2 \tag{21}$$

homogeneous equations.

The condition of surface current steadiness

$$\mathrm{div}\,\mathbf{i} = 0 \tag{22}$$

decreases the polynomial degree by unity, adding

$$N_{\mathrm{div}} = N^{(L-1)} = L^2 \tag{23}$$

more homogeneous equations for the coefficients.

Thus, the number of equations $N_L^{(M)} + N_{\bar{S}} + N_{\mathrm{div}}$ always exceeds the number of unknowns N by unity. It can be shown that one of the homogeneous equations is a linear combination of the rest.

EXAMPLE OF ADEQUATE CURRENTS

In order to illustrate the method, we consider the case when the components of currents (11) and (16) are cubic polynomials. As was noted, this case generally involves 48 unknown coefficients of polynomials (16). Fortunately, as in the case of scalar sources [3], the calculations are substantially simplified, since the resulting system of 48 linear algebraic equations splits into eight independent subsystems. Moreover, as in the case of scalar sources [3], it is sufficient to find the analytical solutions of four subsystems only because of the symmetry of the problem (providing the equivalence of the Cartesian directions). The remaining solutions may be obtained by circular permutations in those found previously.

According to the aforesaid, instead of considering all fifteen independent components of the magnetic multipoles $\mathbf{m}$, m_{ij}, and m_{ijk}, we can restrict our analysis to, e.g., the following ones: (1) m_x, m_{xyy}, and m_{xzz}; (2) m_{xx} and m_{yy}; (3) m_{xy}; and (4) m_{xyz}. These components are split into four groups, each having an independent subsystem of equations for the polynomial coefficients.

Equating the components $m_x^{(\mathbf{i})}$, $m_{xyy}^{(\mathbf{i})}$, and $m_{xzz}^{(\mathbf{i})}$ calculated by formulas (6) and (9)[3] to the corresponding components of the bulk currents $m_x^{(\mathbf{j})}$, $m_{xyy}^{(\mathbf{j})}$, and $m_{xzz}^{(\mathbf{j})}$ and completing the set of inhomogeneous equations obtained by necessary homogeneous equations resulting from (20) and (22), we find the following closed subsystem of seven equations (see Footnote 2):

$$3\gamma_{030} + \gamma_{210} + \gamma_{012} - 3\beta_{003} - \beta_{021} - \beta_{201} = \frac{15c}{2\pi R^6} m_x^{(\mathbf{j})},$$

---

[3] With $\mathbf{j}dV$ replaced by $\mathbf{i}dS$.

$$\begin{gathered} 5\alpha_{111} + 3\beta_{003} - 4\beta_{021} + \beta_{201} \\ + 12\gamma_{030} - 11\gamma_{210} - \gamma_{012} = \frac{35c}{2\pi R^8} m^{(\mathbf{j})}_{xyy}, \\ \alpha_{111} + 2\beta_{021} + 2\gamma_{012} = 0, \quad \alpha_{111} + \beta_{201} + \gamma_{210} = 0, \\ \beta_{003} + \gamma_{012} = 0, \end{gathered}$$

where

$$m^{(\mathbf{j})}_{x} = \frac{2\pi}{15c} a^5 \Big\{ C_{010} - B_{001} - \frac{1}{7} a^2 \times (B_{201} - C_{210} + 4B_{021} - 4C_{012}) \Big\}, \tag{24}$$

$$m^{(\mathbf{j})}_{xyy} = \frac{8\pi}{315c} a^9 (4B_{201} - 14B_{021} + C_{210} - 11C_{012}). \tag{25}$$

Expressions (24) and (25) are the result of integration by formulas (6) and (9) and the subsequent simplification in view of the relationships

$$\begin{gathered} A_{111} + 2B_{021} + 2C_{012} = 0, \\ A_{111} + B_{201} - B_{003} + C_{201} - C_{012} = 0, \\ B_{021} - B_{003} + C_{030} - C_{012} = 0, \\ B_{001} + a^2 B_{003} + C_{010} + a^2 C_{012} = 0 \end{gathered}$$

between the coefficients A, B, and C, which follow from (12) and (13).

The solution of the first subsystem of the equations (for the unknowns α_{111}, β_{201}, β_{021}, β_{003}, γ_{210}, γ_{030}, and γ_{012}) has the form

$$\alpha_{111} = \frac{7c}{12\pi R^8} (m^{(\mathbf{j})}_{xyy} - m^{(\mathbf{j})}_{xzz}), \tag{26}$$

$$\gamma_{030} = -\beta_{021} = \frac{c}{4\pi R^6} \left(3m^{(\mathbf{j})}_{x} + \frac{7}{6R^2} m^{(\mathbf{j})}_{xyy} \right), \tag{27}$$

$$\gamma_{210} = \frac{c}{4\pi R^6} \left(3m^{(\mathbf{j})}_{x} + \frac{7}{2R^2} m^{(\mathbf{j})}_{xyy} - \frac{7}{6R^2} m^{(\mathbf{j})}_{xzz} \right). \tag{28}$$

Similarly, equating the corresponding quadrupole components $m^{(\mathbf{i})}_{xx}$ and $m^{(\mathbf{i})}_{yy}$ found by (7) to $m^{(\mathbf{j})}_{xx}$ and $m^{(\mathbf{j})}_{yy}$ in view of (20) and (22) yields three equations forming the second subsystem:

$$\gamma_{110} - \beta_{101} = \frac{15c}{8\pi R^6} m^{(\mathbf{j})}_{xx}, \quad \langle \alpha_{011} \rangle = 0,$$

where, according to (7),

$$m^{(\mathbf{j})}_{xx} = \frac{8\pi}{105c} a^7 (C_{110} - B_{101}), \tag{29}$$

and $\langle A_{011} \rangle = 0$. Hereafter, the angular brackets $\langle \ldots \rangle$ mean the sum of three terms of the circular permutation.

The solution of the second subsystem of equations (for the unknowns α_{011}, β_{101}, and γ_{110}) has the form

$$\alpha_{011} = \frac{5c}{8\pi R^6} (m^{(\mathbf{j})}_{xx} + 2m^{(\mathbf{j})}_{yy}) = \frac{a^7}{7R^6} A_{011}. \tag{30}$$

Equating the quadrupole moments $m^{(\mathbf{i})}_{xy}$ to $m^{(\mathbf{j})}_{xy}$ gives the value of the coefficient γ_{002} and leads us to the third subsystem of four equations for the remaining unknown coefficients α_{101}, β_{011}, γ_{200}, and γ_{020}:

$$\begin{gathered} \alpha_{101} - \beta_{011} - 2\gamma_{200} + 2\gamma_{020} = \frac{15c}{4\pi R^6} m^{(\mathbf{j})}_{xy}, \\ \alpha_{101} + \gamma_{200} = 0, \quad \alpha_{101} + \beta_{011} = 0, \end{gathered}$$

where

$$m^{(\mathbf{j})}_{xy} = \frac{8\pi}{35c} a^7 A_{101}, \tag{31}$$

and $C_{002} = 0$, $A_{101} = -B_{011} = -C_{200} = C_{020}$. The unknown coefficients are

$$\gamma_{002} = 0, \tag{32}$$

$$\alpha_{101} = \gamma_{020} = \frac{5c}{8\pi R^6} m^{(\mathbf{j})}_{xy} = \frac{a^7}{7R^6} A_{101}. \tag{33}$$

Equating the magnetic octupole components $m^{(\mathbf{i})}_{xyz}$ and $m^{(\mathbf{j})}_{xyz}$ calculated by (10) to each other gives zero values for α_{300}, β_{030}, and γ_{003} and results in the fourth subsystem of six equations

$$\begin{gathered} \langle \alpha_{102} - \alpha_{120} \rangle = \frac{7c}{2\pi R^8} m^{(\mathbf{j})}_{xyz}, \\ \beta_{210} + \gamma_{201} = 0, \quad \alpha_{120} + \beta_{210} = 0 \end{gathered}$$

for the unknown coefficients α_{120}, α_{102}, β_{210}, β_{020}, γ_{201}, and γ_{021}. Here,

$$m^{(\mathbf{j})}_{xyz} = \frac{4\pi}{63c} a^9 \langle A_{102} \rangle, \tag{34}$$

and

$$\begin{gathered} \langle A_{100} \rangle = 0, \\ A_{120} - A_{102} + \beta_{210} - B_{012} + 2C_{003} - C_{021} - C_{201} = 0, \\ A_{300} - A_{102} + C_{003} - C_{201} = 0, \\ 3A_{300} + B_{210} + C_{201} = 0, \\ A_{100} + a^2 A_{102} + C_{001} + a^2 C_{201} = 0, \\ C_{001} + a^2 C_{003} = 0. \end{gathered}$$

The calculations yield

$$\alpha_{300} = 0, \tag{35}$$

$$\alpha_{102} = \frac{7c}{12\pi R^8} m_{xyz}^{(\mathbf{j})} = \frac{a^9}{27R^8} \langle A_{102} \rangle. \tag{36}$$

Thus, the problem is solved, since the other coefficients are found by the circular permutation or interchanging of the coordinates in formulas (26)–(28), (30), (32), (33), (35), and (36).

The example considered involves the currents corresponding to the degrees $L = 2$ and 1 of the polynomials as special cases. In particular, the simplest adequate currents with the components in the form of first-degree polynomials may be represented as

$$j_x = 0, \quad j_y = B_{001}z, \quad j_z = -B_{001}y; \tag{37}$$

$$i_x = 0, \quad i_y = \frac{a^5}{5R^4} B_{001}z, \quad i_z = -\frac{a^5}{5R^4} B_{001}y. \tag{38}$$

CONCLUSION

We complete the paper with two remarks.

(1) Splitting the set of the algebraic equations for the polynomial coefficients of the surface currents into three ($L = 1$), seven ($L = 2$), or eight ($L \geq 3$) independent subsystems simplifies the process of finding the solution. At the same time, this means that, for a fixed L, there exists the corresponding number (3, 7, or 8) of independent systems of currents, each of which is responsible only for its own components of the magnetic 2^L field tensor.

(2) Consider the case when a sphere of radius a with given currents $\mathbf{j}(\mathbf{r})$ is concentrically placed into a spherical cavity of radius $R > a$ that is in the bulk of a superconductor. It is required to find the currents $\mathbf{i}'(\mathbf{r})$ induced on the surface of the cavity. If the problem of the surface currents $\mathbf{i}(\mathbf{r})$ adequate to the given bulk ones $\mathbf{j}(\mathbf{r})$ that is discussed in this paper is solved, the problem regarding the superconductor is solved as well. Namely,

$$\mathbf{i}'(\mathbf{r}) = -\mathbf{i}(\mathbf{r}).$$

REFERENCES

1. Ya. I. Frenkel, *Electrodynamics (General Theory of Electricity): Collection of Selected Works* (Akad. Nauk SSSR, Moscow, 1956), Vol. 1.
2. Ya. I. Frenkel, *Electrodynamics* (ONTI, Leningrad, 1935), Vol. 2.
3. R. Z. Muratov, Astron. Zh. **69**, 604 (1992) [Sov. Astron. **36**, 306 (1992)].
4. R. Z. Muratov, Zh. Tekh. Fiz. **67** (4), 1 (1997) [Tech. Phys. **42**, 325 (1997)].
5. B. V. Medvedev, *Principles of Theoretical Physics* (Nauka, Moscow, 1977).
6. S. P. Efimov, Teor. Mat. Fiz. **39** (2), 219 (1979).

Translated by M. Fofanov

Technical Physics, Vol. 47, No. 4, 2002, pp. 385–388. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 11–14.
Original Russian Text Copyright © 2002 by Perevezentsev, Svirina, Ugol'nikov.

GASES AND LIQUIDS

Model of Local Melting of Grain Boundaries Containing Impurity Atom Clusters

V. N. Perevezentsev, Yu. V. Svirina, and A. Yu. Ugol'nikov
Lobachevsky State University, pr. Gagarina 23, Nizhni Novgorod, 603600 Russia

Received January 16, 2001; in final form, October 2, 2001

Abstract—A model describing the initial stages of the local melting of grain boundaries containing impurity atom clusters is developed. The local melting process is considered as the formation of liquid zones whose geometry and number depend on the initial concentration and thermodynamic characteristics of the impurity atoms, as well as on the material parameters and temperature. The evolution of the liquid–solid interface structure with increasing temperature is discussed. The dependences of the liquid–solid interface parameters on temperature are obtained at various initial concentrations and thermodynamic characteristics of the impurity atoms. *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

It is known that some microcrystalline aluminum alloys and composites are prone to superplastic deformation at anomalously high rates (up to 10^3 s^{-1}) [1–7] at near-solidus temperatures. The study of high-rate superplasticity has shown that this effect is caused by the local melting of the grain boundaries. Experimental investigations into the local melting of grain boundaries using differential scanning calorimetry, transmission electron microscopy, and Auger spectroscopy have revealed the following features of this phenomenon [6]: (1) the local melting of the boundaries is closely related to the aggregation of impurity atoms at the boundaries (in particular, silicon and magnesium clusters at grain boundaries and interfaces play a vital role in aluminum alloys), (2) according to the phase and chemical composition of the alloy, the local melting temperature may be different from the solidus temperature by several degrees to tens of degrees, (3) both liquid and solid zones are observed at the boundaries under local melting conditions, and (4) the thickness of the liquid layer does not exceed 30 mm at the final stages of the local melting. Efforts to theoretically describe the liquid–solid structure of the grain boundaries with regard for the features listed, as well as the application of the theory to the problem of high-rate superplasticity, was undertaken in [8, 9]. A model of the boundary structure at the final stages of the local melting was proposed in [8]. It treats a grain boundary as a liquid layer containing solid streaks. The main disadvantage of this model is the assumption that the liquid–solid interface is flat. A more realistic model that considers conditions of phase and chemical equilibrium at the liquid–solid interfaces was developed in [9]. An approach put forward in [9] was applied to analyze the temperature dependence of the superplastic deformation rate at the initial stages of the local melting [10].

The aim of this study is to elaborate a model of grain-boundary local melting and to analyze the effect of the concentration and thermodynamic characteristics of impurity atoms segregating at the boundaries. It is also of interest to see how the material parameters affect the structure of the boundaries at the initial stages of the local melting. The application of this model for the description of high-rate superplasticity will be the subject of subsequent publications.

MODEL OF LIQUID–SOLID STRUCTURE OF GRAIN BOUNDARY

Let us suppose that, at the initial stage of the local melting, a grain boundary can be represented as a solid layer of thickness δ_0 that contains impurity atoms. Their concentration in this layer is $C_0 \gg C_0^v$, where C_0^v is the volume impurity concentration. It can be assumed that, early in the local melting, liquid islands (droplets) appear and their number increases with temperature. Therefore, we will view the liquid–solid structure of the grain boundary as a solid layer of the grain-boundary phase with uniformly distributed liquid droplets. During the local melting, the impurity is redistributed among the liquid and solid zones of the boundary. Let C_i^L and C_i^b be the concentrations of impurity atoms in the liquid and solid zones of the boundary, respectively. To describe the grain boundary structure, it is necessary to specify parameters characterizing the equilibrium geometry of the liquid droplets, the number of the liquid zones, and the concentration of impurity atoms in the liquid and solid zones of the boundary. Generally, the geometry of the liquid zones in phase equilibrium can be represented as two spherical segments with a radius of curvature R, height $y = (\delta - \delta_0)/2$ (δ is the thickness of the liquid zones), and base radius r_L. These seg-

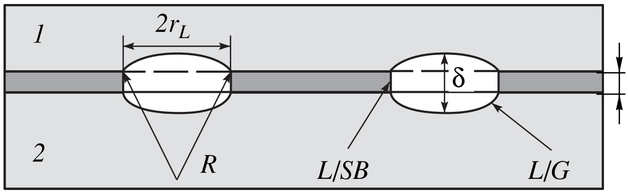


Fig. 1. Liquid–solid interface structure: *1*, grain I; *2*, grain II.

ments are separated by a cylinder of height δ_0 (Fig. 1). In order to determine the parameters describing the geometry of the liquid zones (y, R, and r_L), we consider the balance of forces acting at the interfaces between the liquid and solid phases, namely, at the interface between the grain body and the liquid zone, *L*/*G*, and at that between the liquid zone and the solid grain-boundary phase, *L*/*SB*. The balance condition at the *L/G* interface can be written in the form

$$2\gamma_{L/S}/R + q\Delta T/T = (C_i^L - C_0^v)kT, \quad (1)$$

where $\Delta T = T_m - T$ (T_m is the absolute melting point), $q = \lambda_m\rho$ is the heat of fusion at constant pressure, and $\gamma_{L/S}$ is the energy per unit surface area of the liquid–solid interface.

The first term on the left-hand side of Eq. (1) describes the surface pressure and the second one represents the configurational pressure related to a change in the bulk free energy due to the phase transition. The term on the right-hand side of Eq. (1) is the osmotic pressure of the impurity atoms. The balance condition at the *L*/*SB* interface has the form

$$\begin{gathered}\gamma_{L/S}/r_L + q\Delta T/T \\ + (2\gamma_{L/S}\cos\varphi - \gamma_b)/\delta_0 = (C_i^L - C_i^b)kT,\end{gathered} \quad (2)$$

where γ_b is the grain boundary surface energy (surface tension).

The third term on the left-hand side of Eq. (2) is the pressure acting on the *L*/*SB* interface due to the surface tension at the *L*/*G* interfaces and the *L*/*SB* interface, and φ is the angle between the *L*/*G* interface and its reference (undisturbed) plane (Fig. 1). In view of the relationship $\cos\varphi = 1 - y/R$, Eq. (2) can be written as

$$\begin{gathered}\gamma_{L/S}/r_L + q\Delta T/T + \Delta\gamma/\delta_0 \\ - 2\gamma_{L/S}y/\delta_0 R = (C_i^L - C_i^b)kT,\end{gathered} \quad (3)$$

where $\Delta\gamma = 2\gamma_{L/S} - \gamma_b$.

The parameters y, R, and r_L are related by the relationship

$$r_L^2 = y(2R - y). \quad (4)$$

To describe the structure of the liquid–solid interface, it is also necessary to determine the number of liquid droplets n_L per unit area of the grain boundary. We supplement the system of Eqs. (1), (3), and (4) by the condition of conservation of the number of the impurity atoms in the course of melting. In the limiting case where the outflow of the impurity atoms from the boundary into the grain body can be neglected in the local melting process, this condition becomes

$$C_0\delta_0 = C_i^b\delta_0(1 - \alpha_L) + C_i^L\delta_0\alpha_L + C_i^L 2V_L n_L, \quad (5)$$

where $V_L = (1/3)\pi y^2(3R - y)$ is the volume of the spherical segment, n_L is the number of the liquid droplets per unit area of the grain boundary, and $\alpha_L = \pi n_L r_L^2$ is the fraction of the liquid phase.

The first term on the right-hand side of Eq. (5) is the number of the impurity atoms per unit area of the solid zones at the boundary, and the second and the third ones describe the number of the impurity atoms in the liquid phase. By solving the system of Eqs. (1) and (3)–(5), one finds the analytical dependences of the geometrical parameters n_L, y, R, and r_L on C_i^L, C_i^b, and T.

The temperature dependences of the parameters n_L, y, R, and r_L can be found from those of the impurity atom concentrations C_i^L and C_i^b. To this end, we will take advantage of two more additional relationships. The first of them is the condition of chemical equilibrium between the impurity atoms in the liquid and solid grain-boundary phases:

$$\Psi_i^L + kT\ln C_i^L a^3 = \Psi_i^b + kT\ln C_i^b a^3, \quad (6)$$

where Ψ_i^L and Ψ_i^b are the chemical potentials of the impurity in the liquid and solid phases, respectively, and a is the interatomic spacing.

This condition implies that impurity diffusion between the liquid and solid phases is absent in equilibrium. The second relationship can be written on assumption that thermodynamic equilibrium exists and the free energy of the system remains unchanged ($\Delta F = 0$) during the local melting. The expression for ΔF can be written in the form

$$\Delta F = \Delta F_\gamma + \Delta F_m + \Delta G_i = 0, \quad (7)$$

where ΔF_γ is the change in the surface energy during the local melting, ΔF_m is the change in the bulk energy, and ΔG_i is the change in the chemical potential of the impurity atoms.

In turn, the expressions for ΔF_γ and ΔF_m per unit area of the boundary are written as

$$\Delta F_\gamma = 4\pi Ry n_L\gamma_{L/S} + 2\pi r_L\delta_0 n_L\gamma_{L/S} - \gamma_b\pi n_L r_L^2, \quad (8)$$

$$\begin{gathered}\Delta F_m = (q\Delta T/T - C_i^L kT)2V_L n_L \\ + (q\Delta T/T - (C_i^L - C_i^b)kT)\delta_0\pi n_L r_L^2 + (C_0 - C_i^b)kT\delta_0.\end{gathered} \quad (9)$$

The first and second terms on the right-hand side of Eq. (8) describe the increase in the surface energy of the system when the liquid droplets appear. The third term stands for the decrease in the surface energy of the grain boundary due to the disappearance of the solid zones

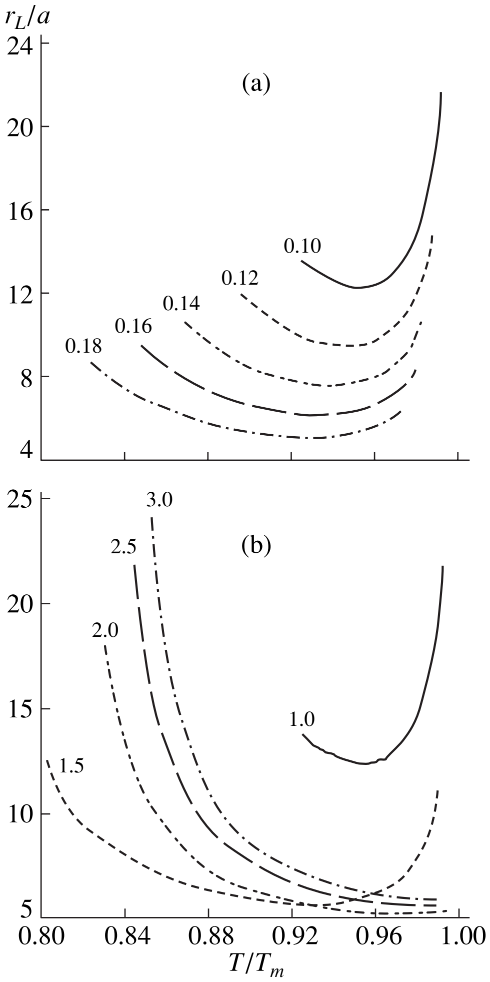


Fig. 2. Temperature dependence of the liquid droplet radius (a) for different initial concentrations of the impurity atoms $C'_0 = C_0a^3$ (indicated by figures at the curves) and (b) for different values of the thermodynamic characteristic $\Delta\Psi' = \Delta\Psi/kT_m$ of the impurity atoms (indicated by figures at the curves).

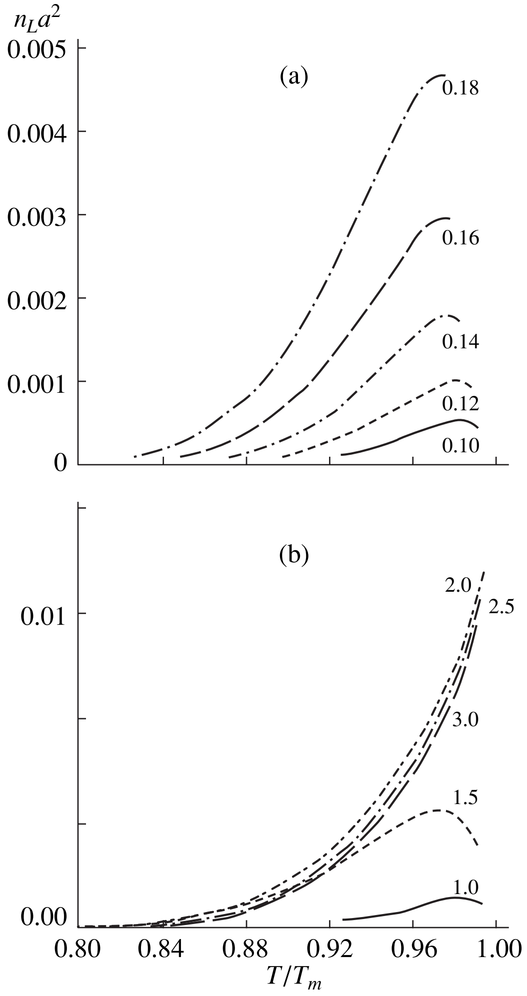


Fig. 3. Temperature dependence of the number of the liquid droplets per unit area of the boundary. (a) Values of C'_0 (indicated by figures at the curves) and (b) values of $\Delta\Psi'$ (indicated by figures at the curves).

during the local melting. On the right-hand side of (9), the first term describes the change in the energy during the melting of the grain body, the second term is the change in the energy when the grain boundary phase melts, and the third term is the change in the chemical potential of the grain boundary atoms due to the impurity redistribution during the local melting.

The expression for the change in the chemical potential of the impurity atoms has the form

$$\Delta G_i = -C_0\delta_0(\Delta\Psi + kT\ln(C_0/C_i^L)), \qquad (10)$$

where $C_0\delta_0$ is the amount of the impurity per unit area of the boundary and $\Delta\Psi = \Psi_i^b - \Psi_i^L$.

System (1), (3)–(7) is a self-consistent system of equations that enables the temperature dependences of the boundary structure parameters (n_L, y, R, r_L, C_0^L, and C_i^b) to be obtained under the local melting conditions for given material parameters ($\gamma_{L/S}$, γ_b, q, and δ_0), parameters characterizing the energy of the impurity atoms in the liquid and solid grain-boundary phases (Ψ_i^L and Ψ_i^b), and initial concentration of the impurity atoms.

NUMERICAL ANALYSIS OF BOUNDARY STRUCTURE PARAMETERS

The numerical simulation was carried out for the model alloy with the following typical values of the parameters: $qa^3 \sim 1.5kT_m$, $\delta_0 \sim 2a$, $\gamma_ba^2 \sim 0.3kT_m$, and $\gamma_{L/S}a^2 \sim 0.2kT_m$. As follows from the analysis, the temperature dependences of the radius of the liquid drop-

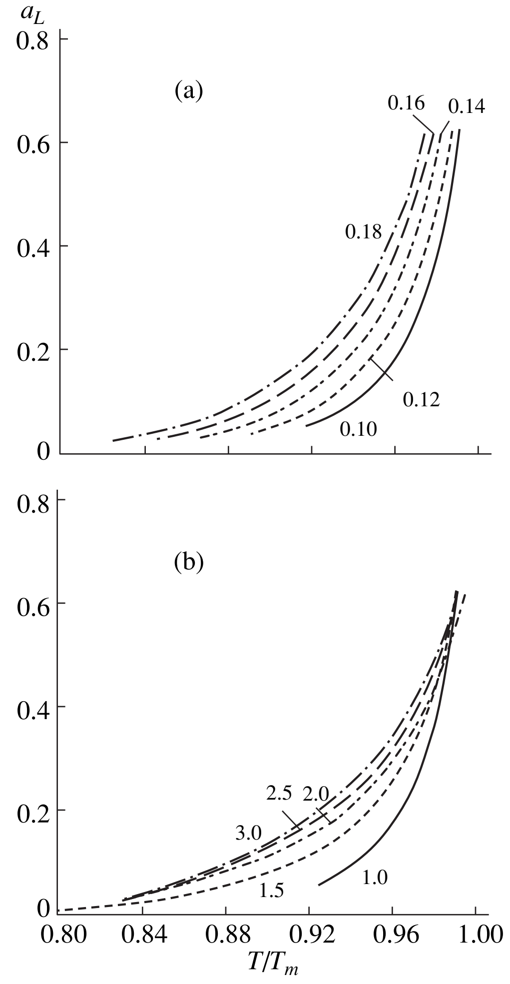


Fig. 4. Temperature dependence of the surface fraction of the liquid phase: (a) the same as in Fig. 2a and (b) the same as in Fig. 2b.

lets r_L (Fig. 2a) and of the number of the liquid droplets n_L per unit area of the boundary (Fig. 3a) are nonmonotonic at fixed $\Delta\Psi = 1kT_m$. For fixed $C_0a^3 = 0.1$, the run of the temperature dependences of r_L and n_L depends on the value of $\Delta\Psi$ (Figs. 2b, 3b). At small $\Delta\Psi$ ($1kT_m$ and $1.5kT_m$), these dependences are also nonmonotonic. At high $\Delta\Psi$, r_L and n_L vary monotonically throughout the temperature range considered. At the same time, for all the values of C_0a^3 and $\Delta\Psi$, the liquid phase fraction (Fig. 4) monotonically increases with temperature. The analysis shows that the temperature at which the liquid phase appears considerably depends on the initial impurity concentration at the boundary and on its energy characteristics. The local melting point T_0 is higher the lower the initial impurity concentration. It is noteworthy that, in the model, the beginning of local melting (the appearance of stable liquid droplets) is characterized by the finite value of the liquid phase fraction. At temperatures $T < T_{i0}$, there are no stable droplets and the liquid phase can exist only due to thermal fluctuations. It should be noted that, at $\alpha_L \approx 0.6$, the overlap of the liquid zones can be expected and the model becomes inapplicable. In this case, the grain boundary structure should be considered as a liquid layer with solid islands [9]. The variation of the material parameters $\gamma_b a^2$ and qa^3 does not affect the temperature dependence of the boundary parameters but considerably changes the melting point: the process starts earlier in the materials with a lower heat of fusion qa^3 and a higher surface energy $\gamma_b a^2$ of the boundary.

CONCLUSION

A model describing the initial stages of the local melting of grain boundaries containing impurity clusters was developed.

The system of equations relating the liquid–solid interface parameters (geometry, dimensions and number of the liquid zones, and impurity content in the liquid and solid zones at the boundary) to temperature, initial concentration, and thermodynamic characteristics of the impurity atoms were derived.

The numerical analysis of the model suggested shows that the local melting starts (stable liquid droplets appear) at a finite fraction of the liquid phase. The fraction of the liquid phase invariably increases monotonically with temperature, whereas the temperature dependences of the size and number of liquid zones vary according to the initial concentration and thermodynamic characteristics of the impurity.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 00-02-16546).

REFERENCES

1. T. G. Nieh, C. A. Henshall, and J. Wadsworth, Scr. Metall. **19**, 1375 (1985).
2. K. Higashi, Mater. Sci. Forum **170–172**, 131 (1994).
3. L. Koike, M. Mabuchi, and K. Higashi, Acta Metall. Mater. **43**, 199 (1995).
4. K. Higashi, M. Mabuchi, and T. G. Langdon, ISIJ Int. **36**, 1423 (1996).
5. K. Higashi, T. G. Nieh, and J. Wadsworth, Acta Metall. Mater. **43**, 3275 (1995).
6. K. Higashi, Mater. Sci. Forum **243–245**, 267 (1997).
7. M. Mabuchi, H. Iwasaki, and K. Higashi, Acta Mater. **46**, 5335 (1998).
8. V. N. Perevezentsev, Mater. Sci. Forum **243–245**, 297 (1997).
9. V. N. Perevezentsev, K. Higashi, and J. V. Svirina, Mater. Sci. Forum **304–306**, 217 (1999).
10. V. N. Perevezentsev, T. G. Langdon, and J. V. Svirina, in *Proceedings of the International Conference "Current Status of Theory and Practice of Superplasticity in Materials," Ufa, 2000*, p. 21.

Translated by M. Lebedev

Technical Physics, Vol. 47, No. 4, 2002, pp. 389–396. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 15–22.
Original Russian Text Copyright © 2002 by Shiryaeva.

GASES AND LIQUIDS

Nonlinear Vibrations of a Charged Drop Due to the Initial Excitation of Neighboring Modes

S. O. Shiryaeva
Demidov State University, Sovetskaya ul. 14, Yaroslavl, 150000 Russia
e-mail: shir@uniyar.ao.ru

Received May 31, 2001

Abstract—The asymptotic analysis of the nonlinear vibrations of a charged drop that are induced by a multimode initial deformation of its equilibrium shape is performed. It is shown that when two, three, or several neighboring modes are present in the initial deformation spectrum, the mode with the number one (translational mode) appears in the second-order mode spectrum. The excitation of the translational mode follows from the requirement of center-of-mass immobility and causes the dipole components (which are absent in the linear analysis) to appear in the spectra of the acoustic and electromagnetic radiation of the charged drop. *© 2002 MAIK “Nauka/Interperiodica”.*

(1) The capillary vibration and the stability of a charged drop are of considerable interest in a number of areas of science and technology where charged drops play an important part (see, e.g., [1–3] and Refs. therein). Because of much interest in a charged drop as a physical object, most related physical problems formulated in the framework of linear models have already been solved. A large number of papers devoted to the nonlinear analysis has appeared in recent years [4–13]. This analysis allows one to gain much more information on the object. In view of the cumbersome calculations associated with the nonlinear problems, many aspects of the nonlinear vibration of a charged drop either have not been considered or are poorly understood. The latter concerns the so-called translational instability of drops and bubbles, which shows up when two neighboring modes are present in the spectrum of initially excited modes [14]. In accordance with [14], the center of mass of a translationally unstable drop undergoing nonlinear oscillations acquires a translational velocity. This statement seems to be incorrect, since it is in contradiction with the well-known law of theoretical mechanics: no motions inside a closed system can result in the motion of its center of mass. Therefore, the problem of nonlinear capillary vibrations of a charged drop when the initial deformation of the equilibrium spherical shape of a drop is induced by the excitation of neighboring modes is of interest.

(2) Consider the time evolution of the shape of a drop of an ideal incompressible perfectly conducting liquid with the density ρ and the surface tension coefficient σ. We assume that the drop, having the charge Q and radius R, is in a vacuum. At the time instant $t = 0$, the equilibrium spherical shape of the drop undergoes an axisymmetric perturbation of fixed amplitude, the perturbation being much smaller than the radius of the drop. Our aim is to find the spectrum of arising capillary oscillations of a drop (its shape) for $t > 0$.

Since the initial perturbation of the surface is axisymmetric and small, we assume that the shape of the drop remains axisymmetric at any time instant. Then, in the polar coordinate system with the origin at the center of mass of the drop, the equation describing the surface of the drop in dimensionless variables where $R = \rho = \sigma = 1$ has the form

$$r(\Theta, t) = 1 + \xi(\Theta, t), \quad |\xi| \ll 1. \tag{1}$$

The flow inside the drop is assumed to be potential. In this case, the flow field in the drop, $\mathbf{V}(\mathbf{r}, t) = \nabla\psi(\mathbf{r}, t)$, is defined only by the velocity potential function $\psi(\mathbf{r}, t)$.

The set of equations for drop evolution consists of the Laplace equations for the field velocity potential $\psi(\mathbf{r}, t)$ and electrostatic potential $\Phi(\mathbf{r}, t)$

$$\Delta\psi(\mathbf{r}, t) = 0, \tag{2}$$

$$\Delta\Phi(\mathbf{r}, t) = 0 \tag{3}$$

(hereafter, Δ is the Laplacian operator) and boundary conditions

$$r \longrightarrow 0: \psi(\mathbf{r}, t) \longrightarrow 0, \tag{4}$$

$$r \longrightarrow \infty: \Phi(\mathbf{r}, t) \longrightarrow 0, \tag{5}$$

$$r = 1 + \xi(\Theta, t): \frac{\partial\xi}{\partial t} = \frac{\partial\psi}{\partial r} - \frac{1}{r^2}\frac{\partial\xi}{\partial\Theta}\frac{\partial\psi}{\partial\Theta}, \tag{6}$$

$$\Delta p - \frac{\partial\psi}{\partial t} - \frac{1}{2}(\nabla\psi)^2 + \frac{1}{8\pi}(\nabla\Phi)^2 = \nabla\cdot\mathbf{n}, \tag{7}$$

$$\Psi(r, \Theta, t) = \Phi_s(t). \tag{8}$$

We also introduce the condition of center-of-mass immobility and the condition of conservation of the total charge and volume of the drop:

$$-\frac{1}{4\pi}\oint_S(\mathbf{n}\cdot\nabla\Phi)ds = Q, \tag{9}$$
$$S = [r = 1+\xi(\Theta, t),\ 0\le\Theta\le\pi,\ 0\le\phi\le 2\pi],$$

$$\int_V r^2 dr\sin\Theta d\Theta d\phi = \frac{4}{3}\pi, \tag{10}$$
$$V = [0\le r\le 1+\xi(\Theta, t),\ 0\le\Theta\le\pi,\ 0\le\phi\le 2\pi],$$

$$\int_V \mathbf{e}_r r^3 dr\sin\Theta d\Theta d\phi = 0, \tag{11}$$
$$V = [0\le r\le 1+\xi(\Theta, t),\ 0\le\Theta\le\pi,\ 0\le\phi\le 2\pi].$$

Note that conditions (10) and (11) have to be satisfied at any instant of time including the initial one. Therefore, at $t = 0$, these conditions define the amplitudes of the zero and first modes in the expansion of the initial perturbation of the equilibrium spherical shape of the drop $\xi(\Theta)$ in Legendre polynomials. In other words, the zero and first modes cannot have arbitrary amplitudes. Their amplitudes depend on the shape of the initial deformation (perturbation).

The initial conditions are set as the initial deformation of the equilibrium spherical shape of the drop and as the zero initial velocity of the surface:

$$t = 0\colon \xi(\Theta) = \xi_0 P_0(\mu) + \xi_1 P_1(\mu) + \varepsilon\sum_{i\in\Xi} h_i P_i(\mu), \tag{12}$$
$$\sum_{i\in\Xi} h_i = 1, \quad \frac{\partial\xi}{\partial t} = 0,$$

where Ξ is the set of the numbers of initially excited modes and $\mu \equiv \cos\Theta$.

In (6)–(12), Δp is the difference in the constant pressures inside and outside of the drop in equilibrium; $\mathbf{n}$, unit normal vector to the surface defined by (1); $\Phi_s(t)$, potential that is constant over the surface of the drop; ε, amplitude of the initial perturbation of the surface shape; $P_i(\mu)$, ith-order Legendre polynomials; h_i, coefficient specifying the partial contribution of the ith vibration mode to the total initial perturbation; and ξ_0 and ξ_1, constants that are determined from conditions (10) and (11) at the initial time instant and are equal to

$$\xi_0 \approx -\varepsilon^2\sum_{i\in\Xi}\frac{h_i^2}{(2i+1)} + O(\varepsilon^3), \tag{13}$$
$$\xi_1 \approx -\varepsilon^2\sum_{i\in\Xi}\frac{9ih_{i-1}h_i}{(2i-1)(1i+1)} + O(\varepsilon^3).$$

up to second-order infinitesimals in ε.

In Eqs. (3), (5), (8), and (9), it is assumed that the electric charge is distributed over the surface and is in equilibrium at any instant of time. This assumption is valid if the characteristic time of the charge redistribution is much smaller than the characteristic hydrodynamic time of surface vibration

$$\varepsilon_*\varepsilon_0\lambda \ll \left(\frac{\rho R^3}{\sigma}\right)^{1/2},$$

where λ is the resistivity of the liquid and ε_* is the permittivity.

For a drop of distilled water ($\lambda = 10^4\ \Omega$ m, $\varepsilon_* = 80.08$, $\rho = 10^3$ kg/m^3, $\sigma = 72.8\times10^{-3}$ N/m, and $\varepsilon_0 = 8.85\times10^{-12}$ F/m), the left-hand side of this inequality is two orders of magnitude smaller than the right-hand one at $R = 1$ mm.

(3) We will take advantage of method of multiple scales [15] to solve the problem up to second-order infinitesimals in ε. We expand the unknown functions $\xi(\Theta, t)$, $\psi(\mathbf{r}, t)$, and $\Phi(\mathbf{r}, t)$ in the power series in the small parameter ε, and consider the series as functions of various time scales (rather than merely of time t) defined via the small parameter ε: $T_m \equiv \varepsilon^m t$. Then,

$$\xi(\Theta, t) = \sum_{m=0}^{\infty}\varepsilon^m\xi^{(m)}(\Theta, T_0, T_1, T_2, \ldots);$$
$$\psi(\mathbf{r}, t) = \sum_{m=0}^{\infty}\varepsilon^m\psi^{(m)}(r, \Theta, T_0, T_1, T_2, \ldots); \tag{14}$$
$$\Phi(\mathbf{r}, t) = \sum_{m=0}^{\infty}\varepsilon^m\Phi^{(m)}(r, \Theta, T_0, T_1, T_2, \ldots).$$

Using standard techniques involved in the method of multiple scales (see, e.g., [13] for details), we find the expression for the time evolution of the surface shape:

$$r(\Theta, t) \approx 1 + \varepsilon\left\{\sum_{i\in\Xi} M_i^{(1)}(t)P_i(\mu)\right\}$$
$$+ \varepsilon^2\left\{\sum_{n=0}^{\infty} M_n^{(2)}(t)P_n(\mu)\right\} + O(\varepsilon^3);$$

$$M_i^{(1)}(t) = h_i\cos(\omega_i t);$$

$$M_0^{(2)}(t) = -\frac{1}{2}\sum_{i\in\Xi}\frac{h_i}{(2i+1)}(1+\cos(2\omega_i t));$$

$$M_1^{(2)}(t) = -\sum_{i\in\Xi}\frac{9ih_{i-1}h_i}{(2i-1)(2i+1)}\cos(\omega_i t)\cos(\omega_{i-1}t);$$

$$M_n^{(2)} = [N_n(t) - N_n(0)\cos(\omega_n t)], \quad n \ge 2;$$

$$N_n(t) = \frac{1}{2}\sum_{i \in \Xi}\sum_{j \in \Xi} h_i h_j \{\lambda_{ijn}^{(+)}\cos((\omega_i + \omega_j)t) + \lambda_{ijn}^{(-)}\cos((\omega_i - \omega_j)t)\};$$

$$\omega_n^2 \equiv n(n-1)[(n+2) - W]; \quad W \equiv \frac{Q^2}{4\pi};$$

$$\lambda_{ijn}^{(\pm)} \equiv [\gamma_{ijn} \pm \omega_i\omega_j\eta_{ijn}][\omega_n^2 - (\omega_i \pm \omega_j)^2]^{-1};$$

$$\gamma_{ijn} \equiv K_{ijn}\Big[\omega_i^2(n - i + 1) + 2n[j(j+1) - 1] + [j(i+1) - i(2i - 2n + 7) + 3]n\frac{W}{2}\Big] + \alpha_{ijn}\Big[\frac{1}{i}\omega_i^2 + n\frac{W}{2}\Big];$$

$$\eta_{ijn} \equiv K_{ijn}\Big(\frac{n}{2} - i + 1\Big) + \alpha_{ijn}\frac{1}{i}\Big(1 + \frac{n}{2j}\Big);$$

$$K_{ijn} \equiv [C_{ijn}^{000}]^2;$$

$$\alpha_{ijn} \equiv -\sqrt{i(i+1)j(j+1)}\,C_{ijn}^{000}C_{ijn}^{-110};$$

$$C_{ijn}^{000} \equiv \begin{cases} 0, \quad \text{if} \quad i + j + n = 2g + 1, \\ \text{where } g \text{ is an integer;} \\ \dfrac{(-1)^{g-n}\sqrt{2n+1}\,g!}{(g-i)!(g-j)!(g-n)!} \\ \times\Big[\dfrac{(2g-2i)!(2g-2j)!(2g-2n)!}{(2g+1)!}\Big]^{1/2}, \end{cases}$$

if $i + j + n = 2g$ (g is an integer);

$$C_{ijn}^{-110} \equiv \sqrt{2n+1}\,n! \times \Big[\frac{(i+j-n)!\,i(i+1)}{(n+i-j)!(n-i+j)!(i+j+n+1)!\,j(j+1)}\Big]^{1/2} \times \sum_z \frac{(-1)^{i+1+z}(i+z-1)!(n+j-z+1)!}{z!(i-z+1)!(n-z)!(j-n+z-1)!}. \tag{15}$$

The summation in the last expression is over all integers z yielding nonnegative values of the expressions under the factorial sign. Note that C_{ijn}^{000} and C_{ijn}^{-110} are the Clebsch–Gordan coefficients [16], which are nonzero only if the subscripts satisfy the conditions $|i - j| \le n \le (i + j)$ and $i + j + n = 2g$ (g is an integer).

From (15) it follows that, if two neighboring vibration modes (even and odd) are present in the mode spectrum describing the initial surface perturbation, the amplitude of the first mode is nonzero. In this case, the center of mass is at rest at the origin for $t \ge 0$ because its displacement induced by the excitation of two neighboring modes is compensated. Note that the translational instability of a bubble due to the excitation of two neighboring surface vibrational modes discussed in [14] is associated with incorrect initial conditions. Namely, the statement of the problem in [14] does not require that the center-of-mass system initially be at rest. In other words, in [14], the velocity of the center of mass of the drop is other than zero with respect to the laboratory system of coordinates at $t = 0$.

The time dependences of the amplitudes of the modes excited by the intermode interaction in the second-order approximation with respect to ε are shown in Figs. 1 and 2 for various initial deformations of the spherical shape. In Fig. 1, the initial deformation is defined by the superposition of two modes of equal amplitudes; in Fig. 2, by that of three modes of the same amplitude. According to these data, the amplitude of the fundamental mode has the highest rate of rise presumably because its excitation energy is the least. According to Figs. 1a–1c, when the initial deformation of the equilibrium shape is due to two various modes, the rate of rise of the fundamental mode amplitude increases with increasing number of initially excited modes. For three initially excited modes, the tendency is the same. It seems likely that, for ε = const, the energy injected into the vibrating system rapidly increases as the numbers of modes responsible for the initial deformation grow.

The number of modes excited in the second-order approximation depends on the fact that the coefficients $\lambda_{iin}^{(\pm)}$ and $\lambda_{jjn}^{(\pm)}$ in (15) are nonzero only for even n in the closed intervals $[0, 2i]$ and $[0, 2j]$, respectively, whereas $\lambda_{ijn}^{(\pm)}$ and $\lambda_{jin}^{(\pm)}$ are other than zero for n from the closed interval $[|i - j|, (i + j)]$ such that $n + i + j$ is even. Therefore, when odd and even modes (with numbers k and p) are excited simultaneously, their interaction will generate all even modes from the closed interval $[0; \max(2k, 2p)]$ and all odd modes from the interval $[|k - p|, (k + p)]$. In the case of the initial excitation of either two even or two odd modes, only even modes from the closed interval $[0; \max(2k, 2p)]$ are excited in the second-order approximation.

(4) Let us consider in greater detail whether the center of charge of a drop is displaced under nonlinear vibration of its surface. The radius vector of the center of charge $\mathbf{R}_q$ is given by

$$\mathbf{R}_q = \frac{1}{Q}\int_S \mathbf{r}\,dQ = -\frac{1}{4\pi Q}\int_S (\mathbf{n}\cdot\nabla\Phi)\mathbf{r}\,dS = -\frac{1}{4\pi Q}\int_\Omega \frac{(\mathbf{n}\cdot\nabla\Phi)}{(\mathbf{n}\mathbf{e}_r)} r^3(\Theta, t)\mathbf{e}_r\,d\Omega,$$

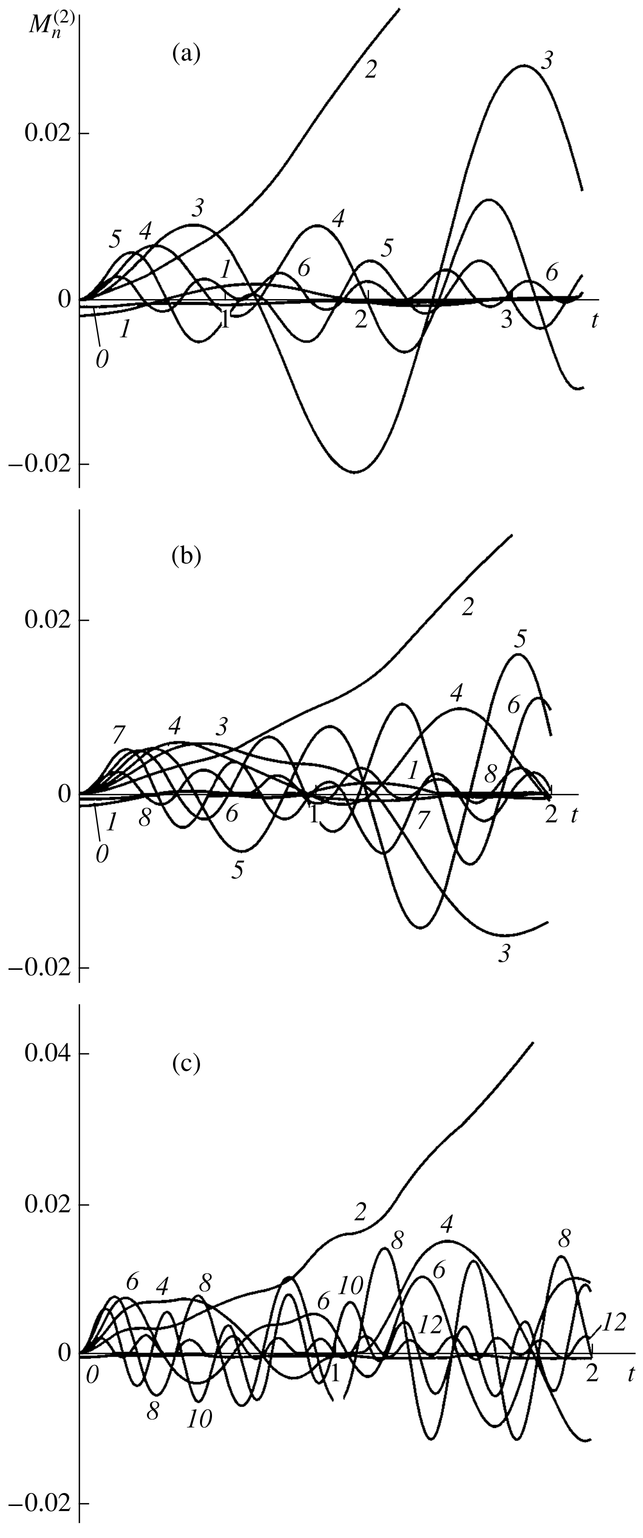


Fig. 1. Time dependences of the dimensionless amplitudes $M_n^{(2)}$ of various modes of the capillary vibrations of a charged drop that appear in the second-order approximation due to interaction between modes at $W = 3.9$ and $\varepsilon = 1$, and the initial deformation in the linear approximation with respect to ε: (a) $\varepsilon[P_2(\mu) + P_3(\mu)]/2$, (b) $\varepsilon[P_3(\mu) + P_4(\mu)]/2$, and (c) $\varepsilon[P_4(\mu) + P_6(\mu)]/2$. The numbers at the curves denote the mode number.

where $\Omega \equiv [0 \le \Theta \le \pi;\ 0 \le \phi \le 2\pi]$, $r(\Theta, t)$ is given by (15); and the parentheses stand for the scalar product of the vectors.

The radial unit vector $\mathbf{e}_r$ of the spherical coordinate system is related to the unit vectors of the Cartesian system as

$$\mathbf{e}_r = \mathbf{e}_x \sin\Theta\cos\phi + \mathbf{e}_y \sin\Theta\sin\phi + \mathbf{e}_z\cos\Theta.$$

Let us express the components of the vector $\mathbf{R}_q$ in the Cartesian coordinates through the spherical functions $Y_n^m(\Theta, \phi)$ in view of the relations

$$\cos\Theta = \sqrt{\frac{4\pi}{3}}Y_1^0(\Theta, \phi);$$

$$\sin\Theta\cos\phi = -\sqrt{\frac{2\pi}{3}}[Y_1^1(\Theta, \phi) + Y_1^{-1}(\Theta, \phi)],$$

$$\sin\Theta\sin\phi = i\sqrt{\frac{2\pi}{3}}[Y_1^1(\Theta, \phi) - Y_1^{-1}(\Theta, \phi)].$$

Then, the components of the vector $\mathbf{R}_q$ will have the form

$$R_{qx} = \frac{1}{2\sqrt{6\pi}Q}\int_\Omega [r(\Theta, t)]^3 \frac{(\mathbf{n}\cdot\nabla\Phi)}{(\mathbf{n}\mathbf{e}_r)} \times [Y_1^1(\Theta, \varphi) + Y_1^{-1}(\Theta, \varphi)]d\Omega,$$

$$R_{qy} = -\frac{i}{2\sqrt{6\pi}Q}\int_\Omega [r(\Theta, t)]^3 \frac{(\mathbf{n}\cdot\nabla\Phi)}{(\mathbf{n}\mathbf{e}_r)} \times [Y_1^1(\Theta, \varphi) - Y_1^{-1}(\Theta, \varphi)]d\Omega, \qquad (16)$$

$$R_{qz} = \frac{1}{2\sqrt{3\pi}Q}\int_\Omega [r(\Theta, t)]^3 \frac{(\mathbf{n}\cdot\nabla\Phi)}{(\mathbf{n}\mathbf{e}_r)} Y_1^0(\Theta, \varphi)d\Omega.$$

In view of (1) and expansions (14), the integrand in (16) can be represented by the power series in the small parameter ε up to terms of the order of $O(\varepsilon^3)$:

$$[r(\Theta, t)]^3\frac{(\mathbf{n}\cdot\nabla\Phi)}{(\mathbf{n}\mathbf{e}_r)} \approx -Q + \varepsilon\left[\frac{\partial\Phi^{(1)}}{\partial r} - Q\xi^{(1)}\right]_{r=1} + \varepsilon^2\left[\frac{\partial\Phi^{(2)}}{\partial r} - Q\xi^{(2)} + \left(\frac{\partial^2\Phi^{(1)}}{\partial r^2} + 3\frac{\partial\Phi^{(1)}}{\partial r}\right)\xi^{(1)} - \frac{\partial\Phi^{(1)}}{\partial\Theta}\frac{\partial\xi^{(1)}}{\partial\Theta}\right]_{r=1} + O(\varepsilon^3).$$

Taking into account the form of the functions $\xi^{(1)}(\Theta, t)$ and $\xi^{(2)}(\Theta, t)$ in (15), as well as the potential $\Phi(\mathbf{r}, t)$ found by solving the boundary problem given by (3), (5), (8), and (9), namely,

$$\Phi^{(1)}(r, \Theta, t) = Q\sum_{i=\Xi} M_i^{(1)}(t) r^{-(i+1)} P_i(\mu),$$

$$\Phi^{(2)}(r,\Theta,t) = Q\sum_{n=1}^{\infty} M_n^{(2)}(t)$$

$$+\sum_{i\in\Xi}\sum_{j\in\Xi} iK_{ijn}M_i^{(1)}(t)M_j^{(1)}(t)r^{-(n+1)}P_n(\mu),$$

we represent the integrand in (16) as an expansion in Legendre polynomials:

$$[r(\Theta,t)]^3\frac{(\mathbf{n}\cdot\nabla\Phi)}{(\mathbf{n}\mathbf{e}_r)} \approx -QP_0(\mu)$$

$$-\varepsilon Q\sum_{i\in\Xi}(i+2)h_i\cos(\omega_i t)P_i(\mu)$$

$$-\varepsilon^2 Q\sum_{n=0}^{\infty}\{(n+2-\delta_{n,0})M_n^{(2)}(t) \qquad (17)$$

$$-\sum_{i\in\Xi}\sum_{j\in\Xi}[((i-1)(i+1)-i(n+1-\delta_{n,0}))K_{ijn}$$

$$-\alpha_{ijn}]M_i^{(1)}(t)M_j^{(1)}(t)\}P_n(\mu)+O(\varepsilon^3).$$

In the derivation of (17), the following expansions coming from the Clebsch–Gordan expansion [16] have been used:

$$P_i(\mu)P_j(\mu) = \sum_{n=0}^{\infty}K_{ijn}P_n(\mu),$$

$$\frac{\partial P_i}{\partial\Theta}\frac{\partial P_j}{\partial\Theta} = \sum_{n=0}^{\infty}\alpha_{ijn}P_n(\mu),$$

where the coefficients K_{ijn} and α_{ijn} are defined in (15).

Substituting (17) into (16), we turn from the Legendre polynomials to the spherical functions [16]

$$P_j(\mu) = \sqrt{\frac{4\pi}{(2j+1)}}Y_j^0(\Theta,\phi)$$

taking into account the orthonormality condition for the spherical functions

$$\int_\Omega Y_j^m(\Theta,\phi)[Y_n^k(\Theta,\phi)]^*d\Omega = \delta_{j,n}\delta_{m,k}. \qquad (18)$$

Then, in the case of axisymmetric (with respect to the z axis) vibrations of the surface, when the perturbation of its equilibrium spherical shape $\xi = xi(\Theta, t)$ is independent of the angle ϕ the integrand in (17) is represented by the series in spherical functions Y_j^0 with the zero superscript and, in view of (18), the center of charge in the xy plane is not displaced:

$$R_{qx} = R_{qy} = 0.$$

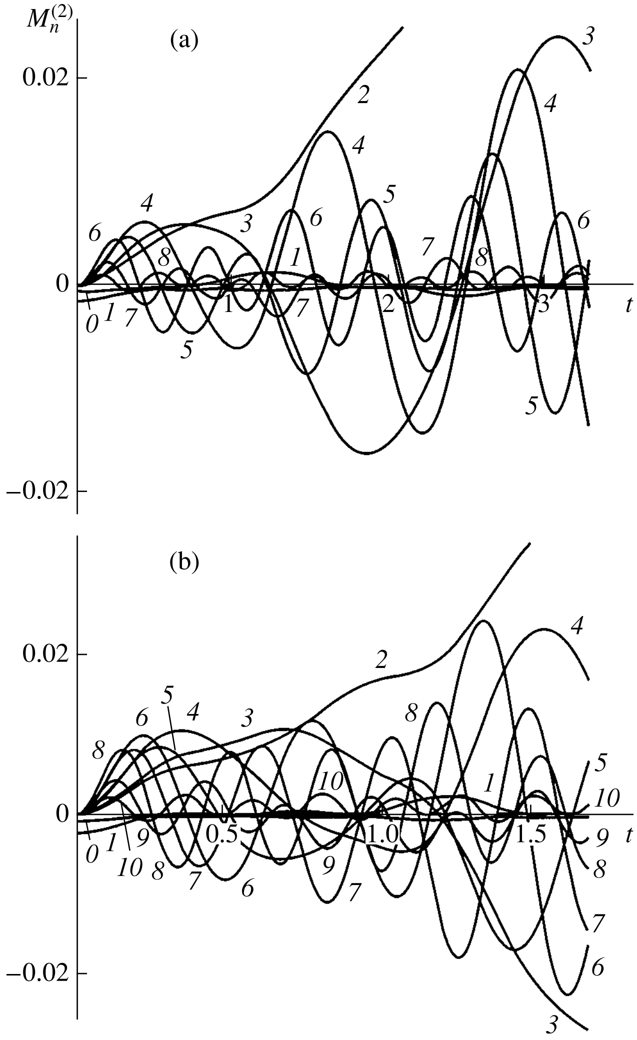


Fig. 2. The same as in Fig. 1 for (a) $\varepsilon[P_2(\mu) + P_3(\mu) + P_4(\mu)]/3$ and (b) $\varepsilon[P_3(\mu) + P_4(\mu) + P_5]/3$.

Its displacement along the Oz axis is described by

$$R_{qz} \approx \varepsilon^2\Big\{M_1^{(2)}(t) + \frac{1}{3}\sum_{i\in\Xi}\sum_{j\in\Xi}[\alpha_{ij1} - ((i-1)(i+1) \qquad (19)$$

$$-2i)K_{ij1}]M_i^{(1)}(t)M_j^{(1)}(t)\Big\} + O(\varepsilon^3).$$

Recall that, according to definitions introduced in (15), the coefficients α_{ijn} and K_{ijn} are expressed via the Clebsch–Gordan coefficients and, hence, are nonzero only if $i + j + n = 2g$ (g is an integer) and $|i - j| \le n \le (i + j)$. In our case, $n = 1$, so that

$$i + j = 2g - 1, \qquad (20)$$

$$|i - j| \le 1 \le i + j. \qquad (21)$$

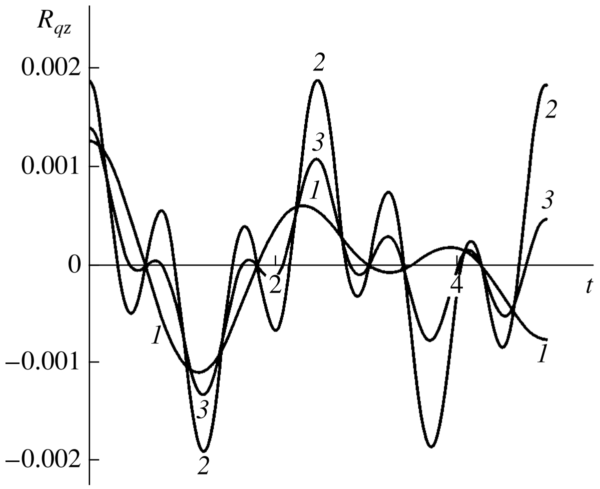


Fig. 3. Dimensionless displacement of the center of charge of a drop R_{qz} vs. dimensionless time t for $W = 3.9$ and $\varepsilon = 0.1$. In the linear approximation with respect to ε, the initial perturbation of the equilibrium spherical shape has the form (*1*) $\varepsilon[P_2(\mu) + P_3(\mu)]/2$, (*2*) $\varepsilon[P_3(\mu) + P_4(\mu)]/2$, and (*3*) $\varepsilon[P_3(\mu) + P_4(\mu)]/3$.

The right-hand inequality in (21) is always satisfied, since, in the spectrum of the vibrational modes Ξ, the zero and first modes are absent. The left-hand inequality in (21) implies that $j = i$ or $i - 1$ for $j \le i$ and $j = i + 1$ for $j > i$. Consequently, the index j should take the values $\{i - 1; i; i + 1\}$ at a fixed i. However, the value $j = i$ contradicts requirement (20). Thus, the excitation spectrum for a given i should also contain the $(i + 1)$th or $(i - 1)$th mode (i.e., a neighbor of the ith mode) in order for the coefficients λ_{ij1} and K_{ij1} to be nonzero. In this case, the center of charge will be displaced along the z axis.

Using the explicit form of the Clebsch–Gordan coefficients [16] and the Kronecker symbol $\delta_{m,k}$, one readily finds

$$K_{ij1} = \delta_{j,i-1}\frac{3i}{(2i-1)(2i+1)} + \delta_{j,i+1}\frac{3(i+1)}{(2i+1)(2i+3)}, \tag{22}$$
$$\alpha_{ij1} = \delta_{j,i-1}(i-1)(i+1)K_{ij1} + \delta_{j,i+1}i(i+2)K_{ij1}.$$

Substituting (22) into (19) and reindexing, we reduce (19) to the more compact form

$$R_{qz} \approx \varepsilon^2\left\{M_1^{(2)}(t) + \sum_{i\in\Xi}\frac{3i}{(2i+1)}M_i^{(1)}(t)M_{i-1}^{(1)}(t)\right\}. \tag{23}$$

Finally, taking into account the expressions for the coefficients $M_i^{(1)}$ and $M_i^{(2)}$ in (15) and turning to dimensional variables, we find the displacement of the center of charge along the z axis in the second-order approximation with respect to the surface perturbation:

$$R_{qz} \approx \varepsilon^2 R\sum_{i\in\Xi}\frac{6i(i-2)}{(2i-1)(2i+1)}h_ih_{i-1}\cos(\omega_i t)$$
$$\times\cos(\omega_{i-1}t) \equiv \varepsilon^2 R\sum_{i\in\Xi}\frac{3i(i-2)}{(2i-1)(2i+1)}h_ih_{i-1} \tag{24}$$
$$\times\,[\cos[(\omega_i + \omega_{i-1})t] + \cos[(\omega_i - \omega_{i-1})t]],$$
$$\omega_i^2 \equiv \frac{\sigma}{\rho R^3}i(i-1)\left[(i+2) - \frac{Q^2}{4\pi\sigma R^3}\right].$$

The plot $R_{qz} = R_{qz}(t)$ for several numbers of initially excited neighboring modes is shown in Fig. 3.

(5) The presence of center-of-charge oscillations causes the drop to emit dipole electromagnetic radiation. Consider the case when two neighboring modes i and $i - 1$ are initially excited. The intensity I_e of the electromagnetic radiation from a single drop has the form [17]

$$I_e = \frac{4}{3c^3}\{|d_{\omega_i+\omega_{i+1}}|^2(\omega_i + \omega_{i+1})^4$$
$$+ |d_{\omega_{i+1}-\omega_i}|^2(\omega_{i+1} - \omega_i)^4\},$$

where c is the velocity of light in a vacuum, d_{ω_j} is the dipole moment of the drop with its center of charge oscillating with a frequency ω_j.

According to (24), in our situation, we have

$$d_{\omega_i+\omega_{i+1}} \equiv d_{\omega_{i+1}-\omega_i} \equiv Q\overline{R_{qz}},$$

where $\overline{R_{qz}}$ is the peak value of R_{qz}.

Hence,

$$I_e = \frac{4}{3c^3}Q^2(\overline{R_{qz}})^2\{(\omega_i + \omega_{i+1})^4 + (\omega_{i+1} - \omega_i)^4\}. \tag{25}$$

In view of (24), expression (25) evaluates the order of magnitude of the intensity of the background (noise) electromagnetic radiation from various (natural and artificial) liquid-drop systems, such as clouds of convection. In [18], the intensity of the electromagnetic radiation from a cumulus cloud was estimated in the framework of linear analysis. The physical model proposed in [18] considers the total multipole (starting from quadrupole) radiation from large, heavily charged hydrometeors (drops of $R = 1$ mm) freely falling in the cloud and coagulating with smaller charged drops of radius 10 μm. The total intensity of the multipole electromagnetic radiation (except for the dipole radiation, which is absent in the linear analysis) of a cloud with a diameter of 5 km, $I = 3 \times 10^{-2}$ W, found in [18] is essentially overestimated. This is because the maximum admissible charges of hydrometeors rather than their mean values (which are several orders of magnitude smaller) were used in calculations and the concentration of hydrometeors of the given size was overstated.

Let us estimate the intensity of the background dipole electromagnetic radiation with (24) and (25) when the displacement of the center of charge is caused by the excitation of only two neighboring modes with $i = 100$ and $i + 1 = 101$. Note that, in this case, the sum in (24), which determines the amplitude of the center-of-charge displacement R_{qz}, is reduced to the single term with $i = 101$. The estimation is performed for a drop with the mean radius $R = 30$ μm and mean charge $Q = 2.5 \times 10^5$ C.G.S.E. The concentration of such drops in a cumulus cloud is $n \approx 10^3$ cm^{-3} according to reference data [19]. We also put $\varepsilon^2 = 0.1$, $h_{100} = h_{101} = 0.5$, $\sigma = 73$ dyn/cm, and $\rho = 1$ g/cm^3. Then, for a cloud with a diameter of 5 km (as in [18]), the intensity of the electromagnetic background induced by the motion of the center of charge (with the excitation of the translational mode) of the drops when their surface vibrates because of physical microprocesses inside the cloud (coagulation with finer drops, evaporation, condensation, as well as hydrodynamic and electrical interaction with neighboring drops) is equal to $I \approx 10^{-6}$ W. The radiation frequency in this case is of the order of megahertz.

If we had assumed that neighboring modes in the range from k to $k + m$ are excited (instead of two modes), the integral intensity of the radiation would have increased in proportion to m.

(6) From expression (15), describing the shape of a vibrating drop at any instant of time, it is seen that, if two modes with sequential numbers are present in the mode spectrum specifying the shape of the initial perturbation, the mode with $n = 1$ corresponding to the translational motion appears in the second-order vibrational spectrum. The appearance of this mode is the result of the requirement for the center of mass of a vibrating drop to be at rest [see (11)]. The amplitude of the mode with $n = 1$ for several numbers of initially excited neighboring modes is plotted in Fig. 4.

In the case of a compressible environment (instead of a vacuum) with the density ρ_*, kinematic viscosity ν, and velocity of sound V, the mode with $n = 1$ due to nonlinear vibrations of the drop results in the generation of dipole acoustic waves with the frequencies $(\omega_i + \omega_{i+1})$ and $(\omega_i - \omega_{i+1})$. Note that the linear analysis of the acoustic radiation from vibrating drops gives multipole radiation (starting with quarupole one related to the fundamental mode $n = 2$). The dipole acoustic radiation associated with the drop is absent in the linear analysis. If the drop radius R is of the same order of magnitude as $(\nu/\omega_i)^{1/2}$, the expression for the intensity I_s of the dipole acoustic radiation from the drop has the form [20]

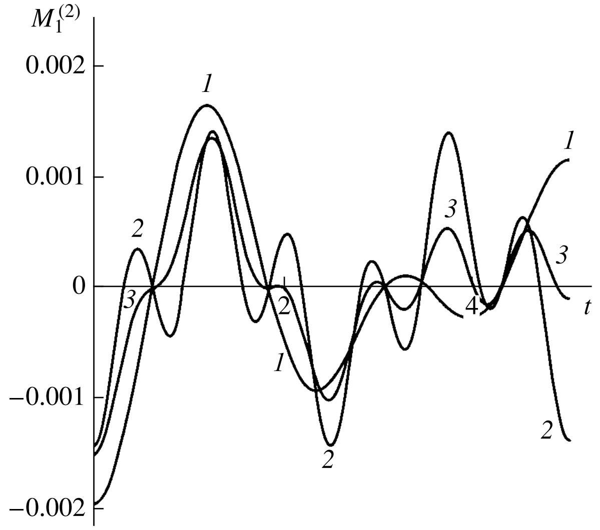


Fig. 4. Dimensionless amplitude $M_1^{(2)}$ of the first mode ($n = 1$) of the drop surface oscillations vs. dimensionless time t. Curves *1–3* show the data corresponding to the approximations in Fig. 3.

$$I_s \equiv \frac{3\pi\rho_* \nu^2 R^2}{2V^2}\{(\omega_i + \omega_{i+1})^2 |U_{\omega_i + \omega_{i+1}}|^2 + (\omega_{i+1} - \omega_i)^2 |U_{\omega_{i+1} - \omega_i}|^2\}, \tag{26}$$

where U_j is the peak velocity of the surface motion at the jth vibration frequency.

Assuming that the translational mode ($i = 1$) appears because of the presence of two neighboring modes (with numbers i and $i + 1$) in the spectrum of modes responsible for the initial deformation, we find, according to (15),

$$|U_{\omega_i + \omega_{i+1}}| \approx \varepsilon^2 R \frac{9ih_{i-1}h_i}{(2i-1)(2i+1)} h_i h_{i-1}(\omega_i + \omega_{i+1}), \tag{27}$$

$$|U_{\omega_i - \omega_{i+1}}| \approx \varepsilon^2 R \frac{9ih_{i-1}h_i}{(2i-1)(2i+1)} h_i h_{i-1}(\omega_i - \omega_{i+1}). \tag{28}$$

Substituting (27) and (28) into (26), we finally come to

$$I_s \equiv \frac{3\pi\rho_* \nu^2 R^2}{2V^2}\left\{\varepsilon^2 R \frac{9ih_{i-1}h_i}{(2i-1)(2i+1)} h_i h_{i-1}\right\}^2 \times \{(\omega_i + \omega_{i+1})^4 + (\omega_{i+1} - \omega_i)^4\}. \tag{29}$$

Let us estimate the intensity of the dipole acoustic radiation related to the translational mode in the spectrum of capillary surface vibrations for a rain drop with $R = 250$ μm when the second and third modes are excited. We put $\sigma = 73$ dyn/cm, $\rho = 1$ g/cm^3, $\rho_* = 1.3 \times 10^{-3}$ g/cm^3, $V = 3.3 \times 10^4$ cm/s, $W = 1$, $\varepsilon = 0.1$, $\nu = 0.15$ cm^2/s, and $h_2 = h_3 = 0.5$. From (29), it is easy to estimate the intensity of the dipole acoustic radiation of a single drop: $I_s \approx 10^{-13}$ erg/s. The integral power density of the acoustic radiation at frequencies of tens of kilohertz is ≈28 dB/km^3 at the boundary of the rain-filled space. Such a density corresponds to the sound intensity of human speech.

CONCLUSION

When the spectrum of modes responsible for the initial perturbation of the equilibrium spherical shape of the drop has two or more neighboring vibrations, the translational mode appears, as follows from asymptotical calculations up to second-order infinitesimals. The translational mode is responsible for dipole acoustic radiation from the drops. In the case of a charged drop, this causes dipole electromagnetic emission. These effects may be of significant importance in analyzing physical phenomena taking place in multiphase liquid-drop systems of both artificial and natural origin: clouds, fog, or rain.

ACKNOWLEDGMENTS

This work was supported by the Foundation of the President of the Russian Federation (project no. 00-15-9925).

REFERENCES

1. A. I. Grigor’ev and S. O. Shiryaeva, Izv. Akad. Nauk, Mekh. Zhidk. Gaza, No. 3, 3 (1994).
2. A. I. Grigor’ev, Zh. Tekh. Fiz. **70** (5), 22 (2000) [Tech. Phys. **45**, 543 (2000)].
3. D. F. Belonozhko and A. I. Grigor’ev, Elektrokhim. Obrab. Met., No. 4, 17 (2000).
4. J. A. Tsamopoulos and R. A. Brown, J. Fluid Mech. **147**, 373 (1984).
5. J. A. Tsamopoulos, T. R. Akylas, and R. A. Brown, Proc. R. Soc. London, Ser. A **401**, 67 (1985).
6. O. A. Basaran and L. E. Scriven, Phys. Fluids A **1**, 795 (1989).
7. R. Natarayan and R. A. Brown, Proc. R. Soc. London, Ser. A **410**, 209 (1987).
8. N. A. Pelekasis, J. A. Tsamopoulos, and G. D. Manolis, Phys. Fluids A **2**, 1328 (1990).
9. Z. Feng, J. Fluid Mech. **333**, 1 (1997).
10. S. O. Shiryaeva, A. I. Grigor’ev, and D. F. Belonozhko, Pis’ma Zh. Tekh. Fiz. **26** (19), 16 (2000) [Tech. Phys. Lett. **26**, 857 (2000)].
11. S. O. Shiryaeva, Pis’ma Zh. Tekh. Fiz. **26** (22), 76 (2000) [Tech. Phys. Lett. **26**, 1016 (2000)].
12. D. F. Belonozhko and A. I. Grigor’ev, Zh. Tekh. Fiz. **70** (8), 45 (2000) [Tech. Phys. **45**, 1001 (2000)].
13. S. O. Shiryaeva, Zh. Tekh. Fiz. **71** (2), 27 (2001) [Tech. Phys. **46**, 158 (2001)].
14. Z. C. Feng and L. G. Leal, Phys. Fluids **7**, 1325 (1995).
15. A.-H. Nayfeh, *Perturbation Methods* (Wiley, New York, 1973; Mir, Moscow, 1976).
16. D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii, *Quantum Theory of Angular Momentum* (Nauka, Leningrad, 1975; World Scientific, Singapore, 1988).
17. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 2: *The Classical Theory of Fields* (Nauka, Moscow, 1967; Pergamon, Oxford, 1971).
18. V. I. Kalechits, I. E. Nakhutin, and P. P. Poluéktov, Dokl. Akad. Nauk SSSR **262**, 1344 (1982).
19. *Clouds and Cloud Atmosphere: Handbook*, Ed. by I. P. Mazin, A. Kh. Khrgian, and I. M. Imyanitov (Gidrometeoizdat, Leningrad, 1989).
20. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics* (Nauka, Moscow, 1986; Pergamon, New York, 1987).

Translated by M.S. Fofanov

Technical Physics, Vol. 47, No. 4, 2002, pp. 397–405. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 23–31.
Original Russian Text Copyright © 2002 by Lapushkina, Bobashev, Vasil'eva, Erofeev, Ponyaev, Sakharov, Van Wie.

Influence of Electric and Magnetic Fields on the Shock Wave Configuration at the Diffuser Inlet

T. A. Lapushkina*, S. V. Bobashev*, R. V.Vasil'eva*, A. V. Erofeev*,
S. A. Ponyaev*, V. A. Sakharov*, and D. Van Wie**

** Physicotechnical Institute, Russian Academy of Sciences,*
Politekhnicheskaya ul. 26, St. Petersburg, 194021 Russia
e-mail: tanuysha@pop.ioffe.rssi.ru
*** Johns Hopkins University, Laurel, Maryland, 20723, USA*

Received September 20, 2001

Abstract—The possibility is investigated of influencing the shock wave configuration in a xenon plasma flow at the inlet of a supersonic diffuser by applying electric and magnetic fields. Flow patterns resulting from interaction of the plasma with external fields in the entire diffusor volume and its different sections are compared. The patterns are obtained by the Schlieren method using two recording regimes: individual frames or a succession of frames. The study focuses on the normal shock wave formation process under strong MHD interaction over the whole diffuser volume. Basic factors affecting the plasma flow velocity in the diffuser under externally applied fields are compared, namely, the ponderomotive force and the Joule heating of the gas by the electric field, which decelerate the supersonic flow, and the heat removal to the external electric circuit producing the opposite effect. It has been shown that the external fields are most effective if applied to the inlet part of the diffuser, while the flow in the diffuser section, where there is a large density of dissipative structures, is not readily responsive to external factors. It is suggested that the measure of response can be estimated by the energy that goes to the shock wave formation as a result of the flow interaction with the diffuser walls. *© 2002 MAIK "Nauka/Interperiodica".*

FORMULATION OF THE PROBLEM

The present work is a continuation of a series of investigations [1, 2] initiated by the AJAX project [3, 4] and aimed at studying the effect of external factors on the shock wave patterns arising in a diffuser with complete internal flow compression. The main part of the diffuser is a linearly converging duct formed by two wedge-shaped walls. The shock wave configuration consists of two attached shock fronts interacting with each other in the diffuser. An external electric field is applied in such a way that it enhances the current induced by the magnetic field. The flow is formed both as a result of the MHD interaction proper and due to heating of the working gas by the external electric field. Both these factors produce the same effect, that is, they decelerate the supersonic flow and accelerate the subsonic one. The experiment was carried out under conditions in which the effects of these two factors were comparable in strength.

The main objective of this work was to trace the changes in the shock wave configuration caused by the external fields and identify the processes and diffuser sections mainly responsible for changes in the flow pattern by comparing flow patterns obtained with the external factors applied over the whole diffuser volume or its particular sections and when only the electric field or both the electric and magnetic fields are applied. To solve this problem, the conditions in the diffuser under applied external fields were chosen in such a way as to have the flow of a strong interaction type in the diffuser as a whole and of a weak interaction type in its different parts. The strong interaction, according to the authors' classification [2], is that producing a normal shock wave in the diffuser due to MHD deceleration of the flow. The formation mechanism of this shock wave is a separate problem also considered in this paper.

EXPERIMENTAL

A schematic of the experimental setup was described in [1, 2]. The setup consists of a vacuum test section integrated with a shock tube. The test section includes a supersonic nozzle accelerating the plasma flow to $M_0 = 4.3$ and a diffuser with a set of electrodes connected as in a segmented Faraday generator. A pulsed magnetic field with a magnetic induction of up to 1.5 T is produced by Helmholtz coils 30 cm in diameter. A slug of the shock-heated gas becomes stagnant at the end of the tube facing the test section, heats up to a stagnation temperature, and flows through the inlet slit into the nozzle and then the diffuser. As a working medium, a xenon plasma is utilized. The initial xenon pressure in the low-pressure part of the shock tube is 20 torr and the Mach number of the incident shock wave is $M = 8.3$. Design parameters at the diffuser inlet are as follows: gas density $\rho_0 = 0.0985$ kg/m^3;

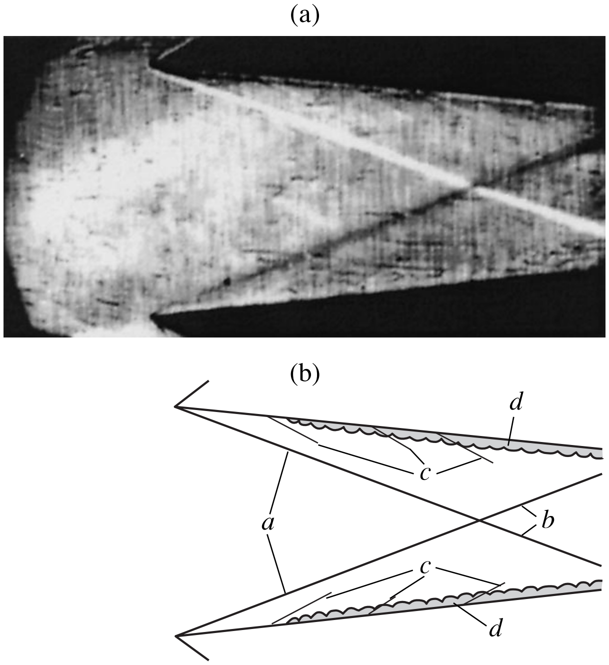


Fig. 1. A Schlieren pattern of the flow in the diffuser and a schematic diagram of gasdynamic nonuniformities without external fields.

concentration of heavy particles $n_0 = 4.5 \times 10^{23}$ m^{-3}; flow velocity $u_0 = 1.6 \times 10^3$ m/s; plasma conductivity $\sigma_0 = 800$ S/m; plasma efflux duration $t_A = 350$ μs; Reynolds number $\mathrm{Re} = (\rho_0 u_0 l)/\mu_0 = 10^6$, where l is the diffuser length and μ_0 is the coefficient of gas viscosity; and the initial magnitude of the Stuart parameter is $\mathrm{St}_0 = (\sigma B^2)/(\rho_0 u_0) = 5$ m^{-1} at $B = 1$ T.

For visualization of the shock wave configuration, two modifications of the Schlieren method are used. The first modification, described in [5], produces a single high-resolution flow pattern at a specified moment in time from the beginning of efflux. As an illuminant, an OGM-20 ruby laser is used with an exposure of 30 ns. In the second method, the development of the shock wave configuration throughout the efflux process is recorded as a succession of frames. To this end, the illuminating and recording units of the Schlieren optical system were modified. In this regime a Podmoshenskiĭ light source capable of maintaining a steady illumination of high intensity for 500 μs was used as an illuminant. For recording we used a VSK-5 high-speed cinecamera, allowing one to take up to 130 frames at intervals from 4 to 25 μs. For integrating the cinecamera with the Schlierenoptics, the viewing-focusing system of the camera was modified to provide illumination of the entire frame by the probing light beam and accurate focusing of the subject under study on the recording film.

Figure 1 shows a Schlieren pattern (a) of a flow in the diffuser obtained in the single-frame recording regime without external fields, as well as (b) a scheme of discontinuities formed. The picture clearly shows input shock fronts *a* and their interaction point located at a distance of $X_c = (44 \pm 1.5)$ mm from the diffuser inlet. This characteristic length will be used in what follows for evaluating the influence of the external fields on the flow structure. The results of the reflection of the attached shocks fronts from one another are the shock fronts (*b*). It should be noted that, in this case, the near-wall layer (*d*) registered poorly and its position in the explanatory scheme is indicated rather arbitrarily. Additional discontinuities (*c*) seen in the picture arise from the roughness of the diffuser walls.

In this setup, because of a large voltage drop across the thick near-wall layers, potential differences had to be applied between the electrodes using an external source of electric power in order to achieve current densities required for MHD interaction. Application of the external voltage was implemented by discharging long lines in a circuit consisting of a plasma gap between pairs of electrodes and a load resistance of 0.1 Ω.

Across pairs of electrodes connected to the long lines, the same voltages of 110 V were applied. The Ohm law for a circuit including a plasma gap between two electrodes, a load resistor, an MHD generator with the electromotive force $\varepsilon = uBh$, and a source of external voltage V has the form

$$V + uBh = I(R_{\mathrm{eff}} + R_L), \qquad (1)$$

where V is the voltage applied to the electrodes; u is the flow velocity; B is the magnetic induction; h is the distance between the electrodes in a pair; I is the current; R_{eff} is the effective internal plasma resistance including a resistance of the flow core and a resistance of the near-wall layer; and R_L is the load resistance.

We studied the flow interaction with external fields in three (overlapping) regions (see Fig. 2) by passing the current between appropriate electrodes. As seen in Fig. 2, with connection *I*, the voltage is applied to pairs of electrodes in the diffuser from the third to the seventh. In this case, the interaction occurs in the whole diffuser volume. The interaction zone length is roughly 90 mm. With connection *II*, the voltage is applied only to electrode pairs from the fourth to the seventh, so that the diffuser inlet is excluded from the interaction zone; the interaction zone length in this case is about 70 mm. With connection *III*, the voltage is applied to the third pair of electrodes, thus confining the interaction to the diffuser inlet; the interaction zone length is approximately 20 mm.

Figure 3 demonstrates current density distributions along the duct for the three schemes of applying the external voltage in the case of MHD interaction and in zero external magnetic field. The current density was determined as the ratio of the current measured to the area of the electrode and the interelectrode gap. The zero of the abscissa corresponds to the front of the diffuser. In the case of connections types *I* and *II*, the average current densities are plotted at the midpoints of the electrodes; for connection *III*, the current density is plotted over the whole width of the electrode. As seen in the plots, the electromotive force induced by the magnetic field, contrary to expectations, did not increase the current density. To all appearances, the reasons are the Hall effect arising because of imperfect sectioning of the electrodes and the increase of the near-wall layer width as a result of MHD interaction. Using *I*–*V* characteristics, we evaluated the magnitudes of R_{eff}. At $B = 0$ $R_{\text{eff}} \approx 0.15\ \Omega$ and at $B = 1.3$ T $R_{\text{eff}} = 0.2$–$0.4\ \Omega$. Accordingly, the load factor $k = R_L/(R_L + R_{\text{eff}})$ varies in the range from 0.2–0.4.

COMPARATIVE ESTIMATIONS OF THE INFLUENCE OF ELECTRIC AND MAGNETIC FIELDS ON THE FLOW PARAMETERS

Elementary impacts on the flow are those of force and energy. The principle force on the flow is exerted by a ponderomotive force $F = jB$. The power developed by the ponderomotive force is $N_1 = juB$.

The energy impact has two components. The first component is the removal to the external electric circuit of an electric power $N_2 = kjuB$ generated through MHD interaction. The second component is the heat supplied from an external source of a power $N_3 = (1 - k)jV/h$. The work done by the ponderomotive force and the heat supplied decelerate the supersonic flow, while the heat removal accelerates it. To evaluate and compare the actions of these factors on the flow parameters, we use a set of equations [6] describing an MHD interaction. An expression for variation of the flow Mach number M along the duct has the form

$$(M^2-1)\frac{d\ln M^2}{dx} = -[2+(\gamma-1)M^2] \times\left[\frac{d\ln(\rho u)}{dx} - \frac{\gamma}{a^2}\frac{jBu}{\rho u}\right] - \frac{(\gamma-1)(1+\gamma M^2)}{a^2}\left[k\frac{juB}{\rho u} + (1-k)\frac{jV}{h\rho u}\right]. \tag{2}$$

This equation can be expressed in the form of a reversal condition for the effect of external factors on subsonic and supersonic flows [6],

$$(M^2-1)\frac{dM^2}{dx} = \sum_k r_k N_k, \tag{3}$$

where N_k are the types of external factors and r_k are coefficients corresponding to them; here we consider the factors of the aforementioned types N_1, N_2, and N_3 and add to them the geometrical factors $N_4 = (d\ln(\rho u))/dx$.

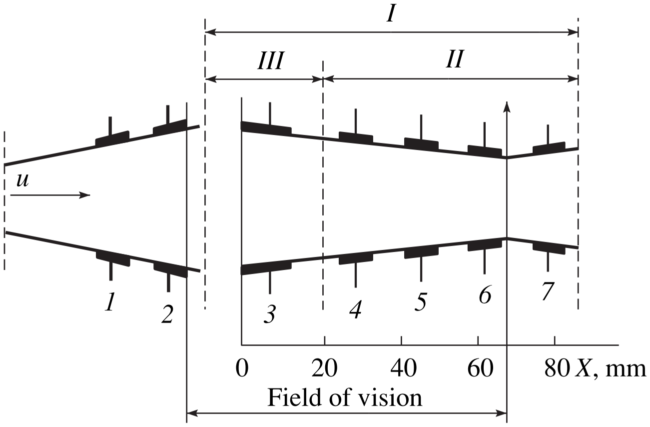


Fig. 2. Schematic of the channel with marked interaction zones for three types of connection to the external voltage.

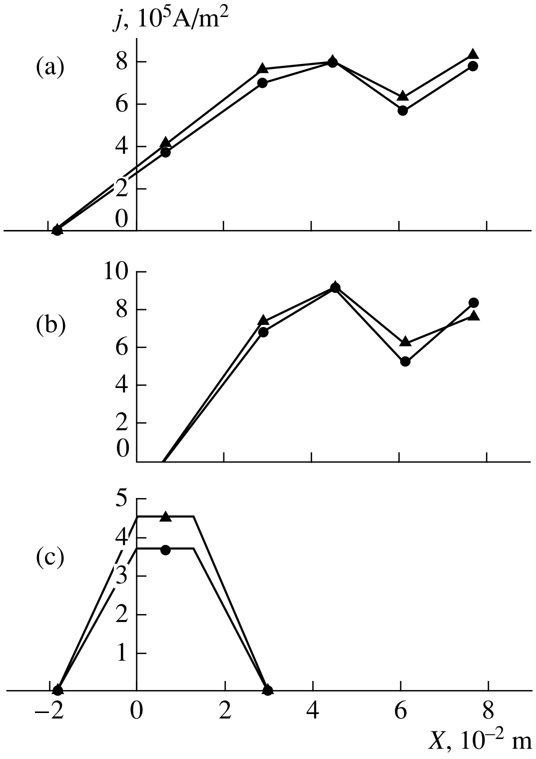


Fig. 3. Current density distribution along the duct. Connections: (a) *I*; (b) *II*; (c) *III*. $B = 0$ (triangles) and 1.3 T (circles).

Using Eq. (3), one can compare the effects of external factors on the flow. For example, it is possible to determine at what values of the ratio between the external electric field and an electric field induced by the magnetic field a particular factor—gas heating or ponderomotive force—is predominant in decelerating the flow. The condition of equality of the actions of these two factors has the form

$$N_1 r_1 = N_3 r_3 \quad \text{or} \quad \frac{uBh}{V} = \frac{(\gamma-1)(1+\gamma M^2)(1-k)}{\gamma[2+(\gamma-1)M^2]}. \tag{4}$$

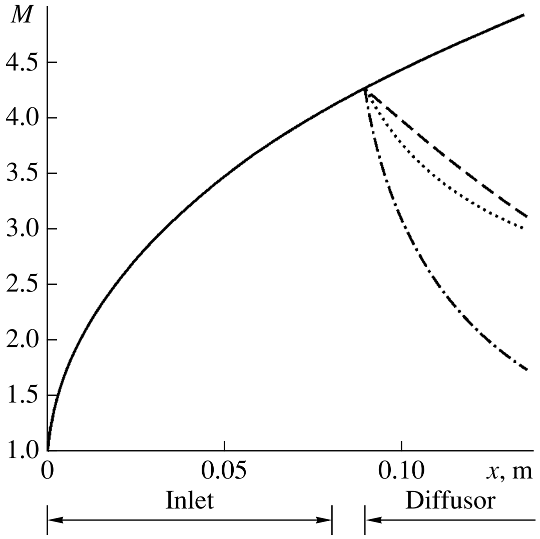


Fig. 4. Variation of the flow Mach number along the axis of the gasdynamic path under the action of magnetic and electric fields. $B = 0$, $V = 0$ (solid line); $B = 1.3$ T, $V = 0$ (dashed line); $B = 0$, $V = 110$ V (dotted line); $B = 1.3$ T, $V = 110$ V (dotted-dashed line).

Consider a case of $k = 0$. For flows with large Mach numbers ($M > 5$) $(uBh)/V \approx 1$, this means that, at $(uBh)/V > 1$, the dominant factor in the flow deceleration is the ponderomotive force, and, at $(uBh)/V < 1$, the major role is played by the Joule heating caused by the external electric field. As the Mach number decreases, the role of the ponderomotive force increases; for example, for $(uBh)/V > 0.4$, it is dominant at $M = 1$, becoming more so at still lower Mach numbers. As seen from (2), the relative role of the ponderomotive force grows with increasing load factor, and at $k \longrightarrow 1$ the flow deceleration is mainly determined by the MHD interaction in the flow.

Figure 4 shows variation of the calculated Mach number along the axis of the gasdynamic path under the action of magnetic and electric fields. The interaction zone occupies the length of the diffuser from the inlet to a point where the attached shock fronts intersect. The calculation results are obtained by solving a set of equations describing variation of the flow velocity, the state variables, and the Mach number and by using the Ohm law [6] for conditions close to those of the present experiment. In the calculation, it was assumed that $X = 0$ corresponds to the critical cross section; $T_{cr} = 7700$ K; $\rho_{cr} = 2.2$ kg/m^3; and $a_{cr} = 940$ m/s, where the subscript a_{cr} refers to the critical cross section parameters. For conductivity in the interaction zone, we used the effective conductivity of $\sigma_{eff} = 150$ S/m obtained in the experiment. The load factor was $k = 0.3$.

In Fig. 4, it is seen that the electric and magnetic fields singly applied affect the flow noticeably and decrease the Mach number by about 25%. Simultaneous application of the external fields, as seen from the calculation data, leads to a drastic decrease in the Mach number over the interaction zone. This may cause, as shown in [7], the formation of a shock wave of MHD deceleration, which transforms a supersonic flow into a subsonic one. However, the variation of the gas parameters along the diffuser axis is hardly an adequate flow characteristic for the diffuser as a whole. An answer to the question of how the flow pattern changes under the influence of external fields should be obtained from experiment.

INTERACTION WITH EXTERNAL FIELDS IN THE WHOLE DIFFUSER VOLUME: CONNECTION *I*

Figure 5 demonstrates flow patterns in the diffuser obtained by the Schlieren method in the single-frame recording regime, as well as schematic diagrams of the shock wave structure. Figure 5.1 shows a flow structure formed under the action of only the electric field. Clearly seen here are the attached shocks *a*, shocks *b* resulting from interaction between the attached shocks, and shocks *f* representing reflections of shocks *b* from the walls or the near-wall layers *d*. The near-wall layers can be clearly distinguished in the Schlieren patterns. It is seen that the layer thickness increases downstream and is considerably larger than it would be without passing the electric current (compare with Fig. 1). The near-wall layer produces a shock wave *l* clearly seen by the lower wall. Gasdynamic discontinuity *e* we identify as a contact surface formed in the region beyond the attached shocks because the gas flow at the diffuser inlet is diverging and therefore the gas passing through different parts of the attached shock front has different densities, Mach numbers, and incidence angles. All this results in a large density gradient, which can be seen in the density field calculated in [8]. Probably, the electric field enhances this gradient, producing the contact surface *e*. In comparison with the flow pattern without external fields (see Fig. 1), the basic distinction of the flow pattern observed with an applied external electric field is that in the latter case the slope angles of the attached shock fronts increase, the shock front intersection point X_c shifts nearer the diffuser inlet, $X_c = (37 \pm 1.5)$ mm, and the points of the shock front reflection from the diffuser walls become closer to the inlet. Thus, it is possible to conclude that the flow pattern formed in the electric field is similar to that emerging under a weak MHD interaction as described in the numerical simulation of the process [8] and in the experiment [2].

Fig. 5.2a shows Schlieren flow pattern at a fixed moment of time, and Fig. 5.2b shows a scheme of the basic gasdynamic nonuniformities observed in the diffuser under the joint action of the electric and magnetic fields. It is seen that the flow pattern is appreciably different. First of all, attention is attracted by a highly developed near-wall layer, especially near the upper wall, where the cathodes are located. Radical changes occurred in the configuration of the inlet shock fronts.

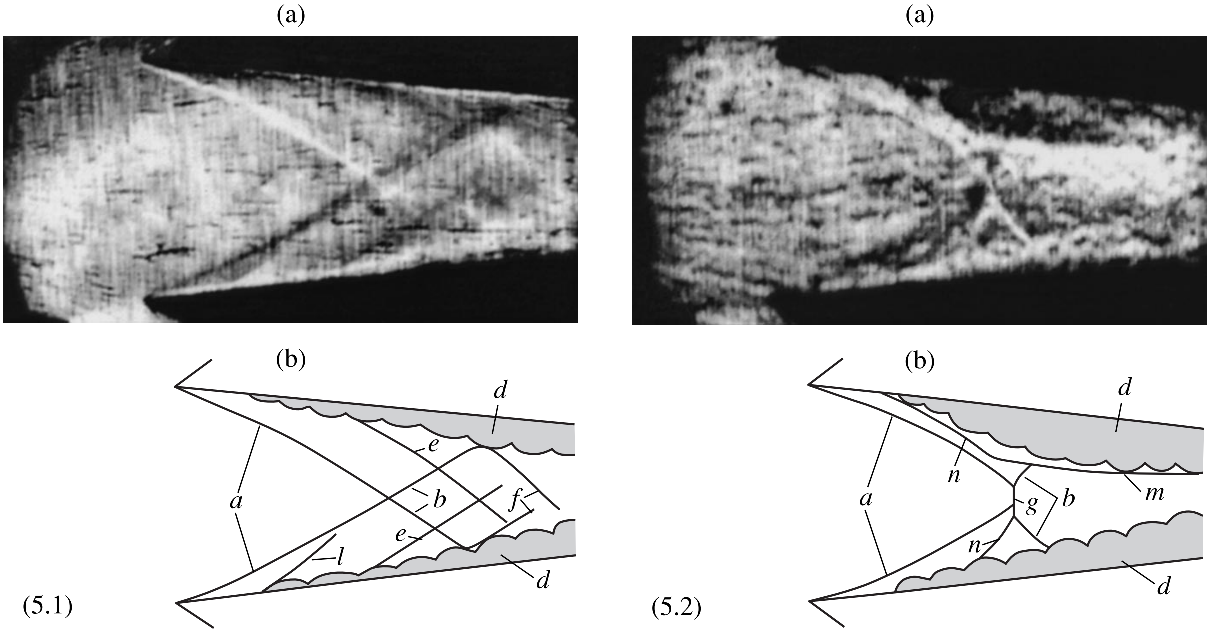


Fig. 5. Connection *I*. (a) Schlieren patterns and (b) schematic diagrams of the flow. B = (5.1) 0 and (5.2) 1.3 T.

At the place where oblique shock fronts intersected before, now we see a triple shock configuration consisting of a normal shock *g* in the flow core and two oblique shocks *a* and *b*. Between them one can also distinguish discontinuities *n*; however, it is very difficult to conclude if they are the same contact surfaces *e* or shock waves due to the near-wall layer. Also, we have not been able so far to identify discontinuity *m*, which can be seen by the upper wall. On the whole, the flow can be characterized as a typical case of strong interaction with a normal shock front of MHD deceleration formed in the flow core; however, the flow pattern is much complicated by the highly developed near-wall layer.

Dynamics of the formation process of the MHD deceleration shock front is illustrated in Fig. 6, which shows a selection of Schlieren patterns obtained with the help of the high-speed cinecamera. The exposure time is 2 μs, and the time interval between the frames is 5.7 μs. Frames 2, 6, and 13 show how the gas enters the diffuser. Seen in frame 22 are the already developed near-wall layer and a second shock front near the upper wall, which is probably caused by the near-wall layer. The subsequent frames demonstrate that all the changes observed in the flow happen to the attached shock waves. Schematically, this transformation is shown in Fig. 7. At some moment of time t_1, a regular reflection of the attached shock fronts from each other takes place. The pattern changes with time, and at the moment t_2, the lower shock front becomes noticeably concave, while the shape of the upper shock front is difficult to judge because it is distorted by the highly developed near-wall layer. However, despite the concavity, which makes the oblique shock front look more like a normal one at the core of the flow, the reflection remains regular. Then, at the moment t_3, the regular reflection of the shock fronts *a* from one another transforms into a Mach reflection. The Mach wave *g* is formed with the reflected shock front *b* attached to it. Thus, a triple configuration arises (see Fig. 5.2b). It should be noted that such a process of flow formation was unexpected. Rather, it was believed that the highest pressure will build up at the end of the interaction zone and the compression waves will produce an upstream shock wave changing the inlet shock wave configuration. The experiment has revealed quite a different way of the formation of the MHD deceleration shock front.

INTERACTION IN VARIOUS DIFFUSER SECTIONS: CONNECTION *II*

To find a way of a more effective influence on the shock wave configuration, let us consider how it changes when an electric current is passed through various diffuser sections. Figure 8 demonstrates the flow structure for connection *II*. An external voltage is applied to all of the electrodes except the first pair. A Schlieren pattern obtained with only the applications of electric field is presented in Fig. 8.1a. A schematic diagram of the shock fronts observed is given in Fig. 8.1b. It is seen that the slope angles of the attached shock fronts did not change much as compared to Fig. 1, and the distance X_c remained practically the same. This is evidence that, in this case, the current flowing through the plasma has little effect on the attached shock fronts *a* formed at the diffuser inlet. The positions of shock fronts *b* changed only slightly; their slope angles increased to such an extent that shock

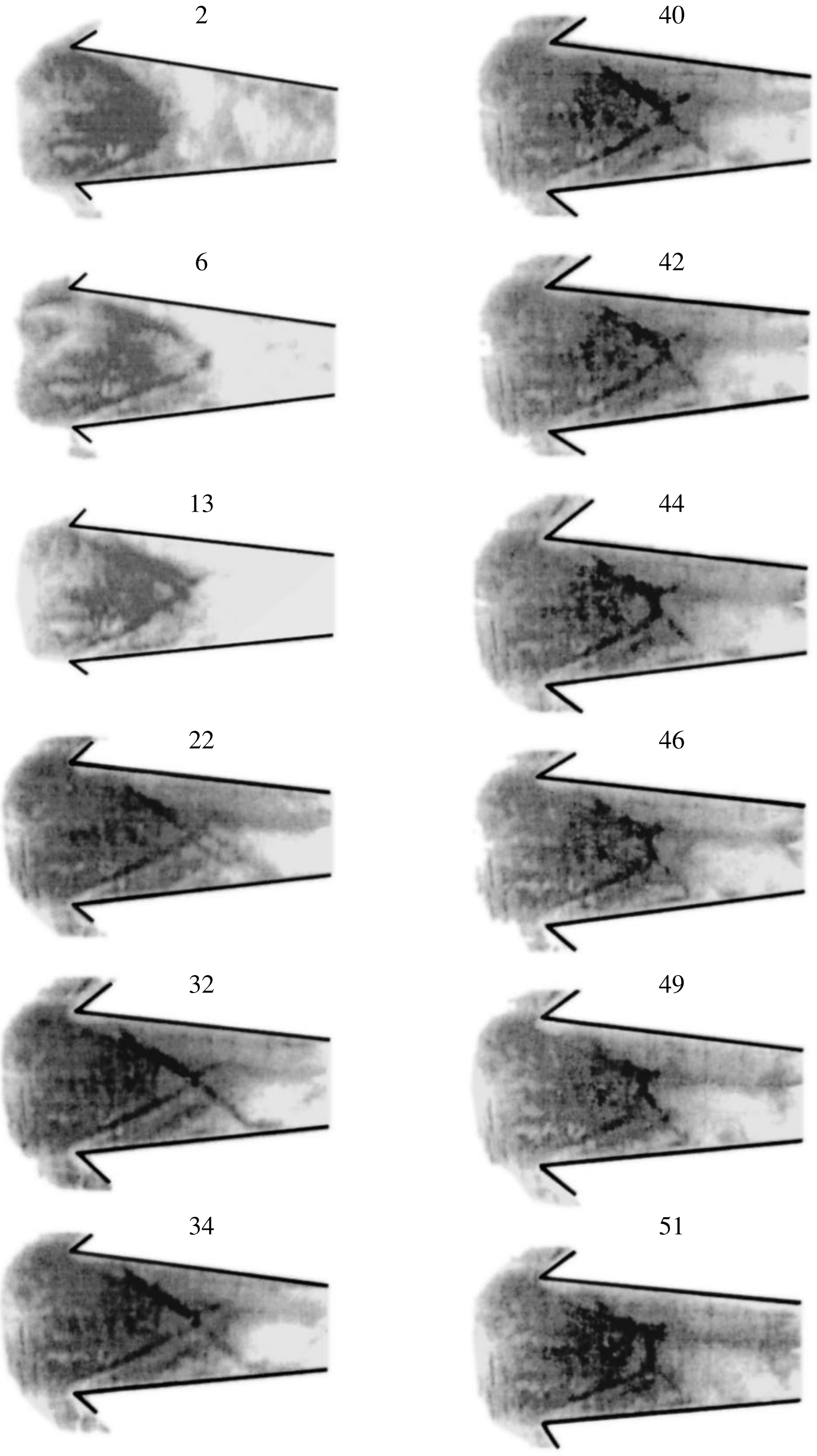


Fig. 6. Connection *I*. Frame-by-frame recording of Schlieren flow patterns. The frames are identified by their numbers.

fronts *f* reflected from the near-wall layer became visible. The current caused a noticeable growth of the near-wall layer starting near the first pair of electrodes in this connection.

Application of the magnetic field in addition to the electric field did not cause noticeable shifts of the attached shock fronts (see Figs. 8.2a and 8.2b), the only observable effect being that their slight bending of contact surface *e* became more clearly seen. In addition, close to the still more developed near-wall layer, gasdynamic discontinuity *m* was observed.

CONNECTION *III*

When an external voltage was applied only to the first pair of electrodes located near the inlet diffuser edges (connection *III*), a shock wave configuration arose, which can be seen in Fig. 9. Passing the electric current without applying the magnetic field (Figs. 9.1a and 9.1b) noticeably changed the slope of the shock fronts, and their intersection point shifted toward the diffuser inlet, $X_c = (36 \pm 1.5)$ mm, and a marked growth of the near-wall layer was observed beginning right at the diffuser inlet. With an onset of the MHD interaction (Figs. 9.2a and 9.2b), no further change of X_c occurred; however, the attached shock fronts became slightly concave, and the wide near-wall layers could be clearly seen, the upper layer being appreciably thicker than the lower one. Contact surface *e* was clearly discernible.

Many more details of the gasdynamic structure were brought out in an experiment carried out with a gas whose density was twice as high. These experimental results are shown in Figs. 10a and 10b. It is seen that, under the joint action of the electric and magnetic fields, changes in the shape of the attached shock fronts are accompanied by additional disturbances in the vicinity of the attached shock fronts.

DISCUSSION OF THE RESULTS

The experimental data show that connection *II*, in spite of creating a more extensive interaction region, has a weaker influence on the shock wave configuration in the diffuser than connection *III*. Let us consider all the factors affecting the flow under applied external fields. This will allow us to select a diffuser section for more effective local impact on the inlet shock fronts.

To compare quantitatively the energy expended for variation of the flow parameters with the two connection types, *II* and *III*, let us evaluate the work of the ponderomotive force and the energy spent for heating the gas by the external electric field. Energy pick-up, according to estimates, plays a minor role, and therefore it is not considered here. The work of the ponderomotive force over the length L is equal to $A = jBL$ and was determined from the experimental data. As a measure of the influence of this work on the flow, the Stuart parameter $\mathrm{St} = jBL/\langle \rho u^2 \rangle$ was used, which is defined as the ratio between the work of the ponderomotive force and twice the average kinetic energy of the flow in the diffuser volume over the length L at $V = 0$ and $B = 0$. From the kinetic energy estimated on the basis of calculation data, we obtain Stuart parameter values of 0.3 for connection *II* and 0.04 for connection *III*. St_{II} is almost an order of magnitude larger than St_{III}, which is evidence of a stronger MHD interaction in the second duct section.

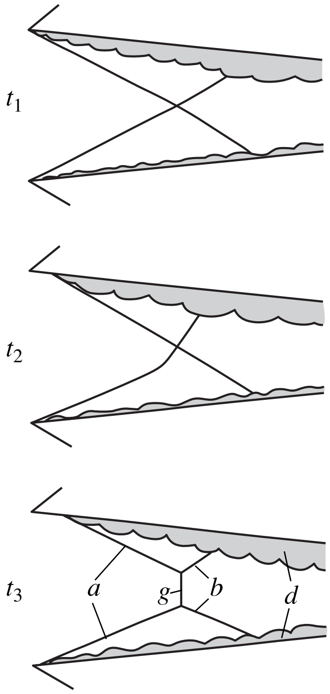


Fig. 7. Schematic diagram of the MHD deceleration shock formation.

Contributions to the flow deceleration of the plasma heating by the electric field for the two connections can be compared if one evaluates the ratio of the amount of Joule heat,

$$Q_c = (1-k)\frac{jLV}{uh},$$

to the flow enthalpy $H = \langle \rho(u^2 + C_pT) \rangle$ averaged over the time $\Delta t = L/u$, that is, the quantity $P = Q_c/H$. Here, C_p is the specific heat of the gas at constant pressure. Using the calculated data for the velocity and temperature of the downstream flow, we obtain the following values of the parameter P for the two interaction zones: $P_{II} = 0.6$ and $P_{III} = 0.07$. It is seen that the effect on the flow of the Joule heating, like that of the ponderomotive force, should be greater if the external voltage is applied using connection *II*. However, comparison of the flow fields in regions *II* and *III* (see Figs. 8 and 9) shows that in region *II*, where the applied impact is greater, changes in the positions of the shock fronts are substantially less than in region *III*, where a weaker influence significantly changes the slope and shape of the attached shock fronts. Thus, external voltage affects the attached shock fronts more significantly if applied to the diffuser inlet section.

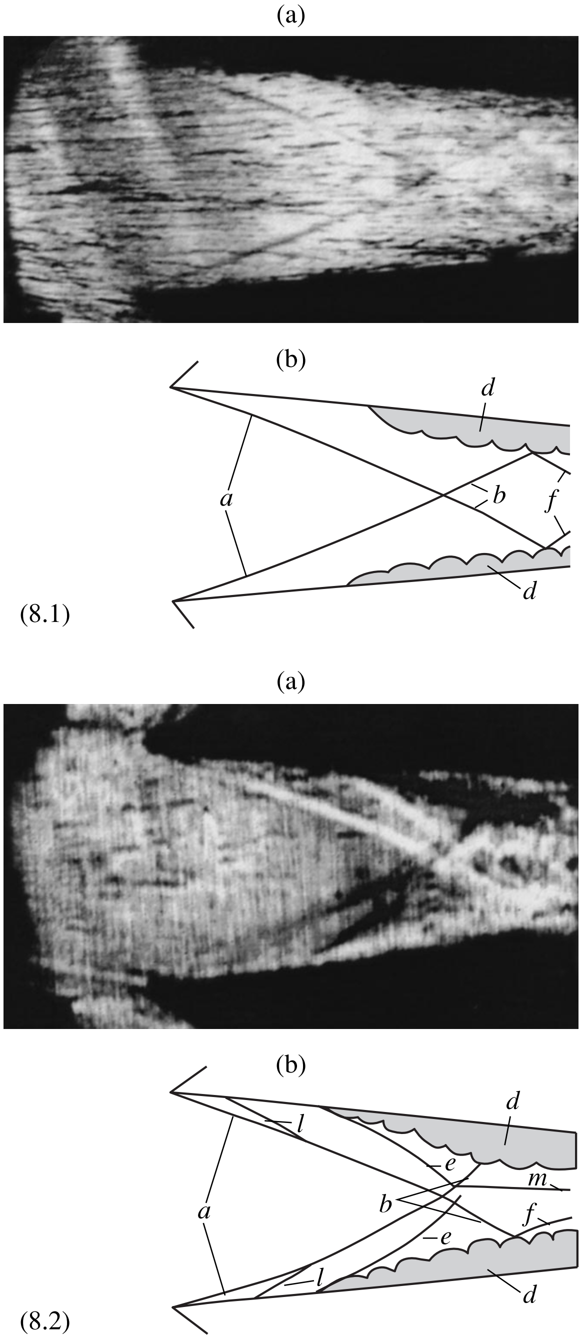


Fig. 8. Connection *II*. (a) Schlieren patterns and (b) schematic diagrams of the flow. B = (8.1) 0 and (8.2) 1.3 T

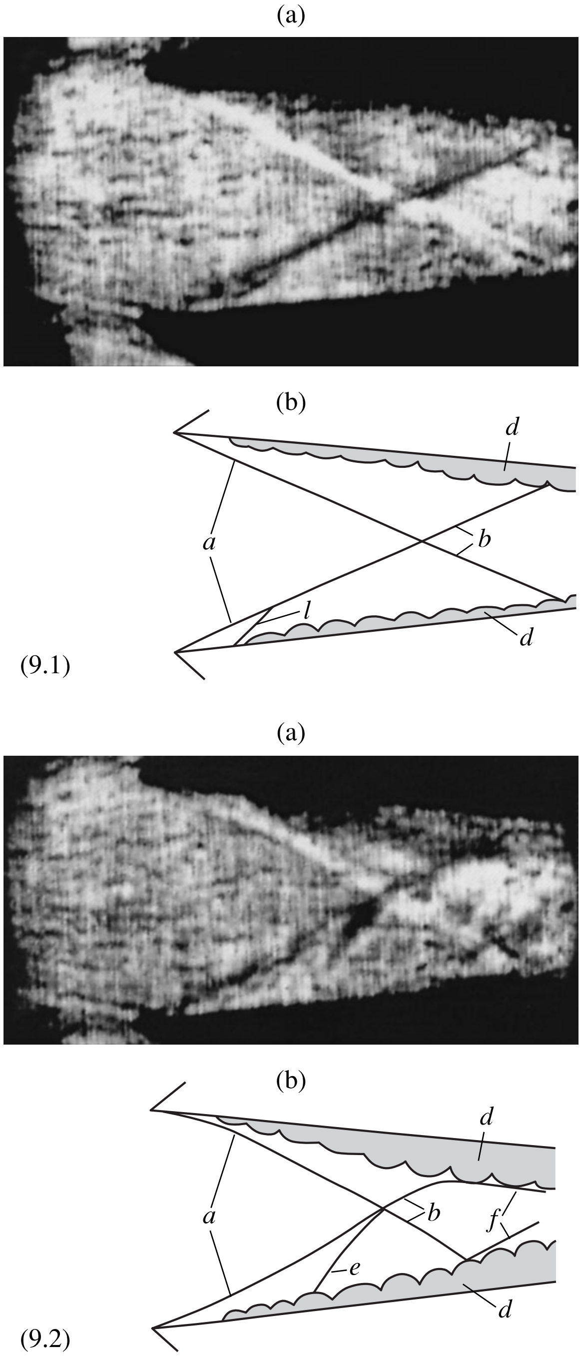


Fig. 9. Connection *III*. (a) Schlieren patterns and (b) schematic diagrams of the flow. B = (9.1) 0 and (9.2) 1.3 T.

In connection to this, differences in the flow structure in regions *II* and *III* should be considered. Only half of the attached shock fronts are found in region *III*, most of the flow volume here being occupied by a continuous flow. In region *II*, where the density of dissipative structures is greater, the attached shock fronts persist; here they are reflected from one another, and practically all the reflected shock fronts are found in this region. As the experiment has shown, the system of shock fronts in region *II* is less responsive to external influences; that is, to change the flow structure shocks, a stronger impact is necessary than in the case of continuous flow. It appears natural to relate the observed unresponsiveness of the shocks to the energy taken from the flow for their formation in the course of interaction with the diffuser walls. The irreversible energy

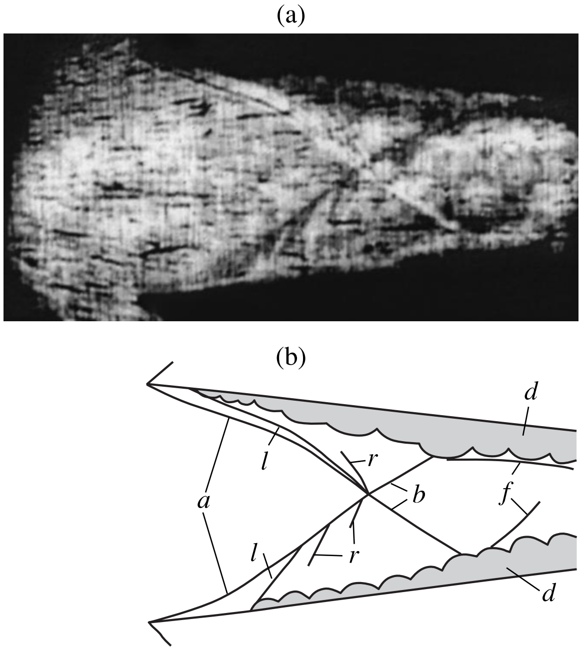


Fig. 10. A Schlieren pattern of the flow and a schematic diagram of gasdynamic discontinuities under the action of electric and magnetic fields at a gas density of ρ_0 = 0.196 g/m^3.

loss Q to a shock is given by the change of the entropy S across the shock is as follows:

$$Q = \int T dS = \langle T \rangle \Delta S,$$
$$\Delta S = S_2 - S_1 = C_v \ln \frac{p_2 \rho_1^\gamma}{p_1 \rho_2^\gamma}, \qquad (5)$$

where indices 1 and 2 denote parameters in front of and behind the shock, respectively.

Quantitatively, this energy loss is about two orders of magnitude less than the total flow enthalpy. A question arises whether the ratio of the energy lost to all shocks in the interaction zone and the energy supplied by the external fields can serve as a measure of the response of the shocks to the external factors aimed at changing their positions.

CONCLUSION

The major results of this study are as follows. It has been found that the near-wall layer has a strong influence on the flow pattern in the diffuser. If the interaction is strong, a shock wave of MHD deceleration is produced due to transition from the regular reflection of the attached shock fronts to the Mach reflection. It has been ascertained that, for effective control of the attached shock fronts, the external impacts should be applied to the diffuser inlet section. The flow containing shock waves is no less responsive to externally applied factors. It is suggested that this effect is associated with the amount of energy that goes to the shock wave formation as a result of the flow interaction with the diffuser walls.

ACKNOWLEDGMENTS

The authors express their sincere gratitude to B.G. Zhukov for helpful discussions. The study was performed with the financial support of the ISTC.

REFERENCES

1. S. V. Bobashev, R. V. Vasil’eva, E. A. D’yakonova, *et al.*, Pis’ma Zh. Tekh. Fiz. **27** (2), 63 (2001) [Tech. Phys. Lett. **27**, 71 (2001)].
2. S. V. Bobashev, E. A. D’yakonova, A. V. Erofeev, *et al.*, in *Proceedings of the 2nd Workshop on Magneto–Plasma–Aerodynamics in Aerospace Applications, Moscow, 2000*, p. 64.
3. V. L. Fraĭshtadt, A. L. Kuranov, and E. G. Sheĭkin, Zh. Tekh. Fiz. **68** (11), 43 (1998) [Tech. Phys. **43**, 1309 (1998)].
4. E. P. Gurijanov and P. T. Harada, AIAA Pap., No. 96-4609 (1996); in *Proceedings of the 7th Aerospace Planes and Hypersonic Technology Conference, Norfolk, 1996.*
5. S. V. Bobashev, E. A. D’yakonova, A. V. Erofeev, *et al.*, AIAA Pap., No. 2000-2647 (2000).
6. L. A. Vulis, A. L. Genkin, and V. A. Fomenko, *Theory and Calculation of Magnetohydrodynamic Stream in Channels* (Atomizdat, Moscow, 1971).
7. R. V. Vasil’eva, A. V. Erofeev, *et al.*, Zh. Tekh. Fiz. **57** (2), 251 (1987) [Sov. Phys. Tech. Phys. **32**, 150 (1987)].
8. Yu. P. Golovachev, Yu. A. Kurakin, A. A. Schmidt, and D. M. Van Wie, AIAA Pap., No. 2001-2883 (2001).

Translated by N. Mende

Technical Physics, Vol. 47, No. 4, 2002, pp. 406–409. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 32–35.
Original Russian Text Copyright © 2002 by Shuaibov, Dashchenko, Shevera.

GAS DISCHARGES, PLASMA

Optical Characteristics of the Steady-State Plasma of a Longitudinal Discharge in a He/CFC-12 Mixture

A. K. Shuaibov, A. I. Dashchenko, and I. V. Shevera
Uzhhorod National University, Uzhhorod, 88000 Ukraine
e-mail: ishev@univ.uzhgorod.ua

Received September 28, 2001

Abstract—The emission from the plasma of a contracted longitudinal dc discharge in a He/CF_2Cl_2 mixture in the wavelength range 130–300 nm is investigated. It is shown that the discharge plasma emits within the range 150–260 nm. The emission consists mainly of the broad bands of Cl_2 molecules and single-charged chlorine ions. The pressure and composition of the working mixtures, the discharge current, and the time during which the emitter can operate on a single gas fill are optimized to attain the best characteristics of UV and VUV radiation. The results obtained are of interest for developing a steady-state source of VUV and UV radiation to be applied in microelectronics, photochemistry, and medicine. *© 2002 MAIK “Nauka/Interperiodica”.*

A dc glow discharge in the mixtures of Xe, Kr, and Ar with chlorine has been successively used as a source of VUV and UV emission from rare gas chlorides with an average radiation power of 3–20 W and efficiency of 8–23% [1–3]. In [2], it was also found that He/Cl_2 mixtures can be employed in low-pressure electric-discharge lamps, which operate on the $Cl_2(D'–A')$ 258-nm band and whose output parameters are close to those of the corresponding excimer lamps. An advantage of such a UV source is that its working mixture does not contain expensive rare gases like Xe or Kr. This provides the possibility of creating a lamp with a longitudinal circulation of the working mixture, where the problem of the gas mixture lifetime becomes unimportant. Lamps of this kind are characterized by the elevated output radiation powers and efficiencies. In [4], we optimized the optical characteristics of a compact emitter operating on a He/Cl_2 mixture pumped by a longitudinal dc discharge. It was found that the emission spectrum of this lamp can be extended to 195–200 nm by operating on the $Cl_2(^1\Sigma–^1\Pi_u)$ bands and using an outer lamp envelope transparent in the wavelength range $\lambda \leq 200$ nm. Some medical and industrial applications (sterilization of medical instrumentation, food conservation, etc.) require powerful sources of short-wavelength radiation in the range 170–220 nm. Thus, in order to further extend the emission spectrum to the VUV spectral range, it is of importance to continue the experiments of [4] with another halogen-containing agent (CFC-12), whose dissociation products emit in the range 150–180 nm. It was found in [5, 6] that the emission from the plasma of a transverse volume discharge in He/Cl_2 and Ar/CF_2Cl_2 mixtures contains VUV radiation from the products of dissociation of chlorine and CFC-12 molecules, which can also be applied to develop a dc emission source operating on a He/CF_2Cl_2 mixture.

In this paper, we present the results of the studies of VUV and UV emission from a longitudinal discharge in a low-pressure He/CF_2Cl_2 mixture.

A longitudinal dc discharge was ignited in a quartz tube with an inner diameter of 5 mm and interelectrode distance of 100 mm. Cylindrical electrodes made of Ni foil were placed in the discharge tube with open ends. The discharge tube was housed in a 10-l buffer chamber connected to a 0.5-m-long vacuum monochromator via a LiF window. The open end of the discharge tube was in front of the entrance slit of the vacuum monochromator. The spectrometer was calibrated in the wavelength range 165–350 nm with the help of continuous hydrogen emission; then, the curve obtained was extrapolated to $\lambda = 130$ nm. The system for recording the characteristics of the longitudinal discharge was identical to that described in [3, 4].

In experiments, we studied the current–voltage characteristics, the power deposited in the discharge plasma, the emission spectra in the 130- to 300-nm wavelength range, the intensities of the emission bands as functions of the pressure and composition of the working mixture, the dependences of the maximum intensities in different emission bands on the discharge current, and the emitter lifetime. The emission band intensity was determined from the area under the corresponding curve on the chart recorder paper with allowance for the relative spectral sensitivity of the vacuum spectrometer. At a CFC-12 pressure higher than 0.1 kPa, a longitudinal dc discharge in pure CFC-12 and He/CFC-12 mixtures becomes contracted. As the CFC-12 partial pressure and the total pressure of the He/CFC-12 mixture increases, the diameter of the plasma filament separated from the inner wall of a discharge tube decreases from 2 to 1 mm. Although the discharge contraction leads to a decrease in the area of the working surface of the cylindrical emitter, it is still of interest from the point of view of increasing the time during which the emitter can operate on a single gas fill. The reason is that the service life of chlorine-containing devices strongly depends on the coefficient of chlorine diffusion into glass, which significantly increases with the temperature of the discharge tube [7]. A contracted longitudinal discharge is the simplest emitter, in which the plasma does not contact the glass envelope; thus, in this respect, it is similar to an excimer lamp with microwave pumping [8]. The ignition of a usual dc glow discharge in CFC-containing mixtures at CFC partial pressures higher than 100 Pa is questionable. In [9], such a discharge was ignited in CCl_4 vapor, but only when the gas medium was excited with a fast longitudinal ionization wave. The ion–ion plasma of a contracted discharge in electronegative gases significantly differs from the common electron–ion plasma [10, 11]; this fact attracts additional interest to studying these discharges. In the short-wavelength spectral range, the optical characteristics of an ion–ion plasma formed of CFC-12 molecules were not investigated, which impedes the development of steady-state sources of VUV–UV radiation.

Figure 1a presents the current–voltage characteristics of a longitudinal discharge in a He/CFC-12 mixture. At a discharge current of I_{ch} = 1–15 mA, the discharge voltage (U_{ch}) is inversely proportional to the discharge current, which is characteristic of the subnormal mode of discharge operation [12]. At $I_{ch} \geq 15$ mA, the longitudinal discharge switches to the normal mode, in which U_{ch} changes only slightly as the current increases. The increase in the CFC-12 content and the pressure of the He/CFC-12 mixture leads to a sharp increase in the discharge ignition voltage and the quasi-steady discharge voltage U_{ch} in the normal mode (Fig. 1a). The power deposited in the discharge plasma increases with the total pressure of the mixture and the CFC-12 content and attains 30–65 W (Fig. 1b).

The plasma emission spectrum corrected with allowance for the relative sensitivity of the recording system is shown in Fig. 2. As in the case of a longitudinal discharge in a He/Cl_2 mixture [4], the 200- and 258-nm emission from chlorine molecules prevails. Moreover, against the background of the Cl_2^{**} VUV emission bands in the 140- to 170-nm range, there are intense spectral lines of single-charged chlorine ions. The identification of this linear spectrum is shown in the table. A group of unresolved Cl^{+*} spectral lines forms the shortest wavelength part of the spectrum with λ 156–162 nm. In the plasma of a transverse discharge in Ar/CF_2Cl_2 mixture, no emission from chlorine ions in the wavelength range 155–165 nm was recorded [6], which is possibly related to the involvement of helium in the production of the excited chlorine ions and the specific features of the plasma of a contracted longitudinal discharge.

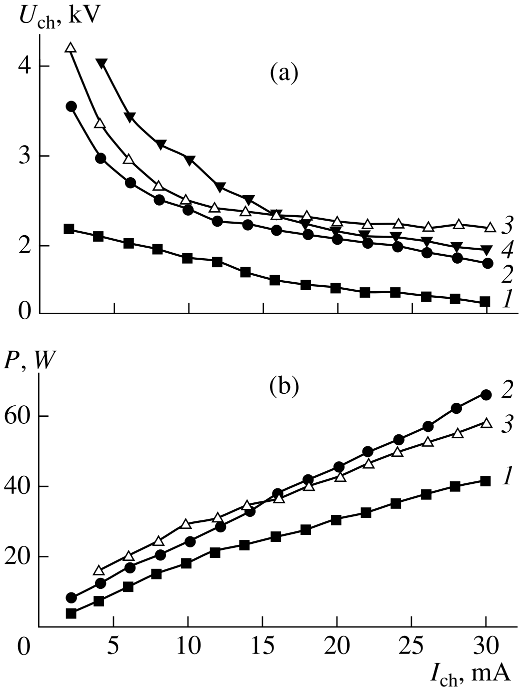


Fig. 1. (a) Current–voltage characteristics of a longitudinal dc discharge in the He/CF_2Cl_2 = (*1*) 1.33/0.13, (*2*) 1.33/0.40, (*3*) 1.33/0.67, and (*4*) 6.70/0.40-kPa mixtures, and (b) the power deposited in plasma vs. discharge current in the He/CF_2Cl_2 = (*1*) 1.33/0.13, (*2*) 1.33/0.67, and (*3*) 6.70/0.40-kPa mixtures.

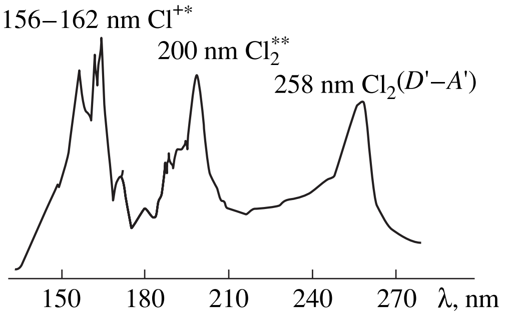


Fig. 2. VUV–UV emission spectrum from the steady-state electric-discharge plasma in a He/CF_2Cl_2 mixture.

The dependences of the VUV and UV emission intensities on the composition of a He/CF_2Cl_2 mixture are shown in Fig. 3. At low pressures of the mixture, the optimum partial pressure of CFC-12 is in the range 0.3–0.5 kPa. Further increase in the CFC-12 content leads to a decrease in the intensities of the majority of the emission bands and the total intensity in the range 130–280 nm. The emission from chlorine ions is the most sensitive to the CFC-12 content in the mixture (Fig. 3a). As for the dependence of the emission intensities on the helium partial pressure, there are a narrow peak at

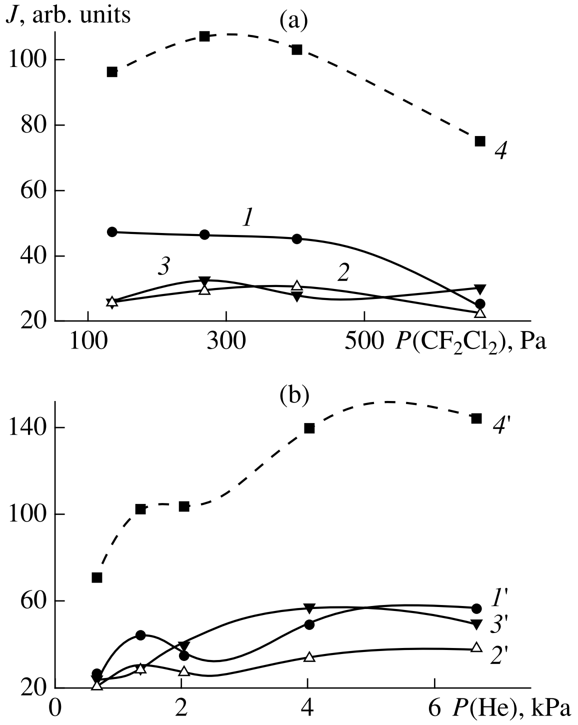


Fig. 3. Emission intensities from a longitudinal discharge in He/CF_2Cl_2 mixtures vs. (a) the CFC-12 partial pressure at P(He) = 1.33 kPa and (b) the helium partial pressure at $P(CF_2Cl_2)$ = 0.40 kPa in the ranges (*1*, *1*') 130–175, (*2*, *2*') 175–215, and (*3*, *3*') 215–280 nm and (*4*, *4*') the total emission intensity in the 130- to 280-nm spectral range.

P(He) = 1.3–1.4 kPa and a broad peak at P(He) = 5–6 kPa in the spectral range where the emission from Cl^{+*} ions prevails. The emission from the $Cl_2(D'–A')$ 258-nm band has a broad peak at P(He) = 4 kPa, and the intensity of the Cl_2^* emission band with a maximum at λ = 200 nm slightly increases with increasing P(He). At low helium pressures, the VUV emission from chlorine ions prevails. The total emission intensity in the spectral range 130–280 nm also increases with the helium partial pressure. The optimum partial pressure of helium in the mixture is within 5–6 kPa. Such behavior of the dependence of VUV and UV emission intensities on P(He) in the ion–ion plasma of a contracted discharge can be related to the significant contribution of energy transfer from the metastable atoms and molecules of helium to the molecules and main dissociation products of CFC-12 in the plasma (Cl_2, Cl^+, etc.), as well as to the efficient recombination of the Cl^+ and Cl^- ions toward the production of $Cl_2(D')$ and Cl_2^{**} molecules.

An increase in the discharge current in the range 2–30 mA enhances the emission intensity from chlorine molecules and ions (Fig. 4). For the shortest wavelength fraction of the emission spectrum, the rate at which the intensity grows as I_{ch} increases is nearly twice as high as that for chlorine molecules. The increase in the helium partial pressure from 0.67 to 4.00 kPa also increases the rate of the intensity growth in the shortest wavelength part of the spectrum by a factor of approximately 1.5. When the discharge tube is cooled only with the working mixture of the contracted discharge under study, no saturation of the VUV and UV emission intensities is observed at the currents $I_{ch} \leq$ 30 mA. In the emitter based on a glow discharge in an Ar/Cl_2 mixture (which is the working mixture of the lampsoperating on ArCl 175-nm and $Cl_2(D'–A')$ 258-nm band system), under similar pumping conditions, the emission intensity dropped already at $I_{ch} \geq$ 22 mA.

The resource characteristics of the VUV–UV emitter pumped with a contracted longitudinal discharge and operating in the gas-static regime with the use of a 10l buffer volume are seen in Fig. 5. Each curve in Fig. 5 was obtained with a fresh He/CFC-12 = 1.33/0.4-kPa mixture at I_{ch} = 20 mA. The lifetime (defined as the time it takes for the emission intensity to fall by a factor of 2) of an emitter operating on either λ = 162 or

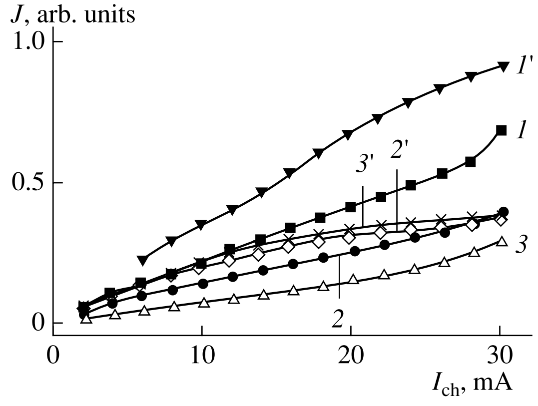


Fig. 4. Maximum emission intensities at the wavelengths λ = (*1*, *1*') 162, (*2*, *2*') 200, and (*3*, *3*') 258 nm vs. discharge current in the He/CF_2Cl_2 = (*1*–*3*) 0.67/0.40-kPa and (*1*'–*3*') 4.00/0.40-kPa mixtures.

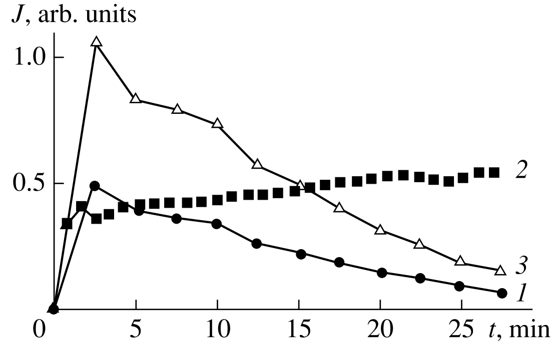


Fig. 5. Maximum emission intensities at the wavelengths λ = (*1*) 258, (*2*) 200, and (*3*) 162 nm vs. discharge run time in a He/CF_2Cl_2 mixture.

Constituents of the VUV emission from chlorine ions in the plasma of a He/CF_2Cl_2 mixture

λ, nm Cl II	E_{low}, eV	E_{up}, eV	Transition
156.5	13.67	21.60	$3d\ {}^5D^0_{0, 1, 2, 3, 4}-5f\ {}^5F_{1, 2, 3, 45}$
156.6	11.70	19.61	$3p^5\ {}^3P^0_0-4p''\ {}^3S_1$
160.4(5)	14.85	22.58	$3d\ {}^3D^0_{2, 3}-4f'\ {}^3G_{3, 4}$
161.1	15.71	23.41	$4s'\ {}^3D^0_2-5p''\ {}^3D_2$
161.9	13.67	21.33	$3d\ {}^5D^0_3-6p\ {}^5P_2$
161.9	13.67	21.33	$3d\ {}^5D^0_4-6p\ {}^5P_3$
173.1	14.85	22.01	$3d\ {}^3D^0_3-5p'\ {}^3D_3$
178.7	11.65	18.59	$3p^5\ {}^3P^0_1-4p'\ {}^3P_1$
178.9	15.65	22.58	$3d'\ {}^3F^0_3-4f'\ {}^3G_4$
179.2	11.65	18.57	$3p^5\ {}^3P^0_1-4p'\ {}^3P_2$
179.3	15.68	22.59	$3d'\ {}^3F^0_4-4f'\ {}^3G_5$
179.4	15.68	22.59	$3d'\ {}^3F^0_4-4f'\ {}^3H_5$
179.7	11.70	18.59	$3p^5\ {}^3P^0-4p'\ {}^3P_1$
185.7	13.67	20.35	$3d\ {}^5D^0_{0, 1, 2, 3, 4}-4f\ {}^5F_{1, 2, 3, 45}$
188.7	11.58	18.14	$3p^5\ {}^3P^0_2-4p'\ {}^3D_2$
192.3	11.70	18.14	$3p^5\ {}^3P^0_0-4p'\ {}^3D_1$
193.7	13.38	19.77	$4s^5\ {}^5S^0_2-5p\ {}^5P_3$
199.6	16.39	22.60	$3d'\ {}^3G^0_4-4f'\ {}^1H_5$
199.7	16.39	22.60	$3d'\ {}^3G^0_5-4f'\ {}^3H_6$
199.8	16.39	22.59	$3d'\ {}^3G^0_3-4f'\ {}^3H_4$
199.9	16.39	22.59	$3d'\ {}^3G^0_4-4f'\ {}^3H_5$

258 nm does not exceed 15 min, whereas for the 200-nm band, it is longer than 60 min. Thus, in the gas-static regime, the emission from the Cl_2^{**} 200-nm band is of most interest. For an emitter operating on λ = 258-nm line or on the chlorine ion transitions, it is reasonable to use the regime of slow circulation of the working mixture in the discharge tube. The He/CF_2Cl_2 flow rate may be within the range 0.3–1.0 l/min. The average output power in the 130- to 280-nm range does not exceed 1–1.5 W with an efficiency of 2–3%.

The results of studying the optical characteristics of a contracted longitudinal discharge in He/CF_2Cl_2 mixtures can be summarized as follows. At a deposited power of 20–60 W, the discharge acts as a source of radiation in the 130- to 280-nm spectral range. The radiation is generated due to the spontaneous decay of the dissociation products of CFC-12 molecules and incorporates the $Cl_2(D'–A')$ 258-nm band, the chlorine emission continuum with a maximum at λ = 200 nm, and the emission from Cl^{+*} ions in the range 130–175 nm. The optimum partial pressures of CFC-12 and helium lie within the ranges 0.3–0.5 and 5–6 kPa, respectively. The increase in the discharge current from 2 to 30 mA enhances the emission intensity throughout the entire VUV–UV spectral range with no tendency to saturation. The longest emitter lifetime is achieved when operating on the 200-nm band. The total emission power attains 1.5 W with an efficiency of ≤3%. The active medium of the emitter is the ion–ion plasma; hence, it is of interest to diagnose this plasma with the aim of studying the mechanisms for the production of the excited chlorine ions and molecules and to perform the relevant numerical simulations. It is promising to use the VUV–UV emitter under study in a repetitive mode with the pumping by a fast longitudinal ionization wave (nanosecond longitudinal low-pressure discharge).

REFERENCES

1. A. P. Golovitskiĭ and S. V. Lebedev, Opt. Spektrosk. **82**, 251 (1997) [Opt. Spectrosc. **82**, 227 (1997)].
2. M. I. Lomaev, A. N. Panchenko, É. A. Sosnin, and V. F. Tarasenko, Zh. Tekh. Fiz. **68** (2), 64 (1998) [Tech. Phys. **43**, 192 (1998)].
3. A. K. Shuaibov, A. I. Dashchenko, and I. V. Shevera, Zh. Tekh. Fiz. **71** (8), 121 (2001) [Tech. Phys. **46**, 1049 (2001)].
4. A. K. Shuaibov, A. I. Dashchenko, and I. V. Shevera, Teplofiz. Vys. Temp. **39**, 833 (2001).
5. A. K. Shuaibov, Zh. Tekh. Fiz. **70** (10), 117 (2000) [Tech. Phys. **45**, 1346 (2000)].
6. A. K. Shuaibov, L. L. Shimon, A. I. Dashchenko, and I. V. Shevera, Teplofiz. Vys. Temp. **38**, 386 (2000).
7. V. I. Svettsov, A. L. Kupriyanovskaya, and A. B. Maryshev, Zh. Prikl. Spektrosk. **35**, 205 (1981).
8. A. P. Golovitskiĭ, Pis'ma Zh. Tekh. Fiz. **24** (6), 63 (1998) [Tech. Phys. Lett. **24**, 233 (1998)].
9. L. M. Vasilyak, S. V. Kostyuchenko, A. V. Krasnochub, and M. E. Kuz'menko, Zh. Prikl. Spektrosk. **65**, 302 (1998).
10. Yu. I. Bychkov, S. L. Gorokhov, and A. G. Yastremskiĭ, Kvantovaya Élektron. (Moscow) **30**, 733 (2000).
11. A. A. Kudryavtsev, A. L. Kuranov, V. G. Mishakov, *et al.*, Zh. Tekh. Fiz. **71** (3), 29 (2001) [Tech. Phys. **46**, 299 (2001)].
12. Yu. P. Raizer, *Gas Discharge Physics* (Nauka, Moscow, 1987; Springer-Verlag, Berlin, 1991).

Translated by N. Ustinovskiĭ

Technical Physics, Vol. 47, No. 4, 2002, pp. 410–414. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 36–40.
Original Russian Text Copyright © 2002 by Shadrov, Tagirov, Boltushkin.

SOLIDS

Structure Characteristics and Magnetization Reversal of High-Coercivity Cobalt-Based Alloy Films

V. G. Shadrov, R. I. Tagirov, and A. V. Boltushkin

Institute of Solid-State and Semiconductor Physics, Belarussian Academy of Sciences,
ul. Brovki 17, Minsk, 220072 Belarus

Received June 14, 2001

Abstract—The structure of “continuous” Co–W and Co–Ni–W alloy films, as well as Co-containing metallic oxide heterogeneous films, covering the aluminum surface is studied in dependence of intergranular magnetic interaction and magnetization reversal processes by taking the angular dependences of the coercive force and irreversible susceptibility and also from δ*M* curves. *© 2002 MAIK “Nauka/Interperiodica”.*

INTRODUCTION

In recent years, many researchers studying magnetic media have been concentrated on Co-based high-coercivity alloy films [1, 2]. The magnetic parameters of these films vary in wide limits, making them suitable for both longitudinal and vertical data writing. To date, however, the processes of magnetization reversal and intergranular magnetic interaction, which are of great scientific and practical interest, have been studied in films condensed in a vacuum [3–5]. In the films obtained electrochemically, these processes are poorly understood [6, 7], although these films admit a wider control of their properties. For example, the electrochemical filling of anodic oxide films (AOFs) covering the aluminum surface by a magnetic material makes it possible to easily control the magnetic separation of needle-like crystals and their diameter by varying the geometric parameters of the AOF cellular structure [8]. Also, the electrochemically obtained films are less expensive.

In this work, we seek a correlation between the structure of “continuous” Co–W and Co–Ni–W alloy films, as well as Co-containing metallic oxide heterogeneous films, covering the aluminum surface and the processes of intergranular magnetic interaction and magnetization reversal. In the Co–W films, the magnetic separation of the grains is less pronounced and is due to higher concentrations of crystal defects, foreign impurities and compounds, etc., in the intergranular layers, hence, a reduced magnetic moment of the layers. By varying the deposition conditions and/or postelectrolysis treatment of the Co–W films, one can change their structure parameters: grain size and shape, texture and texture uniformity, width of the boundary regions. On the one hand, this affects the magnetic properties of the grains and their magnetic separation; on the other, this makes experimental data difficult to interpret.

METHODS OF FILM PREPARATION AND CHARACTERIZATION

Films of Co–W alloy were obtained by electrochemically depositing from sulfate–chloride electrolytes on copper substrates. The electrolyte compositions were close to those described in [7, 9]. The deposition conditions were as follows: pH 6.6, current density 10 mA/cm^2, and electrolyte temperature 18–50°C. The Co–10 wt % Ni–15 wt % W films were obtained under somewhat different conditions: pH 3.5–5.3, current density 55 mA/cm^2, and electrolyte temperature 18–55°C. Co-containing metallic oxide heterostructures were obtained by electrochemical deposition into pores that were formed on the AD-1 aluminum surface by anodizing in a sulfuric acid electrolyte under conditions reported in [10]. Some of the samples were annealed at 240°C under a pressure of 10^{-5} torr. The structure of the films was studied with a DRON-2 diffractometer, EVM-100 LM and JEOL-100CX electron microscopes; magnetic characteristics, with an automated vibration magnetometer under fields of up to 15 kOe.

The difference in the crystal structure of the films causes differences in the processes of intergranular magnetic interaction and magnetization reversal. Unfortunately, methods for the direct observation of magnetization reversal are of limited application (for example, Lorentz microscopy applies only to free thin films); therefore, indirect methods are widely used. Among them are those using the angular dependences of the coercive force, remanent magnetization, rotational hysteresis, and losses due to magnetization reversal [3–6].

If in the crystal consisting of isolated grains magnetization reversal is due to incoherent rotation, the angular dependence of the coercive force is given by [11]

$$H_c = H_0 \frac{1.08S^{-2}(1-1.08S^{-2})}{[(1-1.08S^{-2})^2 - \sin^2\varphi(1-1.16S^{-2})]^{1/2}}, \quad (1)$$

Preparation conditions, composition, texture, and magnetic properties of the Co–W and Co–Ni–W films ($h = 1$ μm)

Sample no.		T, °C	W, wt %/pH	Texture	H_c, Oe	H_r, Oe	H'_r, Oe
Co–W	1*	18	15/6.6	[001]	639	843	848
	2	18	15/6.6	[001]	175	230	490
	3	29	15/6.6	tr. [001]	536	628	634
	4	40	16/6.6	[100]	391	422	583
	5**	19	15/6.6	[001]	168	227	374
CoNi–W	6	19	/3.5	tr. [001]	146	245	239
	7	19	/4.5	[001]	233	292	307
	8	19	/5.3	[001]	460	478	541
	9	47	/5.3	[001]	555	635	694

* Thickness 0.05 mm; ** annealed sample 2.

where $S = R/R_0$, R is the grain radius, and R_0 is the characteristic radius depending on the exchange interaction energy and saturation magnetization.

The character of magnetic interaction can be determined, for example, by finding the $\delta M(H)$ functions and measuring the irreversible susceptibilities χ^d_{irr} and χ^r_{irr} [5]. They are defined as

$$\delta M(H) = I_d(H) - (I_\infty - 2I_r(H)),$$

$$\chi^d_{\text{irr}} = dI_d/dH, \quad \chi^r_{\text{irr}} = dI_r/dH,$$

where I_∞ is the remanent magnetization after the saturating field has been switched off; $I_d(H)$ and $I_r(H)$ are the remanent magnetizations after switching off the saturating field H when a negative field is applied to the sample in the state I_∞ and when a positive field is applied to the demagnetized sample, respectively; and H_r and H'_r are the fields at which $I_d(H) = 0$ and $I_r(H'_r) = 0.5I_\infty$.

According to [12], for noninteracting single-domain particles, the relationships $\delta M(H) = 0$ and $H_r = H'_r$ are valid. In the case of magnetostatic (dipole–dipole) interaction between single-domain particles, $\delta M(H) < 0$ and $H_r > H'_r$. If the grains are divided into domains and/or in the case of exchange interaction between the grains, $\delta M(H) > 0$ and $H_r = H'_r$.

The type of interaction can also be judged from the position of the $\chi_{\text{irr}}(H)$ peaks: in the absence of interaction, the peaks coincide; for dipole–dipole interaction, the peak of $\chi^d_{\text{irr}}(H)$ is shifted relative to that of $\chi^r_{\text{irr}}(H)$ toward weaker fields; and for exchange interaction, the former peak lies in a higher field range.

STRUCTURE EXAMINATIONS

As follows from the X-ray and electron microscopy data, the Co–W films deposited on the copper substrates consist of individual grains, which represent the solid solution based on hcp cobalt [9]. At an electrolyte temperature of 20°C, first the transition layer of thickness about 50 nm forms. It consists of randomly oriented equiaxial grains separated by layers with an increased concentration of defects, inclusions, hydrogen, and $Co(OH)_2$ hydroxide. As the film thickens, columnar grains of diameter 40–50 nm with the [001] texture start to form.

During the growth of the columnar grains, their active surface is gradually filled by adsorbed impurity particles and hydroxide sol particles, which are from time to time captured by moving growth steps. As a result, the columns contain zones with an increased distortion of the crystal lattice, which separate more perfect zones. With a rise in the electrolyte temperature, the relative adsorption of impurities and hydroxides decreases and the column growth rate rises. Eventually, blocks extended in the column growth direction appear [9]. The compositions of the films, their texture, and magnetic parameters vs. electrolyte temperature and pH are listed in the table. As the electrolyte temperature is raised to 32°C, the axis in the Co–W films changes the orientation from the hcp lattice to the film plane and platelet-like grains appear. Simultaneously, the total amount of the impurities drops. At an electrolyte temperature of 40–45°C, all the grains become platelet-like.

In the Co–Ni–W films, the dependence of the texture on the electrolyte temperature is weaker: the [001] texture changes to the [100] texture at an electrolyte temperature of 47°C. The texture is also more stable against pH variations: in the range $3.5 < \text{pH} < 5.3$, the texture remains the same, [001]. One might expect that, with pH varying in the so wide limits, the impurity concentration will considerably change: for example, the impurity concentration in electrodeposited Co films

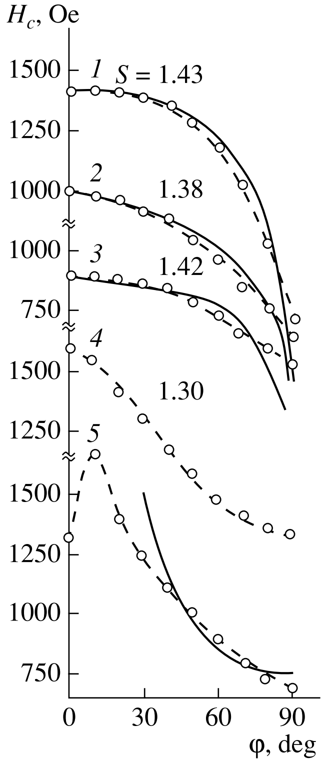


Fig. 1. Coercive force vs. magnetization reversal angle. Continuous lines, calculation; o, data points. Curve *5* was constructed according to the displacement theory [14]. *1*, Co film with pores of diameter 18 nm; *2*, annealed sample 3; *3*, Co film with pores of diameter 30 nm; *4*, Co–20 wt % W film with the [001] texture; and *5*, Co–20 wt % W film with the [100] texture.

changes from 2.3×10^{15} cm^{-3} at pH 1.7 to 1.4×10^{17} cm^{-3} at pH 5.7 [13].

Co films consisting of individual needle-like grains of diameter (AOF pore diameter) 18 and 30 nm and length ≈1 μm have the hcp lattice with the randomly oriented *c* axis (the I_{002}/I_{100} X-ray intensity ratio is 2 to 3). The annealing of the films of both types obtained under the above conditions affects the geometric parameters of the columnar (cellular) structure insignificantly.

MAGNETIC MEASUREMENTS AND DISCUSSION

The "continuous" Co-based structures and AOFs studied in this work consist of grains separated by nonmagnetic layers (boundaries) or those with a decreased magnetization. The sizes of the grains (needles of diameter 18–30 nm on the AOF surface and columns of diameter 40–50 nm in the Co–W films) suggest that the grains are almost single-domain. Such an assumption seems to be quite reasonable: the critical grain size corresponding to purely single-domain cobalt has been estimated at 30 nm [14]. If it is taken into account that the saturation magnetization in Co–W films is lower than in pure cobalt, the critical size will increase. At a large energy of uniaxial magnetic anisotropy, which may be as high as 10^{16} erg/cm^2, individual grains can be viewed as permanent magnets. If the energy of anisotropy appreciably exceeds the energy of magnetostatic interaction between the grains, the film represents an ensemble of isolated noninteracting particles; then, according to [5], $\delta M(H) = 0$, $H_r = H'_r$, and the angular dependence of the coercive force is given by expression (1).

The measurements on the Co-containing AOFs, consisting of individual needle-like grains separated from each other, confirm the above conclusion: at the early stage, the angular dependences are coincident with the calculated ones if *S* is taken to be 1.43, 1.38, and 1.42 for curves *1*–*3*, respectively (Fig. 1). With regard for the finite sizes of cobalt particles (in calculations performed in [11], infinitely long cylinders are considered) and the dispersion of easy-magnetization axis and the anisotropy field, the coincidence will be even better (the associated correction has been made for uniaxial cobalt–chromium films with parameters close to those of the films studied in this work [15]). Note that the values of H_{c1} for the films with different pore diameters differ in accordance with theoretical predictions [8]. With this circumstance taken into account, the associated values of *S* virtually coincide. This can be explained by the lack of adequate experimental data (specifically, for the contribution from the crystallographic component of the magnetic anisotropy) and also by the fact that our model is to some extent conventional. At the same time, an increase in H_{c1} in the annealed samples is attended by a decrease in *S*; in other words, the refinement of the needle-like crystallite structure and the related growth of anisotropy can be associated with the enhanced effect of rotation in magnetization reversal. This is also confirmed by a decrease in the rotational hysteresis integral in the annealed structures [16].

Incoherent rotation may also be responsible for magnetization reversal in the hcp Co–W films, which have the [001] texture and columnar grains. The experiment (Fig. 1, curve *4*) coincides with the calculation at least up to 30°–40° if *S* is set equal to 1.30. At higher angles of magnetization reversal, the experimental and analytical curves diverge considerably, indicating the major contribution from another mechanism of magnetization reversal, namely, domain wall displacement [4, 7].

As follows from the aforesaid, magnetization reversal in the Co and Co–20 wt % W films deposited at room temperature ([001] texture) is described well by the model of separated uniaxial grains. If, however, the energy of uniaxial magnetic anisotropy decreases or the saturation magnetization grows, the effect of magnetostatic interaction becomes more pronounced. Such a situation is observed in the Co–W and Co–Ni–15 wt %

W films, where the magnetic moment lies in the film plane in spite of the columnar morphology and [001] texture. This means that the energy of the demagnetizing field exceeds the magnetostatic energy; in this case, the film cannot be represented as an ensemble of noninteracting particles. The experimental data suggest the presence of magnetostatic interaction, since $\delta M < 0$ and $H_r < H_r'$ for most of the films (see table and Fig. 2).

It should also be stressed that the columns, while having a rather perfect [001] texture, consist of several blocks slightly misoriented relative to each other. These blocks are separated by zones where the crystal lattice is highly distorted [9]. Thus, the blocks inside the columns can be viewed as single-domain magnetostatically interacting particles. The validity of this assumption is confirmed by the measurements of the magnetic characteristics of samples 2 (Co–W) and 8 (Co–Ni–W): for both $\delta M(H) < 0$. The same measurements taken from sample 1 (Co–W), which, like sample 2 (Co–W), was obtained at room temperature, indirectly support the presence of magnetostatic interaction between the blocks: at room temperature, magnetostatic interaction in sample 1 is almost absent. The reason for the absence of magnetostatic interaction in sample 1 is the absence of the block structure, since the thickness of this sample coincides with that of the transition layer, consisting of equilibrium grains.

More evidence in favor of this assumption is the reduced interaction in the annealed Co–W films (h = 1 μm) with the [001] texture (Fig. 2, curve *5*). This reduction is due to the poorer magnetic separation between the substructure elements because of a decrease in the concentration of crystal defects and their redistribution. As a result, the contribution of the substructure elements to the resulting magnetostatic interaction drops (with the intergranular interaction remaining almost unchanged).

As the electrolyte temperature grows, the adsorption of the hydroxides decreases and the probability of the hydroxides being captured by the growing crystal lattice drops. Simultaneously, the [001] texture changes to the [100] texture. Because of this, the impurities are pushed out to the grain boundaries and the lattice in the columns becomes more perfect. Eventually, the blocks in the grains disappear and their contribution to the magnetostatic interaction vanishes (Co–W sample 3). Under these conditions, there is no reason to assume a considerable magnetostatic interaction between the grains, since the grain boundaries remain sufficiently wide. It is also necessary to take into account that the [001] texture axis, along which the energy of magnetic anisotropy is the highest, may change the orientation and lie in the film plane.

Note that the situation in sample 3 appears to be more complicated. From the irreversible susceptibility curves (Fig. 3), it follows that exchange interaction is observed in weak fields (the peak of χ^r lies to the left of that of χ^d) and magnetostatic interaction, in high fields (the peaks change places). Possibly, at an electrolyte temperature of 29°C, the grains not only are refined but also grow. As a result, grains coarser than purely single-domain grains may appear. Simultaneously, magnetostatic interaction between the blocks is partially retained and that between the grains takes place.

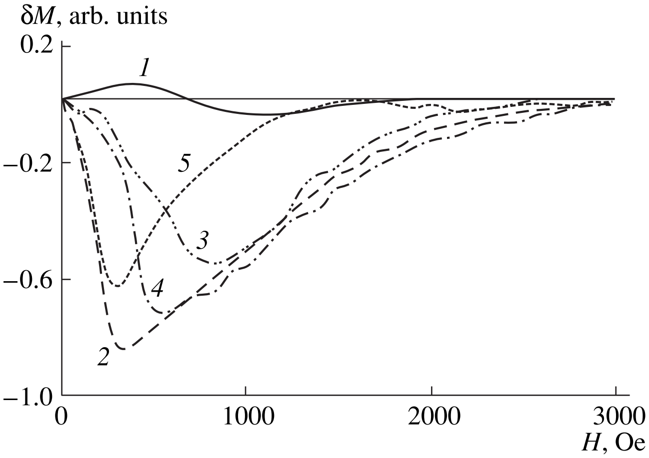


Fig. 2. Field dependences of δ*M* taken from the Co–W films. Curve numbers correspond to sample numbers in the table.

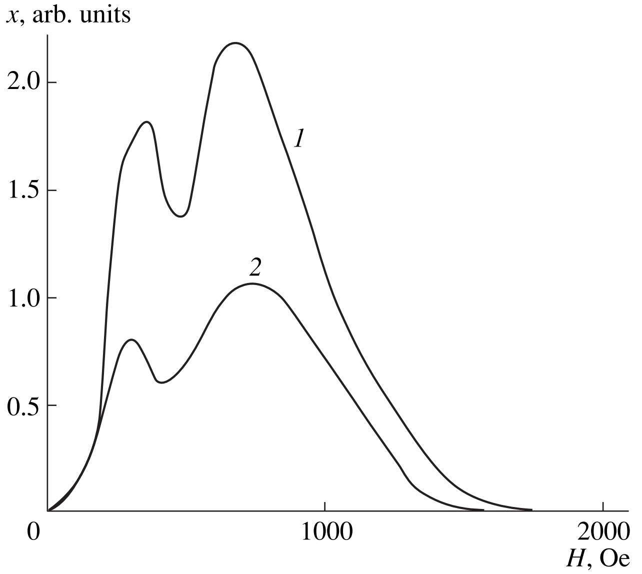


Fig. 3. Field dependences of the irreversible susceptibility for the Co–W film (sample 3): (*1*) χ^d and (*2*) χ^r.

A further rise in the electrolyte temperature causes new structures to appear. The columns are replaced by platelets, and the grain size grows. The impurities are removed from the grains, and the relative volume of the grain boundaries shrinks. As a consequence, the magnetic separation of the grains degrades, and the magnetostatic interaction grows (samples 4, Co–W, and 9, Co–Ni–W).

Studying the Co–Ni–W films deposited at different pH, we could clarify the effect of grain size and the magnetic separation of the grains on magnetization reversal. The decrease in pH from 5.3 to 3.5 drastically (by one to two orders of magnitude) reduces the impurity concentration in the films [13]. This affects the grain boundary thickness and the grain size: the grain boundaries become thinner and the grains, "purer" and coarser. As a result, magnetostatic interaction weakens and exchange interaction appears ($H_r' < H_r$, sample 6).

CONCLUSIONS

(1) Upon the electrodeposition of cobalt into pores in anodized aluminum, films consisting of needle-like ferromagnetic particles form. The particles are close to single-domain grains by size, and magnetization reversal in them is accomplished by incoherent rotation.

(2) Continuous Co–W and Co–Ni–W films deposited at room temperature consist of grains divided into blocks, which are presumably responsible for the magnetostatic interaction observed.

(3) The coarsening of the grains and a decrease in the concentration of adsorbed impurities, which take place when the electrolyte temperature rises and pH declines, violates the conditions under which the grains are single-domain. This shows up as magnetostatic interaction between the grains.

ACKNOWLEDGMENTS

This work was supported by the Belarussian Foundation for Basic Research (grant no. T99-107) and by the NATO grant HTECH LG, no. 940656.

REFERENCES

1. K. E. Johnson, C. M. Mate, J. A. Mertz, *et al.*, IBM J. Res. Dev. **40**, 511 (1996).
2. P. J. Grundy, J. Phys. D **31**, 2975 (1998).
3. J. C. Lodder and L. Cheng-Zhang, J. Magn. Magn. Mater. **74**, 74 (1988).
4. R. Ranjan, J. S. Gau, and N. Amin, J. Magn. Magn. Mater. **89**, 38 (1990).
5. P. I. Mayo, K. O'Grady, R. W. Chantrell, *et al.*, J. Magn. Magn. Mater. **95**, 109 (1991).
6. U. Admon, M. P. Dariel, E. Grunbaum, and J. C. Lodder, J. Appl. Phys. **66**, 316 (1989).
7. V. G. Shadrow, R. I. Tagirov, A. V. Boltushkin, and N. N. Kozich, J. Magn. Magn. Mater. **118**, 165 (1993).
8. D. AlMawlawi, N. Coombs, and M. Moskovits, J. Appl. Phys. **70**, 4421 (1991).
9. A. V. Boltushkin, V. G. Shadrov, T. A. Tochitskiĭ, and Zh. P. Apkhipenko, Elektrokhimiya **9**, 1105 (1990).
10. V. G. Shadrov, A. V. Boltushkin, L. B. Sosnovskaya, *et al.*, Metally, No. 2, 120 (1999).
11. S. Shtrikman and D. Treves, J. Phys. Radium **20**, 286 (1959).
12. E. P. Wohlfarth, J. Appl. Phys. **29**, 595 (1958).
13. S. Nakahara and S. Mahajan, J. Electrochem. Soc. **127**, 283 (1980).
14. S. V. Vonsovskiĭ, *Magnetism* (Nauka, Moscow, 1971; Wiley, New York, 1974).
15. J. Nakamura and S. Iwasaki, IEEE Trans. Magn. **23**, 153 (1987).
16. A. V. Boltushkin, V. G. Shadrov, N. N. Kozich, *et al.*, Vestsi Akad. Navuk Belarusi, Ser. Fiz.-Mat. Navuk, No. 1, 47 (1993).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 4, 2002, pp. 415–418. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 41–45.
Original Russian Text Copyright © 2002 by Ostashev, Shevchenko.

SOLIDS

Formalization of Models of a Deformable Polycrystalline Material in Terms of Mesomechanics

V. V. Ostashev and O. D. Shevchenko

Pskov Polytechnical Institute (Branch of St. Petersburg State Technical University),
Pskov, 180680 Russia

Received June 15, 2001

Abstract—The stage of formalization is necessary for constructing an adequate mathematical model of any phenomenon. By the example of a deformable polycrystalline material, it is shown that the formalization algorithm is based on a closed set of definitions, which represents the material under study as an open nonlinear dynamic structurally stable hierarchical (multilevel) dissipative self-organizing information system. Formalization principles, as applied to models of plastic deformation under static loading, are experimentally verified with polycrystalline fcc materials (MO copper and 08Kh18N10T austenitic steel). © *2002 MAIK "Nauka/Interperiodica".*

GENERAL STATEMENTS

The description of plastic deformation in a polycrystalline material on the mesoscopic level requires the qualitative formalization of their mathematical models and imposes certain demands on the rigor of the terminology used and basic definitions. In the general case, a formalization algorithm must map the physical model onto the mathematical one and vice versa.

It is known that the "hottest" issues in mesomechanics are geometrical image of a deformation-induced defect (hereafter, deformation defect), structural levels of deformation, dissipative structures, and representative volume. The physical mechanisms of plastic deformation on the mesoscopic level have been considered by the authors of mesomechanics [1, 2]. In the general case, they are as follows.

(1) The basic mechanism of deformation is initial glide, which always results in initial material rotation. All other deformation mechanisms are accommodation ones, which provide the relaxation of the field of rotational moments acting on a structural deformation element from the ambient material.

(2) The accommodation mechanisms of deformation are realized through secondary fluxes of defects and can cause both the material rotation (multiple glide) and the crystallographic rotation of a structural element of deformation (grain-boundary glide, grain migration, or fragmentation).

(3) The natural relation between shear and rotation leads to the fact that a translational rotational vortex (TRV) of different scale rather than shear is an elementary act of plastic deformation in mesomechanics.

Within the framework of this approach, a deformable polycrystalline material is treated on the mesoscopic level as an ensemble of interacting elements—mesodefects, which have certain geometrical dimensions and are separated from each other by boundaries. The evolution of the system depends on the dynamics of mesodefects, their mechanical characteristics, and the interaction within the representative volume. The interaction between mobile mesodefects results in the formation of dissipative structures, as a rule, of the correlation type; however, the functional ordering also favors the formation of geometrical structures characteristic of the mesoscopic level. This means that mesodefects containing both the shear and rotational components of the strain make possible the displacement of volume structure elements of different scale (subgrains, grains, their conglomerates, or extended blocks) in a deformable solid.

PRINCIPLES OF FORMALIZATION

Thus, models of physical mesomechanics are based on the description of nonlinear interactions between large-scale deformation defects called mesodefects, and the deformable material refers to open nonlinear dynamic structurally stable, hierarchical dissipative, self-organizing information systems.

An open system always suggests mass transfer during the deformation process, as well as energy and information exchange through the fluxes of deformation defects. If the intensity of these fluxes is sufficiently high, ordered space–time structures arise. The system captures the transfer fluxes even if they are somewhat structured, transforms and organizes them, and endows with its own space–time structure. In essence, this is the way in which a deformable material becomes structured, i.e., self-organizes. One can conclude that a deformable crystalline material adapts to loading conditions by varying the extent of its openness. As an open system, a material can quantitatively be characterized with the equation of entropy balance

$$\frac{ds}{dt} = \frac{ds_i}{dt} + \frac{ds_e}{dt}, \tag{1}$$

where the terms stand for the rate of entropy production in an open system, the rate of entropy production in the system due to internal irreversible processes, and the rate of entropy exchange with the environment, respectively. ds_i/dt is always positive by definition, while ds_e/dt may be both positive and negative.

The nonlinear behavior (nonlinearity) of a polycrystalline material in response to an external action shows up in its integral properties and includes a variety of possible stationary states. A dynamic system is defined as a process or object whose initial state is specified by a set of parameters and for which there always exists an operator that describes the behavior of the system in time and space.

Nonlinear dynamic systems differ from strictly determinate ones in that their behavior is sometimes akin to random. Such a phenomenon is called dynamic chaos. Dynamic chaos results from internal interactions in the system rather than as a consequence of random external actions. In other words, it is due to the intrinsic dynamics of a nonlinear system. The chaos (as an intrinsic property of the system) arises under almost any deformation conditions. However, it may go unnoticed if the range of parameters where it manifests itself is extremely narrow, the time interval of its manifestation is very extended, or it is shaded by more vigorous processes.

A polycrystalline material as a structurally stable system is characterized by the stability of the parameters of its geometrical structure and by the invariance of the functional relations between the dynamic and statistical components of the system. Obviously, any dynamic solution and any ordering in the structure is a fluctuation from the viewpoint of the statistical approach. Thus, the instability of dynamic states means the stability of statistical ones. The reverse is also true: the stability of dynamic solutions means the absence of fluctuation relaxation and, therefore, the instability of the statistical component.

By a multilevel hierarchical system, we mean a set of structures with different levels of hierarchy. Correct structurization of nonlinear systems is one of the most complicated problems in simulating plastic microstrains. It is based on the Leibniz principle of identification of state indistinguishability. As applied to our problem, this principle is stated as follows: upon originating and interacting, deformation defects that define a structural level must have properties so much alike that the defects are indistinguishable when being analytically processed with a given method. Three scale levels are conventionally identified: macrolevel, mesolevel, and microlevel. The hierarchical structure of the mesolevel can in its turn be subdivided into structural levels as follows. A deformable polycrystalline material is represented as a set U of structural levels and a set R of links (relations) between them. A pair of sets U and R is divided into lower level subsystems whose dynamics is defined by the interaction between mesodefects in the representative volume. In this representation, systems of the previous level are mesodefects for next-level systems; i.e., the distribution of mesodefects over the levels is mapped by the set of embedded sets:

$$U_j(R_j) \longrightarrow U_{j+1}(R_{j+1}) \longrightarrow U_{j+2}(R_{j+2}) \longrightarrow SS. \tag{2}$$

Open nonlinear nonequilibrium systems, in which new structures may appear through the dissipation (scattering) of the flux entering the system and propagating in it, refer to self-organizing systems. Basic premises to self-organization in dissipative systems are: an energy or information flux from the outside due to the openness of the system; the nonequilibrium state of the system, enhancing fluctuations to the point where chaos creates order; and the presence of potentiality, which allows stability exchange between the structures inside. The emergence of dissipative structures during deformation follows from general thermodynamic principles.

The representation of a deformable polycrystalline material by an information system is based on the idea that matter and energy are the products of the process of information transmission and storage [3, 4]. Without pretending to the strict definition of the term “information,” we point to its three most general properties: information cannot exist without the interaction of objects (in our case, deformation defects), information is not lost by any of them during the interaction, and the interaction leads to the formation of a new complicated system in which new information properties not found in the initial systems may emerge.

Thus, as applied to our problem, any interaction between the deformation defects during which one of them acquires information and the other does not loose it is called information interaction. Information interactions are asymmetric. We can distinguish conceptually coupled and transmitted informations. Coupled, or structural, information $C(S)$ is defined by a set of statistical data on the state of mesodefects for each structural level. Coupled information is a measure of the complexity of the material structure under deformation:

$$\begin{aligned} C(S) &= -(f_1 \log f_1 + f_2 \log f_2 + f_n \log f_n) \\ &= -\sum f_k \log f_k, \end{aligned} \tag{3}$$

where f_k are the parameters of the frequency distribution of the deformation mode for mesodefects at a structural level considered.

Transmitted information $I(S)$ is represented as the map of lower level coupled information onto a structure of higher hierarchical level and is defined as conventional information

$$I(S) = -\left[\sum f_k(\varepsilon)\log f_k(\varepsilon) - \sum f_{k\omega}(\varepsilon)\log f_{k\omega}(\varepsilon) - \sum f_{k\gamma}(\varepsilon)\log f_{k\gamma}(\varepsilon)\right], \tag{4}$$

where $\sum f_k(\varepsilon)\log f_k(\varepsilon)$ is the coupled information at a structural level considered, and $\sum f_{k\omega}(\varepsilon)\log f_{k\omega}(\varepsilon)$ and $\sum f_{k\gamma}(\varepsilon)\log f_{k\gamma}(\varepsilon)$ is the conventional information due to the linear plastic deformation, which depend on the shear and rotational deformation components.

EXPERIMENTAL BASIS

The experimental validation of the formalization principles discussed in this study is based on the investigation into the field of displacements of the nodes of a dividing network applied directly on a deformable sample. A mesh of the network 10 μm in size is the geometrical image of a mesodefect on a lower structural level. More stable structural levels are studied by selecting a set of model materials (such as MO copper) undergoing static deformation under different conditions (grain size, rate of loading, operating length of the sample, and rigidity of the tester). We designed a 2^4 factorial experiment to find samples with the highest plasticity, those with the highest strength, and those having a set of desired characteristics. The field of displacements is transformed into the field of distortions $\beta_{ij} = \varepsilon_{ij} + \omega_{ij}$ by numerical differentiation. For the plane case, all components of the tensor β_{ij} for the symmetric and antisymmetric components are calculated by the Cauchy formulas.

For a network mesh with the size i, the linear, shear, and rotational components of the plastic microstrains determined over an ensemble of mesodefects within the length x are given by some random functions $\varepsilon_{ij}(x_i|e_j)_i$, where $i = 10, 20, \ldots, 120$ μm and j is the average strain.

The distinctive feature of plastic microstrains on the mesoscopic level is their oscillating character and nonuniform spatial distribution. They can be considered as relatively periodic processes that satisfy the general definition [5]

$$|\varepsilon_i[x + T(x)] - \varepsilon_i(x)| \le \varepsilon, \quad \varepsilon > 0. \tag{5}$$

On a given structural level, the process is described by a number of dynamic components $Y_i(x)$ in form (1) and random quantities $R(x)$ characterizing relations between mesodefects inside the level:

$$H(x) = \sum_n^1 Y_i(x) + R(x). \tag{6}$$

Using such a representation and the resonance search method we calculated the periodograms for the linear, shear, and rotational modes of four discrete strains under the conditions of the factorial experiment. Each of the periodograms characterizes the number of harmonics of certain wavelength and amplitude in the spectrum of corresponding strain. The phase diagram is plotted in the M operator–N operator coordinates; the bifurcation diagram, in the plastic microstrain–period coordinates. The method of sliding frequency window was used to find bifurcations present in the phase diagram and boundaries between domains of attraction for different attractors [6].

From the periodograms of the random functions $\varepsilon_{ij}(x_i|e_j)_i$, the synchronization coefficient for the set of deformation defects associated with mesh sizes from 10 to 120 μm was calculated [7, 8].

Statistical processing included the search for the plastic microstrain distribution, the calculation of the autocorrelation functions, as well as tests for stationarity and ergodicity of the random functions $\varepsilon_{ij}(x_i|e_j)_i$.

DISCUSSION

Given a set of periodograms, the state of a polycrystalline material strained can be characterized on each of the levels as the dynamic evolution of the frequency distribution. It is assumed that the material as a nonlinear system has a spectrum of unrealized but realizable stationary states. The analysis of the periodograms of the translational and rotational deformation modes shows that the frequency spectrum varies during the deformation process: spectral harmonics appear and/or disappear, which is characteristic of only nonlinear systems (Fig. 1). Two, convergent and divergent, stages of self-organization can be discriminated from the periodograms. Analytically, the latter is associated with an increase in the number of harmonics in the spectrum; the former, with a decrease in this number.

The joint analysis of the phase diagrams for linear, rotational, and shear modes in the material as a dynamic system shows that dissipative structures develop in a certain sequence. They show up as bifurcations of different codimensionality along with regular and chaotic attractors. A qualitative modification of the phase portrait is represented by a bifurcation; an established motion, by an attractor. In these terms, for a poly-

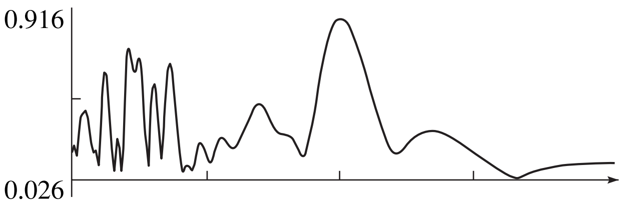


Fig. 1. Periodogram of linear plastic microstrains. The material is MO copper with an average strain of 2%.

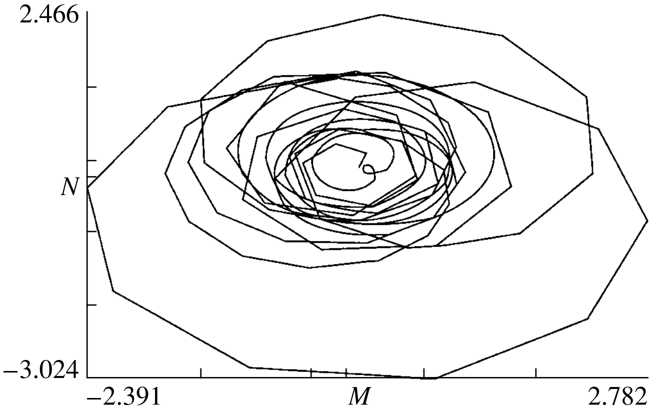


Fig. 2. Phase diagram for the shear deformation mode. The material is MO copper with an average strain of 2%.

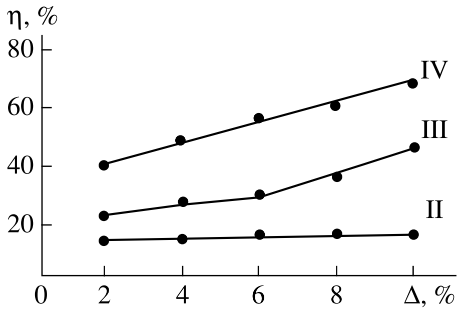


Fig. 3. Graphical identification of the structural levels from the synchronization coefficient η vs. deformation Δ dependence. The defect size is IV, 10–20; III, 40–60; and II, 80–120 μm.

crystalline material deformed under different conditions, one can indicate general rules and distinguishing features.

A material deformed can be represented as a structurally stable system; i.e., the topological structure of the phase diagram is stable in a range of strains up to some critical value at which one of the deformation modes bifurcates.

The stability of dissipative, for example, high-plasticity structures is maintained owing to the strictly dynamic shear–rotation–shear process, so that the stable limit cycle persists on the phase diagram in one of the deformation modes (shear or rotational) at any instant (Fig. 2).

From the topological structure of the phase diagrams on the different levels, one can judge the fulfillment of the scale invariance principle and, hence, the stability of the dissipative structures.

The considerable difference between the representative volumes under the conditions of high plasticity and high strength indicates that plastic deformation develops more steadily.

The coordinated activity of the whole complex system of interacting deformation defects on each of the structural levels is provided by synchronization, which means the most general case of establishing certain frequency relations between individual deformation defects and groups of defects. A quantitative measure of synchronization is the synchronization coefficient (SC) [2]. The SC is calculated from the periodograms of linear plastic microstrains for a set of deformation defects identified with mesh sizes from 10 to 20 μm.

From the variation of the SC, one can divide deformation defects into three groups and specify the hierarchy of structural levels: IV, the level of intragrain plastic strains with defect sizes of 10 and 20 μm; III, the level of intergrain plastic strains with defect sizes of 40 and 60 μm; and II, the level of interaction between groups of grains (80 and 120 μm) (Fig. 3). The classification by SC illustrates the space–time development of plastic deformation. Synchronization, or the occurrence of coherence, is a mechanism of self-organization.

Generally speaking, it is stable self-organization processes which are central to the concept of a dissipative structure in synergetics. The universality of this concept implies the possibility of geometrical characteristics being replaced by their information and numerical analogs.

Thus, the concepts of synergetics offer scope for a unified approach to formalizing all known models of plastic deformation in polycrystals on the mesoscopic level.

REFERENCES

1. V. E. Panin, *Physical Mesotechnology and Computer-Aided Design of Materials* (Nauka, Novosibirsk, 1995), Vols. I, II.
2. V. A. Likhachev, V. E. Panin, and E. É. Zosimchuk, *Cooperative Deformation Processes and Localization of Deformation* (Naukova Dumka, Kiev, 1989).
3. V. V. Rybin, *High Plastic Strains and Metal Failure* (Metallurgiya, Moscow, 1986).
4. S. Ya. Berkovich, *Cellular Automata as a Model of Reality: Search for New Representations of Physical and Informational Processes* (Mosk. Gos. Univ., Moscow, 1993).
5. M. G. Serebryannikov and A. A. Pervozvanskiĭ, *Detection of Latent Periodicities* (Nauka, Moscow, 1965).
6. O. D. Shevchenko, V. V. Ostashev, and V. K. Fedyukin, Tr. Pskov. Politekh. Inst., No. 2, 142 (1998).
7. V. V. Ostashev and O. D. Shevchenko, in *Proceedings of the XXXV Workshop “Topical Problems of Strength,” Pskov, 1999*, Part II, p. 433.
8. V. V. Ostashev and O. D. Shevchenko, Pis’ma Zh. Tekh. Fiz. **24** (16), 50 (1998) [Tech. Phys. Lett. **24**, 645 (1998)].

Translated by D. Zhukhovitskiĭ

Technical Physics, Vol. 47, No. 4, 2002, pp. 419–425. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 46–52.
Original Russian Text Copyright © 2002 by Smolyar, A. Eremin, V. Eremin.

SOLIDS

The Distribution of Liberated Energy and Injected Charge under Normal Incidence of a Fast Electron Beam on a Target

V. A. Smolyar, A. V. Eremin, and V. V. Eremin
Volgograd State Technical University, pr. Lenina 28, Volgograd, 400131 Russia
e-mail: rector@vstu.ru

Received July 23, 2001

Abstract—Formulas for the distributions of the liberated energy and injected charge are derived within the diffusion model of the kinetic equation for an electron beam normally incident on the target. The theory does not involve empirical adjustable parameters. The calculation of the liberated energy for the case of a planar directional electron source in an infinite medium (C, Al, Sn, and Pb) correlates well with the Spencer data, obtained from the exact solution of the Bethe equation, which is also used as the basis in the diffusion model. *© 2002 MAIK “Nauka/Interperiodica”.*

INTRODUCTION

As follows from [1], the diffusion approximation of the kinetic equation can be used as a basis for constructing a backscattering and penetration model for electrons incident on a target. Unlike the starting Bethe diffusion model [2] and semiempirical models based on it [3], our theory does not involve adjustable parameters needed for backscattering coefficients to reasonably agree with associated experimental data. The diffusion model has been extended up to final kinetic expressions for the backscattering and penetration of electrons and supplements, along with other analytical approaches intended to simplify the kinetic equation [4], the available software suites developed for Monte Carlo numerical simulation (such as the ETRAN program [5]) or for numerical simulation with semiempirical algorithms (the EDMULT program [6]). The model suggested in [1] is a mathematically closed diffusion approximation of the kinetic equation. Therefore, the good fit of the analytical backscattering coefficients to the experimental values allows us to hope that the analytical energy and injected charge distributions will not require empirical adjustable parameters as well.

BASIC STATEMENTS OF THE DIFFUSION MODEL OF THE KINETIC EQUATION

In the Bethe designations, the kinetic equation for the electron density $f(\mathbf{r}, \mathbf{\Omega}, t)$, where $\mathbf{r}$ is the coordinate space, $\mathbf{\Omega}$ covers motion directions, and t is a time instant uniquely related to a path s, has the form [2]

$$\frac{\partial f}{\partial s} = \frac{\partial f}{v \partial t} = -\mathbf{\Omega} \cdot \nabla f + \int d\mathbf{\Omega}' N\sigma(v, \mathbf{\Omega}' \cdot \mathbf{\Omega}) \times [f(\mathbf{r}, \mathbf{\Omega}, t) - f(\mathbf{r}, \mathbf{\Omega}, t)], \tag{1}$$

where v is the electron velocity, s is the path length, N is the volume atom concentration, and σ is the differential scattering cross section.

The problem of an electron beam striking a target should be complemented by an electron source located on the target surface and emitting electrons along the z axis, which is aligned with the inner normal $\mathbf{n}$ to the surface:

$$S(\mathbf{r}, \mathbf{\Omega}, s) = S_0 \delta(x)\delta(y)\delta(z)\delta(\mathbf{n} - \mathbf{\Omega})\delta(s), \tag{2}$$

where S_0 is the emission rate of the source, that is, the number of electrons emitted in unit time.

Having penetrated into the target, the beam is scattered and becomes isotropic with time. When the directional electron anisotropy becomes small, the diffusion approximation is applied. With this approximation, only the first two terms can be left in the expansion of $f(\mathbf{r}, \mathbf{\Omega}, t)$ in Legendre polynomials:

$$f(\mathbf{r}, \mathbf{\Omega}, s) = \frac{1}{4\pi}(f(\mathbf{r}, s) + 3\mathbf{\Omega} \cdot \mathbf{J}(\mathbf{r}, s)), \tag{3}$$

where

$$F(\mathbf{r}, s) = \int f(\mathbf{r}, \mathbf{\Omega}, s) d\mathbf{\Omega}$$

is the electron density and

$$\mathbf{J}(\mathbf{r}, s) = \int \mathbf{\Omega} f(\mathbf{r}, \mathbf{\Omega}, s) d\mathbf{\Omega}$$

is the electron flow density vector.

Eventually, kinetic equation (1) is reduced to the diffusion equation

$$\frac{\partial F}{\partial s} = -\frac{\lambda(s)}{3}\Delta F \tag{4}$$

and the equation for the diffusion flow

$$\mathbf{J} = -\frac{\lambda(s)}{3}\nabla F. \tag{5}$$

Here, λ is the transport length defined by the integral

$$\frac{1}{\lambda(s)} = 2\pi\int_0^{\pi}\sin(\Theta)d\Theta(1-\cos(\Theta))N\sigma(\Theta, v(s)), \tag{6}$$

where Θ is the scattering angle.

At the interface, there is no electron flow from the free space into the target other than incident flow (2), so that

$$f(\mathbf{r}, \mathbf{\Omega}, s) = 0, \quad \mathbf{\Omega}\cdot\mathbf{n} > 0, \quad s > 0. \tag{7}$$

Condition (7) cannot be met in diffusion approximation (4); however, it can be satisfied having been direction-averaged in the target. Substituting (4) and (5) into (3) and averaging the electron flow inside the target over angles, we come to the boundary condition

$$F(x, y, 0, s) = \frac{2\lambda}{3}\frac{\partial F}{\partial z}(x, y, 0, s). \tag{8}$$

Thus, the basic statements of the diffusion model are as follows.

(i) Kinetic equation (1) is replaced by Eqs. (4) and (5), which suggests that the model is based on the diffusion approximation.

(ii) Incident flow (2) is changed to an isotropic source, which is placed at a total diffusion depth z_d and emits electrons with an energy E_0 that have an initial range shorter by z_d.

(iii) Additional condition (7) is imposed at the boundary with the free space. It implies the absence of the diffusion flux from the free space into the target.

ANALYTICAL SOLUTION TO THE PROBLEM OF AN ELECTRON BEAM INCIDENT ON A SEMI-INFINITE TARGET

Starting kinetic equation (1) assumes continuous deceleration; therefore, the total path of an electron in the target is obtained by integrating the reciprocal of the mean electron energy loss per unit path:

$$s(E_0, E) = \int_E^{E_0}\frac{dE}{\langle\varepsilon(E)\rangle}. \tag{9}$$

By definition,

$$\langle\varepsilon(E)\rangle = \int_0^{\varepsilon_{\max}}\varepsilon w_{\text{inel.}}(E, \varepsilon)d\varepsilon, \tag{10}$$

where $w_{\text{inel.}}(E, \varepsilon)$ is the number of electron inelastic collisions per unit length with an energy loss ε per collision and $\varepsilon_{\max}$ is the maximal energy loss.

As the total diffusion length, we take the mean displacement of stopped electrons emitted by the planar source, which is inserted into an infinite scattering medium. The displacement is counted normally to the source plane. The formula for the mean displacement of such electrons as a function of residual range was found by Lewis [7]:

$$\langle z(r)\rangle = \frac{R_0}{d+1}(1-r^{d+1}),$$

where $R_0 = s(E_0, 0)$ is the initial range of incident electrons, $r = R(E)/R_0$ is the residual range, and $d = R_0/\lambda$.

With this expression, one can locate an isotropic source that replaces the incident flow, i.e., find the total diffusion length:

$$z_d \equiv \langle z(r=0)\rangle = \frac{R_0}{d+1}. \tag{11}$$

According to statement (ii) of the diffusion model, the path length ξ of diffusing electrons and the total path length of electrons that is counted from the point where the electrons enter the target [formula (9)] relate as

$$\zeta = \frac{R_0 - z_d}{R_0}s(E_0, E).$$

Thus, the diffusion model of the problem of incident electron flow of unit intensity (Eqs. (1) and (2)) is mathematically reduced to the equation of diffusion from δ source:

$$\frac{\partial F}{\partial\zeta} = -\frac{\lambda(\zeta)}{3}\Delta F + \delta(x)\delta(y)\delta(z-z_d)\delta(\zeta), \tag{12}$$

which has the boundary conditions

$$\frac{1}{2}F(x, y, 0, \zeta) = \frac{\lambda(\zeta)}{3}\frac{\partial F}{\partial z}(x, y, 0, \zeta). \tag{13}$$

The problem stated by Eqs. (12) and (13) can be solved numerically with standard techniques. To analytically solve this problem, we introduce a new variable— the “age” of electrons,

$$\tau(\zeta) = \frac{1}{3}\int_0^{\zeta}\lambda(\zeta)d\zeta \tag{14}$$

—and average the diffusion coefficient in boundary condition (13). Then, diffusion equation (12) and boundary condition (13) will take the form

$$\frac{\partial F}{\partial\tau} = \Delta F + \delta(x)\delta(y)\delta(z-z_d)\delta(\tau), \tag{15}$$

$$F(x, y, 0, \tau) = \frac{2}{3}\langle\lambda\rangle\frac{\partial F}{\partial z}(x, y, 0, \tau) \tag{16}$$

and have constant coefficients. Here, ⟨λ⟩ is age-averaged transport length:

$$\langle\lambda\rangle = \frac{1}{\tau_0}\int_0^{\tau_0}\lambda(\tau)d\tau,$$

where τ_0 is the age of the stopped electrons that is calculated by formula (14), where the upper limit of the integral equals $R_0 - z_d$.

The solution of problem (15), (16) yields the density of the electrons with the age τ:

$$F(x, y, z, \tau) = \left\{\frac{1}{2\sqrt{\pi\tau}}\left[\exp\left(-\frac{(z-z_d)^2}{4\tau}\right) + \exp\left(-\frac{(z+z_d)^2}{4\tau}\right)\right] - a\exp(a^2\tau + a(z+z_d)) \times \operatorname{erfc}\left(a\sqrt{\tau} + \frac{z+z_d}{2\sqrt{\tau}}\right)\right\}\left(\frac{1}{2\sqrt{\pi\tau}}\right)^2\exp\left(-\frac{x^2+y^2}{4\tau}\right), \quad (17)$$

where $a = 2\langle\lambda\rangle/3$.

The density of the remaining electrons calculated by formula (17) and multiplied by the electron charge represents the distribution of the injected charge in the target for the unit intensity of the incident beam. For the beam intensity S_0 electrons per second, we obtain the density of the charge injected in unit time:

$$\frac{dQ(x, y, z)}{dV} = S_0 eF(x, y, z, \tau_0), \quad (18)$$

where e is the electron charge and dV is an element of the volume.

The backscattering coefficient is found by subtracting the density of the stopped electrons [formula (17)] integrated over the target volume from unity. The formula for the backscattering coefficient thus obtained has the form

$$b = \operatorname{erfc}\left[\frac{z_d}{2\sqrt{\tau_0}}\right] - \exp(a^2\tau_0 + az_d)\operatorname{erfc}\left(a\sqrt{\tau_0} + \frac{z_d}{2\sqrt{\tau_0}}\right). \quad (19)$$

Now, we will calculate the density of the beam energy liberated in the target (for brevity, the liberated energy density), which is designated as $dW(x, y, z)/dV$. By definition, the liberated energy density is equal to the energy lost by all electrons in unit volume per second. Therefore,

$$\frac{dW}{dV} = \int_0^{E_0} F(x, y, z, E)\xi(E)dE, \quad (20)$$

where $F(x, y, z, E)dE$ is the number of electrons with an energy in the interval $(E - dE, E + dE)$ per unit volume and $\xi(E) = |dE/dt|$ is the mean energy lost by electrons with an energy E per second.

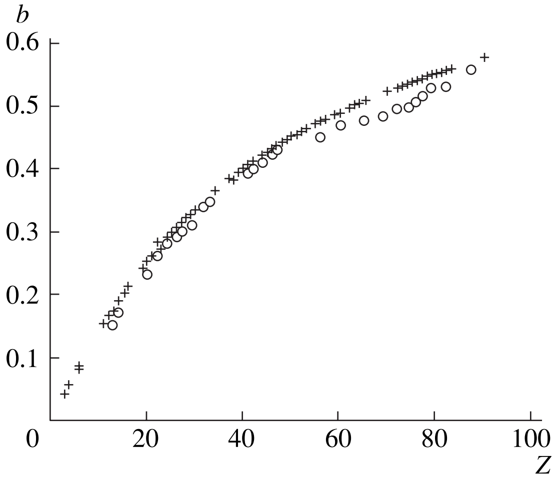


Fig. 1. (+) Analytical and (o) experimental [9] values of the backscattering coefficients for 40-keV electrons normally incident on the target.

With regard for the fact that

$$F(x, y, z, E)dE = F(x, y, z, t(E))dt,$$

we obtain

$$\frac{dW}{dV} = \int_0^{E_0} F(x, y, z, t(E))dE, \quad (21)$$

instead of (20), where $F(x, y, z, t)$ is the solution of kinetic equation (1), which was integrated over all directions and has form (17) in the diffusion approximation.

Until now, we have not used any scattering features. For this reason, the aforesaid equally refers to electrons, positrons, and light ions. To obtain numerical results, it is necessary to specify the elastic and inelastic scattering cross sections in formula (6) for transport length and in formula (10) for energy loss per unit path length for given target material and energy of penetrating particles.

RESULTS AND DISCUSSION

The target material and the electron energy enter into the diffusion model through the transport length and mean energy loss per unit length, rather than through the elastic and inelastic scattering cross sections as in (1). The transport length and the mean energy loss per unit length were calculated with Spencer standard formulas [8]:

$$\frac{1}{\lambda(T)} = 2\pi r_e^2 N(Z+1)\frac{(T+1)^2}{T^2(T+2)^2} \times [\ln(1 + 1/\eta) - 1/(1+\eta)], \quad (22)$$

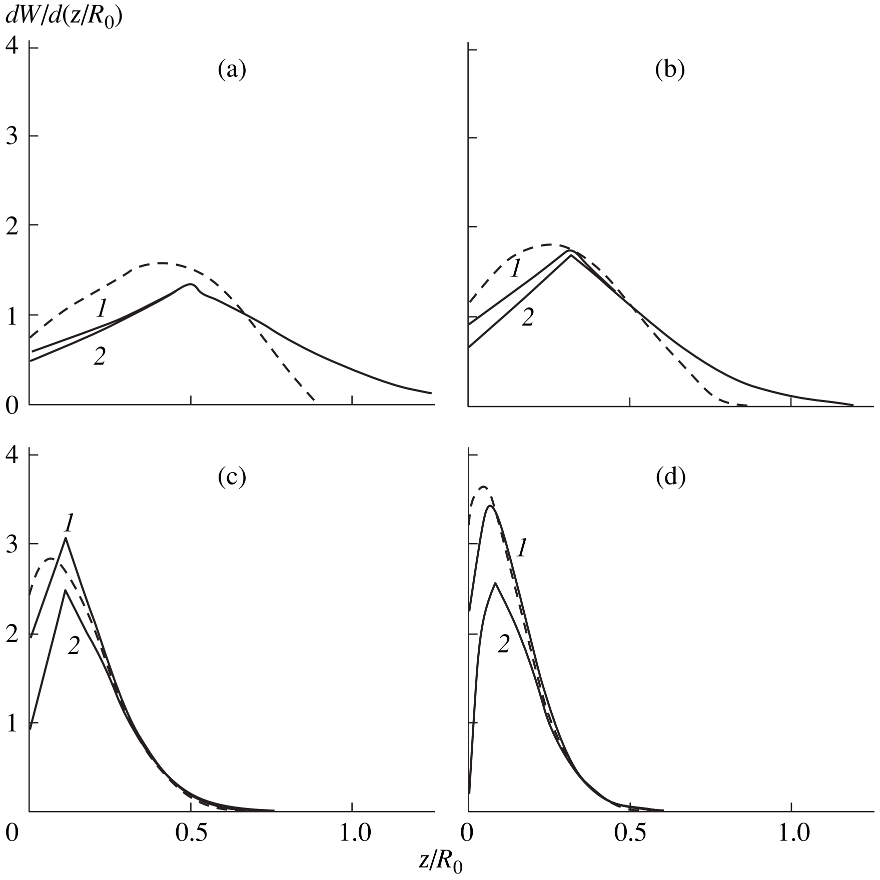


Fig. 2. Liberated energy distribution for a planar directional source of electrons with an energy E_0 placed into an infinite scattering medium (dashed lines, exact solutions [12]; continuous lines *1*, diffusion model) and into a semi-infinite target (continuous curves *2*, diffusion model). The linear energy density $dW/d(z/R_0)$, corresponding to the unit intensity of the source, is converted to dimensional units via the factor E_0/R_0. (a) C, (b) Al, (c) Sn, and (d) Pb. E_0 = (a, b) 25, (c) 50, and (d) 100 keV.

$$\langle \varepsilon(T) \rangle = 4\pi r_e^2 NZ \frac{(T+1)^2}{T(T+2)} \ln[1.166T/I(Z)], \qquad (23)$$

where $T = E/mc^2$ is the energy in units of the rest energy of an electron, r_e is the classical radius of an electron, N is the number of atoms in unit volume,

$$\eta = 1.75 \times 10^{-5} Z^{2/3}/[T(T+2)]$$

is the parameter of nucleus screening by atomic electrons, and

$$I(Z) = \begin{cases} 13.6Z, & Z < 10 \\ (9.76 + 58.8Z^{-1.19})Z, & Z \geq 10 \end{cases}$$

is the mean ionization potential of target atoms (in electronvolts).

Formulas (22) and (23) are valid for energies from $I(Z)$ to 1 MeV; i.e., in the interval where energy losses are defined by the ionization and excitation of atoms and radiation losses can be neglected.

Figure 1 compares the backscattering coefficients calculated by formula (19) for the normal incidence of electrons on the target with experimental data obtained in [9]. The calculations were made for solid elements from lithium to uranium, whose densities are given in [10]. From Fig. 1, it follows that the diffusion model gives fairly exact values of the backscattering coefficient. However, it cannot confirm the dependence of the backscattering coefficient on the atomic number of the target, which was found in [11].

The discrepancy between the model calculations of the liberated energy and the actual values may be checked with the US NBS tables [12], which list the Spencer data [8] for the liberated energy density distributions in graphite, aluminum, tin, and lead. This comparison is of interest also due to the fact that the Bethe kinetic equation is used as basic in the diffusion model suggested. Spencer proceeds from Bethe kinetic equation (1) and solves the problem of energy liberation for a monoenergetic electron source placed into an infinite scattering medium. Its theory does not involve adjustable parameters, and the energy liberation is expressed through the double spatial-range moments of the coefficients of the $f(\mathbf{r}, \mathbf{\Omega}, t)$ expansion in Legendre polynomials. Spencer directly looks for the liberated energy

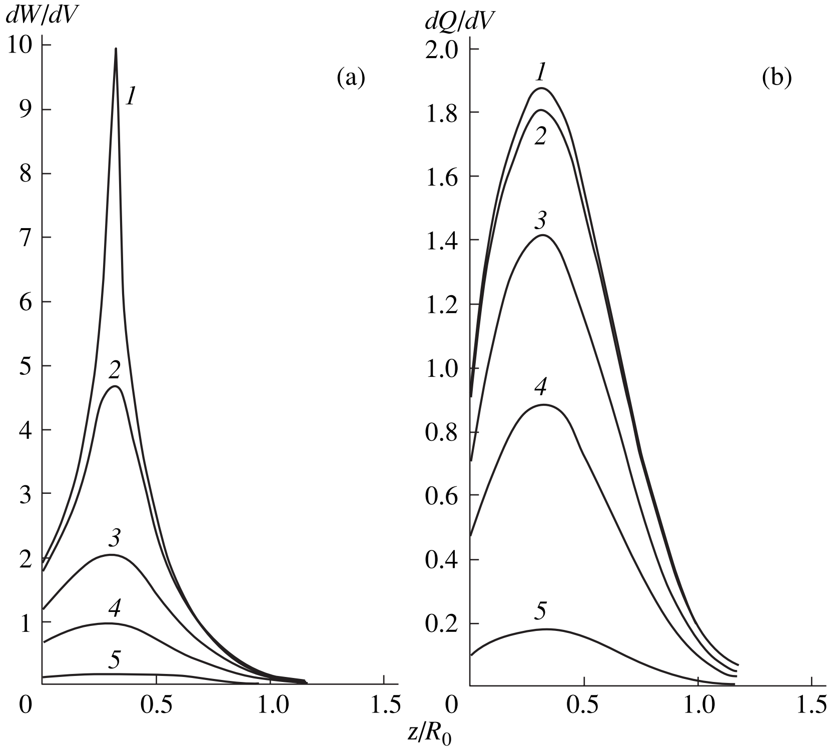


Fig. 3. Distributions of the (a) energy density and (b) injected charge in a semi-infinite aluminum target for a normally incident sharply focused electron beam with an energy E_0 = 25 keV. The distance r from the beam axis is expressed in units of the initial range R_0: r/R_0 = (*1*) 0.05, (*2*) 0.1, (*3*) 0.25, (*4*) 0.4, and (*5*) 0.7. The energy density dW/dV is converted to dimensional units via the factor E_0/R_0^3. The injected charge density dQ/dV is converted to dimensional units via the factor e/R_0^3, where e is the electron charge. Both values correspond to the unit intensity of the source.

density, not the solution of the kinetic equation. Our calculations of the stopping power, total trajectory range, and elastic collision transport length performed for aluminum at an electron energy of 25 keV using the cross sections available from the database of the Ioffe Physicotechnical Institute, Russian Academy of Sciences [13], virtually coincide with the values listed in the Spencer tables [12].

In Fig. 2, the liberated energy density calculated by (21) and integrated over the coordinates transverse to the z axis, that is, the distribution of the linear density of the energy liberated from a wide monoenergetic electron beam, is compared with the Spencer calculations [12]. Work [12] tabulates the distribution of the linear density of the energy liberated when a planar electron source placed into an infinite medium unidirectionally emits monoenergetic electrons normally to the source plane. The discrepancy between our calculations for a semi-infinite target and the tabulated energy values for an infinite medium can partially be explained by the fact that the electrons penetrating into an infinite medium experience multiple scattering and may find

Table

Element	E_0, keV	b	b_E	$\langle E_b\rangle/E_0$	W/E_0	W_s/E_0
C	25	0.073	0.051	0.703	0.949	0.972
Al	25	0.159	0.118	0.741	0.882	0.923
Sn	50	0.415	0.343	0.824	0.656	0.787
Pb	100	0.512	0.437	0.852	0.563	0.725

Note: b, backscattering coefficient; b_E, coefficient of backward dissipated energy (the ratio of the backscattered electron energy to the energy of the incident electron beam); $\langle E_b\rangle/E_0$, ratio of the mean energy of a backscattered electron to the energy of an incident electron; W/E_0, ratio of the energy absorbed by the target per incident electron to the energy of an incident electron (this work); and W_s/E_0, the same taken from the Spencer tables [12].

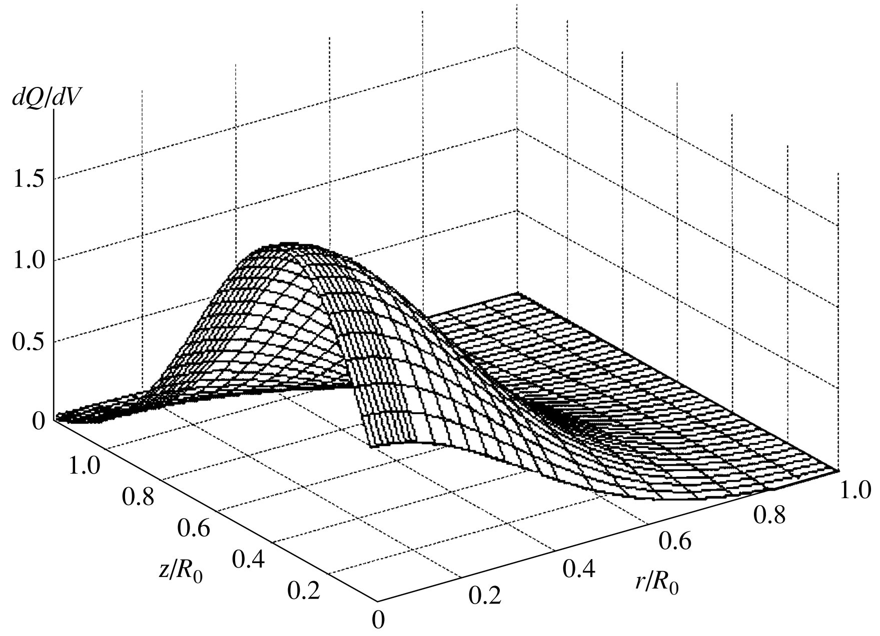


Fig. 4. Injected charge density distribution in a semi-infinite aluminum target for a normally incident sharply focused electron beam with an energy E_0 = 25 keV. The z axis is directed normally to the surface from the point of beam incidence. r is the distance to the z axis. The injected charge density dQ/dV, corresponding to the unit intensity of the source, is converted to dimensional units via the factor e/R_0^3, where e is the electron charge and R_0 is the initial electron range.

themselves on the left of the source, which emits to the right along the z axis, and then again pass to the right half of the space. Such transitions may occur repeatedly; consequently, in an infinite medium, the energy density on the right of the source plane must be larger than that near a semi-infinite target subjected to electron bombardment. The diffusion model does show that the energy density near the surface bombarded is lower than that in an infinite medium, as follows from Fig. 2. The table shows the backscattering coefficients and the fractions of the backward dissipated energy obtained by integrating over the space of energy density and injected charge (see Fig. 2).

The larger the atomic number of an element and the smaller the electron energy compared with the initial energy, the smaller the discrepancy between the diffusion model and the rigorous solutions. This feature of the diffusion model is also observed in the liberated energy distributions shown in Fig. 2.

It can be expected that the injected charge distribution simulated within the diffusion model must be closer to reality than the liberated energy distribution, since the latter characteristic is more integral. As follows from Fig. 3, the distributions of the injected charge and liberated energy for the case of a sharply focused beam incident normally to the target surface (formulas (17) and (18)) greatly diverge. A sharp peak of the energy density at the depth of total diffusion is due to the fact that, according to the diffusion model, an incident beam is replaced by an isotropic electron source at this depth. Here, the discrepancy between the diffusion model and physical reality shows up most vividly. If the beam is wide and planar, the discrepancy decreases, yet remaining appreciable, as follows from the curves in Figs. 2 and 3. The injected charge distribution is, in essence, the distribution of the stopped electrons. That is why the location and the geometry of the source, be it an incident beam or an isotropic source located at the diffusion depth, is of minor significance. Figure 4 demonstrates the detailed three-dimensional injected charge distribution obtained by formula (18). It is seen that this distribution is much smoother than the energy distribution.

Thus, the model suggested in this work is a mathematically closed diffusion approximation of the kinetic equation that describes the problem of an electron beam incident on a target. The model does not require adjustable parameters to bring the backscattering coefficients, the liberated energy distribution, and, apparently, the injected charge distribution into coincidence to experimental data.

ACKNOWLEDGMENTS

The authors thank the developers of the electron scattering archive (Ioffe Physicotechnical Institute, Russian Academy of Sciences) for submission of the database.

REFERENCES

1. V. A. Smolyar and A. V. Eremin, Radiotekh. Élektron. (Moscow) **46** (5), 599 (2001).
2. H. Bethe, M. E. Rose, and L. P. Smith, Proc. Am. Philos. Soc. **78** (4), 573 (1938).
3. K. Kanaya and S. Okayama, J. Appl. Phys. **5**, 43 (1972).
4. L. A. Bakaleĭnikov, S. G. Konnikov, K. Yu. Pogrebitskiĭ, *et al.*, Zh. Tekh. Fiz. **64** (4), 9 (1994) [Tech. Phys. **39**, 354 (1994)].
5. S. M. Seitzer, Appl. Radiat. Isot. **42**, 917 (1991).
6. T. Tabata *et al.*, Radiat. Phys. Chem. **35**, 821 (1990).
7. H. W. Lewis, Phys. Rev. **78**, 526 (1950).
8. L. V. Spencer, Phys. Rev. **98**, 1597 (1955).
9. A. Ya. Vyatskin, A. N. Kabanov, B. N. Smirnov, and V. V. Trunev, Radiotekh. Élektron. (Moscow) **21**, 895 (1976).
10. *Tables of Physical Quantities: Handbook*, Ed. by I. K. Kikoin (Atomizdat, Moscow, 1976).
11. Yu. D. Kornyushkin, Zh. Tekh. Fiz. **69** (6), 40 (1999) [Tech. Phys. **44**, 645 (1999)].
12. L. V. Spencer, *Energy Dissipation of Fast Electrons* (National Bureau of Standards, Washington, 1959), Monograph 1.
13. http://www.ioffe.rssi.ru/ES.

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 4, 2002, pp. 426–431. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 53–59.
Original Russian Text Copyright © 2002 by Mordvintsev, Kudryavtsev.

SOLID-STATE ELECTRONICS

Conditions for the Production of a Single Conducting Nanostructure by Electroforming

V. M. Mordvintsev and S. E. Kudryavtsev

Institute of Microelectronics and Informatics, Russian Academy of Sciences,
ul. Universitetskaya 21, Yaroslavl, 150007 Russia
e-mail: mvm@imras.yar.ru

Received January 22, 2001

Abstract—The process of electroforming (the production of a carboniferous conducting medium when the current flows through an organic material under a high electric field) in open sandwichlike structures with an insulating gap several tens of nanometers in width is considered. It is shown experimentally that there are factors that both favor (external ballast resistor) and prevent (local spreading resistance and the presence of the initial conductivity in the insulating gap) the production of a single conducting element between the electrodes. A simple model of the process in terms of the equivalent electric circuit is proposed. The model helps to find the tradeoff between these factors and to construct an *I–V* diagram, which exhibits a region within which a single conducting nanostructure can be electroformed. An expression that relates the minimum permissible resistance of the nanostructure to its geometric parameters is derived. *© 2002 MAIK “Nauka/Interperiodica”.*

INTRODUCTION

Electroforming in metal–insulator–metal (MIM) structures [1, 2] has been studied for a long time. The analysis of a huge body of experimental data (including those obtained recently) shows the following [3]. Electroforming can be considered as a self-organization process in the nanometer insulating gap in the carboniferous conducting medium between the metallic electrodes of the MIM structure. The carboniferous medium is produced in an organic material coming from the outside (from the gas phase or, e.g., from a photoresist film deposited on the structure) whose conductivity changes when the organic molecules are destroyed by an electron impact as a result of current passage.

In [4], a so-called MIM nanodiode with a carboniferous active medium was designed. It is based on an open sandwichlike MDM (metal–dielectric–metal) structure where an insulating gap exposed to an organic material is formed at the end face of the dielectric film with a thickness of several tens of nanometers by locally etching two upper layers of a conventional sandwichlike MDM structure. Electroforming in open sandwichlike Al–Al_2O_3–W structures with an oxide thickness of 20–40 nm [5] showed that, under certain conditions, the carboniferous conducting medium forms only at one site of the end face perimeter over a length comparable to the insulating gap width. This fact indicates that a single conducting structure with nanometer sizes in all three dimensions arises.

From the practical point of view, the design of the MIM nanodiode in the form of an open sandwichlike Si–SiO_2–W structure [6] is promising for nonvolatile memory. In this structure, electroforming greatly differs in comparison with that in the Al–Al_2O_3–W structure. Here, the carboniferous conducting medium tends to form along the entire perimeter of the insulating gap simultaneously. This makes it difficult to form the single conducting nanostructure. The basic factors responsible for such behavior are the high resistivity of the silicon electrode and the appreciable initial conductivity of the insulating gap surface. In this paper, we discuss mechanisms of the processes and study conditions under which it is possible to fabricate the single conducting nanostructures even in the presence of these unfavorable factors.

EXPERIMENTAL RESULTS AND MECHANISMS OF THE PROCESSES

It was shown [5] that the presence of a series ballast resistor R_b during the electroforming process is favorable to the formation of the single conducting element (nanostructure) in the insulating gap of an Al–Al_2O_3–W open sandwichlike structure where the dielectric is several tens of nanometers wide. The physical mechanism of the process is as follows. A high-value (no less than several megaohms) series-connected resistance provides strong degenerative feedback between the current through and voltage U across the structure when the supply voltage V gradually increases. When the current increases because of a rise in the conductivity of the insulating gap, the voltage across the gap drastically drops. As the field strength in the gap increases, some (first) nanotip on the cathode surface (tungsten) starts to emit, leading to the formation of the carboniferous conducting medium. As soon as this takes place, conditions for emission from other regions worsen, causing the single conducting element to appear. Its growth is accompanied with an increase in the conductivity, current, and, hence, voltage drop across the ballast resistance. Such a process lasts until the voltage U across the MIM structure (at a given R_b) decreases to the threshold value U_{th} (close to 3 V). The formation of the carboniferous conducting medium under these conditions becomes impossible (see below), and the growth of the nanostructure, as well as the increase in its conductivity, stop. When $R_b = 0$, a local electric breakdown develops in the Al–Al_2O_3–W structure.

In the open sandwichlike Si–SiO_2–W structures made on p-type boron-doped silicon wafers with different resistivities, electroforming shows a number of features [6]. First, the gap has an appreciable initial (i.e., existing before the electroforming) conductivity, which is proportional to the structure perimeter. This conductivity is presumably associated with the chemical etching of silicon oxide (setup and fabrication technology are detailed in [6]). Second, the forming can be carried out without the ballast resistance. In this case, no electrical breakdown is observed, unlike the Al–Al_2O_3–W structures, and the current in the highly conducting state is several orders of magnitude larger. This means that the electroforming takes place over the entire gap; i.e., the conducting medium forms along the whole exposed perimeter of the structure.

These features can be explained by the high resistivity of the silicon. Because of this, a large local resistance R_s to the carrier flow toward the silicon electrode is series-connected to every conducting nanostructure arising in the insulating gap. According to estimates [6], R_s is on the order of 1 MΩ for silicon with a resistivity of 10 Ω cm; therefore, the spreading resistance may substantially influence all the processes. In particular, it prevents local electrical breakdowns. Moreover, the voltage drop across the local spreading resistance decreases the voltage across a forming conducting element, leaving the total voltage across the gap unchanged. The growth of the corresponding nanostructure slows down and ceases when the threshold voltage is attained. At the same time, new conducting elements may originate in other regions of the gap. In other words, the local resistance series-connected to a conducting nanostructure stimulates electroforming over the entire gap (parallel electroforming) and suppresses the formation of a single nanostructure. The validity of this statement was confirmed in experiments with Si–SiO_2–W structures based on high-resistivity silicon (with a resistivity of 10 Ω cm). All attempts to produce a single conducting nanostructure under these conditions failed: in all the cases, parallel forming was developed.

The initial conductivity distributed over the insulating gap also favors parallel electroforming, because the electron flows initiate the formation of the conducting medium at many sites simultaneously. If this conductivity is too high, it is impossible to produce a single conducting nanostructure even in structures based on low-resistivity silicon (with a resistivity of 0.1 Ω cm), as follows from experiments.

The presence of the appreciable current before electroforming allows one to determine the threshold voltage U_{th}, which initiates the formation of the carboniferous conducting medium from organic molecules. This voltage switches the electroformed structure from the low- to the high-resistivity state. Usually, this voltage is about 3 V [2]. Unlike Al–Al_2O_3–W structures [5], here it is not associated with the voltage required to induce the current inside the insulating gap. The time dependences of the current at fixed voltages U across the Si–SiO_2–W structure are shown in Fig. 1. The smooth initial current is observed when U is less than some threshold value U_{sf}. At $U > U_{sf}$, specific current jumps with the gradual increase in its mean value are seen, which means the beginning of electroforming. The process starts with a time delay, which is the smaller, the higher U. The value of U_{sf} is in the range between 4 and 5 V and is almost independent of the SiO_2 film thickness in the range from 10 to 40 nm. The fact that it is somewhat higher than the typical value 3 V implies a voltage drop across some structure elements. The value of U_{sf} is also independent of the radius of curvature of the tungsten electrode edge (the radius can be varied by varying tungsten film etching conditions), which seems

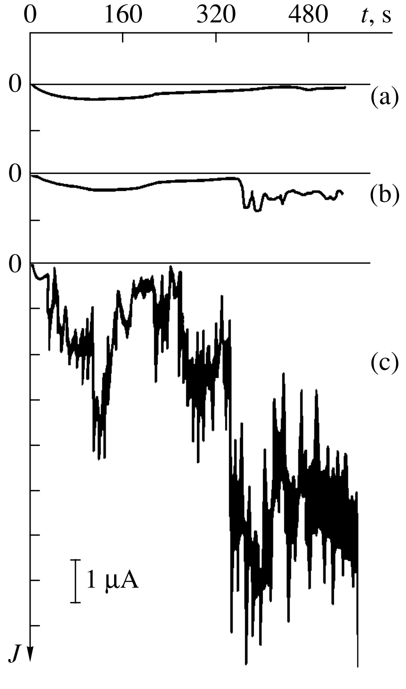


Fig. 1. Time dependences of the current through open sandwichlike Si–SiO_2–W structures at fixed voltages. The structures are covered by a resist layer. The resistivity of the silicon is 10 Ω cm, and the SiO_2 layer thickness is 11.4 nm. U = (a) 3.6, (b) 4.0, and (c) 4.4 V.

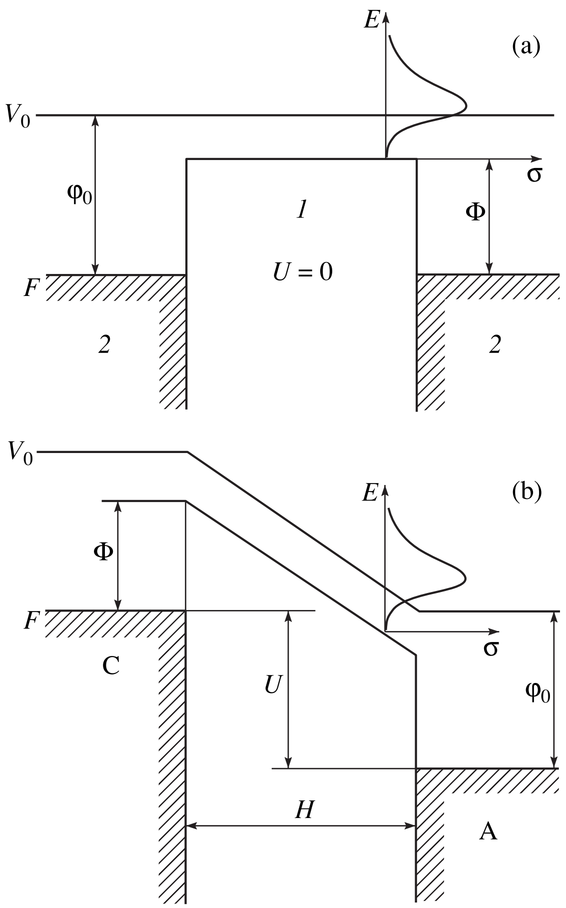


Fig. 2. Simplified potential diagrams of the MIM structure illustrating the mechanism responsible for the appearance of the threshold voltage U_{sf} (3–4 V) during the electroforming. The qualitative dependence of the cross section of dissociative electron attachment to an adsorbed molecule on the electron kinetic energy is shown. *1*, adsorbed organic insulator; *2*, metal; C, cathode; and A, anode.

to be quite natural unlike the forming in the Al–Al_2O_3–W structures.

Thus, we can argue that the initial smooth current is the prerequisite for the electroforming process (i.e., for the formation of the carboniferous conducting medium), which starts when the voltage U exceeds the threshold value U_{sf}. The aforesaid is in good agreement with the model discussed in [7]. The carboniferous conducting medium forms as a result of the destruction of organic molecules in the flow of slow electrons due to dissociative electron attachment. Having captured an electron, the molecule passes to the autodecay repulsive term of a negative ion [8]. The effective cross section σ of this process resonantly depends on the electron kinetic energy E (Fig. 2) and has a peak, which can be very close to $E = 0$ [9]. For the molecules adsorbed on the solid surface, which occurs in our system, the curve σ(E) is somewhat shifted downward from the vacuum level V_0, as qualitatively shown in Fig. 2a. The dissociation of the molecule is possible only if the electron is in an appropriate energy position. It can tunnel to the autodecay repulsive term of the molecule (Fig. 2b) if the voltage U applied to the nanometer gap of width H is somewhat higher than the potential barrier Φ (or somewhat lower than the work function φ_0). Since a typical value of the work function is 4.5–5.0 eV, the threshold voltage U_{sf} must be close to the physical parameter U_{th}, which is somewhat lower than φ_0 and is usually about 3 V. We will find the same value of U_{th} when electrons tunnel through a lower barrier and then are accelerated by the electric field up to a necessary kinetic energy.

The same mechanism apparently defines the threshold voltage for the modification of the solid surface (i.e., for molecular bond destruction) in a scanning tunneling microscope. Usually, this voltage is about 3–4 V [10–12]. This mechanism is also responsible for the specific peak appearing in the quasi-static *I*–*V* characteristics of electroformed MIM structures near 3–4 V [2]. The appearance of the peak at a decreased voltage can be explained by the fact that the formation of the carboniferous conducting medium (hence, the decrease in the insulating gap width and the change in the dielectric composition in the gap) takes place only at higher-than-threshold voltages.

Note that the threshold voltage U_{th} defined by the mechanism described above could not be found at the early stage of electroforming in the Al–Al_2O_3–W structures, because the initial conductivity here is lacking and the formation of the carboniferous conducting medium is limited by the current, which passes only at the voltage (electric field) providing electron emission from the cathode. The value of U_{sf} varied from 5 to 15 V [5] depending on the tungsten electrode geometry. In the Si–SiO_2–W structures, the current passes over the open face end even at very low voltages and does not control the formation of the conducting medium. Here, the process is limited by low electron energies. At higher voltages U, the value of U_{sf} (close to U_{th}) can reliably be detected.

Thus, we can suggest that the primary effect of the silicon substrate resistivity ρ on the electroforming process in the Si–SiO_2–W structures is that it governs the spreading resistance. As ρ decreases, the situation comes close to that in the MIM structures. Experimental data completely confirm this statement. Figure 3 shows the current–voltage characteristics taken in the process of electroforming for two structures with close thicknesses of silicon oxide and greatly differing (by two orders of magnitude) resistivities of the silicon wafers. As is known [7], a decrease in the current with increasing voltage is due to the partial burnout of the carboniferous conducting medium. In the low-resistivity silicon, the averaged current maximum is one order of magnitude lower and shifts to lower voltages and the current essentially oscillates. The distinctions from the case of the high-resistivity silicon are caused just by the decrease in the spreading resistance, because of which parallel forming is weakly pronounced: separate current peaks are the results of the forming in one or several regions of the insulating gap where the conducting medium forms and burns out before it forms in other regions. Therefore, the current amplitudes are separated in time and do not add up as in Fig. 3a.

For low-resistivity silicon, the conditions under which a single conducting nanostructure forms [6], as in the Al–Al_2O_3–W MIM structures, can readily be provided by using a high-value stair-step ballast resistor.

MODEL

Thus, there are several factors variously influencing the formation of the single conducting structure during electroforming in real Si–SiO_2–W structures. Conditions under which they are counterbalanced are of great interest. Strictly speaking, the formation of the conducting medium and the increase in the current should be described by a system of differential equations that constitute a difficult mathematical problem. However, it is possible to describe such objects in terms of the simplified equivalent circuit, which takes into account basic features of the process.

Electroforming is usually carried out under quasi-steady-state conditions, i.e., when the voltage varies slowly. Therefore, the averaged (not instantaneous) state of the conducting medium inside the insulating gap is of interest. Then, as the first approximation, one can use the simplest equivalent circuit, including resistors only. Such an equivalent circuit for a real metal–insulator–semiconductor structure is shown in Fig. 4. The external ballast resistance R_b is series-connected to the entire structure. The local spreading resistance R_s is series-connected to each of the regions of insulating gap *3* where the carboniferous conducting medium may form (nanotips *4* on cathode surface *1*). The initial conductivity distributed over the insulating gap is simulated by the resistance R_0. The spreading resistance series-connected to R_0 is ignored because the current spreads over a large area. The supply voltage V increases gradually (which is typical of electroforming), and the voltage U across the insulating gap is generally a part of the total voltage V.

Let us assume first that the initial conductivity of the structure is negligible; i.e., R_0 approaches infinity. It is clear that the origination (field emission from a randomly appeared nanospike and the subsequent formation of the carboniferous conducting medium) of the first conducting element with increasing V does not depend on R_b and R_s values and is caused by the nonuniform emission from the insulating gap perimeter. Since, in real experiments, the emission is always nonuniform, single nanostructure *5* begins to grow. However, this is accompanied by an increase in the conductivity and in the current through the structure. As a result, the voltage across the resistances drops. Therefore, it is of interest to find the maximum conductivity at which other conducting elements do not originate as yet.

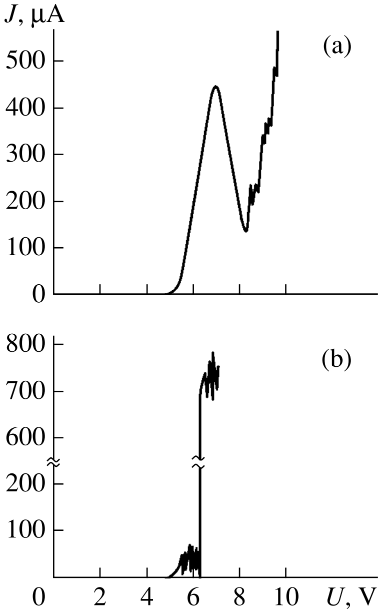


Fig. 3. Quasi-static current–voltage characteristics of open sandwichlike Si–SiO_2–W structures during electroforming for the first stage of voltage increase without the ballast resistance. The structures are covered by a resist layer. The exposed perimeter length is 64 μm. The rate of change of voltage is 0.05 V/s. The resistivity of the silicon is (a) 10 and (b) 0.1 Ω cm. The SiO_2 layer thickness is (a) 16 and (b) 17 nm.

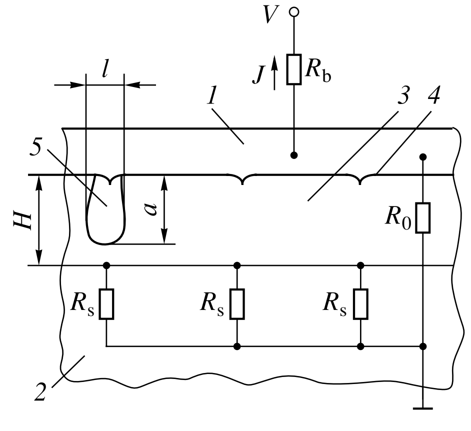


Fig. 4. Insulating gap of the metal–insulator–semiconductor structure and the equivalent electric circuit during the electroforming. *1*, metallic cathode; *2*, semiconductor anode; *3*, insulating gap; *4*, nanospikes on cathode surface; and *5*, conducting nanostructure (carboniferous conducting medium).

The appearance of the first conducting element is characterized by the growth voltage $V = U_{g1} = U_g$ corresponding to the emission initiated at the “weakest” site. As the nanostructure grows and the current increases, the voltage V is redistributed between the insulating gap and the resistances R_b and R_s. If we have

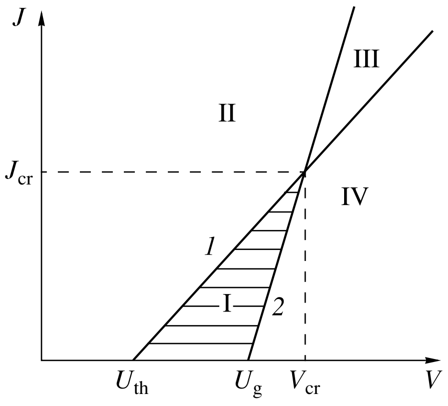


Fig. 5. Current–voltage diagram showing the region (dashed) of existence of the single conducting nanostructure in the case of the zero initial conductivity.

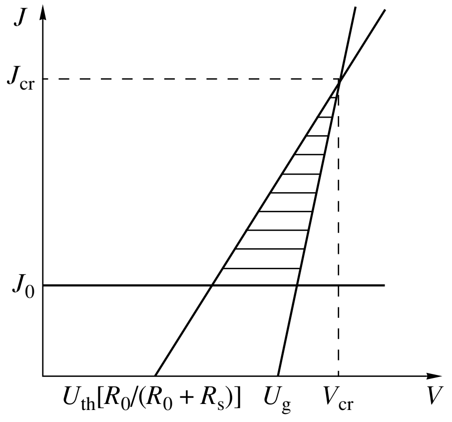


Fig. 6. The same as in Fig. 5 in the general case.

only the first nanostructure, the total current J flows through it, so that the voltage U_{ON} across the nanostructure is given by

$$U_{ON} = V - J(R_b + R_s). \tag{1}$$

Since the current through all other regions of the insulating gap is absent, the voltage U_{OFF} across them is

$$U_{OFF} = V - JR_b. \tag{2}$$

If $U_{OFF} < U_{gi}$, where $i = 2, 3, \ldots$, and U_{gi} are the growth voltages for the second and subsequent elements (the voltages form a monotonically increasing sequence), other conducting elements do not arise. Assuming that the conditions for the emission initiation are everywhere similar, we can take the stronger inequality

$$U_{OFF} < U_g, \tag{3}$$

i.e., assume that the process is characterized by one value $U_g = U_{g1}$, which is the least in the sequence.

The originated nanostructure can grow through the extension of the carboniferous conducting medium if

$$U_{ON} > U_{th}. \tag{4}$$

The diagram obtained from Eqs. (1)–(4) is shown in Fig. 5. The single conducting nanostructure can exist only in the dashed region. The coordinates of the critical point (point of intersection) are obtained from the solution of Eqs. (1) and (2):

$$V_{cr} = U_g + (U_g - U_{th})\frac{R_b}{R_s}, \tag{5}$$

$$J_{cr} = \frac{U_g - U_{th}}{R_s}. \tag{6}$$

Let for given R_b and R_s the maximum voltage V at which the single conducting nanostructure still exists be V_{cr}. The conductivity of the nanostructure can grow until the current exceeds J_{cr}. At the critical point, the resistance of the nanostructure is determined as the ratio U_{cr}/J_{cr} and is given by

$$r_{cr} = \frac{U_{th}}{U_g - U_{th}}R_s. \tag{7}$$

This resistance depends neither on the external voltage V nor on the external ballast resistance R_b. Since the conductivity of the nanostructure is defined (on assumption that the resistivity of the conducting medium is constant) by its geometric characteristics, such as $h = H - a$ and l (Fig. 4), condition (7) is the limitation imposed on the sizes a and l of the single conducting nanostructure. The slopes of lines *1* and *2* in Fig. 5 are $1/(R_b + R_s)$ and $1/R_b$, respectively. Hence, in the special case $R_s = 0$, corresponding to the MIM structure, the lines in the diagram become parallel, the critical point disappears, and the conductivity of the single nanostructure can reach any value.

In region II (Fig. 5), other conducting elements cannot arise but, in contrast to region I, the growth of the first nanostructure is also impossible. In region III, other conducting elements can arise and the growth of the first nanostructure is impossible. Finally, in region IV, both processes can take place.

When the structure has an appreciable initial conductivity, i.e., R_0 is finite, we must consider two additional effects. First, the growth voltage U_g decreases because the current distributed over the insulating gap initiates the formation of the carboniferous conducting medium. Second, R_0 inhibits other processes. The first effect facilitates the formation of all the conducting elements, while the second one hampers the appearance of even the first of them.

It should be emphasized that the growth voltage U_g is not a physical constant. It depends on the insulating gap width, cathode geometry on the nanometer scale, distribution of the initial current along the insulating gap, and rate of rise of the voltage. However, U_g can easily be found experimentally from the current jump when voltage gradually increases; therefore, it has a certain value for a specific structure.

Prior to the origination of the first conducting element, the total current flows through R_0. Since it arises when the voltage across the gap reaches the value U_g, the following inequality holds:

$$J > \frac{U_g}{R_0} = J_0. \tag{8}$$

Once the first conducting element has arisen, the following equations become valid:

$$\begin{aligned} V &= JR_b + U + J_1R_s, \\ U + J_1R_s &= J_2R_0, \quad J = J_1 + J_2. \end{aligned} \tag{9}$$

Here, J_1 is the current through the conducting nanostructure in the “on” state, J_2 is the current through the resistance R_0, and U is the voltage across the insulating gap.

A solution of system (9) in view of inequality (4), which is the condition for the growth of the nanostructure, yields the inequality

$$J < \frac{1}{R_b + \dfrac{R_sR_0}{R_s + R_0}}\left(V - \frac{R_b}{R_s + R_0}U_{th}\right). \tag{10}$$

Finally, the condition when none of the conducting elements but the first one can originate is derived from Eqs. (2) and (3) and has, as before, the form of the inequality

$$J > \frac{1}{R_b}V - \frac{1}{R_b}U_g. \tag{11}$$

For the growth of the single conducting nanostructure to proceed, inequalities (8), (10), and (11) must be fulfilled simultaneously (dashed region in Fig. 6). This region is similar to that in Fig. 5 except for the fact that the current J_0 is its lower boundary. The critical point, where straight lines (10) and (11) intersect, has the coordinates

$$U_{cr} = \frac{R_bR_s + R_bR_0 + R_sR_0}{R_sR_0}U_g - \frac{R_b}{R_s}U_{th}, \tag{12}$$

$$J_{cr} = \frac{R_s + R_0}{R_sR_0}U_g - \frac{1}{R_s}U_{th}. \tag{13}$$

As before, the critical current J_{cr} depends neither on V nor on R_b. Hence, if the initial conductivity is so large that $J_0 > J_{cr}$, the single nanostructure cannot form in principle for any values of the external parameters. The minimum resistance of the single conducting nanostructure, which corresponds to its maximum size, is generally given by the expression

$$r_{cr} = \frac{U_{th} + U_g\dfrac{R_s}{R_0}}{U_g - U_{th}}R_s.$$

This expression and Eqs. (12) and (13) turn to corresponding formulas for the case when the initial conductivity is absent if R_0 approaches infinity.

Our analytical results are in good agreement with the experimental data presented above. The single conducting nanostructure forms if the electroforming parameters remain within the dashed region in Fig. 6.

CONCLUSION

Our experimental results have shown that there are factors both favoring (external ballast resistance) and preventing (local spreading resistance and initial conductivity) the production of a single conducting element by electroforming open sandwichlike structures with an insulating gap several tens of nanometers in width. A simple model using the equivalent electric circuit allows us to find the trade-off between these factors and construct a current–voltage diagram that has a region within which the single conducting nanostructure can be reliably electroformed. The conductivity, as well as the sizes of the conducting element, are basically limited from above. The expression for the extreme value of the nanostructure resistance was derived.

ACKNOWLEDGMENTS

This work was supported by the State Scientific and Technical Subprogram “Promising Technologies and Devices for Micro- and Nanoelectronics” of the Ministry of Science of the Russian Federation.

REFERENCES

1. G. Dearnley, A. M. Stoneham, and D. V. Morgan, Rep. Prog. Phys. **33**, 1129 (1970).
2. H. Pagnia and N. Sotnik, Phys. Status Solidi A **108** (1), 11 (1988).
3. K. A. Valiev, V. L. Levin, and V. M. Mordvintsev, Zh. Tekh. Fiz. **67** (11), 39 (1997) [Tech. Phys. **42**, 1275 (1997)].
4. K. A. Valiev, S. E. Kudryavtsev, V. L. Levin, *et al.*, Mikroélektronika **26**, 3 (1997).
5. V. M. Mordvintsev, S. E. Kudryavtsev, and V. L. Levin, Zh. Tekh. Fiz. **68** (11), 85 (1998) [Tech. Phys. **43**, 1350 (1998)].
6. V. M. Mordvintsev and S. E. Kudryavtsev, Mikroélektronika (in press).
7. V. M. Mordvintsev and V. L. Levin, Zh. Tekh. Fiz. **64** (12), 88 (1994) [Tech. Phys. **39**, 1249 (1994)].
8. A. V. Eletskiĭ and B. M. Smirnov, Usp. Fiz. Nauk **147**, 459 (1985) [Sov. Phys. Usp. **28**, 956 (1985)].
9. J. P. Johnson, L. G. Christophorou, and J. G. Carter, J. Chem. Phys. **67**, 2196 (1977).
10. P. N. Luskinovich, V. D. Frolov, A. E. Shavykin, *et al.*, Pis’ma Zh. Éksp. Teor. Fiz. **62**, 868 (1995) [JETP Lett. **62**, 881 (1995)].
11. E. F. Golov, G. M. Mikhaĭlov, A. N. Red’kin, *et al.*, Mikroélektronika **27**, 97 (1998).
12. R. M. Penner, M. J. Heben, and N. S. Lewis, Appl. Phys. Lett. **58**, 1389 (1991).

Translated by M. Astrov

Technical Physics, Vol. 47, No. 4, 2002, pp. 432–437. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 60–65.
Original Russian Text Copyright © 2002 by Karachinov.

SOLID-STATE ELECTRONICS

Negative Crystals of Silicon Carbide

V. A. Karachinov

Ya. Mudryĭ State University, Sankt-Peterburgskaya ul. 41, Novgorod, 173003 Russia

Received May 10, 2001

Abstract—Systems of negative silicon carbide crystals are classified and studied by experimental methods. The crystal structure and morphology forming during growth, etching, and erosion are discussed. *© 2002 MAIK “Nauka/Interperiodica”.*

INTRODUCTION

The problem of implementing quantum-size effects resulting in nanostructures with a single electronic system, as well as in nanocomposites, stimulates the search for new materials and methods of their fabrication [1]. New materials are frequently prepared by producing a regular or quasi-regular set of voids. In a semiconductor crystal, such voids can be pores made by etching; in composites, natural spaces between the balls [2]. Among wide-gap semiconductors, silicon carbide is of particular interest for nanoelectronics. Its properties have already found application in microwave electronics, optoelectronics, and power electronics [3]. However, a set of voids in SiC crystals may be used in areas untraditional for this material, such as ultrasonics, micromechanics, chemistry, thermal electronics, and even jewellery [3–5]. Generally, the problem here is to controllably produce a system of negative crystals in SiC and to study their properties [6, 7].

This work is devoted to the classification and experimental investigation of such systems.

EXPERIMENTAL TECHNIQUES

The objects of experiment were 6H-SiC single crystals grown in a vacuum or argon at growth temperatures T_g = 2100–3000 K with the conventional Lely and LÉTI (the abbreviation from Leningradskiĭ ÉlektroTekhnicheskiĭ Institut) techniques [8]. The concentration of uncompensated donors was $N_d – N_a = 5 \times 10^{17}$–$5 \times 10^{18}$ cm^{-3}; the dislocation density, $N_D = 1 \times 10^4$–1×10^7 cm^{-2}. The electrical erosion processing of the SiC crystal was carried out with the well-known sewing method and by filamentary electrode cutting [4, 9] on both commercial and laboratory equipment. The structure and the morphology of the crystals were examined by the Lang X-ray diffraction method and anisotropic chemical etching in a KOH solution at T = 450–600°C, as well as in a metallurgical microscope and a BS-340 scanning electron microscope (SEM) operating in the secondary electron mode.

RESULTS AND DISCUSSION

By negative crystal (NC), one means a faceted cavity inside a positive crystal that is usually filled with a mother medium (solution or gas). NCs are viewed as three-dimensional crystal defects [6, 10]. It is necessary to distinguish opened NCs, those communicating with a crystallization medium (etch pits or striation), and closed NCs (inclusion of the mother medium, foreign inclusions, etc.). NCs with an artificially created surface (faceted or unfaceted) will be considered as pseudo-NCs. To date, a large number of NCs has been identified in SiC crystals. They can be classified by their specific features.

1. NC Classification

Experimental data show that various SiC crystallization conditions, as well as various treatments of SiC crystals, such as etching, mechanical processing, and erosion, may lead to the formation of different NCs and even NC systems (see table).

It is known that there is a relation between SiC crystallization and decrystallization; namely, nucleation mechanisms, as well as growth kinetics and mechanisms, have common features [7]. Under certain conditions, this may result in the simultaneous growth of positive and negative crystals. Together with individual NCs, NC systems may form, where NCs are distributed randomly or along particular crystallographic directions. Lely- and LÉTI-grown SiC positive crystals may contain NCs distributed along the directions perpendicular to the (0001) basis planes or along the oblique directions, NCs of variable (arbitrary) orientation, and those growing parallel to the (0001) basis planes. Typical examples of crystallographically oriented NCs are shown in Fig. 1. In spite of high SiC crystallization temperatures, NCs grew by the solid–liquid–vapor mechanism [10]. Silicon or impurities (metallic solvents), which may fall into the reaction zone from the graphite fixtures or be specially introduced into the vapor source [16], serve as the liquid phase. SiC positive crystals obtained by the Lely method grow in an unbounded space, as a result of which they take the form of thin

Systems of negative crystals in 6H-SiC

	NC production methods	NC shape	NC system	Crystallographic orientation, localization
1	SiC sublimation growth [8, 11]	Holes (pits)	Irregular	Growth faces, arbitrary
		Disks		
		Unfaceted cavities (pores)	Quasi-regular	[0001]
		Prisms		
		Pyramids	Irregular	Arbitrary
		Pseudo-NC of ordinary shape	Regular	[0001]
2	SiC etching [11–13]: chemical, electrochemical	Hexagonal pits	Quasi-regular	(0001)
		Circular flat-bottomed pits	Irregular	
	thermal, by molten metals or silicon, through mask (stencil)	Filamentary (whiskers)	Irregular	
		Unfaceted cavities (pores)	Quasi-regular	
		Pseudo-NC (pits)	Regular	
3	SiC erosion [9, 14, 15]: electrical, beam, mechanical, ultrasonic	Pits	Quasi-regular	(0001), (1010), $(11\bar{2}0)$
		Filamentary	″	
		Pseudo-NC of regular and irregular shape	Regular	Any

platelets. In such crystals, NCs are usually oriented along the direction of maximal growth rate ($[11\bar{2}0]$). Fig. 1a is illustrative in the sense that it demonstrates the growth of NCs in the direction parallel to the (0001) basis plane through the development of a filamentary NC. Such a mechanism is consistent with the conventional model of negative crystal evolution [7].

The use of seeds and also the limited growth space in the LÉTI method (sublimation profiling [11]) provide the oriented growth of the positive and negative crystals. Much lower crystallization temperatures (T_g = 1700–2000°C for the growth in a vacuum) compared with the Lely method raise the liquid phase formation probability, particularly at the early stage of growth, hence, a high density of NCs extended in the [0001] direction. Such an NC system may also form at crystallization temperatures above 2100°C, where the density of "upright" dislocations, along which voids (usually decorated by graphite) are produced, is high. High growth rates of the positive crystals and the depletion of the SiC vapor source intensify this process. Figure 1b shows the traces of solidified melt inside the NC.

Unfortunately, the fact that NCs forming in SiC during growth have different shapes and sizes are today treated as process errors [17]. Only regular systems of NCs in SiC are of practical value. However, they are difficult to produce during crystallization, since the growth of NCs has a random character and has been poorly studied. The exception is quasi-regular sets of etch pits and regular systems of pseudonegative crystals. Some of their features are discussed below.

2. Quasi-regular Systems of NCs

(i) Etch pits. The chemical etching of SiC crystals in salt melts and oxide melts, as well as in alkaline solutions, readily generate etch pits on the (0001)Si face. The sizes and shapes of the etch pits depend largely on growth crystal defects and etching conditions [12]. Earlier [11, 12], it has been shown that small hexagonal etch pits uniquely correspond to edge and screw dislocations; circular flat-bottomed pits, to dislocation loops due to vacancies; and large hexagonal pits, to growth NCs. Of most interest for us are hexagonal etch pits related to upright dislocations (UDs). This is because these dislocations readily form when SiC crystals grow in the [0001] direction. As follows from the X-ray diffraction data, the dislocation lines are straight (Fig. 2) and are aligned primarily with the C axis. In the crystals grown at temperatures below 2300°C, the UD density exceeds ~10^5 cm^{-2} and varies along the crystal insignificantly (Fig. 2b). UDs form a rather stable system; however, well-known methods, such as complex profiling, isovalent doping, and the use of polytype gates (layers), may provide the control of the UD inheritance process and form quasi-regular dislocation (hence, etch-pit) ensembles [11].

The cooling rate after the crystallization has a great effect on the UD distribution over the crystal cross section. It affects the temperature field and, consequently, thermoelastic stresses generated by standard dislocation distributions. Fig. 3a shows the experimentally found distribution of the etch pit density along the radius of the SiC crystal. At a cooling rate of ≈970 K/h

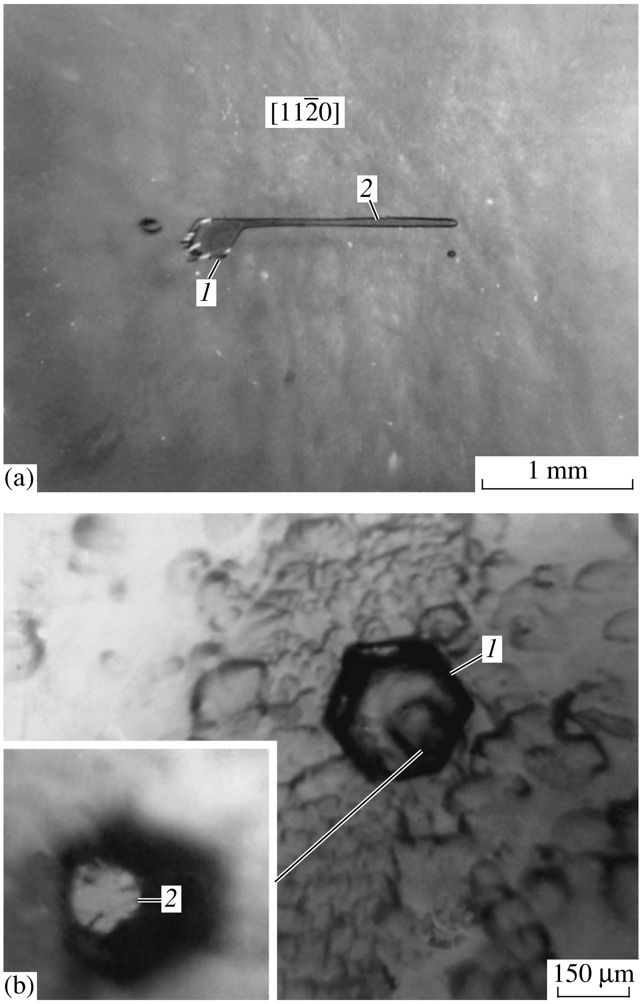


Fig. 1. Growth negative crystals. (a) Lely method, growth in the [0001] direction: (*1*) hexagonal prism and (*2*) filamentary NC; (b) LÉTI method, growth in the [0001] direction: (*1*) hexagonal prism and (*2*) traces of melt crystallized on the cavity bottom.

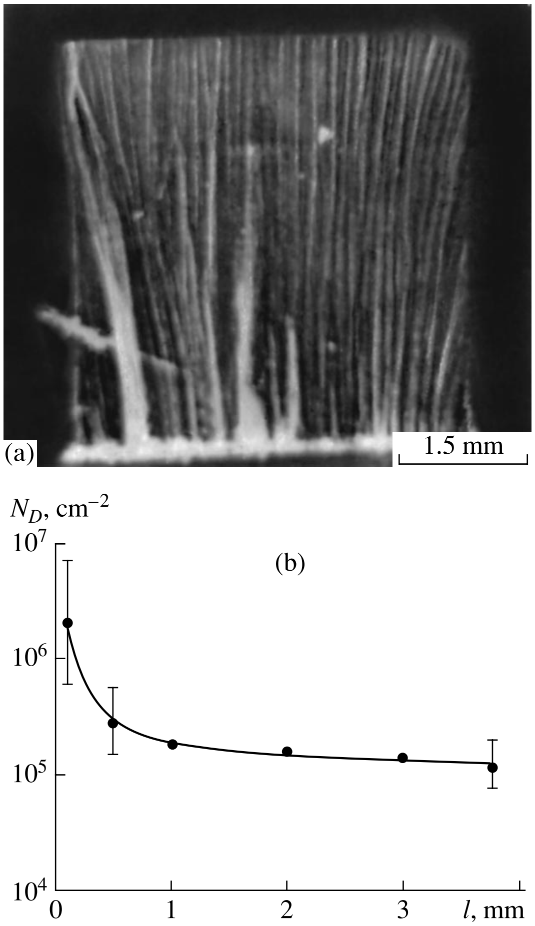


Fig. 2. Dislocation structure of the 6H-SiC crystals (LÉTI method, $T_g \cong 2100$°C). (a) UD pattern obtained by the Lang method from the (0012) plane, Mo$K_{\alpha 1}$ radiation; (b) variation of the etch pit (dislocation) density along the crystal.

or lower, one can observe the weak tendency to a *V*-shaped distribution. The etch pit distribution markedly changes starting from a cooling rate of ≈1650 K/h or higher. It becomes *W*-shaped with more or less distinct boundaries at the center and at the edges of the crystal (Fig. 3b). Since the cooling rate is finite, the temperature over the cross section has no time to equalize, which generates stresses depending also on the crystal size (Fig. 3c).

The SEM examination of the etch pit structure on the silicon faces indicates the following (Fig. 4). It is clearly seen that the pits have the form of a pyramid with a regular hexagonal base. Because of a difference between the normal and tangential components of the SiC dissolution rate at the sites where dislocations emerge on the surface, the terrace structure forms. No asymmetry of the etch pits and the base facet was revealed. In most of the pits observed in the crystals with the flat undistorted growth front, the pyramid tip was symmetric relative to the base. This observation is corroborated by the X-ray topographic data, which indicate that the UD lines in the crystals studied cross the (0001) growth face largely at right angles.

(ii) Breakdown figures. It is known that chemical etching under strong electric fields generates a set of pits that are inclined relative to the basic position. For quartz crystals, for example, this inclination may reach ≈10° [12]. It is also known that a discharge generated by a strong electric field can produce (without using chemicals) specific figures on the SiC crystal surface. The figures may appear as usual erosion holes or as a set of negative filaments known as EFT defects [14].

In the case of individual discharges, a quasi-regular set of EFT defects usually forms. It, however, may contain ordered groups of extended (coupled) defects (chains or mosaic). Spikes oriented in the $\langle 11\bar{2}0 \rangle$ directions are dominant coupling elements (bridges) here. The morphology of EFT defects was refined with the SEM (Fig. 5) and was found to be a combined result

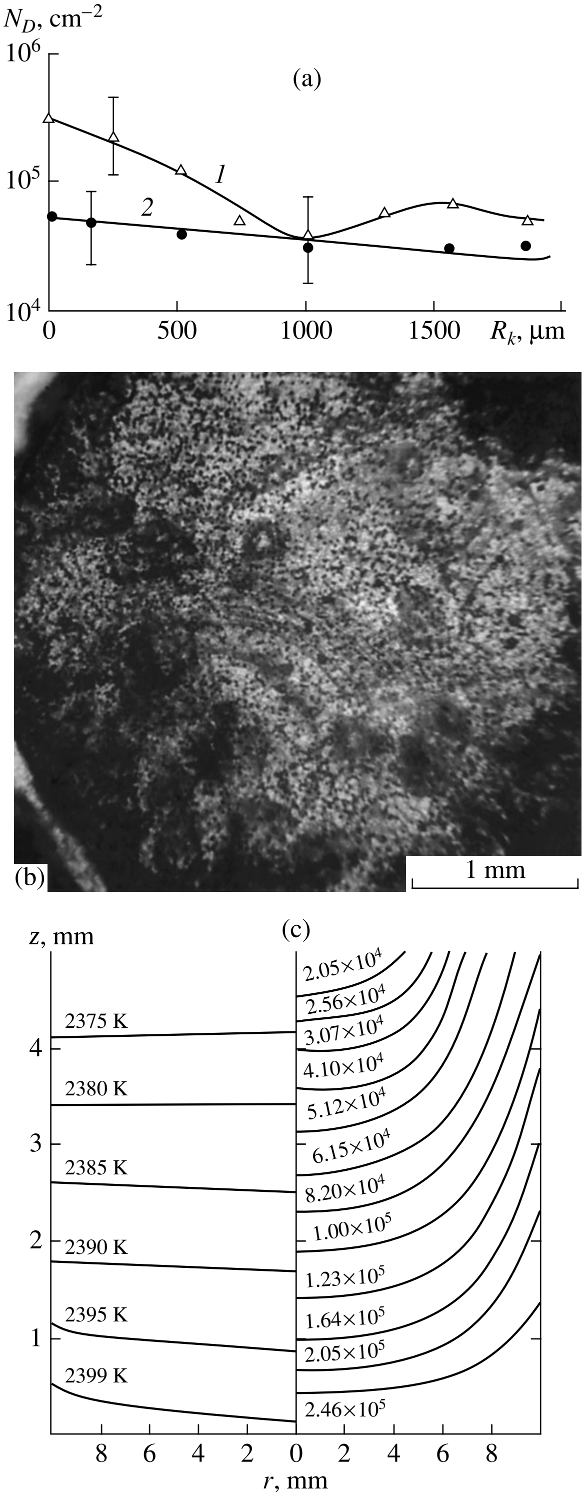


Fig. 3. Effect of the cooling rate on the dislocation structure. (a) Experimental radial distribution of the etch pit density, cooling rate is (*1*) ≈1650 and (*2*) ≈970 K/h; (b) natural (0001)Si face (etch pit pattern), cooling rate is ≈1650 K/h; and (c) calculated distributions of the temperature and thermoelastic stresses in the crystal (figures on the right are stresses in N/m^2).

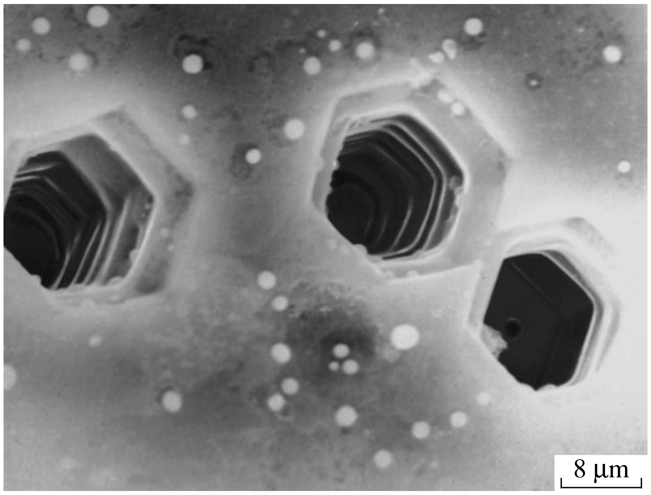


Fig. 4. SEM micrograph of the hexagonal etch pit structure on the (0001)Si face. LÉTI growth method.

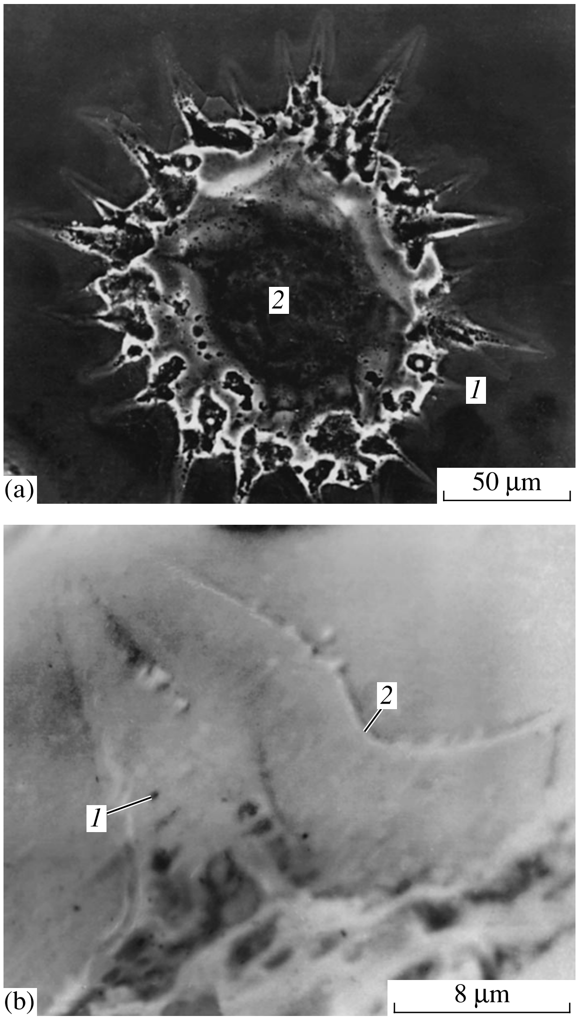


Fig. 5. SEM micrograph of the EFT defect structure on the (0001)Si face in SiC. (a) General view: (*1*) spikes and (*2*) hole; (b) growth step: (*1*) spike and (*2*) step.

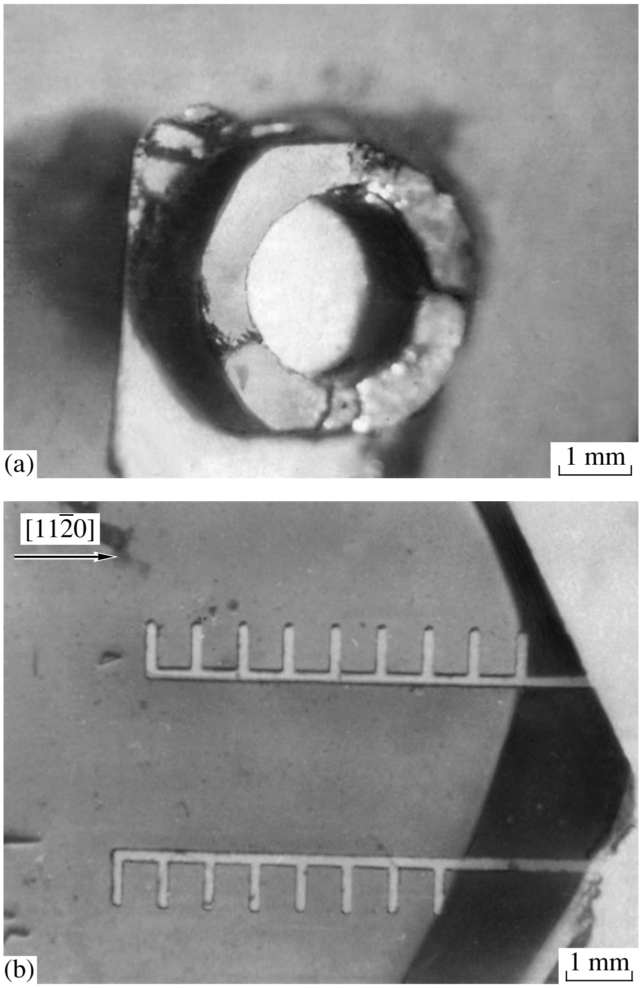


Fig. 6. (a) Pseudonegative crystals: (a) hollow crystal grown in argon at $T_g \cong 2100$°C and (b) crystal obtained by erosion in water (filamentary brass electrode of diameter $d \approx$ 100 μm).

of erosion and partial surface breakdown of the crystal. By its formation conditions, this defect can be classified as a negative breakdown figure [18]. It has the 6*mmm* point symmetry group. The effects of orientation and pinching of electron avalanches spreading over the SiC facet during the discharge and acting as extended heat sources produce a system of oriented cavities (spikes). Careful examination shows that the spikes oriented in the $[11\bar{2}0]$ and $[10\bar{1}0]$ directions differ in shape and size. Sharp-front spikes are oriented mostly in the former direction. This is consistent with the general idea that the crystallographic direction influences SiC evaporation channels with the participation of the silicon liquid phase [7]. Along with ordinary spikes, recuperative partition-divided spikes also appear. As a rule, recuperative spikes branch. High temperature gradients and the presence of the silicon-based liquid phase favor the epitaxial growth of SiC. Typical growth steps of height from 0.5 to 1.0 μm around the spikes are shown in Fig. 5b. The steps remained on the surface even after the long-term etching of the crystals in a boiling aqueous solution of KOH.

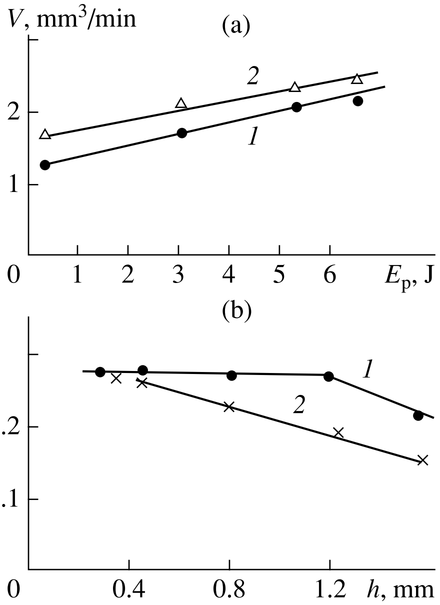


Fig. 7. Growth parameters of cylindrical pseudo-NCs obtained by the sewing method in transformer oil. (a) Experimental growth rate vs. electric pulse power: growth in the (*1*) [0001]Si direction and (*2*) [0001]C direction; (b) growth rate variation along the pseudo-NC: (*1*) tubular electrode and (*2*) cylindrical electrode (steel).

(iii) Regular sets of pseudo-NCs. The sewing principle, which is widely used in the size profiling of materials [19], may serve as a basis for methods that produce regular sets of voids—pseudonegative crystals—in both single-crystal and polycrystalline SiC. Figure 6a demonstrates a pseudo-NC in the form of a hollow crystal obtained by the self-sewing of a positive single crystal during its growth from the vapor.

Erosion methods, such as electrical erosion in liquid insulators, offer wide applicability and are easy to implement. Note that the early pseudo-NC obtained in SiC was threaded. Figure 6b shows a regular NC system; Fig. 7, typical growth parameters for a cylindrical pseudo-NC. The sewing depth is traditionally a parameter of great interest. Experiments with bulk single crystals and layer structures (sets of platelets) demonstrated that the NC growth rate monotonically decreases as a rule. This is presumably explained by an increase in the evacuation length for SiC erosion products [19].

It is known that the basic geometrical characteristics of SiC pseudo-NCs are directly related to the shape of the profiling electrode. The profiling of a material having a high erosion resistance at an effective diameter $d \le 1$ mm is a challenge. Therefore, the use of a set of positive filamentary crystals seems to be promising for solving this problem. However, this statement calls for further experimental verification.

CONCLUSION

Special technological actions on SiC matrix crystals may form crystallographically oriented systems of negative crystals in them. The growth of NCs during the electrical erosion of SiC is accompanied by the epitaxial growth of a positive crystal layer. One can produce regular systems of pseudo-NCs and carry out the size profiling of SiC with any degree of complexity.

REFERENCES

1. V. N. Bogomolov, D. A. Kurdyukov, A. V. Prokof'ev, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **63**, 496 (1996) [JETP Lett. **63**, 520 (1996)].
2. M. E. Kompan and N. Yu. Shibanov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **29**, 1859 (1995) [Semiconductors **29**, 971 (1995)].
3. P. A. Ivanov and V. E. Chelnokov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **29**, 1921 (1995) [Semiconductors **29**, 1003 (1995)].
4. V. A. Karachinov, in *Proceedings of the 3rd International Conference "Crystals: Growth, Properties, Real Structure, Application"* (VNIISIMS, Aleksandrov, 1997), Vol. 2, p. 240.
5. Yu. V. Deryugin, *Single Crystals of Corundum in Jewellery Industry* (Mashinostroenie, Leningrad, 1984).
6. V. A. Mokievskiĭ, *Crystal Morphology* (Nedra, Leningrad, 1983).
7. E. I. Givargizov, *Growth of Filamentary and Scaly Crystals from Vapor* (Nauka, Moscow, 1977).
8. V. I. Levin, Yu. M. Tairov, M. G. Travadzhyan, and V. F. Tsvetkov, Izv. Akad. Nauk SSSR, Neorg. Mater. **14**, 1062 (1978).
9. V. A. Karachinov, Kristallografiya **43**, 1097 (1998) [Crystallogr. Rep. **43**, 1038 (1998)].
10. R. S. Wagner, J. Cryst. Growth **3**, 159 (1968).
11. V. A. Karachinov, Author's Abstrct of Candidate's Dissertation (Leningrad, 1985).
12. K. Sangwal, *Etching of Crystals. Theory, Experiment and Application* (Elsevier, Amsterdam, 1987; Mir, Moscow, 1990).
13. A. G. Ostroumov, M. I. Abaev, and M. I. Karklina, Izv. Akad. Nauk SSSR, Neorg. Mater. **15**, 1497 (1979).
14. V. A. Karachinov, in *Proceedings of the 3rd International Conference "Crystals: Growth, Properties, Real Structure, Application"* (VNIISIMS, Aleksandrov, 1997), Vol. 2, p. 154.
15. V. G. Ral'chenko, V. I. Konov, A. A. Smolin, *et al.*, in *Proceedings of the International Workshop "Silicon Carbide and Related Materials"* (Novg. Gos. Univ., Novgorod, 1997), p. 41.
16. V. A. Il'in, V. A. Karachinov, Yu. M. Tairov, and V. F. Tsvetkov, Pis'ma Zh. Tekh. Fiz. **11**, 749 (1985) [Sov. Tech. Phys. Lett. **11**, 312 (1985)].
17. B. A. Kirillov, Sb. Nauchn. Tr. S-Peterb. Gos. Élektrotekh. Univ., No. 488, 87 (1995).
18. *Progress in Dielectrics*, Ed. by J. B. Birks and J. H. Schulman (Wiley, New York, 1959; GÉI, Moscow, 1962).
19. N. K. Mitskevich and I. G. Nekrashevich, *Electroerosion Machining of Metals* (Nauka i Tekhnika, Minsk, 1988).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 4, 2002, pp. 438–443. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 66–71.
Original Russian Text Copyright © 2002 by Babichev, Kozlovskiy, Romanov, Sharan.

SOLID-STATE ELECTRONICS

A Silicon Stress-Sensitive Unijunction Transistor

G. G. Babichev, S. I. Kozlovskiy, V. A. Romanov, and N. N. Sharan

Institute of Semiconductor Physics, National Academy of Sciences of Ukraine,
Kiev, 03028 Ukraine

e-mail: mickle@semicond.kiev.ua

Received May 14, 2001

Abstract—The performance of a silicon stress-sensitive unijunction transistor is investigated. The transistor is classed as a stress-sensitive semiconductor lateral bipolar device with an *S*-type input (emitter) *I*–*V* characteristic. The optimal layout of the device and its basic parameters are determined. The device can serve as a basis for designing relaxation oscillators with the physically integrated function of mechanical stress-to-signal frequency conversion at the output. © *2002 MAIK “Nauka/Interperiodica”.*

A unijunction transistor (or double-base diode) is a semiconductor device with one *p*–*n* junction and two base contacts. Its static (*S*-shaped) *I*–*V* characteristic has a region with negative differential resistances. Because of this, unijunction transistors are used in pulse generators, sawtooth generators, threshold devices, and converters. Circuits based on unijunction transistors are simpler and more reliable than those built around diode-connected and bipolar transistors [1–3]. Therefore, they are promising for transducers of physical quantities. For example, transducers where the output signal frequency depends on magnetic field strength have been designed on the basis of magnetic-field-sensitive unijunction transistors (unijunction magnetotransistors) [2, 4].

In this work, we consider the design and basic characteristics of a unijunction strain-sensitive transistor, around which transducers of mechanical quantities, primarily pressure transducers with the frequency output, can be built. It is known [1, 2, 5] that frequency-output transducers can be readily interfaced with digital systems and offer a high noise immunity because frequency-modulated signals are weakly sensitive to interference and crosstalk.

Unijunction transistors can be made by conventional methods of integrated technology and be placed on the planar side of a profiled silicon membrane of *EE* type (having two rigid central areas) [6]. This membrane converts a load uniformly distributed over its surface to uniaxial elastic compressive–tensile strain in that its part where the transistor is placed. The layout, electrical circuit, and crystallographic orientation of the unijunction transistor are shown in Fig. 1.

In the absence of elastic strain, the transistor operates like a conventional unijunction transistor [1–3]. With a potential difference U_{bb} applied to the base contacts and in the absence of a voltage between the emitter and the base electrode B_1 ($U_e = 0$), a small inverse emitter current $I_e = I_s$ (I_s is the saturation current) passes through the *p*–*n* junction. This current is due to the voltage $U_{bb}\eta$, where $\eta = l_e/l_x$ is the voltage transfer coefficient, or the internal coefficient of voltage division, and l_e is the spacing between the emitter and the base electrode B_1. We assume that the geometric sizes of the emitter are small compared with the base region. As the negative potential difference U_e grows, the junction becomes directly biased and the emitter starts to inject nonequilibrium electrons, which drift to the base electrode B_1 in the sweeping field. Because of this, the resistance R_{eb1} of the first base decreases and the direct bias of the emitter junction increases further. This process is of an avalanche character and passes the transistor to the on state. The input (emitter) *I*–*V* characteristic of the unijunction transistor is schematically shown in Fig. 2. It has two points of practical value: the peak (U_p, I_p) and the valley (U_v, I_v) of the voltage, where the

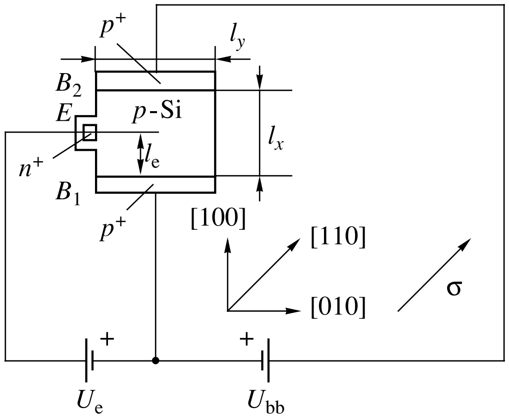


Fig. 1. Layout, connection diagram, and crystallographic orientation of the unijunction stress-sensitive transistor. B_1 and B_2, base electrodes; *E*, emitter.

derivative $\partial U_\mathrm{e}/\partial I_\mathrm{e}$ vanishes:

$$\frac{\partial U_\mathrm{e}}{\partial I_\mathrm{e}}(I_\mathrm{e} = I_\mathrm{p}) = 0 \quad \text{for} \quad \frac{\partial^2 U_\mathrm{e}}{\partial I_\mathrm{e}^2}(I_\mathrm{e} = I_\mathrm{p}) < 0, \tag{1}$$

$$\frac{\partial U_\mathrm{e}}{\partial I_\mathrm{e}}(I_\mathrm{e} = I_V) = 0 \quad \text{for} \quad \frac{\partial^2 U_\mathrm{e}}{\partial I_\mathrm{e}^2}(I_\mathrm{e} = I_V) > 0. \tag{2}$$

The parameters U_p, I_p and U_V, I_V are accordingly referred to as peak (switch-on) voltage and current and valley (saturation) voltage and current, respectively [1, 2].

In the case of uniaxial elastic strain in the crystallographic direction [110], the initially isotropic mobility μ_p of minority carriers (holes) in the base region of a strain-sensitive transistor becomes anisotropic. Therefore, the holes, drifting in the longitudinal sweeping field of the base (in the x direction; Fig. 1), shift to the transverse direction, and the transverse potential difference V_y appears. The strain-induced transverse potential difference modulates the injection from the emitter *p–n* junction and, hence, the voltage U_p of switching the unijunction stress-sensitive transistor to the high-conductivity state.

The operation of a unijunction stress-sensitive transistor is formally the same as the operation of a unijunction magnetotransistor, in which the Hall field modulates the emitter injection. This effect has been named injection modulation [7]. Note that the operation of early unijunction magnetotransistors was based on another, so-called deflection effect [2, 4]. Here, a magnetic field modulates the resistance of the first base, deflecting emitter-injected nonequilibrium carriers toward the lateral surfaces of the base, which have a low or high surface recombination rate.

A unijunction stress-sensitive transistor can be characterized through the absolute and relative strain sensitivities in terms of switching-on voltage (S_{AP} and S_{RP}) and saturation (valley) voltage (S_{AV} and S_{RV}):

$$S_{AP(AV)} = \left|\frac{\partial U_{p(v)}}{\partial \sigma}\right|_{\sigma \to 0}, \qquad S_{RP(RV)} = \frac{1}{U_{p(v)}}\left|\frac{\partial U_{p(v)}}{\partial \sigma}\right|_{\sigma \to 0}. \tag{3}$$

Here, σ is the mechanical stress in the base region of the stress-sensitive transistor. By writing $\sigma \longrightarrow 0$, it is meant that the above parameters are determined at small σ. In order to optimize the layout and the geometry of the unijunction stress-sensitive transistor, we will analyze the emitter characteristic $U_e(I_e)$ and evaluate the parameters U_p, I_p, U_V, and I_V. The analysis will be performed in the low uniform strain approximation [6, 8], i.e., when the anisotropy parameter is small ($|a| \ll 1$):

$$a(\sigma) = \Pi_{44}\sigma/2, \tag{4}$$

where Π_{44} is the shear piezoresistive coefficient for *p*-Si.

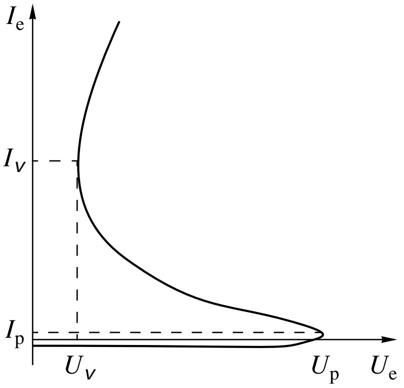


Fig. 2. Input *I–V* characteristic of the unijunction stress-sensitive transistor.

Let the base of the unijunction stress-sensitive transistor be bounded by the segments

$$0 < x < l_x, \quad 0 < y < l_y, \quad 0 < z < l_z. \tag{5}$$

We will consider the unijunction stress-sensitive transistor operating under two limiting conditions: low and high injection levels in the first base. In the former case, the emitter characteristic of the device is given by

$$U_\mathrm{e}(I_\mathrm{e}) = \frac{T}{e}\ln\left(1 + \frac{I_\mathrm{e}}{I_\mathrm{s}}\right) + I_\mathrm{e}R_\mathrm{eb1}(I_\mathrm{e}) + U_\mathrm{bb}[1 + R_\mathrm{eb2}(I_\mathrm{e})/R_\mathrm{eb1}(I_\mathrm{e})]^{-1} + V_y(\sigma)/2. \tag{6}$$

Here, T is temperature in energy units, e is the charge of an electron, and R_eb2 is the resistance of the second base (i.e., of the area between the emitter and the base electrode B_2). As follows from (6), the elastic strain modulates the emitter injection by the transverse potential difference $Y_y(\sigma)$. For the uniform strain, the transverse potential difference is expressed as [8]

$$V_y(\sigma) = a(\sigma)U_\mathrm{bb}F(l_\mathrm{e}, l_x, l_y), \tag{7}$$

where $F(l_\mathrm{e}, l_x, l_y)$ is a geometric factor and

$$F(l_\mathrm{e}, l_x, l_y) = \frac{4}{\pi^2}\sum_{n=1}^{\infty}\frac{[1-(-1)^n]}{n^2} \times \sin\left(\frac{\pi n l_\mathrm{e}}{l_x}\right)\tanh\left(\frac{\pi n l_y}{2l_x}\right). \tag{8}$$

In view of (7), the emitter characteristic of the device takes the form

$$U_\mathrm{e}(I_\mathrm{e}, \sigma) = \frac{T}{e}\ln\left(1 + \frac{I_\mathrm{e}}{I_\mathrm{s}}\right) + I_\mathrm{e}R_\mathrm{eb1}(\mathrm{I_e}) + U_\mathrm{bb}\left\{\frac{R_\mathrm{eb1}(I_\mathrm{e})}{[R_\mathrm{eb1}(I_\mathrm{e}) + R_\mathrm{eb2}(I_\mathrm{e})]} + \frac{a(\sigma)}{2}F(l_\mathrm{e}, l_x, l_y)\right\}. \tag{9}$$

Let us analyze expression (9). According to experimental data [1], the resistance of the second base R_{eb1} can be assumed to be independent of the emitter current in the first approximation:

$$R_{\mathrm{eb2}}(I_{\mathrm{e}}) \cong R_{\mathrm{eb2}}(I_{\mathrm{e}} = 0) = R_{\mathrm{eb2}}. \tag{10}$$

At the same time, the dependence $R_{\mathrm{eb1}}(I_{\mathrm{e}})$ is governed by a number of factors [1, 2], such as base geometry, boundary conditions at the contacts, nonequilibrium carrier distribution along the base segment $0 \le x \le l_{\mathrm{e}}$, etc. For the steady-state case, the set of equations for the distribution of nonequilibrium carrier concentration in the quasi-neutrality approximation has the form

$$\frac{1}{e}\mathrm{div}(\mathbf{j}_{\mathrm{p}}) + R - G_0 = 0, \tag{11}$$

$$\mathrm{div}(\mathbf{j}) \equiv \mathrm{div}(\mathbf{j}_n + \mathbf{j}_{\mathrm{p}}) = 0, \quad \mathrm{curl}(\mathbf{E}) = 0, \tag{12}$$

$$\mathbf{j}_{\mathrm{p}} = e\mu_{\mathrm{p}} p\mathbf{E} - eD_{\mathrm{p}}\nabla\mathbf{p}, \quad \mathbf{j}_n = e\mu_n n\mathbf{E} - eD_n\nabla\mathbf{n}, \tag{13}$$

$$n - n_0 = p - p_0 = \Delta p. \tag{14}$$

Here, n and p are the concentrations of electrons and holes, respectively; n_0 and p_0 are their equilibrium values; μ_n, D_n and μ_{p}, D_{p} are the mobilities and the diffusion coefficients of electrons and holes, respectively; $\mathbf{j}_n$ and $\mathbf{j}_{\mathrm{p}}$ are the densities of the electron and hole currents, respectively; R and G_0 are the rates of volume recombination and generation, respectively; and $\mathbf{E}$ is electric field. Eliminating $\mathbf{E}$ from Eqs. (13), we obtain the expression for the hole current density

$$\mathbf{j}_{\mathrm{p}} \cong \frac{\mathbf{j}}{b+1} - eD\nabla\mathbf{p}, \tag{15}$$

where $D = 2bD_{\mathrm{p}}/(b + 1)$ and $b = \mu_n/\mu_{\mathrm{p}}$.

We assume that the nonequilibrium carrier distribution in the base is quasi-one-dimensional and take into account the boundedness of the base in the yz plane by introducing the effective lifetime τ_{eff} and the associated diffusion length $L = \sqrt{D\tau_{\mathrm{eff}}}$ [10, 11]. According to [12], the effective value of L in our case (l_x, $l_y \gg l_z$) is limited by the base thickness: $L \approx l_z$. Substituting expression (15) for the hole current density into Eq. (11), we find, under the above assumptions, the equation for the hole concentration:

$$\frac{\partial^2 p}{\partial x^2} - \frac{p - p_0}{L^2} = 0. \tag{16}$$

The boundary conditions are taken in the form

$$\frac{\partial p}{\partial x}(x = l_{\mathrm{e}}) = \frac{I_{\mathrm{e}}}{eS_{\mathrm{e}}D(b+1)}, \tag{17}$$

$$p(x = 0) = p_0, \tag{18}$$

where S_{e} is the surface area of the emitter junction. Condition (17) follows from expression (14) for the hole current density and means the absence of the hole component at the boundary of the emitter p–n junction. The solution of problem (16)–(18) has the form

$$p(x) = p_0 + \frac{I_{\mathrm{e}} L \sinh(x/L)}{eD_n S_{\mathrm{e}}(b+1)\cosh(l_{\mathrm{e}}/L)}. \tag{19}$$

Since the injection is low, the emitter current in expression (19) is limited:

$$I_{\mathrm{e}} \ll eD_n S_{\mathrm{e}}(b+1)p_0/L\tanh(l_{\mathrm{e}}/L). \tag{20}$$

The resistance of the first base is found by integration:

$$R_{\mathrm{eb1}}(I_{\mathrm{e}}) = \frac{\rho_0}{S_{\mathrm{b}}}\int_0^{l_{\mathrm{e}}} \frac{dx}{[1 + (b+1)\Delta p(x, I_{\mathrm{e}})/p_0]}, \tag{21}$$

where ρ_0 is the base resistivity in the absence of injection and S_0 is cross-sectional area of the base. Integrating (21) yields

$$\begin{aligned} R_{\mathrm{eb1}}(I_{\mathrm{e}}) &= \frac{R_{\mathrm{eb1}}(I_{\mathrm{e}} = 0)L}{l_{\mathrm{e}}\sqrt{1 + P(I_{\mathrm{e}})^2}} \\ &\times \ln\left\{\frac{1 + [P(I_{\mathrm{e}}) + \sqrt{1 + P(I_{\mathrm{e}})^2}]\tanh(l_{\mathrm{e}}/2L)}{1 + [P(I_{\mathrm{e}}) - \sqrt{1 + P(I_{\mathrm{e}})^2}]\tanh(l_{\mathrm{e}}/2L)}\right\}, \end{aligned} \tag{22}$$

where $P(I_{\mathrm{e}}) = I_{\mathrm{e}}/I_0$ and $I_0 = eDS_{\mathrm{e}}p_0\cosh(l_{\mathrm{e}}/L)/L$.

To find the peak voltage U_{p} by formula (9), one must determine, along with the dependence $R_{\mathrm{eb1}}(I_{\mathrm{e}})$ and the current I_{p}, the geometric factor $F(l_{\mathrm{e}}, l_x, l_y)$, which is shown in Fig. 3 as a function of the base size and the emitter position. Its peak value is observed at $l_x \cong l_y$ and $l_{\mathrm{e}} = l_x/2$. These relationships for l_x, l_y, and l_{e} correspond to the case of uniform (constant) strain within the transistor base. The transverse potential difference $V_y(\sigma)$ and the geometric factor for the nonuniform case are given in [8, 13].

To simplify calculations, we will take advantage of the following experimental observation [1]: in a wide range of I_{e}, the dependence of the resistance of the first base on the emitter current can be approximated by the expression [2, 4]

$$R_{\mathrm{eb1}}(I_{\mathrm{e}}) = \frac{R_{\mathrm{eb1}}(I_{\mathrm{e}} = 0)}{1 + \xi I_{\mathrm{e}}}, \tag{23}$$

where ξ is an approximating factor.

The R_{eb1} vs. emitter current dependence calculated by formulas (22) and (23) is shown in Fig. 4. The curve is seen to be well approximated by expression (23) with $\xi = 1.11 \times 10^4\ \mathrm{A}^{-1}$. With formula (22), the curve $R_{\mathrm{eb1}}(I_{\mathrm{e}})$ was constructed for the silicon transistor base region measuring $l_x \times l_y \times l_z = 100 \times 100 \times 10\ \mu\mathrm{m}$ and silicon electrophysical parameters (at 300 K) $D_n = 35\ \mathrm{cm}^2\ \mathrm{s}^{-1}$, $D_{\mathrm{p}} = 13\ \mathrm{cm}^2\ \mathrm{s}^{-1}$, $p_0 = 5 \times 10^{15}\ \mathrm{cm}^{-3}$, $n_i = 1.4 \times 10^{10}\ \mathrm{cm}^{-3}$.

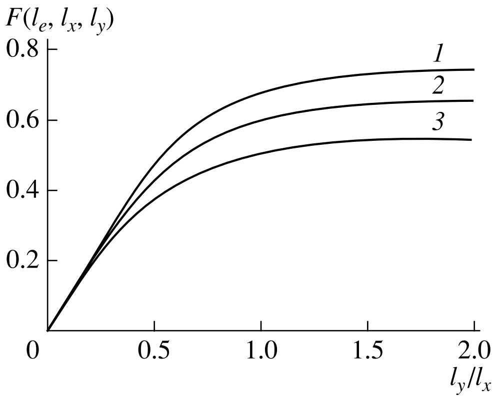


Fig. 3. Geometrical factor vs. transistor base size. The emitter position l_e/l_x = (*1*) 0.5, (*2*) 0.3, and (*3*) 0.8.

An expression for the current peak is found by substituting expression (23) into expression (9) for the input characteristic and equating the first derivative of the emitter characteristic ((expression (1)) to zero. At low injection levels ($I_e\xi \ll 1$), the expression for the current peak I_p is

$$I_p \cong \frac{T}{2[U_{bb}\eta(1-\eta)\xi - R_{eb1}(I_e = 0)]}. \tag{24}$$

Given I_p, the peak voltage can be easily found from (9). Figure 5 shows the analytic dependence of the peak voltage on the anisotropy parameter a. The curve $U_p(a)$ is linear, which is of importance for applications.

According to (3) and (9), the absolute strain sensitivity in terms of switching-on voltage, S_{AP}, does not depend on the emitter current:

$$S_{AP} = \frac{\Pi_{44}}{4}U_{bb}F(l_e, l_x, l_y), \tag{25}$$

with $\Pi_{44} = 1.4 \times 10^{-9}$ Pa^{-1} and $U_{bb} = 4$ V, we have $S_{AP} \cong 10^{-9}$ V/Pa.

The relative strain sensitivity in terms of switching-on voltage S_{RP} vs. potential difference U_{bb} is shown in Fig. 6. As U_{bb} grows, S_{RP} tends to saturation (curve *1*), while S_{AP} linearly increases (curve *2*).

Because of the shortening effect under high injection levels ($I_e\xi \gg 1$), the transverse potential difference V_y in (6) can be neglected, $V_y(\sigma) \approx 0$, so that the emitter characteristics of the stress-sensitive transistor takes the form

$$U_e(I_e) = \frac{T}{e}\ln\left(1 + \frac{I_e}{I_s}\right) + (I_e + I_{bb})R_{eb1}(I_t), \tag{26}$$

where $I_{bb} = U_{bb}/(R_{eb2} + R_{eb1}) \cong U_{bb}/R_{eb2}$ is the interbase current.

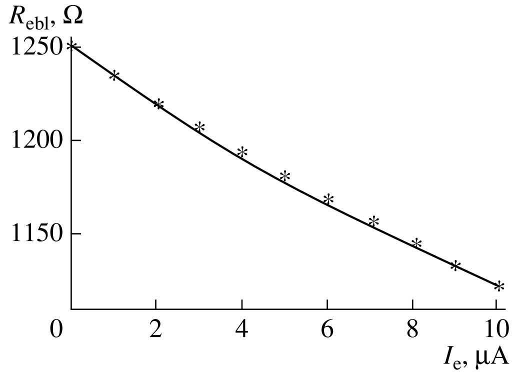


Fig. 4. Resistance of the first base vs. emitter current. Continuous curve, calculation by formula (22); symbols, approximation by formula (23).

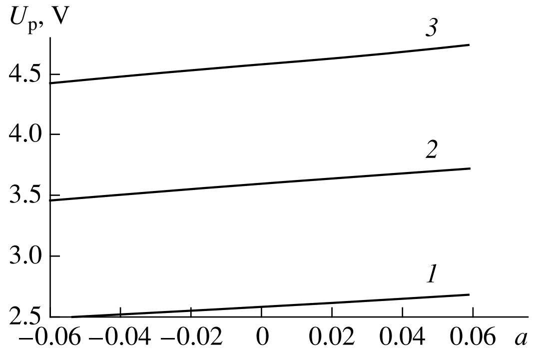


Fig. 5. Peak voltage vs. anisotropy parameter. U_{bb} = (*1*) 4, (*2*) 6, and (*3*) 8 V.

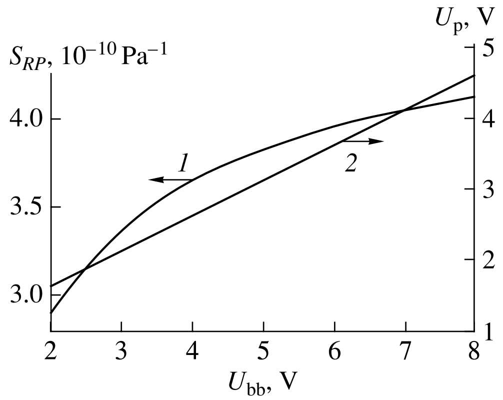


Fig. 6. (*1*) Relative strain sensitivity in terms of switching-on voltage and (*2*) switching-on (peak) voltage vs. potential difference U_{bb}.

Substituting approximating dependence (23) into formulas (26) and (2), we obtain the valley current

$$I_v \cong \frac{2[U_{bb}\xi\eta - R_{eb1}(I_e = 0)(1-\eta)]}{T\xi^2(1-\eta)}. \tag{27}$$

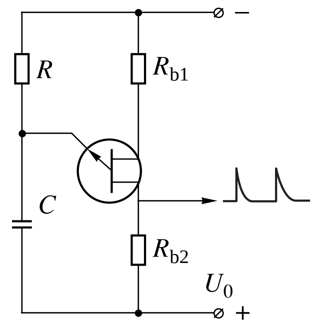


Fig. 7. Simple multivibrator based on the unijunction stress-sensitive transistor.

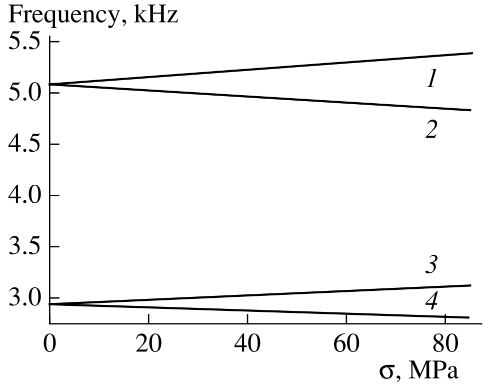


Fig. 8. Oscillation frequency of the multivibrator in Fig. 7 vs. elastic uniaxial mechanic (*1*, *3*) compressive and (*2*, *4*) tensile stress in the base of the stress-sensitive transistor. U_{bb} = (*1*, *2*) 6 and (*3*, *4*) 4 V. $RC = 10^{-3}$ s, U_0 = 10 V.

The valley voltage U_v is easy to find by substituting I_v into expression (26) for the emitter characteristic.

When the injection level in the stress-sensitive device is high, the uniaxial elastic strain modulates the resistance R_{eb1} of the first base, as with unijunction magnetotransistors [2, 4], by virtue of the deflection effect. Because of the strain-induced anisotropy of the hole mobility, nonequilibrium holes are deflected toward either the nearest or farthest (relative to the emitter) lateral side of the base (Fig. 1). As a consequence of the deflection effect, the valley voltage and current may depend on the elastic strain. The necessary condition for such a dependence is an excess of the transverse field E_y over the diffusion one [9, 10, 14]:

$$E_y \cong aE_x > \frac{T}{eL}, \tag{28}$$

where E_x is the sweeping field.

According to inequality (28), anisotropy parameter values at which the drift field exceeds the diffusion one obey the inequality

$$a > \frac{l_e T}{eLU_v} \cong 0.1. \tag{29}$$

Since the maximal value of the anisotropy parameter is somewhat smaller than $a = 0.06$ in our case (which corresponds to an elastic stress of 86 MPa), the effect of mechanical stress on U_v and I_v will be negligible.

Let us briefly consider the principal electrical circuit of mechanical quantity (pressure, force, and acceleration) transducers based on unijunction stress-sensitive transistors. Like related devices built around conventional unijunction transistors [1–3, 15], the strain-sensitive transducers also admit any of standard operating conditions: amplification, generation of harmonic and relaxation oscillations, switching, detection, etc. Unijunction stress-sensitive transistors are best suited to be used in switches and relaxation oscillators.

Figure 7 shows the circuit of the simplest multivibrator based on a unijunction stress-sensitive transistor. The oscillation frequency of the multivibrator depends on the elastic mechanical stress in the base of the transistor. In the steady-state operating mode, the oscillation period of the multivibrator is given by [1–4]

$$T(\sigma) = RC\ln\left[\frac{U_v - U_0}{U_p(\sigma) - U_0}\right]. \tag{30}$$

The output signal frequency of the multivibrator vs. elastic mechanical stress in the base of the unijunction stress-sensitive transistor is depicted in Fig. 8.

Thus, we have obtained the expressions for peak (switching) and valley currents (voltages) in the input *I–V* characteristic of the silicon unijunction stress-sensitive transistor. Their dependence on the mechanical stress in the base, potential difference between the base electrodes, and base geometry have been evaluated.

REFERENCES

1. I. G. Nedoluzhko and E. F. Sergienko, *Unijunction Transistors* (Énergiya, Moscow, 1974).
2. I. M. Vikulin and V. I. Stafeev, *Physics of Semiconductor Devices* (Radio i Svyaz’, Moscow, 1990).
3. S. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1981; Mir, Moscow, 1984), Vol. 1.
4. I. M. Vikulin, L. F. Vikulina, and V. I. Stafeev, *Galvanomagnetic Devices* (Radio i Svyaz’, Moscow, 1983).
5. S. V. Gumenyuk and B. I. Podlepetskiĭ, Zarubezhn. Élektron. Tekh. **12** (343), 3 (1989).
6. G. G. Babichev, I. P. Zhad’ko, S. I. Kozlovskiy, *et al.*, Zh. Tekh. Fiz. **64** (9), 84 (1994) [Tech. Phys. **39**, 908 (1994)].
7. H. P. Baltes and R. S. Popovic, Proc. IEEE **74**, 1107 (1986).

8. I. I. Boĭko, I. P. Zhad'ko, S. I. Kozlovskiy, and V. A. Romanov, Optoélektron. Poluprovodn. Tekh. **27**, 94 (1994).
9. I. I. Boĭko and V. A. Romanov, Fiz. Tekh. Poluprovodn. (Leningrad) **11**, 817 (1977) [Sov. Phys. Semicond. **11**, 481 (1977)].
10. Z. S. Gribnikov, G. I. Lomova, and V. A. Romanov, Phys. Status Solidi **28**, 815 (1968).
11. M. A. Lampert and P. Mark, *Current Injection in Solids* (Academic, New York, 1970; Mir, Moscow, 1973).
12. G. G. Babichev, V. N. Guz', I. P. Zhad'ko, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **26**, 1244 (1992) [Sov. Phys. Semicond. **26**, 694 (1992)].
13. I. P. Zhadko, G. G. Babichev, S. I. Kozlovskiy, *et al.*, Sens. Actuators **90**, 89 (2001).
14. Z. S. Gribnikov and R. N. Litovskiĭ, Fiz. Tekh. Poluprovodn. (Leningrad) **14**, 675 (1980) [Sov. Phys. Semicond. **14**, 397 (1980)].
15. S. A. Garyainov and I. D. Abezgauz, *Negative Resistance Semiconductor Devices* (Énergiya, Moscow, 1967).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 4, 2002, pp. 444–447. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 72–75.
Original Russian Text Copyright © 2002 by Ryabukho, Ul'yanov, Minenkova.

OPTICS, QUANTUM ELECTRONICS

Formation of a Quasi-Regular Interference Pattern in a Speckle-Modulated Laser Beam Reflected from a Transparent Optically Nonuniform Layer

V. P. Ryabukho, S. S. Ul'yanov, and I. F. Minenkova
Saratov State University, Moskovskaya ul. 155, Saratov, 410026 Russia
Institute of Problems of Precision Mechanics and Control, Russian Academy of Sciences,
Rabochaya ul. 24, Saratov, 410028 Russia
e-mail: rvp@sgu.ssu.runnet.ru

Received June 18, 2001

Abstract—Conditions for the formation of a quasi-regular interference pattern in the speckle-modulated radiation reflected from a rough transparent layer irradiated by a focused laser beam are found. A laser method for determining the optical thickness of the transparent coating deposited onto the rough surface is proposed. *© 2002 MAIK "Nauka/Interperiodica".*

Popov [1] and Kiedron [2] considered an optical method for measuring the thickness of relatively thin transparent plates and layers with optically smooth surfaces. The method is based on the interference of optical waves reflected by both sides of the object at various angles of incidence of a laser beam. The spacing of isoclinic fringes in this case depends on the optical thickness of the plate or the layer reflecting the beam. Priezzhev *et al.* [3] discussed the use of a similar method employing a focused laser beam for determining the thickness of the cornea. In the above papers, the layers studied had optically smooth surfaces, which do not introduce irregular phase perturbations into the waves reflected. If at least one surface of the layer is rough, a speckle structure, disturbing the regular interference pattern, forms in the laser beam reflected [4]. The purpose of this work is to consider conditions for the formation of a quasi-regular interference pattern in the speckle-modulated field reflected from an optically nonuniform transparent layer. Also, we will discuss a technique for measuring the local thickness of a rough transparent layer that is based on the precise focusing of the incident laser beam on the scattering surface of the layer.

Let us consider a focused laser beam striking a transparent layer with the rough bottom surface (Fig. 1). The field reflected represents the superposition of the laser beam with the uniform wave front and the speckle-modulated scattered beam. For the diffraction field away from the focal spot on the layer surface, the angular distribution of the intensity in the resulting reflected field is given by

$$I(\bar{\alpha}, z) = I_1(\bar{\alpha}, z) + I_2(\bar{\alpha}, z) + 2\sqrt{I_1(\bar{\alpha}, z) I_2(\bar{\alpha}, z)} \times \cos[\varphi(\bar{\alpha}) + \Delta\psi(\bar{\alpha})], \qquad (1)$$

where $I_1(\bar{\alpha}, z)$ and $I_2(\bar{\alpha}, z)$ are the angular distributions of the intensities in the beams reflected from the smooth (top) and rough (bottom) surfaces of the layer, respectively; $\bar{\alpha}(\alpha, \beta)$ is the direction of reflection; $\varphi(\bar{\alpha})$ is a random spatial–angular distribution of the phase of the speckle-modulated field; and $\Delta\psi(\bar{\alpha})$ is the determinate phase difference introduced by a transparent layer with the mean thickness $\bar{h}$ and the refractive index n.

In the plane of incidence (α, $\beta = 0$), the determinate phase difference is written as [5]

$$\Delta\psi(\alpha) = \frac{4\pi\bar{h}}{\lambda}\sqrt{n^2 - \sin^2\alpha}. \qquad (2)$$

It follows from Eq. (1) that the resulting interference

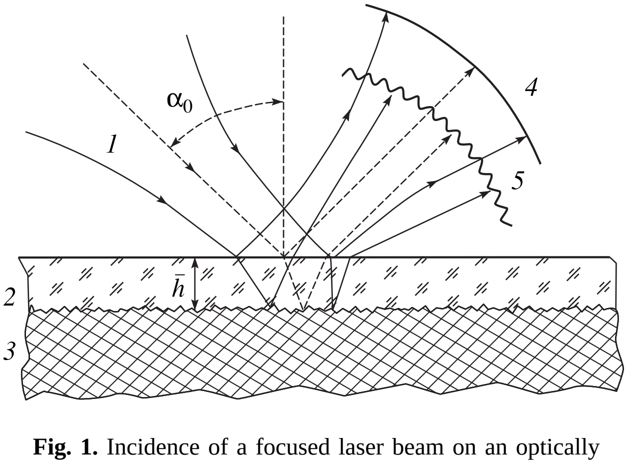


Fig. 1. Incidence of a focused laser beam on an optically nonuniform transparent layer and the reflection of uniform and speckle-modulated beams. *1,* Incident Gaussian laser beam; *2,* transparent layer with the rough bottom surface; *3,* substrate; *4,* reflected beam with the uniform wave front; and *5,* speckle-modulated reflected beam.

1063-7842/02/4704-0444$22.00 © 2002 MAIK "Nauka/Interperiodica"

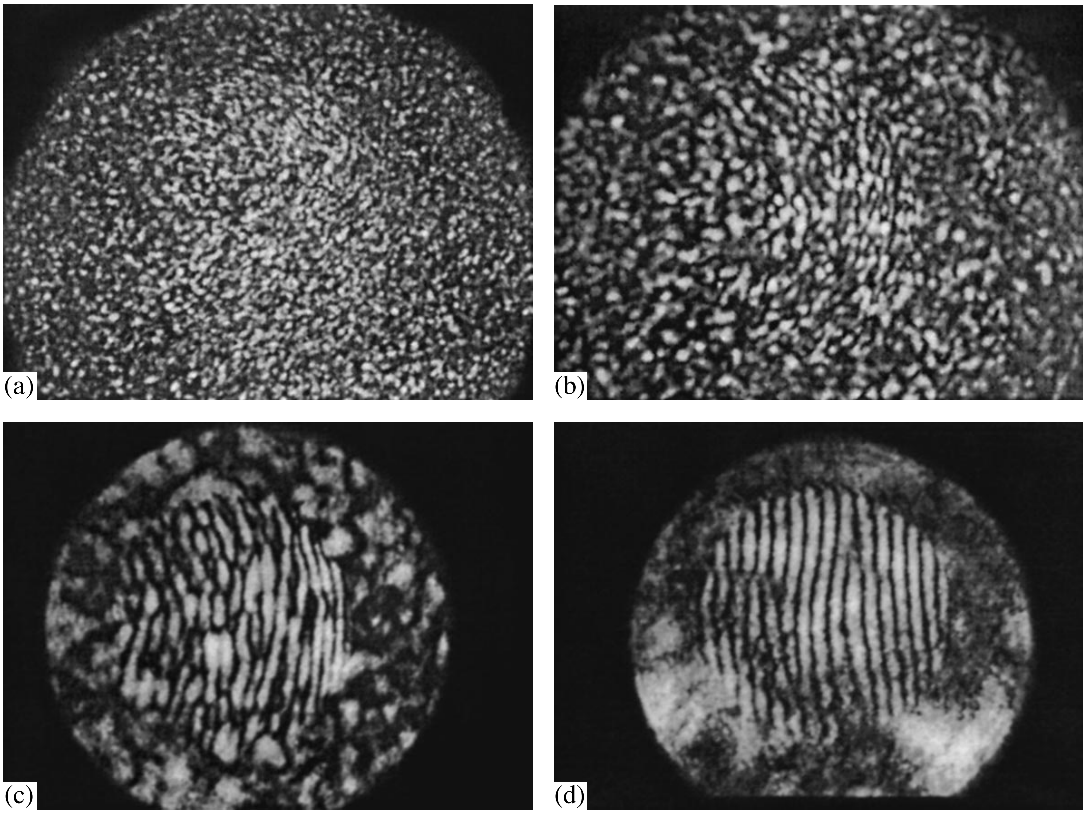


Fig. 2. Development of quasi-regular interference fringes against the background of the laser speckle structure as the beam waist approaches the rough surface of a thin transparent layer with an optical thickness $n\bar{h} \approx 1.6 \times 30$ µm. (a, b) transverse size ε_a of the speckles is smaller than the spacing Λ_a of the interference fringes, (c) $\varepsilon_\alpha \approx \Lambda_\alpha$, and (d) $\varepsilon_\alpha > \Lambda_\alpha$.

pattern in the beam reflected is qualitatively specified by the relationship between the phase distribution parameters $\varphi(\bar{\alpha})$ and $\Delta\psi(\bar{\alpha})$. The determinate component of the phase difference $\Delta\psi(\bar{\alpha})$ shows up in the spatial distribution of the intensity of the resulting speckle-modulated field as regular interference fringes if the spatial (angular) rate of change of the phase difference $\partial\Delta\psi(\alpha)/\partial\alpha$ is substantially greater than that of the random phase difference $\partial\varphi(\alpha)/\partial\alpha$ in the speckle-modulated field. The fastest variation of $\Delta\psi(\bar{\alpha})$ takes place in the plane of incidence. For this plane, the formal condition for regular fringe appearance is written as

$$\varepsilon_\alpha > 2\pi[\partial(\Delta\psi(\bar{\alpha}))/\partial\alpha]^{-1} = \Lambda_\alpha(\alpha), \tag{3}$$

where ε_α is the mean angular size of speckles in the field reflected from the rough surface and $\Lambda_\alpha(\alpha)$ is the angular spacing of regular interference fringes expected.

In other words, it is necessary that the transverse size of the speckles in the light scattered by the rough surface be larger than the spacing of the interference fringes expected. For a Gaussian laser beam, the transverse size of speckles in the plane of incidence can be estimated by the expression based on the Van Cittert–Zernicke theorem for the speckle field [6, 7]:

$$\varepsilon_\alpha \cong (\lambda\cos\alpha_0)/(\pi w\cos\alpha), \tag{4}$$

where α_0 is the angle of incidence of the axial beam (Fig. 1) and w is the radius of the Gaussian beam on the scattering surface of the layer. The coefficient $\cos\alpha_0/\cos\alpha$ takes into account the variation in the speckle size away from the optical axis of the reflected laser beam ($\alpha \neq \alpha_0$).

The speckle size is increased and, consequently, condition (3) is satisfied if the laser beam waist lies on the rough surface of the layer. In this case, w in Eq. (4) is minimized. Figure 2 illustrates the effect of the determinate component of the phase difference $\Delta\psi(\bar{\alpha})$ on the interference pattern in the reflected laser beam. Quasi-regular interference fringes develop against the background of the speckle structure as the waist of the incident beam (with angle of incidence of about 45°) approaches the scattering surface of the layer. Figure 2d shows the interferogram corresponding to the minimum size of the laser spot on the rough surface and, accordingly, to the maximum size of the speckles in the beam reflected.

In this case, the angular spacing Λ_α of the almost parallel interference fringes is given by an expression

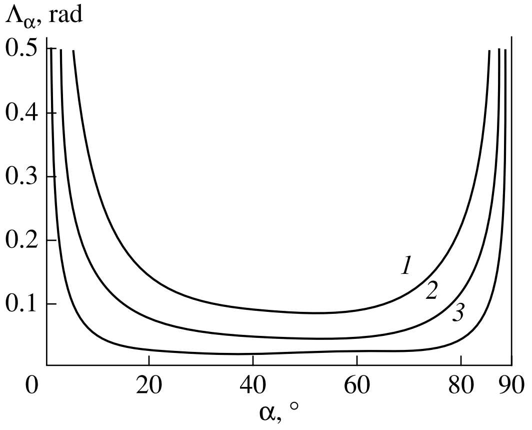


Fig. 3. Angular spacing Λ_a of interference fringes versus the angle of reflection α of laser radiation ($\lambda \approx 0.63$ μm) from the surface of the transparent layer with the refractive index $n = 1.5$ and the thickness h = (*1*) 10, (*2*) 20, and (*3*) 50 μm.

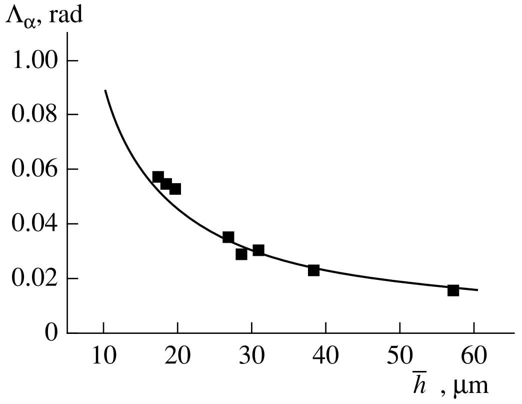


Fig. 4. Angular spacing of interference fringes versus the thickness of the lacquer coating applied on the rough metal surface. Squares, data points; continuous line, theory ($\alpha_0 \approx 45°$, $\lambda \approx 0.63$ μm, and $n \approx 1.577$).

derived from Eqs. (2) and (3):

$$\Lambda_\alpha(\alpha) = \frac{\lambda}{\bar{h}}\frac{\sqrt{n^2 - \sin^2\alpha}}{\sin 2\alpha}. \tag{5}$$

Figure 3 shows the variation of the fringe spacing with the angle of reflection α in the plane of incidence for transparent layers of various optical thickness $n\bar{h}$. The spacing tends to infinity at normal ($\alpha = 0$) and grazing incidence ($\alpha \approx 90°$). In the former case, interference rings centered on the normal to the layer surface are observed in the reflected beam. Almost straight fringes with the least spacing appear when the angles of incidence are close to 45°. It is seen from Fig. 3 that in a wide angular range centered at 45°, the spacing of the fringes varies insignificantly. The fringes are equispaced and run parallel to each other. For given n, the constancy of the spacing allows the simple calculation of the layer thickness $\bar{h}$ if the value of Λ_α is measured:

$$\bar{h} = \frac{\lambda}{\Lambda_\alpha}\sqrt{n^2 - 0.5}. \tag{6}$$

One can determine the value of Λ_α from several fringes observed in a rather wide angular aperture of the laser beam reflected (Fig. 2d).

In experiments, we used samples with randomly irregular interfaces, in particular, transparent lacquer layers of various thickness applied onto the rough metal surface. The refractive index of the lacquer ($n = 1.577$) was determined by the immersion technique [8]. The laser beam was focused on the bottom rough surface of the layer with microobjectives with $NA = 0.1$–0.2. We pre-expanded the laser beam with an auxiliary microobjective to completely fill the aperture of the focusing objective. This makes it possible to diminish the radius of the beam waist w_0 and to increase the transverse size ε_α of the speckles and the angular aperture of the field reflected. One can observe a rather large number of quasi-regular interference fringes in the cross section of the reflected beam if the angular aperture $\theta = 2\lambda/\pi w_0$ of the laser beam reflected from the smooth surface of the layer is larger than the angular spacing Λ_α of the fringes. Using Eq. (6), we obtain an additional condition for the observation of the fringes when the focused laser beam is incident at an angle of $\alpha \approx 45°$:

$$\bar{h} \geq \frac{\pi w_0}{2}\sqrt{n^2 - 0.5}. \tag{7}$$

One can use this relationship for estimating the minimum layer thickness $\bar{h}$ at which the regular interference pattern still can be observed in the reflected laser beam at a given numerical aperture NA of the focusing objective. For example, if $NA = 0.2$, $w_0 \cong \lambda/NA$, and $n = 1.5$, we find $\bar{h}_{\min} \cong 6$ μm. The maximum value of $\bar{h}_{\min}$ that can be measured in practice is limited by the resolution of the detectors and analyzers of the interference pattern with the minimum spacing $\Lambda_{\min}$.

Figure 4 shows experimental and theoretical variations of the angular spacing of the interference fringes with the thickness of the transparent lacquer coating ($n = 1.577$) covering the rough metal surface. The angle of incidence of the focused laser beam was $\alpha_0 \approx 45°$. The theoretical curve $\Lambda_\alpha(h)$ was calculated by Eq. (6). We also determined the optical thickness $\Delta = h\bar{h}$ of the lacquer layer (Fig. 4) by the method of low-coherence interferometry [9, 10] using a MII-4 Linnik interferometer equipped with a source of white light. The error of measurement was no greater than 3 μm.

Figure 5 compares the thickness $\bar{h}$ of the lacquer layers measured by the methods of focused laser beam and low-coherence interferometry. Note that the use of both these methods, as well as of low-coherence inter-

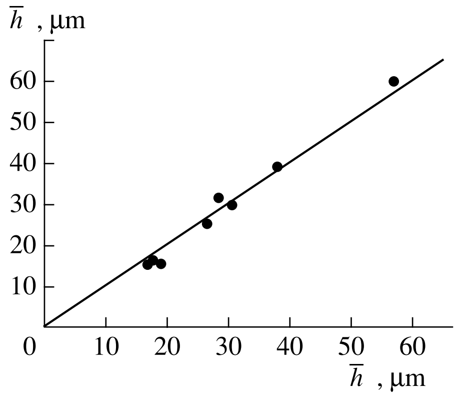


Fig. 5. Comparison of the lacquer thickness values measured by low-coherence (horizontal axis) and laser (vertical axis) interferometry.

ferometry and confocal microscopy [11–13], enables one to independently determine the geometrical thickness $\bar{h}$ of the layer and its refractive index n by jointly solving equations (6) and $\Delta = n\bar{h}$. From these equations, n and $\bar{h}$ can be approximated as $n^2 \approx (\Lambda_\alpha\Delta/\lambda) + 0.25$ and $\bar{h}^2 \approx (\lambda\Delta/\Lambda_\alpha)$ if the optical thickness Δ is sufficiently large.

Thus, the focusing of the laser beam with a small beam waist and, accordingly, a large angular aperture (divergence) on the scattering surface of a transparent layer makes it possible to observe quasi-regular interference fringes in the speckle-modulated radiation reflected. From their spacing, one can measure the local thickness of the layer on the rough surface. The method presented can be used for the contactless optical monitoring of the thickness of transparent coatings (e.g., protective lacquer coatings in microelectronics) applied on optically rough surfaces.

ACKNOWLEDGMENTS

This work was supported in part by the Russian Foundation for Basic Research (grant no. 00-15-96667), the Program "Leading Scientific Schools of the Russian Federation," and the U.S. Civilian Research and Development Foundation for the Independent States of the Former Soviet Union (CRDF) (grant no. REC-006).

REFERENCES

1. Yu. N. Popov, Prib. Tekh. Éksp., No. 3, 237 (1978).
2. P. W. Kiedron, Proc. SPIE **62**, 103 (1986).
3. A. V. Priezzhev, V. V. Tuchin, and L. P. Shubochkin, *Laser Diagnostics in Biology and Medicine* (Nauka, Moscow, 1989).
4. M. Francon, *Laser Speckle and Applications in Optics* (Academic, New York, 1979; Mir, Moscow, 1980).
5. M. Born and E. Wolf, *Principles of Optics* (Pergamon, Oxford, 1969; Nauka, Moscow, 1973).
6. J. W. Goodman, in *Laser Speckle and Related Phenomena*, Ed. by I. Dainty (Springer-Verlag, Berlin, 1975), pp. 9–75.
7. J. W. Goodman, *Statistical Optics* (Wiley, New York, 1985; Mir, Moscow, 1988).
8. N. M. Melankholin, *Methods for Studying Optical Properties of Crystals* (Nauka, Moscow, 1970).
9. D. A. Usanov, O. N. Kurenkova, A. V. Skripal', and V. B. Feklistov, USSR Inventor's Certificate No. 1772627, Byull. Izobret., No. 40 (1992).
10. D. A. Usanov and A. V. Skripal', *Television-Controlled Measuring Microscopy* (Saratovskiĭ Univ., Saratov, 1996).
11. T. Fukano and I. Yamaguchi, Opt. Lett. **21**, 1942 (1996).
12. M. Haruna, M. Ohmi, T. Mitsuyama, *et al.*, Opt. Lett. **23**, 966 (1998).
13. K. Yoden, M. Ohmi, Y. Ohnishi, *et al.*, Opt. Rev. **7**, 402 (2000).

Translated by A. Chikishev

Technical Physics, Vol. 47, No. 4, 2002, pp. 448–452. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 76–80.
Original Russian Text Copyright © 2002 by Vardanyan, Oksuzyan.

ELECTRON AND ION BEAMS, ACCELERATORS

Cerenkov Radiation Fields in an Inhomogeneously Filled Circular Waveguide

A. S. Vardanyan and G. G. Oksuzyan
Yerevan Physics Institute, Yerevan, 375036 Armenia
e-mail: vashot@jerewan1.yerphi.am

Received August 28, 2000; in final form, September 17, 2001

Abstract—Cerenkov radiation fields excited in a partially filled circular waveguide with a channel in which an electron beam propagates are studied. The surface of the channel is covered by a thin quasi-conducting layer to remove the electrostatic charge due to the beam passage. The effect of the thickness of the cylindrical dielectric layer is investigated. The thin quasi-conducting layer is shown to attenuate the Cerenkov waves generated in the lossless medium only slightly. *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The Cerenkov radiation produced by a single electron bunch and by a periodic train of bunches in a circular waveguide loaded by a lossless cylindrical layered dielectric has been studied in [1]. In particular, the cross-sectional distribution of the longitudinal component of the electric field produced by a periodic train of bunches was shown to be quasi-stepwise, remaining quasi-constant in gaps between the dielectric layers. This behavior is observed because, when the bunches follow periodically, the single-mode field in the waveguide is set up at the bunch repetition rate, i.e., with a wavelength much greater than the spacing between the dielectric layers. The waveguide radius is chosen so that this mode alone is excited (this radius will be referred to as the resonance radius), with the field intensity remaining sufficiently high. In fact, as early as in 1940, Mandelstam, and in 1947, Ginzburg and Frank [2], demonstrated that, in a channel with a radius smaller than the wavelength, Cerenkov radiation is generated as in a continuum. Such a field pattern [3–5] makes the Cerenkov radiation mechanism promising for the acceleration of charged particle bunches.

This paper discusses the problem of removing the electrostatic charge accumulated on the surface of the channel when the electron beam passes through it. This charge can be removed by applying a thin conducting layer whose parameters are chosen so as to minimize the attenuation.

Expressions for the fields are obtained here by the method developed in [1].

CERENKOV RADIATION IN THE PRESENCE OF A CONDUCTING LAYER

Let a periodic train of N identical cylindrical bunches of radius r_0, charge q, and spatial period d pass along the z axis of a circular waveguide at a speed $v = \beta c$. The charge distribution along each bunch is Gaussian: $f(\xi) = (\pi\overline{\xi^2})^{-1/2}\exp(-\xi^2/\overline{\xi^2})$; in the cross section, the bunches are homogeneous.

The waveguide is filled with a dispersed dielectric having an empty channel of radius b. A thin conducting film of thickness $\delta = b - b_1$ ($b_1 > r_0$) is applied on the interior surface of the channel (Fig. 1).

The permittivity of the dielectric insert is $\varepsilon = \varepsilon_1$ (in the region $b \le r \le a$). In the channel, $\varepsilon_{ch} = \varepsilon_3 = 1$. In the conducting layer, the permittivity is complex-valued: $\varepsilon_2 = \varepsilon' + i\varepsilon''$.

We will seek solutions for the potential φ in the cylindrical coordinate system (r, φ, z) in the form

$$\varphi = \int_{-\infty}^{\infty} \Phi(\omega, r)\frac{2I_1(kr_0)}{kr_0}\phi_p\left(\frac{\omega}{v}\right)\exp\left(i\frac{\omega}{v}(z - vt)\right)d\omega, \quad (1)$$

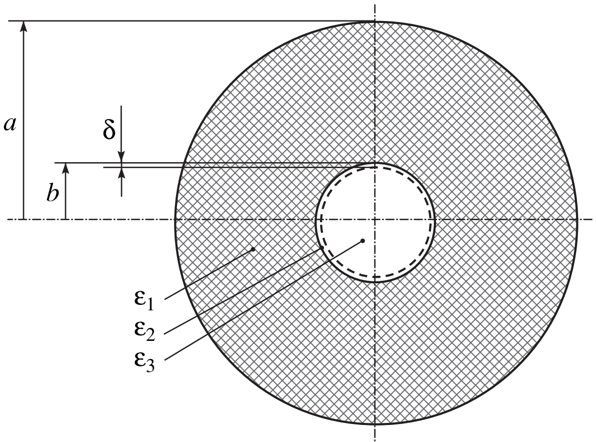


Fig. 1. Waveguide cross section: a is the waveguide radius, b is the channel radius, and δ is the thickness of the conducting layer.

1063-7842/02/4704-0448$22.00 © 2002 MAIK "Nauka/Interperiodica"

where

$$\Phi(\omega, r) = \begin{cases} \dfrac{q}{\pi v}[K_0(kr) + \alpha I_0(kr)] & 0 < r \le b - \delta \\ \dfrac{q}{\pi v \varepsilon_2}[\eta K_0(sr) + \gamma I_0(sr)] & b - \delta \le r \le b \\ \dfrac{q}{\pi v \varepsilon_1}[\Theta K_0(-ipr) + \xi I_0(-ipr)] & b \le r \le a. \end{cases} \quad (2)$$

Here,

$$k = \frac{\omega}{v}\sqrt{1-\beta^2}; \quad s = \frac{\omega}{v}\sqrt{1-\varepsilon_2\beta^2};$$

$$p = \frac{\omega}{v}\sqrt{\varepsilon_1\beta^2 - 1};$$

I_0, I_1, and K_0 are the modified Bessel functions; and

$$\phi_p\left(\frac{\omega}{v}\right) = \sum_{n=1}^{N} \phi\left(\frac{\omega_n}{v}\right)\exp\left(i\frac{\omega_n}{v}(z - vt)\right), \quad (3)$$

where the function

$$\phi\left(\frac{\omega}{v}\right) = \left(\int f(\xi)d\xi\right)^{-1}\int f(\xi)e^{-i\frac{\omega}{v}\xi}d\xi$$

describes the radiation of a single bunch and summation is performed over positive and negative frequencies.

The coefficients α, η, γ, Θ, and ξ in (2) are determined from the following boundary conditions for the Fourier components of the fields and potentials:

$$E_{z\omega 3} = E_{z\omega 2} \quad \text{or} \quad (1-\beta^2)\varphi_{\omega 3} = (1-\varepsilon_2\beta^2)\varphi_{\omega 2} \quad (\text{at } r = b_1);$$

$$H_{\varphi\omega 3} = H_{\varphi\omega 2} \quad \text{or} \quad \frac{\partial\varphi_{\omega 3}}{\partial r} = \varepsilon_2\frac{\partial\varphi_{\omega 2}}{\partial r} \quad (\text{at } r = b_1);$$

$$E_{z\omega 1} = E_{z\omega 2} \quad \text{or} \quad (1-\varepsilon_1\beta^2)\varphi_{\omega 1} = (1-\varepsilon_2\beta^2)\varphi_{\omega 2} \quad (\text{at } r = b); \quad (4)$$

$$H_{\varphi\omega 1} = H_{\varphi\omega 2} \quad \text{or} \quad \varepsilon_1\frac{\partial\varphi_{\omega 1}}{\partial r} = \varepsilon_2\frac{\partial\varphi_{\omega 2}}{\partial r} \quad (\text{at } r = b);$$

$$E_{z\omega 1} = 0 \quad \text{or} \quad \varphi_{\omega 1} = 0 \quad (\text{at } r = a).$$

The longitudinal electric field component E_z in the waveguide can be written as

$$E_z = -\frac{q(1-\beta^2\varepsilon)}{v}\mathrm{Re}\int_{-\infty}^{\infty}\Phi(\omega, r) \times \frac{2I_1(kr_0)}{kr_0}\phi_p\left(\frac{\omega}{v}\right)e^{i\frac{\omega}{v}(z-vt)}i\omega d\omega. \quad (5)$$

Let the medium that fills the waveguide meet the Cerenkov condition $\varepsilon_1\mu_1\beta^2 > 1$. In this case, the coefficient a is given by

$$\alpha = \left(K_1(kb_1) + \frac{\varepsilon_2}{fs}kK_0(kb_1)\right)D^{-1}, \quad (5a)$$

where

$$D = I_1(kb_1) - \frac{\varepsilon_2}{fs}kI_0(kb_1),$$

$$f = \frac{I_0(sb_1) + \chi K_0(sb_1)}{I_1(sb_1) + \chi K_1(sb_1)}, \quad (5b)$$

$$\chi = \frac{p\Psi_1\varepsilon_2 I_1(sb) - s\Psi_2\varepsilon_1 I_0(sb)}{p\Psi_1\varepsilon_2 K_1(sb) + s\Psi_2\varepsilon_1 K_0(sb)}.$$

Without going into details of calculating the coefficients η, γ, Θ, and ξ, we give the final formulas for the combinations of the functions I_0 and K_0 that enter into expressions (2):

$$\eta K_0(sr) + \gamma I_0(sr) = -\frac{\varepsilon_2 k}{s^2 p}\frac{I_0(sr) + \chi K_0(sr)}{I_0(sb_1) + \chi K_0(sb_1)}D^{-1}, \quad (5c)$$

$$\Theta K_0(-ipr) + \xi I_0(-ipr) = -\frac{\varepsilon_1 k}{b_1 p}\frac{I_0(sb) + \chi K_0(sb)}{I_0(sb_1) + \chi K_0(sb_1)}\frac{\Psi_1(r)}{\Psi_1(b)}D^{-1}, \quad (5d)$$

where Ψ_1 and Ψ_2 are the Abel functions

$$\Psi_1(b) = J_0(pb)N_0(pa) - J_0(pa)N_0(pb), \quad (5e)$$

$$\Psi_2 = J_1(pb)N_0(pa) - J_0(pa)N_1(pb). \quad (5f)$$

Integration over frequency in (5) is performed by the method of residues on the complex plane ω as described in [1]. The poles are calculated from the equation $D = 0$. Taking into account that solutions of this equation are complex-valued, we obtain the following expressions for the longitudinal electric field component (for the three regions of the waveguide):

$$E_{z_1} = -\frac{2q\sqrt{1-\beta^2}}{b_1 v}\sum_{\lambda}\frac{I_0(sb) + \chi K_0(sb)}{I_0(sb_1) + \chi K_0(sb_1)}\frac{\Psi_1(r)}{\Psi_1(b)} \times \frac{2I_1(kr_0)}{kr_0}F\left(\frac{\omega}{v}\right)\Big/\frac{dD}{d\omega}\bigg|_{\omega = \omega_\lambda}, \quad (6a)$$

$$E_{z_2} = -\frac{2q\sqrt{1-\beta^2}}{b_1 v}\sum_{\lambda}\frac{I_0(sr) + \chi K_0(sr)}{I_0(sb_1) + \chi K_0(sb_1)} \times \frac{2I_1(kr_0)}{kr_0}F\left(\frac{\omega}{v}\right)\Big/\frac{dD}{d\omega}\bigg|_{\omega = \omega_\lambda}, \quad (6b)$$

$$E_{z_3} = -\frac{2q\sqrt{1-\beta^2}}{b_1 v}\sum_{\lambda}\frac{I_0(sr)}{I_0(sb_1)}\frac{2I_1(kr_0)}{kr_0} \times F\left(\frac{\omega}{v}\right)\Big/\frac{dD}{d\omega}\bigg|_{\omega=\omega_\lambda}, \tag{6c}$$

where

$$F\left(\frac{\omega}{v}\right) = \frac{2\exp\left(-\frac{\omega^2-w^2}{4v^2}\overline{\xi^2}-\frac{w}{v}|z-vt|\right)}{1-\exp\left(-\frac{wd}{v}\right)\cos\frac{\omega d}{v}+\exp\left(-\frac{2wd}{v}\right)}$$
$$\times\left\{\cos\frac{\omega}{v}\left(z-vt+\frac{w\overline{\xi^2}}{2v}\right)-\exp\left(-\frac{wd}{v}\right)\right.$$
$$\times\cos\frac{\omega}{v}\left(z-vt+\frac{w\overline{\xi^2}}{2v}\right)-\exp\left(-\frac{wNd}{v}\right) \tag{7}$$
$$\times\left[\cos\frac{\omega}{v}\left(z-vt+\frac{w\overline{\xi^2}}{2v}-Nd\right)-\exp\left(-\frac{wd}{v}\right)\right.$$
$$\left.\left.\times\cos\frac{\omega}{v}\left(z-vt+\frac{w\overline{\xi^2}}{2v}-(N-1)d\right)\right]\right\}.$$

Here, N is the number of bunches in the train, $\mathrm{Re}\,\omega \longrightarrow \omega$, and $\mathrm{Im}\,\omega \longrightarrow w$. In the conducting layer, the real part of the permittivity is taken to be equal to 1.

The radiation is maximum when the fields produced by individual bunches add in phase. If the thickness of the conducting layer and the radius of the channel are given, the waveguide radius can be chosen so as to satisfy the condition for in-phase radiation of the bunches $(\omega_{nl}/2v)d = \pi l$ (l = 1, 2, …). In this case, the single-mode regime sets up at a given frequency (mode) that is equal to or is a multiple of ($l \neq 1$) the bunch repetition rate. When $n = 1$ and $l = 1$, the fundamental mode grows (at $\omega = 2\pi v/d$) and we have

$$F_p\left(\frac{\omega_1}{v}\right) = 2\exp\left(-\left[1-\left(\frac{w_1}{\omega_1}\right)^2\right]\frac{\pi\overline{\xi^2}}{d^2}-\frac{w_1}{v}|z-vt|\right)$$
$$\times\frac{1-\exp\left(-N\frac{w_1 d}{v}\right)}{1-\exp\left(-\frac{w_1 d}{v}\right)}\cos\frac{\omega_1}{v}\left(z-vt+\frac{w_1\overline{\xi^2}}{2v}\right). \tag{7a}$$

When $w = \mathrm{Im}\,\omega \longrightarrow 0$,

$$\lim_{w\to 0}F_p\left(\frac{\omega_1}{v}\right) = 2N\exp\left(-\frac{\omega_1^2\overline{\xi^2}}{4v^2}\right)\cos\frac{\omega_1}{v}(z-vt).$$

Table

l	λ, cm	δ, cm	$\|E_1\|$, kV/m
1	10.0	3.1	48.490
2	5.0	1.4	42.230
3	3.33	0.81	33.750
4	2.5	0.65	24.950
5	2.0	0.4	17.200
6	1.66	0.3	11.150
7	1.43	0.25	6.850
8	1.25	0.2	4.050
9	1.11	0.17	2.350
10	1.0	0.14	1.400
11	0.91	0.12	1.300
12	0.83	0.1	0.130

Further, to clarify the effect of the conducting layer on Cerenkov radiation, we will consider the field in a very thin (in terms of wavelength) cylindrical dielectric layer coaxial with the waveguide. Having studied the field in this layer, we will turn to the effect of a thin conducting layer applied on the interior surface of the channel.

LOSSLESS DIELECTRIC CYLINDRICAL INSERT IN THE WAVEGUIDE

The study of Cerenkov radiation in a cylindrical dielectric layer is of interest for the waveguide technology and also helps to make some useful estimates. The formulas obtained in [1] can be used to develop an algorithm for numerically analyzing the effect of layers with various thicknesses, in particular, very thin layers. In order to perform a comprehensive analysis of the field, we calculated the resonance radius of the waveguide (at the bunch repetition rate f = 3 GHz and the permittivity $\varepsilon = 2.1$), a = 3.8 cm. With this value, the main condition for the Cerenkov radiation to have a discrete spectrum at harmonics of the bunch repetition frequency, i.e., the condition for synchronism between the charge motion and the phase velocity of the wave generated ($v_{\mathrm{ph}} = \omega/\gamma = v$), can be satisfied by choosing the appropriate radius and thickness of the dielectric pipe. It was found that, when the dielectric cylinder is very thin, the particle velocity can be equal to the phase velocity of the wave only for modes with very large numbers (for very short wavelengths). In other words, with a thin dielectric insert, the structure becomes a below-cutoff waveguide at longer wavelengths (smaller mode numbers). It should be noted here that, at very short wavelengths, the condition $\beta^2\varepsilon > 1$ is violated because the permittivity tends to 1. At a wavelength $\lambda \approx$ 10 cm, the waveguide with a cylindrical insert less than 1 cm thick is evanescent. For the second harmonic of the bunch repetition frequency ($\lambda \approx 5$ cm) to propagate

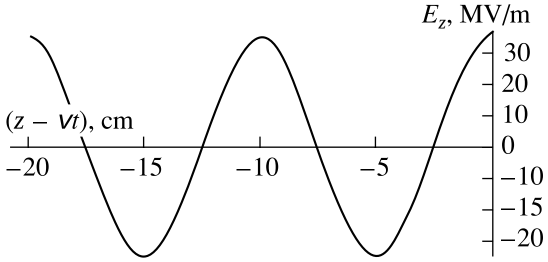


Fig. 2. E_z component of the field produced at the axis of the channel by a train of $N = 10000$ bunches versus parameter $(z - vt)$.

in the waveguide, the thickness must be 0.5 cm or less; for the fifth harmonic ($\lambda \approx 2$ cm), ≈1 cm; and so on. The field amplitude decreases with increasing harmonic number. In our calculations, the waveguide radius $a = 3.8$ cm and the channel radius $b = 0.7$ cm were constant and the thickness δ was taken such that this waveguide radius ($a = 3.8$ cm) was resonant for various harmonics of the bunch repetition frequency.

The table lists the electric field amplitude for the fundamental mode versus thickness of the dielectric cylinder. The data were calculated by formulas obtained in [1] for a single bunch ($N = 1$). As follows from the table, the amplitude of the fundamental mode rapidly decreases with decreasing thickness δ. In particular, at $\delta = 0.05$ cm, the amplitude $|E_1|$ is as small as 0.07 V/m; therefore, the generation of the Cerenkov waves in a thin (≈70 Å thick) conducting layer can be neglected and the real part of the permittivity can be assumed to be equal to 1 ($\varepsilon' = 1$).

INTENSITY OF THE CERENKOV RADIATION (QUASI-CONDUCTING LAYER)

The thin conducting layer may significantly decrease the field amplitude and reduce the efficiency of the Cerenkov radiation mechanism in the two-beam acceleration scheme to zero. In fact, calculations by formulas (6) and (7) for a 10-Å-thick high-conductivity ($\varepsilon'' \approx 10^7$) layer show that the field attenuation is intolerably high. At the same time, to remove the static charge accumulated on the surface of the channel, it is not necessary to use a good conductor: a quasi-conducting material would suffice for this purpose.

Figure 2 shows the field produced by 10 000 bunches, each comprising 3×10^9 electrons, at the waveguide axis for a 70-Å-thick layer and the permittivities $\varepsilon' = 1$ and $\varepsilon'' = 10^3$. The field was calculated by formulas (6a) and (7). Estimates made for eleven modes show that only the fundamental mode is enhanced in the periodic train of bunches and the quasi-monochromatic wave with an amplitude of ≈35 MV/m sets up, as illustrated in Fig. 2.

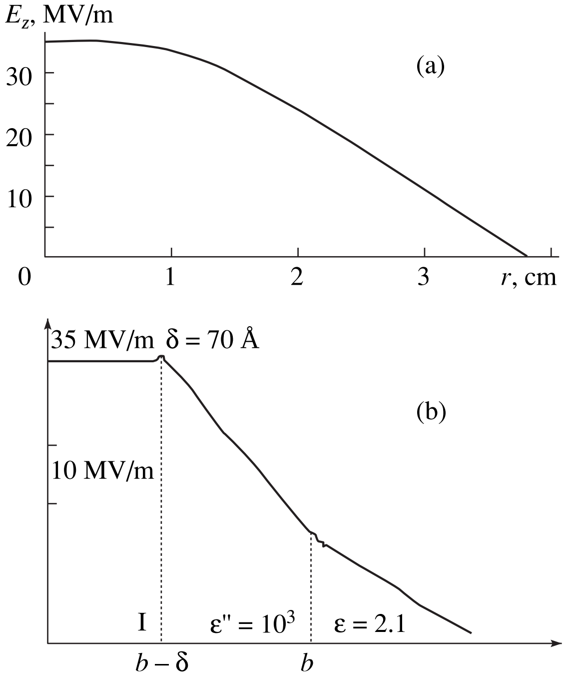


Fig. 3. Radial distributions of the field component E_z over the cross section $(z - vt) = 0$ in the (a) waveguide and (b) conducting layer I. $N = 10000$ bunches.

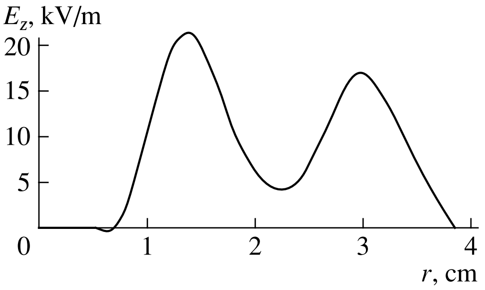


Fig. 4. Radial field distribution of eleven higher modes over the cross section $(z - vt) = -2.5$ cm at $N = 10000$ bunches (at the axis, the field is zero).

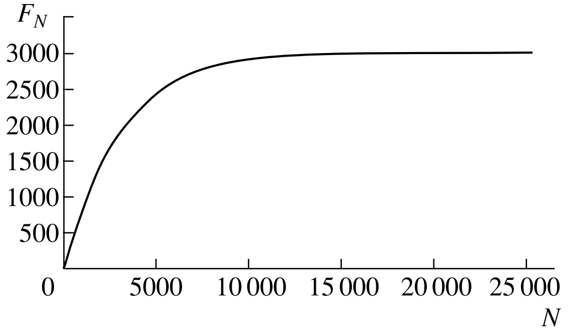


Fig. 5. Interference factor F_N versus number of bunches. The curve rapidly saturates.

Thus, the 70-Å-thick quasi-conducting layer does not prevent the generation of a sufficiently strong field, while removing the static charge.

Figure 3a plots the cross-sectional field distribution in the presence of the 70-Å-thick quasi-conducting layer. Inside the layer, the field decreases more rapidly than in the dielectric; in the channel, it is quasi-uniform as before. For illustration, Fig. 3b represents the field distribution in the layer on an enlarged (by a factor of 10^6) scale.

Figure 4 shows the radial distribution of the total field from 11 higher modes at the zero phase of the first mode ($z - vt = -2.5$ cm). The field amplitude is less than 0.1% of the fundamental mode amplitude (20 kV/m and 35 MV/m) in this case.

In Fig. 5, the interference gain factor

$$F_N = \frac{1 - \exp\left(-N\frac{w_1 d}{v}\right)}{1 - \exp\left(-\frac{w_1 d}{v}\right)}$$

is plotted against N for the case when the fields produced by N bunches add in phase. As is seen, the lossy film causes the curve to saturate. At $N = 3000$ bunches, the gain is ≈70%; at $N = 10\,000$, ≈99%.

CONCLUSION

The channel through which electron bunches pass insignificantly decreases the field intensity at the bunch repetition frequency when the wavelength of the radiation (≈10 cm) is much greater than the channel diameter (1.4 cm). At the fundamental harmonic, the longitudinal electric field component is almost constant across the channel.

The study of the Cerenkov radiation in the thin cylindrical layer has shown that Cerenkov waves are efficiently generated at very high harmonics of the fundamental frequency and that the dielectric layer exerts a minor effect on the Cerenkov radiation at the repetition frequency of the electron bunches. Therefore, with the complex permittivity, the attenuation due to the dielectric loss will primarily be observed.

The Cerenkov radiation in a circular dielectric-loaded waveguide with an axial channel covered by a high-conductivity layer has been considered. Even if this layer is very thin (≈10 Å), it introduces a strong attenuation into the system, making it low-efficient. At the same time, a low-conductivity layer can provide both a small attenuation and the effective removal of the static charge from the wall.

Finally, in a waveguide with high losses (nontransparent medium), the condition for Cerenkov radiation (synchronism between the phase velocity of the wave generated and the velocity of the emitting charge) becomes uncertain, because the phase velocity itself turns out to be complex. Therefore, the effect of the real part of the permittivity should be neglected, especially because, for a very thin (in terms of wavelength) dielectric layer, $\varepsilon \longrightarrow 1$ at high frequencies and the condition for Cerenkov radiation is violated as we noted above.

ACKNOWLEDGMENTS

We are grateful to É.D. Gazazyan for the valuable discussions. This work was supported by the INTAS (grant no. A-087).

REFERENCES

1. A. S. Vardanyan, Izv. Akad. Nauk Arm., Fiz. **34**, 323 (1999).
2. V. L. Ginzburg and I. M. Frank, Dokl. Akad. Nauk SSSR **56**, 699 (1947).
3. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1984).
4. B. M. Bolotovskiĭ, Usp. Fiz. Nauk **75**, 295 (1961) [Sov. Phys. Usp. **4**, 781 (1962)].
5. É. D. Gazazyan and É. M. Laziev, Izv. Akad. Nauk Arm. SSR, Ser. Fiz.-Mat. Nauk **16** (2), 79 (1963).

Translated by A. Khzmalyan

Technical Physics, Vol. 47, No. 4, 2002, pp. 453–458. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 81–87.
Original Russian Text Copyright © 2002 by Kozlov.

The Laser-Induced Formation of Superhard Structures and Phase Transformations of Carbon in the Near-Surface Layer of Cast Iron

G. I. Kozlov

Institute of Problems of Mechanics, Russian Academy of Sciences, Moscow, 117526 Russia

Received May 10, 2001; in final form, August 20, 2001

Abstract—Superhard structures with a microhardness of (2–5) × 10^{10} Pa formed in the near-surface layer of gray iron covered by a thin copper inductor and processed by laser radiation are experimentally studied. The copper favors the formation of globular (or even spherical) carbon and stimulates the saturation of the near-surface layer of iron by carbon. The elemental analysis is performed, and the Raman spectrum from carbon globules is recorded. Based on this spectrum and the spectra from other carbon structures, the author argues that the laser processing converts the graphite of the near-surface layer to pyrolytic carbon, which offers a high hardness. A possible mechanism behind this process is discussed. *© 2002 MAIK “Nauka/Interperiodica”.*

Laser quenching is a most effective way of hardening the metal and alloys surfaces. This process is highly nonequilibrium because the characteristic time of crystal structure rearrangement in metals and alloys subjected to laser processing is comparable to the heating and cooling times. Physically, the hardening of metals and alloys implies the production of structures making plastic deformation (or, in other words, the motion of dislocations) difficult. Basically, this can be achieved in different ways, e.g., by producing phase inhomogeneities in the crystal structure of metals and alloys, decreasing the grain size, or fabricating harder and more perfect crystal structures.

In this work, which elaborates upon the studies [1, 2], an intriguing phenomenon is reported: the induced formation of superhard structures in the near-surface layer of gray iron when the material is covered by a thin copper inductor and then laser-processed. The term “induction” here means that the crystal structure of the inductor controls (during cooling) the postirradiation phase transformations in near-surface layers of metals and alloys being in contact with the inductor.

EXPERIMENTAL SETUP AND METHODS OF INVESTIGATION

The structure and phase transformations taking place in the near-surface layer of gray iron under the laser irradiation were studied with the special setup incorporating a cw multibeam 6-kW CO_2 laser with diffusion cooling [3], a focusing system, and a special control device that provides scanning over the surface being processed. As a focusing means, a NaCl lens with a focal distance of 1.1 m was used. The lens could move along the laser beam axis, thus making it possible to vary the size of the focal spot and, hence, the radiation intensity on the surface. The combination of the multibeam CO_2 laser and the long-focus lens provided the uniform distribution of the laser intensity over the central part of the beam in the focal plane and a natural decline of the intensity to the beam periphery.

To trigger the mechanism inducing the phase transformations, a thin copper layer was applied on the iron surface immediately before the processing. Copper as an inductor was taken because it, like γ-Fe, has the fcc lattice and is therefore expected to strongly influence the phase transitions, the structure and properties of various constituents, and eventually the physical and mechanical behavior of the iron surface. After a number of trials, we succeeded in developing the effective technique for applying a thin copper layer on iron with a special paste. The paste consists of copper powder and the emulsion readily absorbing the radiation of a CO_2 laser. It is easy to apply on the surface, and its thickness in these experiments was 100–150 μm. Once the paste had been dried, the surface was irradiated. The radiation power was kept constant at 5.5 kW, and the radiation intensity was selected in such a way that the thickness of the molten iron was small and the iron–copper mixing was insignificant. With this beam power, the above condition was met when the focal spot diameter was about 8 mm and the scan rate over the surface, 0.67 cm/s.

After the processing, the copper layer on the surface was removed by milling to a depth of 0.5 mm (first step) and 0.9 mm (second step). Then, a 10-μm-thick layer was chemically etched off to minimize the effect of the milling-deformed zone. The milling exposed the specular matrix within which superhard structures formed during the laser processing, as will be evidenced by further examination. We studied the phase composition, microhardness, and microstructure on the surface of the steps and on the end faces of specially prepared specimens.

The phase and elemental compositions, as well as the microstructure, were studied by X-ray diffraction (monochromatic FeK_α radiation) and in a CAMSCAN CS44C-100S scanning electron microscope equipped with a LINC ISIS-L200D energy dispersion analyzer. The size of the excitation zone was about 0.3 mm^2. The microhardness distribution across the laser radiation track was found with a PMT-3 hardness meter at an indentation load of 0.1 kg. The microhardness was also measured in separate regions with various optical contrast.

MICROHARDNESS, PHASE COMPOSITION, AND MICROSTRUCTURE OF THE LASER IRRADIATION ZONE

The microhardness of the matrix was naturally of primary interest early in this study. Its distribution (0.25-mm pitch) across the hardened zone at the first step of the specular matrix is shown in Fig. 1. The distribution is nonuniform: in places, the microhardness is as high as $(2–5) \times 10^{10}$ Pa. This means that superhard structure components have formed. The highest value, 5.15×10^{10} Pa, was observed within a regularly shaped light gray region of size about 20 μm. In an optical microscope, one can distinguish several structure components differing in color, shape, and hardness. We failed to measure the microhardness of most dark globular precipitates (as will be referred to later), because they do not give optical contrast (the indentations against the dark background are not seen). In deeper layers (second step), the microhardness values were lower; however, here too, superhard regions, $(1.3–1.5) \times 10^{10}$ Pa, were observed. Such high values cannot be associated with laser thermal hardening, which provides the microhardness in the range of $(0.8–1.1) \times 10^{10}$ Pa [4].

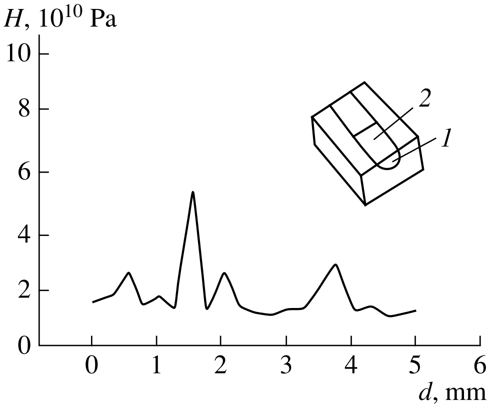


Fig. 1. Microhardness H across the hardened zone (d is the width of the zone). A cast iron specimen with a laser irradiation track is shown. The composition, microstructure, and microhardness were studied on the (*1*) end face and (*2*) on the surface.

The superhard structures formed by the laser processing of the iron surface covered by the thin copper film is undeniably related to the inducing effect of copper on the mechanism and the kinetics of postirradiation phase transformations (i.e., those proceeding during cooling). It is important to identify structure components responsible for the so high microhardness values. For this purpose, we performed the phase composition analysis.

In this analysis, the diffraction pattern from the area processed was compared with that from the unprocessed region, which incorporated largely α-Fe, Fe_3C, and carbon. The pattern taken from the first step in the specular matrix indicates that here the material consists of ferrite, austenite, cementite, and carbon. All reflections are strongly diffuse in comparison with those for the unprocessed region, indicating the disperse structure of these components. In the range of angles corresponding to the martensite doublet, the split of the reflection is faint. In deeper layers (second step), the intensity of martensite doublet reflection grows and the reflections become narrower. From these results, we can conclude, with surprise, that the hardening of gray iron is not related to the formation of martensite. It seems that the inducing effect of the copper is in that it acts as an austenite-stabilizing agent, reducing the austenite transformation temperature and thus greatly refining pearlite, which forms during cooling. However, grain refining is only one reason for iron hardening by laser irradiation under our experimental conditions. We failed to determine structure components responsible for the so high microhardness values with the help of the phase analysis. It was established, however, that martensite forms in small amounts and is not responsible for the laser-induced hardening of inductor-covered iron.

The microstructure of the specimen was examined on its end face. The optical image of the specular zone is shown in Fig. 2, where the boundary of the molten zone is distinctly seen. Within this zone, the surface layer of the iron is heavily saturated by carbon. Along with black carbon precipitates, dark gray needles (presumably bainite) and light ferrite regions with round-ended dispersed cementite precipitates are observed. The microstructure varies with distance from the surface. The concentration of the graphite precipitates decreases, and lamellar pearlite becomes a dominant component. The number of the light ferrite regions with cementite precipitates inside also decreases.

More detailed information on the near-surface microstructure of the iron in the matrix zone was derived with the electron microscopic pattern taken from the central region of the laser spot (Fig. 3). Here, three layers can be distinguished: a ≈200-μm-thick dark near-surface layer with the dispersed structure, a ≈100-μm-thick transition layer with a minor concentra-

Fig. 2. Microstructure of the specular matrix zone. (*1*) Boundary of the molten zone.

tion of precipitates, and a light layer with carbon in the globular and lamellar forms. As follows from the elemental analysis, the dark layer consists mainly of carbon and copper, while the transition layer includes mainly iron and carbon. At the point marked by the cross in Fig. 2, the element concentrations are as follows: carbon, 37.82%; iron, 58.56%; silicon, 2.88%; and copper, 0.74%. Thus, having studied the phase composition and the microstructure of the matrix, we conclude that the inducing effect of the copper is twofold: the considerable enrichment of the surface by carbon and the stabilization of austenite. The latter appreciably reduces the austenite transformation temperature and, accordingly, refines structures resulting during cooling (including carbides and pearlite).

THE TRANSFORMATION OF GRAPHITE INTO PYROLYTIC CARBON DURING LASER PROCESSING

The question remains to be answered as to which structure constituents are responsible for the high microhardness values. The only way to clarify the situation seems to be the in-depth analysis of the iron surface microstructure. The micrographs taken both before and after the laser irradiation (Fig. 4) distinctly show that the molten zone contains a large amount of globular carbon precipitates. Note first of all that, during the crystallization of the iron from the melt, the graphite nucleates at one site and then grows, taking on the form of heavily curved flakes. Therefore, in the plane of the metallographic section (Fig. 4b), graphite in iron is usually seen as straight and curved lamellas, which are different sections of the graphite flakes. After the laser irradiation of the iron surface, the carbon structure in the near-surface layer radically changes, as follows from Fig. 4a. The initial variously shaped lamellas virtually transform into globules, some of which having a nearly perfect spherical shape (Fig. 5).

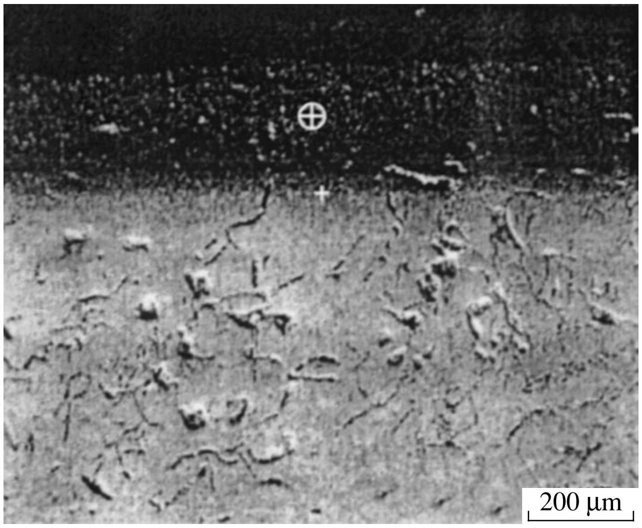


Fig. 3. Microstructure of the iron surface layer across its depth after laser processing.

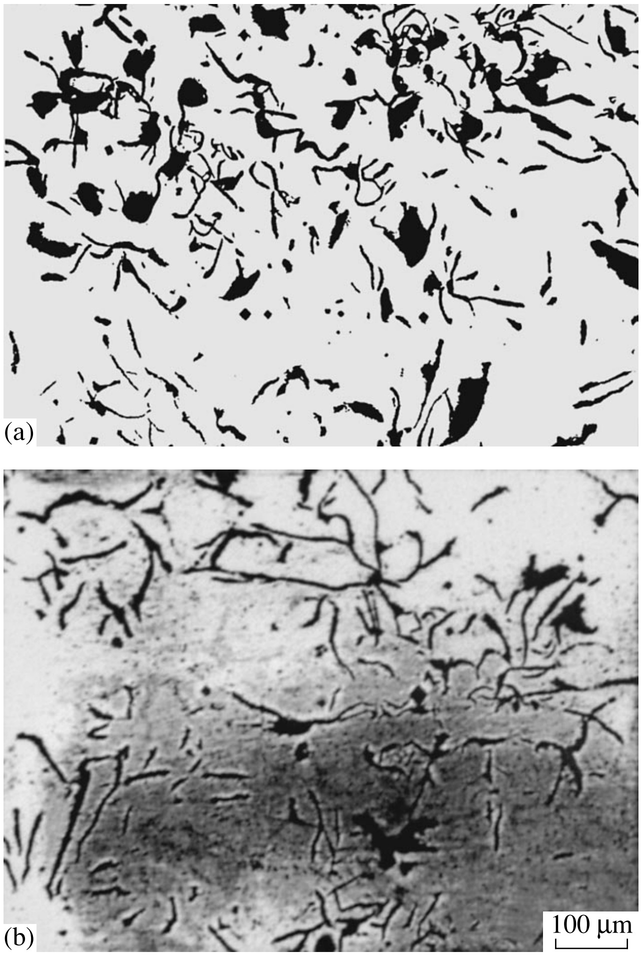


Fig. 4. Microstructure of the iron surface layer (×100) (a) after and (b) before laser processing.

Fig. 5. Microstructure of the laser-processed zone with spherical carbon globules (×400).

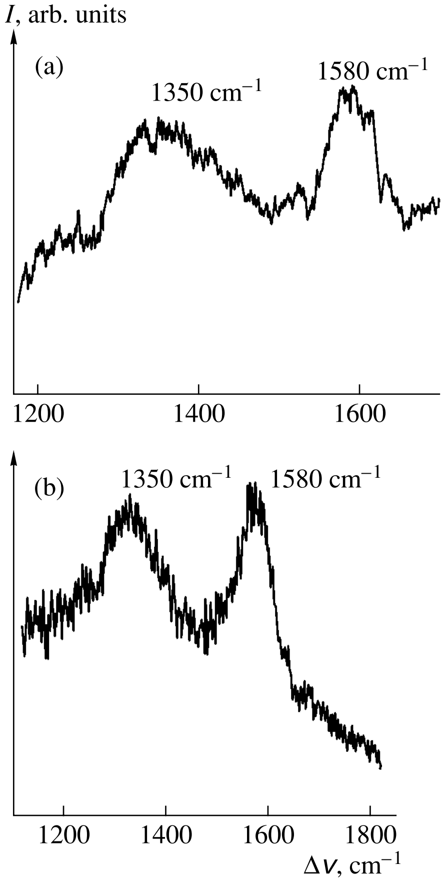


Fig. 6. Raman spectrum intensity vs. Raman line shift $\Delta\nu$: (a) carbon inclusions in the laser irradiation zone and (b) pyrocarbon.

Next, it is necessary to elucidate the structure and composition of the globular carbon. Whether it is graphite or has some another structure? To answer this question, we performed appropriate investigations. The elemental analysis showed that some globules consist of nearly pure graphite, while others contain silicon (to 7.1%), chromium (4.5%), iron (1.1%), and oxygen (20.7%) impurities. Such a difference in the globule composition is likely to be associated with the dynamics of the phase transformations and diffusion processes during the laser irradiation. The concentration of the first three impurities drops with irradiation time. As for the oxygen, it is by far adsorbed on carbon from air after the metallographic section has been prepared and the globules have been exposed. Thus, the globules can be thought of as carbon structures forming from the melt.

The structure of the globular carbon was judged from the Raman spectra for carbon inclusions in the irradiation zone at the first step in the specular matrix. Unlike the Raman spectra for pyrolytic graphite and diamond, of which narrow peaks at 1580 and 1332 cm^{-1} are typical, the spectrum of the globular carbon (Fig. 6) has a specific two-humped structure. This circumstance and also the widths of the peaks indicate that the laser processing of the copper-covered gray iron surface causes a strong disorder in the structure of the graphite contained in the near-surface layer of the gray iron. As a result, long-range order, typical of crystal structures, practically disappears, and the molecular carbon structures become fragmented and chaotic; that is, carbon becomes amorphous. The widths of the peaks in the Raman spectrum imply that the particles are small, 50–80 nm in size.

It was of interest to compare the Raman spectrum of the globular carbon with that of CVD pyrocarbon [5]. The latter spectrum is shown in Fig. 6b. The surprising thing is that these spectra, which were taken from the structures prepared under entirely different conditions, are almost identical. One therefore can conclude that the laser processing of the copper-covered iron converts the near-surface graphite to pyrocarbon, offering a high hardness.

GRAPHITE-TO-PYROCARBON CONVERSION MECHANISM

A mechanism of graphite-to-pyrocarbon conversion deserves special discussion. It can be suppose that, over a period of time after the laser irradiation, the near-surface carbon is in the liquid state. This supposition is based on the analysis of the iron microstructure in the irradiation zone and explains the shrinkage of the carbon into globules and the subsequent formation of the amorphous pyrocarbon. In other words, it may be conceived that the amorphous pyrocarbon is the supercooled liquid phase passed to the condensed state. It is known that if the cooling rate of some alloys is sufficiently high, the viscosity of the liquid metal increases to such a degree that crystallization centers have no time to arise and the metal solidifies into an amorphous structure with short-range order. If short-range order extends to 1.3–1.8 nm, such a structure is considered to be amorphous [4]. If its characteristic size is between 2 and 10 nm, the structure is fine-grained. Since dislocations cannot glide over slip systems in amorphous alloys, they usually offer a very high strength.

Until recently, it was however believed that the liquid phase of graphite (for example, due to laser processing) can form and exist only under high pressures. It is easy to show that high-pressure conditions may occur in our experiments. First, we note that the temperature and phase stresses induced by the laser processing of the copper-covered iron in the surface and subsurface layers of the metal develop in a complex manner. After the irradiation, the surface layer cools and shrinks last. When shrinking, it acts on the subsurface layer, causing compressive temperature stresses in the latter, while the surface is under tensile stresses. If the phase transformations in the subsurface layer cause an increase in the specific volume (for example, martensitic transformations), the compressive stresses will inevitably grow because of the phase component. Analysis of the temperature and phase stresses arising upon laser processing of copper-covered cast iron is a complex problem, which calls for special consideration.

Of interest is to estimate temperature stresses arising in the iron when the iron–copper system solidifies by cooling from the laser processing temperature to room temperature. The reason for the stresses is the great difference between the thermal expansion coefficients of the metals. For copper and iron plates of length l in the free state, the difference in the linear expansion coefficients would cause a difference Δl in the linear dimension. Let α_1 and α_2 be the thermal expansion coefficients of copper and iron, respectively, and T_1 and T_0 be the melting point of copper and the normal temperature. Then, the value of Δl is given by

$$\Delta l = l(T_1 - T_0)(\alpha_2 - \alpha_1). \tag{1}$$

Under experimental conditions, the linear sizes will not of course differ because the copper and iron layers are fused into each other.

Using the Hooke law and taking into account the fact that in equilibrium compressive stresses in the iron must be balanced by tensile stresses in the copper, we arrive at the approximate expression for the compressive stress σ_c in the near-surface iron layer:

$$\sigma_c = -\frac{E_1 E_2}{E_1 + E_2}(T_1 - T_0)(\alpha_2 - \alpha_1), \tag{2}$$

where E_1 and E_2 are the respective elastic moduli of copper and iron.

Substituting the numerical values of the quantities involved in Eq. (2), we find that stresses arising in the iron under the experimental conditions may reach 4×10^8 Pa. This value is consistent with the literature data [4] obtained by directly measuring temperature stresses due to laser processing. Thus, the simplified one-dimensional estimate of the possible stresses and the associated literature data indicate that the crystallization of the copper–iron system from the melt after laser processing causes stresses as high as 4×10^8 Pa due to the considerable difference in the temperature coefficients of linear expansion. So high stresses may, in principle, favor the transition of graphite to the liquid state.

The question of whether or not liquid graphite exists under similar conditions has arisen many times but still remains unresolved. It has been believed until recently that graphite heated to temperatures above 3800 K under the atmospheric pressure vaporizes without passing to the liquid phase. However, if the pressure is increased to 10^7 Pa, graphite passes to the liquid (molten) state. Indeed, the signs of thermal reflowing were detected in experiments on laser heating of graphite at pressures $p > 10^7$ Pa. It was supposed [6] that the graphite reflowing also takes place at $p < 10^7$ Pa. In this case, however, the reflowing signs cannot be detected, since the film of the melt is very thin and quickly evaporates after the heating. The parameters of the triple (solid–liquid–vapor) point for carbon were determined by extrapolating the curves of phase equilibrium for graphite melting and evaporation [6]. They were found to be $p_T = 10^5$ Pa and $T_T = 4000$ K. These values considerably differ from those established earlier ($p_T = 10^7$ Pa, $T_T = 5000$ K).

Unfortunately, the possibility of graphite-to-carbyne phase transition during graphite heating was not considered in [6]. Therefore, it is not quite clear to which solid and liquid phases the triple point found corresponds. One can assume that the above-mentioned considerable spread in the triple point parameters is associated with a change in the phase composition. The change in the phase composition may also be the reason for the dependence of the melting point of “graphite” on the heating time of the specimen, which is observed in experiments carried out by various researchers.

From the phase diagram of carbon (Fig. 7) suggested by Whittaker [7] and complemented by the triple point parameters obtained in [6], it follows that at temperatures above 2600 K graphite transforms into carbyne; hence, no graphite liquid can exist. The diagram also shows that carbyne is stable between 2600 and 3800 K in a wide pressure range up to the diamond formation pressure. The triple point parameters for carbyne somewhat differ from those reported above: $p_T = 2 \times 10^4$ Pa and $T_T = 3800$ K. Hence, at temperatures above 3800 K, carbyne passes to the liquid state. It was reported [7] that carbyne liquid is transparent and colorless, as well as has a low emissivity [7]. This correlates with the comment [8] that films of molten carbon were “sufficiently transparent.” One can therefore conclude that the triple point parameters obtained in [6] may also refer to carbyne. Thus, carbyne, not graphite, liquid must exist! Similarly, at still higher pressures, one might expect the existence of diamond liquid at the

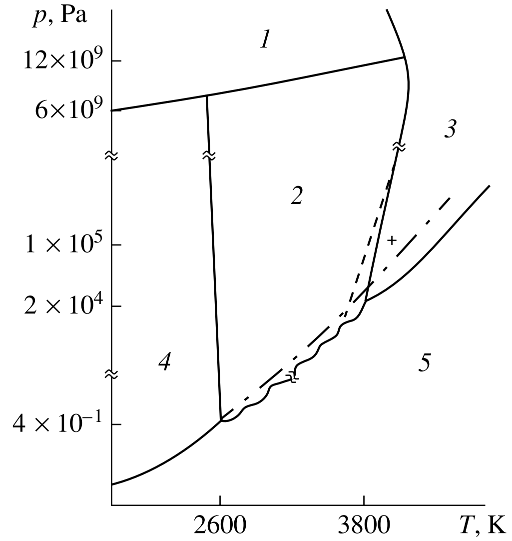


Fig. 7. *P–T* phase diagram of carbon suggested by Whittaker and complemented by the triple point (+) found in [6]. (*1*) Diamond; (*2*) carbyne; (*3*) liquid; (*4*) graphite; and (*5*) vapor.

boundary of the diamond region! Of course, this reasoning is valid if the Whittaker phase diagram is true, rather than purely hypothetical, and the carbyne region does exist.

It should be noted that the existence of carbyne is presently virtually beyond question. It has been found in the natural form and is produced as thin films in the laboratory conditions. From the phase diagram given, it follows that carbyne can be obtained from carbon if one, first, reaches the region where carbyne is thermodynamically stable and, second, retains it during cooling to the normal conditions. In other words, it is necessary that quenching be fast, since, if the temperature is reduced smoothly, reverse processes take place and carbyne readily decomposes, turning to carbon. The ratio of the forward and back reaction rates (i.e., the rates of carbyne formation and decomposition) is 1 : 500 [7]. For this reason, carbyne has been obtained only as thin films to date. As for the molecular structure of carbyne, we note that the dissociation of graphite at temperatures above 2600 K implies the breaking of the single bonds and the production of triple-bond chains in the form of long molecules. These chains may variously combine to form the hexagonal structure of one of the carbyne's modification.

As a result of the triple bond formation, atoms in the chain move closer together, so that the carbyne density far exceeds that of graphite. The density of various carbyne modifications varies between 2.68 and 3.43 g/cm^3, while that of graphite is no more than 2.25 g/cm^3. The formation of triple bonds indicates that liquid carbyne must consist of polyine molecules of type $(–C{\equiv}C–)_p$, which are stable at high temperatures.

Thus, based on the body of the data reported, we can argue that when the cast iron surface is subjected to laser processing, the graphite of the surface layer heats up, takes the globular shape, and turns to carbyne (or carbyne liquid if the temperature is sufficiently high). During the subsequent fast free cooling of the surface layer, the carbyne decomposes, turning to opaque pyrocarbon.

In summary, we stress that the above effects are of a general character. The laser processing of various metal–alloy systems provides extreme experimental conditions and offers considerable scope for the control of phase transformations. Researchers are given an opportunity to modify the surface and improve its physical and mechanical parameters, as well as develop materials with unique properties.

ACKNOWLEDGMENTS

The author is indebted to Prof. V.T. Bublik and T.B. Sagalova for the fruitful discussions and assistance.

This work was partially supported by the Russian Foundation for Basic Research (grant no. 00-01-00212).

REFERENCES

1. G. I. Kozlov, Pis'ma Zh. Tekh. Fiz. **25** (24), 61 (1999) [Tech. Phys. Lett. **25**, 997 (1999)].
2. G. I. Kozlov, Pis'ma Zh. Tekh. Fiz. **26** (11), 84 (2000) [Tech. Phys. Lett. **26**, 490 (2000)].
3. G. I. Kozlov and V. A. Kuznetsov, Kvantovaya Élektron. (Moscow) **16**, 1360 (1989).
4. A. G. Grigor'yants and A. N. Safonov, *Principles of Laser Thermostrengthening of Alloys* (Vysshaya Shkola, Moscow, 1988).
5. K. J. Huttinger, Chem. Vap. Deposition **4**, 151 (1998).
6. É. I. Asinovskiĭ, A. V. Kirillin, and A. V. Kostanovskiĭ, Teplofiz. Vys. Temp. **35**, 716 (1997).
7. A. G. Whittaker, Nature **276**, 695 (1978); Science **200**, 763 (1978).
8. É. I. Asinovskiĭ, A. V. Kirillin, A. V. Kostanovskiĭ, and V. E. Fortov, Teplofiz. Vys. Temp. **36**, 740 (1998).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 4, 2002, pp. 459–464. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 88–93.
Original Russian Text Copyright © 2002 by Popok, Azarko, Khaibullin.

The Effect of High Implant Doses and High Ion Current Densities on Polyimide Film Properties

V. N. Popok*, I. I. Azarko*, and R. I. Khaibullin**
** Belarussian State University, Minsk, 220050 Belarus*
e-mail: popok@bsu.by
*** Kazan Physicotechnical Institute, Russian Academy of Sciences, Kazan, 420029 Russia*
e-mail: rik@kfti.knc.ru

Received September 17, 2001

Abstract—Ar^+ and Ar^{2+} ions with energies of 40 and 80 keV are implanted into thin polyimide films. The implant doses and the ion current densities are varied in a wide range between 2.5×10^{14} and 1.5×10^{17} cm^{-2} and between 1 and 16 μA/cm^2, respectively. The effect of the implantation parameters on the electrical, paramagnetic, and optical properties of the ion-modified near-surface polymer layer is studied. It is shown that the radiation-stimulated thermolysis of polyimide and its chemical constitution are responsible for a monotonic growth of the electrical conductivity of the layer with increasing ion current at a given implant dose. When the ion current density is fixed, the conductivity grows stepwise with implant dose, whereas the concentration of paramagnetic centers and the optical transmission of the modified layer decrease. The dependences observed are treated within a model of the structural reconfiguration of the polymer carbonized phase formed during the implantation. *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

In recent years, conductive polymers have become a popular micro- and optoelectronic material in both scientific research and applications [1]. The reason is the advances in the synthesis of new conductive polymers and, hence, their extended potentialities in these areas of technology. Polymer films with the metallic type of conduction have found application as corrosion inhibitors, compact capacitors, and antistatic coatings for photographic films and monitor screens. Moreover, the methods of synthesizing polyconjugated systems (which have evolved from the early 1970s) have made it possible to produce polymers of *n*- or *p*-conductivity by doping. These materials are used in fabricating organic electronic components, such as light-emitting diodes, transistors, solar cells, storage batteries, etc. [1, 2].

A promising way for producing conductive polymers is ion implantation [3–5]. As a result of irradiation, a nanostructured system of carbon atoms with conjugated bonds is formed in the implanted layer of a polymer [6]. The conductivity of the implanted layer monotonically rises with dose and may vary within 10 to 15 orders of magnitude [3, 7, 8]. The use of high implantation doses to produce high-conductivity polymer layers implies that the process is carried out at increased ion current densities, which cuts the irradiation time and improves the process efficiency. However, a high ion current inevitably means a high power being released in the layer irradiated; therefore, this parameter is critical for organic materials because of their poor radiation hardness and thermal stability. On the other hand, an increase in the ion current density at a fixed dose raises the conductivity of the implanted polymer layer [9]. Thus, the optimization of the ion implantation conditions for the formation of conductive layers with desired electrical properties seems to be a topical problem in ion-beam modification of polymers.

In this work, we study the effect of argon ion implantation on the electrical, paramagnetic, and optical properties of polyimide. Polyimide is a representative of the most thermally stable polymers and allows the ion current density to be varied in wide limits up to 16 μA/cm^2. To date, such a high value of the ion current density, as applied to implantation into polymers, has not been reported in the literature. The use of argon ions eliminates the doping effect of impurities and makes it possible to study the purely radiation effect of ions on the polymer. In addition, argon implantation into polyimide under other process conditions (ion beam parameters) has been extensively studied [9–11], which allows us to comparatively analyze surface modifications.

EXPERIMENTAL

The object of investigation was 40-μm-thick polyimide films of density 1.43 g/cm^3 (Fig. 1a). The advantage of polyimide is that it does not exhibit the transition to the viscoelastic state with increasing temperature. The softening temperature depends on the concentration of imide groups (degree of imidization) *s*: at $s = 1$, this temperature exceeds 500 K. The glass

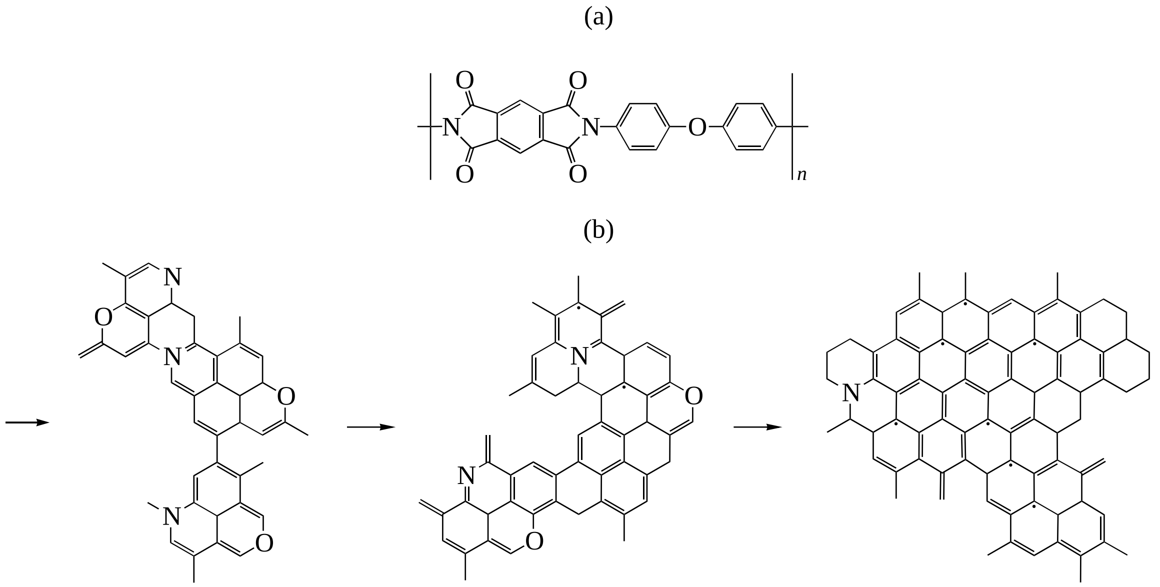


Fig. 1. (a) Chemical formula of a polyimide elementary unit and (b) polymer structure transformation upon implantation.

transition temperature is 590 K [12], so that polyimide can be considered as fairly stable against radiation-induced thermal destruction. The implantation of Ar^+ and Ar^{2+} ions was carried out with an ILU-3 ion-beam accelerator at room temperature at a residual pressure of 10^{-5} torr. The process parameters were the following: the energy 40 and 80 keV, respectively; doses $D = 2.5 \times 10^{14}$–1.5×10^{17} cm^{-2}; and the ion current density $j = 1$–16 μA/cm^2. To avoid sample breakdown and overheating during the implantation, the samples were placed into a special cassette that provided efficient ion charge drainage from the polymer and tight contact between the films and the metal substrate, which was cooled by running water. The films implanted were visually free of any serious thermal damages (charring, surface corrugation, etc.) under all the implantation conditions. However, at some critical value of the ion current density (18 μA/cm^2 at a dose of 1.0×10^{16} cm^{-2} in our case), local fused spots and blisters were detected on the sample surface.

Electron paramagnetic resonance (EPR) studies were carried out with a Varian E112 spectrometer in the X band (9.3 GHz) at room temperature. The concentration N_{PC} of paramagnetic centers in the implanted layer was calculated by the conventional technique [13] with regard for the thickness of the radiation-damaged layer. According to the calculations by the SRIM-2000 computer program [14] and to Rutherford backscattering data [15], the thickness of the polymer layer modified is 80 ± 10 and 150 ± 10 nm for 40- and 80-keV argon ions, respectively. The sheet resistivity of the polymer samples modified was measured by the standard two-point probe technique [8] with an E7-14 emittance meter. The linear contacts were made of fine-grained silver paste applied on the sample surface. Optical transmission spectra from the implanted films were recorded in the visible range with a Hitachi-330 spectrophotometer at room temperature.

ION-BEAM MODIFICATION OF POLYMER STRUCTURE

To better understand results obtained in this work, we will first briefly consider the concept of polymer structure modification that has been generally accepted to date [4, 6–8]. Because of the electronic and nuclear stopping of the ions, the molecular structure of the polymer is damaged along the trajectories of the ions. Note that unlike inorganic materials, the breakage of chemical bonds in polymers takes place not only because of nuclear collisions but also the electronic excitation [16]. The effect of electron stopping on the polymer structure is most pronounced when the ions are implanted in compounds with heteroatom-containing functional groups (polyimide in our case). The destruction of the electron shells of atoms and the breakage of chemical bonds are selective processes, aimed primarily at breaking weaker bridge bonds between macromolecules and at transforming functional chemical groups with a high degree of electron delocalization. However, the core of the polymer damaged region (latent track), which has the highest concentration of radiation-induced defects, forms largely as a result of collisions with nuclei. For the doses used in this work, the latent tracks overlap [4]. This, as well as developed cascades of the electronic excitation and ionization of a polymer (so-called penumbras [17]), causes the considerable fragmentation of polymer chains, followed by radiation-induced cross-linking.

During the stopping of the ions in polymers, the layer irradiated loses highly volatile components and the latent tracks serve as channels through which gases diffuse to the polymer surface. As a result of outgassing, the near-surface layer becomes enriched by carbon atoms and carbon clusters form. The nucleation of carbon nanoclusters in polymers upon implantation has been confirmed by neutron scattering [18], electron microscopy [19], optical spectroscopy [20], and EPR studies [21]. Today, this fact is considered to be commonly accepted. As the implantation dose increases, the clusters usually grow, interlink, and coagulate up to the formation of a quasi-continuous carbonized layer [22]. The type of conduction in this layer may vary from semiconducting to metallic, depending on the irradiation conditions.

ELECTRICAL, OPTICAL, AND PARAMAGNETIC PROPERTIES OF POLYIMIDE VS. ION CURRENT DENSITY

Figures 2–4 show the electrical, optical, and paramagnetic properties of the polyimide layer irradiated as functions of the ion current density *j* at a fixed dose of 1.0×10^{16} cm^{-2} of singly charged argon ions Ar^{+}. It follows from Figs. 2 and 3 that, as *j* grows, the resistivity ρ linearly decreases and the optical transmission of the modified layer monotonically declines. A rise in the conductivity with growing *j* is generally consistent with data obtained in [9]. However, it is impossible to compare the absolute resistivity values, since in [9] they are given in units Ω/□. In all polyimide samples studied, the singlet isotropic EPR line with a *g* factor of 2.0025 ± 0.0005, which is close to the *g* factor of a free electron, is observed. This EPR signal is related to unpaired π electrons, entering into the electron system of the carbon clusters, which are produced during the implantation [21]. Figure 4 demonstrates the effect of the ion current density on the EPR signal parameters. It is seen that N_{PC} monotonically drops, while the linewidth ΔH increases with growing *j*.

Thus, as the ion current density increases at a fixed dose, the conductivity grows, the transmission and the concentration of paramagnetic centers decrease, and the EPR line broadens. This implies the structure modification of the polyimide carbonized phase under the conditions of a growing energy effect. As the ion current grows, the composition and the structure of the surface layer are affected not only by radiation-induced phenomena (fractionation, cross-linking, atom displacements, etc.) but also by the local heating of the polymer near the latent track due to the transfer of the ion kinetic energy to atoms of the polymer matrix (thermal spike effect) [23]. As a result of the implantation, unified radiation-induced thermolysis takes place. The characteristic time of this process after the ion–target collision is 10^{-12}–10^{-10} s, and the space around the track heats up to temperatures above 1000 K [24]. It is known that, upon implantation into polyimide, ether links

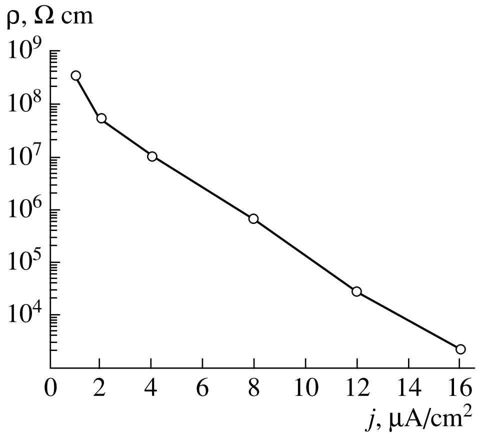


Fig. 2. Resistivity vs. ion current density *j* for the samples irradiated by 40-keV Ar^{+} ions with a constant dose of 1×10^{16} cm^{-2}.

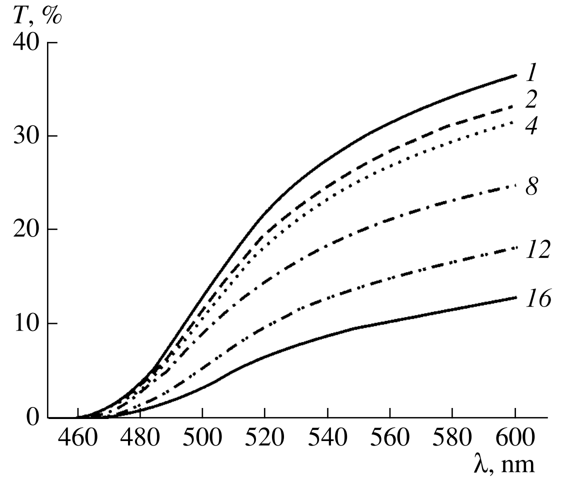


Fig. 3. Optical transmission curves for the same experimental conditions as in Fig. 2. Figures at the curves show the ion current density values in μA/cm^2.

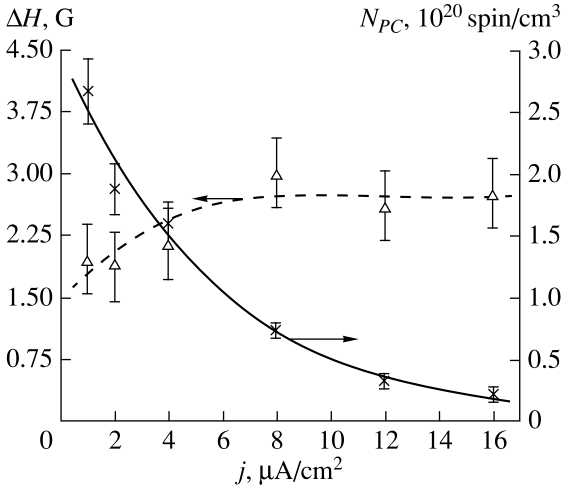


Fig. 4. EPR linewidth and concentration of paramagnetic centers vs. ion current density. The experimental conditions are the same as in Fig. 2.

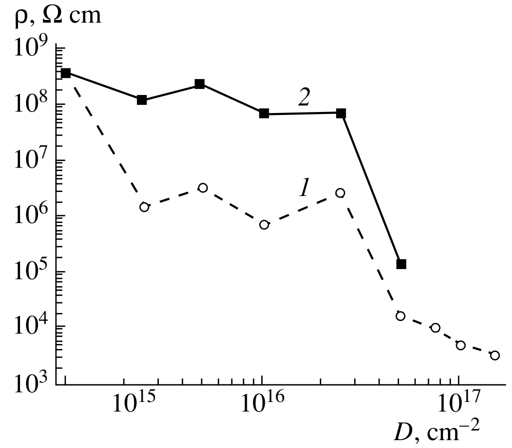


Fig. 5. Dose dependences of the resistivity for the samples irradiated by (*1*) singly charged (E = 40 keV, j = 8 μA/cm^2) and (*2*) doubly charged (E = 80 keV, j = 4 μA/cm^2) argon ions.

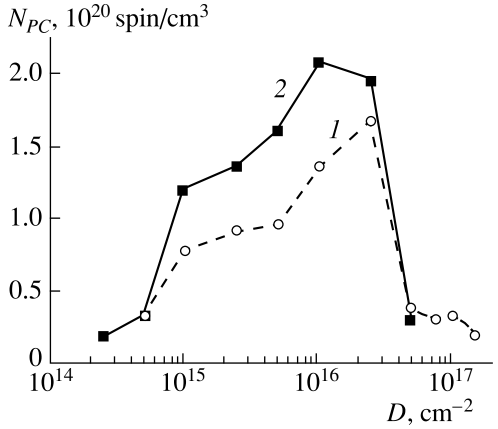


Fig. 6. Dose dependences of the paramagnetic center concentration for the same experimental conditions as in Fig. 5.

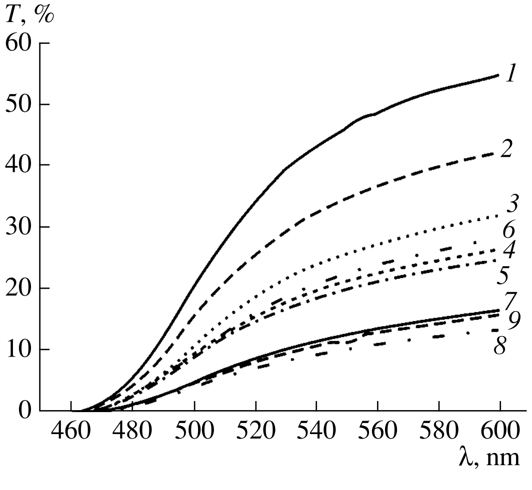


Fig. 7. Optical transmission curves for the samples irradiated by Ar$^+$ ions (40 keV, j = 8 μA/cm^2). The dose D = (*1*) 5 × 10^{14}, (*2*) 1 × 10^{15}, (*3*) 2.5 × 10^{15}, (*4*) 5 × 10^{15}, (*5*) 1 × 10^{16}, (*6*) 2.5 × 10^{16}, (*7*) 5 × 10^{16}, (*8*) 7.5 × 10^{16}, and (*9*) 1.5 × 10^{17} cm^{-2}.

between aromatic rings degrade first, which is accompanied by the release of hydrogen molecules and carbon oxide [8, 25]. During polymer thermolysis, imide groups may convert to imine and amide groups [11]. As the ion current density increases, so does the power dissipated in the polymer. Therefore, it is logical to assume that extended polycondensed aromatic structures may form (Fig. 1b) by analogy with the structure transformations in polyimide under pyrolysis (1000–1200 K) [26]. The growth of the carbonized phase with increasing j, which is accompanied by the development of systems of conjugated links, is responsible for the growth of the conductivity of the layer irradiated, as well as for the decrease in its optical transmission and in the number of paramagnetic centers.

POLYIMIDE PARAMETERS VS. IMPLANT DOSE

Figures 5–8 show that the experimentally found dependences of the electrical and paramagnetic parameters of the polyimide on the implant dose at a constant ion beam power $P = Ej$ = 0.32 W for the singly charged (E = 40 keV, j = 8 μA/cm^2) and doubly charged (E = 80 keV, j = 4 μA/cm^2) argon ions are nonmonotonic. For doses between 5.0 × 10^{14}–2.5 × 10^{16} cm^{-2}, we observe a slight decline in ρ (Fig. 5); a rise in N_{PC} (Fig. 6); an almost constant ΔH (2.5–3.0 G for Ar$^+$ and 1.5–2.0 G for Ar^{2+}); and a decrease in the optical transmission, which tends to saturation (Fig. 7). For doses above 2.5 × 10^{16} cm^{-2}, the conductivity rises stepwise for the samples irradiated by both Ar$^+$ and Ar^{2+} ions. This rise is accompanied by a drastic drop of the optical transmission, a decrease in the concentration of paramagnetic centers (Figs. 5–7), and a broadening of the EPR line (to 4.0 G for Ar$^+$ and 2.5 G for Ar^{2+} ions). Similar dose dependences of the EPR spectrum parameters were observed in polyimide irradiated by argon ions that had energies close to those mentioned above (40–90 keV) [10].

These experimental dependences of the polyimide parameters on the implantation dose are easy to understand within the above concept of implantation-induced structure modification of polymers. Namely, the increase in the concentration of paramagnetic centers with dose in the interval 5.0 × 10^{14}–2.5 × 10^{16} cm^{-2} (Fig. 6) indicates the growth of the carbon clusters in the polymer irradiated. The constancy of the EPR linewidth in this interval implies the absence of qualitative changes in the electron subsystem of the carbonized phase. The evidence for this fact is that ρ for these doses varies only slightly (Fig. 5). The reduction of the optical transmission with increasing implant dose also implies the growth of the carbon clusters, which are basic absorbing centers in the polymer irradiated. The tendency of the transmission to saturate (curves *4*–*6* in Fig. 7) indicates that the processes of radiation-induced cross-linking and carbon cluster agglomeration are coming to an end.

As has been mentioned, the resistivity, as well as the paramagnetic and optical parameters of the samples, vary stepwise at doses $>2.5 \times 10^{16}$ cm^{-2}. This can be explained by the transition of the carbonized phase, which is characterized by the presence of extended carbon cluster agglomerates and conjugated macrorings in the polymer-like structure, to the phase of amorphous carbon or graphite-like material with individual polymer chain fragments. Such a conclusion correlates well with the results obtained in [10], where it has been concluded from the temperature dependence of the conductivity that the conduction type in Ar^+- and N^+-irradiated polyimide changes from semiconducting to metallic at doses exceeding 3×10^{16} cm^{-2}, that is, when the phase change is expected. A similar percolation transition has been also observed in other polymers, for example, after the low-current (0.1 μA/cm^2) implantation of light ions with doses higher than 5×10^{16} cm^{-2} [22]. Therefore, we can conclude that the phase transitions observed are dose-related, rather than current-related, processes.

Thus, the variations of the EPR spectrum parameters and the resistivity correlate, leading us to conclusion that the electrical and paramagnetic properties of ion-irradiated organic semiconductors are closely related.

However, the electrical, paramagnetic, and optical parameters of the samples irradiated by singly charged argon ions differ from the parameters observed on the samples irradiated by doubly charged ions (Figs. 5, 6, 8). One reason may be that the effects of electronic deceleration on polymer structure modification differ for ions with differing charge and energy. Another reason is the enhanced effect of radiation-stimulated thermolysis. As was noted, the effect of electron stopping on the molecular structure of polyimide is considerable because of weak ether links between aromatic rings [8]. When being decelerated, 80-keV Ar^{2+} ions cause the significant radiation-induced ionization of the polymer, which, in turn, produces much more radical changes in its structure than in the case of the singly charged ions. This is reflected by higher values of N_{PC} (Fig. 6) and lower values of the optical transmission (Fig. 8). The conductivity also drops (Fig. 5) because of the increase in the concentration of radiation-induced defects. It is not improbable, however, that the decrease in the optical transmission for the case of Ar^{2+} ions is due to the thickening of the modified polyimide layer because of an increase in the projected ion range (see the section EXPERIMENTAL). When Ar^+ ions are implanted into polymers, a higher value of j favors thermolysis effects, which, as was noted, lead to the restructuring of the carbonized phase; consequently, the concentration of paramagnetic centers decreases and the conductivity of the layer modified grows in comparison with the irradiation by Ar^{2+} ions.

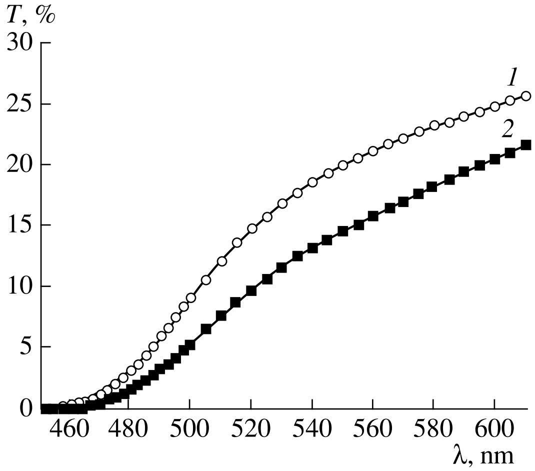


Fig. 8. Transmission curves for polyimide irradiated as in Figs. 5 and 6. The dose is 1×10^{16} cm^{-2}.

CONCLUSION

In this article, we report a complex study on the electrophysical parameters of polyimide films as functions of argon ion implantation conditions. It is shown that, as the ion current density increases from 1 to 16 μA/cm^2 at a high fixed dose (1.0×10^{16} cm^{-2}), the conductivity of polyimide grows by five orders of magnitude without signs of radiation- or temperature-induced destruction of the polymer. This allows us to conclude that high-current implantation is an efficient practical means of controlling the conductivity of this polymer. The variation of the electrical parameters of polyimide irradiated correlates with that of the paramagnetic and optical characteristics. Based on this fact, the idea of polymer structure modification and carbonized phase formation under the action of high-power ion beams and due to the polyimide specific chemical structure is put forward. It is found that the electrophysical parameters of the polymer films depend on the dose stepwise when it exceeds a certain value (~2.5×10^{16} cm^{-2} in our case). Such behavior finds an explanation within our model of carbonized phase restructuring. It is shown that the change in the ion charge at the constant ion beam power causes only a quantitative discrepancy between the polymer parameters, because they produce different effects of ionization and radiation-induced thermolysis; qualitatively, however, the dose dependences of the polyimide parameters remain the same.

ACKNOWLEDGMENTS

The authors are indebted to I.A. Karpovich for the assistance in the electrical measurements.

R.I. Khaibullin thanks the Russian Foundation for Basic Research (grant no. 99-03-32548) and the Leading Scientific Russian Schools Organization (grant no. 00-15-96615) for the financial support.

REFERENCES

1. The 2000 Nobel Prize in Chemistry (for the Discovery and Development of Conductive Polymers), Press Release, www.nobel.se/chemistry/laureates/2000/press.html.
2. C. K. Chiang, C. R. Fisher, Y. W. Park, *et al.*, Phys. Rev. Lett. **39**, 1098 (1977).
3. T. Venkatesan, L. Calcagno, B. S. Elman, *et al.*, in *Ion Beam Modification of Insulators*, Ed. by P. Mazzoldi and G. W. Arnold (Elsevier, Amsterdam, 1987), pp. 301–379.
4. G. Marletta and F. Iacona, in *Materials and Processes for Surface and Interface Engineering*, Ed. by Y. Pauleau (Kluwer, Dordrecht, 1995), pp. 597–640.
5. R. E. Gied, M. G. Moss, J. Kaufmann, *et al.*, in *Electrical and Optical Polymer Systems*, Ed. by D. L. Wise, G. E. Wnek, D. J. Trantolo, *et al.* (Marcel Dekker, New York, 1998), pp. 1011–1030.
6. V. N. Popok, Poverkhnost, No. 6, 103 (1998); Surf. Investigation **14**, 843 (1999).
7. E. H. Lee, Nucl. Instrum. Methods Phys. Res. B **151**, 29 (1999).
8. V. B. Odzhaev, I. P. Kozlov, V. N. Popok, *et al.*, *Ion Implantation of Polymers* (Belaruss. Gos. Univ., Minsk, 1998).
9. M. Iwaki, Nucl. Instrum. Methods Phys. Res. B **175–177**, 368 (2001).
10. A. N. Aleshin, A. V. Gribanov, A. V. Dobrodumov, *et al.*, Fiz. Tverd. Tela (Leningrad) **31** (1), 12 (1989) [Sov. Phys. Solid State **31**, 6 (1989)].
11. A. De Bonis, A. Bearzotti, and G. Marletta, Nucl. Instrum. Methods Phys. Res. B **151**, 101 (1999).
12. C. Feger, M. M. Khojasteh, and J. E. McGrath, *Polyimides, Chemistry and Characterization* (Elsevier, Amsterdam, 1989).
13. J. E. Wertz and J. R. Bolton, *Electronic Spin Resonance: Elementary Theory and Practical Applications* (McGraw-Hill, New York, 1972; Mir, Moscow, 1975).
14. J. F. Ziegler, in *The Sopping and Range of Ions in Matter (SRIM-2000)*, www.research.ibm.com/ionbeams.
15. V. N. Popok, I. I. Azarko, R. I. Khaibullin, *et al.*, Appl. Phys. A (in press).
16. J. Davenas, X. L. Xu, G. Boiteux, *et al.*, Nucl. Instrum. Methods Phys. Res. B **39**, 754 (1989).
17. J. L. Magee and A. Chattejee, J. Phys. Chem. **84**, 3529 (1980).
18. D. Fink, K. Ibel, P. Goppelt, *et al.*, Nucl. Instrum. Methods Phys. Res. B **46**, 342 (1990).
19. G. R. Rao, Z. L. Wang, and E. H. Lee, J. Mater. Res. **8**, 927 (1993).
20. I. P. Kozlov, V. B. Odzhaev, I. A. Karpovich, *et al.*, Zh. Prikl. Spektrosk. **65**, 377 (1998); J. Appl. Spectr. **65**, 390 (1998).
21. I. P. Kozlov, V. B. Odzhaev, V. N. Popok, *et al.*, Zh. Prikl. Spektrosk. **65**, 562 (1998); J. Appl. Spectr. **65**, 583 (1998).
22. V. N. Popok, I. A. Karpovich, V. B. Odzhaev, *et al.*, Nucl. Instrum. Methods Phys. Res. B **48**, 1106 (1999).
23. A. Miotello and R. Kelly, Nucl. Instrum. Methods Phys. Res. B **122**, 458 (1997).
24. H. De Cicco, G. Saint-Martin, M. Alurralde, *et al.*, Nucl. Instrum. Methods Phys. Res. B **173**, 455 (2001).
25. V. Švorčik, R. Endršt, V. Rybka, *et al.*, Eur. Polym. J. **31**, 189 (1995).
26. C. Z. Hu, K. L. De Vris, and J. D. Andrade, Polymer **28**, 663 (1987).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 4, 2002, pp. 465–469. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 94–98.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

The Operation of High-Speed Emergency Protection Based on Uranium Hexafluoride

Zh. S. Takibaev*, S. A. Spitsyna, G. K. Potrebenikov**,**
A. V. Yushkov, and K. A. Kopzhasarova****
** Kazakh National Nuclear Center, Almaty, 480082 Kazakhstan*
e-mail: takibaev@ietp.alma-ata.su
*** Al-Farabi State University of Kazakhstan, ul. Vinogradova 95, Almaty, 480012 Kazakhstan*

Received March 17, 1999; in final form, April 12, 2001

Abstract—The dynamic properties of a gas–liquid one-time direct emergency protection device used in nuclear power plants are studied using computer experiments within a model suggested. In the case of reactivity emergency, an absorbent is introduced into the reactor core automatically without human intervention and any external control. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

In modern reactor designs, the internal safety of a nuclear reactor is the issue of fundamental importance. This concept of internal safety means that, using negative feedbacks, a reactor can be stabilized or shut down if there is a real heavy accident hazard. Direct emergency protection (DEP) functioning in the reactor core unattended and responding to a deviation of any controllable parameter from a preset value is an example. Such devices must feature a high speed, i.e., quickly sense an impermissible change in a controllable parameter and quickly perform a safety action, i.e., to introduce an efficient neutron absorbent.

In a number of works [1–4], it has been suggested to improve the DEP speed by introducing compounds having a large thermal-neutron fission cross section into the sensitive elements. Such DEP devices sense the primary sign of reactivity emergency (increased neutron flux density), and the resulting temperature effects are used by the control means of the safety system to provide the delivery of a neutron adsorbent to the reactor core.

In this work, we consider a version of gas–liquid DEP built around a one-time element. The sensitive medium here is a gas containing a fissile component. The actuator responds when the pressure in the volume exceeds a preset value and provides the delivery of a liquid neutron absorbent into the working volume of the DEP element located in the reactor core.

GAS–LIQUID DEP ELEMENT

For the DEP element to be placed in the reactor core, its shape and sizes must be compatible with those of standard reactor components: fuel element, rod guide tube in the control and protection system (CPS), or fuel assembly.

The actuator is a simple pneumatic relay, which responds when the pressure in the sensing element exceeds a preset value. When introduced, the absorbent is kept outside the core. If the actuator comes into action, the metallic bellows moves upward, thereby allowing the absorbent to enter into the working volume.

The leak tightness and the accuracy of the relay [5, 6] specify the operating pressure, which varies between 0.12 and 0.18 MPa in automated pneumatic devices [7]. The sensitivity threshold of pneumatic relays is usually 20 to 40% lower. Let the sensitivity threshold pth for the relay of the DEP element actuator be 10^5 Pa. The bellows (in the pneumatic relay), which activates the control valve and works in tension, may be replaced by a piston. However, the tight joint of the bellows to the sensor body better prevents fissioner leakage.

It has been suggested [1, 2] to use gaseous UF_6 as a sensitive medium in DEP devices. At typical intrareactor temperatures, this compound can be considered as perfect gas [8]. In addition, its thermophysical properties are well understood [9–11].

When uranium in UF_6 is enriched by the ^{235}U isotope, the volume concentration of fissile nuclei in the sensing element is given by

$$\gamma_5 = \frac{N_A \rho_U \chi}{\mu_U}, \tag{1}$$

where N_A is the Avogadro number, ρ_U is the UF_6 density, μ_U is the UF_6 molecular mass, and χ is the enrichment (hereafter, the subscript 5 refers to ^{235}U and the subscript U, to UF_6; the molecular mass is calculated without allowance for enrichment).

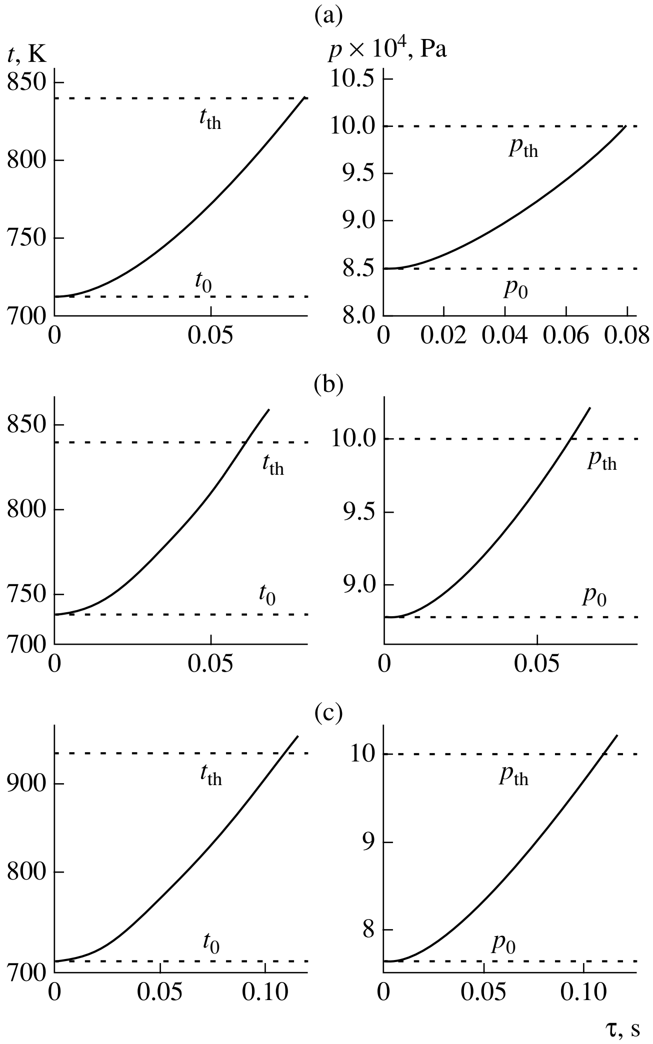


Fig. 1. Mean temperature and pressure in the sensing medium (element) at $\rho = 0.9\beta$ and $\tau = 0$. (a) $\rho_U = 5000$ kg/m^3 and $\chi = 0.045$, (b) $\rho_U = 5000$ kg/m^3 and $\chi = 0.055$, and (c) $\rho_U = 4500$ kg/m^3 and $\chi = 0.05$.

The sensing element (medium) can be viewed as a volume source with an energy release volume density

$$q_v = E_f \Sigma_f \Phi, \qquad (2)$$

where E_f is the energy release per fission event, $\Phi = \Phi_0 n$, n is the relative neutron flux density, Φ_0 is the neutron flux density in the steady state, $\Sigma_f = \gamma_5 \sigma_f$, and Σ_f and σ_f are the respective thermal-neutron macro- and microscopic fission cross sections for ^{235}U nuclei.

The quasi-stationary heat conduction equation for an infinite cylinder is given by [12]

$$\frac{dt}{d\tau} = \frac{\Theta - t}{\tau_0} + an, \qquad (3)$$

where t is the mean temperature of the fuel element, Θ is the coolant temperature, $a = q_v/\rho' C$, and $\tau_0 = \xi RC\rho'/2$.

For a rod-type sheath-free fuel element of radius R, density ρ', and specific heat capacity C, the thermal resistance ξ is related to the thermal conductivity λ of the fuel kernel material as $\xi = R/4\lambda$. For a sheathed fuel element, $\xi = R/4\lambda + \delta/\lambda_1$, where δ is the sheath thickness and λ_1 is the thermal conductivity. This relationship holds if heat transfer from the sheath to the coolant is perfect. Usually, $\lambda_1 \gg \lambda$, so that the second term in the expression for ξ is neglected in the estimates. Two geometric parameters (r, the sensing element radius; R_1, outer radius of the DEP element sheath) are used in the dynamics simulation: $a = q_v/\rho_U C_U$, $\tau_0 = \xi R_1 C_U \rho_U/2$, and $\xi = r/4\lambda_U$. In the steady state ($\tau = 0$), the mean temperature of the sensing element is $t_0 = \Theta + a\tau_0$. The corresponding pressure p_0 in the volume of the sensing element should not exceed the pressure p_{th}, at which the DEP actuator comes into action:

$$p_0 = \frac{\rho_U R_m}{\mu_U} t_0 < p_{th}, \qquad (4)$$

where R_m is the gas constant.

Since p_{th} is given in specifications for actuators, the range of the UF$_6$ density is bounded from above by the inequality $\rho_U < \mu_U p_{th}/R_m\Theta$. At $p_{th} = 10^5$ Pa and $\Theta = 579$ K, $\rho_U < 7250$ kg/m^3. The DEP element operates at the temperature $t_{th} = \mu_U p_{th}/\rho_U R_m$, which must be attained at the threshold neutron flux density n_{th}. The temperature growth lags behind the growth of the neutron flux. The thermal delay of the DEP element depends on its time constant τ_0, which, in turn, depends on the UF$_6$ density in the sensing medium. For a given $\rho_U p_{th} = p_0 t_{th}/t_0$.

Let the DEP element be placed into the rod guide tube of a 1000 VVÉR water-moderated water-cooled power reactor. Then, for $n_{th} = 1.4$ at $E_f = 3.14 \times 10^{-11}$ J [13], $\sigma_f = 580.2 \times 10^{-28}$ m^2. In this case, the mean neutron flux density is $\Phi_0 = 10^{18}$ s^{-1} m^{-2}, the coolant temperature in the core of the 1000 VVÉR reactor is $\Theta = 579$ K [14], $r = 0.004$ m, and $R_1 = 0.0063$ m. If the temperature dependences of the parameters a and τ_0 are taken into account upon solving Eq. (3), it can be solved only numerically.

With a positive reactivity introduced, the DEP element behavior can be described by solving heat conduction equation (3) jointly with the set of equations of point kinetics [12]

$$\frac{dn}{d\tau} = \frac{\rho - \beta}{l} n + \Sigma \lambda_i C_i, \quad n(\tau = 0) = 1,$$

$$\frac{dC_i}{d\tau} = \frac{\beta_i}{l} n - \lambda_i C_i, \quad C_i(\tau = 0) = \frac{\beta_i}{l\lambda_i}, \quad i = 1, \dots, 6, \quad (5)$$

$$\frac{dt}{d\tau} = \frac{\Theta - t}{\tau_0} + an, \quad t(\tau = 0) = t_0,$$

where t_0 is a solution of the equation $t - \Theta = a(t)\tau_0(t)$.

For the positive reactivity $\rho = 0.9\beta$, the solutions of set (5) at the initial time instant ($\tau = 0$) are depicted in Fig. 1a, where the mean temperature and pressure in the sensing medium are shown for $\rho_U = 5000$ kg/m^3 and $\chi = 0.45$. The dashed lines parallel to the abscissa axis indicate the values corresponding to the steady-state conditions and to the instant of actuator operation. The time delay of the DEP element actuator is several tens of milliseconds, starting from the instant of reactor runaway. Note that this time for the automated emergency protection of the reactor may be as high as several hundred of milliseconds [15]. To further cut the DEP element time delay, one can increase the enrichment of UF_6 by the ^{235}U isotope. From Fig. 2b, it follows that this can be done if the temperature and the pressure in the sensing medium are elevated.

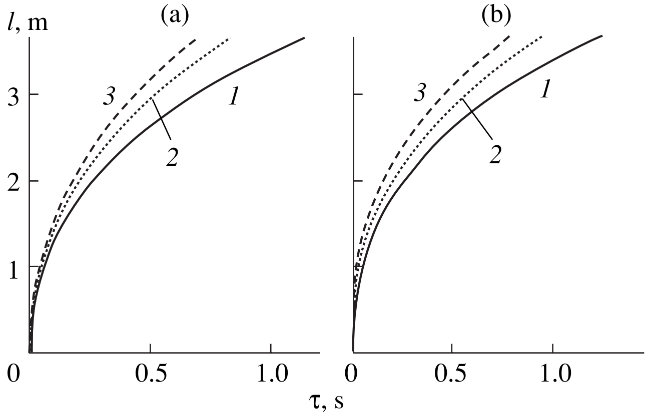


Fig. 2. DEP control member level in the reactor core after the actuator operation. (a) $P_0 = 1.5$ MPa; ε = (*1*) 0.1, (*2*) 0.15, and (*3*) 0.2. (b) $\varepsilon = 0.1$; P_0 = (*1*) 1.4, (*2*) 1.8, and (*3*) 2.2 MPa.

RESPONSE TIME OF DEP

Liquid neutron absorbents are usually used for eliminating excess reactivity, for example, in boron systems of VVÉR reactor control [16]. With boron dissolved in the moderator [17], one can control only small reactivity variations. It has been suggested [15] to apply CPSs where liquid absorbents circulate through hermetically sealed pipelines in special channels made in the core.

A strong absorbent of thermal neutrons is gadolinium ($\sigma_a = 4.6 \times 10^{-24}$ m^2), for example, in the form $Gd(NO_3)_3$ [18]. With a density $\rho_{Gd} > 3000$ kg/m^3, the macroscopic thermal-neutron absorption cross section of $Gd(NO_3)_3$ exceeds 250 cm^{-2}. Then, the DEP element working volume filled with this material can be considered as a black absorbing rod [19] introduced into the reactor core. In terms of hydraulics, a neutron absorbent container with an inner radius R can be represented as a circular channel with a hydraulic diameter $d_l = 2R$ and the working volume of the absorbent, as an annular channel with a hydraulic diameter $d = 2(R - r)$, where r is the inner radius of the annular channel. The absorbent flow is laminar and becomes steady in areas comparable in size to the hydraulic diameters. The friction drags for circular and annular channels of length l and h, respectively, are [20]

$$\Delta P_1 = \xi_1 \frac{l}{d_1}\left(\frac{\rho_{Gd} v_1^2}{2}\right), \quad \Delta P = \xi \frac{h}{d}\left(\frac{\rho_{Gd} v^2}{2}\right), \qquad (6)$$

where v_1 and v are the cross-section-averaged velocities of the steady flow in the circular and annular channels, respectively, and ξ_1 and ξ are the coefficients of hydraulic resistance for the circular and annular channels, respectively. The coefficients are given by

$$\xi_1 = 64/\mathrm{Re}_1,$$

$$\xi = 64F(\Theta)/\mathrm{Re}.$$

Here,

$$F(\Theta) = (1 - \Theta)^2/[1 + \Theta^2 + (1 - \Theta)^2/\ln\Theta]$$

$$\Theta = r/R,$$

and Re_1 and Re are the Reynolds numbers for the circular and annular channels, respectively [21]. Substituting ξ_1 and ξ into (6), we obtain

$$\Delta P_1 = 8\eta l v_1/R^2,$$

$$\Delta P = 8\eta l v F(\Theta)/[R^2(1 - \Theta^2)],$$

where η is the dynamic viscosity. The Poiseuille formulas for the circular and annular channels are obtained by expressing the velocities through the friction drags:

$$v_1 = R^2 \Delta P_1/8 v l,$$

$$v = R^2(1 - \Theta^2)\Delta P/8\eta h F(\Theta).$$

For a DEP element with $R = 0.0059$ m, the reactor core height is $H = 3.66$ m. Let $r = 0.004$ m, as before. At the coolant temperature, $\eta = 0.085$ Pa s and $\rho_{Gd} = 3430$ kg/m^3. The absorbent filling the working volume throughout the core depth fills the container to a level $L = (1 - \Theta^2)H$. It produces a hydrostatic head of $\rho g L$. This head of the absorbent in the container makes it difficult to provide a necessary rate of its delivery to the working volume of the DEP element. The rate of delivery can be increased if an extra pressure in the container is provided by a compressible gas.

Let P_0 be the initial gas pressure in the container. The gas layer thickness will be expressed through the core height: $L_0 - L = \varepsilon H$. Within a time τ after the actuator has come into action, a part of the absorbent enters from the container into the working volume, filling it to a level h; accordingly, its level in the container will decrease from the initial level L to a level $l = L - (1 - \Theta^2)h$. Then,

$$P = P_0(L_0 - L)/(L_0 - l) = P_0(L_0 - L)/[L_0 - L + (1 - \Theta^2)h].$$

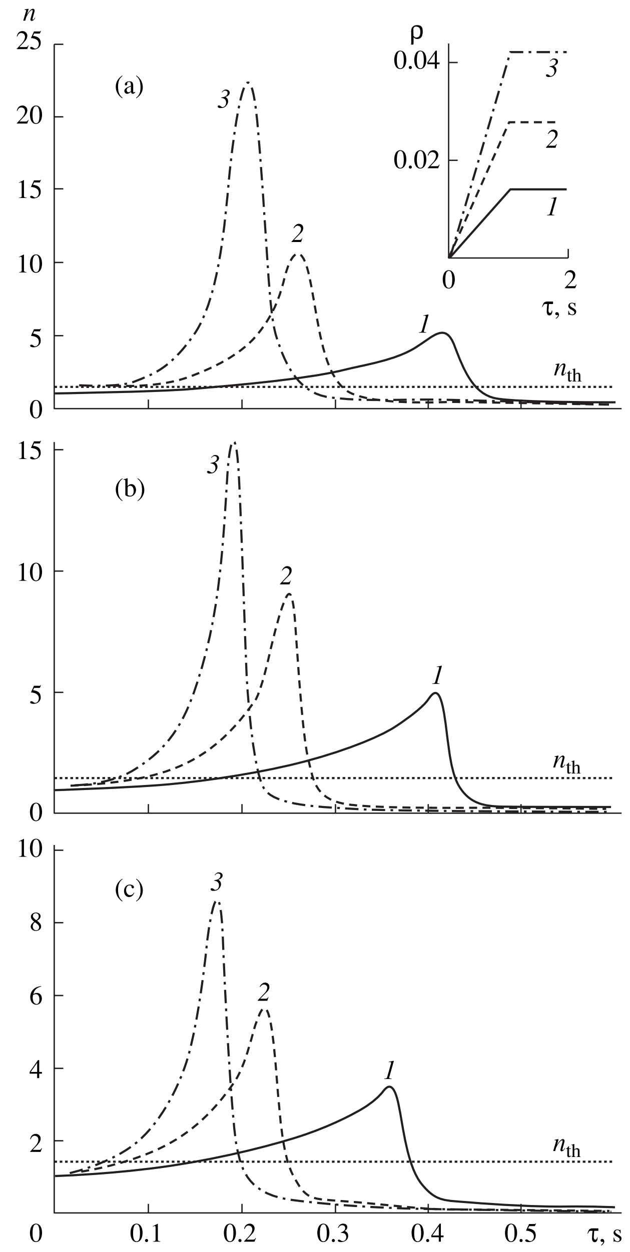
Since the instantaneous flow rate of an incompressible liquid is the same in any cross section, $v_1 = (1 - \Theta^2)v$. Hence, we have

$$v(h) = \frac{P_0 \varepsilon R^2}{8\eta(1-\Theta^2)H[\varepsilon + (1-\Theta^2)h/H]\{1 - [1 - F(\Theta)/(1-\Theta^2)^2]h/H\}}. \tag{7}$$

Treating $v(h)$ as the rate of displacement of the absorbent level in the working volume, we come to the equation for the depth to which the DEP control member descends after the actuator operation:

$$\frac{dh}{d\tau}v(h), \quad h(\tau = 0) = 0. \tag{8}$$

Figure 2 demonstrates the solutions of Eq. (8) for different initial thicknesses ε of the gas layer and various initial gas pressures P_0 in the container. It is seen that the actuator can be introduced into the reactor core for a time comparable to that for which the rods of automated emergency protection are introduced when they remedy a reactivity accident.



REACTIVITY ACCIDENT SIMULATION

In a reactivity accident, the DEP element behavior is described by the set of equations including the equations of point kinetics and Eqs. (3) and (8):

$$\begin{gathered}
\frac{dn}{d\tau} = \frac{\rho - \delta\rho(h) - \beta}{l}n + \Sigma\lambda_i C_i, \quad n(\tau = 0) = 1, \\
\frac{dC_i}{d\tau} = \frac{\beta_i}{l}n - \lambda_i C_i, \quad C_i(\tau = 0) = \frac{\beta_i}{l\lambda_i}, \quad i = 1, \ldots, 6, \\
\frac{dt}{d\tau} = \frac{\Theta - t}{l} + an, \quad t(\tau = 0) = t_0, \quad t(\tau \geq \tau_{th}) = t_{th}, \\
\frac{dh}{d\tau} = v(h), \quad h(\tau \leq \tau_{th}) = 0, \quad h(\tau > \tau_{th} + \tau_{1,}) = H,
\end{gathered} \tag{9}$$

where τ_{th} is the time instant at which the actuator comes into action and $\tau_{1\,th}$ is the time for which the absorbent is introduced into the DEP element working volume.

The DEP system supplements the automated emergency protection of the reactor [22]. In [14], the threshold neutron flux density is taken to be $n_{th} = 1.2$; we put $n_{th} = 1.4$.

A reactivity accident is simulated by introducing the function $\rho(t)$ that specifies the positive reactivity growth in the equations of point kinetics. If the positive reactivity grows with a constant rate, $\rho(\tau) = v_\beta \beta\tau$ $(\tau \geq 0)$, where v_β is the rate of positive reactivity growth in terms of the total fraction β of delayed neutrons, it is

Fig. 3. Remedy of the reactivity accident due to the growth of the positive reactivity with different rates: v_1 = (*1*) 2, v_2 = (*2*) 4, and v_3 = (*3*) 6 s^{-1}. The DEP parameters are $\varepsilon = 0.2$ and $\rho_U = 5000$ kg/m^3. (a) $P_0 = 1$ MPa; $\chi = 0.05$; $\tau_{1\,th} = 1.08$ s; and τ_{th} = (*1*) 0.4, (*2*) 0.24, (*3*) 0.18 s. (b) $P_0 = 2$ MPa; $\chi = 0.05$; $\tau_{1\,th} = 0.54$ s; and τ_{th} = (*1*) 0.4, (*2*) 0.24, and (*3*) 0.18 s. (c) $P_0 = 2$ MPa; $\chi = 0.06$; $\tau_{1\,th} = 0.54$ s; and τ_{th} = (*1*) 0.35, (*2*) 0.22, and (*3*) 0.16 s.

reasonable to limit the reactivity:

$$\rho(\tau) = \begin{cases} \rho_0 v_\beta \beta t, & 0 \le \tau \le \tau_{\max} \\ \rho_0 v_\beta \beta \tau_{\max}, & \tau > \tau_{\max}. \end{cases} \quad (10)$$

At fixed $\tau_{\max}$, to small v_β, there corresponds a slow growth of a relatively small reactivity; at high v_β, a large reactivity grows with a high rate. We put $\tau_{\max} = 1$ s.

Figure 3 demonstrates the solutions of Eqs. (9) when a reactivity accident is simulated by different growth rates of a positive reactivity. From Fig. 3a, it is seen that the gas–liquid DEP operates the faster, the more dangerous is the reactor runway. Figure 3b shows the solutions of Eqs. (9) at a high initial pressure in the neutron absorbent container. The speed of the gas–liquid DEP rises because the time of absorbent introduction into the core decreases. Figure 3c shows the solutions of Eqs. (9) for the case where the gas–liquid DEP spends a lesser power to remedy the reactor runaway because of a greater enrichment of the uranium in the sensing medium. Here, the DEP speed increases because the delay time of the actuator is smaller.

The results of reactivity accident simulation imply that the emergency protection system considered offers a high speed.

REFERENCES

1. Zh. S. Takibaev, Preprint No. 3-96, NYaTs RK (Almaty, 1996).
2. Zh. S. Takibaev, Izv. Akad. Nauk Resp. Kaz., Ser. Fiz.-Mat., No. 2, 54 (1997).
3. J. S. Takibaev, in *Proceedings of Fujii-e Symposium "Various Methods to Operate and Control Processes in an Active Zone in Condition of Extraordinary Situation," Tokyo, 1996*, p. 225.
4. Zh. S. Takibaev, in *Proceedings of the International Scientific and Practical Conference "Nuclear Power Engineering in Kazakhstan: Advanced Aspects," Aktau, Kazakhstan, 1996*, p. 71.
5. I. S. Balakirev and A. É. Sofiev, *Application of Pneumo- and Hydro-Automation Facilities in Chemical Industry* (Khimiya, Moscow, 1973).
6. M. D. Lemberg, *Relay Systems Pneumo-Automation* (Énergiya, Moscow, 1968).
7. V. N. Dmitriev and V. G. Gradetskiĭ, *Foundations of Pneumo-Automation* (Mashinostroenie, Moscow, 1979).
8. G. I. Verkhivker, S. D. Tetel'baum, and G. P. Konyaeva, At. Énerg. **24**, 158 (1968).
9. A. G. Morachevskiĭ and I. B. Sladkov, *Physicochemical Properties of Molecular Inorganic Compounds* (Khimiya, Moscow, 1987).
10. V. V. Malyshev, in *Thermophysical Properties of Gases* (Nauka, Moscow, 1976), pp. 97–105.
11. S. S. Bakulin, S. A. Ulybin, and E. P. Zherdyaev, in *Thermophysical Properties of Gases* (Nauka, Moscow, 1976), pp. 40–45.
12. I. A. Kuznetsov, *Emergency and Transient Processes in Fast Reactors* (Énergoatomizdat, Moscow, 1987).
13. B. S. Petukhov, L. G. Genin, and S. A. Kovalev, *Heat Transfer in Nuclear Power Plants* (Énergoatomizdat, Moscow, 1986).
14. B. A. Dement'ev, *Nuclear Power Reactors* (Énergoatomizdat, Moscow, 1990).
15. I. Ya. Emel'yanov, P. A. Gavrilov, and B. N. Seliverstov, *Control and Safety of Nuclear Power Plants* (Atomizdat, Moscow, 1975).
16. V. A. Sidorenko, *Safety Problems of Water-Moderated Water-Cooled Power Reactor Works* (Énergoatomizdat, Moscow, 1977).
17. G. G. Bartolomeĭ, G. A. Bat', V. D. Baĭbakov, and M. S. Alkhutov, *Foundations of Theory and Computing Methods of Nuclear Power Reactors* (Énergoatomizdat, Moscow, 1989).
18. J. M. Harrer, *Nuclear Reactor Control Engineering* (Van Nostrand, Princeton, 1963; Atomizdat, Moscow, 1967).
19. B. A. Dement'ev, *Kinetics and Regulation of Nuclear Reactors* (Énergoatomizdat, Moscow, 1986).
20. P. A. Kirillov, Yu. S. Yur'ev, and V. P. Bobkov, *Handbook of Heat-hydraulic Calculations (Nuclear Reactors, Heat Exchanger, Steam Generators)* (Énergoatomizdat, Moscow, 1990).
21. B. S. Petukhov, *Heat Transfer and Resistance in Laminar Fluid Flow in Pipes* (Énergiya, Moscow, 1967).
22. R. R. Ionaĭtis and N. L. Shvedov, At. Tekh. Rubezhom, No. 1, 10 (1988).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 4, 2002, pp. 470–473. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 99–102.
Original Russian Text Copyright © 2002 by Belyaev, Drokin, Shabanov, Shepov.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

High-Frequency Dielectric Spectra from Liquid Crystals of Series *n*CB and *n*OCB

B. A. Belyaev, N. A. Drokin, V. F. Shabanov, and V. N. Shepov
Kirenskiĭ Institute of Physics, Siberian Division, Russian Academy of Sciences,
Akademgorodok, Krasnoyarsk, 660036 Russia
e-mail: belyaev@iph.krasn.ru

Received July 5, 2001

Abstract—The design of a resonant frequency-tunable high-sensitivity microstrip sensor is suggested. The permittivity dispersion of liquid crystals of two homologic series, alkylcyanobiphenyls (7CB and 8CB) and alkyloxycyanobiphenyls (7OCB and 8OCB), is studied at frequencies of 100–900 MHz. The dielectric spectra are shown to be the sum of the Debye relaxation and dielectric resonances observed at $f \approx 160$, 280, 360, 450, 550, and 650 MHz. The dielectric resonances are present in the spectra of all the samples in both the nematic and isotropic phase. The substitution of an oxygen atom (series *n*OCB) for a carbon atom (series *n*CB) in liquid crystal molecules has a minor effect on the dielectric resonance frequencies but changes the resonance intensities and splits some of the resonance lines. *© 2002 MAIK “Nauka/Interperiodica”.*

INTRODUCTION

The Cole–Cole diagrams are very suitable for evaluating the deviation of liquid crystal dielectric spectra from the Debye frequency dependence of the dielectric constants [1, 2]. This requires measurements to be made at several widely separated frequencies in the high-frequency and microwave ranges. The deviations are usually described by a set of Debye relaxation regions. However, because of the lack of adequate experimental data, the reason for a number of spectral features remains unclear. In [3], the dielectric spectra in a wide frequency range (100 kHz–10 GHz) were studied by time-domain spectroscopy. It was shown that the dielectric spectrum taken from liquid crystals of the *n*CB series shows individual faint non-Debye dispersion regions at frequencies of 39, 225, and 550 MHz for 7CB and 34.3, 225, and 600 MHz for 8CB. It is of interest that these regions are observed in both the mesomorphic state and the isotropic phase.

The existence of extra low-intensity narrow-band dispersion regions was also reported in [4], where the dielectric spectra of liquid crystals placed in porous media were studied. In that work, the dispersion features of 5CB and 8CB crystals were observed for the perpendicular component of the permittivity in the nematic phase at 100 MHz.

In [5], the dispersion region of the permittivity was studied in *n*CB liquid crystals at frequencies from 50 to 500 MHz with special microstrip resonant sensors. Distinct extra dispersion regions imposed on the conventional Debye relaxation pattern were found. It is worthy to note that the frequency of the most intense region at 300 MHz does not depend on a specific representative from this homologic series, as well as on temperature and type of liquid-crystal ordering (nematic or smectic). At the same time, the intensity of this extra region strongly varies with temperature and homologic type.

In this work, we further refine the method of measuring the dielectric dispersion with microstrip resonant sensors and study the dielectric spectra of two types of liquid crystals: *n*-alkylcyanobiphenyls (7CB and 8CB) and *n*-alkyloxycyanobiphenyls (7OCB and 8OCB) at frequencies from 100 to 900 MHz. Also, we compare these spectra and discuss the effect of oxygen (*n*OCB series), which substitutes carbon (*n*CB series) in liquid crystal molecules, on the relaxation and resonance characteristics of the dielectric spectra.

MICROSTRIP INSTRUMENT SENSOR AND EXPERIMENTAL METHODS

High-sensitivity frequency-tunable (in discrete steps) resonance sensors built around microstrip structures have proved themselves a promising tool for measuring the dispersion in the decimeter range [6]. Such sensors operate with small amounts (volumes) of materials tested and take the dielectric spectra by measuring the real and imaginary components of the permittivity with a small frequency step.

For taking the dielectric spectra from the liquid crystals, we used the modified design of the microstrip resonant sensor [6] (Fig. 1). The sensor substrate is a 15×30-mm alumina plate with a permittivity $\varepsilon = 9.6$ and a thickness $h = 1$ mm. The lower plated side of the substrate is soldered to the metallic base of the sensor, which serves as a screen. On the upper side of the substrate, microstrip conductors forming a half-wave rectangular microstrip resonant loop are patterned by chemical etching. Input and output microstrip lines

(A and B) are connected to the opposite sides of this loop through coupling capacitors C_c. A capacitive measuring interdigitation C_x with a spacing of 0.15 mm is placed at the antinode of the microwave field of the resonator. Three breaks to connect lumped elements controlling the operation of the sensor were etched in the microstrip conductor at the maximum of the microwave field. The resonant frequency of the sensor can be varied in wide limits. This is achieved by inserting either calibrated inductances L, which lower the frequency of the sensor, or an electrically controlled capacitor (varactor V), which raises the frequency, into the central break of the microstrip conductor. In this way, the frequency of the sensor can be increased from 100 to 900 MHz in 20 MHz steps. The capacitors C_1 and C_2 and resistors R_1–R_4 serve to isolate the controlling circuits and to apply a dc orienting electric field to the sample. In the nematic phase, the long axis of the molecules rotate toward the microwave field when a voltage of ≈25 V is applied to the sample. In this case, the longitudinal component $\varepsilon_{\parallel}(f)$ of the permittivity is measured. The transverse component $\varepsilon_{\perp}(f)$ is measured by applying a permanent magnetic field $H = 2500$ Oe perpendicularly to the microwave electric field.

A 0.20-mm-thick liquid crystal sample was placed directly on the interdigitation, which was bounded by a glass rim, and covered by a Teflon film to prevent contamination. Identical expendable measuring cells were made for either liquid crystal homologue by a specially developed precise lithography technique that offers high reproducibility. The cells were connected to the mutual resonance loop. Such an approach provides a quick replacement of the samples and also prevents them from being contaminated during measurements, which inevitably occurs when a single measuring cell is used.

The real component of the permittivity was calculated from the difference in the resonance frequencies of the sensors with and without the sample. The imaginary component was calculated from the change in the loaded Q factor of the resonator when the liquid crystal was placed in it. The amplitude–frequency responses of the microstrip sensors were recorded with an R4-37 automated meter of complex transfer coefficients. For the nematic samples, the cell temperature was kept at $T = (T_{n-i} - 4)$ K, where T_{n-i} is the temperature of the nematic-to-isotropic phase transition, for their dielectric spectra to be convenient to compare. For the samples in the isotropic phase, the temperature was kept at $T = (T_{n-i} + 4)$ K. The absolute accuracy of temperature maintenance was $\Delta T = \pm 0.1$ K or higher.

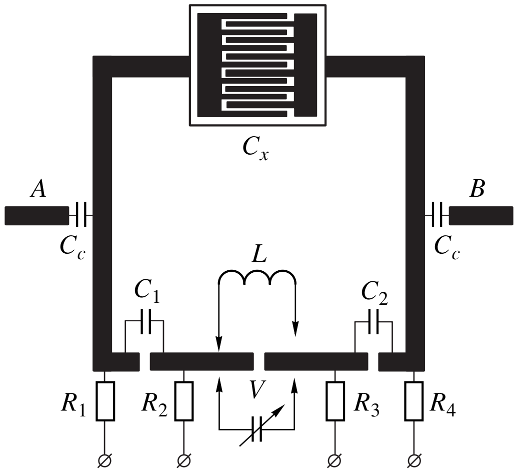


Fig. 1. Microstrip resonant sensor.

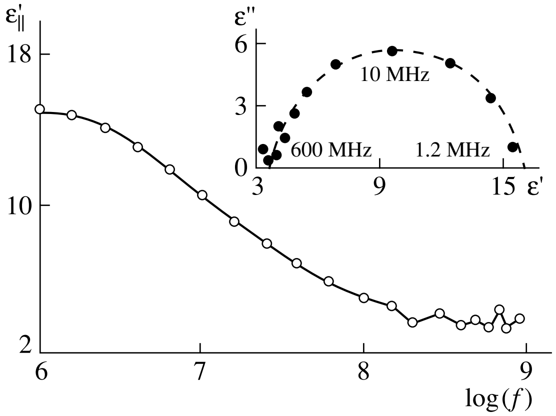


Fig. 2. Frequency dependence of $\varepsilon'_{\parallel}(f)$ in semilogarithmical coordinates for 7CB. The inset shows the Cole–Cole diagram.

RESULTS

Along with taking the high-frequency spectra, we measured the permittivity at some frequencies in the megahertz range. It was necessary for providing the desired accuracy in numerically approximating the frequency dependence of the permittivity with the Debye relaxation formula. By way of example, Fig. 2 demonstrates the dielectric spectrum $\varepsilon'_{\parallel}(f)$ for the 7CB sample. The inset here shows the Cole–Cole diagram with data points. In the right-hand side of this diagram, the data points are fairly well described by a semicircle. In the left-hand side, however, the measurements markedly deviate from the semi-circle both to the left and to the right. As follows from the dispersion curve $\varepsilon'_{\parallel}(f)$ in Fig. 2, such a deviation occurs because the dielectric spectrum contains faint yet well distinguishable non-Debye dispersion regions at $f > 100$ MHz.

Figure 3 shows the high-frequency part of the dielectric spectrum for the real, $\varepsilon'_{\parallel}(f)$, and imaginary, $\varepsilon''_{\parallel}(f)$, components of the permittivity of the 7CB liquid crystal sample. The dashed curve corresponds to the Debye approximation of $\varepsilon'_{\parallel}(f)$ with a single relaxation

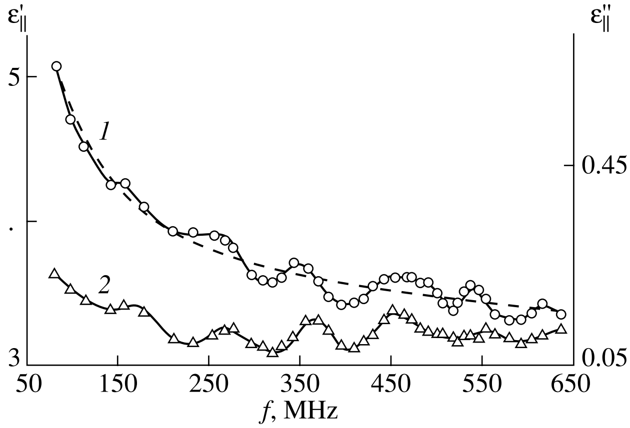


Fig. 3. High-frequency spectra (*1*) $\varepsilon'_{\parallel}(f)$ and (*2*) $\varepsilon''_{\parallel}(f)$ for 7CB sample.

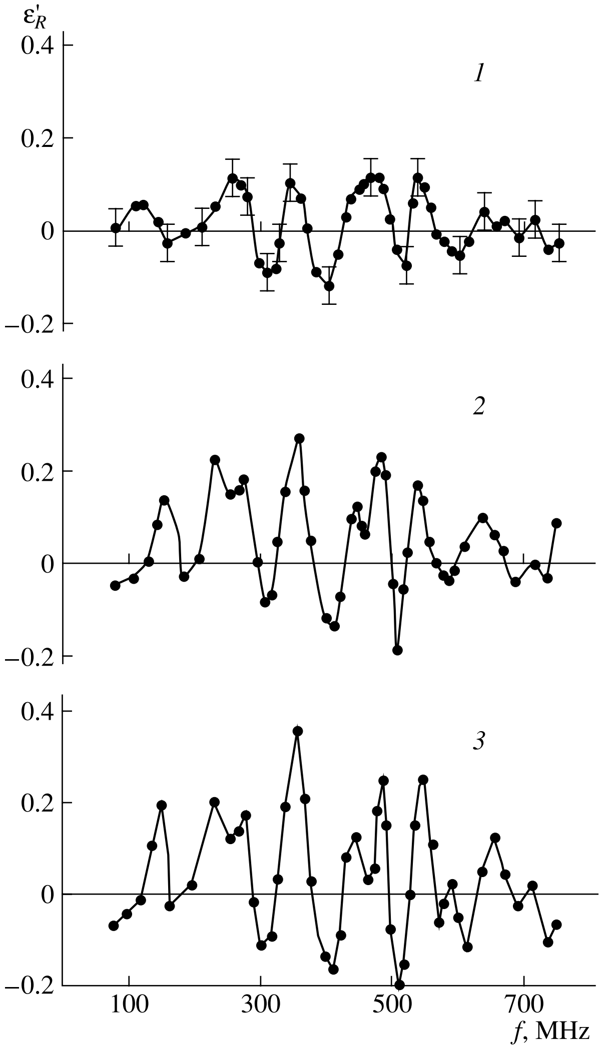


Fig. 4. Resonance spectra ε'_R obtained from $\varepsilon'_{\parallel}(f)$ for (*1*) 7CB, (*2*) 7OCB, and (*3*) 8OCB samples.

time $\tau = 3.25 \times 10^{-8}$ s. The extra dispersion regions with $\varepsilon''_{\parallel}(f)$ peaks at $f \approx 160$, 280, 360, 450, 550, and 650 MHz. Several hardly distinguishable peaks at higher frequencies up to 900 MHz are also observed. Comparing curves *1* and *2*, one can see that the peaks of $\varepsilon''_{\parallel}(f)$ in Fig. 3 coincide with the points where the dispersion curve $\varepsilon'_{\parallel}(f)$ meets the dashed curve, which corresponds to the Debye approximation. With the narrow extra dispersion regions, such behavior of $\varepsilon'_{\parallel}(f)$ and $\varepsilon''_{\parallel}(f)$ is typical of resonance processes. These may be microwave-field-induced intramolecular vibrations of mobile molecular fragments. A phenomenological dielectric spectrum can be described by the sum of the Debye relaxation and dielectric resonances:

$$\varepsilon'(\omega) - \varepsilon'_{\infty} = \frac{(\varepsilon'_0 - \varepsilon'_{\infty})}{1 + \omega^2\tau^2} + \frac{1}{2}\sum_i \Delta\varepsilon_i \times \left[\frac{1 + \omega_{0i}(\omega + \omega_{0i})g_i^2}{1 + (\omega + \omega_{0i})^2 g_i^2} + \frac{1 - \omega_{0i}(\omega - \omega_{0i})g_i^2}{1 + (\omega - \omega_{0i})^2 g_i^2}\right], \tag{1}$$

where ε'_0 and ε'_{∞} are the real components of the static and high-frequency permittivities, $\tau = 1/2\pi f_D$ is the relaxation time (f_D is the Debye resonance frequency), $\Delta\varepsilon$ is the resonance intensity, ω_0 is the resonance frequency, g is the damping factor, i is the resonance number, and $\omega = 2\pi f$.

For more detailed comparison of the frequencies and heights of the resonance peaks, we proceeded as follows. From the spectrum $\varepsilon'_{\parallel}(f)$ measured, the numerically approximated Debye relaxation contribution was subtracted. The “resonance spectra” $\varepsilon'_R(f)$ thus calculated for the 7CB, 7OCB, and 8OCB samples are given in Fig. 4 for the nematic phase. Even slight distortions of the lines are discovered, and the behavior of the resonance frequencies and peak intensities is clearly seen. Note that the resonance spectrum for 8CB differs from that for 7CB insignificantly and is omitted in Fig. 4. It follows from the spectra that the typical resonance frequencies determined from Fig. 3 ($f \approx 160$, 280, 360, 450, 550, and 650 MHz) do not depend on the mesogene type for both *n*CB and *n*OCB series. However, the presence of oxygen in an alkyloxycyanobiphenyl molecule (7OCB and 8OCB) between the rigid core and loose H–C–H alkyl groups increases the intensity of the resonance peaks and even causes some of them to split.

In the isotropic phase, the shape of the spectra remains virtually unchanged for both series, but the resonance line intensity increases almost twice. Such behavior was also observed in [5, 7] for 5CB and 6CB samples. Comparing the Debye relaxation parameters obtained by approximation, one can see that the substitution of an oxygen atom for a carbon atom in liquid crystal molecules strongly affects the Debye parameters: the relaxation time decreases, the high-frequency component $\varepsilon'_{\parallel}(\infty)$ of the permittivity grows, and the dielectric anisotropy increases. The numerical values of these parameters for the liquid crystals studied are very close to those obtained in [1, 2, 8].

It was shown [8] that an angle β between the dipole moment of an alkyl group and the axis of symmetry of the molecule increases with n in the nCB series and still stronger in the nOCB series. In other words, as the length of alkyl groups in the nCB and nOCB series increases, so does the normal component of the total dipole moment of the molecules. It seems likely that the change in the resonance line intensity and the split of the lines are also associated with an increase in the normal component of the dipole moment of the molecules. Of course, this assumption needs experimental verification; however, it should be taken into account in studying the dielectric spectra considered in this work.

CONCLUSION

With special microstrip frequency-tunable resonant sensors, we studied the high-frequency dielectric spectra of alkylcyanobiphenyl (7CB and 8CB) and alkyloxycyanobiphenyl (7OCB and 8OCB) liquid crystals at frequencies between 100 and 900 MHz. The dielectric spectra of each of the crystals are described by the sum of the Debye relaxation and dielectric resonances at $f \approx$ 160, 280, 360, 450, 550, and 650 MHz. For the longitudinal component $\varepsilon'_{\parallel}(f)$ of the permittivity, the relaxation is characterized by a single relaxation time for each of the crystals. By comparing the dielectric spectra, we conclude that oxygen in the nOCB liquid crystals has a minor effect on the dielectric resonance frequencies but changes the resonance intensities and causes some of the resonance line to split.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant no. 00-03-32206).

Shepov greatly appreciates the financial support from the expert group in the young scientist competition (Siberian Division, Russian Academy of Sciences) devoted to Akademician Lavrent’ev’s 100th birthday.

REFERENCES

1. D. Lippens, J. P. Parneix, and A. Chapoton, J. Phys. (Paris) **38**, 1465 (1977).
2. J. M. Wacrenier, C. Druon, and D. Lippens, Mol. Phys. **43**, 97 (1981).
3. T. K. Bose, B. Campbell, S. Yagihara, and J. Thoen, Phys. Rev. A **36**, 5767 (1987).
4. G. P. Sinha and F. M. Aliev, Phys. Rev. E **58**, 2001 (1998).
5. B. A. Belyaev, N. A. Drokin, V. F. Shabanov, and V. N. Shepov, Fiz. Tverd. Tela (St. Petersburg) **42**, 956 (2000) [Phys. Solid State **42**, 987 (2000)].
6. B. A. Belyaev, N. A. Drokin, and V. N. Shepov, Zh. Tekh. Fiz. **65** (2), 189 (1995) [Tech. Phys. **40**, 216 (1995)].
7. B. A. Belyaev, N. A. Drokin, V. F. Shabanov, and V. N. Shepov, Pis’ma Zh. Éksp. Teor. Fiz. **66**, 251 (1997) [JETP Lett. **66**, 271 (1997)].
8. S. Urban, B. Gestblom, and A. Wurflinger, Mol. Cryst. Liq. Cryst. **331**, 1973 (1999).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 4, 2002, pp. 474–477. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 103–106.
Original Russian Text Copyright © 2002 by Gilev, Trubachev.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Generation of Magnetic Field by Detonation Waves

S. D. Gilev and A. M. Trubachev

Lavrent'ev Institute of Hydrodynamics, Siberian Division, Russian Academy of Sciences,
pr. Akademika Lavrent'eva 15, Novosibirsk, 630090 Russia
e-mail: Gilev@hydro.nsc.ru

Received July 30, 2001

Abstract—The generation of a magnetic field by a system of detonation waves in a condensed explosive is reported. The convergence of the detonation waves, which exhibit a high conductivity in the chemical reaction zone, increases the magnetic field at the axis of the system. The fact of magnetic field generation is demonstrated experimentally. Features of the new method of magnetic cumulation are discussed. A simple compression model that qualitatively agrees with experimental data is proposed. *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Magnetic flux cumulation by a conducting shell set in motion by detonation products of high chemical explosives (HEs) has been used to obtain extrahigh magnetic fields up to 2×10^3 T and energy densities on the order of 10^6 J/cm^3 [1–5]. Difficulties associated with this method are conducting boundary instability, conducting layer overheating, and complexity of the system that initiates the magnetic flux. The shock-wave magnetic cumulation method [6–9] largely eliminates these problems. In this method, a material placed into the compression region becomes a good conductor under the action of a shock wave. A system of shock waves converging towards the axis traps and compresses the magnetic flux. The strongest magnetic fields produced by small-size shock-wave generators with an HE charge of about 0.5 kg were about 400 T [7].

In this paper, it is proposed to compress the magnetic flux and generate strong magnetic fields with the help of detonation waves in condensed explosives. The conducting layer of the detonation products acts as a piston, which compresses the magnetic flux. This technique retains many advantages of the shock-wave method and still does not require an additional power supply. The efficiency of the new magnetic cumulation method depends on the conductivity of the detonation products, which must be high enough.

The conductivity of the detonation products from most condensed HEs is $\sigma \sim 1$ $(\Omega\ \mathrm{cm})^{-1}$ [10–13]. Early measurements, however, were erroneous: it has been shown [14] that for trinitrotoluene (TNT) detonation products, $\sigma \approx 25$ $(\Omega\ \mathrm{cm})^{-1}$. A new measuring procedure that significantly improves the time resolution and raises the upper limit of conductivity measurement has been proposed [15]. When applied to cast TNT, this procedure revealed the complex nature of detonation conduction associated with the chemical reaction zone and the subsequent expansion of the detonation products. The peak conductivity in the chemical reaction zone (≈1 mm) was measured to be ≈250 $(\Omega\ \mathrm{cm})^{-1}$ and ≈30 $(\Omega\ \mathrm{cm})^{-1}$ outside it.

HIGH-CONDUCTIVITY ZONE OF DETONATION PRODUCTS

The peak electrical conductivity of the TNT detonation products [15] is the highest among HEs. However, it is insufficient for magnetic cumulation. The conductivity can be increased by introducing additives (a metal or semiconductor powder) into HEs. The metallic additive is inherently conductive. The semiconductor additive is initially nonconducting but acquires good conducting properties (i.e., passes to the metallic state) under the action of a shock wave due to the thermal excitation of the semiconductor subjected to strong compression.

We measured the electrical conductivity of the detonation products of HE–metal mixtures. The experiment was similar to that described in [15]. The shunt was of the same width as the space occupied by the HE. It was demonstrated that increasing the weight content of the metal in the mixture to ≈50% augments the detonation product conductivity by more than three orders of magnitude compared with HEs without the additive. The maximum content of the conducting additive in the mixture is limited because its too great amount impairs the detonation power. For mixtures with linear dimensions of about 1 cm, the resistance of the detonation products may be of several milliohms.

Figure 1 shows the conductivity profile of the cyclotetramethylenetetramine (HMX)/Al = 60/40 mixture in a detonation wave. As follows from the figure, the detonation conductivity profile is highly nonuniform. The maximal conductivity is observed in the ≈2-mm-thick zone adjacent to the detonation front.

This high-conductivity zone travels through the HE together with the wave front and is essentially coincident with the chemical reaction zone. The aluminum oxidation reaction and the expansion of the detonation products rapidly decrease the conductivity. Such a specific object (a high-conductivity zone propagating at a speed of ≈5 km/s) is of interest for high-power electromagnetic systems. In particular, the high-conductivity zone of the HE chemical reaction can be used for compressing the magnetic flux and generating a magnitude field.

DETONATION GENERATOR OF MAGNETIC FIELD

Figure 2 is the cross-sectional view of a detonation magnetic field generator. The generator comprises 12 initiators *1*, equispaced in a circle. The initiators detonate auxiliary HE *2*. The central region of the generator (62 mm in diameter and 50 mm in height) is filled with operating HE *3* (HMX/Al = 60/40 mixture of density ≈1.4 g/cm^3). The detonation waves excited in the operating HE converge to the axis, at which inductive pickup *4* is placed.

The degree of convergence of the detonation waves to the axis was estimated by detecting the optical radiation from the generator surface with the help of a VFU streak camera operating in a frame-by-frame mode. Figure 3 shows a sequence of photographs with the operating region magnified to 150 mm. The time increases from top to bottom and from left to right. The interval between frames is 1.6 μs. The maximum luminance corresponds to the region near the detonation front. The efficiency of wave initiation and the degree of convergence can be evaluated as satisfactory.

Prior to the detonation, a magnetic field of 2 T was created in the operating space. To this end, a bank of capacitors was discharged through external Helmholtz coils. The current through the coils was monitored with a reference inductive gage; the signal from measuring pickup *4*, with an S9-27 oscilloscope.

Figure 4a shows an integrated record of the voltage from the measuring pickup. This signal is proportional to the magnetic field at the pickup location. The magnetic field at the axis of the system increases as the detonation waves in the operating HE converge. This means that a conducting plug, which compresses the magnetic flux, arises in the HE. In this experiment, the magnetic field intensity grew by a factor of 2.2.

ANALYSIS OF THE DETONATION GENERATOR PERFORMANCE

Consider a converging cylindrical detonation wave propagating in a transverse magnetic field. Let the HE conductivity change in a stepwise manner from zero to σ at the infinitely thin front of the wave. The one-dimensional Maxwell equations and the Ohm law yield the system of equations

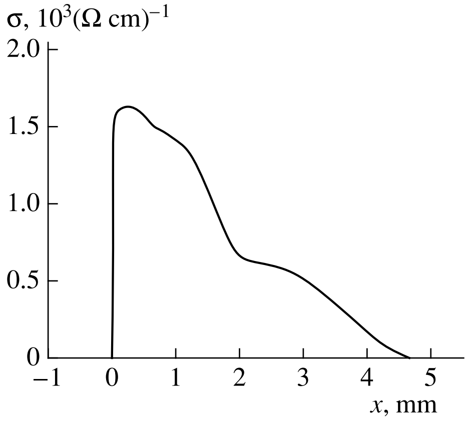


Fig. 1. Conductivity profile of the HMX/Al = 60/40 mixture in a detonation wave. *x* = 0 is the position of the detonation front.

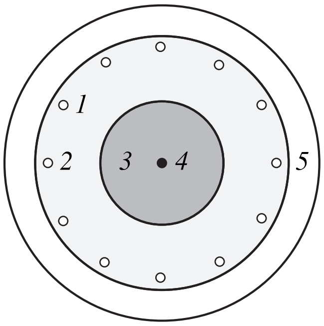


Fig. 2. Detonation generator of magnetic field: *1*, initiation points; *2*, boost explosive; *3*, operating explosive; *4*, inductive pickup; and *5*, shell.

$$\frac{\partial B}{\partial r} = -\mu_0 j, \quad (1)$$

$$\frac{1}{r}\frac{\partial(rE)}{\partial r} = -\frac{\partial B}{\partial t}, \quad (2)$$

$$j = \sigma(E + uB), \quad (3)$$

where *B* is the magnetic induction, *j* is the electric current density, *E* is the electric field, and *u* is the mass velocity. The magnetic field at the front of the wave is continuous (the front is infinitely thin and, therefore, does not disturb the current). Hence, the electric field is also continuous. Equations (1)–(3) and the continuity condition for the electric field lead to the relationship

$$\frac{r}{2}\frac{\partial B}{\partial t}\bigg|_{r = r_f} = uB + \frac{1}{\mu_0\sigma}\frac{\partial B}{\partial r}\bigg|_{r = r_f}, \quad (4)$$

which is valid at the wave front. Here, r_f is the current radius of the front. If $\sigma \longrightarrow \infty$ and the mass velocity *u*

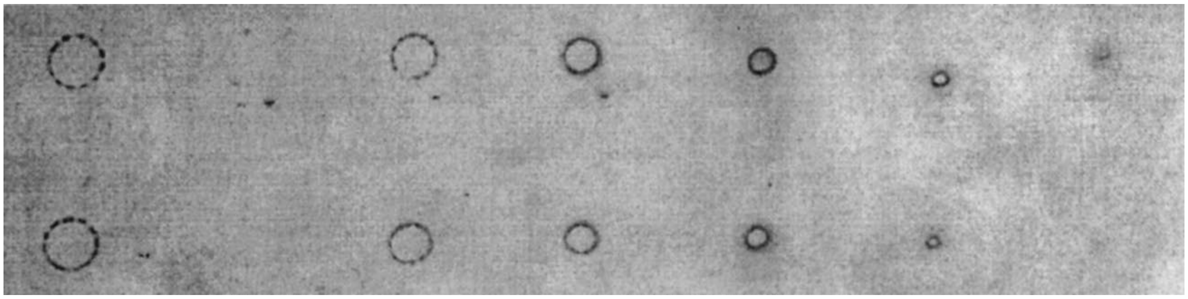


Fig. 3. Convergence of detonation waves (a sequence of photographs) in the generator.

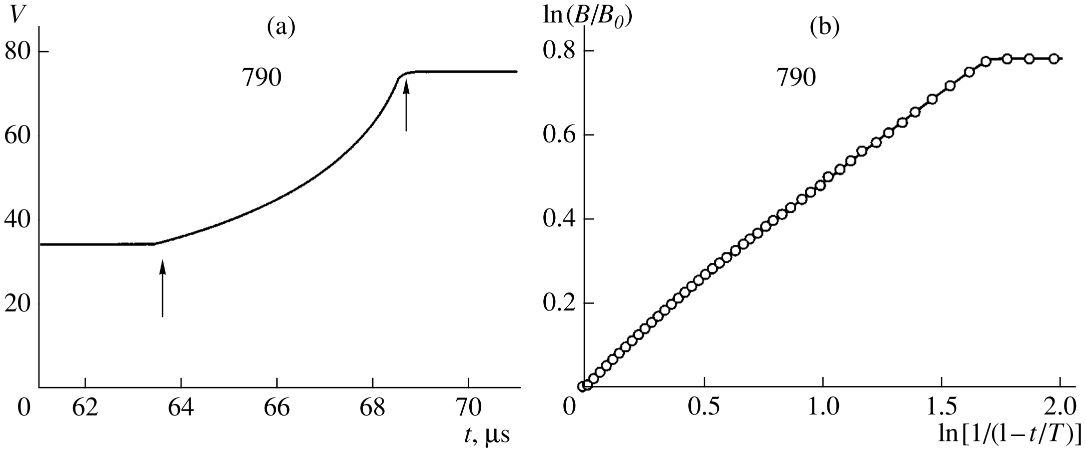


Fig. 4. Magnetic field generation by a detonation wave in the condensed explosive. (a) Record of the inductive pickup (the signal is proportional to the magnetic field at the axis); the arrows show the time instants when the conducting layer is formed in the explosive and when the wave arrives at the pickup. (b) Experimental data processing; the slope of the curve determines the effective magnetic field growth index α.

and the wave velocity D are constant, the following relationship can easily be obtained from (4):

$$\frac{B(t)}{B_0} = \left(\frac{R_0}{r_f}\right)^{2\frac{n}{D}}.$$

Here, B_0 is the initial magnetic field and R_0 is the initial radius of the region occupied by the HE. This formula was obtained in [6] for the shock-wave compression of the magnetic field. The u/D ratio depends on the compressibility of the material and tends to unity in highly porous substances. In the case of a detonation generator, the value of u/D can be varied in narrow limits. For normal Chapman–Jouguet detonation of condensed explosives, $u/D \approx 0.25$ [16]. The second term on the right-hand side of Eq. (4) is negative; hence, the ohmic loss decreases the rate of field rise. Let the field be compressed by a uniform conducting layer with a finite conductivity that moves at a constant velocity. A numerical solution to this model problem shows that the magnetic field at the axis can be described by the formula

$$\frac{B(t)}{B_0} = \left(\frac{R_0}{r_f}\right)^{2\alpha}, \tag{5}$$

where α is an effective exponent ($\alpha < u/D$ at finite σ and $\alpha \longrightarrow u/D$ as $\sigma \longrightarrow \infty$).

If formula (5) is valid, the representation of experimental data in the coordinates ($\ln(B/B_0)$ and $\ln[1/(1 - t/T)]$) (Fig. 4b) must yield a straight line (t is the time from the beginning of magnetic compression and T is the total compression time). From the slope of this line, one determines the exponent 2α. In Fig. 4b, the time T was calculated from the diameter of the region occupied by the HE and the detonation wave velocity D found in other experiments. As can be seen from the figure, the curve is close to a straight line, which indicates the validity of formula (5). As a result, the average exponent thus found is $\alpha \approx 0.24$. This value is close to u/D typical of condensed HE detonation.

The experiments show that the increase in the magnetic field significantly depends on the detonation regime. For example, for the damped detonation, the field was observed to increase by a factor of ≈8. This effect occurs because the compression (and the effective exponent α) increases when the detonation wave transforms into the shock wave. The use of damped detonation also extends the operating life of the generator.

A thorough analysis of the initial stage of the cumulation process reveals the presence of overcompressed detonation, which occurs when the mixture is initiated by a high booster explosive. This effect in Fig. 4b shows up in the larger slope of the curve at the beginning of the process. Thus, the transient detonation regimes affect the magnetic cumulation efficiency.

DISCUSSION

Our experiments have shown that detonation waves can be used for generating a magnetic field. The efficiency of the detonation generator is limited by two factors. First, the ratio u/D for a detonation wave is relatively small. This leads to high convection losses of the magnetic flux through the wave front. Second, the field is compressed by the conducting region of nonzero resistance. In our experiments, the magnetic Reynolds number of the conducting region $\mathrm{Re}_m = \mu_0 \sigma u x_0$ (x_0 is the dimension of the region) was estimated as $\mathrm{Re}_m \approx 0.4$. HEs with a higher conductivity of the detonation products and with a wider chemical reaction zone can provide better results.

An advantage of the detonation method of magnetic cumulation is that it extends the potentialities of the device. The use of the detonation wave gives more freedom in choosing the compression geometry and allows for its transformation in the cumulation process. This is impossible to achieve in the method of compression with a metallic shell [1, 2] and is difficult to realize in the shock-wave compression method [6–9], where rarefaction waves become significant because of the edge effects. No fundamental limitations are imposed on the size of the initial field-generating areas. The magnetic field generator can also integrate different approaches to magnetic cumulation.

Note that, since the output signal depends on the state of the detonation products, experiments with a detonation generator may be useful for studying the detonation process. An increase in the field significantly depends on the transient detonation regimes. This means that the magnetic cumulation principle can be the basis for the noninvasive diagnostics of detonation processes.

CONCLUSION

The method of generating a magnetic field by a system of detonation waves in a condensed HE is shown to be efficient. Unlike shock wave generation, detonation is a self-sustained process, which imparts new properties to cumulation systems (energy independence, flexibility, and the ability to dynamically change the compression geometry).

ACKNOWLEDGMENTS

We are grateful to A.M. Ryabchun for developing a program that simulates magnetic flux compression by a conducting layer.

This work was supported by the Russian Foundation for Basic Research (grant no. 99-02-16807).

REFERENCES

1. C. M. Fowler, W. B. Garn, and R. S. Caird, J. Appl. Phys. **31**, 588 (1960).
2. A. D. Sakharov, Usp. Fiz. Nauk **88**, 725 (1966) [Sov. Phys. Usp. **9**, 294 (1966)].
3. H. Knoepfel, *Pulsed High Magnetic Fields* (North-Holland, Amsterdam, 1970; Mir, Moscow, 1972).
4. A. S. Lagutin and V. I. Ozhogin, *Pulsed High Magnetic Fields in Physical Experiment* (Énergoatomizdat, Moscow, 1988).
5. *Megagauss and Megampere Pulsed Technology and Applications: Proceedings of the 7th International Conference on Generation of Megagauss Magnetic Fields and Related Experiments* (VNIIÉF, Sarov, 1997), Vols. 1, 2.
6. S. D. Gilev and A. M. Trubachev, Pis'ma Zh. Tekh. Fiz. **8**, 914 (1982) [Sov. Tech. Phys. Lett. **8**, 396 (1982)].
7. E. I. Bichenkov, S. D. Gilev, A. M. Ryabchun, *et al.*, Prikl. Mekh. Tekh. Fiz., No. 3, 15 (1987).
8. K. Nagayama, Appl. Phys. Lett. **38**, 109 (1981).
9. K. Nagayama and T. Mashimo, J. Appl. Phys. **61**, 4730 (1987).
10. A. A. Brish, M. S. Tarasov, and V. A. Tsukerman, Zh. Éksp. Teor. Fiz. **37**, 1543 (1959) [Sov. Phys. JETP **10**, 1095 (1960)].
11. R. L. Jameson, S. J. Lukasik, and B. J. Pernick, J. Appl. Phys. **35**, 714 (1964).
12. A. P. Ershov, P. I. Zubov, and L. A. Luk'yanchokov, Fiz. Goreniya Vzryva **10**, 864 (1974).
13. A. G. Antipenko, A. N. Dremin, and V. V. Yakushev, Dokl. Akad. Nauk SSSR **225**, 1086 (1975).
14. A. P. Ershov, N. P. Satonkina, O. A. Dibirov, *et al.*, Fiz. Goreniya Vzryva **36** (5), 97 (2000).
15. S. D. Gilev and A. M. Trubachev, Zh. Tekh. Fiz. **71** (9), 123 (2001) [Tech. Phys. **46**, 1185 (2001)].
16. *Physics of Explosion*, Ed. by K. P. Stanyukovich (Nauka, Moscow, 1975).

Translated by A. Khzmalyan

Technical Physics, Vol. 47, No. 4, 2002, pp. 478–483. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 107–112.
Original Russian Text Copyright © 2002 by Noskov, Malinovskiĭ, Cooke, Wright, Schwab.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Modeling of Discharges in Spatially Charged Dielectrics

M. D. Noskov*, A. S. Malinovskiĭ*, C. M. Cooke,**
K. A. Wright, and A. J. Schwab*****
** Tomsk Polytechnical University, Tomsk, 634034 Russia*
e-mail: noskov@hvd.tpu.ru
*** Massachusetts Institute of Technology, MA 02139-4307 Cambridge, USA*
e-mail: cmcooke@mit.edu
**** University of Karlsruhe, 76128 Karlsruhe, Germany*
e-mail: schwab@ieh.etec.uni-karlsruhe.de

Received August 2, 2001

Abstract—A stochastic–deterministic model of a discharge induced by space charge in dielectric is developed. The spatiotemporal and current characteristics of the discharge are simulated numerically. The simulation results are compared with the experimental data on discharges in polymethylmethacrylate samples charged by means of an electron beam. The interrelation among the growth of the current-carrying channels, energy transfer, and energy release during the discharge is discussed. *© 2002 MAIK “Nauka/Interperiodica”.*

INTRODUCTION

In low-conductivity solid dielectrics, long-lived space charges are possible to produce. A rather high local space charge density in dielectric can be achieved, e.g., with the help of moderately intense electron beams [1–4]. The electric field generated by space charge can provoke a discharge between the charged region and the dielectric surface. The growth of the discharge channels results in the formation of branching and twisting stochastic discharge structures. The use of an electron beam allows one to control the amount and spatial distribution of space charge. This makes it possible to evaluate the effect of the electric charge distribution inside the dielectric on the spatiotemporal and current characteristics of the discharge structure [5, 6].

It is reasonable to use numerical simulations to quantitatively describe a discharge induced by space charge. The greatest success in describing the stochastic growth and branching of the discharge channels was achieved using a fractal model of dielectric breakdown proposed in [7]. In this model, the growth of the channel was assumed to depend randomly on the electric field and the probability of the channel formation was assumed to be proportional to a certain degree of the local electric field. The electric field was calculated from the Laplace equation, and the discharge structure was regarded as an electrode continuation. In [8], the model was generalized by introducing the threshold electric field for the channel growth, assuming that the voltage drop along the channels is constant. Various modifications of the fractal model of breakdown in dielectric were used to describe the formation of dendritic and bushy structures [8, 9] and the influence of the nonuniform permittivity [9, 10], conductivity [11], and breakdown strength [12] on the discharge development. However, the interrelation between the spatiotemporal and current characteristics of the discharge structure cannot be adequately described without allowance for the dynamics of the charge propagation along the discharge channels and the change in the channel conductance. With this in mind, stochastic–deterministic models were proposed [13, 14], in which the discharge channel growth and the charge propagation were described self-consistently. In our study, this approach is developed and employed to describe a discharge induced by space charge. The results from computer simulations of the discharge structure are compared with the experimental data obtained when studying discharges in Plexiglass (polymethylmethacrylate).

EXPERIMENTAL SETUP

Space charge was produced by an electron beam accelerated up to an energy of 3 MeV with the help of a Van de Graaff generator. The electron beam was incident on a rectangular 12.7-mm-thick polymethylmethacrylate sample (Fig. 1a). At such a sample thickness, 3-MeV electrons produced a 2-mm-thick charged layer in the middle of the sample. The value and spatial distribution of space charge were controlled by changing the beam intensity and masking the sample surface. An electron beam with a current of several microamperes produced space charge with a density on the order of 1 μC/cm^2. The discharge was ignited when the tip of a grounded needle touched the sample lying on a grounded surface (Fig. 1b). The current flowing through the needle during the discharge was measured with the help of a Rogowski coil and an oscillograph.

STOCHASTIC–DETERMINISTIC MODEL OF THE DISCHARGE DEVELOPMENT

The model describes the growth of the discharge channels, the charge propagation along the channels, the dynamics of the electric field, and the change in the channel conductance. The discharge channel grows due to the destruction of dielectric and the onset of the conduction phase. Up to now, no quantitative theory of the formation of the conducting channels has been developed. This is related to both the complicated stochastic behavior of the channel growth and the great number of physicochemical processes involved (heating, cracking, gas release, dissociation, ionization, recombination, etc.) that occur in a strong electric field and lead to the onset of the conduction phase in dielectric. Since the phase transition occurs due to the electric field energy, one can attempt to describe the discharge structure growth by relating the probability of the conducting channel formation to the local density of the electric field energy. The form of this stochastic dependence can be found either by analyzing physicochemical processes in dielectric or from experimental data on the discharge development. In our model, a stepwise dependence of the growth probability on the local energy density is used as the first approximation. Inside the discharge structure, the probability density ω for the discharge channel to grow along the **n** direction is assumed to be proportional to the squared projection E_n of the local electric field onto this direction, provided that the projection magnitude exceeds a certain critical value E_c,

$$\omega = \alpha\Theta(E_n - E_c)E_n^2, \qquad (1)$$

where α is the growth rate constant, E_c is the critical electric field for the discharge channel growth, and $\Theta(x)$ is the step function ($\Theta(x) = 1$ for $x > 0$ and $\Theta(x) = 0$ for $x \le 0$).

The electric field potential is calculated as a superposition of the fields created by the charges located along the discharge structure L, on the sample surface S, and inside the dielectric volume V,

$$\varphi = \int_L \frac{\lambda dl}{4\pi\varepsilon\varepsilon_0 r} + \iint_S \frac{\sigma dS}{4\pi\varepsilon\varepsilon_0 r} + \iiint_V \frac{\rho dV}{4\pi\varepsilon\varepsilon_0 r}, \qquad (2)$$

where ε is the dielectric constant; ε_0 is the permittivity of free space; λ, σ, and ρ are the linear, surface, and space charge densities, respectively; and r is the distance between the observation point and the point where the charge is located.

The densities of the charges located along the discharge channels and on the dielectric surface alter during the discharge. It is necessary to take into account the surface charges because, when producing the space charge inside the dielectric, the ions of the opposite sign are deposited onto the sample surface. Moreover, the sample lies on the grounded surface. In modeling, the initial distribution of the surface charge density was set such that, in accordance with the experimental conditions, the potential on the sample surface be zero. It was assumed that the charge density changed only on the sample side that contacted the grounded surface. At every instant, the calculated charge distribution on this surface was adjusted so that its potential remained zero.

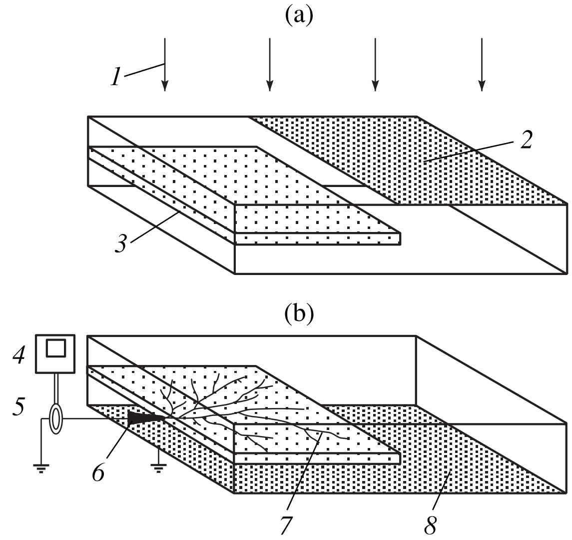


Fig. 1. Schematic of (a) the formation of space charge in dielectric and (b) the discharge ignition: (*1*) electron beam, (*2*) mask, (*3*) charged layer, (*4*) oscillograph, (*5*) Rogowski coil, (*6*) needle igniting the discharge, (*7*) discharge channels, and (*8*) grounded surface.

The charge propagation along the discharge channel is described by Ohm's law

$$I = \gamma E_l, \qquad (3)$$

where I is the current, γ is the channel conductance per unit length (the product of the conductivity by the channel cross-sectional area), and E_l is the projection of the electric field onto the channel direction.

During the discharge, the channel conductance per unit length changes due to the channel widening, ionization, recombination, electron attachment to the channel wall and gas molecules, etc. These processes depend on both the energy released in the channel due to Joule heating and the energy dissipated into the surrounding medium. Hence, it is reasonable, as the first approximation, to use the following equation for the time evolution of the conductance per unit length:

$$\frac{d\gamma}{dt} = \chi\gamma E_l^2 - \xi\gamma, \qquad (4)$$

where χ and ξ are the parameters governing the increase and decrease in the channel conductance, respectively.

The first term on the right-hand side of Eq. (4) relates the growth of the channel conductance per unit length to Joule heating. It can be regarded as an analogue of the Rompe–Weizel formula for the conduc-

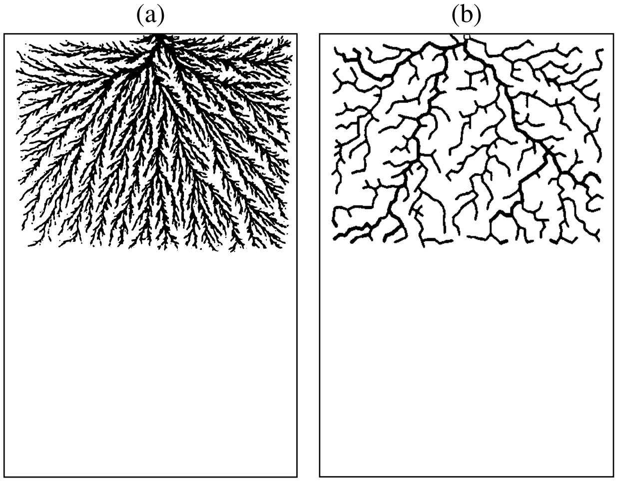


Fig. 2. (a) Experimental and (b) computed pictures of the discharge structure in the space charge region of a sample 90 × 135 mm in size for a charge density of 0.8 μC/cm^2 at the sample top.

tance of a spark channel [15] expressed in the differential form. The second term describes the decrease in the channel conductance per unit length caused by energy dissipation. The conductance per unit length of a newly created channel is set equal to γ_0.

Based on the above model, we developed a three-dimensional numerical code, which enabled computer simulations of discharges in spatially charged dielectrics. In most of the numerical models of the discharge structure growth (see, e.g., [7–13]), a rectangular lattice was used to calculate the electric field and to describe the channel growth. The possible growth directions were limited by the lattice edges and diagonals. In our model, a nonlattice algorithm is used, which allows us to avoid this limitation. The discharge structure is represented as a dendrite chain of point charges set at a distance d from each other. The d value determines the minimum spatial scale in modeling. The smaller d, the more detailed the description of the discharge structure. However, the reduction in d significantly increases the computation time. Here, we set $d = 2$ mm. The electric field was calculated using the charge method [16], based on the principle of superposition of the fields created by individual charges. In computer simulations, the time step was set at $\Delta t = 2 \times 10^{-9}$ s. At every time step, the following procedures were performed in succession: (i) the calculation of the field using Eq. (2) with allowance for change in the charge surface density; (ii) the simulation of the discharge structure growth by adding new points to either the existing structure or the initiating needle according to the distribution of the probability density (1); (iii) the calculation of the change in the charge distribution along the discharge structure in accordance with the charge conservation law and conduction equation (3); and (iv) the calculation of the change in the channel conductance per unit length according to Eq. (4). The simulations were stopped when the discharge structure growth terminated and the channel conductance dropped to zero.

Thus, the model incorporates five parameters: E_c, α, χ, ξ, and γ_0, which, together with a sample shape and the magnitude and spatial distribution of space charge, determine the spatiotemporal and current characteristics of the discharge. The physical meaning of the model parameters is quite clear. The critical electric field E_c is related to the existence of the threshold field for the formation of a conducting channel in dielectric. An increase in E_c reduces the dendricity of the discharge structure. The growth rate constant α relates the probability of the channel formation to the local density of the electric field energy. An increase in α stimulates the discharge structure growth. The parameter χ governing the increase in the conductance per unit length relates the rate at which the conductance grows to the Joule heating power in the channel. The parameter ξ governing the reduction in the conductance per unit length determines the rate at which the conductance decreases (the quantity reciprocal of ξ can be treated as the characteristic time during which the channel remains in the conducting state after the energy release terminates). An increase in χ and decrease in ξ enlarge the discharge channel conductance per unit length and reduce the voltage drop required to sustain the channel in the conducting state. This is accompanied by an increase in the electric field at the ends of the channels and the accelerated growth of the discharge structure. The initial conductance per unit length γ_0 characterizes the processes leading to the formation of conducting channels in dielectric. A decrease in γ_0 decelerates the conductance growth in the newly created channels but hardly affects the conductance of the existing channels. The model parameters can be determined by analyzing either the processes in dielectric and the discharge channels or the experimental data on the discharge development. The parameter values used in our study ($\alpha = 7 \times 10^{-6}$ m^2/V^2 s, $E_c = 5 \times 10^7$ V/m, $\chi = 4 \times 10^{-5}$ S m^2/J, $\xi = 10^8$ s^{-1}, and $\gamma_0 = 2 \times 10^{-9}$ S m) were determined by comparing the simulation results with the experimental data on discharges in polymethylmethacrylate.

RESULTS AND DISCUSSION

The growth of the discharge structure begins with the formation of one or a few channels that start at the needle tip initiating the discharge. The channels form only within the charged layer inside the dielectric sample. The shape of the region occupied by the discharge structure corresponds to the shape of the space charge distribution. If a mask is used and the charge is implanted only in one-half of the sample, then the discharge structure grows only within this zone and stops at its edge (Fig. 2). If the charge is implanted in a belt

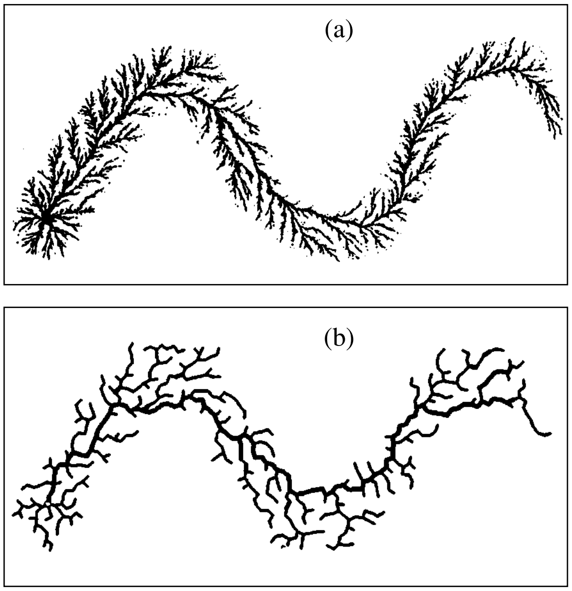


Fig. 3. (a) Experimental and (b) computed pictures of the discharge structure growing along the charged sinusoidal belt on a sample 80 × 160 mm in size for a charge density of 0.8 μC/cm^2 inside the belt.

shaped like a sinusoid, then the discharge structure grows within this belt (Fig. 3).

The growth rate of the discharge structure and the current flowing through the needle depend on the magnitude and spatial distribution of space charge in dielectric. The higher the charge density, the faster the discharge structure growth and the higher the maximum current. If the sample is charged uniformly, then, as the discharge develops, the current decreases. Figure 4a presents the simulated and measured time evolutions of the discharge current for a uniformly charged sample 50 × 140 mm in size (the charge density is 0.6 μC/cm^2). The growth rate of the discharge structure decreases with time. The simulated time evolution of the discharge structure length L (the distance from the needle tip to the structure farthest point) is shown in Fig. 4b (curve *1*). The channel conductance per unit length decreases simultaneously with the deceleration of the structure growth (Fig. 4b, curve *2*). When the space charge is distributed nonuniformly, the discharge characteristics become correlated with the distribution of the charge density in the channel regions. Figure 5a presents the time evolution of the current in a 50 × 140 mm sample, in which five transverse 50 × 24 mm belts with a charge density of 0.8 μC/cm^2 are separated by 5-mm-wide uncharged belts. The current maxima and minima correspond to the moments when the channels pass through the charged and uncharged belts, respectively. When the channels pass through the uncharged belts, the growth rate of the discharge structure and the channel conductance per unit length also decrease (Fig 5b).

The results obtained show that the charge transport along the channels and the energy released in the chan-

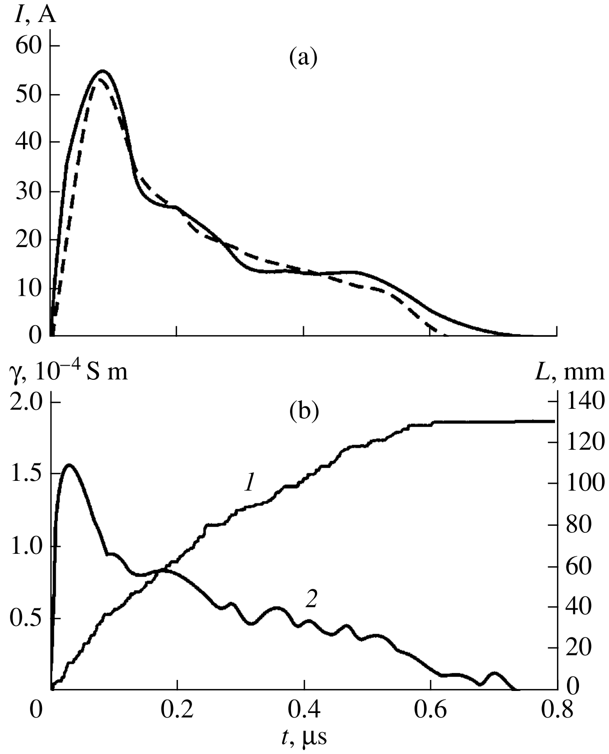


Fig. 4. Time evolution of the discharge in a uniformly charged sample: (a) the simulated (solid line) and measured (dashed line) discharge currents and (b) the discharge structure length (*1*) and the maximum conductance per unit length in the channels (*2*).

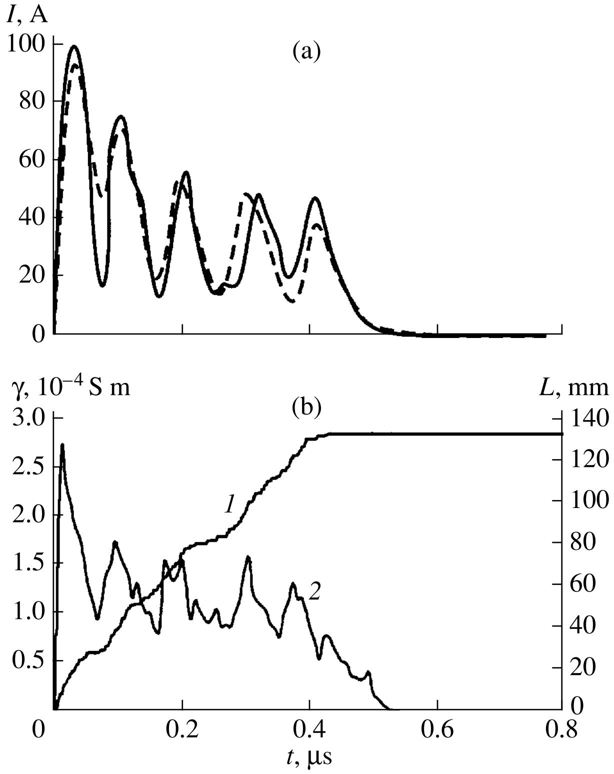


Fig. 5. Same as in Fig. 4 for a nonuniformly charged sample.

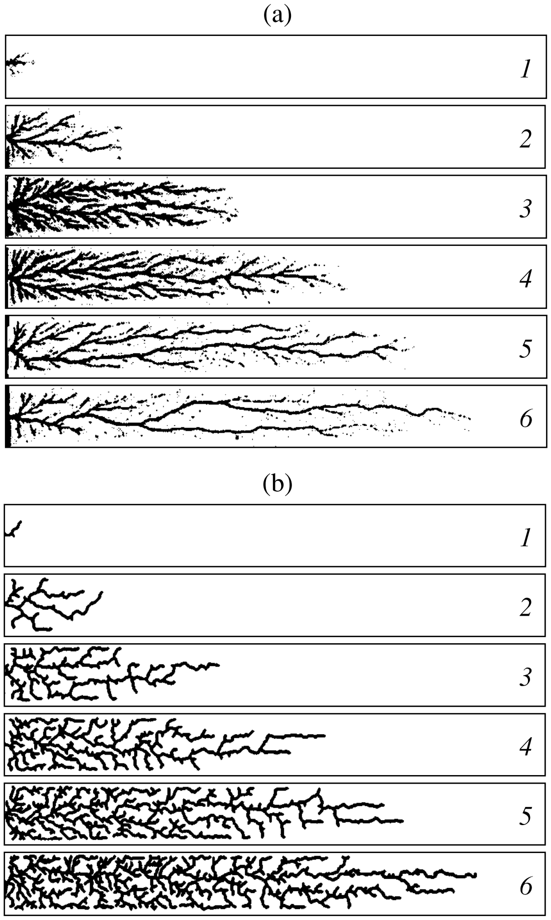


Fig. 6. Discharge structures in samples with different charge densities: (a) experimental results for $n = \rho/\rho_0 =$ (*1*) 1, (*2*) 1.5, (*3*) 2.0, (*4*) 2.5, (*5*) 3.0, and (*6*) 3.5 and (b) simulation results for the charge density equal to (*1*) 0.25, (*2*) 0.38, (*3*) 0.51, (*4*) 0.63, (*5*) 0.76, and (*6*) 0.88 $\mu C/cm^2$.

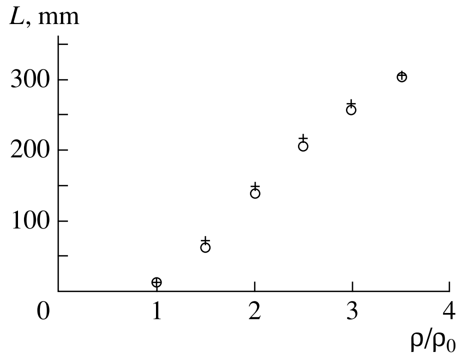


Fig. 7. Measured (crosses) and simulated (circles) discharge structure length vs. the reduced space charge density.

nels vary self-consistently with the growth of the discharge channels, which results in the formation of the discharge structure in a spatially charged dielectric. The necessary condition for the channel growth is the presence of a sufficiently high electric field at the channel ends. Hence, for the voltage drop along the channel to be insignificant, the discharge structure should be well conducting. To sustain the discharge channel in the conducting state, the energy should be permanently released in the channel. The electrostatic electric field created by the space charge in dielectric serves as an energy source. The growth of the discharge structure ensures the conversion of the field energy into the energy of Joule heating. Since the energy density distribution correlates with the space charge density distribution, the discharge channels tend to occupy the entire charged region. If there is an uncharged belt, the channels can cross it due to the field created by the charges located at the opposite side of the belt. However, in this case, the released power decreases, which leads to a decrease in the conductance, current, and growth rate of the channels. If the uncharged belt is sufficiently wide, the channel conductance can drop to zero and the discharge structure stops growing. A similar effect was observed in both simulations and experiments.

The lengthening of the discharge structure leads to a decrease in the power released per unit length of the channel. Hence, the structure growth is accompanied by a decrease in the channel conductance per unit length. This results in the reduction of the electric field at the channel ends and the deceleration of the growth. The latter leads to a decrease in the power released in the channel, which, in turn, causes a further decrease in the channel conductance. As a result, the conductance drops to zero and the discharge structure stops growing. The higher the space charge density in dielectric, the higher the energy density of the electrostatic field. Hence, the dimensions of the discharge structure should increase with the charge density. This conclusion agrees with the results of simulations and experiments with a discharge in a sample 42 × 350 mm in size at various charge densities. The discharge structure starts to grow when the charge density reaches a certain critical value ρ_0. Increasing the charge density in a stepwise manner with a step of 0.5 ρ_0 results in a significant enlargement of the discharge structure (Fig. 6). The dependence of the structure length on the charge density (normalized to ρ_0) is shown in Fig. 7.

A comparison of the experimental data with the simulation results shows that the model adequately describes the main features of a discharge induced by space charge in dielectric. This allows us to conclude that the main assumptions underlying the model reflect the real physical processes governing the development of the discharge. Since the model is rather general, it can also be applied to describe discharges in various materials under different conditions. The results obtained can be used to study the discharge phenomena in dielectric components of the apparatus that are exposed to charged particle beams (e.g., accelerators, space crafts, etc.).

CONCLUSION

A numerical model of a discharge induced by space charge in dielectric is proposed. The model is based on both a stochastic approach to describing the growth of the discharge channels and the use of deterministic equations for calculating the dynamics of the electric field and charge transport. The three-dimensional numerical model allows one to perform computer simulations of the spatiotemporal and current characteristics of the discharge. A quantitative agreement between the simulation results and the experimental data on the discharge development in polymethylmethacrylate samples charged with the help of electron beams indicates the reliability of the model. The results obtained show that the charge transport and the energy released in the channels vary self-consistently with the growth of the conducting channels, which results in the formation of the discharge structure in a spatially charged dielectric.

ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft and the Russian Foundation for Basic Research.

REFERENCES

1. B. Gross and K. A. Wright, Phys. Rev. **114**, 725 (1959).
2. H. Lackner, I. Kohlberg, and S. V. Nablo, J. Appl. Phys. **36**, 2064 (1965).
3. J. Furuta, E. Hiraoka, and A. Okamoto, J. Appl. Phys. **37**, 1873 (1966).
4. B. Gross and S. V. Nablo, J. Appl. Phys. **38**, 2272 (1967).
5. C. M. Cooke, E. R. Williams, and K. A. Wright, in *Proceedings of the IEEE International Symposium on Electrical Insulation, Philadelphia, 1982*, p. 95.
6. C. M. Cooke, E. R. Williams, and K. A. Wright, in *Proceedings of the International Conference on Properties and Applications of Dielectric Materials, Xian, China, 1985*, Vol. 2, p. 1.
7. L. Niemeyer, L. Pietronero, and H. J. Wiesmann, Phys. Rev. Lett. **52**, 1033 (1984).
8. H. J. Wiesmann and H. R. Zeller, J. Appl. Phys. **60**, 1770 (1986).
9. V. R. Kukhta, V. V. Lopatin, and M. D. Noskov, Zh. Tekh. Fiz. **65** (2), 63 (1995) [Tech. Phys. **40**, 150 (1995)].
10. O. S. Gefle, A. V. Demin, V. R. Kukhta, *et al.*, Élektrichestvo, No. 7, 61 (1994).
11. D. I. Karpov, V. V. Lopatin, and M. D. Noskov, Élektrichestvo, No. 7, 59 (1995).
12. P. J. J. Sweeney, L. A. Dissado, and J. M. Cooper, J. Phys. D **25**, 113 (1992).
13. V. Lopatin, M. D. Noskov, R. Badent, *et al.*, IEEE Trans. Dielectr. Electr. Insul. **5**, 250 (1998).
14. M. D. Noskov and A. S. Malinovskiĭ, in *Proceedings of the VI International Conference "Modern Problems of Electrophysics and Electrohydrodynamics of Fluids," St. Petersburg, 2000*, p. 159.
15. R. Rompe and W. Weizel, Z. Phys. B **122**, 9 (1944).
16. H. Singer, H. Steinbigler, and P. Weiss, IEEE Trans. Power Appar. Syst. **93**, 1660 (1974).

Translated by N. Ustinovskiĭ

Technical Physics, Vol. 47, No. 4, 2002, pp. 484–490. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 113–119.

Carbon Diffusion between the Volume and Surface of (100) Molybdenum

N. R. Gall', E. V. Rut'kov, and A. Ya. Tontegode

Ioffe Physicotechnical Institute, Russian Academy of Sciences,
ul. Politekhnicheskaya 26, St. Petersburg, 194021 Russia

Received September 18, 2001

Abstract—The diffusion of carbon atoms between the volume and the surface of (100) molybdenum is directly studied at temperatures between 1400 and 2000 K (i.e., at process temperatures) for the first time. The balance of carbon atoms in the system is determined. The difference in the activation energies of carbon dissolution and precipitation, $\Delta E = E_{s1} - E_{1s}$, is found for the case when the diffusion fluxes of dissolved and precipitated carbon atoms are in equilibrium. This difference defines the enrichment of the surface by carbon relative to the bulk. The experimentally found activation energy of carbon dissolution is E_{s1} = 3.9 eV. The activation energy of carbon precipitation is estimated at E_{1s} = 1.9 eV. The latter value is close to the energy of bulk diffusion of carbon in molybdenum. © *2002 MAIK "Nauka/Interperiodica".*

Particle diffusion in solids has been the subject of much investigation [1–4]. However, atom diffusion between the volume and the bulk of a metal has escaped the attention of researchers. This kind of diffusion is of importance in physical and chemical processes taking place on heated solid surfaces, because relations between the dissolution and precipitation of impurities must be taken into consideration in this case. Fields where diffusion between surface and bulk plays a great part are heterogeneous catalysis, metallurgy, semiconductor electronics, corrosion of metals, etc. Yet, we are familiar only with one work devoted to this issue [5], where the dissolution of carbon on the (100)W surface was studied on a quantitative basis. The authors of [5] have established the extremely intriguing fact: the energy of activation of carbon dissolution is high, 5.0 eV, and far exceeds the energy of activation of bulk carbon diffusion in tungsten, 2.55 eV. This means that, even knowing the activation energy of bulk diffusion, one cannot comprehensively describe the transport of impurities between the bulk and surface of different metals. Moreover, in some cases of practical significance, surface-to-bulk diffusion may control the overall rate of impurity or dopant diffusion in metals. It is therefore necessary to find the energy parameters of surface-to-bulk and bulk-to-surface transport in solids from direct experiments and tabulate the findings in relevant reference books together with the energy of bulk diffusion.

Unfortunately, work [5] has not been elaborated upon. In review [6], where gas adsorption, desorption, and reactions on metal surfaces are considered, diffusion over the surface, rather than normally to it, is discussed. The aim of this article is to study the diffusion of carbon atoms between the surface and the bulk in thin textured molybdenum tapes. Processes associated with carbon adsorption on the molybdenum surface at temperatures between 300 and 2000 K in a wide range of surface and bulk carbon concentrations have been traced from the very beginning of the adsorption to the complete carbonization (Mo_2C) across the tape in [7, 8].

EXPERIMENTAL

The experiments were carried out under superhigh vacuum conditions ($P \leq 1 \times 10^{-10}$ torr) in a high-resolution ($\Delta E/E \leq 0.1\%$) Auger spectrometer equipped with a prism energy analyzer [9]. Directly heated textured molybdenum tapes measuring 1 × 0.02 × 40 mm with the work function constant over the tape (e_φ = 4.45 eV) were used as substrates. The work function uniformity was verified with a mass spectrometer by thermionic emission methods and by surface ionization of Na and K atoms [10]. The surface and bulk of the tape were purified by ac heating to T = 2200 K in the oxygen atmosphere ($P_{O_2} \sim 10^{-5}$ torr) and in superhigh vacuum. After the purification, only Auger peaks of molybdenum were observed. During the purification, the tapes were so textured that the (100) face was exposed. The accuracy of substrate orientation was as high as ≈99.5%, as follows from the X-ray diffraction data. The mean grain size in the tape measured in a scanning tunnel microscope was ≈20 μm. Both our and literature data [11] indicate that such a treatment of metallic samples can effectively remove nonmetal impurities (C, S, N, O, P, etc.) from both the surface and the bulk.

Temperatures above 900 K were measured with an optical micropyrometer; below 900 K, by extrapolating the dependence of the temperature on tape heating current. The temperature nonuniformity over the tape was less than 10 K within 40 mm. Note that the Auger spectra were taken directly from the heavily heated (up to 2100 K) tapes. This is of fundamental importance in studying equilibrium high-and medium-temperature processes occurring between surface and bulk with the participation of carbon atoms, because excess carbon does not precipitate on the surface in this case (unlike [5], where it precipitated upon cooling).

Carbon was applied on both the working and back surface of the tape from a calibrated carbon atom source [12]. At room temperature, the sticking coefficient of carbon was assumed to equal unity [13]. The surface carbon concentration N_C, that is, the amount of carbon per 1 cm^2 of the surface, was estimated as

$$N_C = \nu_C t,$$

where ν_C is the calibrated atom flux density and t is the deposition time.

Note that carbon desorbs from the refractory metal surfaces only at temperatures above 2000 K [14].

RESULTS

(1) *Carbon Distribution between the Surface and Bulk of the Tape*

If the tape covered by carbon on its working side to the concentration $N_C = 1 \times 10^{16}$ cm^{-2} at 300 K is heated to 1450 K, a major portion of the carbon will dissolve in the bulk. The carbon remaining on the surface with a concentration $N_s = 1 \times 10^{15}$ cm^{-2} forms surface molybdenum carbide MoC [7]. If we deposit the same amount of carbon (1×10^{16} cm^{-2}) on the back side of the tape at 300 K and then heat up the tape to 1450 K, Auger signals of carbon will be observed after a time. The peak is the highest at $t > 40$ s corresponding to surface carbide with $N_s = 1 \times 10^{15}$ cm^{-2} (Fig. 1). Thus, irrespective of the side on which the carbon is applied to a concentration $N_C = 1 \times 10^{16}$ cm^{-2} at 300 K, the carbon concentration on both sides turns out to be equal, $N_s = 1 \times 10^{15}$ cm^{-2}, after heating the tape to 1450 K and the carbon in an amount $N_C = 1 \times 10^{16} – 2 \times 10^{15} = 8 \times 10^{15}$ cm^{-2} dissolves in the bulk of the molybdenum tape. It follows that bulk diffusion in the 20-μm-thick tape proceeds with a high rate at 1450 K.

We will show that our results for bulk diffusion are consistent with those available in the literature. The coefficient of bulk diffusion of carbon in molybdenum is given by [15]

$$D[\mathrm{cm}^2/\mathrm{s}] = 3.4 \times 10^{-2} \exp\{-1.78[\mathrm{eV}]11\,600/T\}.$$

For unidirectional diffusion, the diffusion length λ varies with time as $\lambda = (2Dt)^{1/2}$. Putting $\lambda = h = 20$ μm, we find the time it takes for carbon to diffuse from one side of the tape to the other at 1450 K: $t \approx 90$ s, which is in good agreement with our experimental data (Fig. 1).

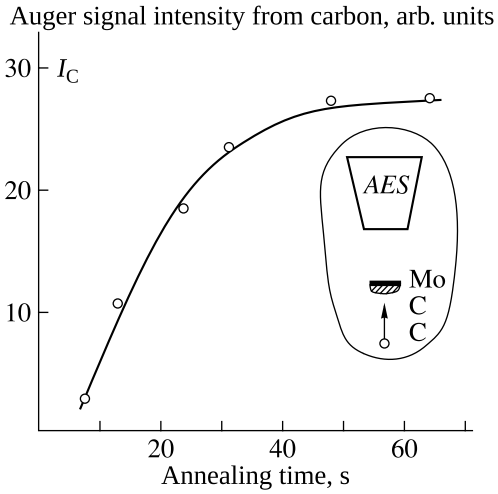


Fig. 1. Auger signal from carbon on the working molybdenum surface vs. time of annealing at 1450 K. Carbon with a concentration $N_C = 1 \times 10^{16}$ cm^{-2} is deposited on the back side of the tape at 300 K. Saturation corresponds to the formation of surface carbide with $N_C = 1 \times 10^{15}$ cm^{-2}.

The use of the calibrated flux makes it possible to find the carbon distribution between the surface and the bulk of the molybdenum tape (see table). To do this, we applied carbon in amounts corresponding to N_C on either of the sides (at 300 K) and drove it in the tape by heating to 1600 K. Then, the tape was slowly (for several minutes) cooled to room temperature and Auger peaks of carbon were recorded. The table lists the carbon concentration on the working side (or back side, since the concentrations are the same) at 300 K and the amount of carbon that passed into the bulk from 1 cm^2 of the surface, $N_b = N_C – 2N_s$. Since the carbon is uniformly distributed over the molybdenum tape, one can calculate its concentration in one layer of interstitials:

$$N_1 = N_b/m,$$

Carbon distribution between the surface and the bulk of molybdenum at different doses N_C

N_C	N_s	N_b	N_1
5×10^{14}	$<1 \times 10^{14}$	$\sim 4 \times 10^{14}$	3×10^{9}
1×10^{15}	2.5×10^{14}	5×10^{14}	3.4×10^{9}
3×10^{15}	7×10^{14}	1.6×10^{15}	1.1×10^{10}
4.5×10^{15}	9×10^{14}	2.7×10^{15}	2×10^{10}
7.5×10^{15}	1×10^{15}	5.5×10^{15}	4×10^{10}
8×10^{16}	1×10^{15}	7.8×10^{16}	5.5×10^{11}

Note: N_s, surface concentration of carbon (on both sides); $N_b = N_C – 2N_s$, its bulk concentration; N_1, carbon concentration in the first subsurface layer. The concentrations are given in atoms/cm^2, and the last two quantities are calculated under the assumption that carbon is uniformly dissolved in the bulk of the substrate.

where $m = 2h/d \sim 1.4 \times 10^5$ is the number of layers of interstitials in the tape and $d = 3.06$ Å is the lattice constant in the [100] direction [16]. The factor 2 takes account of the fact that the unit cell of the bcc lattice has two equivalent layers of interstitials that are parallel to the (100) face.

The calculated value of N_1 is also given in the table. At small concentrations of the carbon deposited, $N_C < 5 \times 10^{14}$ cm^{-2}, its major portion is in the bulk, so that $N_s < 1 \times 10^{14}$ cm^{-2}. For $N_C = 1 \times 10^{14}$ cm^{-2}, the amount of the carbon grows both on the surface (to $N_s = 2.5 \times 10^{14}$ cm^{-2}) and in the bulk ($N_b = 5 \times 10^{14}$ cm^{-2}). At still higher N_C, $N_C = 7.5 \times 10^{15}$ cm^{-2}, the surface concentration becomes maximal, $N_s = N_{s\,max} = 1 \times 10^{15}$ cm^{-2} (that is, MoC carbide forms on the surface), while the bulk concentration of the carbon reaches $N_b = 5.5 \times 10^{15}$ cm^{-2}. It is of interest that a further rise in N_C by one order of magnitude, $N_C = 8 \times 10^{16}$ cm^{-2}, does not increase the surface concentration of the carbon: surface carbide MoC with $N_C = 1 \times 10^{15}$ cm^{-2} is observed on both sides.

(2) Carbon Diffusion into the Metal at Small $N_C < 1 \times 10^{15}$ cm^{-2}

Carbon impurity was carefully removed from the tape, and carbon was applied on the tape to the concentration $N_C < 5 \times 10^{14}$ cm^{-2} at 300 K. Then, the tape temperature was stepwise increased and the kinetics of carbon dissolution in the bulk of the metal was examined with Auger electron spectroscopy. It turned out that up to $T \approx 1250$ K, the surface carbon did not dissolve. The dissolution became noticeable at $T = 1350$ K, and its characteristic time was $t \approx 1$ min. To find the kinetic parameters that characterize the diffusion of the surface carbon into the bulk, we will take advantage of the results obtained in [17], where the similar problem was solved. A decrease in the surface carbon concentration, dN_s, for a time dt is related to the dissolution flux ν_{s1} and the mean lifetime τ_{s1} of the carbon adatoms to dissolution as

$$dN = -\nu_{s1}dt = -Ndt/\tau_{s1}. \quad (1)$$

Integrating (1) when the initial surface concentration N_0 drops to N for the dissolution time t yields

$$N = N_0 \exp\{-t/\tau_{s1}\}. \quad (2)$$

From (2), one can estimate the energy of activation E_{s1} of carbon diffusion into the molybdenum volume:

$$E_{s1} = kT \ln\{[t \ln(N_0/N)]/\tau_0\}. \quad (3)$$

Putting $T = 1350$ K and $\tau_0 = 10^{-13}$ s, we obtain $E_{s1} = 3.9$ eV from the initial portion of the dissolution curves. Such a value of this energy for (100)Mo is unexpectedly too high, exceeding the energy of activation of bulk diffusion, $E_0 = 1.78$ eV [15], by $\Delta E \approx 2$ eV. Note that E_{s1} appreciably exceeds E_0 for all surface carbon coatings on molybdenum if $N_s < 1 \times 10^{15}$ cm^{-2}.

As was mentioned, the great difference between E_{s1} and E_0 has been reported in [5] for the (100)W–C system and in our work [18] for the (100)Ta–C system. We believe that this effect is typical of many Me–C systems.

(3) Diffusion of Surface Carbon into the Metal at $N_C > 1 \times 10^{15}$ cm^{-2}

Experiments show that surface carbide MoC offers good thermal stability. The Auger signal of the carbon taken directly from the heated surface carbide decreases only at $T > 1400$ K. We were interested in the kinetics of the dissolution of the surface carbon in the metal at concentrations higher than 1×10^{15} cm^{-2}. For this purpose, pure molybdenum was covered by a carbon film with a concentration of 2×10^{16} cm^{-2} at 300 K. Then, heating the sample stepwise, we followed the dissolution of the surface carbon using the Auger technique. Of great interest is the fact that the excess carbon (i.e., the amount of carbon exceeding that needed for the formation of the surface carbide) completely dissolves at 900–1000 K, while the surface carbide remains intact up to $T = 1400$ K. If carbon of concentration $N_s = 5 \times 10^{15}$ cm^{-2} is applied on the previously formed surface carbide MoC at 300 K, the excess carbon also dissolves in the tape upon heating to 900–1000 K.

It appears that there exist two types of adsorption sites on the (100)Mo surface, which greatly differ in the energy of dissolution activation. The surface carbide originates at sites with strong bonds, and carbon atoms are arranged in “depressions” between four molybdenum atoms. When all such sites are occupied (which corresponds to a surface carbon concentration of 1×10^{15} cm^{-2}), sites with weak bonds come into play and the temperature threshold of marked carbon diffusion into the metal shifts by ≈400 K toward lower temperatures. For $T = 900$ K and the dissolution time $t \approx 1$ min, we estimated the activation energy E_{s1} of the dissolution of the carbon occupying the weakly bonding sites. Using the method described in the previous section and again putting $\tau_0 = 10^{-13}$ s, we found $E_{s1} \approx 2.5$ eV.

Now let us discuss the nature of the weakly bonding sites. It is easy to evaluate that the area under all carbon atoms in the surface carbide is insignificant. Indeed, if the diameter of a carbon atom is $d = 1.4$ Å, the total area occupied by the carbon atoms entering into MoC amounts to $S = (\pi d^2/4) \times 10^{15}$ Å$^2 \approx 1.6 \times 10^{15}$ Å^2 per 1 cm^2 or ≈16% of the total surface area of the metal. Therefore, if the carbon of the surface carbide occupies all depressions between four surface atoms of molybdenum (from these depressions, the carbon diffuses into the bulk with the higher energy of activation, $E_{s1} \approx 3.9$ eV), carbon atoms covering the surface carbide are adsorbed around the strongly bonding sites, from which they diffuse into the metal with the lower energy of activation, $E_{s1} \approx 2.5$ eV. Note that the strong dependence of the activation energy of carbon dissolution on the carbon concentration on the molybdenum surface, which is observed in our work, is discovered for the first time. It seems likely that such a dependence is typical of other metal–carbon systems and of other adsorbates. In particular, different adsorption sites have been also discovered in (100)W–Si [19] and (100)Mo–Si [20] systems: the silicon of surface silicide remains on the surface up to high temperatures (T = 1400–1500 K), whereas the silicon applied on the surface silicide readily dissolves in the metal at 900–1000 K.

(4) Equilibrium Diffusion of Carbon Atoms between the Surface and the Bulk

Consider the results of experiments performed under equilibrium conditions, i.e., when the carbon atom flux from the surface, v_{s1}, equals the flux from the nearest subsurface layer to the surface, v_{1s}. It is of interest which information on the (100)Mo–C system can be gained from these experiments. Carbon with a concentration $N_C = 1.5 \times 10^{16}$ cm^{-2} was applied on the working side of the tape at 300 K. Then, it was driven in the metal by heating to 2000 K. After cooling to 300 K, both sides of the tape exhibited the surface carbide with $N = 1.5 \times 10^{15}$ cm^{-2}, while the concentration of the dissolved carbon in the bulk per 1 cm^2 was found to be $N_b = 13 \times 10^{15}$ cm^{-2}. In other words, the carbon concentration per layer of interstitials was

$$N_1 = N_b/m = 13 \times 10^{15}/1.4 \times 10^5 \approx 1 \times 10^{11} \text{ cm}^{-2}.$$

Next, the tape temperature was increased stepwise from 300 to 2000 K, and the Auger signal directly from the heavily heated sample was recorded. The temperature dependence of the carbon concentration on the surface normalized to the carbon concentration in the surface carbide with $N_s = 1 \times 10^{15}$ cm^{-2} ($\Theta = 1$) is shown in Fig. 2 (curve *1*). Up to 1400 K, the surface concentration of the carbon remains constant and then the carbon of the surface carbide starts to dissolve. At 1620 K, half the surface carbon is in the dissolved state. The bulk concentration of the carbon is low ($N_1 \approx 10^{11}$ cm^{-2}), so that this portion of carbon makes a minor contribution to the Auger signal. With a further increase in the temperature, the surface concentration of the carbon continues to fall ($N_s \approx 5 \times 10^{13}$ cm^{-2} at 1900 K). The curve $\Theta = f(T)$ is well reproduced with a temperature rise or fall. Also, no time changes in the surface concentration were observed. The steady states in our experiments are apparently associated with equilibrium conditions, under which the surface–bulk mass transfer of the carbon is absent; that is, the oppositely directed diffusion fluxes become equal to each other $V_{1s} = V_{s1}$. The great difference in the carbon concentration on the surface (10^{14}–10^{15} cm^{-2}) and in the near-surface layer of interstitials (~10^{11} cm^{-2}) in equilibrium stands out. This may occur if the rate of carbon precipitation far exceeds that of dissolution; that is, $E_{1s} \ll E_{s1}$ may be expected in this case.

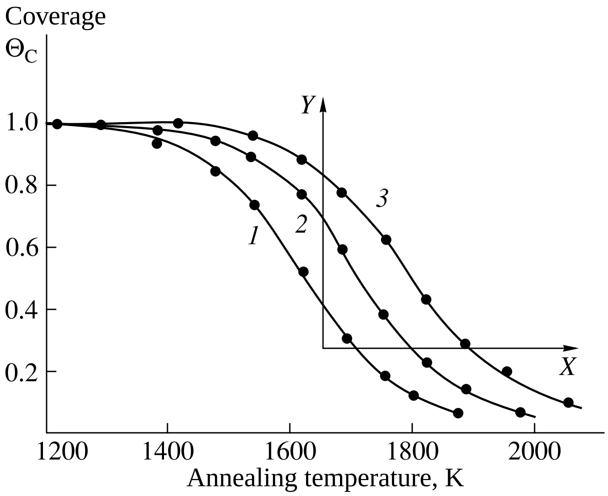


Fig. 2. Equilibrium coverage of the (100)Mo surface by carbon vs. substrate temperature. The concentration of carbon atoms in the layer of interstitials is (*1*) 1.0×10^{11}, (*2*) 2×10^{11}, and (*3*) 4×10^{11} cm^{-2}. $\Theta = 1$ corresponds to the carbon concentration in the surface carbide ($N_s = 1 \times 10^{15}$ cm^{-2}).

If another portion of carbon with $N_C = 1.5 \times 10^{16}$ cm^{-2} is now deposited on the working side at 300 K and then driven in the metal by heating to 2000 K, we will obtain a new equilibrium curve $\Theta = f(T)$ (Fig. 2, curve *2*). Since for each given T = const the diffusion flux of the carbon to the surface v_{1s} increases (because the bulk concentration of the carbon rises), so does the equilibrium concentration N_s and the curve as a whole shifts toward higher temperatures, retaining its shape (Fig. 2). Curve *3* in Fig. 2 corresponds to a new, still greater increase in the bulk concentration of the carbon.

Let us write expressions for the dissolution, v_{s1}, and precipitation, v_{1s}, fluxes when the surface and bulk of the tape exchange particles [4]:

$$\begin{aligned} v_{s1} &= N_s^*(1 - N_1/N_{1m})W_{s1} \\ &= N_s(1 - N_1/N_{1m})C_{s1}\exp\{-E_{s1}/kT\}, \end{aligned} \tag{4}$$

$$v_{1s} = N_1 W_{1s} = N_1 C_{1s} \exp\{-E_{1s}/kT\}(1 - N_s/N_{sm}). \tag{5}$$

Here, N_s and N_{sm} are the instantaneous and highest possible surface concentrations of the atoms, respectively; N_1 is the atom concentration in the layer of interstitials that is the closest to the surface; N_{1m} is the concentration corresponding to the limiting solubility; W_{s1} and W_{1s} are the dissolution and precipitation probabilities, respectively; E_{s1} and E_{1s} are the energies of dissolution and precipitation, respectively; and C_{s1} and C_{1s} are preexponentials. The factor $(1 - N_1/N_{1m})$ in Eq. (4) reflects

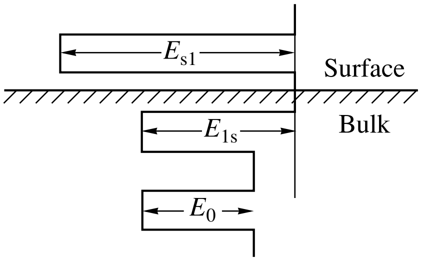


Fig. 3. Carbon atom energy on the surface and in the bulk as a function of distance normal to the surface.

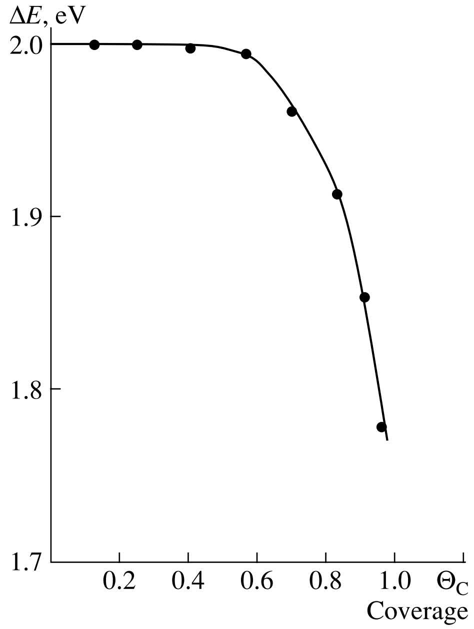


Fig. 4. Dependence $\Delta E = E_{s1} - E_{1s} = f(\Theta)$ for the C/(100)Mo adsorption system when the diffusion fluxes from and to the surface are in equilibrium. $\Theta_C = 1$ corresponds to surface carbide with $N_s = 1 \times 10^{15}$ cm^{-2}.

the fact that the dissolution flux of the carbon into the metal vanishes when the limiting solubility of the carbon in the metal is reached.

The factor $(1 - N_s/N_{sm})$ in Eq. (5), the so-called Langmuir factor, calls for special discussion. The appearance of this factor in this equation is basically due to historical reasons. Its primary goal is to show that the surface can receive a limited number of adatoms (no more than N_{sm}). As applied to precipitation, its meaning is not quite clear; it is anyway obvious that N_{sm} is much greater than the concentration of the carbon in the surface carbide, 1×10^{15} cm^{-2}. Therefore, if the carbon concentration is equal to or lower than that in the surface carbide, this factor can be ignored in calculations, since it is close to unity and its contribution is negligible.

When studying the interaction of carbon atoms with the heated (111)Pt surface, we found [21] that the surface carbon entirely dissolves in platinum if $N_1 < N_{1m}(T)$. At $N_1 = N_{1m}(T)$, the dissolution of the surface carbon completely ceases ($\nu_{s1} = 0$) and newly deposited carbon is accumulated in the layer adsorbed. It is interesting that if the carbon occupies even a small fraction, $\sim 10^{-5}$, of interstitials in the platinum, its diffusion into the bulk is terminated. The results for the (100)Mo–C system were processed for $N_1 \ll N_{1m}$; hence, the factor $(1 - N_1/N_{1m})$ could be neglected.

The mass transfer between the surface and bulk of a metal specifies the energy of dissolution activation E_{s1} and the energy of precipitation activation E_{1s}. In general, both may differ from E_0, the energy of activation of bulk diffusion (Fig. 3). We will try to determine the difference $\Delta E = E_{s1} - E_{1s}$ in the activation energies from the carbon equilibrium distribution in the molybdenum tapes with various carbon content. To do this, we cut the family of the curves in Fig. 2 at N_s = const (*X* section). Then, as follows from Eqs. (4) and (5) for the case $\nu_{1s} = \nu_{s1}$ and N_s = const,

$$\ln N_1(T) = \ln\{N_s C_{1s}/C_{s1}\} - \Delta E/kT. \qquad (6)$$

From the slope of the straight line $\ln N_1(T) = f(1/kT)$, we found ΔE for each of the sections N_s = const. In the surface coverage interval $0 < \Theta < 0.5$, ΔE = const = 2.0 ± 0.1 eV. At $\Theta > 0.5$, ΔE slightly diminishes, reaching 1.8 eV at $\Theta = 0.95$ (Fig. 4). Knowing ΔE, one finds from (6) that $C_{s1}/C_{1s} = 380$.

Of interest is the fact that at $\Theta < 0.5$ (Fig. 2), the surface coverage at T = const is directly proportional to the amount of the carbon dissolved in the bulk (*Y* section), of which the constancy of ΔE in this coverage range is an indication. For $\Theta > 0.5$, ΔE drops and the surface coverage grows much more slowly compared with the bulk concentration of the carbon (Fig. 2). It should be stressed that the data in Fig. 4 refer to surface carbon concentrations $N_s < 1 \times 10^{15}$ cm^{-2} ($\Theta < 1$), when the carbon is in the adsorbed layer and enters into the carbide, so that E_{s1} involved in the relationship $\Delta E = E_{s1} - E_{1s}$ is the energy of dissolution activation for the carbon in the surface carbide.

Let us consider to what extent the value $\Delta E = 2.0$ eV found for $\Theta < 0.5$ ($E_{s1} = 3.9$ eV in this case) is adequate. We have $E_{1s} = E_{s1} - \Delta E = 3.9 - 2.0 = 1.9$ eV, which is close to the energy of activation of carbon bulk diffusion in molybdenum. Consequently, in our Mo–C system, the activation energy of carbon precipitation from the nearest subsurface layer is close to the activation energy of bulk diffusion $E_0 = 1.78$ eV [15].

The question arises as to why the surface coverage equals $N_s = 1 \times 10^{15}$ cm^{-2} (MoC surface carbide), being independent of the temperature and the carbon concentration in the bulk of the molybdenum at $T < 1400$ K (Fig. 2). It might be expected that at $T < 1400$ K, the bulk diffusion is “frozen,” preventing the carbon from being accumulated in the adsorbed layer over the surface carbide. This, however, contradicts both to the experimental data (carbon atoms rapidly leave the surface for the bulk through the surface carbide at lower temperatures, 900–1000 K) and to the calculation (as follows from the available literature data [15], freezing the diffusion processes is out of the question). On the other hand, extrapolating the known data for the limiting solubility of carbon in molybdenum at $T > 1800$ K [22] to lower temperatures, one finds that the volume concentrations of carbon found experimentally are close to the limiting values at 900–1400 K. Direct experiments have shown that the deposition of carbon at 1000 K results in the formation of bulk molybdenum carbide when the carbon interplanar concentration reaches $N_1 \approx 7 \times 10^{11}$ cm^{-2}, or ≈0.07 at. %.

The closeness of the carbon concentration to the solubility limit may alter the kinetic and energy parameters of the bulk diffusion of carbon in molybdenum. Such effects are known and have been reported in the literature [23]. For example, we demonstrated with the C/(10–10)Re system [24] that the activation energy of the bulk diffusion of carbon significantly grows when approaching the saturated Re–C solid solution and the diffusion is frozen at $T < 1100$ K. In pure rhenium, however, carbon quickly migrates in the bulk of the metal at T = 800–900 K. One can expect that the activation energy of the bulk diffusion of carbon in the Mo–C system at $T < 1400$ K exceeds $E_0 = 1.78$ eV, since the bulk concentration of the carbon dissolved approaches the solubility limit.

To clarify the situation, we carried out the following experiment. First, the solid solution of carbon in molybdenum corresponding to curve *1* in Fig. 2 was produced. Then, the temperature of the tape covered by the MoC surface carbide was lowered to the room value and oxygen was adsorbed nearly to saturation ($P_{O_2} \approx 1 \times 10^{-7}$ torr, $t \approx 30$ s). Once the tape had been heated to 1000 K, we observed the simultaneous decrease in the Auger peaks for oxygen and carbon. Since oxygen remains on the surface of pure carbon-free molybdenum up to high temperatures, $T \geq 1800$ K, the simultaneous reduction of the intensity of the Auger peaks for O and C can be explained by the production of CO or CO_2 molecules and their low-temperature desorption. In this way, we markedly decreased the surface concentration of carbon from 1×10^{15} to 2×10^{14} cm^{-2} without changing the carbon concentration in the volume. Then, we raised the temperature to 1100–1200 K and observed the precipitation of the carbon on the surface up to the formation of the MoC surface carbide. The early portions of the precipitation curves taken by the AES method are shown in Fig. 5. First, it is seen that the bulk diffusion is not frozen. Second, the early portions are linear, indicating that the carbon precipitation flux is constant (ν_{1s} = const); in other words, the depletion of the near-surface region is compensated for by the rapid delivery of the carbon from the interior. Third, using expression (5) for the flux, we determined the activation energy of carbon precipitation for three temperatures, which was found to be $E_0 = E_{1s} = 2.5 \pm 0.2$ eV. This value seems to be close to the activation energy of the bulk diffusion. Knowing E_{1s} and the value $N_1 = 5 \times 10^{11}$ cm^{-2}, one can estimate from (5) the preexponential: $C_{1s} = 2 \times 10^{12}$ s^{-1}.

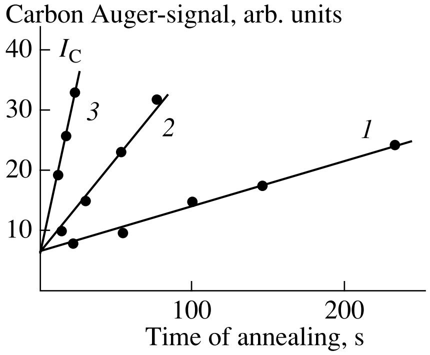


Fig. 5. Temperature dependence of the Auger signal intensity from carbon for the carbonized Mo tape ($N_C = 5 \times 10^{16}$ cm^{-2}): (*1*) 1100, (*2*) 1160, and (*3*) 1210 K. Prior to annealing, the surface carbide was removed in the oxygen atmosphere. $I_C = 7$ corresponds to residual carbon on the surface.

Thus, our assumption that the diffusion is frozen at $T < 1400$ K is invalid. Therefore, the constancy of the coverage in this temperature range can reasonably be associated with the equalization of the potential barriers for carbon dissolution in and precipitation from the bulk.]; that is, $\Delta E \longrightarrow 0$.

DISCUSSION

The scarcity of reliable literature data on transport processes at the surface–volume interface in metals can be related to several reasons. First, to obtain quantitative results, it is necessary to know the exact concentration of dissolved carbon in a metal and carefully vary it, since even trace amounts of impurities (~10^{-3} at. %) may drastically influence experimental findings. In our work, this difficulty was obviated by properly selecting samples and carbon deposition methods. The use of textured metallic tapes makes it possible to easily and effectively remove impurities, including carbon, from the sample and, in addition, quickly reach uniform equilibrium between the surface carbon and carbon in the bulk. In terms of homogeneity and orientation accuracy, the surface of the tape approaches the single-crystal surface; however, the fabrication of single-crystal tapes is as yet a technological challenge. The use of the absolutely calibrated flux of carbon atoms allows us to completely cover the sample and evaluate the impurity particle balance in the system. For 40-mm-long tapes and a carbon deposition area of 30 mm, the migration of carbon to the cold edges can be neglected. The feasibility of carbon deposition on both sides of the tape is also of great importance. This often made it possible to verify physical concepts put forward.

Surface carbon detection was the basic issue. It was found that the Auger spectra of adsorbed carbon taken directly from the sample heated up to 2000 K provides a better insight into the physics of the processes, since the high-temperature state of the surface is affected by high-rate precipitation when the heating current is switched off.

CONCLUSION

The direct study of high-temperature carbon diffusion between the (100) surface and the bulk of molybdenum was performed for the first time. The kinetic parameters of carbon transport from the surface (dissolution) and to the surface (precipitation) were determined. The nature of thermally stable molybdenum carbide on the surface at $T < 1400$ K was discussed. It was inferred that its thermal stability is apparently associated with the equalization of the activation energies for both processes. It should be taken into account that a considerable difference between the activation energies of bulk diffusion (≈1.8 eV) and dissolution (≈3.9 eV) may cause great (by many orders of magnitude) differences in the rates of diffusion processes. This is of importance in metallurgy, heterogeneous catalysis, and other applications. These values, along with the kinetic parameters for bulk diffusion, must undeniably be tabulated in reference books on metal physics.

ACKNOWLEDGMENTS

This work was supported by the Program “Atomic Surface Structures” of the Ministry of Science of the Russian Federation (project no. 4.6.99).

REFERENCES

1. R. M. Barrer, *Diffusion in and Through Solids* (Cambridge Univ. Press, Cambridge, 1941).
2. P. G. Shewmon, *Diffusion in Solids* (McGraw-Hill, New York, 1963; Metallurgiya, Moscow, 1966).
3. M. P. Seach, *Practical Surface Analysis by Auger and X-ray Photoelectron Spectroscopy* (Wiley, London, 1983), pp. 273–317.
4. A. Ya. Tontegode, Zh. Tekh. Fiz. **43**, 1042 (1973) [Sov. Phys. Tech. Phys. **18**, 657 (1973)].
5. K. J. Rawlings, S. D. Foulias, and B. J. Hopkins, Surf. Sci. **109**, 513 (1981).
6. S. J. Lombard and A. T. Bell, Surf. Sci. Rep. **13**, 1 (1991).
7. A. Ya. Tontegode, M. M. Usufov, E. V. Rut’kov, and N. R. Gall’, Zh. Tekh. Fiz. **62** (10), 148 (1992) [Sov. Phys. Tech. Phys. **37**, 1038 (1992)].
8. E. V. Rut’kov, A. Ya. Tontegode, M. M. Usufov, and N. R. Gall, Appl. Surf. Sci. **78**, 179 (1994).
9. V. N. Ageev, E. V. Rut’kov, A. Ya. Tontegode, and N. A. Kholin, Fiz. Tverd. Tela (Leningrad) **23**, 2248 (1981) [Sov. Phys. Solid State **23**, 1315 (1981)].
10. É. Ya. Zandberg, E. V. Rut’kov, and A. Ya. Tontegode, Zh. Tekh. Fiz. **46**, 2610 (1976) [Sov. Phys. Tech. Phys. **21**, 1541 (1976)].
11. B. M. Shepilevskiĭ and V. G. Glebovskiĭ, Poverkhnost, No. 7, 26 (1982).
12. N. R. Gall, E. V. Rut’kov, A. Ya. Tontegode, *et al.*, Chem. Vap. Deposition **5**, 72 (1997).
13. J. H. Leck, *Pressure Measurement in Vacuum Systems* (Chapman and Hall, London, 1964), pp. 70–100.
14. A. Ya. Tontegode, Prog. Surf. Sci. **38**, 201 (1991).
15. P. S. Rudman, Trans. Metall. Soc. AIME **239**, 1949 (1967).
16. C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1975; Nauka, Moscow, 1978).
17. A. Ya. Tontegode and F. K. Yusifov, Poverkhnost, No. 4, 20 (1994).
18. N. R. Gall, E. V. Rut’kov, and A. Ya. Tontegode, Surf. Sci. **472**, 187 (2001).
19. V. N. Ageev, E. Yu. Afanas’eva, N. R. Gall’, *et al.*, Poverkhnost, No. 5, 7 (1987).
20. N. R. Gall, E. V. Rut’kov, A. Ya. Tontegode, and M. M. Usifov, Phys. Low-Dimens. Semicond. Struct. **9/10**, 137 (2000).
21. E. V. Rut’kov and A. Ya. Tontegode, Fiz. Tverd. Tela (St. Petersburg) **38**, 635 (1996) [Phys. Solid State **38**, 351 (1996)].
22. E. Fromm and E. Gebhardt, *Gase und Kohlenstoff in Metallen* (Springer-Verlag, Berlin, 1976; Metallurgiya, Moscow, 1980).
23. G. Schulze, *Metallphysik* (Akademic, Berlin, 1967; Mir, Moscow, 1971).
24. N. R. Gall, S. N. Mikhailov, E. V. Rut’kov, and A. Ya. Tontegode, Surf. Sci. **191**, 185 (1987).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 4, 2002, pp. 491–494. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 120–123.
Original Russian Text Copyright © 2002 by Belyaev, Rubets, Nuzhdin, Kalinkin.

BRIEF COMMUNICATIONS

Effect of Thermodiffusion on Perfection of Crystal Structures Formed by Condensation from Vapor Phase

A. P. Belyaev, V. P. Rubets, M. Yu. Nuzhdin, and I. P. Kalinkin
St. Petersburg State Technological Institute, St. Petersburg, 198013 Russia
e-mail: Belyaev@TU.SPB.RU

Received February 5, 2001

Abstract—Results are reported of an investigation of the effect of nonuniform synthesis conditions stimulating thermodiffusion on the perfection of the crystalline structure of cadmium telluride films synthesized in vacuum by condensation from the vapor phase. Results of technological, geometrical, electorgraphical, and electron-microscopic studies are given. It is shown that the positive effect of the nonuniform conditions on the perfection of the crystalline structure formed has a thresholdlike character. In the case where thermodiffusion takes place, the substrate temperature has an unusual effect on the crystalline perfection of the film obtained. Against expectations, raising the substrate temperature tends to increase the disorder in the structure. It has been found that thermodiffusion affects the duration of the Ostwald ripening phase and therefore provides a means of controlling the structure formation processes. The experimental data are shown to agree with the theory of the film formation processes. *© 2002 MAIK “Nauka/Interperiodica”.*

INTRODUCTION

The latest years have seen an increased interest of researchers and technologists in nonstandard conditions of film synthesis [1, 2]. One of these is obviously synthesis in a graded temperature field. It is reported [3] that such fields improve the crystalline perfection. However, subsequent studies of the film growth mechanisms under noniniform conditions revealed that their influence has a thresholdlike character and is beneficial only under certain regimes. Given below are new research results on the vapor-phase growth processes of cadmium telluride films in a graded temperature field, which show the threshold character of the nonuniform conditions and are consistent with the theory of film growth.

SAMPLES AND EXPERIMENTAL TECHNIQUE

Film formation processes both under uniform (same temperature at all substrate points) and nonuniform (in a graded temperature field) conditions were studied on cadmium telluride films synthesized on muscovite substrates. The film thickness for different samples did not exceed 1 μm. The films were synthesized by a quasi-closed (hot wall) method [3] under a vacuum of $\approx 10^{-3}$ Pa at a substance flux incident onto a substrate of $\Phi = 1.32 \times 10^{16}$ cm^{-2} s^{-1}. The temperature gradient of the thermal field along the substrate (nonuniform synthesis conditions) was produced by heating its central part and cooling the periphery [3]. The substrate was heated by a heater at the center of the substrate holder. The temperature was measured by chromel–alumel thermocouples. Film thickness and its uniformity were measured using a MII-4 microinterferometer accurate to ±0.03 μm. Structural studies were made with an ÉMR-100 electronograph and a PÉM-100 electron microscope. All structural data presented below refer to the central area of the films.

EXPERIMENTAL RESULTS

The effect of the graded temperature field on the perfection of the crystalline structure formed was studied with the use of technological, geometrical, and structural methods. The main results are given in Figs. 1–3.

Figure 1 shows the effect on the film structure perfection of the substrate temperature and the temperature field. Electronograms in Figs. 1a and 1b demonstrate degradation of the crystalline perfection of the films at lower temperatures under uniform conditions. Comparing Figs. 1a and 1b, it can be seen that while at a temperature of $T_s = 523$ K (Fig. 1a), the film structure is close to epitaxial; at $T_s = 473$ K (Fig. 1b), it is already polycrystalline. Electronograms in Figs. 1c–1e refer to nonuniform conditions. The electronograms in Figs. 1c–1e were obtained from films synthesized at temperatures T_s of 473, 490, and 523 K, respectively. It is seen in these electronograms that the influence of the temperature gradient on crystalline perfection can be both beneficial and detrimental. The effect is positive at low substrate temperatures. In the absence of the temperature field at $T_s = 473$ K, the films obtained are polycrystalline (Fig. 1b), whereas in the opposite case the structure formed is close to epitaxial (Fig. 1c). The positive effect is a function of the substrate temperature. From Figs. 1c (473 K) and 1d (490 K), it is seen that at

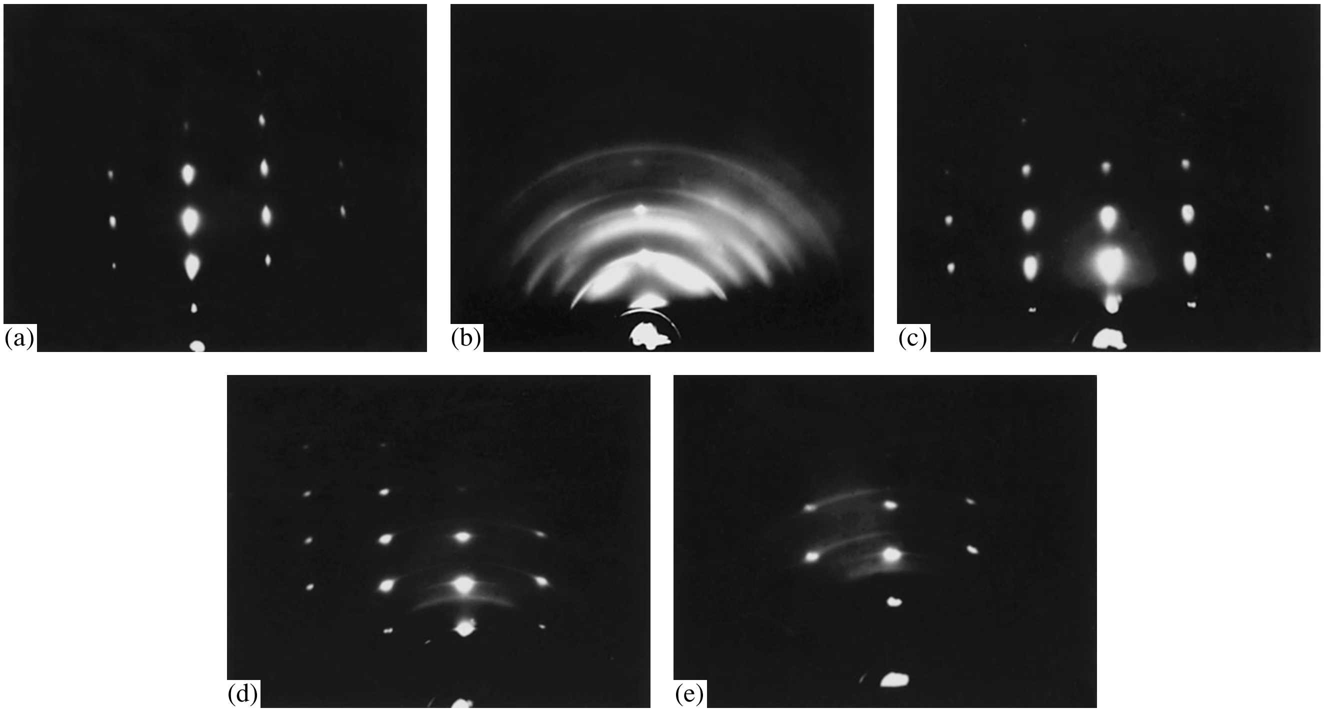


Fig. 1. Electronograms from cadmium telluride films synthesized under uniform conditions (a, b) and under temperature gradient (c–e) at substrate temperatures of T_s = 523 (a, e), 473 (b, c), and 490 K (d).

a higher temperature there is less ordering in the structure.

The effect of nonuniform temperature is negative at high temperatures, as seen in the electronograms in Figs. 1a and 1e. At T_s = 523 K, the films grown with a uniform temperature profile are close to epitaxial (Fig. 1a) and have a significantly worse structure if grown under a temperature gradient (Fig. 1e).

Figure 2 demonstrates the effect of the temperature field on uniformity of the growth-rate distribution along the radial coordinate *x*. Curve *1* refers to growth under uniform conditions at T_s = 473 K; curves *2* and *3* refer to growth in a nonuniform temperature field at T_s = 523 and 473 K, respectively. The figure clearly demonstrates the effect of the nonuniform temperature field on the growth rate, which is lower in the central area and higher at the periphery. The growth rate and its nonuniformity in the radial direction depend on the substrate temperature, both being higher at elevated temperatures.

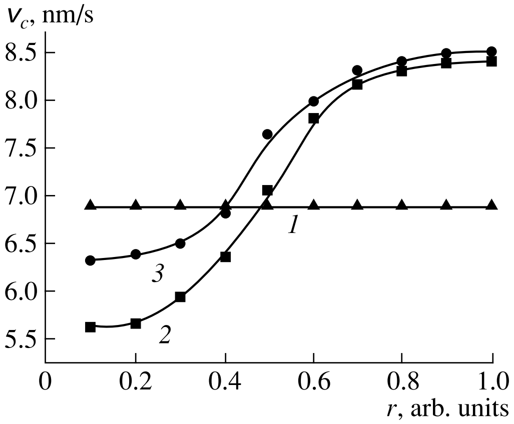


Fig. 2. Dependence on the coordinate of the growth rate of layers of cadmium telluride (the *x* axis originates at the substrate center). Curve *1* refers to growth under uniform conditions at T_s = 473 K; curves *2* and *3* refer to growth in a temperature field at T_s = 523 and 473 K, respectively.

The effect of the temperature field on surface morphology is seen in Fig. 3. Films in Figs. 3a and 3b were grown at T_s = 523 K, and the film in Fig. 3c at 473 K. Electron micrographs in Figs. 3a and 3c show the surfaces of films synthesized under nonuniform conditions and in Fig. 3b, under uniform conditions. From the electron-microscopic studies, it follows that with the temperature field the size of film crystallites is smaller if the substrate temperature is high. At lower substrate temperatures under nonuniform growth conditions, the crystallites grow to a larger size.

RESULTS AND DISCUSSION

According to current concepts, the process of film synthesis from a vapor phase on the surface of a solid consists of several stages. These are nucleation induced by fluctuations, the Ostwald ripening (OR) or coalescence, and merging into a continuous film. Perfection of the crystal structure formed is usually determined by

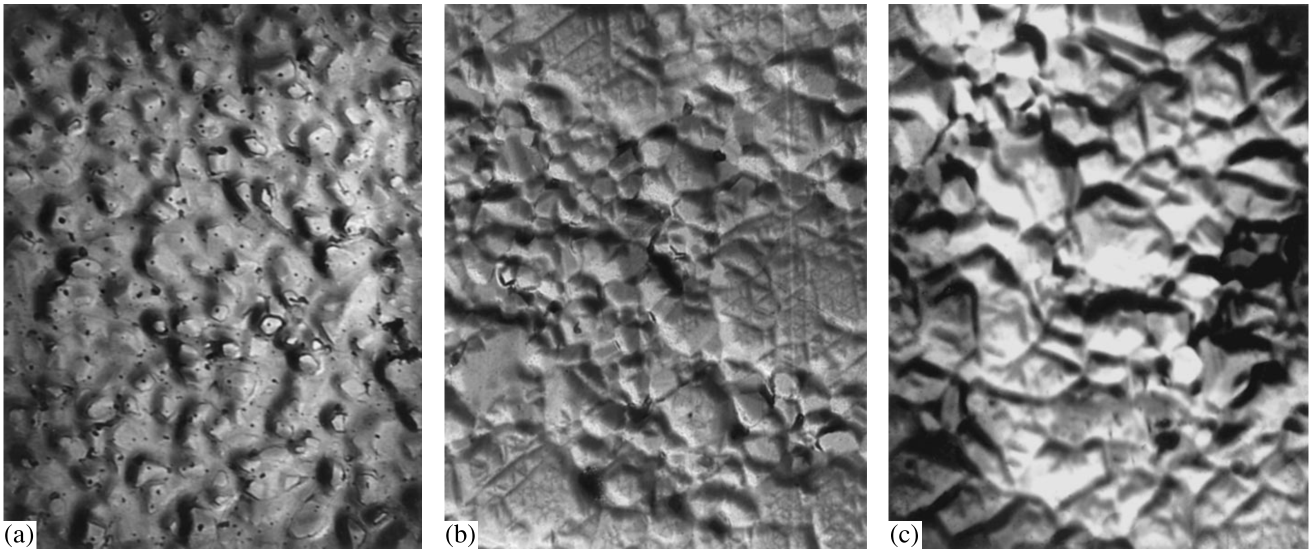


Fig. 3. Surface morphology of cadmium telluride films grown in uniform conditions (b) and with the temperature gradient (a, c) at substrate temperatures $T_s = 523$ (a, b) and 473 K (c) (×2000).

the OR stage [4]. At this stage, by means of the generalized diffusion field, there takes place a correlated redistribution of the atoms of the condensed substance between disperse particles (DP) formed, as a result of nucleation and formation of correlated chemical bonds (incorporation of atoms into the crystal lattice of DPs). Hence, it can be supposed that the deterioration of the crystal structure due to the temperature field at high substrate temperatures (Figs. 1a, 1e), as well as the higher perfection obtained in the case of relatively low temperatures (Figs. 2b, 2c), occur at the OR stage. In this connection, let us consider the effect of temperature on OR in a nonuniform temperature field.

The growth rate of a DP in a generalized field is affected by other DPs and depends on the concentration of the components in the system and on temperature. In particular, if the DP growth rate is limited by incorporation, the expression for the growth rate has the form [5]

$$\vartheta = \frac{2\sigma\beta V_m^2 \psi_1(\Theta)\alpha(\Theta)}{k_B T R}\left(\frac{R}{R_{\text{cr}}} - 1\right), \tag{1}$$

where σ is the surface tension, β is the unit boundary flux onto a DP, V_m is the volume of an atom in the solid phase, $\psi_1(\Theta)$ and $\alpha(\Theta)$ are the parameters accounting for the DP shape, k_B is the Boltzmann constant, T is the temperature, R is the DP radius, and R_{cr} is the critical radius of a DP.

The concentration of components on the substrate surface (entering (1) implicitly through R_{cr}) is determined by the substance fluxes from/to sources/sinks and their initial distribution function over the size space, which, in addition to temperature, are again determined by the substance sources/sinks. The flux from the source in all the above experiments was the same; therefore, in what follows we consider only the sinks of a substance.

In the experiment in question, as shown in [3], the sinks of a substance are formed by a nonuniform temperature field. Their capacity can be estimated from the equation [6]

$$i = -\rho D\left(\nabla c + \frac{k_T}{T}\nabla T\right). \tag{2}$$

Here, i is the substance flux due to thermodiffusion, ρ is the effective density of adatoms on the surface, D is the coefficient of the surface diffusion of atoms, Δc is the adatom concentration, ∇T is the temperature gradient, k_T is the thermodiffusion ratio, and Dk_T is the coefficient of thermodiffusion.

According to [3], at low temperature, sinks of this kind lengthen the OR stage and thus serve to make the film crystal structure more perfect. At higher temperatures, as testified by the experimental results under discussion, the effect of thermodiffusion becomes negative. The reason can now be easily seen from Eqs. (1) and (2). Increasing the temperature decreases both the DP growth rate v, which ensures the correlated rearrangement of atoms in a DP at the OR stage, and the capacitance of substance sinks affecting the OR process duration. This conclusion is further supported by the decrease of the size of crystallites (Figs. 3a, 3c) and increase of the film growth rate (Fig. 2).

In addition, this conclusion helps to explain the negative influence of the sinks on the film structure perfection in comparison with uniform conditions. As mentioned above, at high temperatures, when under uniform conditions the films growth is epitaxial, the introduction of thermodiffusion flows are found to degrade the crystalline structure of the film (compare Figs. 1a and 1e). Evidently, the reason is that, in the presence of thermodiffusion sinks, the atoms at the OR stage do not have enough time to be incorporated into the crystal lattice of a DP and be involved into secondary nucleation, which, naturally, impairs the crystalline perfection of the growing film (the effect predicted theoretically in [1]). Additional evidence of this is provided by morphology studies (Figs. 3a, 3b). Sizes of the crystallites in films grown in the presence of sinks happen to be considerably less than in films grown at the same substrate temperature but under uniform conditions. Therefore, although the theory distinguishes three stages in the process of first-kind phase transformation on the surface of a solid, these stages may, in fact, be concurrent.

CONCLUSIONS

From the results presented above, the following conclusions can be made.

(1) Thermodiffusion sinks, spontaneous or intentional, occurring during vapor-phase growth on the surface of films having nondecaying sources of a substance can have totally different effects on the crystalline perfection of the growing film.

(2) Weak thermodiffusion sinks interfere with the incorporation of atoms into the crystal lattice and produce disorder in the crystalline film growing on the substrate. Higher capacity sinks stimulate formation of a perfect crystalline structure by extending the Ostwald ripening stage and compensating for the negative effect of the sinks on the incorporation process.

(3) During the normal layer-by-layer film growth from the vapor phase with nondecaying sources of substance, both nucleation due to fluctuations and the Ostwald ripening take place.

ACKNOWLEDGMENTS

We are grateful to S.A. Kukushkin for helpful suggestions and discussions. This work was supported by the Russian Foundation for Basic Research, project no. 99-03-32676.

REFERENCES

1. D. A. Grigor’ev and S. A. Kukushkin, Zh. Tekh. Fiz. **68** (7), 111 (1998) [Tech. Phys. **43**, 846 (1998)].
2. A. P. Belyaev, V. P. Rubets, and I. P. Kalinkin, Zh. Tekh. Fiz. **71** (4), 133 (2001) [Tech. Phys. **46**, 495 (2001)].
3. A. P. Belyaev, V. P. Rubets, M. Yu. Nuzhdin, and I. P. Kalinkin, Fiz. Tverd. Tela (St. Petersburg) **43**, 745 (2001) [Phys. Solid State **43**, 778 (2001)].
4. S. A. Kukushkin and A. V. Osipov, Usp. Fiz. Nauk **168**, 1083 (1998) [Phys. Usp. **41**, 983 (1998)].
5. S. A. Kukushkin, Fiz. Tverd. Tela (St. Petersburg) **35**, 1582 (1993) [Phys. Solid State **35**, 797 (1993)].
6. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics* (Nauka, Moscow, 1986; Pergamon, New York, 1987).

Translated by B. Kalinin

Technical Physics, Vol. 47, No. 4, 2002, pp. 495–496. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 124–125.

Shock Wave Strength and Discharge Energy in Gases

K. V. Korytchenko, Yu. Ya. Volkolupov, M. A. Krasnogolovets,
M. A. Ostrizhnoĭ, and V. I. Chumakov

Uskoritel' Research Complex, Kharkov, 61108 Ukraine

e-mail: dovbnya@nik.kharkov.ua

Received April 2, 2001; in final form, September 11, 2001

It is demonstrated that taking advantage of the dynamic properties of a gas stream and applying the technique for directional shock wave formation, one can generate an intense shock wave at low gas discharge energies. In our experiments, we used the setup schematically shown in the figure. It operates as follows. Cooled air from an external source is forced through inlet *5* made in conductor *3* and then through hole *7* into cavity *4* of insulator *1*. The dimensions of outlet *6* in the conductor, inlet *5*, and holes *7* are chosen so that the pressure in cavity *4* of insulator *1* is higher than the ambient pressure. The pressure smoothly drops to the ambient value in and outside channel *6*. The gas flow velocity in channel *6* cannot exceed the critical velocity of sound C_{*2}. Thus, the velocity of disturbance propagation in the gas flow varies from C_0 to C_{*2} (C_0 is the velocity of sound in the stagnant gas flow) [1].

Closing switch *S* causes the capacitive energy storage to discharge. Part of the discharge energy is released as the kinetic energy, which raises the pressure in cavity *4* and gives rise to a compression wave propagating along channel *6*. The incoming gas pressure prevents the gas from flowing towards channel *5*. Since the rate of vibrational-to-kinetic energy conversion is directly proportional to the gas pressure, deactivation processes in the cavity lead to a further rise in the pressure and faster increasing-amplitude compression waves begin to propagate in the somewhat heated gas. This results in the formation of a shock wave. This process is similar to the shock wave formation in detonation tubes [2].

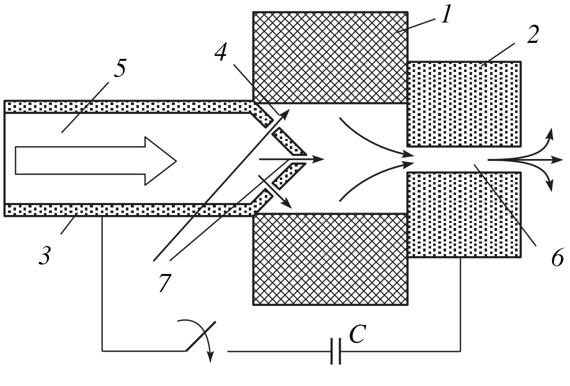
Based on the theory of vibrational deactivation of gas molecules [3], a technique for evaluating the design parameters of the setup described above has been elaborated. It takes into account shock wave formation conditions in detonation tubes and the gas properties.

The design of the setup extends the time of shock wave formation, which allows the more efficient use of the discharge energy through an increase in the duration of the discharge and also due to the fact that the discharge is initiated in the higher-pressure region. The parameters ensuring the formation of the shock wave under the conditions outlined above and estimated for $P_0 = 10^5$ Pa and $T_0 = 293$ K were the following: the temperature $T = 3840$ K, the velocity $D = 2935$ m/s, the pressure $P_1 = 100 \times 10^5$ Pa, and the density $\rho_1/\rho_2 = 6.76$.

The necessary condition for the shock wave generation in denotation tubes is the laminarity of the flow in any outflow mode. As is known from experiments, with disturbances near the inlet eliminated, the flow remains laminar up to the Reynolds number $R = 10^5$, where $R = U_{\max}d/2\nu$, $U_{\max}$ is the velocity of the fluid at the tube axis, d is the diameter of the tube, and ν is the fluid kinetic viscosity. In our case, $U_{\max}$ corresponds to the velocity D of the shock wave front and d, to the diameter of the outlet. The maximal heating temperature T in the insulator cavity can be determined by equating the maximal velocity of the air flow heated to the critical velocity of sound C_{*2}:

$$T = \frac{C_{*2}^2 M(\gamma + 1)}{2\gamma R} = \frac{D^2 M(\gamma + 1)}{2\gamma R},$$

where M is the molar mass of the gas in kg/mol and R is the gas constant.



Setup for the shock wave formation by a gas discharge.

The discharge duration is found from the time t of shock wave formation:

$$t = \frac{l}{C_{01}} - \frac{l}{C_{02}},$$

where l is the outlet length and C_{01} and C_{02} are the velocities of sound in the stagnant flow under the initial conditions and in the gas heated by the electrical discharge, respectively.

For example, if $d = 0.001$ m and $l = 0.0015$ m, we find that $R \approx 10^4$, $T = 11100$ K, and $t = 4 \times 10^{-5}$ s. For comparison, the time of shock wave formation in open air under normal initial conditions is no less than 10^{-5} s.

REFERENCES

1. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics* (Nauka, Moscow, 1986; Pergamon, New York, 1987).
2. F. A. Baum, K. L. Stanyukovich, and B. I. Shekhter, *Physics of Explosion* (Fizmatgiz, Moscow, 1959).
3. Ya. B. Zel'dovich and Yu. P. Raizer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena* (Nauka, Moscow, 1963; Academic, New York, 1966).

Translated by A. Sidorova-Biryukova

Technical Physics, Vol. 47, No. 4, 2002, pp. 497–498. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 126–127.
Original Russian Text Copyright © 2002 by Savintsev, Temrokov.

BRIEF COMMUNICATIONS

On the Surface States in Magnesia and Baria

A. P. Savintsev[1] and A. I. Temrokov[1, 2]

[1] *Kabardino-Balkar State University, ul. Chernyshevskogo 173, Nalchik, 360004 Russia*
e-mail: savinal@mail.ru
[2]*Research Institute of Applied Mathematics and Automation,*
Russian Academy of Sciences, Nalchik, 360000 Russia

Received May 18, 2001

Abstract—The band gap narrowing due to the Tamm states, as determined from the spectroscopy data for magnesia and baria single crystals, are compared with analytical results for the (100) face of these crystals. © *2002 MAIK “Nauka/Interperiodica”.*

It is well known that the interface between a solid and a vacuum or another medium is usually characterized by specific energy states, so-called surface states. When the surface disturbance is sufficiently high, bound surface states, such as Tamm states [1], may appear at the surface. In a number of cases, surface states merge with surface energy bands. As a rule, the energy distribution of surface states is difficult to reliably establish. It is commonly supposed that, in ionic and partially ionic crystals, there exist acceptor surface bands below the conduction band and donor surface bands above the valence band. It is also assumed that the surface bands in these crystals merge with the conduction and valence bands (Fig. 1). Therefore, the presence of surface (Tamm) states in dielectric crystals may alter the band gap at their surface. We compare the predicted effect of Tamm states on the band structure in BaO and MgO crystals with spectroscopic data. There is not much data today on the surface properties of BaO and MgO compounds, having the rock-salt structure [1, 2].

The band gap narrowing due to Tamm states (ΔE) at the dielectric surface was calculated for the (100) crystal face. According to Fig. 1, $\Delta E = E_c + E_2 - E_v - E_1$. The quantum-mechanical calculation of the band gap narrowing for the BaO (100) face gives $\Delta E = 0.6$ eV [3].

In the quasi-classical approximation, the band gap narrowing due to surface states is related to the face specific surface energy σ [3, 4] as

$$\Delta E = 4\sigma/n,$$

where n is the number of particles per unit face surface area.

Setting $\sigma = 957$ mJ/m^2 for the MgO (100) face and the lattice constant to be equal to 0.43 nm [5], we find that $\Delta E = 1.11$ eV.

The results of calculation are compared with UV spectroscopy data for BaO and MgO single crystals with a thickness of 0.38 and 1.2 mm, respectively [2]. The oxides have the crystal lattice of rock-salt type with the (100) cleavage plane, which features the smallest σ [2, 6].

As follows from Fig. 2, the transmission band edge in the BaO and MgO crystals is not coincident with the fundamental absorption edge E_a. The fundamental absorption edge $E_c - E_v$ is 4.0 ± 0.2 and 6 eV in BaO and MgO, respectively [2]. The transmission band edge

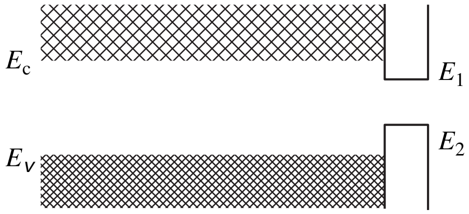


Fig. 1. Bands of ionic surface states [1]: E_c, conduction band edge; E_v, valence band edge; E_1, acceptor surface-state band edge; and E_2, donor surface-state band edge.

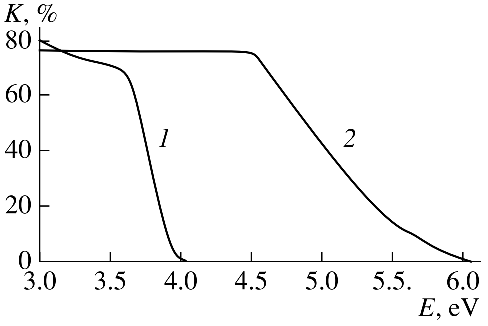


Fig. 2. Transmission of (*1*) BaO and (*2*) MgO single crystals versus photon energy.

$E_1 - E_2$ is 3.5 ± 0.05 and 4.55 ± 0.10 eV in BaO and MgO, respectively. From these values, the band gap narrowing

$$\Delta E = E_c + E_2 - E_v - E_1 = E_a - E_t$$

is estimated at 0.5 ± 0.2 and 1.45 ± 0.10 eV for BaO and MgO, respectively.

Thus, we conclude that (1) the energy difference between the fundamental absorption edge and the transmission band edge can serve as a measure of the band gap narrowing due to surface states and (2) the band gap narrowing calculated for the (100) face of baria and magnesia is in good agreement with the transmission vs. photon energy spectra of these single crystals.

REFERENCES

1. S. Davison and J. Levine, *Surface States* (Academic, London, 1970; Mir, Moscow, 1973).
2. A. A. Vorob'ev, *Physical Properties of Ionic Crystalline Dielectrics* (Tomskiĭ Univ., Tomsk, 1960), Vol. 1.
3. T. M. Taova, A. I. Temrokov, and A. Yu. Kishukov, *Effect of High-Power Energy Fluxes on Substance* (Inst. Vys. Temp. Akad. Nauk, Moscow, 1992), pp. 66–77.
4. A. P. Savintsev and A. I. Temrokov, in *Proceedings of the 15th International Conference "Equations of State," Elbrus, 2000*, p. 25.
5. A. I. Temrokov, Doctoral Dissertation, Leningrad, Leningrad Gos. Univ., 1982.
6. V. D. Kuznetsov, *Surface Energy of Solids* (GITTL, Moscow, 1954).

Translated by A. Sidorova-Biryukova

Technical Physics, Vol. 47, No. 4, 2002, pp. 499–501. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 128–130.
Original Russian Text Copyright © 2002 by Spitsyn.

BRIEF COMMUNICATIONS

A Relationship for the Surface Impedance at the Skin Effect for an Arbitrary Kernel of Integral

A. I. Spitsyn
Kharkov State Technical University of Radio Engineering, Kharkov, 433053 Ukraine

Received June 21, 2001

Abstract—For the case of the skin effect in metals and an arbitrary function $J(R)$ entering into the nonlocal relation between the current density and the electric field, general distributions of the electric field and surface impedance are found. The results obtained are applicable to both normal and superconducting metals and are represented as the Fourier transform of the function $J(R)$. *© 2002 MAIK "Nauka/Interperiodica".*

In determining the surface impedance of both normal metals and superconductors, one uses the nonlocal relation between the current density **j** and the electric field **E** at the skin effect, which generally has the form [1, 3]

$$\mathbf{j}(\mathbf{r}_0) = C_E \int_V \frac{\mathbf{R}(\mathbf{R}\cdot\mathbf{E})}{R^4} J(R)\,dV, \qquad (1)$$

where integration is performed over the whole space, $\mathbf{R} = \mathbf{r} - \mathbf{r}_0$ is the difference between the radius vectors of the point of integration and the point at which the current density is determined, and (…) is the scalar product.

We set $J(0) = 1$ and represent the proportionality coefficients C_E for a harmonically time-varying field ($\sim l^{i\omega t}$) in the form

$$C_E = \frac{3}{4\pi i\omega\mu_0\xi\delta^2}, \qquad (2)$$

where δ is the complex penetration depth, which equals $\delta_{\text{c.l}}$ if the local limit exists [4].

Relationship (2) contains the parameter

$$\xi = \int_0^\infty J(R)\,dR, \qquad (3)$$

which, in the case of normal metals, is equal to $l/(1 + i\omega\tau)$, where l is the mean free path of a conduction electron and τ is the relaxation time [2].

For superconductors (in the stationary case), the parameter ξ coincides with the coherence length introduced into the BCS theory [3].

In the literature, instead of representation (1), one may encounter the relationship between the current density vector **j** and the vector potential **A**:

$$\mathbf{E} = -\frac{\partial\mathbf{A}}{\partial t} = -i\omega\mathbf{A},$$
$$\mathbf{j} = -C_A \int_V \frac{\mathbf{R}(\mathbf{R}\cdot\mathbf{A})}{R^4} J(R)\,dV. \qquad (4)$$

The coefficients C_E and C_A relate as

$$C_A = i\omega C_E. \qquad (5)$$

We also assume that the function $J(R)$ has the asymptotics as $R \longrightarrow \infty$:

$$J(R) \sim l^{-R/\alpha}, \quad \mathrm{Re}\,\alpha > 0. \qquad (6)$$

For normal metals, $\alpha = \xi$; for superconductors in the stationary case, $\alpha = \xi_F$ [5].

Consider a metal bounded by a planar surface S and assume that $\mathbf{E} = \mathbf{E}(z)$. Next, introduce the rectangular coordinate system with its origin on the plane (see figure) and with the z axis directed inward to the metal. Let

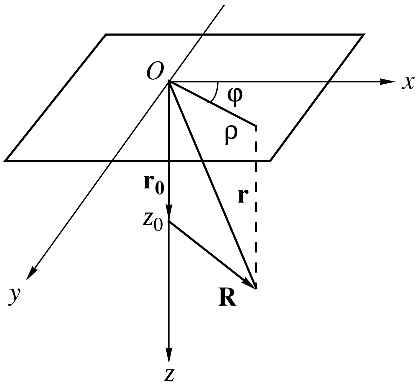


Figure.

1063-7842/02/4704-0499$22.00 © 2002 MAIK "Nauka/Interperiodica"

the x axis be aligned with the vector **E**. Then, from relationship (1), $j_x = j_z = 0$ and

$$j_x = j = C_E \int_V \frac{x^2 E}{R^4} J(R) dV, \quad E_x = E. \tag{7}$$

Now, introduce the cylindrical coordinate system related to the initial one (see figure):

$$x = \rho \cos\varphi, \quad y = \rho \sin\varphi,$$

$$R^2 = \rho^2 + (z - z_0)^2, \quad dV = \rho d\rho d\varphi dz.$$

After introducing the variable $u = R/|z - z_0|$, relationship (7) can be written in the form

$$j(z_0) = \int_{-\infty}^{+\infty} \Omega_E(z - z_0) E(z) dz, \tag{8}$$

$$\begin{aligned}\Omega_E(z) &= \pi C_E \int_1^{\infty} \left(\frac{1}{u} - \frac{1}{u^3}\right) J(u|z|) \\ &= \pi C_E [E_{J1}(|z|) - E_{J3}(|z|)],\end{aligned} \tag{9}$$

where

$$E_{Jn}(z) = \int_1^{\infty} \frac{J(uz)}{u^n} du. \tag{10}$$

With regard for relation (4) between **j** and **A**, expression (8) is replaced by

$$j(z_0) = -\int_{-\infty}^{+\infty} \Omega_A(z - z_0) A(z) dz. \tag{11}$$

In the relation between j and A through the plane kernel of integral Ω_A, we use, as is customary, the minus sign before the integral. It is evident that

$$\Omega_A = i\omega\Omega_E. \tag{12}$$

The corresponding equation for the electric field E inside the metal,

$$\frac{d^2 E}{dz_0^2} - i\omega\mu_0 \int_{-\infty}^{+\infty} \Omega_E(z - z_0) E(z) dz = 0, \tag{13}$$

will be solved for the case of the mirror reflection of electrons from the boundary. In this case, the solution is symmetric about the plane S [4]:

$$E(z_0) = E(-z_0). \tag{14}$$

Setting the boundary conditions in the form

$$\frac{d}{dz} E(z)|_{z=0} = E'(0), \quad E(\infty) = 0$$

and applying the Fourier transformation to the terms of Eq. (13), we find

$$-2E'(0) - k^2 \tilde{E}(k) - i\omega\mu_0 \tilde{\Omega}_E(k) = 0. \tag{15}$$

Here, $\tilde{E}(k)$ and $\tilde{\Omega}_E(k)$ are the Fourier transforms of the functions $E(z)$ and $\Omega_E(z)$:

$$\tilde{E}(k) = \int_{-\infty}^{+\infty} E(z_0) e^{ikz_0} dz_0, \quad \tilde{\Omega}_E(k) = \int_{-\infty}^{+\infty} \Omega_E(z_0) e^{ikz_0} dz_0.$$

From (15),

$$\tilde{E}(k) = -\frac{2E'(0)}{k^2 + i\omega\mu_0 \tilde{\Omega}_E(k)}, \tag{16}$$

which, after the inverse Fourier transformation, leads to the solution

$$E(z_0) = -\frac{2E'(0)}{\pi} \int_0^{\infty} \frac{\cos kz_0 dk}{k^2 + i\omega\mu_0 \tilde{\Omega}_E(k)}. \tag{17}$$

From relationship (9) for $\tilde{\Omega}_E(k)$, we have

$$\tilde{\Omega}_E(k) = \pi C_E \int_1^{\infty} \left(\frac{1}{u^2} - \frac{1}{u^4}\right) \tilde{J}\left(\frac{k}{u}\right) du, \tag{18}$$

where $\tilde{J}(k)$ is the Fourier transform of the function $J(|z|)$.

The surface impedance is given by

$$Z = -\frac{i\omega\mu_0 E(0)}{E'(0)} = \frac{2i\omega\mu_0}{\pi} \int_0^{\infty} \frac{dk}{k^2 + i\omega\mu_0 \tilde{\Omega}_E(k)}, \tag{19}$$

and the complex penetration length $\delta_c = Z/(i\omega\mu_0)$ has the form

$$\delta_c = \frac{2}{\pi} \int_0^{\infty} \frac{dk}{k^2 + i\omega\mu_0 \tilde{\Omega}_E(k)}. \tag{20}$$

In general, the skin depth is defined by the expression [4]

$$\delta_s = \mathrm{Re}\,\delta_c. \tag{21}$$

Now, we will show that the choice of the coefficient C_E in form (2) is in agreement with nonlocal relationship (1). Since the characteristic distance of $\Omega_E(z)$ variation is on the order of $|\xi|$ and $E(z)$ varies over a distance on the order of the skin depth δ_s, then, in the nonlocal limit $|\xi| \ll |\delta_{\text{c.l}}|$ ($\delta_c = \delta_{\text{c.l}}$), the electric field in Eq. (8) can be factored out from the integral:

$$j(z_0) \approx E(z_0) \int_{-\infty}^{+\infty} \Omega_E(z - z_0) dz = \tilde{\Omega}_E(0) E(z_0).$$

Using this relation in Eq. (13), we obtain

$$\delta_{c.l}^2 = \frac{1}{i\omega\mu_0\tilde{\Omega}_E(0)}. \qquad (22)$$

From (18), it follows that

$$\tilde{\Omega}_E(0) = \frac{4}{3}\pi C_E \xi; \qquad (23)$$

consequently, the choice of the coefficient C_E in form (2) corresponds to the limit $|\xi| \ll |\delta_{c.l}|$.

To conclude, we give the expressions for the function $J(R)$ in different cases. When calculating the surface impedance of normal metals [6], one uses the function $J(R)$ from the Reuter–Sondheimer theory: $J(R) = \exp(-R/\xi)$, where $\xi = l/(1 + i\omega\tau)$. In the case of the Pippard kernel in the phenomenological theory of superconductivity, $J(R) = \exp(-R/\xi)$ at the coherence length $\xi = \xi_p > 0$ [7]. In the microscopic theory of superconductivity (the stationary case),

$$J(R) = \frac{4a}{\pi \tanh\frac{\pi}{2a}} \sum_{n=0}^{\infty} \frac{\exp(-R/\xi_n(T, l))}{1 + (2n+1)^2 a^2}, \qquad (24)$$

with the associated relationships for a and ξ_n given in [4, 8]. The terms of series (24) are real and have the same form as in the Pippard theory of superconductivity but different values of the parameter ξ.

REFERENCES

1. D. C. Mattis and J. Bardeen, Phys. Rev. **111**, 412 (1958).
2. G. E. H. Reuter and E. H. Sondheimer, Proc. R. Soc. London, Ser. A **195**, 336 (1948).
3. R. Meservey and B. B. Schwartz, *Superconductivity* (Marcel Dekker, New York, 1969), pp. 17–191.
4. F. F. Mende and A. I. Spitsyn, *Surface Impedance of Superconductors* (Naukova Dumka, Kiev, 1985).
5. J. Halbritter, Z. Phys. **243**, 201 (1971).
6. A. I. Spitsyn, Radiotekh. Élektron. (Moscow) **38**, 2152 (1993).
7. A. B. Pippard, Proc. R. Soc. London, Ser. A **216**, 547 (1953).
8. A. I. Spitsyn, Zh. Tekh. Fiz. **64** (4), 68 (1994) [Tech. Phys. **39**, 385 (1994)].

Translated by Yu. Vishnyakov

Technical Physics, Vol. 47, No. 4, 2002, pp. 502–503. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 131–132.
Original Russian Text Copyright © 2002 by Spitsyn.

BRIEF COMMUNICATIONS

The Possibility of Gaining Information on the Crystallographic Orientation of Cylindrical Samples from Surface-Impedance Measurements at Microwave Frequencies

A. I. Spitsyn
Kharkov State Technical University of Radio Engineering, Kharkov, 433053 Ukraine

Received June 21, 2001

Abstract—From measurements of the Q factors or frequency shifts of H_1 polarized modes in a coaxial cavity information on the crystallographic orientation of a single-crystal sample of which the cavity inner cylinder is made can be derived. The choice of H_1 modes in the coaxial cavity and its sizes at microwave frequencies is optimized. It is found that the determination of the crystallographic orientation of a uniaxial crystal at the H_{121} mode is much more accurate than that at E_1 modes. *© 2002 MAIK “Nauka/Interperiodica”.*

Measurements of the mean surface impedance $\bar{Z} = \bar{R} + i\bar{X}$ of single crystals at microwave frequencies provide a possibility of gaining information on the crystallographic orientation of samples with the anisotropic surface impedance [1]. In [1], on the assumption of the local tensor relation between the electric field and the current density, we calculated the mean surface impedance and its components for cylindrical samples. The calculation was performed for the E_0 and E_1 modes (for example, of a cylindrical or coaxial cavity) with the first subscripts 0 and 1 for the case of a uniaxial crystal. However, the corresponding results for the H modes are more appropriate for determining the crystallographic orientation. In this study, particular numerical results are presented for the H_1 mode in the coaxial cavity where the coaxial is made of a sample under investigation.

In the expansion of the Fourier component of the Z_{xx} tensor in the polar angle φ on the cylindrical of a uniaxial crystal, the coefficient a_2 defines surface the value of $\bar{Z}_H$ - for the H_1 mode and has the form [1]

$$a_2 = \frac{1}{2}(Z_{01} - Z_{03})\sin^2\theta, \qquad (1)$$

where Z_{03} and Z_{01} are the surface impedances with the resistivities equal to the principal values of the resistivity tensor, ρ_3 and ρ_1, in the directions parallel and perpendicular to the C principal axis of the uniaxial crystal, respectively, and θ is the angle between the cylindrical axis Z and the C axis. Note that the ratio $a_2/\bar{Z}_H$ for the H_1 mode may amount to several tens of percent, whereas the corresponding ratio for the E_1 modes is as low as several percent.

In determining the angle θ for a given sample, one can measure the Q factor of the TEM mode or that of any of the E or H modes with the zeroth subscript and then, using the determined value of the partial Q factor for the cylindrical surface, find the angle θ. For this purpose, it is more appropriate to use the TEM or E_0 mode of the coaxial cavity. In the figure, curve *1* represents the ratio $Z_{TEM}(\theta)/Z_{01}$ as a function of θ. This ratio is equal to that of the partial Q factors for the cylindrical surface of the sample, $Q(\pi/2)/Q(\theta)$, or to the ratio of the frequency shifts in the coaxial cavity with respect to the frequency of a cavity with the perfectly conducting walls (the shifts are due to the cylindrical surface), $\Delta\omega(\theta)/\Delta\omega(\pi/2)$. The ratio $\gamma = Z_{01}/Z_{03}$ was taken to be

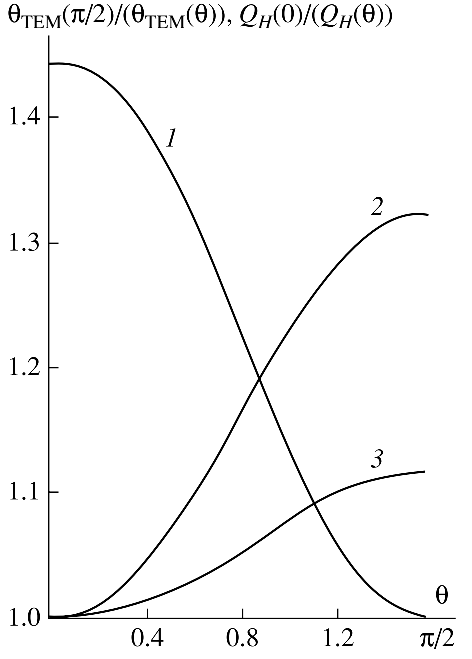


Figure.

1063-7842/02/4704-0502$22.00 © 2002 MAIK “Nauka/Interperiodica”

equal to $1/\sqrt{2}$. This dependence is specified by the sample material.

The direction of the *C* axis can be determined by measuring the *Q* factors (or frequency shifts) of the two orthogonal polarized H_1 modes in the coaxial cavity. From [1], it follows that the mean surface impedance of the cylindrical surface in the case of the *H* modes is represented as

$$\bar{Z}_H = \alpha\bar{Z}_{zz} + \nu\bar{Z}_{xx}, \qquad (2)$$

that is, depends on both the mean component of the surface impedance tensor $\bar{Z}_{zz}$ and $\bar{Z}_{xx}$. Since the maximum difference between $\bar{Z}_{zz}$ and $\bar{Z}_{\mathrm{TEM}}$ is much less than between $\bar{Z}_{xx}$ and $\bar{Z}_0$, where $\bar{Z}_0$ is the mean surface impedance of the cylinder measured for the H_0 mode, it is desirable that $\alpha/\nu \ll 1$. For the H_{1ml} mode,

$$\frac{\alpha}{\nu} = \left(\frac{\pi l R_2}{h}\right)^2 \frac{1}{p^2\mu_{1m}^4(p)}, \quad p = \frac{R_1}{R_2}. \qquad (3)$$

Here, R_1 and R_2 are the inner and outer radii of the coaxial cavity, respectively, and μ_{1m} is the *m*th root of the equation

$$J_1'(\mu_{1m})N_1'(p\mu_{1m}) - J_1'(p\mu_{1m})N_1'(\mu_{1m}) = 0,$$

where J_1 and N_1 are the Bessel and Neumann functions, respectively.

In order that the ratio α/ν be small, it is necessary to take the mode subscript $l = 1$. If the cavity sizes are such that $\pi R_2/h \le 1$, the mode subscript is $m = 2$, and $p = 1/2$ (hence, the root $\mu_{12} = 6.565$), this ratio in the 3-cm range will be $\sim 10^{-3}$. In this case, the term $\alpha\bar{Z}_{zz}$ in relationship (2) can be neglected and it is possible to consider that $\bar{Z}_H$, in our case, depends only on $\bar{Z}_{xx}$.

As was shown in [2], in the absence of irregularities in a coaxial cavity, the anisotropy is the main factor that influences the orientation of the polarized modes. Therefore, the resulting polarized modes, as in the case of planar samples [3, 4], will be oriented in such a way that one of them has the polarization direction along the projection of the *C* axis onto the plane perpendicular to the cylindrical axis of symmetry, and the other is oriented perpendicularly to this projection. Consequently, knowing the polarization directions of the steady-state polarized modes, one can find the plane where the *C* axis lies. Thus, the direction of the *C* axis of a uniaxial cylindrical crystal is defined by the two directions. The figure (curves *2* and *3*) shows the θ dependences of the ratio $Q(0)/Q_{1,2}(\theta)$ for two polarized modes H_{121} in the above coaxial cavity when the ratio $Z_{01}/Z_{03} = 1/\sqrt{2}$.

In a coaxial cavity, there always exist irregularities (holes, pins, etc.), which, together with anisotropy, affect the resulting directions of polarized modes. If the orientations of polarized modes are specified in the main by irregularities, there additionally appears the possibility of changing the directions of mode polarizations in the single-crystal sample, for example, by rotating the sample about the *z* axis in the presence of the irregularities. In this case, the ratio of the partial *Q* factors of the polarized modes is [1, 2]

$$\frac{Q(0)}{Q_{1,2}(\theta)} = \frac{\bar{Z}_0}{Z_{01}} \pm \frac{1}{2}\frac{a_2}{Z_{01}}\cos 2\varphi_0, \qquad (4)$$

where φ_0 is the angle between the polarization direction of one of the polarized modes and the projection of the crystal *C* axis. At $\varphi_0 = 0$ or $\pi/2$, the ratio $Q(0)/Q_{1,2}(\theta)$ for one of the polarized modes will be maximal and for the other, minimal.

Attaining the maximum and minimum *Q* factors of the polarized modes by rotating the cylindrical sample about the *z* axis, one can find the plane containing the *C* axis from the directions of their polarizations. Note that the values of Q_1 and Q_2 for the resulting polarized modes can also be used for the independent determination of the angle θ.

Such a procedure of gaining information on the crystallographic orientation of cylindrical samples is also applicable in the general case of triaxial crystals.

REFERENCES

1. A. I. Spitsyn, Zh. Tekh. Fiz. **64** (11), 105 (1994) [Tech. Phys. **39**, 1146 (1994)].
2. A. I. Spitsyn, Available from the State Scientific and Technical Library of Ukraine No. 653-Uk (Khar'kov, 1996), p. 97.
3. F. F. Mende, A. I. Spitsyn, A. V. Skugarevski, and L. A. Maslova, Cryogenics **25** (2), 92 (1985).
4. F. F. Mende, A. I. Spitsyn, and N. N. Dubrov, Zh. Tekh. Fiz. **50**, 1609 (1980) [Sov. Phys. Tech. Phys. **25**, 938 (1980)].

Translated by Yu. Vishnyakov

Technical Physics, Vol. 47, No. 4, 2002, pp. 504–506. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 133–135.
Original Russian Text Copyright © 2002 by Loskutov, Levitin.

BRIEF COMMUNICATIONS

The Structure and Fatigue Life of Titanium Alloys Processed by Electrical Pulses

S. V. Loskutov and V. V. Levitin
Zaporozh'e State Technical University, Zaporozh'e, 69060 Ukraine
e-mail: svl@zstu.edu.ua

Received August 14, 2001

Abstract—The structure and properties of VT3-1 titanium alloy processed by electrical pulses are investigated. The electric pulse processing improves the fatigue life of titanium specimens, as follows from high-cycle fatigue tests. *© 2002 MAIK "Nauka/Interperiodica".*

The effect of high-power current pulses on the metal structure was considered in [1–4]. It was shown that the electric current of density on the order of 100 MA/m^2 not only thermally affects the metal crystal lattice but also has a specific (electroplastic) effect on lattice defects. For instance, the fatigue life of steels was extended after they had been processed with electric current pulses [1]. The same processing seems to be promising for titanium alloys as well.

VT3-1 and VT8 titanium alloys, containing α and β phases, offer high operating characteristics and especially high strength. The reason is the metastability of the α and particularly β modifications, as well as the complex kinetics of the phase transformations, which are accompanied by the formation of new metastable phases [5]. These transformations proceed in local volumes, and the kinetics of these processes is related to the type and concentrations of alloying elements. The phase composition and the phase transformation kinetics depend on the thermal treatment of the titanium alloys, the value and duration of dynamic and static loads, surface plastic strains, etc. Different hardening treatments cause high (up to 850 MPa) residual compression stresses in the surface layer of parts made of titanium alloys. The residual stresses change the properties of the material that influence the fatigue strength. Experience suggests that the application of high-density current pulses as the final stage of treatment may bring parts subjected to load cycling into a more energy favorable near-equilibrium state. This effect is associated with the enhanced diffusion mobility of atoms in the vicinity of lattice defects because of preferred heat evolution around them.

Our aim was to perform X-ray diffraction investigation into structure changes in the titanium alloys after electric pulse processing and to reveal its effect on the high-cycle fatigue life.

Test specimens (shovels) made of VT3-1 titanium alloy had the working part measuring $2 \times 10 \times 60$ mm. The surface was cold-hardened with small balls vibrating in an ultrasonic field, as well as by vibration treatment. In the process of vibration treatment, the specimens were fixed in a cell containing ceramic granules. The cell was put in vibration motion, and the vibration time and amplitude were varied.

To process the specimens with unipolar current pulses, we used the discharge of a capacitor bank from a T1220 press generating electrohydraulic pulses. The current through the specimen was controlled by varying the length and the thickness of a fine copper wire to be exploded. For the most part, the specimens were processed with current pulses of height $\approx$170 MA/m^2 and duration of 10 μs. The heating temperature was determined with a Chromel–Copel thermocouple elastically pressed against the middle of the specimen. It is obvious that the temperature thus measured is somewhat underestimated. During the processing with current pulses, the temperature was varied in the range between 80 and 250°C. Changes in the surface state of the specimens were estimated by taking the contact potential difference (CPD) distribution.

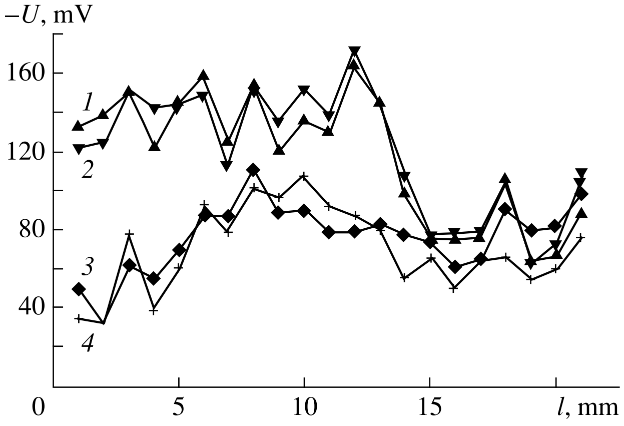
The specimens were investigated before and after the current pulse processing with a DRON-3M computerized diffractometer [6]. The CoK_β monochromatic radiation was employed. Residual macrostresses were measured with the 2θ–$\sin^2\psi$ method, and microstrains and coherent scattering block sizes were approximately found from the diffraction peak broadening.

A small piezoelectric shaker [7] was employed in the high-cycle fatigue tests. The vibration frequency and the dynamic load were varied, and the number of cycles to failure was noted. The critical load at which the specimens fail the test base (the prescribed number of cycles was 2×10^6) was looked for all types of mechanical processing. Each batch comprised six specimens mechanically processed under identical conditions. Then, the specimens were subjected to the electric pulse processing and tested at this critical load to estimate their lifetime. For each of the batches, the mean number of cycles to failure is listed in table. Deviations from the mean were within 30% for all of the specimens.

Oxide films colored from light yellow to brown appear on the surface of the specimens after the electric pulse processing. However, the CPD profile (or the energy relief of the surface) changes insignificantly, as follows from the figure, where the shift of the CPD distribution is shown. Near the clamp (l = 22 mm), where the specimen cross section is the largest and the current passes, the change in the CPD distribution is also minor. Within the operating part, the CPD tends to grow (the electron work function increases), which is typical of oxidizing titanium alloy surfaces.

To find residual stresses in strained metals, one selects X-ray reflections with Bragg angles close to $\pi/2$. The accuracy of this method depends on strain-induced lattice oriented microdistortions, which may violate the linear dependence of the interplanar spacing d on $\sin^2\psi$, where ψ is the angle between the normal to the specimen surface and the normal to a reflecting crystallographic plane. It was found that the plane stress state of the surface layer is improved after the current pulse processing (the plots in the 2θ–$\sin^2\psi$ method become more linear) [8]. Presumably, this means that the defect distribution in the surface layer becomes more uniform and the plane stress state sets up.

The short processing time and relatively low heating temperatures raise the question regarding a mechanism of structure transformations when high-power current pulses disturb the crystal lattice. To reveal structural changes caused by the dc heating and vacuum annealing of the specimens, we carried out a number of experiments. The specimens were heated by direct current to 350°C and annealed at 650°C for 3 h in a vacuum (the residual pressure was 2×10^{-5} torr or below).



CPD distribution over the specimen surface: (*1*) initial, (*2*) in 48 h, (*3*) after single-pulse processing (8 kV, heating to 250°C), and (*4*) in 48 h after the pulse.

In accordance with the X-ray diffraction data, in these specimens, the intensity of the diffraction peaks grows insignificantly. The phase composition remains almost unchanged. However, the residual macrostresses become tensile, the microstrains are reduced, and mosaic blocks increase in size. The heating of the specimens to 350°C in air results in the formation of an oxide film similar to that formed during the electric pulse processing.

Characteristics of the titanium alloy structure and the fatigue test results

Batch number	Processing	Residual macrostresses σ, MPa	Microstrains ε, 10^{-3}	Block size D, nm	Mean number N of cycles to failure, $\times 10^6$
08	Mechanical grinding and annealing	–82	2.5	17	0.1
	The same with subsequent current pulse application (t = 250°C)	–103	1.7	19	0.3
07	Mechanical grinding and annealing	–37	2.8	18	0.5
	The same with subsequent current pulse application (t = 250°C)	–38	2.5	16	1.2
04	Ultrasonic processing with balls 1.9 mm in diameter, 3 min	–530	1.6	13	1.3
	The same with subsequent current pulse application (t = 80°C)	–560	1.5	14	1.8
05	Vibration grinding and polishing, 30 min	–590	1.2	12	1.8
	The same with subsequent current pulse application (t = 225°C)	–360	2.0	20	2.6
06	Vibrational hardening with ceramic granules 1.9 mm in diameter	–320	0.9	14	2.7
	The same with subsequent current pulse application (t = 250°C)	–285	2.4	22	5.0

After the vacuum annealing, the diffraction peak intensity grows, the resolution of closely spaced peaks from different phases is improved, and the residual macrostresses relax. For the initially stressed surfaces, the growth of microstrains and mosaic blocks after the annealing is more significant than in the case of the stress-free surfaces. In the specimens subjected to the ultrasonic hardening and vibration grinding, the annealing reduces the spread of the structure parameters from specimen to specimen and macro- and microstresses become nearly equal in all the specimens.

The measurements and the test results are given in the table. The electric pulse processing is seen to extend the fatigue life of the specimens subjected to the different pretreatments. In the case of the surface-hardened specimens, the electric pulse processing reduces the residual macrostresses. However, the residual microstresses rise and mosaic blocks somewhat increase.

The X-ray diffraction data also point to the fact that the titanium alloys processed by electric pulses change their phase composition. The diffraction peaks $(200)_\beta$ and $(110)_\beta$ alter to the greatest extent. The intensity of the former grows, which may be explained by the $\alpha'' \longrightarrow \beta$ low-temperature transformation [9].

As follows from the high-cycle fatigue tests, the lifetime of the specimens subjected to a single current pulse (when the residual compression stresses are reduced by 100–150 MPa) increases two or three times at the same levels of the dynamic load. Conceivably, the partial reduction of the residual stresses is attended by changes in the structure that improve the fatigue strength of the specimens.

From the experimental data obtained, we can conclude that the titanium specimens subjected to a high-power current pulse undergo structure modifications that are distinguished from those appeared after the similar dc heating or after the vacuum annealing. The plane stress state of the surface layer after the current pulse processing is improved (the plots in the 2θ–$\sin^2\psi$ method become more linear). As a result of the electric pulse processing, the fatigue life of the titanium alloy specimens surface-hardened by different procedures rises.

REFERENCES

1. L. B. Zuev, O. V. Sosnin, S. F. Podboronnikov, *et al.*, Zh. Tekh. Fiz. **70** (3), 24 (2000) [Tech. Phys. **45**, 309 (2000)].
2. A. I. Pinchuk and S. L. Shavreĭ, Fiz. Tverd. Tela (St. Petersburg) **43**, 1416 (2001) [Phys. Solid State **43**, 1476 (2001)].
3. N. E. Kir'yanchev, O. A. Troitskiĭ, and S. A. Klevtsur, Probl. Prochn., No. 5, 101 (1983).
4. N. V. Dubovitskaya, S. M. Zakharov, and L. N. Larikov, Fiz. Khim. Obrab. Mater., No. 3, 128 (1980).
5. M. A. D'yakova, E. A. L'vova, and S. Z. Khabliev, Fiz. Met. Metalloved. **44**, 1254 (1977).
6. B. A. Serpetskiĭ, V. V. Levitin, S. V. Loskutov, and V. K. Man'ko, Zavod. Lab. **64** (3), 28 (1998).
7. S. V. Loskutov, V. V. Levitin, V. K. Man'ko, *et al.*, Zavod. Lab. **65** (7), 43 (1999).
8. D. M. Vasil'ev and V. V. Trofimov, Zavod. Lab. **50** (7), 20 (1984).
9. E. A. L'vova, M. A. D'yakova, and S. Z. Khabliev, Izv. Akad. Nauk SSSR, Met., No. 1, 154 (1979).

Translated by B. Malyukov

Technical Physics, Vol. 47, No. 4, 2002, pp. 507–509. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 136–138.
Original Russian Text Copyright © 2002 by Narkhinov.

BRIEF COMMUNICATIONS

Detection and Measurement of the Current of Ribbon-Shaped Electron Beams at a Particular Section of a Converging Emitter System

V. P. Narkhinov

Department of Physical Problems, Presidium of Buryat Scientific Center,
Siberian Division, Russian Academy of Sciences, Ulan-Ude, 670047 Russia
e-mail: burne@bsc.buryatia.ru

Received August 22, 2001

Abstract—A method and an associated device for directly measuring the current distribution over 28 radially converging electron beams are considered. The operation of the device, which can easily be installed in the coaxial system of a gas-discharge electron source, is described. The short-focus radial beams are found to feature a relatively high current uniformity. *© 2002 MAIK “Nauka/Interperiodica”.*

INTRODUCTION

A number of design solutions are employed in gas-discharge electron injectors in order to form beams propagating along a certain direction. Coaxially arranged electron sources can produce both diverging [1] and radially converging [2] charged particle beams. Annular plasma emitters with a radially converging beam, which are used for pumping lasers, significantly increase the lasing efficiency and power because the electron beam energy is uniformly applied to the gaseous mixture [2].

The radially converging beams are also used in electron–ion processing of materials and for curing lacquers and resins on wires and cables. For a radially converging electron beam to efficiently act on an object, its azimuthal power density distribution should be as uniform as possible.

Electron beams are complex subjects of study. Therefore, there are no universal methods that can comprehensively characterize any beam, although a vast variety of well-established diagnostics tools exist [3].

This paper proposes a technique for evaluating the azimuthal current uniformity of converging electron beams over a given section of the transport path.

DESCRIPTION OF THE EXPERIMENTAL SETUP

Figure 1a shows gas-discharge electron source *1*, which produces 28 ribbon-shaped radially converging electron beams (with viewing lid *2* removed); Fig. 1b, the device developed for conducting the experiment. Electrons produced by a gas-discharge plasma, which is localized in 28 elementary Penning cells *3*, are extracted through adjustable slit *4* by applying a voltage from a high-voltage rectifier across the gap between anode rings *5* and collector electrode *6*. The collector has 28 1.2-mm-diam. holes *7* and is arranged so that the centers of the holes and of 28 cells *3* coincide.

The design of the device is illustrated in Fig. 2. 60-mm-long 5-mm-diam. Faraday cup *1*, enclosed in protective housing *2*, is fixed to holder *3* by a stud and nut. Conducting copper bushing *5* between dielectric holder *3* and intermediate shaft *4* provides a sliding contact between pressure terminal *6* and rotating Faraday cup *1*. Shaft *4* (intermediate bushing) is fitted to the shaft of synchronous motor *7*. Holder *8* of pressure terminal *6* and the SD-54 synchronous motor are mounted on table *9* (aluminum disk), which is installed in a recess made in support insulator *11*. When the device is energized through wires *10* and vacuum-tight connectors, the Faraday cup starts rotating (n = 2.24 rpm) inside collector electrode *6* (Fig. 1).

Some of the electrons penetrating through the small holes are captured by the Faraday cup and produce electric signals, which are measured.

EXPERIMENTAL RESULTS AND DISCUSSION

The currents of the electrons captured are proportional to the current densities near the holes:

$$J(r) = \frac{I(r)}{S},$$

where $I(r)$ is the current through the hole and S is the cross-sectional area of the hole.

The measurement error due to the error δ in measuring the currents is given by

$$\Delta J = \frac{\delta}{S};$$

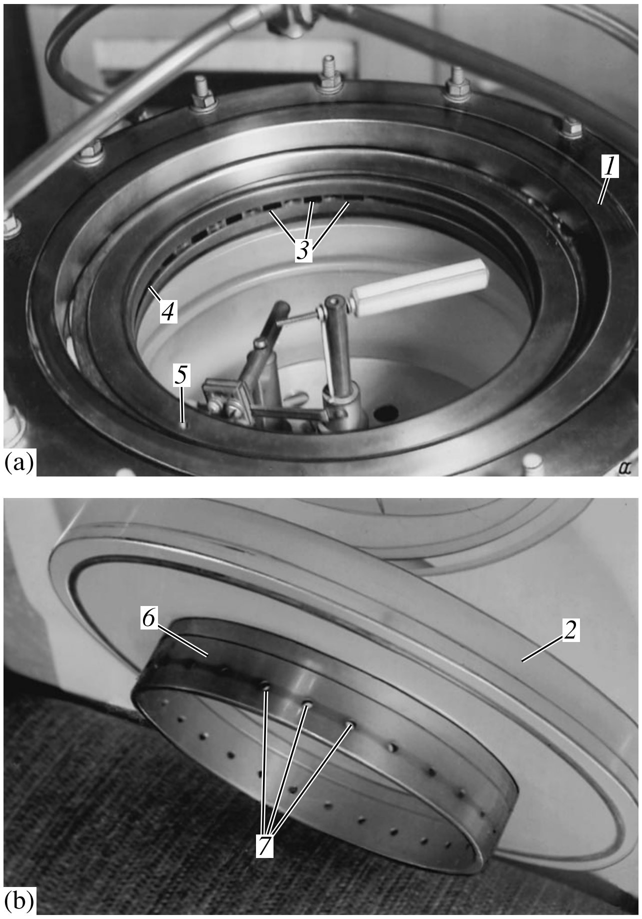


Fig. 1. Plasma electron source with the built-in experimental device for the diagnostics of beam currents uniformity of: (*1*) plasma electron source, (*2*) viewing lid, (*3*) Penning gas-discharge cells, (*4*) emission slit, (*5*) anode rings, (*6*) collector electrode, and (*7*) holes.

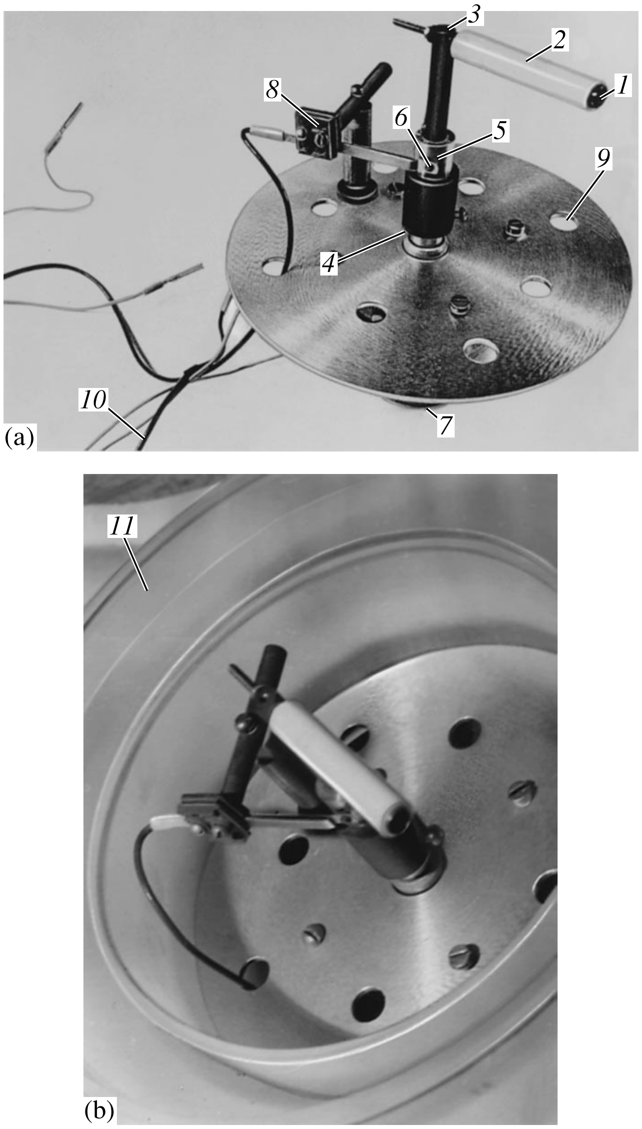


Fig. 2. (a) Design and (b) installation of the device with one Faraday cup: (*1*) Faraday cup, (*2*) protective insulating housing, (*3*) dielectric holder, (*4*) intermediate bushing (shaft), (*5*) conducting bushing, (*6*) pressure terminal, (*7*) electric motor, (*8*) pressure terminal holder, (*9*) table, (*10*) wires, and (*11*) insulating support.

i.e., the error in determining the current density depends on the absolute error δ in measuring the current of the electrons that pass through a hole of cross-sectional area S.

Our estimates showed that the scatter in the currents is relatively small (no more than 5% including the measurement error). The average density of the azimuthal electron current from all the 28 radially converging beams was 10^{-5} A/cm^2.

In the next series of experiments, the cross section of the ribbon beams was varied from 15 to 40 mm^2 by changing the height of the exit window of the emission slit in 0.5-mm steps with the beam width kept constant (10 mm). It turned out that the current uniformity is independent of the beam cross section. The current densities calculated in six experiments were constant within one tenth.

The experimental results can be considered quite reliable, because the electron emission was stable and the rotating Faraday cup, one of the most accurate instruments, measured the currents repeatedly.

The well-known traveling-slit, perforated-chamber, vibrating-probe [3], and rotating-probe [4] techniques are inapplicable to this problem, because they cannot handle a large number of beams simultaneously. Our method for detecting and measuring the current distribution over the variable cross sections of 28 radially converging ribbon electron beams can be used for finding actual currents along the circumference of a 140-mm-diam. cylindrical body being processed.

The use of a single rotating Faraday cup greatly simplifies measurements in a multibeam radially converging system. Due to its simple design, the device is easy to fabricate and maintain, which favors a further development of the method for charged particle beam diagnostics.

Varying the diameter of the collector electrode and the radial position of the Faraday cup with the remaining components kept unchanged, one can measure the current distribution in any cross section of the electron beams. When the anode–collector spacing is longer than 10 mm, one should supply the source with an accelerating electrode and apply the above technique, placing the perforated collector coaxially with the accelerating electrode.

By displacing the centers of the holes from the beam axes, one can find the current density distribution over the beam cross section. Thus, there are vast potentialities for the improvement and modification of both the method and the device.

REFERENCES

1. A. M. Efremov, B. M. Koval’chuk, Yu. E. Kreĭndel’, *et al.*, Prib. Tekh. Éksp., No. 1, 167 (1987).
2. S. P. Bugaev, L. G. Vintizenko, V. I. Gushenets, *et al.*, in *Proceedings of the VII All-Union Workshop on High-Current Electronics, Tomsk, 1988*, Part II, p. 174.
3. S. I. Molokovskiĭ and A. D. Sushkov, *Intense Electron and Ion Beams* (Énergoatomizdat, Moscow, 1991).
4. O. K. Nazarenko, V. E. Lokshin, and K. S. Akop’yants, Élektron. Obrab. Mater., No. 1, 87 (1979).

Translated by A. Khzmalyan

Technical Physics, Vol. 47, No. 4, 2002, pp. 510–512. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 4, 2002, pp. 139–142.
Original Russian Text Copyright © 2002 by Rapis.

BRIEF COMMUNICATIONS

A Change in the Physical State of a Nonequilibrium Blood Plasma Protein Film in Patients with Carcinoma

E. Rapis

Tel-Aviv University, Ramat-Aviv, 64239 Israel

Received September 24, 2001

Abstract—Pronounced changes in the physical phase and in the phase transition dynamics of proteins in drying blood plasma are observed. The morphologies (topologies) of these nonequilibrium films in donors and in patients with various types of metastatic carcinoma qualitatively differ by the process of protein film self-assembly and by symmetry type. New types of defects and solid crystals appear, liquid crystals persist for a long time, etc. The microscopic examination of the drying protein plasma topology can be used for diagnosing metastatic carcinoma. *© 2002 MAIK “Nauka/Interperiodica”.*

INTRODUCTION

The functionality and the structure of material in both organic and inorganic nature depend basically on its physical state (phase) and phase transition features. Yet, a correlation between the formation of functioning protein at the supramolecular (cell) level and its phase transitions under various condensation conditions (for example, the sol–gel transition) has not been adequately studied to date, although phase transitions in biological systems specify the protein behavior. They attend the most important dynamic phenomena in a living organism, such as cell and nucleus division; protein synthesis; fermentative, immune, and locomotive processes; etc. [1]. Thus, this offers scope for devising protein phase tests.

Our early experiments [2–5] have clearly demonstrated the possibility of *in vitro* observing the phase transition dynamics in the protein–water system. These observations have been compared with the *in vivo* behavior of the drying blood protein plasma, which is normally the formation of the “protos” protein modification. Upon the drying, the biological self-organization of the protein film with the specific geometry (topology) and different symmetry types and scales takes place (Fig. 1).

However, these data have not been used to compare the physical state (phase transitions, etc.) of normally and pathologically self-organized protein in medicine and biology. However, it is well known that the growth of malignant cells breaks the order of symmetry at the cell level. In the absence of ordered film growth, the space distribution of the cells changes, which eventually causes the generalization of the process and the death of the organism [6–8]. It is of interest to trace whether these pathological changes are related to self-organization in biological systems. It is believed that the formation, division, and multiplication of cells, as well as their systematic arrangement in space and time, are direct signs of biological self-organization. If so, the breaking of protein self-organization at the cell level should be considered as the basic stage in the pathogenesis of malignant tumors. However, the study of the cancer pathology is being concentrated largely on the cell morphology, immunology issues, search for oncogenes, etc. For example, there is evidence for biochemical, colloidal, and other changes in the blood indices [9–12], as well as for the appearance of glutamic acid and its protein constituents in blood (Savina and Tuev, 1998). However, protein and its self-organization are being investigated at the molecular and atomic levels (the study of molecular chain spiralization [13] is an example). It appears that a distinct tendency in studying the pathology of malignant tumors (specifically carcinomas) in terms of protein self-organization at the supramolecular level is absent.

To consider carcinoma from this standpoint using data for self-organization reported in [2–5], we performed a special investigation.

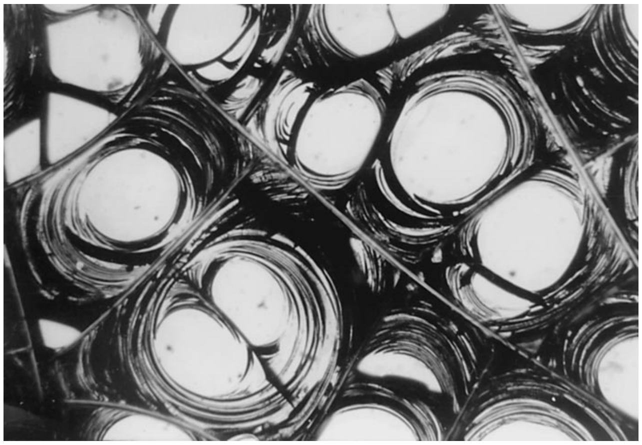


Fig. 1. Typical protein symmetry.

MATERIALS AND METHODS

Our aim was to visualize the liquid–solid transition in the fresh plasma separated out from patient's blood prior to experiments and in the blood plasma kept in a refrigerator at +1.0°C for three days. The plasmas were placed on a solid substrate in equal amounts (for details, see [2–5]).

The objects of investigation were plasma samples taken from five patients (45 samples) with metastatic cancer (mammary gland carcinoma, neck-of-womb carcinoma, and lungs carcinoma) (group 1). They were compared with blood samples taken from 20 healthy donors (45 samples) (group 2) (Fig. 2).

To date, we performed only phenomenological studies; however, even the qualitative data are rather convincing. The morphologies (topologies) of the drying blood films for the two groups distinctly differ in a number of physical parameters: phase transition, symmetry type and scale, protein phase, etc. In group 2, the disturbances showed up as irregular division into protein blocks (films) with symmetry and similarity violation and increased cell nuclei. The central part of the cell (nucleus) becomes too large compared with the irregular movable (fluxible) edge, at which a great array of asymmetric defects is observed (Fig. 3).

The contours of the defects and films only tend to flex; there are no regular parallel flexed lines or films entering into each other (Fig. 2). The films lose the regular orbital (spiral) form, looking like leaves or warts with irregular edges and hyperfine structures inside. Small randomly arranged tapered protrusions are frequently observed. The most surprising thing is the clustering of well-faceted cubic crystals, which usually do not form under nonequilibrium conditions. They are as if slipped over straight thin filaments, which are parallel to each other. With time (in several months), the crystals form large agglomerates, retaining the regular shape.

Another surprising thing was that these crystals appeared near the protein films (in an optical microscope, one half of the film had usual black color; the other, with crystal clusters seen inside, had an irregular shape and was colored white).

DISCUSSION

Thus, we succeeded to visualize the profound changes in the phase transition and the phase state of blood plasma proteins in patients with metastatic cancer. Under nonequilibrium conditions, instead of the solid-state modification "protos" [2–5], identified in the plasma of the donors, we revealed the coexistence of three protein states: (1) the long-lived (up to six to eight months) liquid-crystal phase with a very high mobility ("fluidity"); in this case, the great disturbance of optical effects (hence, of the optical axis orientation) was observed; (2) the strongly modified protos state with the violation of symmetry, self-similarity, tridimensionality, fractality, nucleation, and adhesion to the substrate; (3) the equilibrium solid-crystal phase, appearing under nonequilibrium conditions of condensation.

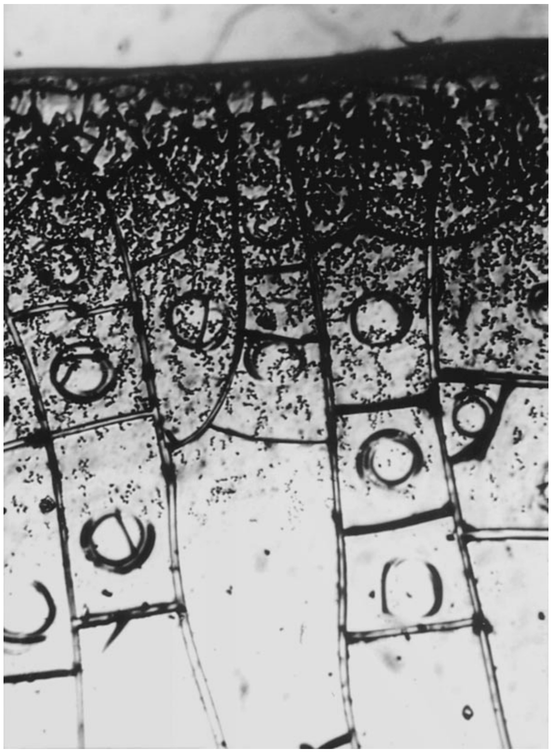


Fig. 2. Micrograph of the drying blood plasma of a healthy donor.

This means that we established changes in the blood plasma proteins taking place in patients with cancer when the protein condenses during self-assembling (phase transitions).

This is because our approach makes it possible to determine, from morphological (topological) patterns, normal protein self-organization with certain symmetries and scales in a complex system, such as the blood plasma. Such behavior of protein is based on its property to compete in activity with other plasma ingredients. It follows from our results that the protein concentration in the blood plasma of a healthy person provides the regular morphology of a drying protein film (Fig. 2).

The profound qualitative changes in the physical properties of the blood plasma proteins in patients with metastatic carcinoma highlight a new aspect of the problem: extracellular pathological change in the proteins with violation of self-organization (upon phase transitions), which alters their phase state.

Therefore, the need for detailed investigation into the phases and symmetries of the solid modification of the protein, as well as of its liquid-crystal phase, for various forms of cancer is obvious. The investigation should be performed by medical men in cooperation with physicians engaged in different field of physics

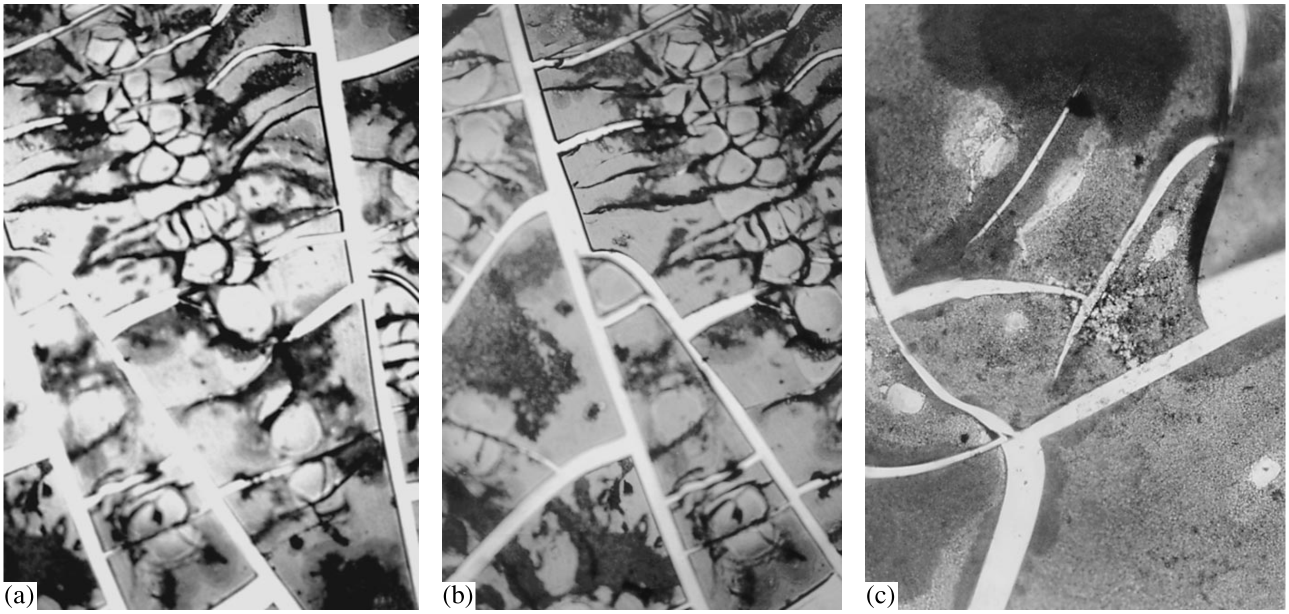


Fig. 3. Micrograph of the drying blood plasma of a patient with cancer.

using contemporary biological and physical technologies. This problem becomes of primary importance in the treatment of cancer, since little is known about chemical and morphological variations in a cancer cell and also about changes in the physical phases of blood plasma proteins. It is necessary to find reasons for those physical changes found experimentally. They can hardly be ignored in studying the pathogenesis of malignant transformations.

Even our pioneering observations, as yet statistically unreliable, must attract the attention of the scientific society.

ACKNOWLEDGMENTS

I sincerely thank Profs. M. Amus'ya, A. Arel', E. Braudo, V. Buravtsev, V. Volkov, A. Zaikin, M. Klinger, L. Manevich, S. Moiseev, Yu. Noeman, I. Prigogine, and M. Safro for the fruitful discussions, valuable comments, and encouragement.

REFERENCES

1. B. M. Alberts *et al.*, Mol. Biol. Cell **3**, 1 (1992).
2. E. G. Rapis and G. Yu. Gasanova, Zh. Tekh. Fiz. **61** (4), 158 (1991) [Sov. Phys. Tech. Phys. **36**, 406 (1991)].
3. E. G. Rapis, Pis'ma Zh. Tekh. Fiz. **21** (9), 13 (1995) [Tech. Phys. Lett. **21**, 321 (1995)].
4. E. G. Rapis, Pis'ma Zh. Tekh. Fiz. **23** (7), 29 (1997) [Tech. Phys. Lett. **23**, 263 (1997)].
5. E. G. Rapis, Zh. Tekh. Fiz. **70** (1), 121 (2000) [Tech. Phys. **45**, 121 (2000)].
6. D. S. Coffey, Nat. Med. **4**, 1342 (1998).
7. V. Backman *et al.*, Nature **406** (6), 36 (2000).
8. T. Goetze and J. Briokman, Biophys. J. **61**, 109 (1992).
9. A. B. Classman and E. Jones, Ann. Clin. Lab. Sci. **24** (1), 1 (1994).
10. G. Buccheri, D. Ferringino, C. Ginardi, *et al.*, Eur. J. Cancer **33**, 50 (1977).
11. B. Ya. Gurvits and K. G. Korotkov, Consciousness and Physical Reality. JSSN. No. 1029-4716, Vol. 1, No. 1 (Publishing Company Folium, 1998), pp. 84–90.
12. J. A. Forrester, E. J. Ambrose, and M. G. P. Stoker, Nature **201**, 945 (1964).
13. P. G. Wolynes and W. A. Eaton, Phys. World **12** (9), 39 (1999).

Translated by V. Isaakyan