Technical Physics, Vol. 50, No. 8, 2005, pp. 1006–1015. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 44–53.

GASES AND LIQUIDS

Effect of the Initial Deformation of a Charged Drop on Nonlinear Corrections to Critical Conditions for Instability

A. I. Grigor'ev, A. N. Zharov, and S. O. Shiryaeva

Demidov State University, Sovetskaya ul. 14, Yaroslavl, 150000 Russia

e-mail: grig@uniyar.ru

Received December 1, 2004

Abstract—A nonlinear (proportional to the vibration amplitude squared) decrease in the critical (in terms of instability) charge of a vibrating drop is found to be limited, as follows from third-order asymptotic calculations. This effect occurs when the spectrum of modes specifying the initial deformation of the drop contains, along with the fundamental mode, higher modes. The influence of the environment density on nonlinear corrections to the critical conditions for instability is analyzed. © *2005 Pleiades Publishing, Inc.*

(1) The critical conditions for instability of an isolated drop of a conducting ideal incompressible liquid against its charge were theoretically derived at the end of the 19th century by Rayleigh in the form of the relationship $W \equiv (Q^2/4\pi\sigma R^3) \geq 4$, where R and Q are the radius and charge of the drop, respectively, and σ is the surface tension coefficient of the liquid [1]. Throughout the 20th century and at the beginning of the current one, this relationship was verified many times in a variety of laboratory experiments: (i) in a vertical electrostatic field between plane-parallel plates (i.e., in an electrostatic suspension similar to the one used by Millikan in the experiments on determining the electron charge) [2], (ii) in a nonuniform time-periodic electric field between electrodes of complicated geometry (a combination of rings with cylindrical and spherical surfaces) [3], in a combined electric suspension made up of an electrostatic field and periodic electric fields between three plane electrodes [4], in an air flow [5], and in an electrodynamic suspension made up of two annular electrodes [6]. In these experiments, the size of drops was varied in a wide range: several hundreds of micrometers [2, 5], several tens of micrometers [3, 4], and several micrometers [6] (it should be noted, however, that the physical mechanisms of excess charge removal from drops several hundreds of micrometers in diameter and one micrometer in diameter are radically different; for details, see [7, 8]). The validity of the Rayleigh criterion was confirmed in all the cases. The highest experimental accuracy was attained in [4], where the Rayleigh criterion was found to be accurate up to 4%, and in [6], where the inaccuracy was close to 5%.

Nonlinear investigations of the vibration and stability of a charged drop [9–12] have shown that quadratic-in-amplitude corrections to vibration frequencies appear in calculations of the third order of smallness in initial deformation amplitude. Since the critical conditions for instability of the drop against its charge are derived from the requirement that the squared frequency of the fundamental mode of vibrations pass through zero and become negative, the critical value of the Rayleigh parameter, $W = W_{cr}$, turns out to be dependent on the squared initial deformation of the equilibrium (spherical) shape. This result is valid so long as the asymptotic expansions used in analyses [9–12] are applicable and a nonlinear correction to the critical value of the Rayleigh parameter meets the only requirement that it be much smaller than four. In other words, depending on the experimental conditions (on the amplitude of the initial deformation of the drop), the measured critical value of the Rayleigh parameter may deviate from the value $W_{cr} = 4$, which is predicted by the linear theory, by 10–20%. In view of the aforesaid, the measurements made in [4, 6] cast some doubt, because the experimental setups used in those works unduly excited the fundamental mode of the drop (its amplitude was very high). This work is aimed at clarifying this point. Note that the investigation takes into account real experimental conditions, namely, the presence of the environment, which is simulated by an ideal incompressible insulating liquid.

(2) Let a spherical drop of an ideal incompressible perfectly conducting liquid of density $\rho_{(i)}$ have radius R and charge Q and be surrounded by an ideal incompressible liquid of density $\rho_{(e)}$ and permittivity ε_d (the gravitational field is absent). The surface tension coefficient is denoted by σ. We assume that the liquid flow in the drop and environment is potential with velocity potentials $\psi_{(i)}$ and $\psi_{(e)}$, respectively. The electric field potential near the drop is denoted by ϕ. It is assumed that the shape of the drop is axisymmetric at any (including the initial) time instant. In dimensionless variables such that $\rho_{(i)} = 1$, $R = 1$, and $\sigma = 1$, the equa-

1063-7842/05/5008-1006$26.00

tion for the interface at any time instant t is given by

$$r = 1 + \xi(\vartheta, t).$$

The mathematical statement of the problem of nonlinear capillary vibrations of a charged drop has the form

$$\Delta\psi_{(i)} = 0; \quad \Delta\psi_{(e)} = 0; \quad \Delta\phi = 0;$$

$$r \longrightarrow 0: \psi_{(i)} \longrightarrow 0;$$

$$r \longrightarrow \infty: \psi_{(e)} \longrightarrow 0; \quad \nabla\phi \longrightarrow 0;$$

$$r = 1 + \xi(\vartheta, t): \partial_t\xi = \partial_r\psi_{(i)} - \frac{1}{r^2}\partial_\vartheta\psi_{(i)}\partial_\vartheta\xi$$

$$= \partial_r\psi_{(e)} - \frac{1}{r^2}\partial_\vartheta\psi_{(e)}\partial_\vartheta\xi;$$

$$\partial_t\psi_{(i)} + \frac{1}{2}(\nabla\psi_{(i)})^2 - \rho\left(\partial_t\psi_{(e)} + \frac{1}{2}(\nabla\psi_{(e)})^2\right)$$

$$= p_0 - p_\infty + p_q - p_\sigma; \quad \phi = \phi_S(t);$$

$$\int_V r^2 \sin\vartheta\, dr d\vartheta d\varphi = \frac{4\pi}{3};$$

$$V = \{r, \vartheta, \varphi | 0 \le r \le 1 + \xi; 0 \le \vartheta \le \pi; 0 \le \varphi \le 2\pi\};$$

$$\int_S \mathbf{n}\cdot\nabla\phi\, dS = -4\pi Q;$$

$$S = \{r, \vartheta, \varphi | r = 1 + \xi; 0 \le \vartheta \le \pi; 0 \le \varphi \le 2\pi\};$$

$$t = 0: \xi = \xi_0 P_0(\cos\vartheta) + \varepsilon \sum_{m\in\Omega} h_m P_m(\cos\vartheta);$$

$$\partial_t\xi = 0.$$

In these equations, p_∞ and p_0 are the pressure in the environment at infinity and a constant pressure in the drop, respectively; p_q and p_σ are the electric field pressure and the pressure of surface tension forces, respectively; $\mathbf{n}$ is the normal vector to the surface of the drop; ϕ_S is the surface electric potential of the drop; $\rho \equiv (\rho_{(e)}/\rho_{(i)})$; ε is a small parameter that has the meaning of the initial perturbation amplitude; $P_m(\cos\vartheta)$ is the mth-order Legendre polynomial; ξ_0 is a constant chosen in such a way that the initial volume of the drop coincides with the volume of an equilibrium sphere; symbol ∂_t means partial differentiation with respect to time t; Ω is the set of indices of initially excited modes; constants h_m take into account the contribution of an mth mode to the initial shape of the drop and satisfy the condition $\sum_{m\in\Omega} h_m = 1$; and Δ is the Laplacian.

(3) Solving the problem by the method of many scales in the third order of smallness, one arrives at an analytical expression for the generatrix of a conducting incompressible liquid charged drop nonlinearly vibrating in an insulating incompressible medium [13],

$$r(\vartheta, t) = 1 + \varepsilon \sum_{n\in\Omega} M_n^{(1)}(t) P_n(\cos\vartheta)$$
$$+ \varepsilon^2 \sum_{n=0}^{\infty} (M_n^{(2)}(t) + \varepsilon M_n^{(3)}(t)) P_n(\cos\vartheta), \quad (1)$$

where

$$M_n^{(1)}(t) = h_n \cos(\omega_n t + \varepsilon^2 b_n t);$$

$$\omega_n^2 = \left(\frac{(n^2-1)n(n+2-W)}{1+n(1+\rho)}\right);$$

$$W \equiv (Q^2/4\pi\varepsilon_d\sigma R^3);$$

$$b_n = \frac{1}{2\omega_n}\left\{\frac{h_n^2(\Xi_n^0 + 2\omega_n^2(\Xi_n^1 - 2\Xi_n^2))}{4(2n+1} + \sum_{k\in\Omega}\frac{h_k^2\Xi_n^0}{2(2k+1)}\right.$$

$$- \sum_{k\in\Omega}\frac{h_k^2}{4}[H_{nkkn}^{1(-)(+)} + H_{knkn}^{2(+)(+)} + H_{knkn}^{2(-)(-)}$$

$$\left. + (1-\delta_{kn})(H_{kknn}^{1(-)(+)} + H_{kknn}^{2(+)(+)} + H_{nkkn}^{2(-)(-)})]\right\}.$$

The explicit form of the coefficients that appear in the expression for factor b_n in the nonlinear correction to the frequency can be found in the Appendix. The awkward analytical expressions for coefficients $M_n^{(2)}$ and $M_n^{(3)}$ are omitted as being of no interest here, since we are considering the critical conditions for instability of the drop against its charge that are directly related to nonlinear corrections to the vibration frequencies.

(i) It has been mentioned above that the corrections to the frequencies of capillary vibrations of the drop modify the critical conditions for instability of an mth mode against the charge. The square of the capillary vibration frequency is known to decrease with increasing charge and vanish at a certain critical value (W_{cr}) of the Rayleigh parameter. A further increase in the charge makes the frequency squared negative, i.e., gives rise to imaginary frequencies and causes the amplitudes of surface capillary vibration of the drop to grow exponentially, which means the onset of instability. With regard to the nonlinear correction to the frequency, the critical condition for instability of an nth mode can be written in the form

$$(\omega_n + \varepsilon^2 b_n)^2 \approx \omega_n^2 + 2\varepsilon^2\omega_n b_n + O(\varepsilon^4) = 0. \quad (2)$$

The parameters involved in this expression must satisfy the condition $|\omega_n| \gg |\varepsilon^2 b_n|$. Basically, substituting the expressions for frequency ω_n and coefficient b_n into

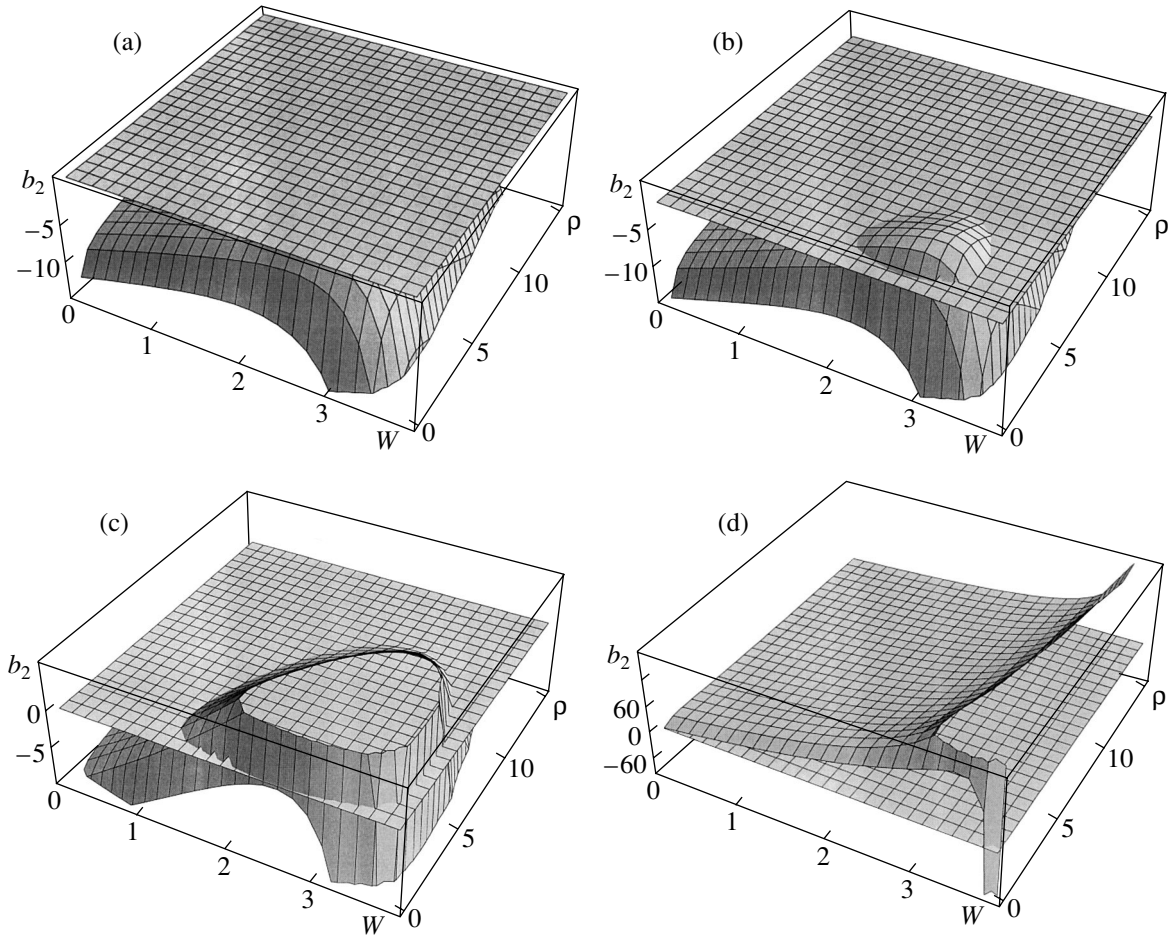


Fig. 1. Coefficient $b_2 = b_2(W, \rho)$ determining the value of the nonlinear correction to the fundamental mode frequency vs. Rayleigh parameter W and dimensionless environmental density ρ. The calculations were made for various types of initial deformation specified by superposition of the fundamental mode ($n = 2$) and a higher mode with number $m > 2$: m = (a) 11, (b) 12, (c) 14, and (d) 30 at $h_2 = 0.75$ and $h_m = 0.25$.

(2), one can obtain an awkward expression relating critical value W_{cr} of the Rayleigh parameter to the spectrum of initially excited modes and their amplitudes expressed in terms of small parameter ε squared. However, relation (2) is also worth analyzing.

As is evident from (2), the calculations of the second order of smallness in ε imply that, as the amplitude of the initial deformation (i.e., ε^2) grows, critical value W_{cr} of the Rayleigh parameter decreases (compared with $W_{cr} = 4$ predicted by the linear theory) if $b_n < 0$ and increases if $b_n > 0$. In [13], various situations with single-mode initial deformations were analyzed and it was demonstrated that $b_n < 0$ everywhere except for the neighborhoods of the frequencies corresponding to internal nonlinear resonances; thus, the conclusion drawn in [9–12] that W_{cr} decreases with increasing amplitude of the initial deformation was substantiated. Actually, however, the sign of coefficient b_n is a complicated function of the spectrum of modes specifying the initial deformation, dimensionless environmental density ρ, and Rayleigh parameter W and is not necessarily negative.

(ii) Figures 1a–1d present the dependence $b_2 = b_2(W, \rho)$ in the parameter range $0 \le W \le 3.9$ and $0 \le \rho \le 15$. The plots are constructed for various initial deformations specified by a superposition of the fundamental mode ($n = 2$) and a higher mode with number $m > 2$, i.e., for two-mode initial deformations of an equilibrium spherical drop with partial contributions $h_2 = 0.75$ and $h_m = 0.25$. Calculations show that coefficient b_2 appearing in the nonlinear correction to the fundamental mode frequency is negative if $m \le 11$. According to Fig. 1b, when $m \ge 12$, parameters W and ρ can be taken such that $b_2 \ge 0$. One can see from Figs. 1a–1d that the ranges of W and ρ where $b_2 \ge 0$ expand with m. Figure 1c demonstrates that, if $m \ge 14$, the leading boundary of this domain reaches the plane ρ = 0, which physically corresponds to the zero environmental density, i.e., to a drop vibrating in vacuum. For $m \ge 30$, ρ = 0, and $W < 3.8$, coefficient b_2 is positive (Fig. 1d), which, according to the aforesaid, means that the critical value of the

Rayleigh parameter, W_{cr}, grows with vibration amplitude. In other words, if the initial deformation of the drop is specified by a superposition of the fundamental mode and a mode with a number $m \geq 30$, a highly charged drop with $W \leq 3.8$ remains stable (in terms of the asymptotic theory) even if the amplitude of its nonlinear vibrations is significant. This implies that, for the initial deformation mentioned above, the critical value of the Rayleigh parameter theoretically cannot deviate from 4 by more than 5%. It seems that this very fact explains the results of the experiments in [4, 6], where the fundamental mode was probably excited together with several higher modes.

Figure 2 plots coefficient b_2 appearing in the correction to the fundamental mode frequency versus partial contributions h_2 and h_m of the fundamental mode (n = 2) and a higher mode with number m (here m = 20) to the initial deformation of the drop. It can be easily seen that b_2 depends on h_m much more strongly than on h_2 and that the fundamental mode and the higher mode contribute to b_2 with different signs.

Figures 3a and 3b plot coefficient b_2 calculated in a more complicated situation where the initial deformation is specified by a superposition of three, rather than two, modes: the fundamental mode (n = 2) and two higher modes. The partial contribution of the fundamental mode is h_2 = 0.75, i.e., the same as in Figs. 1a–1d. The partial contributions of two higher modes add up to $h_m + h_k$ = 0.25, i.e., to the same value as for one higher mode in Fig. 1. The qualitative and quantitative similarity of the data presented in Figs. 1 and 3 is obvious.

Remarkably, Figs. 1d and 3b imply that, in the more general case of nonlinear vibrations of a drop in an environment with $\rho \geq 0.1$, i.e., when the fundamental mode is initially excited together with one or several higher modes with numbers $m > 30$ at subcritical (in the sense of the linear theory) values of the Rayleigh parameter, Rayleigh decomposition of the drop is barely, if at all, possible.

(ii) The physical meaning of the phenomenon discovered (increased stability of the drop when higher vibration modes are excited) is intimately related to the physical mechanism of instability of a heavily charged drop (for details, see [14–16]). According to [14–16], when the Rayleigh parameter attains a critical value, the fundamental mode of the drop becomes unstable; as a result, the drop elongates, taking the shape of a prolate spheroid. As the drop elongates, its charge is redistributed over the surface and the charge density at the tops rises to the point where higher modes are excited. Accordingly, the curvature of the tops of the drop increases further and still higher modes are excited. Eventually, emissive asperities arise at the tops of the drop, which release an excess charge by emitting heavily charged droplets. If a higher mode is imposed on the surface of the drop with the already excited fundamental mode (i.e., on the weakly spheroidal surface),

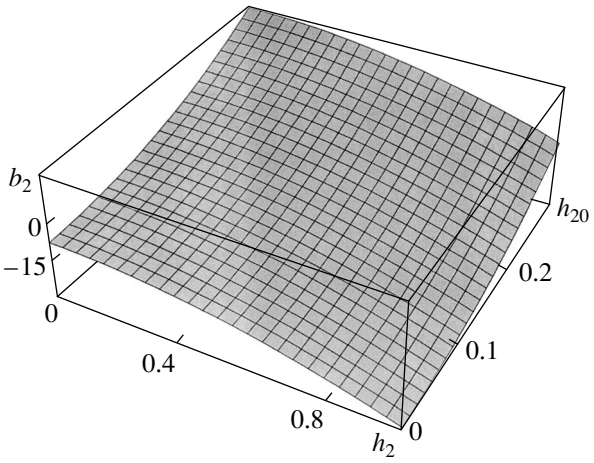


Fig. 2. Coefficient b_2 in the correction to the fundamental mode frequency vs. partial contributions h_2 and h_m of the fundamental mode (n = 2) and the 20th mode (m = 20) to the initial deformation of the drop. The calculation was made for W = 1, ρ = 0.1, and ε = 0.3.

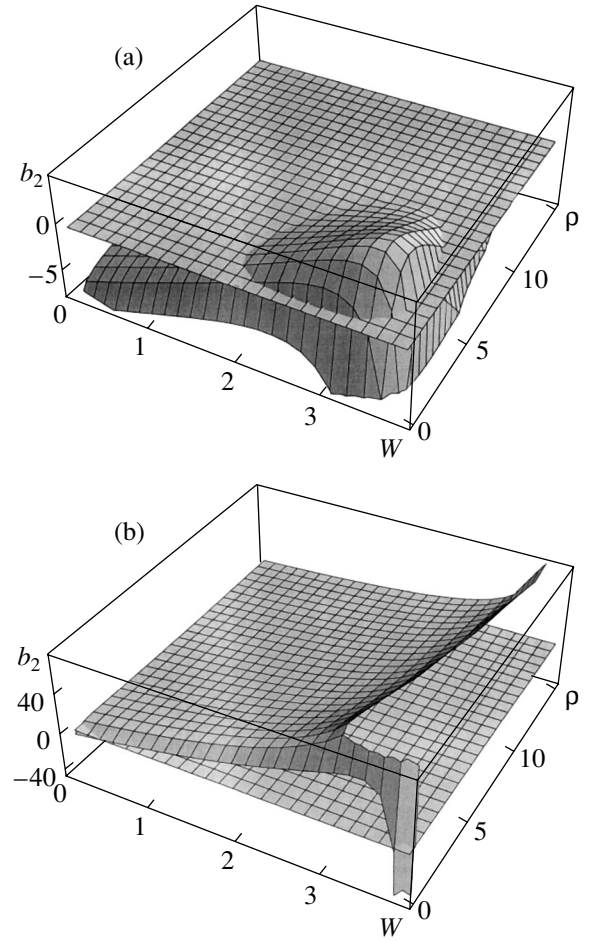


Fig. 3. Coefficient $b_2 = b_2(W, \rho)$ vs. Rayleigh parameter W and dimensionless environmental density ρ. The calculations were made for three-mode initial deformations specified by superposition of the fundamental mode with h_2 = 0.75 and a pair of higher modes with numbers m and $k > 2$: (a) m = 14, k = 20, h_{14} = 0.15, and h_{20} = 0.10 and (b) m = 17, k = 34, h_{17} = 0.1, h_{34} = 0.15.

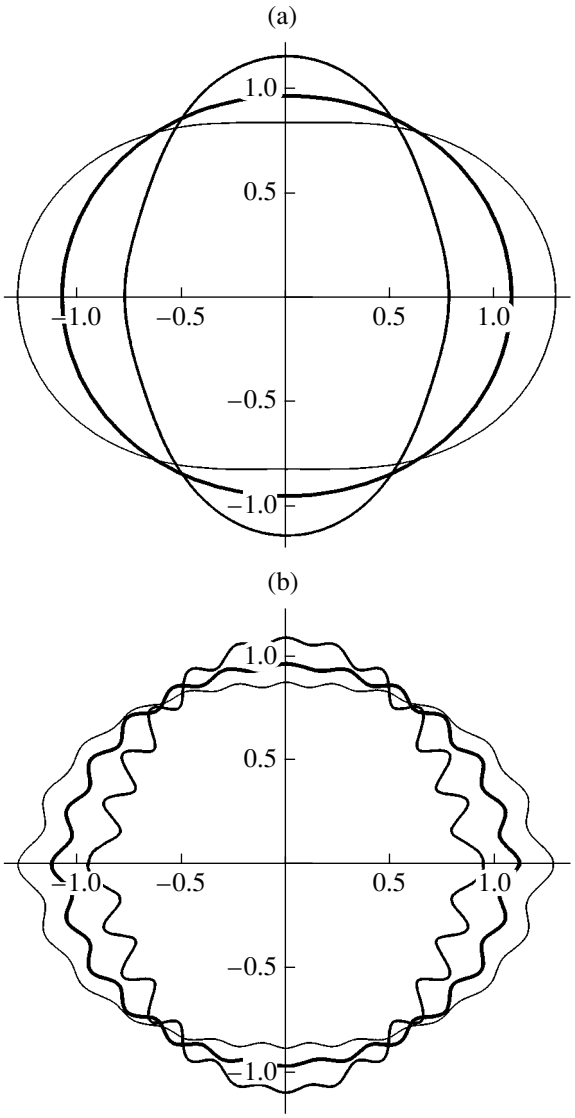


Fig. 4. Generatrix of the drop calculated for $W = 1$ and $\rho = 0$ at time instants $t = 0$ (thin line), $0.25T$ (thick line), and $0.5T$ (the middle-thickness line). T is the vibration period of the fundamental mode. (a) Fundamental mode alone is initially excited with $\varepsilon = 0.3$ and (b) fundamental mode and the 24th mode are initially excited with $\varepsilon = 0.3$, $h_2 = 0.75$, and $h_{24} = 0.25$.

the smooth surface of the spheroid will exhibit a small-scale relief (ripple) (see Fig. 4) with an increased surface charge density at the ridges, where the curvature is very high. As a result, the average charge density at the spheroid's vertices decreases, since the total charge on the drop remains constant, and the fundamental mode amplitude builds up more slowly.

The numbers of modes specifying the initial deformation of the drop were chosen from the following considerations. First, as follows from full-scale experiments [17], the fundamental mode with an amplitude several times smaller than the radius of the drop is typically excited in cloud drops because of a turbulent flow around the drop suspended in the updraft. The same phenomenon was observed in the experiments described in [8]. In experimental works [4, 6], the fundamental mode was excited by the variable component of a suspending electric field and particular emphasis was placed on the observed excitation of the fundamental mode [6]. Excitation of higher modes in large (of diameter ≈200 μm) drops of stormy clouds is due to collisions with fine droplets (of diameter ≈10 μm), whose concentration in clouds is maximal [18]. In the experiments [2–6], higher modes are excited from interaction between the drop and the flow of vapor and clustered charged molecules evaporating from the surface: in all the cases [2–6], drops bearing a charge smaller than the Rayleigh critical value and having an initially large radius evaporated from the surface, tending to reach the critical value of the Rayleigh parameter.

To summarize the aforesaid, we note that the magnitude and sign of a nonlinear correction to the fundamental mode frequency depend on the spectrum of modes responsible for the initial deformation of the drop, $b_n = b_n(\Omega)$. As a consequence, nonlinear corrections to the critical conditions for instability of a charged drop (or to the critical value of the Rayleigh parameter) depend on the initial shape (deformation) of the drop, $W_{cr} = W_{cr}(\Omega)$.

(4) It has already been noted that coefficients $b_n(\Omega)$ have a resonance form. To analyze this circumstance, we recast the expressions for the nonlinear corrections to the frequencies in the form of polynomials in powers of Rayleigh parameter W, as was done in [9].

Consider a free charged drop in vacuum ($\rho = 0$) the initial deformation of which is specified by the excitation of a single mth mode; i.e., $\Omega = \{m\}$ and $\xi = \xi_0 P_0(\cos\vartheta) + \varepsilon P_m(\cos\vartheta)$ at $t = 0$. If one of the second, third, or fourth modes is initially excited, coefficients b_m characterizing the correction to the frequency of the corresponding mode can be represented by the expressions

$$b_2 = -\frac{4}{245}\frac{1}{\omega_2^3\omega_4^2(\omega_4^2 - 4\omega_2^2)}\sum_{i=0}^{4} A_i^{(2)} W^i,$$

$$b_3 = \frac{144}{11011}\frac{1}{\omega_2^2\omega_3\omega_4^2\omega_6^2(\omega_2^2 - 4\omega_3^2)(\omega_6^2 - 4\omega_3^2)}\sum_{i=0}^{6} A_i^{(3)} W^i,$$

$$b_4 = -\frac{1244160}{2433431} \times \frac{1}{\omega_2^2\omega_4^3\omega_6^2\omega_8^2(\omega_2^2 - 4\omega_4^2)(\omega_6^2 - 4\omega_4^2)(\omega_8^2 - 4\omega_4^2)} \times \sum_{i=0}^{8} A_i^{(4)} W^i. \quad (3)$$

Table 1. Coefficients $A_i^{(m)}$

i	$m = 2$	$m = 3$	$m = 4$
0	6606528	421421024000	55735591155609600
1	–4961440	–4128178176000	–77949491906388480
2	1419804	1682216124000	45351373912349312
3	–177168	–362825358328	–14555386948486656
4	7945	43303979512	2840187292166640
5	–	–2679419780	–345863703031648
6	–	66094721	25619763735024
7	–	–	–1049270108016
8	–	–	18006768899

Table 2. Coefficients $B_i^{(m)}$

i	$m = 2$	$m = 3$
0	28507064560128	5478617683875840
1	–31524733738560	–7819057467750912
2	15054864474528	4925997961586016
3	–4000146404140	–1794842151186848
4	634385531392	415673928885262
5	–59666243685	–63222887884663
6	3064663192	6287045064547
7	–66028600	–392114665003
8	–	13835223520
9	–	–208973864

Coefficients $A_i^{(m)}$, $m = 2$–4, are given in Table 1. The denominators of all the coefficients are seen to contain expressions vanishing at $\omega_m^2 = 2^2\omega_n^2$. Thus, a degenerate resonance can occur.

If several modes of capillary vibrations of the drop are initially excited in the absence of the environment, the value of b_m characterizing a nonlinear frequency shift can also be represented in the form of a polynomial in Rayleigh parameter, while with other numerical coefficients. For instance, if modes 2 and 3 are initially excited, i.e., if $\Omega = \{2; 3\}$ and

$$\xi = \xi_0 P_0(\cos\vartheta) + \frac{\varepsilon}{2}(P_2(\cos\vartheta) + P_3(\cos\vartheta)),$$

for $t = 0$, the values of b_m characterizing the corrections to the frequencies of modes 2 and 3 have the form

$$b_2 = \frac{8}{2695}\frac{1}{\omega_2^3\omega_4^2(\omega_2^2 - 4\omega_3^2)(\omega_4^2 - 4\omega_2^2)((\omega_2 - \omega_3)^2 - \omega_5^2)((\omega_2 + \omega_3)^2 - \omega_5^2)}\sum_{i=0}^{7} B_i^{(2)} W^i,$$

$$b_3 = -\frac{20736}{11011}\frac{1}{\omega_2^2\omega_3\omega_4^2\omega_6^2(\omega_2^2 - 4\omega_3^2)(\omega_4^2 - 4\omega_3^2)(\omega_6^2 - 4\omega_3^2)}$$

$$\times \frac{1}{((\omega_2 - \omega_3)^2 - \omega_5^2)((\omega_2 + \omega_3)^2 - \omega_5^2)}\sum_{i=0}^{9} B_i^{(3)} W^i, \qquad (4)$$

Coefficients $B_i^{(m)}$, $m = 2, 3$, are given in Table 2.

It is easy to check that, as compared to (3), the number of resonance situations has increased and secondary combined resonances occur along with degenerate resonances. If one of the factors in the denominators of expressions (3) and (4) is close to zero, the frequency correction may become significant (which actually restricts the applicability of asymptotic expansion (1)).

With a charged drop placed in an dielectric environment, expressions (3) and (4) change. In this case, corrections to the frequencies of capillary vibrations of the drop can be represented by a double series in Rayleigh parameter W and environmental density ρ.

For example, if mode m alone is initially excited, i.e., if $\Omega = \{m\}$ and

$$\xi = \xi_0 P_0(\cos\vartheta) + \varepsilon P_m(\cos\vartheta),$$

for $t = 0$, the value of b_m characterizing a nonlinear frequency shift for $m = 2$, 3, or 4 can be represented in the form

$$b_2 = -\frac{108}{49(3 + 2\rho)^5(5 + 4\rho)^2\omega_2^3\omega_4^2(\omega_4^2 - 4\omega_2^2)}$$

$$\times \sum_{i=0}^{3}\sum_{j=0}^{4} A_{ij}^{(2)}\rho^i W^j,$$

$$b_3 = -\frac{311040}{1573(3 + 2\rho)^2(4 + 3\rho)^6(5 + 4\rho)^2(7 + 6\rho)^2}$$

Table 3. Coefficients $A_{ij}^{(2)}$

j	$i = 0$	$i = 1$	$i = 2$	$i = 3$
0	64×4645215	64×7062426	64×5046448	64×1507136
1	-96×2325675	-96×3487186	-96×2616248	-96×819616
2	4×15972795	4×23556618	4×18378864	4×5935648
3	-8×5935648	-8×1414089	-8×1133192	-8×373424
4	357525	464946	374928	124576

Table 4. Coefficients $A_{ij}^{(3)}$

j	$i = 0$	$i = 1$	$i = 2$
0	$960000 \times 629326659584$	$960000 \times 1983408150528$	$960000 \times 2712246080896$
1	$-8000 \times 92856752701440$	$-8000 \times 289533317285888$	$-8000 \times 39603182031120$
2	$400 \times 972791023411200$	$400 \times 2999075394326528$	$400 \times 4100298992386944$
3	$-40 \times 2807631731351552$	$-40 \times 8538676163356672$	$-40 \times 11634727972545728$
4	$2 \times 9605859672657920$	$2 \times 28671363515022336$	$2 \times 38689881616678592$
5	$-3 \times 645378748456960$	$-3 \times 1872016310675456$	$-3 \times 2470970580690368$
6	32×3297054503680	32×9124083418880	32×11489572520672
7	-4×592208700160	-4×1507906405696	-4×1707665679216
j	$i = 3$	$i = 4$	$i = 5$
0	$960000 \times 2061912758832$	$960000 \times 862494305220$	$960000 \times 153319441275$
1	$-8000 \times 304727092678800$	$-8000 \times 129734419287384$	$-8000 \times 23437700153415$
2	$400 \times 3189974081262768$	$400 \times 1379743042615128$	$400 \times 252718279917453$
3	$-40 \times 9125310799251600$	$-40 \times 3998375748074412$	$-40 \times 740549126229705$
4	$2 \times 30423104453043120$	$2 \times 13451299310165652$	$2 \times 2512567396279269$
5	$-3 \times 1928538595397616$	$-3 \times 855641056887444$	$-3 \times 160764059925333$
6	32×87119177625952	32×3843268025499	32×724042258758
7	-4×1190325182124	-4×506557263330	-4×94989141585

$$\times \frac{1}{\omega_2^2\omega_3\omega_4^2\omega_6^2(\omega_2^2-4\omega_3^2)(\omega_4^2-4\omega_3^2)(\omega_6^2-4\omega_3^2)} \times \sum_{i=0}^{5}\sum_{j=0}^{7} A_{ij}^{(3)}\rho^i W^j, \quad (5)$$

$$\times \frac{1}{\omega_2^2\omega_4^3\omega_6^2\omega_8^2(\omega_2^2-4\omega_4^2)(\omega_6^2-4\omega_4^2)(\omega_8^2-4\omega_4^2)} \times \sum_{i=0}^{5}\sum_{j=0}^{8} A_{ij}^{(4)}\rho^i W^j.$$

$$b_4 = \frac{7558272000}{347633(3+2\rho)^2(5+4\rho)^7(7+6\rho)^2(9+8\rho)^2}$$

Coefficients $A_{ij}^{(m)}$, $m = 2$–4, are given in Tables 3–5. If several modes of capillary vibrations of the drop

Table 5. Coefficients $A_{ij}^{(4)}$

j	$i = 0$	$i = 1$	$i = 2$
0	$2764800 \times 1322934089115625$	$2764800 \times 4614839816039500$	$2764800 \times 6807052318263600$
1	$-23040 \times 222024106178678125$	$-23040 \times 757301895008626500$	$-23040 \times 1106228391650769200$
2	$384 \times 7750478940098759375$	$384 \times 25994171837128242 7500$	$384 \times 37652477054961494800$
3	$-512 \times 1865619665028196875$	$-512 \times 6131134506449743750$	$-512 \times 8841970685535035200$
4	$16 \times 11649205690527234375$	$16 \times 37502380729932297500$	$16 \times 53536654328692102800$
5	$-32 \times 709290797232871875$	$-32 \times 2223032342251267500$	$-32 \times 3112803543724853600$
6	$16 \times 105081062194434375$	$16 \times 316698999015302500$	$16 \times 426967575505092800$
7	$-16 \times 4303646927409375$	$-16 \times 12184316425980000$	$-16 \times 15205018812313400$
8	1181694208996875	2999546960427500	3126793665757200
j	$i = 3$	$i = 4$	$i = 5$
0	$2764800 \times 5553451626253376$	$2764800 \times 2509695611457536$	$2764800 \times 484147691814912$
1	$-23040 \times 903633663255048128$	$-23040 \times 410913465860444160$	$-23040 \times 79705136823156736$
2	$384 \times 309193025535933212 80$	$384 \times 142023262317492510 72$	$384 \times 2777978728682921984$
3	$-512 \times 730155913776865529 60$	$-512 \times 3389390045560944128$	$-512 \times 668777032767583232$
4	$16 \times 443071633883332636 80$	$16 \times 207396142186855813 12$	$16 \times 41234019454435368 96$
5	$-32 \times 25606723830793385 60$	$-32 \times 12028427474555978 24$	$-32 \times 2404939626121973 76$
6	$16 \times 343281924045307200$	$16 \times 160381146213516032$	$16 \times 32163463331921920$
7	$-16 \times 11485079526504640$	$-16 \times 5226886014537856$	$-16 \times 1046883289309184$
8	1927439719822400	779373860226048	153211941490688

placed in an environment are initially excited, corrections to the frequencies of arbitrary initially excited modes can be represented by series of form (5), while with other denominators and numerical coefficients.

For instance, if initially modes 2 and 3 of capillary vibrations of the drop are excited, the values of b_m characterizing corrections to the frequencies of modes 2 and 3 can be written in the form of series

$$b_2 = \frac{1944}{539}\frac{1}{(2+\rho)(3+2\rho)^8(4+3\rho)^5(5+4\rho)^2(6+5\rho)^2}$$
$$\times \frac{1}{\omega_2^3\omega_4^2(\omega_2^2-4\omega_3^2)(\omega_4^2-4\omega_2^2)((\omega_2-\omega_3)^2-\omega_5^2)((\omega_2+\omega_3)^2-\omega_5^2)}$$
$$\times \sum_{i=0}^{11}\sum_{j=0}^{7} B_{ij}^{(2)}\rho^i W^j,$$

$$b_3 = -\frac{559872}{1573}$$
$$\times \frac{1}{(2+\rho)(3+2\rho)^5(4+3\rho)^8(5+4\rho)^2(6+5\rho)^2(7+6\rho)^2}$$

$$\times \frac{1}{\omega_2^2\omega_3\omega_4^2\omega_6^2(\omega_2^2-4\omega_3^2)(\omega_4^2-4\omega_3^2)(\omega_6^2-4\omega_3^2)}$$

$$\times \frac{1}{((\omega_2-\omega_3)^2-\omega_5^2)((\omega_2+\omega_3)^2-\omega_5^2)}\sum_{i=0}^{11}\sum_{j=0}^{9}B_{ij}^{(3)}\rho^i W^j.$$

Coefficients $B_{ij}^{(2)}$ and $B_{ij}^{(3)}$ are omitted, because the related expressions are too awkward.

Since the vibration frequencies depend on the drop-to-environment density ratio, the number of resonance situations (i.e., those where one of the conditions $\omega_m^2 = 2^2\omega_n^2$ and $\omega_m^2 = \omega_n^2 \pm \omega_k^2$ holds) increases significantly compared to the case of the drop vibrating in a vacuum.

The effect of nonlinear corrections to the critical conditions for instability of the drop against its charge near resonance situations deserves separate study. Here, we only state that corrections $b_n(\Omega)$, being of a resonance character, may influence the critical conditions of stability when the resonance value of parameter W is close to its critical value W_{cr}.

CONCLUSIONS

The value of the Rayleigh parameter that is critical in terms of instability of a drop against its charge depends on the vibration amplitude of the drop within the limits imposed by the contributions of each mode specifying the initial deformation of the drop on a nonlinear correction to the Rayleigh parameter.

APPENDIX

Expressions for the Coefficients of Expansion (1)

$$H_{kmln}^{2(\pm)(\pm)} = \sum_{g=1}^{\infty}(\beta_{kmgln}^{2(\pm)}\lambda_{lmg}^{(\pm)} + \mu_{kmgln}^{1(\pm)}) + \sum_{g=0}^{\infty}\mu_{kmgln}^{0(\pm)};$$

$$H_{kmln}^{1(\pm)(\mp)} = \sum_{g=1}^{\infty}(\beta_{kmgln}^{1(\pm)}\lambda_{lmg}^{(\pm)} + \mu_{kmgln}^{1(\mp)}) + \sum_{g=0}^{\infty}\mu_{kmgln}^{0(\mp)};$$

$$\lambda_{mln}^{(\pm)} = (\gamma_{mln} \pm \omega_m\omega_l\eta_{mln})/(\omega_n^2 - (\omega_m \pm \omega_l)^2);$$

$$\gamma_{mln} = (n+1)\chi_n K_{mln}(\omega_m^2(n-m+1+\rho n \times (n-m-1)/(n+1)) + 2n(l(l+1)-1) + ((l(m+1) - m(2m-2n+7)+3)nW/2) + (n+1)\chi_n\alpha_{mln}((1/m - n\rho/((n+1)(m+1)))\omega_m^2 + nW/2);$$

$$\eta_{mln} = (n+1)\chi_n K_{mln}(n/2 - m + 1 + \rho n(2m+3-n)/(2(n+1))) + (n+1)\chi_n\alpha_{mln}((1+n/(2l))/m - n\rho(n+2l+3)/(2(m+1)(l+1)(n+1)));$$

$$K_{mln} = (C_{m0l0}^{n0})^2;$$

$$\alpha_{mln} = -C_{m0l0}^{n0}C_{m(-1)l1}^{n0}\sqrt{m(m+1)l(l+1)};$$

$$\chi_n = (1+n(1+\rho))^{-1};$$

$$\Xi_n^0 = \omega_n^2 + n(n+1)\chi_n(n-1)(4+2n-5W);$$

$$\beta_{kmgln}^{1(\pm)} = \Pi_{kgn}^0 - \Pi_{kgn}^1\omega_k(\omega_l \pm \omega_m) - \Pi_{kgn}^2(\omega_l \pm \omega_m)^2;$$

$$\Xi_n^1 = ((n+1)(n-1) - 3n\rho)\chi_n;$$

$$\beta_{kmgln}^{2(\pm)} = \Pi_{kgn}^0 + \Pi_{kgn}^1\omega_k(\omega_l \pm \omega_m) - \Pi_{kgn}^2(\omega_l \pm \omega_m)^2;$$

$$\Xi_n^2 = n(n-1)\chi_n;$$

$$\mu_{kmgln}^{1(\pm)} = \Lambda_{kmgln}^1 \pm \Gamma_{kmgln}^1\omega_m\omega_k;$$

$$\mu_{kmgln}^{0(\pm)} = \Lambda_{kmgln}^0 \pm \Gamma_{kmgln}^0\omega_m\omega_k;$$

$$\Lambda_{kmgln}^0 = (n+1)\chi_n\omega_k^2 K_{gln}(\alpha_{kmg}(k-2)/k + (k-1)(n-k+2)K_{kmg}/2) + \rho n\chi_n\omega_k^2\Big(((g+1-n)K_{gln} - \alpha_{gln}/(g+1))((k+2)K_{kmg} - \alpha_{kmg}/(k+1)) + ((k+3)\alpha_{kmg}/(k+1) + (k+2)(n-2-k)K_{kmg}/2)K_{gln}\Big)$$

$$+ n(n+1)\chi_n\Big(WK_{gln}((k^3 - 2(m+1)(m+2) - k^2(n-9) - k(2m(m+3)+3n-22))K_{kmg} - 2(k+2)\alpha_{kmg})/2 - ((3k(k+1)-2)K_{kmg} - l(l+1)\alpha_{kmg}/2)K_{gln} + \alpha_{kmg}\Big(l^2K_{lgn} - \sum_{\nu=1}^{[l/2]}(2l-4\nu+1)K_{l-2\nu,g,n}\Big)\Big);$$

$$\Lambda_{kmgln}^1 = n(n+1)\chi_n Wk K_{kmg}((g+1)(l-2-g+n)K_{lgn} + \alpha_{lgn}) + (n+1)\chi_n((\alpha_{lgn}/g + (n+1-g)K_{lgn}) \times (\alpha_{kmg}/m + (1-m)K_{kmg}))\omega_m^2;$$

$$\Gamma^0_{kmgln} = (n+1)\chi_n((k-2)(K_{gln}((k-1)K_{kmg}/2 - \alpha_{kmg}/k) + K_{mgn}((k-1)K_{klg}/2 - \alpha_{klg}/k)) - n(k-1)K_{gln} \times (\alpha_{kmg}/(mk) + K_{kmg}) - \rho n\chi_n(((g+2)K_{mgn} - \alpha_{mgn}/(g+1))((k+2)K_{klg} - \alpha_{klg}/(k+1)) + (k+3)K_{mgn}(\alpha_{klg}/(k+1) - (k+2)K_{klg}/2) + ((g+2)K_{gln} - \alpha_{gln}/(g+1))((k+2)K_{kmg} - \alpha_{kmg}/(k+1)) + (k+3)K_{gln}(\alpha_{kmg}/(k+1) - (k+2)K_{kmg}/2) - (n+1)((\alpha_{mgn}/((m+1)(g+1)) + K_{mgn})((k+2)K_{klg} - \alpha_{klg}/(k+1)) + K_{gln}((k+2)K_{kmg} - \alpha_{kmg}/(k+1)) - (\alpha_{kmg}/((k+1)(m+1)) + K_{kmg})(k+2)K_{gln}));$$

$$\Gamma^1_{kmgln} = (n+1)\chi_n((\alpha_{lgn}/g + (n+1-g)K_{lgn}) \times ((m-1)K_{kmg} - \alpha_{kmg}/m) + ((k+n)\alpha_{kgn}/(gk) + (n+1-g)K_{kgn})((m-1)K_{mlg} - \alpha_{mlg}/m));$$

$$\Pi^0_{kmn} = (n+1)\chi_n(nK_{kmn}(2((k-1)(k+2) + m(m+1)) + W(k-1)(n-5-k)) + (\alpha_{kmn}/k + (n+1-k)K_{kmn})\omega_k^2) - \rho n\chi_n\omega_k^2((n-1-k)K_{kmn} + \alpha_{kmn}/(k+1)) + n(n+1)\chi_n W((m+1)(k+n-m-2)K_{kmn} + \alpha_{kmn});$$

$$\Pi^1_{kmn} = (n+1)\chi_n((k+m-n-2)K_{kmn} - (n+k+m)\alpha_{kmn}/(mk)) + n\rho\chi_n((n-k-m-3)K_{kmn} + (k+m+n+3)\alpha_{kmn}/((m+1)(k+1)));$$

$$\Pi^2_{kmn} = (n+1)\chi_n((m-n-1)K_{kmn} - \alpha_{kmn}/m) + n\rho\chi_n((n-m-1)K_{kmn} + \alpha_{kmn}/(m+1)).$$

Here, C^{n0}_{m0l0} and $C^{n0}_{m(-1)l1}$ are the Clebsch–Gordan coefficients and δ_{kn} is the Kronecker delta.

ACKNOWLEDGMENTS

The work was supported by the Russian Foundation for Basic Research (grant no. 03-01-00760) and by a grant from the President of the Russian Federation (grant no. MK-2946-2004-1).

REFERENCES

1. J. W. Strutt (Lord Rayleigh), Philos. Mag. **14**, 184 (1882).
2. A. Doyle, D. R. Moffet, and B. Vonnegut, J. Colloid Sci. **19**, 136 (1964).
3. T. G. O. Berg, R. J. Trainor, and U. Vaughan, J. Atmos. Sci. **27**, 1173 (1970).
4. J. W. Schweizer and D. N. Hanson, J. Colloid Interface Sci. **35**, 417 (1971).
5. M. Roulleau and M. Desbois, J. Atmos. Sci. **29**, 565 (1972).
6. D. Duft, H. Lebius, B. A. Huber, *et al.*, Phys. Rev. Lett. **89** (8), 084503 (2002).
7. A. I. Grigor'ev and S. O. Shiryaeva, J. Aerosol Sci. **25**, 1079 (1994).
8. A. I. Grigor'ev, Zh. Tekh. Fiz. **71** (10), 1 (2001) [Tech. Phys. **46**, 1205 (2001)].
9. J. A. Tsamopolous and R. A. Brown, J. Fluid Mech. **147**, 373 (1984).
10. J. A. Tsamopolous, T. R. Akylas, and R. A. Brown, Proc. R. Soc. London, Ser. A **401**, 67 (1985).
11. Z. C. Feng, J. Fluid Mech. **333**, 1 (1997).
12. A. N. Zharov, S. O. Shiryaeva, and A. I. Grigor'ev, Zh. Tekh. Fiz. **73** (12), 9 (2003) [Tech. Phys. **48**, 1511 (2003)].
13. A. N. Zharov, S. O. Shiryaeva, and A. I. Grigor'ev, Zh. Tekh. Fiz. **74** (7), 19 (2004) [Tech. Phys. **49**, 824 (2004)].
14. A. I. Grigor'ev, Zh. Tekh. Fiz. **55**, 1272 (1985) [Sov. Phys. Tech. Phys. **30**, 736 (1985)].
15. A. I. Grigor'ev and S. O. Shiryaeva, Zh. Tekh. Fiz. **61** (3), 19 (1991) [Sov. Phys. Tech. Phys. **36**, 258 (1991)].
16. A. I. Grigor'ev, Zh. Tekh. Fiz. **71** (10), 1 (2001) [Tech. Phys. **46**, 1457 (2001)].
17. K. V. Beard, Radio Sci. **19**, 67 (1984).
18. I. P. Mazin, A. Kh. Khrgian, and I. M. Imyanitov, *Clouds and Cloud Atmosphere: A Handbook* (Gidrometeoizdat, Leningrad, 1989) [in Russian].

Translated by A. Pankrat'ev

Technical Physics, Vol. 50, No. 8, 2005, pp. 1016–1020. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 54–58.

GAS DISCHARGES, PLASMA

Interaction of Electromagnetic Oscillations with Charged Particle Flows in Nonuniform Plasma-Like Structures

S. I. Khankina*, V. M. Yakovenko*, and I. V. Yakovenko**

** Usinkov Institute of Radiophysics and Electronics, National Academy of Sciences of Ukraine, Kharkov, 61085 Ukraine*

*** Molniya Research and Design Institute, Kharkov, 61013 Ukraine*

Received November 3, 2004

Abstract—A hydrodynamic theory is developed of the interaction of electromagnetic oscillations with a monoenergetic charged particle beam propagating through a structure consisting of plasma and dielectric layers bounded by perfectly conducting planes. It is shown that, in such a system, plasma oscillations are excited due to the transformation of space charge oscillations of the particle flow into plasma oscillations at the layer boundaries. The regions of generation and decay of plasma oscillations are determined. © *2005 Pleiades Publishing, Inc.*

An important line of investigation in modern radiophysics is the exploration and utilization of millimeter and submillimeter electromagnetic wavelength ranges. The foremost problem in this field of research is the development of sources of such oscillations. Different approaches have been used to solve this problem [1, 2]. One of them involves the search for unstable states in solid plasmalike media [3, 4]. It should be emphasized that modern technology allows one to create various conducting solid structures: films, semiconductors with a superlattice and two-dimensional (2D) electronic gas, metal–dielectric–semiconductor structures, etc. In such sub-μm structures, a peculiar type of plasma oscillations caused by the presence of boundaries can arise; moreover, ballistic mechanisms for charge transfer can also operate in them. This leads to the onset of instabilities caused by the Cherenkov effect, as well as by the transition and deceleration radiation from charged particles [4, 5].

In the present study, we consider the interaction of electromagnetic oscillations with a charged particle (electron) beam propagating through a nonuniform plasmalike medium. It is assumed that an unbounded monoenergetic nonrelativistic electron beam passes through a structure consisting of layers with different electromagnetic properties and bounded by perfectly conducting planes.

The aim of our study is to determine the eigenfrequencies of electromagnetic oscillations in such a structure, as well as the rates with which these oscillations are excited (damped) due to the transformation of the beam kinetic energy into electromagnetic energy and vice versa.

The interaction of a particle beam with a cold bounded plasma can be described in terms of either particle–wave or wave–wave interaction. In the first case, the electron beam is described by a kinetic equation and the perturbed particle flow is regarded as an ensemble of individual perturbations [6, 7]. In the second case, a hydrodynamic approach is used and the beam interacts with the medium through space charge waves (SCWs). It is the latter approach (i.e., one based on the electrostatic and hydrodynamic equations) that is used in this paper. The use of the electrostatic approximation is justified by the assumption that the phase velocities of the excited waves are small compared to the speed of light.

If we ignore the thermal motion of the beam particles, then the beam–medium interaction can be described by the following equations:

$$\nabla \times \mathbf{E} = 0, \quad \nabla \cdot \mathbf{D} = 4\pi e n, \tag{1}$$

$$\frac{\partial n}{\partial t} + \nabla \cdot (n_0 \mathbf{v} + n \mathbf{v}_0) = 0, \quad \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}_0 \cdot \nabla)\mathbf{v} = \frac{e}{m_0}\mathbf{E}. \tag{2}$$

Here, e and m_0 are the charge and mass of the beam electrons; n_0, n, $\mathbf{v}_0$, and $\mathbf{v}$ are the unperturbed (constant) and perturbed values of their density and velocity; and $\mathbf{v}_0 \parallel 0Y$.

The electromagnetic induction $\mathbf{D}$ in each layer is related to the electric field $\mathbf{E}$ by the material equation. In the plasma layer, it has the form

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + 4\pi \int_{-\infty}^{t} \mathbf{j}(t')dt', \tag{3}$$

where ε_0 is the permittivity of the lattice.

The current of the conduction electrons $\mathbf{j} = eN_0\mathbf{u}$ satisfies the continuity equation

$$e\frac{\partial N}{\partial t} + \nabla \cdot \mathbf{j} = 0. \tag{4}$$

1063-7842/05/5008-1016$26.00

Poisson's equation can be written as

$$\nabla \cdot (\varepsilon_0 \mathbf{E}) = 4\pi e(N + n). \quad (5)$$

The electron velocity **u** can be found from the equation of motion

$$\frac{\partial \mathbf{u}}{\partial t} = \frac{e}{m}\mathbf{E}. \quad (6)$$

Here m, N_0, and N are the effective mass and the equilibrium and perturbed densities of the conduction electrons, respectively. In the dielectric layer, we have $N_0 = 0$ and the dielectric constant ε is assigned an index corresponding to the layer number.

The coordinate system is chosen as follows: The planes $y = -d_1$ and $y = d_2$ are perfectly conducting. Regions 1 ($-d_1 \le y \le -a/2$) and 2 ($a/2 \le y \le d_2$) are occupied by media with permittivities ε_1 and ε_2, respectively. Between them ($-a/2 \le y \le a/2$), there is a plasma layer. The electron beam propagates along the Y axis from the cathode ($y = -d_1$) toward the anode ($y = d_2$).

We assume that all the variables in Eqs. (1)–(6) depend on the spatial coordinates and time as $\sim\exp[i(\mathbf{q}\cdot\mathbf{r} - \omega t)]$, where ω is the frequency and $\mathbf{q} = (q_x, q_y, 0)$ is the wave vector. Along the Z axis, the oscillations are assumed to be uniform. The solution to set of Eqs. (1)–(6) in medium 1 has the form (the factor $\exp[i(q_x x - \omega t)]$ is omitted)

$$n(y) = \sum_{k=1}^{2} A_k \exp(iq_{yk}y), \quad (7)$$

$$E_x(y) = E_x^l(y) + E_x^t(y); \quad E_y(y) = -\frac{i}{q_x}\frac{\partial E_x}{\partial y};$$

$$E_x^l(y) = -\frac{4\pi i e q_x}{\varepsilon_1}\sum_{k=1}^{2}\frac{A_k}{q_{yk}^2}\exp(iq_{yk}y); \quad (8)$$

$$E_x^t(y) = \sum_{k=1}^{2} B_k \exp(\chi_k y),$$

$$v_x(y) = v_x^l(y) + v_x^t(y), \quad v_y(y) = -\frac{i}{q_x}\frac{\partial v_x}{\partial y},$$

$$v_x^l(y) = \frac{1}{n_0}\sum_{k=1}^{2}\frac{(\omega - q_{yk}v_0)}{q_{yk}}A_k\exp(iq_{yk}y); \quad (9)$$

$$v_x^t(y) = \frac{ie}{m_0\omega}E_x^t(y).$$

In Eqs. (7)–(9), the following notation is introduced: $q_{y1,2} = \omega/v_0 \pm q_1$, $q_1 = \omega_b/v_0\sqrt{\varepsilon_1}$, $\chi_{1,2} = \pm q_x$, $\omega_b^2 = 4\pi e^2 n_0/m_0$, and A_k and B_k are arbitrary constants.

The solutions in the plasma layer and medium 2 are analogous to that in medium 1. It is only necessary to introduce new arbitrary constants and replace ε_1 with ε_2 in medium 2 and with

$$\varepsilon(\omega) = \varepsilon_0 - \frac{\omega_0^2}{\omega^2}\left(\omega_0^2 = \frac{4\pi e^2 N_0}{m}\right)$$

in the plasma layer.

It can be seen that there are two types of electrostatic fields in the structure: transverse fields ($\mathbf{q}\cdot\mathbf{E}^t = 0$) and longitudinal fields $\mathbf{E}^l$ (fast and slow SCWs), for which $\mathbf{q}\cdot\mathbf{E}^l \ne 0$.

To solve the problem, set of Eqs. (1)–(6) should be supplemented with boundary conditions. Obviously, at the perfectly conducting boundaries, the tangential components of the electric field are equal to zero; i.e.,

$$E_x\big|_{y=-d_1} = 0, \quad E_x\big|_{y=d_2} = 0. \quad (10)$$

Since SCWs are present in the structure, it is also necessary to formulate boundary conditions for hydrodynamic quantities (along with conventional boundary conditions for electrodynamic quantities), i.e., for the perturbed density and velocity of the electron beam. These are the conditions under which the electron density and the normal component of the electron velocity are zero on the cathode surface ($y = -d_1$):

$$n_1\big|_{y=-d_1} = 0, \quad v_{y1}\big|_{y=-d_1} = 0. \quad (11)$$

These conditions are known as the Pierce boundary conditions [8].

At the dielectric–plasma interface, the tangential component of the electric field and the normal component of the electric induction are continuous. Moreover, the mass and momentum fluxes must be continuous there; i.e., the perturbed density and the normal component of the beam velocity must be the same on both sides of the dielectric–plasma interface. To avoid laborious calculations, we will consider a thin plasma layer, such that $q_x a \ll 1$ and $a\omega/v_0 \ll 1$.

In this case, the boundary conditions at $y = -a/2$ and $y = a/2$ can be replaced with those at $y = 0$; i.e., E_{x1}, n, and v_y are continuous at $y = 0$:

$$E_{x1} = E_{x2}, \quad n_1 = n_2, \quad v_{y1} = v_{y2}. \quad (12)$$

However, the normal component of the electric induction D_y undergoes discontinuity caused by the charges induced in the thin plasma layer.

To find the boundary conditions for D_y, we integrate Poisson's equation over the plasma layer (from $-a/2$ to $a/2$). As a result, we obtain

$$D_y\left(\frac{a}{2}\right) - D_y\left(-\frac{a}{2}\right) + iq_x D_x(0)a = -4\pi e[N(0) - n(0)]a. \quad (13)$$

Here, the quantities D_x, N, and n are replaced with their values at $y = 0$.

Integrating the continuity equations over the plasma layer and taking into account the equations of motion, we obtain the following expressions for $N(0)$ and $n(0)$:

$$N(0) = \frac{ieq_x N_0}{m\omega}E_x(0); \quad n(0) = \frac{ieq_x n_0}{m_0\omega}E_x(0). \quad (14)$$

For $q_x a \ll 1$, $\varepsilon_0 \sim \varepsilon_1 \sim \varepsilon_2$, and $\omega_b^2/\omega^2 \ll 1$, expression (11) can be written as

$$D_{y2}(0) - D_{y1}(0) = \frac{i\omega_0^2}{\omega^2}q_x a E_{x1}(0). \quad (15)$$

From the boundary conditions, we find the dispersion relation describing the interaction of plasma oscillations with the beam electrons:

$$\omega^2(\varepsilon_1 \coth q_x d_1 + \varepsilon_2 \coth q_x d_2) - \omega_0^2 q_x a = 2\frac{i\omega_b^2 q_x v_0}{\omega}\Gamma, \quad (16)$$

where

$$\Gamma = -\Gamma_0 + \Gamma_1 \exp[i\omega\tau_1] + \Gamma_2 \exp[i\omega\tau_2] + \Gamma_3 \exp[i\omega\tau],$$

$$\Gamma_0 = \coth^2 q_x d_1 + \coth^2 q_x d_2 + \coth q_x d_1 \coth q_x d_1 - 1,$$

$$\Gamma_j = (\coth q_x d_1 + \coth q_x d_2)\frac{s_j}{\sinh q_x d_j},$$

$$s_j = \cos q_j d_j - i\frac{\omega \sin q_j d_j}{2q_j v_0} \ (j = 1, 2),$$

$$\Gamma_3 = \frac{1}{\sinh q_x d_1 \sinh q_x d_2}\left[\frac{i\omega}{2q_1 v_0}\sin q_1 d_1 - s_2 \cos q_1 d_1\right],$$

$$d = d_1 + d_2, \quad \tau_j = \frac{d_j}{v_0} \ (j = 1, 2), \quad \tau = \frac{d}{v_0}.$$

In deriving Eq. (16), the following inequalities were used:

$$q_x \ll \frac{\omega}{v_0}, \quad \omega_b^2 \ll \omega^2. \quad (17)$$

The left-hand side of Eq. (16) determines the spectrum of electromagnetic oscillations in the structure in the absence of a beam.

If the permittivities of media 1 and 2 are independent of the frequency (i.e., $\varepsilon_1 = \text{const}$ and $\varepsilon_2 = \text{const}$), then the frequency of electromagnetic eigenmodes is determined by

$$\omega^2 = \frac{\omega_0^2 q_x a}{\varepsilon_1 \coth q_x d_1 + \varepsilon_2 \coth q_x d_2}. \quad (18)$$

For $a = 0$, eigenmodes exist only if one of the permittivities ε_1 and ε_2 is negative. This can occur, e.g., in a metal–dielectric–semiconductor–metal structure ($a = 0$, $\varepsilon_1 = \varepsilon_d = \text{const}$, and $\varepsilon_2 = \varepsilon_0 - \Omega^2/\omega^2$). In this case, we have

$$\omega^2 = \frac{\Omega^2}{\varepsilon_d \coth q_x d_1 + \varepsilon_0 \coth q_x d_2}. \quad (19)$$

The right-hand side of Eq. (16) describes the interaction of plasma oscillations with a charged particle beam propagating through a nonuniform medium. The first term does not contain exponential terms and describes the transformation of transverse oscillations into SCWs at the boundary $y = 0$. Indeed, for an isolated boundary ($q_x d_{1,2} \gg 1$, $q_{1,2} d_{1,2} \gg 1$), all the terms on the right-hand side of Eq. (16) (except for the first one) disappear. In this case, the density of a particle beam incident on a dielectric–semiconductor interface is smooth. The modulation of the beam by a transverse field occurs at the interface. The electron density perturbations are then carried away by the particle beam deep into the medium ($y > 0$); as a result, plasma oscillations decay. The frequency and decay rate of the plasma oscillations ($\omega = \omega' + \delta\omega$, $\delta\omega/\omega' \ll 1$) are equal to

$$\omega' = \omega_0\left(\frac{q_x a}{\varepsilon_1 + \varepsilon_2}\right)^{1/2}, \quad (20)$$

$$\mathrm{Im}\,\delta\omega = -\frac{2\omega_b^0 v_0}{\omega_b^2 a}. \quad (21)$$

Introducing the surface density $N_s = N_0 a$, we obtain

$$\mathrm{Im}\,\delta\omega = -\frac{2n_b v_0}{N_s}. \quad (22)$$

Note that collisionless decay takes place only when the condition $2\omega_b^2 v_0/\omega_0^2 a \geq \nu/2$ (where ν is the characteristic frequency of the conduction electrons) is satisfied.

If the widths of layers 1 and 2 are finite, feedback between SCWs and transverse oscillations arises due to their mutual transformations at the boundaries $y = -d_1$ and $y = d_2$. These transformations are characterized by the corresponding exponential factors. Depending on the relation between the oscillation period and the time during which the particles fly between the boundaries, the oscillations can grow or decay.

Let us consider the effect of the boundaries $y = -d_1$ and $y = d_2$ in some particular cases. We assume that the conditions $q_x d_1 \gg 1$, $q_1 d_1 \gg 1$, and $q_2 d_2 \ll 1$ are satisfied. In this case, the frequency and growth (decay) rate of plasma oscillations are

$$\omega'^2 = \frac{\omega_0^2 q_x a}{\varepsilon_1 + \varepsilon_2 \coth q_x d_2}, \quad (23)$$

$$\mathrm{Im}\delta\omega = \frac{\omega_b^2 v_0}{\omega^2 a}\Gamma, \qquad (24)$$

where

$$\Gamma = \frac{1+\coth q_x d_2}{\sinh q_x d_2}\varphi(q_x, \omega\tau_2);$$

$$\varphi(q_x, \omega\tau_2) = \cos\omega\tau_2 + \frac{\omega\tau_2}{2}\sin\omega\tau_2 - \cosh q_x d_2.$$

The function $\varphi(q_x, \omega\tau_2)$ can assume both positive and negative values. In regions where $\varphi(q_x, \omega\tau_2)$ is positive, the amplitude of oscillations increases (the generation zones). Thus, $\varphi(q_x, \omega\tau_2) > 0$ if $\sin\omega\tau_2 > 0$ and $\omega\tau_2/2 \gg \cosh q_x d_2$.

Figure 1 shows the function $\varphi(\omega\tau_2)$ at $q_x d_2 \ll 1$. It can be seen that the first generation zone lies within the interval $2\pi < \omega\tau_2 < 3\pi$. Within each generation zone, the function $\varphi(\omega, \tau_2)$ reaches its maximum and the growth rate of oscillations is maximum at $\omega\tau_2 = \pi/2 + 2\pi l$ (where $l = 1, 2, 3, \ldots$ is the number of the generation zone). This can occur, e.g., when the electron beam passes through a thin metal plate with openings. It is known [9] that, in this case, plasma oscillations are excited in the plate.

If $a = 0$ and the permittivity of medium 2 is a function of frequency ($\varepsilon_2 = \varepsilon_2(\omega)$), then, at $q_x d_1 \gg 1$, the electron beam propagating through a vacuum–semiconductor–metal structure excites oscillations with the frequency and growth rate

$$\omega' = \frac{\Omega}{(\varepsilon_0 + \varepsilon_d \tanh q_x d_2)}, \qquad (25)$$

$$\mathrm{Im}\delta\omega = \frac{i\omega_b^2 q_x v_0}{\Omega^2}\Gamma. \qquad (26)$$

The interaction between a particle beam and the oscillations with frequency (25) can be realized experimentally by passing the beam through openings in a semiconductor plate placed on a metal substrate. Since the hydrodynamics boundary conditions at the planes $y = -d_1$ and $y = 0$ are nearly the same, a similar process takes place in a metal–dielectric–semiconductor structure at $q_x d_2 \gg 1$ and $q_2 d_2 \gg 1$. In this case, the frequency and growth (decay) rate of oscillations are described by expressions (23)–(26), in which d_2 should be replaced with d_1, ε_1 with ε_2, and ε_2 with ε_1.

Finally, the eigenfrequency of a metal–dielectric–plasma layer–dielectric–metal structure ($\varepsilon_1 = \mathrm{const}$, $\varepsilon_2 = \mathrm{const}$) at $q_x d_{1,2} \ll 1$ and $q_{1,2} d_{1,2} \ll 1$ is equal to

$$\omega' = \omega_0 q_x \left(\frac{a d_1 d_2}{\varepsilon_1 d_2 + \varepsilon_2 d_1}\right)^{1/2}. \qquad (27)$$

When the current passes through such a structure,

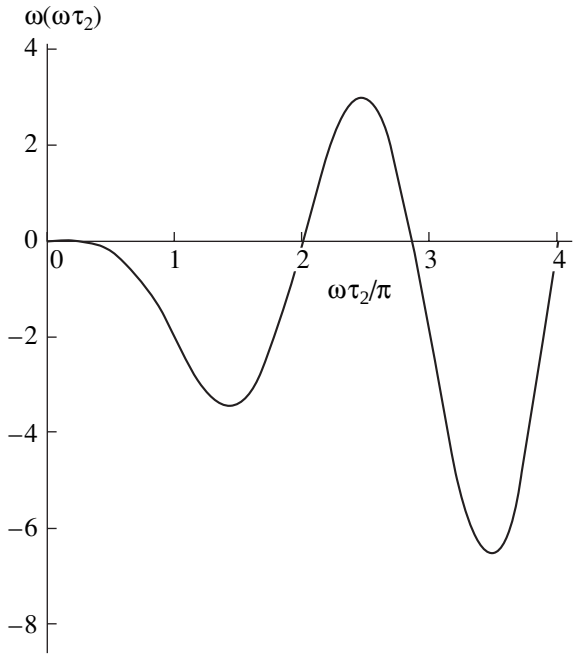


Fig. 1. Regions of the growth and decay of oscillations in a metal–dielectric–semiconductor–dielectric–metal structure at $d_1 \longrightarrow \infty$ and $q_x d_2 \ll 1$.

we have

$$\mathrm{Im}\delta\omega = \frac{\omega_b^2 v_0}{\omega_0^2 a q_x^2 d_1 d_2}\psi(\omega\tau_1, \omega\tau_2; \omega\tau), \qquad (28)$$

where

$$\psi(\omega\tau_1, \omega\tau_2; \omega\tau) = \frac{d}{d_1}\left(\cos\omega\tau_1 + \frac{\omega\tau_1}{2}\sin\omega\tau_1 - 1\right)$$
$$+ \frac{d}{d_2}\left(\cos\omega\tau_2 + \frac{\omega\tau_2}{2}\sin\omega\tau_2 - 1\right)$$
$$- \left(\cos\omega\tau + \frac{\omega\tau}{2}\sin\omega\tau - 1\right).$$

The function $\psi = \psi(\omega\tau_2)$ at different ratios between d_1 and d_2 is shown in Fig. 2. A comparison between the functions ψ and $\varphi(\omega\tau_2)$ at $d_1 \longrightarrow \infty$ (see Fig. 1) shows that the finite value of d_1 leads to more beneficial conditions for the generation of plasma oscillations, because, in this case, the first generation zone corresponds to the lower values of $\omega\tau_2$: $0 < \omega\tau_2 < \pi$.

Let us make some estimates. At a frequency of $\omega' \approx 3 \times 10^{12}$ s^{-1}, $q_x = 10^3$ cm^{-1}, $a \sim d_1 \sim d_2 \approx 10^{-4}$ cm, and $\varepsilon_1 = \varepsilon_2 = 4$, the density of the conduction electrons should be $N_0 \sim 10^{20}$ cm^{-3}. When injecting electrons with a velocity of $v_0 \approx 3 \times 10^8$ cm/s and density of $n_b \sim 10^{13}$ cm^{-3}, the growth rate of plasma oscillations is equal to $\mathrm{Im}\delta\omega \sim 10^{10}$ s^{-1}. In order for the instability to

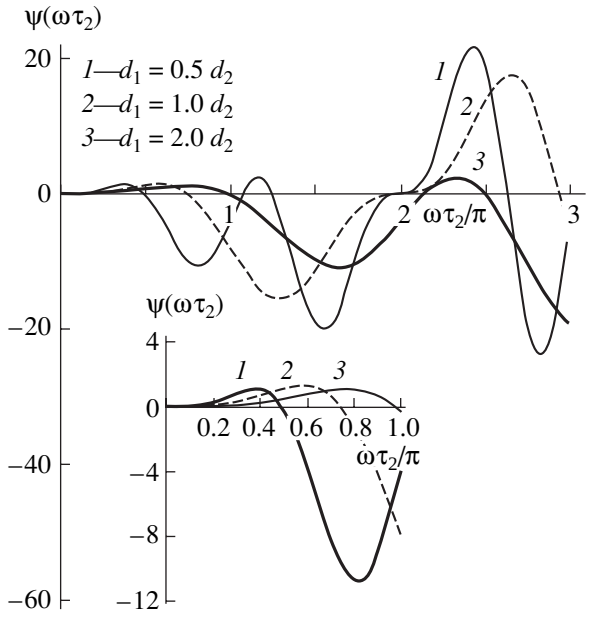


Fig. 2. Comparison of the growth (decay) rates of oscillations at different ratios between the dimensions of the dielectric layers separating the conducting plate from the metal substrate.

develop, it is necessary that $\mathrm{Im}\,\delta\omega$ be larger than the collisional decay rate ($\mathrm{Im}\,\delta\omega > \nu/2$).

Therefore, a charged particle beam propagating through a structure consisting of layers with different electromagnetic properties can generate plasma oscillations. Whether the oscillations will be excited depends on the ratio of the oscillation period to the time during which the particles fly between the boundaries of the layers. The oscillation frequency depends on the properties of the plasma layer. The generation mechanism is determined by the transformation of the SCW energy into plasma oscillations at the layer boundaries.

REFERENCES

1. A. V. Andronov, A. O. Zakhar'in, I. N. Yassievich, and N. N. Zinov'ev, Pis'ma Zh. Éksp. Teor. Fiz. **79**, 448 (2004) [JETP Lett. **79**, 365 (2004)].
2. Yu. A. Morozov, I. S. Nefedov, and V. Ya. Aleshkin, Zh. Tekh. Fiz. **74** (5), 71 (2004) [Tech. Phys. **49**, 592 (2004)].
3. A. B. Mikhailovskii, *Theory of Plasma Instabilities* (Atomizdat, Moscow, 1971; Consultants Bureau, New York, 1974).
4. N. N. Beletskiĭ, A. A. Bulgakov, S. I. Khankina, and V. M. Yakovenko, *Plasma Instabilities and Nonlinear Phenomena in Semiconductors* (Naukova Dumka, Kiev, 1984) [in Russian].
5. N. N. Beletskiĭ, V. M. Svetlichnyĭ, D. D. Khalameĭda, and V. M. Yakovenko, *Microwave Phenomena in Heterogeneous Semiconductor Structures* (Naukova Dumka, Kiev, 1991).
6. V. M. Yakovenko and I. V. Yakovenko, Radiofiz. Radioastron. **4**, 376 (1999).
7. V. M. Yakovenko and I. V. Yakovenko, Dokl. Akad. Nauk Ukr., No. 1, 70 (2000).
8. J. J. Pierce, Appl. Phys. **15**, 7211 (1944).
9. T. W. Ebbesen *et al.*, Nature **391**, 667 (1998); Usp. Fiz. Nauk **169**, 1272 (1999) [Phys. Usp. **42**, 1174 (1999)].

Translated by B. Chernyavskiĭ

Technical Physics, Vol. 50, No. 8, 2005, pp. 1021–1025. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 59–62.

SOLIDS

Nonlinear Dynamics of Cubic Ferromagnets in the Range of Magnetoelastic Resonance

R. M. Vakhitov and O. G. Ryakhova

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

e-mail: VakhitovRM@bsu.bashedu.ru

Received June 25, 2004

Abstract—The spectrum of nonlinear magnetoelastic waves in cubic garnet ferrite crystals with induced uniaxial anisotropy along the [111] axis is theoretically investigated. It is demonstrated that taking into account a voltage mismatch between neighboring domains raises the magnetostriction energy and makes description of the magnetoelastic dynamics in such magnets more involved. For magnetoelastic waves propagating along the [111] axis, the problem reduces to the solution of the triple sin-Gordon equation. The topology of possible solitary waves is analyzed. The conditions are found under which the velocity of stationary motion of domain walls may exceed the sound velocity in the material. © *2005 Pleiades Publishing, Inc.*

It is known that magnetoelastic (ME) interaction, being rather weak in most magnets, can be significantly enhanced under certain conditions and give rise to a number of dynamic ME effects [1, 2]. Specifically, it was observed [3] that the dynamics of domain walls (DWs) in weak ferromagnets, such as $YFeO_3$, exhibits ME anomalies influencing the DW motion. Such behavior of DWs is due to the fact that their limiting velocity in these materials reaches values close to $V_{\lim} = 2 \times 10^4$ m/s, which provides conditions for an ME resonance. A similar situation may occur in other magnets of this class, in which $V_{\lim}$ exceeds the sound velocity.

At present, it is customary to assume that, normally, the limiting velocity of DWs in ferromagnets (such as garnet ferrites) does not exceed 2×10^3 m/s, which is considerably below the sound velocity in these compounds [4]. However, under certain conditions, $V_{\lim}$ in ferromagnetic crystals may overcome this threshold. Accordingly, the nonlinear ME effects that arise in weak ferromagnets may also occur in materials like garnet ferrites.

In this paper, we analyze the range of DW limiting velocities (versus the material parameters) in cubic crystals such as garnet ferrites with induced uniaxial anisotropy (IUA) along the [111] axis (an analogue of epitaxially grown garnet ferrite films with a developed (111) surface is a (111) wafer [4]). Also, possible types of solitary ME waves are analyzed and their structure is determined.

Consider the propagation of ME waves in a (111) ferromagnetic plate. We assume that $\mathbf{M} = \mathbf{M}(z, t)$ and the coordinate system is arranged so that $OX \parallel [11\bar{2}]$, $OY \parallel [\bar{1}10]$, and $OZ \parallel [111]$. The energy density of such a magnet can be represented in the form

$$\begin{aligned} E &= A[(\theta')^2 + \sin^2\theta(\varphi')^2] + K_u\sin^2\theta \\ &+ K_1\left[\frac{1}{4}\sin^4\theta + \frac{1}{3}\cos^4\theta + \frac{\sqrt{2}}{3}\sin^3\theta\cos\theta\cos 3\varphi\right] \\ &+ \frac{K_2}{54}\left[\sin^3\theta\cos 3\varphi - \frac{\sqrt{2}}{2}\cos\theta(3\sin^2\theta - 2\cos^2\theta)\right]^2 \\ &+ \frac{1}{3}B_1\left[u_x'\frac{\sqrt{2}}{2}(-\sin^2\theta\cos 2\varphi + \sqrt{2}\sin 2\theta\cos\varphi)\right. \\ &+ u_y'\sqrt{2}\sin\theta\sin\varphi(\sin\theta\cos\varphi \\ &+ u_y'\sin\theta\sin\varphi(\cos\theta - \sqrt{2}\sin\theta\cos\varphi) \\ &\left.+ u_z'(3\cos^2\theta - 1) + \sqrt{2}\cos\theta) + u_z'\right] \\ &+ \frac{1}{3}B_2\left[u_x'\frac{\sqrt{2}}{2}\left(\sin^2\theta\cos 2\varphi + \frac{\sqrt{2}}{2}\sin 2\theta\cos\varphi\right)\right. \\ &+ \frac{1}{6}((u_x')^2 + (u_y')^2)(C_{11} - C_{12} + C_{44}) \\ &\left.+ \frac{1}{6}(u_z')^2(C_{11} + 2C_{12} + 4C_{44})\right], \end{aligned} \tag{1}$$

where θ and φ are the polar and azimuth angles of magnetization vector **M**, A is the exchange parameter, K_u is the IUA constant, K_1 and K_2 are the first and second constants of cubic anisotropy (CA), u_i are the components of displacement vector **u**; B_i and C_{ij} are magnetoelastic and elastic constants, and M_s is the saturation magnetization. The prime means differentiation with respect to the z coordinate.

1063-7842/05/5008-1021$26.00

To investigate the ME dynamics of the magnets, we proceed from the Landau–Lifshitz equations and equations for an elastic medium,

$$\dot{\theta}\sin\theta = -\frac{\gamma}{M_s}\left\{2A\frac{d}{dz}\left(\left(\frac{\partial\varphi}{\partial z}\right)\sin^2\theta\right) + \frac{\partial E_0}{\partial\varphi}\right\},$$

$$\dot{\varphi}\sin\theta = \frac{\gamma}{M_s}\left\{2A\left(\frac{d^2\theta}{dz^2} - \sin\theta\cos\theta\left(\frac{d\varphi}{dz}\right)^2\right) + \frac{\partial E_0}{\partial\theta}\right\}, \quad (2)$$

$$\rho\ddot{u}_i = \frac{\partial}{\partial x_k}\left(\frac{\partial E}{\partial u_{ik}}\right),$$

where E_0 is the energy density minus the exchange deposition and γ is the gyromagnetic ratio.

A solution to (2) will be sought in the form of stationary ME waves propagating along the $OZ \parallel [111]$ axis with constant velocity V (i.e., in the form $\theta = \theta(\xi)$, $\varphi = \varphi(\xi)$, and $u_i = u_i(\xi)$, where $\xi = z - Vt$) subject to the boundary conditions

$$\theta_0(\xi \longrightarrow \pm\infty) = \pi/2, \quad \varphi_0(\xi \longrightarrow \pm\infty) = \pm\pi/2,$$
$$\frac{\partial u_i}{\partial\xi}(\xi \longrightarrow \pm\infty) = 0. \quad (3)$$

Considering the asymptotic behavior of solutions to (2) in the vicinity of the equilibrium point ($\sim\exp(-k|\xi|)$), where k is a parameter) specified by conditions (3), we obtain a characteristic equation for k [5],

$$4Ak^4 - k^2\Delta_0^2[2A(H_2 + H_3) - V^2M_s^2/\gamma^2] + \Delta_0^4[H_2H_3 - H_1^2] = 0. \quad (4)$$

Here, $\Delta_0 = 6\sqrt{A/K_2}$ is the DW width in the absence of ME interaction [6],

$$H_1 = -\sqrt{2}(K_1 + K_2/6 + 3\sqrt{2}b_1b_2/2\zeta_1),$$
$$H_2 = -K_1 + \frac{K_2}{6} - \frac{2b_1^2 + b_2^2}{\zeta_1} - 2K_u^* - 2B_2(B_1 + B_2)/3\zeta_2, \quad (5)$$
$$H_3 = K_2/3 + 4b_2^2/\zeta_1,$$

where $K_u^* = K_u - 2\pi M_s^2$ is the IUA constant renormalized with regard to demagnetizing fields in the plate,

$$b_1 = (2B_1 + B_2)/3,\ b_2 = \sqrt{2}\,(B_1 - B_2)/3;$$

$$\zeta_1 = \frac{2}{3}(V^2/S_t^2 - 1)(C_{11} - C_{12} + C_{44});$$

$$\zeta_2 = \frac{1}{3}(V^2/S_l^2 - 1)(C_{11} + 2C_{12} + 4C_{44}),$$

and

$$S_l = \sqrt{C_{11} + 2C_{12} + 4C_{44}/3\rho},$$

$$S_t = \sqrt{C_{11} - C_{12} + C_{44}/3\rho}$$

are the longitudinal and transverse velocities of sound.

Equation (4) is biquadratic and has the following solutions:

$$k_{1,2}(V) = (\pm1)\frac{M_s\Delta_0}{8A\gamma}[\sqrt{z_+} \pm \sqrt{z_-}], \quad (6)$$

where

$$z_\pm = V_\pm^2 - V^2,$$
$$V_\pm^2 = \frac{2A\gamma^2}{M_s^2}[H_2 + H_3 \pm 2\sqrt{H_2H_3 - H_1^2}]. \quad (7)$$

Let us consider the roots of characteristic equation (4).

(1) Assume that ME interaction is weak. In this case, there exist two characteristic values of parameter V, V_+ and V_-. For $V > V_+$, k is purely imaginary. This domain corresponds to the spectrum of spin waves, and parameter V_+ is their minimal phase velocity. For $V < V_-$, k is real and there may exist a solution in the form of a solitary DW; here, V_- is the limiting velocity of its stationary motion [5].

For $K_u < 0$ (easy-plane magnet) and $|K_u| \gg K_{1,2}$, obtainable solutions follow from the system of equations [6]

$$\theta = \frac{\pi}{2} - \frac{1}{2|K_u^*|}\left[\frac{M_sV}{\gamma}\varphi' - \frac{\sqrt{2}}{3}\left(K_1 + \frac{1}{6}K_2\right)\cos3\varphi\right],$$
$$\varphi'' + \frac{1}{\Delta_0^2}\sin6\varphi = 0. \quad (8)$$

In the resonance case ($\theta = \pi/2$) with $K_2 > 0$, a solution to system (8) is a wave of magnetic moment reversal (kink or antikink) and describes the motion of a 60° DW between two domains, one of them with $\mathbf{M} \parallel [10\bar{1}]$ ($\varphi_0|_{\xi\to-\infty} = -\pi/6$) and the other with $\mathbf{M} \parallel [01\bar{1}]$ ($\varphi_0|_{\xi\to+\infty} = \pi/6$) [6]. If $K_2 < 0$, $\mathbf{M} \parallel [11\bar{2}]$ ($\varphi_0|_{\xi\to-\infty} = 0$) and $\mathbf{M} \parallel [\bar{1}2\bar{1}]$ ($\varphi_0|_{\xi\to+\infty} = \pi/3$).

It should be noted that, both for the general case, $\theta \neq \pi/2$, and for the case $\theta = \pi/2$, the type of solution and the DW width (Δ_0) depend on only the value of K_2 and are independent of K_1. This is because the first constant of CA does not contribute to the direction anisot-

ropy of **M** in the (111) plane: such an anisotropy arises only when the second constant of CA is taken into account.

In the absence of ME interaction, the DW limiting velocity (V_-) and the spin wave minimal phase velocity (V_+) in garnet ferrites, which are determined by the material parameters, may exceed the sound velocity in them. This is due to the fact that, in such materials, M_s and parameter K_u may be varied in a wide range (from 0 to $\geq 1.5 \times 10^5$ J/m and from -10^4 to 10^4 J/m^3, respectively [4]).

Estimates of V_- for $A \approx 6 \times 10^{-10}$ J/m^2, $\gamma \approx 3 \times 10^5$ s^{-1} (A m)$^{-1}$, $M_s \approx 8 \times 10^2$ A/m, and $|K_u| \approx 10^4$ J/m^3 [4, 7] (under the condition $|K_u| \gg K_{1,2}$, which is typical of a number of epitaxially grown garnet ferrite films) give $V_- \approx 1 \times 10^4$ m/s, whereas the longitudinal sound velocity in ferrites is $S_l \approx 7 \times 10^3$ m/s under the same temperature conditions.

Note that, for $Sm_3Fe_5O_{12}$ and $Tb_3Fe_5O_{12}$ garnet ferrites, K_2 may exceed K_u by one order of magnitude (for instance, in $Sm_3Fe_5O_{12}$, $K_2 = 21 \times 10^4$ J/m^3 at $T = 77$ K [7]). Thus, K_2 becomes a key parameter defining the value of V_-, which now comes equal to $V_- \approx 2 \times 10^4$ m/s (with the values of parameters A, K_u, M_s, and γ remaining the same). Near the point of magnetic moment compensation, V_- may be several times higher than the value mentioned above. Thus, under certain conditions, V_- may significantly exceed the sound velocity in the materials considered.

(2) Consider the roots of the characteristic equation with allowance for ME interaction. In the velocity interval $0 < V \ll V_l$, the terms that contain ζ_2^{-1} are much smaller than $K_{1,2}$ and K_u. Then, parameters H_1, H_2, and H_3 in the expression for $V_\pm$ take the form

$$H_1 = -\sqrt{2}(K_1 + K_2/6 + 3\sqrt{2}b_1b_2/2\zeta_1),$$

$$H_2 = -K_1 + K_2/6 - (2b_1^2 + b_2^2)/\zeta_1 - 2K_u^*, \qquad (9)$$

$$H_3 = K_2/3 + 4b_2^2/\zeta_1.$$

In the interval $S_t \ll V < S_l$, we have $\zeta_1^{-1} \ll 1$ and Eq. (5) becomes

$$H_1 = -\sqrt{2}(K_1 + K_2/6),$$

$$H_2 = -K_1 + K_2/6 - 2K_u^* - 2B_2(B_1 + B_2)/3\zeta_2, \quad (10)$$

$$H_3 = K_2/3.$$

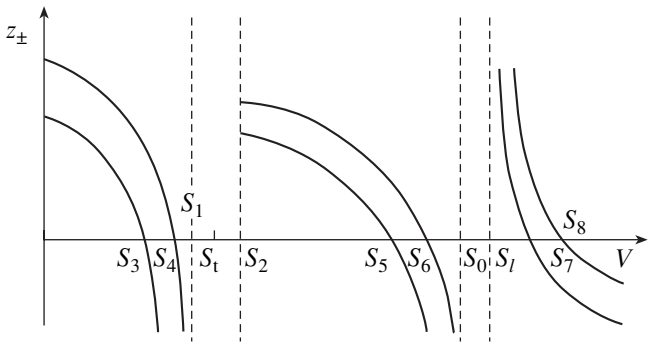


Fig. 1. Schematic dependence of $z_\pm$ on the stationary motion velocity.

The dependence of $z_\pm$ on the stationary motion velocity is schematically shown in Fig. 1, where

$$S_3 = S_1 - O_1(h), \quad S_4 = S_1 - O_2(h),$$

$$S_5 = S_2 + O_3(h), \quad S_6 = S_0 - O_4(h), \qquad (11)$$

$$S_7 = V_- + O_5(h), \quad S_8 = V_+ + O_6(h).$$

Here,

$$S_0 = S_l[1 - B_2(B_1 + B_2)/K_u^*(C_{11} + 2C_{22} + 4C_{44})]^{1/2},$$

$$S_1 = S_t[1 - 3(3b_2^2 - 3b_1^2)/K_u^*(C_{11} - C_{12} + C_{44})]^{1/2},$$

$$S_2 = S_t[1 + (18b_2^2/K_2(C_{11} - C_{12} + C_{44})) + (1 + K_2(2b_2^2 - 3b_1^2)/12b_2^2K_u^*)]^{1/2}$$

(at $|K_u^*| \gg K_{1,2}$).

The roots of Eq. (4) imply that the magnets considered may support ME waves of the same type as in weak ferromagnets (provided that $V_- > S_{t,l}$) [8]. These waves are the following.

(i) A moving DW (MDW) may be observed in the velocity intervals (see Fig. 1) $0 \leq V \leq S_3$, $S_2 < V < S_5$, and $S_1 < V < S_7$; (ii) an ME soliton may observed at $S_3 < V < S_4$, $S_5 < V < S_6$, $S_7 < V < S_8$, $0 \leq V \leq S_3$, $S_2 < V < S_5$, and $S_1 < V < S_7$ (see Fig. 1); (iii) a Cherenkov-type radiating solitary ME wave may be observed at $S_4 < V < S_1$, $S_6 < V < S_0$, and $S_1 < V < S_t \cup S_t < V < S_2$ (see Fig. 1); and (iv) a coupled ME wave may be observed at $V > S_8$ (see Fig. 1).

Solutions to system (2) with boundary conditions (3) can be found from the equations

$$\theta = \frac{\pi}{2} - \frac{1}{2|K_u^*|}\left[\frac{M_s V}{\gamma}\varphi' - \frac{\sqrt{2}}{3}\left(K_1^* + \frac{1}{6}K_2\right)\cos 3\varphi - \frac{2b_1b_2}{\zeta_1}\cos\varphi\right], \qquad (12)$$

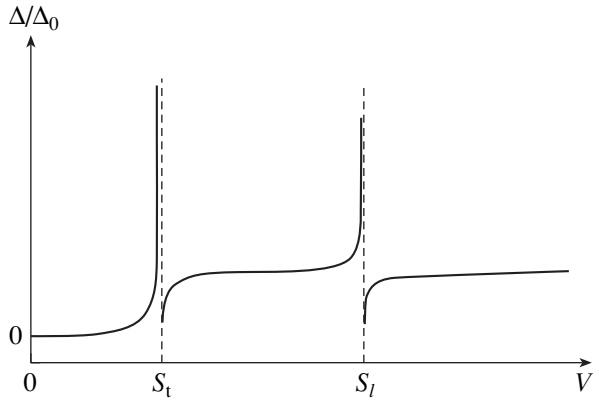


Fig. 2. Profiles of solitary waves that may arise in case (i) for $A \approx 6 \times 10^{-10}$ J/m, $M_s \approx 8 \times 10^2$ A/m, $|K_u| \approx 10^4$ J/m^3, $K_1 \approx -10^5$ J/m^3, and $K_2 \approx 2 \times 10^5$ J/m^3.

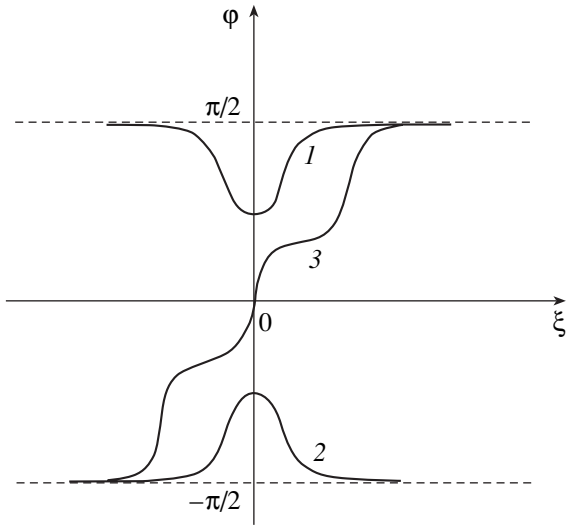


Fig. 3. Schematic dependence of the DW width on the DW stationary motion velocity.

$$\left(\frac{V^2}{\tilde{s}^2} - 1\right)\varphi'' = d_1 \sin 6\varphi - \frac{2}{3} d_2 \sin 4\varphi - \left(d_3 + \frac{d_2}{3}\right)\sin 2\varphi,$$

where $d_1 = 1/\Delta^2$; $d_2 = \sqrt{2}\,\beta b_1 b_2 / A\zeta_1 |K_u^*|$; $d_3 = (b_1 b_2/|K_u^*|\zeta_1^2 - b_2^2/\zeta_1)^2/A$; $b = K_1^* + K_2/6$; $\tilde{s} = 2\gamma\sqrt{A|K_u^*|}/M_s$; and $K_u^* = K_u - 2\pi M_s^2 - B_2(B_1 + B_2)/3\zeta_2$ and $K_1^* = K_1 + 3b_1 b_2/\sqrt{2}\,\zeta_1$ are the IUA and CA constants, respectively, renormalized with regard to ME interaction. The prime means differentiation with respect to new variable ξ, and Δ appearing in the expression for coefficient d_1,

$$\Delta = \Delta_0[1 - \beta(6K_1^*/K_2 + 1)/|K_u^*|]^{-1/2}, \qquad (13)$$

specifies the DW effective width.

One can see from (13) that the dependence of Δ on the stationary motion velocity has singularities in the regions where the DW moves with velocities close to S_3, S_5, and S_7. In these regions, the DW structure changes drastically up to collapse when moving at a velocity close to velocities $S_{t,l}$ (see Fig. 2).

It follows from Eq. (12) that ME interaction essentially complicates the dynamics in the magnet. With allowance for ME, the dynamics is described with the triple sin-Gordon equation (unlike case (1)). This is because of an additional magnetostriction contribution to energy (1) arising from a potential mismatch between neighboring domains. This, in turn, imposes restrictions (in the form of boundary conditions (3)) on the range of waves that might exist according to the wide spectrum of solutions to (12).

Considering nonlinear ME waves that may be observed in the velocity spectrum mentioned above, one can see that, in the ranges corresponding to case (i) (Fig. 3, curve *3*), there may exist, along with MDWs, solutions corresponding to ME solitons (Fig. 3; curves *1*, *2*). It should be noted that, in contrast to paper [8], the topology of the MDW-related solution is a magnetic inhomogeneity corresponding to a 180° DW with inflexion points near the metastable axes (which are easy axes in the absence of ME interaction [6]). Such waists in the DW structure appear when ME interaction is taken into account: it generates three coupled 60° DWs with waists at the sites where solutions are sewed together.

When ME waves move with velocities close to S_3, S_5, or S_7 (case (ii)), there arises a solution corresponding to a kink-type soliton. In the ranges close to the sound velocity, Eq. (12) has (in the dissipation-free approximation) bifurcation points where solutions satisfying condition (3) disappear.

Investigation of the solutions to the characteristic equation showed that the velocity spectrum for solitary ME waves exhibits, as in [8], ME gaps near the sound velocities that are defined as follows:

$$h_{1t} = 3(2b_1^2 + b_2^2)/K_u^*(C_{11} - C_{12} + C_{44}),$$

$$h_{2t} = (18b_2^2/K_2(C_{11} - C_{12} + C_{44})) \times (1 + K_2(2b_2^2 - 3b_1^2)/12b_2^2 K_u^*)$$

near S_t and

$$h_l = B_2(B_1 + B_2)/K_u^*(C_{11} + 2C_{22} + 4C_{44})$$

near S_l.

With the values of the ME gaps falling into the domains indicated, coupled ME waves self-sustained at infinity occur in the absence of decay [8].

The above relations also imply that parameters K_2 and K_u contribute considerably to the value of ME gaps. If IUA is induced by stress $\boldsymbol{\sigma} \parallel [111]$ applied along the

crystallographic direction [111], there appears the possibility of controlling the parameters of ME gaps.

Similarly to [8], it can be assumed that no ME gaps will be observed in the velocity spectrum of ME waves in ferrites under certain conditions if dissipative processes in magnetic or elastic subsystems are taken into account. However, the stationary motion velocity-versus-external magnetic field curve may exhibit anomalies.

Thus, our investigation demonstrates that allowance for ME interaction makes the description of the dynamics in magnets more involved. In this case, analysis of the realizable types of ME waves reduces to studying solutions to the triple sin-Gordon equation with boundary conditions that narrow the spectrum of solutions.

Unlike the topology of waves that may arise in easy-plane magnets [8] (for instance, a 180° DW), the topology of a 180° DW in materials with a more complicated symmetry may exhibit anomalies (waists). Along with a solution like a moving DW, soliton-like solutions may appear for which equilibrium magnetization vector **M** is $\mathbf{M} \parallel [1\bar{1}0]$ ($\varphi_0|_{\zeta \to \infty} = \pi/2$) and $\mathbf{M} \parallel [\bar{1}10]$ ($\varphi_0|_{\zeta \to -\infty} = -\pi/2$).

From the analysis of the asymptotic behavior of the solutions to the system of equations describing the ME dynamics in the magnets, it follows that the limiting velocity of DW stationary motion may be comparable to, or even exceed (for example, at the point of magnetic moment compensation), the well-known values [4] for such materials.

REFERENCES

1. E. A. Turov and V. G. Shavrov, Usp. Fiz. Nauk **140**, 429 (1983) [Sov. Phys. Usp. **26**, 593 (1983)].
2. O. Yu. Belyaeva, L. K. Zarembo, and S. N. Karpachev, Usp. Fiz. Nauk **162**, 107 (1992) [Sov. Phys. Usp. **35**, 106 (1992)].
3. V. G. Bar'yakhtar, B. A. Ivanov, and M. V. Chetkin, Usp. Fiz. Nauk **146**, 417 (1985) [Sov. Phys. Usp. **28**, 563 (1985)].
4. V. V. Randoshkin and A. Ya. Chervonenkis, *Applied Magnetooptics* (Énergoizdat, Moscow, 1990) [in Russian].
5. V. G. Bar'yakhtar, B. A. Ivanov, and A. L. Sukstanskiĭ, Fiz. Tverd. Tela (Leningrad) **20**, 2177 (1978) [Sov. Phys. Solid State **20**, 1257 (1978)].
6. R. M. Vakhitov, Izv. Vyssh. Uchebn. Zaved. Fiz., No. 10, 35 (2002).
7. S. Chikazumi, *Physics of Ferromagnetism*, 2nd ed. (Clarendon, New York, 1997; Mir, Moscow, 1987) [translated from Japanese].
8. V. D. Buchel'nikov and V. G. Shavrov, Fiz. Met. Metalloved. **58**, 845 (1984).

Translated by A. Pankrat'ev

Technical Physics, Vol. 50, No. 8, 2005, pp. 1026–1033. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 63–70.

SOLIDS

Grain-Boundary Diffusion and Solubility of Helium in Submicrocrystalline Palladium

A. N. Zhiganov and A. Ya. Kupryazhkin

Ural State Technical University (UPI), ul. Mira 19, Yekaterinburg, 620062 Russia

e-mail: kupr@dpt.ustu.ru

Received August 13, 2004; in final form, December 15, 2004

Abstract—The diffusion and solubility of helium in palladium with a submicrocrystalline structure are investigated by thermal desorption of helium from He-saturated specimens at temperatures T = 293–508 K and saturation pressures P = 0.1–35 MPa. As the saturation pressure rises, the effective diffusion coefficient increases, exhibits a plateau, and then decreases to its initial value. Along with the four plateaus discovered earlier, the solubility versus saturation pressure dependence in the range 25.5–35.0 MPa demonstrates a fifth plateau, where the solubility is as high as $(3.0 \pm 0.4) \times 10^{17}$ cm^{-3}. It is shown that the helium diffuses along grain boundaries, at which clusters (traps) consisting of eight to ten vacancies are localized, and dissolves in these clusters. The high value of C_{eff} in the fifth plateau is explained by pairwise merging of adjacent vacancy clusters. From the $D_{\text{eff}}(P)$ dependences, the vacancy clusters concentration is estimated as $C^* = 2.32 \times 10^{16}$ cm^{-3}. Within the experimental error, this value coincides with that obtained from the solubility data. Calculations of the energy of helium–defect interaction in submicrocrystalline Pd that are made using the molecular dynamics method support the experimental results. © *2005 Pleiades Publishing, Inc.*

INTRODUCTION

Interaction of helium with defects in metals is a decisive factor causing helium embrittlement. However, little is known about how defects influence the solubility and diffusion of helium in metals, since this problem poses significant experimental and theoretical difficulties. Grain-boundary diffusion and solubility are the hardest to analyze in this case.

It was shown [1, 2] that, in submicrocrystalline palladium, boundaries and vacancy clusters assist helium dissolution and diffusion. In those works, the parameters of helium transfer and the concentration of defects, such as vacancy clusters, in submicrocrystalline palladium were determined at low temperatures and saturation pressures.

Of interest is computer simulation of helium transfer mechanisms and the structure of defects in SMC palladium specimens in a range of saturation pressures wider than that considered in [1, 2].

EXPERIMENTAL AND SIMULATION TECHNIQUES

In this work, the diffusion and solubility of helium was studied on the same palladium specimens as in [1–3]. The specimens had the form of a disk with surface area $S = (1.25 \pm 0.1)$ cm^2, mass $m = (90.9 \pm 0.1)$ mg, and thickness $h = (6.1 \pm 0.3) \times 10^{-3}$ cm. The examination of the specimen by digital optical microscopy showed that it consists of grains with an average size of (3 ± 2) μm. The size of constituent subgrains, which arise from severe plastic deformation during specimen preparation, equaled 150 nm, as determined earlier by transmission electron microscopy [4].

The grain average size remains unchanged upon annealing at temperatures of up to 1300 K, whereas the subgrains grow to 1 μm or more even upon annealing at 553 K [4].

The solubility of helium in palladium was studied by thermal desorption of the gas with the setup described in [1]. The specimen was saturated by helium in a helium environment (saturation chamber) at a given temperature and pressure and then was placed into the measuring chamber. The ranges of saturation temperatures and pressures were, respectively, T = 293–508 K and P = 0.1–35.0 MPa. The solubility measurements were taken until the specimen was fully degassed. The error in solubility determination was no more than 10%. The effective diffusion coefficients were obtained from the time variation of the helium desorption flux from the specimen to a vacuum at the temperature equal to the saturation temperature. The effective diffusion coefficients were determined accurate to 15%.

The energies of helium interaction with defects in the submicrocrystalline palladium were calculated by the molecular dynamics method. To determine the energy of helium dissolution in a defect, we used the following procedure [3]. Prior to simulation, we set the initial configuration of the system (a defect of appropriate structure with dissolved helium atoms inside). The particles of the system were assigned randomly directed momenta so that the temperature of the system

was equal to a given initial temperature of simulation (100 K). Then, simulation with a necessary number of steps was performed under the condition of continuous cooling. During cooling, lattice relaxation near the defect takes place and helium atoms in the defect are rearranged so that the dissolution energy is minimal. The simulation time was taken such that the temperature of the system decreased to a value at which the thermal contribution to the total energy of the system becomes negligible. In our case, a total of 5000 time steps were used and the temperature of the system at the end of simulation was lowered to $\sim 3 \times 10^{-9}$ K. The total energy of the system was determined by averaging over the last 500 steps.

The total energy of the system with a defect free of helium, E_0, was calculated by averaging over the full simulation time. The total energy of the system with a defect occupied by a desired number of helium atoms, E, was calculated in a similar way. The dissolution energy is equal to the difference between the total potential energies of the system with the defect occupied and unoccupied by helium at absolute zero,

$$E^{\mathrm{d}} = E - E_0. \tag{1}$$

In the simulation, we used periodic boundary conditions for a cubic crystal with eight fcc palladium lattice spacings on an edge (lattice spacing $a_0 = 0.388$ nm, the number of particles in an ideal crystallite is 2048). Integration of the equations of motion for the system consisting of N particles was performed by the Euler method with a half-step,

$$\begin{cases} r_j^i(t + \Delta t) = r_j^i(t) + \Delta t\, v_j^i(t + \Delta t/2) \\ v_j^t(t + \Delta t/2) = v_j^i(t - \Delta t/2) + \Delta t m_j^{-1} F_j^i(t) \\ F_j^i(t) = -\sum\limits_{k=1,\, k\neq j}^{N} \partial\varphi(r_j(t) - r_k(t))/\partial r_j^i. \end{cases} \tag{2}$$

Here, $\Delta t = 5 \times 10^{-15}$ s is the time step of integration; $r_j^i(t)$ are the coordinates of the particles; $v_j^i(t) = r_j^i(t)$ are the momenta of the particles; m_j are the masses of the particles; superscripts $i = 1, 2, 3$ refer to the vector components; subscripts $j = 1, \ldots, N$ are the numbers of the particles; and $\varphi(r_{jk})$ is the interaction potential for a pair of particles separated by distance $r_{jk} = |r_j - r_k|$.

The Lenard–Jones pair central interaction potential of type (n–m) for the Pd–Pd system was taken from [5, 6] with parameters $\varepsilon = 0.41$ eV, $n = 10$, $m = 5.5$, and $r_m = 0.2824$ nm. For the He–He pair, we used the (6–12) Lenard-Jones potential with parameters $\varepsilon = 8.815 \times 10^{-4}$ eV and $\sigma = 0.2556$ nm [7].

The potential for the Pd–He pair was taken from [8]. The source potential [8] had only the repulsive component. Therefore, to simulate the energy of helium atom adsorption on the surface of a vacancy cluster, we modified this potential by adding dispersion interaction in the form of the Buckingham HFD1 potential [9], which takes into account fading of dispersion interaction at small distances,

$$\varphi_{\mathrm{Pd-He}}(r) = A\exp(-\beta r) - \frac{C_6}{r^6}\exp\left(-\left(1.28\frac{r_m}{r} - 1\right)^2\right), \tag{3}$$

where $A = 214\,898$ eV, $\beta = 4.154 \times 10^{10}$ m^{-1}, $C_6 = 8.73 \times 10^{-6}$ eV nm^6, and $r_m = 0.5$ nm.

The validity of the chosen attractive component of the potential was checked by comparing the binding energy of a helium atom with the metal surface ((111) plane) with the results of density functional calculations in terms of the jelly model. Since relevant data for the He–Pd system are lacking, we used the data for the He–Au system. The binding energy of a helium atom with the gold surface is $E^{\mathrm{b}} = -8$ meV [10]. The selection of the potential for the Pd–He system is justified in [3], where four potentials for the Pd–He system obtained by different methods are compared.

RESULTS AND DISCUSSION

Earlier, it was shown [1] that the flux of the helium desorbed from the submicrocrystalline palladium specimen can be described by the effective diffusion coefficient; therefore, analysis was performed in the approximation of effective transport coefficients.

Figure 1 demonstrates the solubility isotherms $C_{\mathrm{eff}}(P)$ (together with the data taken from [2]) for temperatures of 403, 433, and 508 K (the other isotherms are omitted to make the figure more illustrative). All the isotherms have clear-cut steps with extended plateaus. The effective solubility in the plateaus with the same number coincide within the error of C_{eff} measurement

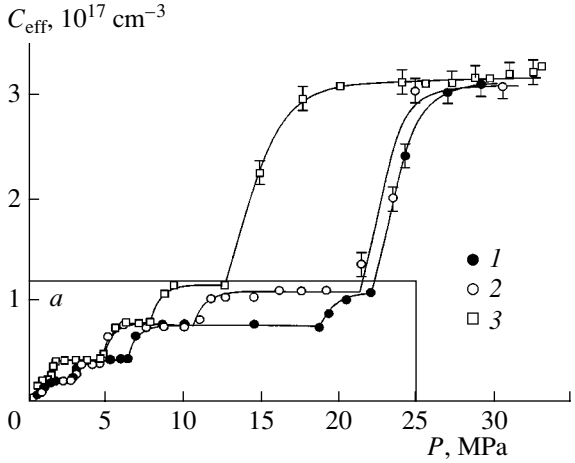


Fig. 1. Helium solubility in the submicrocrystalline palladium vs. saturation pressure. T = (*1*) 403, (*2*) 433, and (*3*) 508 K. Solid lines, approximation. (*a*) Solubility data for the first four steps [2].

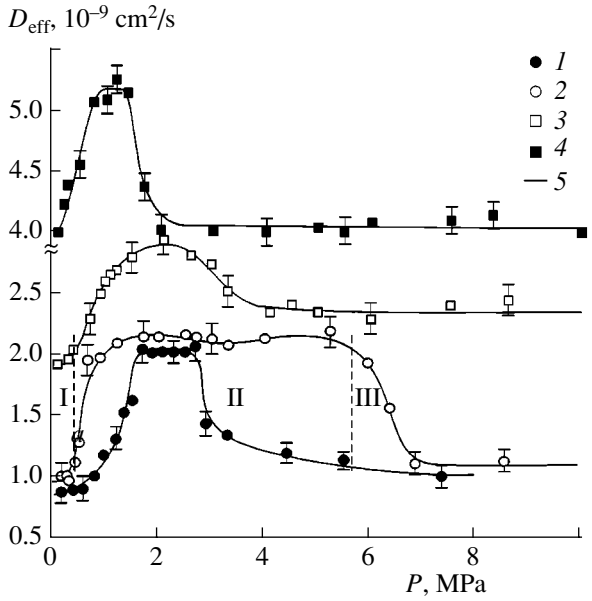


Fig. 2. Effective diffusion coefficient of helium in the submicrocrystalline palladium vs. saturation pressure. T = (*1*) 387, (*2*) 403, (*3*) 433, and (*4*) 483 K; curve *5* is the fitting curve. I, II, III are portions of curve *2* (T = 403 K).

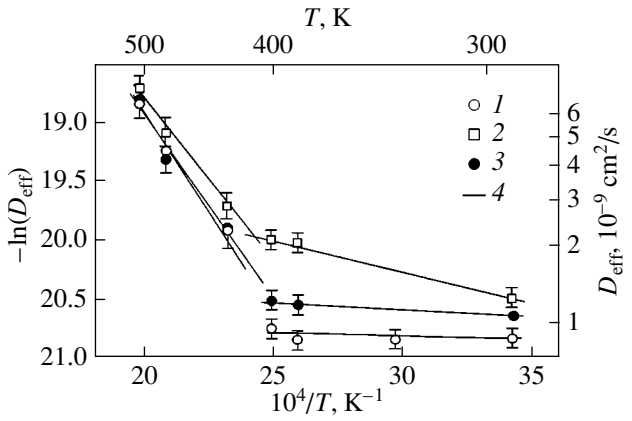


Fig. 3. Temperature dependence of the effective diffusion coefficient. (*1*) Portion I (low saturation pressures), (*2*) portion II (plateau of the curve $D_{eff}(P)$), (*3*) portion III (the postplateau region, high saturation pressures), and (*4*) approximation.

in all the isotherms. From the first to fourth plateau, the solubility varies in the ratio 1 : 2 : 3 : 4, as shown in [2]. The solubility ratio between the fourth and fifth plateaus is 1 : 3. Such an anomalous increase in the solubility at high saturation pressures after the fourth plateau was observed for all the isotherms.

Figure 2 plots the pressure dependences of effective diffusion coefficient $D_{eff}(P)$ at temperatures of 387, 403, 433, and 483 K (for convenience, the pressure range here is narrower than in Fig. 1). The curves $D_{eff}(P)$ can be subdivided into three portions (see, for example, curve *2* taken at T = 403 K, where these portions are the most distinct). In portion I, corresponding to low saturation pressures (P = 0–0.03 MPa), the effective diffusion coefficient is pressure-independent within the measurement error (see, e.g., curve *3*). Portion II appears with increasing saturation pressure: here, D_{eff} first grows with saturation pressure and then remains constant (plateau). In portion III, $D_{eff}(P)$ starts decreasing as the pressure rises and decreases to a value close to D_{eff} in portion I. Subsequently, D_{eff} remains unchanged up to the end of the pressure range (P = 10–35 MPa).

Comparing the pressure dependences of the effective diffusion coefficient and effective solubility shows that the first portion in the isotherms $D_{eff}(P)$ corresponds to the initial ascending (linear) portion of the first stage in the corresponding curves $C_{eff}(P)$ at all temperatures (at temperatures exceeding 400 K, this portion is hardly discernible). The beginning of the first plateau in the curve $C_{eff}(P)$ and the plateau itself correspond to portion II in the curve $D_{eff}(P)$, where D_{eff} grows with saturation pressure. At low temperatures (387 and 403 K), the end of the plateau in the curve $D_{eff}(P)$ correlates with the initial pressure of the third stage. At high temperatures (433, 483, and 508 K), this plateau is terminated at the initial pressure of the second step. The beginning of portion III in the curve $D_{eff}(P)$ (the decrease in D_{eff} with increasing P) corresponds to the linear portion of the third (at 387 and 403 K) or second (at 433, 483, and 503 K) stages in the effective solubility isotherms.

Figure 3 presents the temperature dependences of D_{eff} for portions I–III in the $D_{eff}(P)$ dependence (Fig. 1). The value of D_{eff} in the respective portion was obtained by averaging constant values of D_{eff}, $D_{eff}(P)$ = const. All the $D_{eff}(T)$ dependences consist of two portions (Fig. 3), either obeying the Arrhenius law $D_{eff} = D_{0eff}\exp(-E^D_{eff}/kT)$. Activation energies E^D_{eff} and preexponentials D_{0eff} obtained from the curves $D_{eff}(T)$ are given in Table 1.

To elucidate the mechanism of diffusion, it is necessary to correlate the experimental energies of helium dissolution in the specimen, $E_i^{d(exp)}$, Table 3) with the dissolution energy of a helium atom occupying various positions in the polycrystalline Pd. The dissolution energy estimated by the molecular dynamics method with potential (3) for the He–Pd pair was found to be E_i^d = 3.17 eV for helium in an interstitial, E_v^d = 0.25 eV for helium in a vacancy, and E_d^d = 0.87 eV for helium at a dislocation. Accurate to potential setting, these values coincide with the data from [8] (E_i^d = 3.68 eV, E_v^d = 0.52 eV). This allows us to exclude interstitials, vacancies, dislocations, and subgrain boundaries as

centers of dissolution. Such a conclusion is also substantiated by the fact that the growth of subgrains at annealing [4] does not influence the solubility value corresponding to the plateaus. Thus, one can argue that diffusion along large-angle grain boundaries is the dominating mechanism of helium transfer in the Pd polycrystal. The low values of $D_{0\text{eff}}$ and activation energy of diffusion, $E^{\text{D}}_{\text{eff}}$, are consistent with this inference about the diffusion mechanism.

The run of the effective diffusion coefficient with saturation pressure (Fig. 2, portions I and II) indicates that diffusion follows the trapping mechanism [2], helium atoms being trapped by vacancy clusters localized at grain boundaries and at joints between grain boundaries. It is seen in Fig. 3 that the mechanism of helium diffusion in the specimen changes at a temperature of ≅400 K. It can be assumed that the helium is transferred along grain boundaries via diffusion of vacancies (true grain-boundary diffusion) and divacancies.

Thus, according to the experimental data (Fig. 3), low-temperature ($T \le 403$ K) helium transfer in the specimen is of a low activation energy (Table 1) and may follow two mechanisms simultaneously (diffusion along grain boundaries and grain-boundary divacancies) with vacancy clusters serving as traps. At temperatures above 403 K, grain-boundary divacancies may dissociate. Accordingly, diffusion by divacancies ceases, the effective activation energy of diffusion increases, and the trap parameters change (the clusters become smaller at the same temperatures [2]).

To analyze the high-temperature (433–508 K) behavior of the curve $D_{\text{eff}}(P, T)$, it is assumed that the traps (vacancy clusters) are randomly filled with helium atoms. Then, in the local equilibrium approximation, we can write the expression for D_{eff}, using the approach adopted in [11–13], in the form

$$D_{\text{eff}} \cong \frac{D_{\text{gb}}}{1 + (C_{\text{c}}^* - C_{\text{eff}})^2 \dfrac{k_{\text{c}}}{C_{\text{c}}^*} \exp(B_{\text{c}}/kT)}, \qquad (4)$$

where $D_{\text{gb}} = D_{0\text{gb}}\exp(-E^{\text{D}}_{\text{gb}}/kT)$ is the coefficient of grain-boundary diffusion, C_{c}^* is the density of traps (vacancy clusters), C_{eff} is the solubility of helium in the specimen, B_{c} is the binding energy of a diffusing atom in a vacancy cluster relative to grain-boundary positions of dissolution, and k_{c} is a factor slightly varying with temperature.

In view of the fact that dissolution proceeds primarily in vacancy clusters, which serve as traps, the mechanism of the high-temperature behavior of $D_{\text{eff}}(P)$, according to (4), seems to be the following. At a low fill of traps, $C_{\text{eff}} \ll C_{\text{c}}^*$ (portion I of the curves $D_{\text{eff}}(P)$ in Fig. 2, the linear portion of the first stage of the curves $C_{\text{eff}}(P)$ in Fig. 1a [2, 3]), the diffusion coefficient depends on the trap parameters (see (4)) and does not depend on the fill of traps (saturation pressure). In this case, expression (4) coincides with that obtained in [12, 13] for the case of a low fill. As the traps are filled, $C_{\text{eff}} \approx C_{\text{c}}^*$, the curves $C_{\text{eff}}(P)$ (Fig. 1) and $D_{\text{eff}}(P)$ (Fig. 2) exhibit a plateau. The plateaus in the curves $D_{\text{eff}}(P)$ are related to smoothing out of the potential relief as vacancy clusters are filled with helium (Fig. 4b).

Table 1. Parameters appearing in the dependence of the effective diffusion coefficient of helium in palladium on desorption temperature

Portion of curve $D_{\text{eff}}(P)$	T, K	D_0, cm^2/s	E^{D}, eV
I	293–403	$(0.98^{+1.1}_{-0.9}) \times 10^{-9}$	0.0036 ± 0.0015
	403–508	$(8.3^{+6.8}_{-3.7}) \times 10^{-6}$	0.31 ± 0.02
II	293–403	$(8.2^{+0.6}_{-0.6}) \times 10^{-9}$	0.047 ± 0.002
	403–508	$(2.3^{+0.7}_{-0.4}) \times 10^{-6}$	0.25 ± 0.02
III	293–403	$(1.69^{+0.14}_{-0.13}) \times 10^{-9}$	0.011 ± 0.002
	403–508	$(8.3^{+6.8}_{-3.7}) \times 10^{-6}$	0.28 ± 0.02

On a further increase in the saturation pressure, one more helium atom may dissolve in the trap (the linear portion of the second stage of the curves $C_{\text{eff}}(P)$ in Fig. 1a [2, 3]). Accordingly, D_{eff} declines (Fig. 2, portion III), because a larger number of helium atoms in a cluster may create a potential barrier in the path of diffusion. This barrier may be either overcome or by-

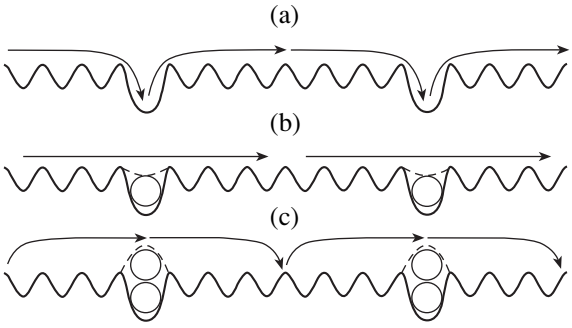


Fig. 4. Potential relief for a helium atom diffusing in polycrystalline palladium. (a) Low saturation pressures (a low fill of traps), (b) complete fill of traps, and (c) fill of traps at high saturation pressures (emergence of traps of a new type).

passed along alternative diffusion paths (Fig. 4c). As the traps are filled with helium atoms, the effective helium solubility eventually exceeds the concentration of vacancy clusters, $C_{\text{eff}} \geq C_c^*$. The current concentration of traps of a new type (vacancy clusters that have transformed into barriers due to the dissolution of an extra helium atom) can be denoted as $C_{c2} = C_{\text{eff}} - C_c^*$. According to (4), at $2C_c^* \geq C_{\text{eff}} \geq C_c^*$ (the second stage of dissolution), the trapping mechanism switches to traps of the new type and we again are dealing with only one type of traps.

One can assume that the parameters of the new traps differ from those of the former; namely, B_c should be replaced by B_{c2} (the barrier height for a diffusing atom relative to grain-boundary positions of dissolution) and k_c should be replaced by k_{c2} (a coefficient weakly varying with temperature). Then, since vacancy clusters are filled independently (in terms of the dissolution model), the high-temperature behavior of $D_{\text{eff}}(P, T)$ can be expressed, using (4), in the form

$$D_{\text{eff}} = \frac{D_{0\text{gb}} \exp(-E_{\text{gb}}^{\text{D}}/kT)}{1 + (C_c^* - C_{\text{eff}})^2 \dfrac{k_c}{C_c^*} \exp(B_c/kT)\theta(P_0 - P) + C_{c2}^2 \dfrac{k_{c2}}{C_c^*} \exp(B_{c2}/kT)\theta(P - P_0)}. \quad (5)$$

Here, P is the saturation pressure, P_0 is the boundary between portions II and III, $C_{c2} = C_{\text{eff}} - C_c^*$ is the concentration of new traps, and

$$\theta(x) = \begin{cases} 0, & x \leq 0 \\ 1, & x > 0. \end{cases}$$

Note that the low-temperature decline in D_{eff} (the beginning of portion III in the curve $D_{\text{eff}}(P)$, Fig. 2) corresponds to the beginning of the third stage, whereas the high-temperature decline corresponds to the beginning of the second stage of the curve $C_{\text{eff}}(P)$ (Fig. 1a) [2, 3]. This can be associated with the fact that, at temperatures above 403 K, vacancies are detached from vacancy clusters, where dissolution occurs, and the energy of helium dissolution in these clusters rises [2].

Thus, at low temperatures, the volume of clusters is large and can readily accommodate two helium atoms without producing a potential barrier in the path of a helium atom diffusing along grain boundaries. At high temperatures, vacancies being detached diminish the cluster's volume and only one helium atom can be placed in a cluster. In this case, adding the second helium atom creates a potential barrier, which is consistent with the model of helium dissolution in palladium put forward in [2]. Portion III in the curve $D_{\text{eff}}(P)$ taken at temperatures above 403 K corresponds to the dissolution of the second atom in a vacancy cluster (second stage), whereas at temperatures below 403 K this portion corresponds to the dissolution of the third atom in a vacancy cluster (the third stage of the curve $C_{\text{eff}}(P)$, Fig. 1a) [2, 3].

The parameters of grain-boundary diffusion, $D_{0\text{gb}}$ and E_{gb}^{D}, and those of traps, B_c, B_{c2}, k_c, and k_{c2}, were estimated using the experimental data for the diffusion and solubility of helium at high temperatures (T = 403, 483, and 508 K). Such estimates can be made in the limiting cases (negligibly low and complete filling of vacancy clusters with helium) of expression (4), which describes the high-temperature behavior of $D_{\text{eff}}(T)$ (Fig. 3). The estimates are given in Table 2. In the calculations, the concentration of vacancy clusters was taken to be $C_c^* = (2.3 \pm 0.3) \times 10^{16}$ cm^{-3}, as follows from the solubility data obtained in [2].

Since the parameters mentioned above were determined from the curve $D_{\text{eff}}(T)$ in the limiting cases (complete or negligibly low filling), it is of interest to evaluate them using a more general dependence of D_{eff} on temperature and saturation pressure. Expression (5) describes only the high-temperature behavior of $D_{\text{eff}}(P, T)$ (T = 403, 483, and 508 K) in the range of saturation pressures from zero to the value corresponding to the beginning of the third stage (40 data points) in the curves $C_{\text{eff}}(P)$. In (5), the concentration of vacancy clusters, C_c^*, remains an unknown parameter. The diffusion parameters and parameters of defects obtained by minimizing the discrepancy between calculated values and data points are presented in Table 2. The dependence of D_{eff} with the parameters obtained by minimization of expression (5) gives a good fit to the experimental data, reproducing the observed nonmonotonic behavior of $D_{\text{eff}}(P, T)$.

Grain-boundary diffusion parameters $D_{0\text{gb}}$ and E_{gb}^{D} found by different methods coincide within the processing error, indicating that this model adequately describes helium diffusion in the specimen. Binding energies B_c in the traps determined by various processing techniques are also in nice agreement. The difference in the values of the parameters of new traps (k_{c2} and B_{c2}) can be explained by sparse data points for the $D_{\text{eff}}(T)$ curves. The value of C_c^* obtained independently from diffusion experiments (see (5)) is in good agreement with the concentration of vacancy clusters, C^* =

$(2.3 \pm 0.3) \times 10^{16}$ cm^{-3}, that was found from measurements of solubility C_{eff} [2].

Analysis of the low-temperature (293–403 K) behavior of $D_{eff}(T)$ needs additional experimental and theoretical data. A greater insight into switching of the diffusion mechanisms also calls for further investigation, specifically, simulation of grain-boundary diffusion of helium in palladium.

SOLUBILITY OF HELIUM AT HIGH SATURATION PRESSURES

According to the earlier model [1, 2], helium in polycrystalline palladium dissolves in vacancy clusters localized at grain boundaries. The presence of the stages (steps), as well as the fact that the heights of plateaus of the first four steps vary in the ratio 1 : 2 : 3 : 4 (within the experimental error), is explained by successive and independent filling of vacancy clusters with helium atoms in amounts from one to four. In terms of this model, other positions of dissolution were assumed to contribute negligibly (within the error of solubility determination) to the effective solubility.

The above hypothesis for the dissolution mechanism was supported by simulating the dissolution of 1–4 helium atoms in clusters consisting of 8–12 vacancies (the simulation was carried out by the molecular dynamics method) [3]. The calculated energies of dissolution of i helium atoms, E_i^d (i = 1–4), in a cluster of nine to ten vacancies are in good agreement with experimental values $E_i^{d(exp)}$, where the energies of dissolution for the first and the second stages correspond to the low-temperature part of the curve $C_{eff}(T)$ (Table 3).

As the saturation pressure increases, an anomalous rise in the solubility is observed (the fifth stage). Throughout the temperature range studied (T = 293–508 K), the temperature dependence of the solubility for the fifth stage obtained from its linear portion (see [2]) is described by the formula

$$C_{eff}(T) = C_{05}\exp(-E_5^d/kT). \quad (6)$$

Preexponential C_{05} is equal to $(5.37^{+2.1}_{-2.2}) \times 10^{15}$ cm^{-3}. According to the dissolution model proposed, the energy of dissolution of helium atoms at the fifth stage of the curve $C_{eff}(P)$ corresponds to dissolution of eight more helium atoms in a vacancy cluster where four helium atoms have already dissolved. This energy is equal to $E_5^{d(exp)} = (1.3 \pm 1.6) \times 10^{-3}$ eV. Energy $E_5^{d(exp)}$ is seen to be considerably lower than the energy of dissolution of the fourth helium atom in the vacancy cluster, $E_4^{d(exp)} = (0.070 \pm 0.007) \times 10^{-3}$ eV (Table 3). This leads us to conclude that, at high saturation pressures, the mechanism of helium dissolution in a vacancy cluster changes, resulting in an anomalous increase in the solubility at nearly zero dissolution energy.

Table 2. Comparison of the parameters appearing in the effective diffusion coefficient of helium in palladium in the temperature interval 433–508 K

Parameter	Processing of experimental curves $D_{eff}(T)$ (Fig. 3), formula (4)		Fitting of the parameters in formula (5) for $D_{eff}(T, P)$
C_c^*, cm^{-3}	$(2.3 \pm 0.3) \times 10^{16}$*		2.32×10^{16}
D_{0gb}, cm^2/s	$C_{eff} \approx C_c^*$	$(2.3^{+0.7}_{-0.4}) \times 10^{-6}$	2.24×10^{-6}
E_{gb}^D, eV		0.25 ± 0.02	0.251
k_c, cm^3	$C_{eff} \ll C_c^*$	1.60×10^{-19}	0.259×10^{-19}
B_c, eV		0.168	0.251
k_{c2}, cm^3	$C_{eff} \approx 2C_c^*$	3.55×10^{-20}	1.79×10^{-17}
B_{c2}, eV		0.228	0.0180

* From the solubility data obtained in [2].

The anomalous increase in the solubility upon transition from the fourth to fifth plateau can be explained by association of adjacent vacancy clusters at grain boundary as the specimen is saturated by helium at high pressures (≈30 MPa). Such an association reduces the dissolution energy, which may raise the solubility by a factor of 2–3 under the same saturation conditions.

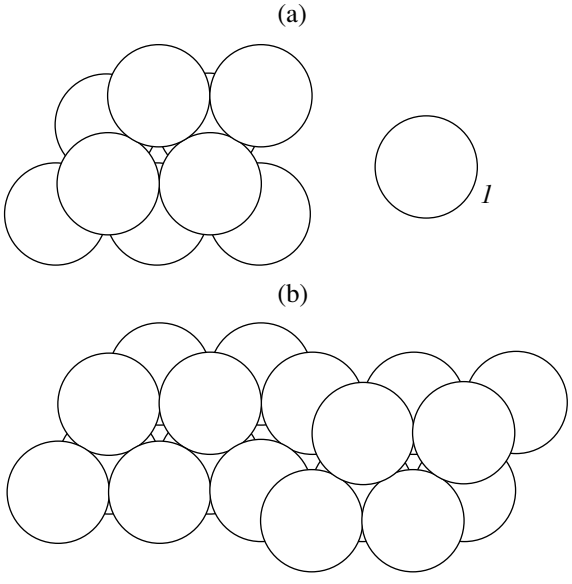


Fig. 5. Vacancy clusters in palladium ((111) plane): (a) cluster consisting of nine vacancies; (b) cluster consisting of 18 vacancies (an association of two clusters either consisting of nine vacancies. (*1*) Vacancy.

Table 3. Calculated and experimental values of the energy of helium atom dissolution in a vacancy cluster of nine vacancies (Fig. 5a)

Number of helium atoms in cluster	ΔT, K	Cluster size (number of vacancies)	E_i^{d}, eV	$E_i^{d(\exp)}$, eV
1	273–400	9	−0.02593	-0.028 ± 0.005
	400–508	7*	+0.15220	$+0.086 \pm 0.009$
2	293–430	9	−0.01300	-0.010 ± 0.003
	430–508	7*	+0.15540	$+0.077 \pm 0.009$
3	293–508	9	+0.01448	$+0.050 \pm 0.006$
4	293–508	9	+0.05208	$+0.070 \pm 0.07$
12**	293–508	9**	−0.01625	$(1.3 \pm 1.6) \times 10^{-3}$

* Dissolution of one or two helium atoms in a cluster from which a divacancy was removed.
** Dissolution of 24 helium atoms in a cluster consisting of 18 vacancies (Fig. 5b).

Upon helium desorption from the specimen, the association forces disappear, the associates (clusters) decompose, and the capacity of dissolution sites is recovered, providing the reproducibility of lower pressure results.

To check the proposed mechanism of anomalous increase of solubility at high pressures, we performed molecular dynamic simulation of helium dissolution in nearby but unassociated clusters consisting of nine vacancies (Fig. 5a), as well as for associated clusters (Fig. 5b).

In order to determine the energy of helium dissolution at the fifth stage of the curve $C_{\mathrm{eff}}(P)$, the process was simulated for two clusters consisting of nine vacancies, either containing four dissolved helium atoms, and for a combined cluster consisting of 18 vacancies, in which 24 helium atoms were dissolved (Fig. 5b). The energy of formation of the combined cluster was also evaluated.

The results obtained were the following: the energy of association of the two vacancy clusters is –3.276 eV and the energy of dissolution of 16 extra helium atoms (a total of 24 helium atoms dissolved) in the combined cluster consisting of 18 vacancies is –0.0325 eV (with regard to the association energy for the two initial vacancy clusters).

The energy of dissolution of extra helium atoms calculated per initial cluster is equal to $E_5^{\mathrm{d}} = -0.016$ eV (Table 3, last row). This means that the suggested mechanism of anomalous increase of helium solubility at high saturation pressures can be used for description of experimental data.

CONCLUSIONS

In this paper, we studied the diffusion and solubility of helium in submicrocrystalline palladium at temperatures T = 293–508 K and saturation pressures P = 0.1–35.0 MPa. As the saturation pressure increases, the helium solubility anomalously increases and the effective diffusion coefficient of helium first grows, exhibits a plateau, and then decreases to the initial value. Simulation of helium dissolution in submicrocrystalline palladium and phenomenological analysis of the high-temperature behavior of $D_{\mathrm{eff}}(P, T)$ indicate that the helium dissolves in clusters consisting of eight to ten vacancies, which are localized at grain boundaries, whereas diffusion is accomplished via helium transfer along grain boundaries, where vacancy clusters serve as traps. At low pressures, traps are empty clusters of eight to ten vacancies; at high saturation pressures, helium-filled vacancy clusters. The anomalous increase in the solubility, which is observed at saturation pressures of 25–35 MPa, is due to reversible pairwise association of adjacent vacancy clusters at grain boundaries when the specimen becomes saturated by helium at high pressures.

The mechanism of low-temperature diffusion invites further investigation.

ACKNOWLEDGMENTS

The authors are grateful to R.R. Mulyukov for submission of the specimens and to V.D. Risovanyĭ for assistance.

REFERENCES

1. A. N. Zhiganov, A. Ya. Kupryazhkin, R. R. Mulyukov, and K. Kh. Bitkulov, Zh. Tekh. Fiz. **72** (11), 96 (2002) [Tech. Phys. **47**, 1440 (2002)].
2. A. Ya. Kupryazhkin, A. N. Zhiganov, R. R. Mulyukov, *et al.*, Zh. Fiz. Khim. **78**, 1422 (2002).
3. A. N. Zhiganov and A. Ya. Kupryazhkin, in *Proceedings of the 7th All-Russia Conference on Reactor Materials*

Science, Dimitrovgrad, 2004, Vol. 3, Chap. 3, pp. 185–198.

4. R. Würschum, A. Kübler, *et al.*, Ann. Chim. (Paris) **21**, 471 (1996).
5. Shu Zhen and G. J. Davies, Phys. Status Solidi A **78**, 595 (1983).
6. A. N. Varaksin and V. S. Kozyaichev, Fiz. Met. Metalloved., No. 2, 45 (1991).
7. A. Michels and H. Wouters, Physica **8**, 923 (1941).
8. W. D. Wilson and R. A. Johnson, in *Interatomic Potentials and Simulation of Lattice Defects*, Ed. by P. C. Gehlen, J. R. Beeler, and R. I. Jaffee (Plenum, New York, 1972), pp. 375–385.
9. E. Zaremba and W. Kohn, Phys. Rev. B **15**, 1769 (1977).
10. K. M. Smith, A. M. Rulis, G. Scoles, *et al.*, J. Chem. Phys. **67**, 152 (1977).
11. A. Ya. Kupryazhkin, V. A. Gubanov, R. N. Pletnev, and G. P. Shveikin, *Defects and Diffusion of Gases in Crystals* (UNTs AN SSSR, Sverdlovsk, 1985) [in Russian].
12. A. B. Lidiard, Radiat. Eff. **53**, 133 (1980).
13. A. Ya. Kupryazhkin, I. R. Shein, and E. V. Popov, Zh. Tekh. Fiz. **53**, 1578 (1983) [Sov. Phys. Tech. Phys. **28**, 970 (1983)].

Translated by M. Lebedev

Technical Physics, Vol. 50, No. 8, 2005, pp. 1034–1037. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 71–74.

SOLIDS

Temperature Dependence of Spall Strength and the Effect of Anomalous Melting Temperatures in Shock-Wave Loading

Yu. V. Petrov and Y. V. Sitnikova

Research Institute of Mathematics and Mechanics, St. Petersburg State University, Universitetskaya nab. 7/9, St. Petersburg, 198504 Russia

e-mail: yp@YP1004.spb.edu

Received November 19, 2004

Abstract—The effects of temperature anomalies in materials subjected to the action of shock waves are studied. The spall failure of aluminum single crystals and polycrystals at various temperatures was experimentally studied in [1]. An analysis of the experimental data for polycrystalline aluminum shows that the breaking strength only weakly changes with temperature when it increases from room temperature to 90% of the melting temperature and, then, drops sharply to zero with a further increase in the temperature. For aluminum single crystals, the effect of anomalously high temperatures was experimentally detected; i.e., their strength remained high in the state where melting was expected during tension. The criterion of incubation time of failure is used to obtain an analytical expression for the temperature dependence of the spall strength of the materials. A new melting criterion, which relates the instant of a phase transition to the melting incubation period, is introduced. This criterion allows one to naturally explain the effect of anomalously high melting temperatures detected during the pulsed action. *© 2005 Pleiades Publishing, Inc.*

TEMPERATURE DEPENDENCE OF SPALL STRENGTH

The temperature dependence of the spall strength of materials can be obtained using the incubation time criterion [2–4]. This criterion is based on the principles of structural macromechanics of fracture and the concept of incubation time of failure, which is related to the dynamics of a relaxation process that prepares fracture. With this criterion, one can calculate the effects of unstable behavior of dynamic strength characteristics, which are observed in experiments on fracture of solids. The time dependence of strength detected under spalling conditions is a typical example of the complex behavior of the dynamic strength of solids.

In the case of spall failure, the incubation time criterion has the form

$$\int_{t-\tau}^{t} \sigma(t')dt' \le \sigma_C\tau, \qquad (1)$$

where $\sigma(t)$ is the time dependence of the local stress at the site of rupture; σ_C is the static strength; and τ is the incubation time of failure, which can be temperature-dependent.

In [1], measurements were performed when samples were subjected to plane shock waves created in the samples by an impact with an aluminum plate, and the thickness ratio of the flyer plate and a sample ensured a near-triangular wave profile. The problem of the reflection of a triangular compressive-stress pulse from the free end of a semi-infinite rod was considered in [2]. The x axis was assumed to be directed along the rod that was located to the right of zero ($x > 0$). The incident pulse was written in the form

$$\sigma_- = -P\left(1 - \frac{ct + x}{ct_i}\right)[H(ct + x) - H(ct + x - ct_i)].$$

Here, P is the pulse amplitude, t_i is the pulse duration, $H(t)$ is the Heaviside function, and c is the maximum wave velocity. The stress profile reflected by the free end was

$$\sigma_+ = +P\left(1 - \frac{ct + x}{ct_i}\right)[H(ct - x) - H(ct - x - ct_i)].$$

The total stress was written as $\sigma = \sigma_- + \sigma_+$. Then, to determine the breaking amplitude, we applied structure–time criterion (1). We analyzed threshold pulses, i.e., breaking pulses of a given duration and the minimum amplitude. At a pulse duration t_i, the minimum breaking amplitude P_* was calculated from the condition

$$\max I = \sigma_C, \quad I = \frac{1}{\tau}\int_{t-\tau}^{t} \sigma(t')dt'. \qquad (2)$$

As a result, we obtained equations to describe the time dependence of strength over the entire loading-

time range by making allowance for the incubation time of failure:

$$P_* = \begin{cases} \dfrac{2\sigma_C\tau(T)}{t_i}, & t_i \le \tau(T) \\ \dfrac{\sigma_C}{1-\dfrac{\tau(T)}{2t_i}}, & t_i \ge \tau(T). \end{cases} \quad (3)$$

The authors of [5] derived an analytical expression for the temperature dependence of the incubation time of fracture and showed that this incubation time can be expressed through a constant τ_0, which is equal to the period of atomic stretching vibrations. Taking this result into account, we propose to consider the temperature dependence of the incubation time of failure as

$$\tau = \tau_0\frac{G}{kT}, \quad (4)$$

where $k = 1.3807 \times 10^{-23}$ J/K is the Boltzmann constant, T is the temperature, and $\tau_0 = 10^{-13}$ s is the period of atomic stretching vibrations in a solid (the period of "elementary fluctuation").

As is known, kT is the energy of the vibrational degree of freedom in the equilibrium state. This energy is required to break an elementary bond, namely, an interatomic bond. The quantity G should be interpreted as an elementary energy portion required for destroying a structure cell.

We will obtain an expression for spall strength if the incubation time of failure τ in Eq. (3) for the time dependence of strength is assumed to depend on temperature T according to law (4). This result can be applied to analyze the experimental data of [1]. In experiments [1], samples were subjected to a triangular pulse and the pulse duration t_i was the same (0.12 μs) for all test temperatures. The tests were performed in the temperature range from 15 to 650°C, which is only 10°C lower than the melting temperature. For each material, we fit the elementary cell fracture energy G to the experimental threshold amplitude obtained at the minimum test temperature (15°C) [1]. Then, using Eq. (4), we automatically obtain the corresponding incubation time of failure for each material. The table gives the experimental threshold amplitudes at 15°C, the corresponding elementary fracture energies G calculated by Eq. (4), the incubation times of failure, and the static strength used in the calculations for both materials.

The temperature dependences of the spall strength calculated by Eqs. (3) with regard for Eq. (4) for aluminum single crystals and polycrystals are given in the figure; symbols demonstrate experimental data. The experimental data have a significant scatter; generally speaking, this is characteristic of such measurements with a high spatial resolution (the measurements were carried out interferometrically). More exact agreement between the calculated and experimental values can be reached as experimental data are accumulated and as the experimental techniques and calculation procedures are refined.

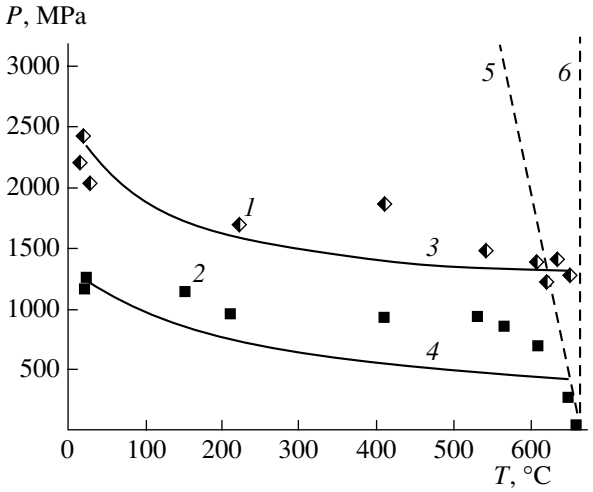
Table

	P_*(15°C), MPa	G, Y	σ_C, MPa	τ, μs
Polycrystalline aluminum	1240	2.97×10^{-14}	100	0.75
Single-crystal aluminum	2400	0.5727×10^{-14}	1000	0.15

Note that the incubation time of failure for polycrystalline aluminum at room temperature (0.75 μs) coincides with the value obtained in [5] for the experimental data of [6]. Bellendir [6] experimentally studied the brittle fracture of solids in a tensile-stress wave using plane collisions between a flyer plate and a target plate. Thus, the experiment performed in [1] for polycrystalline aluminum at room temperature agrees well with the classic experiments carried out at the Ioffe Physicotechnical Institute, and it can be described using the incubation time theory.

EFFECT OF ANOMALOUS TEMPERATURES

When sufficiently high stresses are accumulated in a material, it begins to melt at temperatures that are lower than the melting temperature of the material at zero pressure. To estimate the tensile stresses at which melt-



Temperature dependence of the breaking threshold amplitude of a 0.12-μs pulse for aluminum single crystals ((*1*) experimental data of [1], (*3*) calculation by Eq. (3)) and aluminum polycrystals ((*2*) experimental data of [1], (*4*) calculation by Eq. (3)). The tensile stresses corresponding to the onset of melting are calculated by (*5*) classic criterion (11) and (*6*) Eq. (10) at $\tau_m = 0.7$ μs.

ing begins, the authors of [1] considered the condition for the intersection of the melting curve of a solid in the form

$$\frac{dV}{dp} = \left(\frac{\partial V}{\partial T}\right)_p \frac{dT_m}{dp} + \left(\frac{\partial V}{\partial p}\right)_T, \quad (5)$$

where V is the volume, p is the pressure, and T is the temperature, and the isentrope of expansion of a solid, whose linearized equation of state is

$$V = V(T_{m0}) + (T_0 - T_{m0})\left(\frac{\partial V}{\partial T}\right)_{p=0} + \left(\frac{\partial V}{\partial p}\right)_S, \quad (6)$$

where T_0 is the initial test temperature (at $p = 0$), T_{m0} is the melting temperature at zero pressure, and S is the entropy.

We performed simple transformations: we multiplied both sides of Eq. (6) by p and transposed $V(T_{m0})$ from the right-hand side of Eq. (7) to its left-hand side. Then, we used the approximate formula

$$V - V(T_{m0}) \approx p\frac{dV}{dp}.$$

By equating the right-hand sides of the transformed equations, we obtain the desired relationship between temperature and pressure, which corresponds to a first-order phase transition:

$$p\alpha\frac{dT_m}{dp} - \frac{p}{K_T} = \alpha(T_0 - T_{m0}) - \frac{p}{K_S}, \quad (7)$$

where $\alpha = 1/V(\partial V/\partial T)_{p=0} = 1.12 \times 10^{-4}$ K^{-1} is the volume thermal expansion coefficient; $K_T = -V(\partial p/\partial V)_T = 56.7$ GPa and $K_S = -V(\partial p/\partial V)_S = 71.1$ GPa are the isothermal and isentropic bulk moduli, respectively; and $T_{m0} = 933.2$ K.

This dependence corresponds to a rather smooth (quasi-static) action and implies that melting is considered as an instantaneous process.

We express pressure in terms of temperature in Eq. (7) using the experimental value $dT_m/dp = 64.1 \times 10^{-3}$ K/MPa borrowed from [7] and take into account the fact that tensile stresses correspond to a negative pressure. Then, we obtain

$$P_m = \frac{T_{m0} - T}{0.032}. \quad (8)$$

In [1], the tensile stresses initiating melting of a material were estimated with this expression. For polycrystalline aluminum, the experimental data are concentrated below this estimate, whereas the experiments for single-crystal aluminum show that the intersection of the calculated boundary of the melting region is not accompanied by a sharp decrease in the breaking strength of the material. On the contrary, the material has a high strength.

To explain this effect, we introduce a new fracture criterion that corresponds to melting. We assume that melting is a process having an incubation period τ_m rather than an instantaneous process. The melting criterion is taken to be

$$\int_{t-\tau_m}^{t} \sigma(S)dS \le P_m\tau_m, \quad (9)$$

where P_m is the average ("equilibrium") stress that must appear within the incubation period τ_m to induce melting.

When the equality in Eq. (9) holds true, a material melts and the corresponding temperature is taken to be the melting temperature $T = T_m$. In its form, this criterion is seen to coincide with the incubation time criterion for brittle fracture. Using this criterion to solve the problem for triangular wave stress pulses considered above, we obtain the following dependence of the dynamic threshold amplitude $P_{*\text{melt}}$ (which causes melting) on the pulse duration t_i, the melting incubation period τ_m, and the temperature T:

$$P_{*\text{melt}} = \begin{cases} \dfrac{2P_m(T)\tau_m}{t_i}, & t_i \le \tau_m \\ \dfrac{P_m(T)}{1 - \dfrac{\tau_m}{2t_i}}, & t_i \ge \tau_m. \end{cases} \quad (10)$$

In a classical approach (when melting is considered to be an instantaneous event), i.e., at $\tau_m = 0$, we have

$$P_{*\text{melt}}(T) = P_m. \quad (11)$$

To determine conditions for the onset of melting, the authors of [1] used a classic melting criterion in the form of Eq. (11). High-temperature data for single crystals are above this limit; therefore, to calculate the stresses and temperatures corresponding to the onset of melting of aluminum single crystals, we applied criterion (10) where $\tau_m = 0.7$ μs. In the case of single-crystal aluminum, the melting temperatures that are calculated by model (10) of incubation time and correspond to the threshold breaking amplitudes are seen to be noticeably above the values predicted by classic melting model (11).

We estimated the melting incubation period τ_m for aluminum single crystals from the known overheating of these crystals with respect to the estimate obtained by the classic criterion. In other words, we assumed that experiment gives the temperature at which the spall strength of the single crystals begins to drop sharply due to the onset of melting. The difference between this temperature and its value corresponding to the onset of melting according to the classic criterion (the latter value is determined at the point of intersection of curves *3* and *5*) is the overheating for the aluminum sin-

gle crystals. As follows from [1], the overheating is 30°C, and the melting incubation time chosen by us corresponds to this overheating.

Thus, the model constructed by us explains the effect of a sharp drop in the strength at near-melting temperatures by the competition of two processes, namely, fracture and melting, having their specific incubation times. We can estimate the pressures and temperatures of this transition. It is determined as the point of intersection of the curves that correspond to each of these processes. The model of melting incubation time can account for the effect of anomalously high melting temperatures that was experimentally detected for aluminum single crystals in [1].

CONCLUSIONS

(1) We used a structure–time criterion and proposed the temperature dependence of the incubation time of failure to obtain an analytical expression for the temperature dependence of spall strength.

(2) We were the first to construct a model that takes into account the structural characteristics of a phase transition under wave loading. To determine the instant of a phase transition during shock-wave loading, we applied a new incubation time criterion.

(3) Using the melting criterion introduced in this work, we explained the effect of anomalously high melting temperatures that was experimentally detected in aluminum single crystals.

REFERENCES

1. G. I. Kanel and S. V. Razorenov, Fiz. Tverd. Tela (St. Petersburg) **43**, 839 (2001) [Phys. Solid State **43**, 871 (2001)].
2. Yu. V. Petrov, Preprint IPMASh RAN (Institute of Problems in Machine Science, Russian Academy of Sciences, St. Petersburg, 1996).
3. N. F. Morozov and Yu. V. Petrov, *Problems of Fracture Mechanics of Solids* (*Foundations of Engineering Mechanics*) (St. Petersburg. Gos. Univ., St. Petersburg, 1997; Springer-Verlag, Berlin, 2000).
4. Yu. V. Petrov, Dokl. Akad. Nauk **395**, 621 (2004) [Dokl. Phys. **49**, 246 (2004)].
5. P. A. Glebovskiĭ and Yu. V. Petrov, Fiz. Tverd. Tela (St. Petersburg) **46**, 1021 (2004) [Phys. Solid State **46**, 1051 (2004)].
6. E. N. Bellendir, Candidate's Dissertation (Ioffe Physicotechnical Institute, USSR Academy of Sciences, 1990).
7. T. Gorecki, High Temp.-High Press. **11**, 683 (1979).

Translated by K. Shakhlevich

Technical Physics, Vol. 50, No. 8, 2005, pp. 1038–1042. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 75–79.

SOLIDS

Dynamic Fatigue

S. N. Buravova, S. M. Gavrilkin, and Yu. A. Gordopolov

Institute of Structural Macrokinetics, Russian Academy of Sciences, Chernogolovka, Moscow Oblast, 142432 Russia

e-mail: gordop@ism.as.ru

Received March 29, 2004; in final form, December 28, 2004

Abstract—Dynamic fatigue differs from quasi-static fatigue in the spall damage nature. A consequence of the spall nature is the alternate appearance of longitudinal cracks, each of which becomes the source of rarefaction. The focusing or interference of rarefaction waves specifies the sites of nucleation and growth of channel and ring cracks; therefore, initially existing surface defects are not operative. Another consequence of the spall damage nature is the determining effect of the orientation of the lateral faces of flyer plates having finite dimensions, since these faces specify the intensity of the appearing lateral rarefaction wave. © *2005 Pleiades Publishing, Inc.*

INTRODUCTION

Pulsed loading differs from quasi-static loading in its wave character. Damage is of a spallation nature and is localized at the sites of meeting rarefaction waves accompanying a compression pulse. One-dimensional spallation has been studied rather comprehensively, and damage has been shown to be accumulated in time during tension of a sample by single rarefaction waves [1]. However, there are many engineering objects that are subjected to the action of flyer plates having finite dimensions and low velocities, when a discontinuity occurs after damage accumulation as a result of multiple loading. For example, such objects are steam turbines, an aircraft skin, the screws of fast ships, pulverized-fuel burners, and railroad rails. One of the most important problems of track facilities is contact–fatigue damage of rail heads in the region of compressive stresses. The cause of longitudinal and transverse cracks appearing at the sites where they should not appear according to all the well-known rules is still unknown [2]. Surface loading at impact velocities below 200–300 ms^{-1} is usually considered in the quasi-static approximation, which is a generally accepted approach. Apparently, this is one of the reasons why researchers cannot explain the appearance of defects in rails, the appearance of transverse cracks in the steady-state stage of drop erosion, the formation of pits in craters during cavitation, the formation of spalling cracks in blades, and so on. The purpose of this work is to reveal the specific features of surface fatigue fracture under conditions of multiple pulsed low-rate loading.

EXPERIMENTAL DATA

The propagation of a wave load was experimentally studied using pulsed laser irradiation. Detonation 500-μm-thick coatings of chromium carbide with a nickel binder (the particle size was 40–60 μm) were subjected to 1.06-nm pulsed laser irradiation at a pulse duration of 8–10 ms. The coatings were sprayed in a Korund automatic setup using the products of detonation of a mixture of propane, butane, and oxygen. The irradiated samples were analyzed metallographically on a Neophot-30 microscope. The shock-wave intensity appearing during laser irradiation was close to that created by the collisions of particles having a velocity of 50–250 ms^{-1} (M_0 = 0.01–0.05 is the Mach number, which is equal to the ratio of the impact velocity to the sound velocity). Figure 1 shows transverse polished sections of the irradiated samples. It is seen that irradiation creates a crater and a longitudinal crack under it on the target surface and that a transverse crack is localized near the loaded zone (Fig. 1a). An increase in the irradiation power leads to the formation of a system of smaller longitudinal cracks accompanying the main crack (Fig. 1b). A system of longitudinal cracks was also observed during the irradiation of brittle melted coatings of aluminum oxide. The irradiation of the detonation coatings of tungsten carbide with a nickel binder (with a particle size of about 100 μm) is also accompanied by the formation of longitudinal cracks. However, their penetration depth is substantially lower, since coarser particles in the coatings can hinder crack extension deep into the target.

SURFACE EROSION DAMAGE INDUCED BY LOW-RATE LOADING

The action of flyer plates having finite dimensions results in compression waves in the interacting bodies and a centered rarefaction wave, whose center is located at the interface of the contacting bodies, and the rarefaction wave originates from the space subjected to a compression pulse. The rarefaction wave weakens

1063-7842/05/5008-1038$26.00

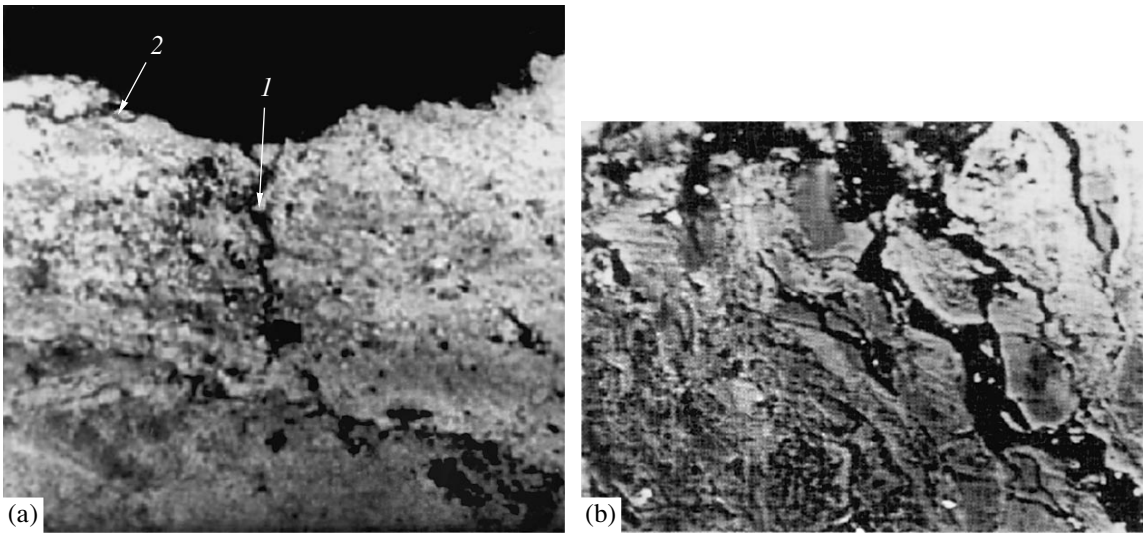


Fig. 1. Transverse polished sections of the sample with the detonation coating of chromium carbide with a nickel binder after pulsed laser irradiation: (a) single (*1*) longitudinal and (*2*) transverse crack and (b) set of longitudinal cracks.

and distorts the shock wave [3, 4]. Wave propagation is considered in the two-dimensional approximation. The set of equations that describes the interaction of deformed bodies is solved by the method of characteristics using a small parameter $\sigma = 0.5(n + 1)u/c_0 = 0.5(n + 1)P/\rho_0 c_0^2$, in terms of which all propagation characteristics can be expressed. Here, u is the particle velocity; P is the pressure; n is the isentropic exponent; and ρ_0 and c_0 are the density and sound velocity of the unperturbed material, respectively. Each state in the distorted front, which is characterized by a parameter σ, moves in the space along straight-line trajectories whose angles with the target face surface are determined from the expression

$$\tan\psi = \frac{1 + 2\sqrt{\sigma\sigma_0} - 0.5\sigma_0 - \sigma}{2\sqrt{\sigma} - \sqrt{\sigma_0}}. \qquad (1)$$

Here, $\sigma_0 = 0.5(n + 1)u_0/c_0$ and u_0 is the particle velocity of the ingoing shock wave. The shock-wave intensity σ_b at a lateral face with an arbitrary slope φ_0 is determined from Eq. (1) by the substitution $\psi = \varphi_0$. It is also seen from Eq. (1) that the compression pulse propagates inside a conical surface whose angle ψ_0 is calculated using the substitution $\sigma = 0$ in Eq. (1). Inside this cone surface, the material is hardened. The intensity σ_b of the distorted shock wave moving along a particle lateral face is lower than the initial one. For lateral faces oriented normally to the sample surface, the pressure is one-fourth of the initial pressure ($\sigma_b = \sigma_0/4$). The rest of the pressure is decreased by the second rarefaction wave resulting from the particle lateral face. The interaction of flyer plates having finite dimensions with a massive target is accompanied by the formation of two rarefaction waves when the slope of the lateral face does not exceed the opening angle of the compression cone.

Meeting the lateral rarefaction waves in the axis of symmetry leads to the formation of a channel zone of tensile stresses. Tension is caused by the second rarefaction wave, since it is the wave that imparts a translational velocity $u_b/c_0 = 0.5(n + 1)\sigma_b\sin\varphi_0$ to the particle lateral faces. Upon low-rate loading, the stress in the tension zone is substantially lower than the spall strength of the material P_s, $u_s/c_0 = P_s/\rho_0 c_0^2$. Therefore, a discontinuity appears after deformation has accumulated under conditions of dynamic fatigue. The number of loading cycles N_0 required to reach the spall strength can be calculated from the expression

$$N_0 = \frac{1}{\xi}\left(\frac{\sigma_s}{\sigma_b\sin\varphi_0} - 1\right), \qquad (2)$$

where ξ is the elastic–plastic hysteresis. The number N_0 characterizes the erosion wear resistance of the material. A discontinuity (pore) appears on the surface; upon the following loading, it grows deep into the material as a channel crack. However, the channel crack becomes the source of a rarefaction wave, which creates conditions for damage accumulation in a ring zone. As a result, by the instant when the number of pulses doubles ($N = 2N_0$), a first-generation crack appears around the channel crack and the channel crack ceases to grow, which is related to a decrease in the distance between rarefaction sources. The maximum depth of the channel crack L_0 (which a dimensionless quantity reduced to the particle radius) is determined from the accumulation condition of a critical deformation at the tip of a growing crack. Since tension is caused by the radial component of the rarefaction rate, the crack depth is $L_0 = \tan\beta$.

The angle β is calculated from the expression $\cos\beta = \sigma_s/(1 + \xi N_0)\sigma_b \sin\varphi_0$. The following cyclic loading leads to the formation of a system of smaller ring spalling cracks.

The character of the longitudinal damage changes when the slope of the particle lateral face exceeds the opening angle of the compression cone ($\varphi_0 \geq \psi_0$). In this case, the shock-wave intensity at the lateral face becomes zero and the second rarefaction wave does not appear. This situation takes place for a spherical particle. However, in this case, tensile stresses σ_a induced by flow divergence, $\sigma_a = 0.5\sigma_0 \tan 0.5(\psi_0 - 90°)$, appear under the contact zone. The critical number of cycles required for the appearance of a channel crack is determined from Eq. (2), where σ_a should substitute for $\sigma_b \sin\varphi_0$. As a rule, $\sigma_a \leq \sigma_b$, and a discontinuity appears at a larger number of loading cycles. It is known from experiments [5–8] that a plane-face eroding agent destroys a target faster than a spherical particle of the same size.

The presence of many longitudinal cracks in the hardened layer results in the fragmentation of the material during loading and in the removal of debris from the contact zone. The processes of erosion and removal of material from the contact zone have not yet been studied. As a result of material loss from the contact zone, a valley forms on the surface, and this formation ends when the slope of the wall becomes lower than the opening of the compression cone ψ_0. In this case, an impact with the bottom of the valley induces a shock wave that can pass through the wall to the target [3]. The exit of this pulse to the face surface is accompanied by the formation of a reflected rarefaction wave. Its interference with the incident wave creates a transverse zone of tensile stresses. At a distance of $\delta = 0.125\sigma_0 h$ from the face surface, a face spalling crack begins to form (for the case of normal wall orientation), where h is the height of the valley lateral wall. The crack length λ depends on angle ψ^* of the trajectory of the state where deformation reaches a critical value $\sigma^* = \sigma_s/(8\sigma_s/\sigma_0 - 1)$ at $N = 2N_0$ cycles. The angle ψ^* can be found via the substitution of σ^* into Eq. (1), and the transverse crack length in this case is $\lambda = h\tan(\psi^* - 90)$. The values of λ and δ specify the size of the forming fragment.

Thus, multiple action of a flyer plate of finite dimensions leads to the formation of micro- and macrocracks. In the initial stage, channel and ring cracks are oriented longitudinally, and, in the final stage, face spalling is oriented transversely.

DYNAMIC NATURE OF TARGET DAMAGE AT LOW-RATE LOADING

A comparison of the surface damage morphologies induced by pulsed laser irradiation and high-rate loading indicates that they are similar [9]. Indeed, the interaction of a flyer bar with a metallic target at an impact velocity of 1 km s^{-1} ($M_0 = 0.2$) [10] results in the formation of a valley on the target surface; this valley has a funnel at its center and a longitudinal crack propagating from this funnel deep into the target. A transverse crack is localized near the crater, under the edge of the surface relief. The same similarity is exhibited by the experimental study of erosion induced by a low-velocity particle flux [5–8]. Single small particles (70 μm in size) that do not fail during an impact ($M_0 = 0.02$) [11] create voids at the bottom of the crater, and the voids propagate deep into the material as longitudinal microcracks. As noted by the authors of [11], the cause of homogeneous crack nucleation is still unknown. In the steady-state erosion stage, the material is rapidly removed, and the characteristic feature of this process is the formation of transverse cracks inclined at a small angle with the target face surface. As a result, wear products acquire a lamellar, flaky shape [6–8]. Cavitation damage is characterized by the formation of isolated valleys having deep pits at their centers and longitudinal cracks penetrating deep into the material from these pits [12]. The pressure induced in the target by periodic surface loading by spherical waves going from the zone of collapsing cavitation bubbles is estimated to be 50–1000 MPa [12–14]. The same pressure appears in the target subjected to an impact with a drop moving at velocities $M_0 = 0.005$–0.03 [8, 14]. Cavitation damage in the steady-state stage is characterized by the appearance of large dints, and material is removed through the fracture of the edges of such pitting aggregations [12–14]. Longitudinal cracks also form during detonation spraying, where the velocity of a molten particle flux is 300–500 ms^{-1}. After removing the coating, one can see stuck grains of the sprayed material in round dints on the substrate surface that are caused by collisions with drops. A longitudinal crack forms before the molten material solidifies. Upon spraying, transverse cracks are only localized at the peaks of large ridges on the surface relief.

Thus, the analysis of the experimental data indicates that, irrespective of the type of pulsed loading (cavitation loading, drop loading, dust loading, laser irradiation, or detonation spraying), the result of a multiple pulsed low-velocity action on a target is identical to the fracture induced by high-rate loading. This finding serves as a basis for the application of the wave mechanics to low-rate loading. Dynamic fatigue should reflect the specific features of propagating wave processes. Under conditions of surface loading by flyer plates having finite dimensions, the damage is caused by non-one-dimensional spallation and results from the focusing or interference of rarefaction waves appearing at the lateral faces of a flyer plate.

NUMERICAL EXAMPLE

We calculate interaction with barlike particles having a square section and a trapezoidal section; the slope

of the lateral faces of the trapezoidal particles coincides with the opening angle of the compression cone, $\varphi_0 = \psi_0$. In this case, the faces do not move and tensile stresses in the channel zone are only caused by flow divergence. We assume that fracture occurs instantaneously, as soon as tensile stresses reach a critical value. The calculation results are shown in Figs. 2–4. Numerals *1* and *2* belong to the square and trapezoidal sections, respectively. The calculation was performed at $\sigma_s = 0.0313$. The loading rate range includes the ranges of dynamic fatigue and a transition to high-rate loading, when fracture appears upon single loading at $N_0 = 0$. As is seen, the reaction of a target material to pulsed loading depends substantially on the particle geometry. The number of loading cycles depends on the particle shape and can be significantly different at the same loading rates. For each orientation of the lateral faces of the particles with finite dimensions, there is a certain critical velocity M_{cr} of a transition to the high-velocity range. When the particles and the target are made of the same material, $M_{cr} = 2u_b/c_0 = 4\sigma_b/(n + 1)$ and a discontinuity nucleus in the form of a void appears on the target surface. An increase in the impact velocity at the same number of loading cycles is accompanied by an increase in the crack depth; however, in the range of a low-velocity impact, an excess increase in the critical number of cycles results in an increase in the longitudinal crack depth as the pulsed load decreases (Fig. 3). The behavior of a transverse crack, which is the main cause of material removal in the steady-state stage of wear, is of interest. At low impact velocities, erosion wear products have a lamellar shape. As the impact velocity increases, the ratio of the face crack length to the distance δ of its nucleation under the surface (λ/δ) decreases and becomes zero at the critical impact velocity (the slope of the transverse crack $\gamma = \delta/\lambda$ behaves oppositely). A further increase in the velocity leads to an increase in the ratio λ/δ. Thus, the fluky shape of the wear products, which was determined experimentally, can be revealed out of a short velocity range near the critical velocity. It should be noted that nonunidimensional spallation has not been studied and that spall strength and an elastic–plastic hysteresis under conditions of the focusing effect of rarefaction waves are unknown. Therefore, the calculation given above is approximate. Nevertheless, it can reveal the specific features of dynamic fatigue. The response of a material to the action of a flyer plate having finite dimensions manifests itself in the creation of a channel crack, and it is this crack that determines the maximum depth of the growing valley. The number of longitudinal cracks continues to increase until conditions for the appearance of a face spalling crack are satisfied.

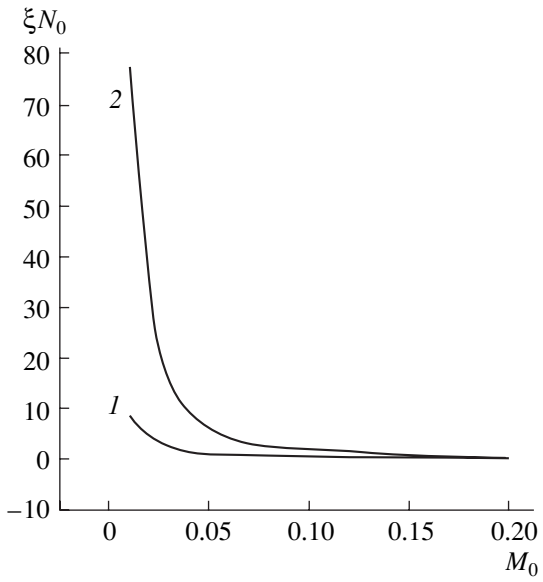


Fig. 2. Dependence of the critical number of cycles on the Mach number.

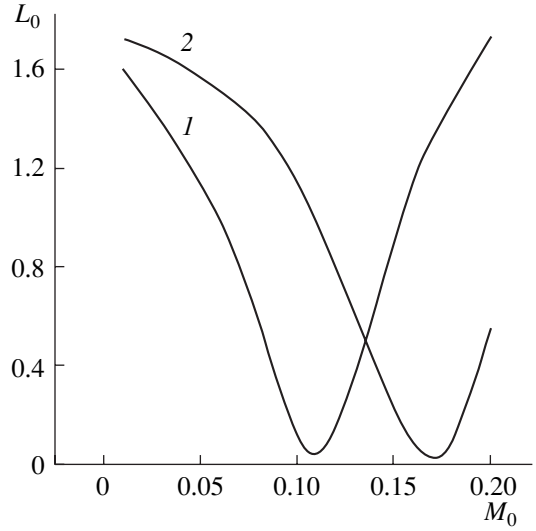


Fig. 3. Dependence of the depth of a channel spalling crack on the Mach number.

DIFFERENCES BETWEEN DYNAMIC FATIGUE AND QUASI-STATIC FATIGUE

The basic result of our work is the detection of dynamic fatigue, which differs from quasi-static fatigue. During quasi-static loading, the characteristic size of a load (the product of the sound velocity into the loading time) exceeds the size of a sample subjected to loading. The penetration of an indenter into a target is accompanied by the formation of a tensile-stress fan and the nucleation of radial cracks at defects and stress concentrators existing on the surface. The cracks extend simultaneously during cyclic loading.

Under conditions of many-cycle dynamic loads, the characteristic size of a load is well below the sample size and cracks result from the interference or focusing of rarefaction waves. Spalling cracks are strictly localized at the sites of meeting rarefaction waves, and they

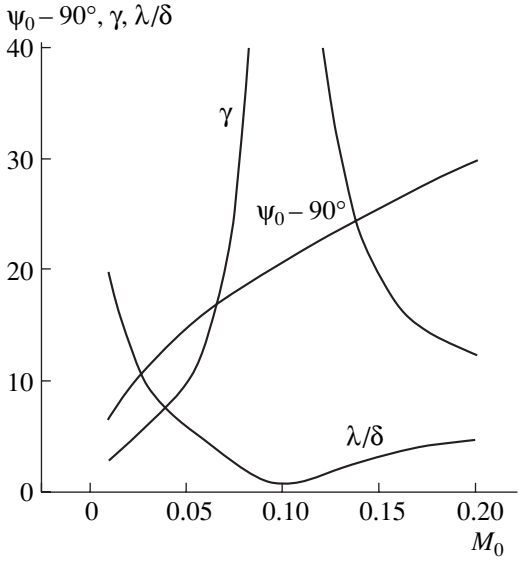


Fig. 4. Dependence of the parameters of transverse damage on the Mach number.

are mainly oriented normally to the target face surface. As a consequence of the spall damage nature, stress concentrators (defects) do not operate during dynamic loading and, thus, cannot be sources for the nucleation of a fracture. Under conditions of dynamic fatigue, spalling cracks form alternately. In the initial stage, a channel crack is formed. A particle "stings" the surface during an impact. A longitudinal crack becomes the source of rarefaction waves; therefore, ring cracks nucleate and grow sequentially around the channel crack. Due to a decrease in the distance between the sources of rarefaction waves, which are longitudinal cracks, the depth of ring cracks of a certain generation decreases as compared to that of the previous generation. Another consequence of the spall nature of dynamic damage is the determining role of the particle geometrical shape, since the orientation of the particle lateral faces specifies the intensity of the appearing rarefaction wave. The fact of a substantial effect of the face angle of an angular particle on the damage during erosion was supported experimentally. The authors of [6, 7] believe that the face angle of an angular particle is as important as the velocity and impact direction.

ACKNOWLEDGMENTS

We thank A.F. Belikova, A.A. Goncharov, and Yu.N. Kiselev for their assistance in the experimental work and for discussions of the results.

REFERENCES

1. N. A. Zlatin, G. S. Pugachev, L. D. Volovets, and S. A. Leont'ev, Zh. Tekh. Fiz. **51**, 1507 (1981) [Sov. Phys. Tech. Phys. **26**, 863 (1981)].
2. V. S. Lysyuk, V. N. Sazonov, and L. V. Bashkatova, *Firm and Reliable Railway Track* (Akademkniga, Moscow, 2003) [in Russian].
3. S. Buravova, Wear **157**, 359 (1992).
4. A. A. Grib, A. G. Ryabinin, and S. A. Khristianovich, Prikl. Mat. Mekh. **20**, 532 (1956).
5. A. G. Evans, in *Erosion*, Ed. by C. Preece (Academic, New York, 1979; Mir, Moscow, 1982), pp. 11–79.
6. A. W. Ruff and S. M. Wiederhorn, in *Erosion*, Ed. by C. Preece (Academic, New York, 1979; Mir, Moscow, 1982).
7. A. H. Cousins and I. M. Hutching, Wear **28**, 335 (1983).
8. J. H. Branton and M. K. Rochester, in *Erosion*, Ed. by C. Preece (Academic, New York, 1979; Mir, Moscow, 1982), pp. 140–200.
9. S. N. Buravova, A. A. Goncharov, and Yu. N. Kiselev, Tribol. Int. **29**, 357 (1996).
10. G. A. Adadurov, A. F. Belikova, and S. N. Buranova, Fiz. Goreniya Vzryva, No. 4, 95 (1992).
11. R. Brown, S. Kosko, and E. J. Jun, Wear **88**, 181 (1983).
12. I. Hanson and K. J. Kristensen, J. Phys. D **11**, 891 (1978).
13. E. P. Georgievskaya, *Cavitation Erosion of Screw Propellers* (Sudostroenie, Leningrad, 1978) [in Russian].
14. *Erosion*, Ed. by C. Preece (Academic, New York, 1979; Mir, Moscow, 1982), pp. 269–330.

Translated by K. Shakhlevich

Technical Physics, Vol. 50, No. 8, 2005, pp. 1043–1047. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 80–84.

SOLID-STATE ELECTRONICS

Fabrication of Semiconductor- and Polymer-Based Photonic Crystals Using Nanoimprint Lithography

E. M. Arakcheeva*, E. M. Tanklevskaya*, S. I. Nesterov*, M. V. Maksimov*, S. A. Gurevich*, J. Seekamp, and C. M. Sotomayor Torres****

** Ioffe Physicotechnical Institute, Russian Academy of Sciences, Politekhnicheskaya ul. 26, St. Petersburg, 194021 Russia*

e-mail: kathy.quantum@mail.ioffe.ru

*** Department of Electrical and Information Engineering, Institute of Materials Science, University of Wuppertal, Gauss-Strasse 20, D-42097 Wuppertal, Germany*

Received October 4, 2004

Abstract—The technology of fabricating photonic crystals with the use of nanoimprint lithography is described. One- and two-dimensional photonic crystals are produced by direct extrusion of polymethyl methacrylate by Si moulds obtained via interference lithography and reactive ion etching. The period of 2D photonic crystals, which present a square array of holes, ranges from 270 to 700 nm; the aperture diameter amounts to the half-period of the structure. The holes are round-shaped with even edges. One-dimensional GaAs-based photonic crystals are fabricated by reactive ion etching of GaAs to a depth of 1 μm through a mask formed using nanoimprint lithography. The resulting crystals have a period of 800 nm, a ridge width of 200 nm, and smooth nearly vertical side walls. © *2005 Pleiades Publishing, Inc.*

INTRODUCTION

The production and investigation of photonic crystals are a rapidly developing area of present-day optoelectronics. In view of the unique properties of photonic crystals, their dispersion may be hundreds of times that of conventional prisms [1] and can provide effective collimation of a laser beam [2]. Furthermore, photonic structures hold much promise as elements of waveguide optoelectronic devices, such as beam splitters [3] and filters [4]. Their size may be only several micrometers, which is by orders of magnitude smaller than in the case of traditional strip waveguides. It is expected that further advances in designing both active [5, 6] and passive elements on the basis of semiconductor photonic crystals will lead to the advent of optical integrated circuits in the future.

In view of the achievements in the production of highly efficient light-emitting devices based on polymeric and organic compounds, these materials are also of considerable interest for the fabrication of photonic crystals. A drawback of such photonic crystals relative to their semiconductor analogues is typically lower refractive indices. However, they have the advantages of comparatively low price and simplicity of production.

The need for industrial fabrication of polymeric and organic photonic crystals stimulates search for high-output, comparatively inexpensive "soft" technologies. Among the most promising methods is nanoimprint lithography, which was proposed in 1995. In essence, this technique is extrusion of a thin polymer film by a mould and, thus, transfer of the mould pattern to it. The high output of this technique at a comparatively low cost, along with the possibility to provide for sub-25-nm structures, make it suitable for commercial production of optical integrated circuits on a photonic crystal basis. In this study, we describe a production technology of one- and two-dimensional photonic crystals intended for passive elements in polymer- and GaAs-based semiconductor structures.

PRINCIPLES OF NANOIMPRINT LITHOGRAPHY

Figure 1 presents the main stages of the nanoimprint lithography process [7]. At first, the substrate surface is covered by a polymer layer with appropriate glass transition temperature T_g and molecular weight. The polymer layer thickness is usually about several hundreds of nanometers. Then, the polymer layer is heated to a temperature above T_g and subjected to pressing (imprinting) by a mould with a desired pattern. During the nanoimprint process, which typically lasts for a few minutes, the liquid polymer redistributes, filling the bulk volume defined by the mould profile (Fig. 1b). The mould pattern is transferred to the film. After cooling the mould and the polymer, the former is detached from the substrate (Fig. 1c).

1063-7842/05/5008-1043$26.00

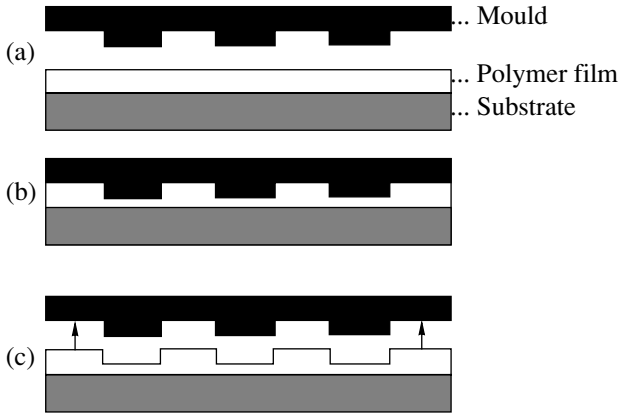


Fig. 1. Nanoimprint schematic representation: (a) mould and polymer film substrate, (b) imprinting, and (c) mould detachment.

In a first approximation, the ultimate spatial resolution of nanoimprint lithography depends on the minimal feature size of the mould. Moulds with features smaller than 100–200 nm are usually prepared by electron-beam lithography and dry etching from high-strength materials (metals, silicon, etc.). The production of moulds with larger features conventionally employs the optical lithography methods, including interference lithography.

The configuration of the mould and the properties of the polymer to be imprinted should fit each other. One of the widely used polymers is polymethyl methacrylate (PMMA). This material proves to be an excellent candidate for imprinting, since it has a low thermal expansion coefficient of $\sim 5 \times 10^{-5}$/°C and a low compression ratio under pressure, $\sim 5.5 \times 10^{-11}$ Pa^{-1}. The polymer properties define the pressure and temperature during imprinting. For PMMA with a glass transition temperature of about 105°C, the imprinting temperature usually ranges from 140 to 180°C and the pressure lies between 4 and 13 MPa. In these conditions, the thermal contraction of PMMA is under 0.8% and the compression ratio is under 0.07% (the lowest value is attained at higher pressures); therefore, the imprinted profile must correspond to the mould profile. As a rule, the thickness of the deposited resist exceeds the maximum profile depth, which makes it possible to preclude contact between substrate and mould and to prolong the service life of the mould. The ratio between the imprint depth and the residual resist thickness is typically 3 : 1.

To reduce the number of air bubbles appearing in the resist, we carried out the process in vacuum conditions. The equipment used for nanoimprinting should ensure an extremely high accuracy of the temperature and pressure control, as well as the parallel orientation of mould relative to the substrate throughout the process, which is necessary to obtain a residual resist layer of uniform thickness. Another problem is to avoid a temperature gradient; this can be done by using an appropriate combination of heating and cooling elements. The requirements mentioned above are the qualifying standards for lithography equipment to suit the nanoimprint applications.

The reproducibility of the imprint procedure and the service life of the mould are two key factors that define the nanoimprint potential for industrial application. The same mould can be used for more than 30 routine imprint procedures without any noticeable change either in the PMMA profile or in the mould.

EXAMPLES OF NANOIMPRINT APPLICATIONS

Fabrication of photonic crystals on the basis of polymeric materials. Nanoimprints can be used for the direct production of polymeric or organic photonic structures on substrates of various types. By way of example, we consider the production of PMMA photonic crystals on a Si substrate.

To obtain Si moulds in the form of periodic stripes or square networks with round holes in the nodes, the appropriate etch masks were prepared by means of interference photolithography, i.e., by exposing the photoresist to a fringed interference pattern resulting from the interference of two He–Cd laser beams (wavelength 442 nm). The interference pattern period is defined by the laser wavelength and the angle between the interfering beams and can be controlled with a high accuracy. A single exposure and the following development of the resist yield a mask in the form of parallel periodic stripes (Fig. 2a). If a specimen after the first exposure is not developed immediately but is turned by 90° around the normal to its surface and is repeatedly exposed, a mask with a square network appears after development. As a photoresist, we used 150-nm-thick As_2S_3 chalcogenide glass layers thermally evaporated in vacuum. The advantages of this inorganic photoresist are high resolution and high development contrast ensuring the verticality of the mask edges. Electron-beam evaporation was used to deposit a 50-nm-thick layer of nickel on the primary mask of the chalcogenide glass, which was obtained after development. After that, the primary mask was dissolved in an alkali solvent, which also removed the nickel layer deposited on the chalcogenide mask. As a result, only the substrate areas that were uncovered by the primary mask retained nickel on the silicon surface. In this manner, striped nickel masks (Fig. 2b) or square nets of metallic disks were formed on the silicon substrates. The pattern period ranged between 270 and 700 nm, and the stripe widths, or the disk diameters, were nearly equal to half of this period. The obtained masks had a sufficient plasma resistance to be used in subsequent reactive ion etching of silicon.

In this study, we employed an Alcatel RDE-300 plasma etcher, which is a diode-type computer-controlled setup providing reactive ion etching in rf plasmas. When selecting the etching parameters, we proceeded from the requirement that the side walls of the resulting structure should be as close to the vertical as possible. Silicon etching was carried out in an SF_6 plasma [8] with 10-sccm (standard cubic centimeters per minute) SF_6 flow, 0.6-Pa gas pressure in the reactor (the reactor was preliminarily pumped down to 5×10^{-4} Pa), and 200-V plasma self-bias voltage. The average etching rate in these conditions was ≈140 nm/min.

Figures 2c and 3a show the pictures of the Si moulds, which present one-dimensional and two-dimensional photonic crystals, respectively. The sizes of the stripes and microcolumns correspond to those of the mask, and the side walls are sufficiently vertical to ensure high-quality imprinting.

The nanoimprinting procedure was carried out in the University of Wuppertal, Germany. To increase the difference between the refractive indices of the polymer and the substrate, the substrate surface was thermally oxidized to form a layer of silicon oxide with a refractive index of 1.45. Then a PMMA layer was deposited, and nanoimprinting was performed. Figure 3b presents a picture of a two-dimensional polymer-based photonic crystal obtained using the nanoimprint method. It is seen that the holes are round-shaped, with fairly smooth edges.

NANOIMPRINT LITHOGRAPHY

Using the method of nanoimprint lithography to prepare etch masks holds special promise for opto- and microelectronic submicronmeter semiconductor device technologies. Nanoimprint lithography is beneficially distinguished from conventional nanotechnologies by a comparative simplicity of implementation and the absence of expensive technological procedures, such as electron-beam lithography. Unlike the standard lithographic techniques, nanoimprinting involves no energy beams. Because of this, its resolution is free from limitations associated with wave diffraction, scattering, reflection from the substrate, and interference in the resist. An additional advantage of nanoimprinting over electron-beam lithography lies in that a specimen is not exposed to high-energy electrons, which penetrate the surface layer and may deteriorate the material quality.

We studied the potential of nanoimprint lithography in the production of semiconductor photonic crystals in collaboration with a scientific team of the University of Wuppertal, Germany. German colleagues supplied a variety of patterns imprinted in the PMMA resist on the GaAs substrate: 5-μm-wide stripes spaced by 5, 0.2, and 0.08 μm, as well as some other patterns with linear sizes from 1 to 20 μm.

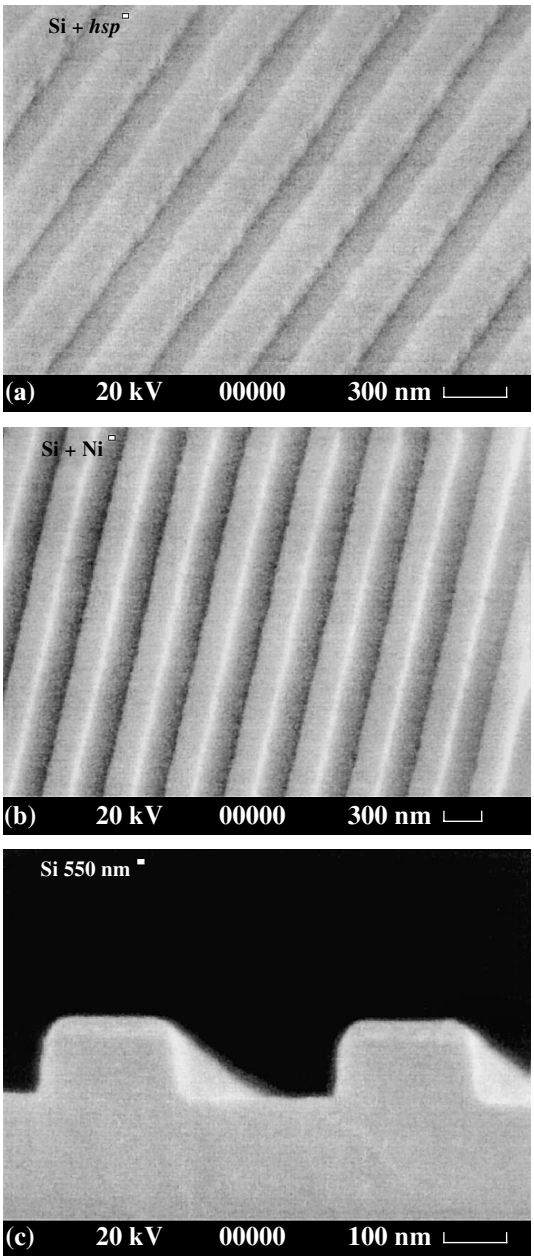


Fig. 2. (a) Si substrate covered by the primary chalcogenide mask with 550-nm period, (b) Si covered by a Ni mask, and (c) the silicon mould in the form of a one-dimensional photonic crystal as obtained after the Si etching through the Ni mask.

Our task was to transfer the contrast imprinted in the PMMA resist into the GaAs layer as deeply as possible without a loss of resolution. With the help of a Dektak profilometer, we determined the contrast: the PMMA thickness was 400 and 100 nm at the nonimprinted regions and in the bottom layers, respectively. Then, the specimens were subjected to reactive ion etching in oxygen plasma with the following parameters: the self-bias voltage 50 V, the O_2 flow 20 sccm, and the gas pressure in the reactor 1.0 Pa. This regime corresponds to

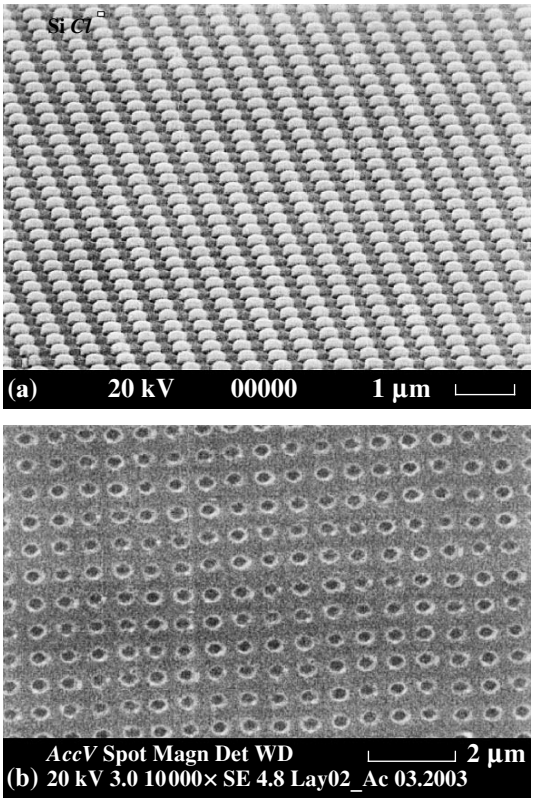


Fig. 3. (a) Silicon mould in the form of a two-dimensional photonic crystal with a 700-nm period, as obtained after the Si etching through the Ni mask and (b) two-dimensional polymeric photonic crystal.

the PMMA etch rate ≈40 nm/min. As a result, the bottom layers were etched off and the patterned resist with a thickness of ≈300 nm remained on the substrate. It is known that PMMA is not a plasma-resistant material; therefore, it seems impossible to etch GaAs through such a mask. The proper mask was obtained by the lift-off method, which consisted in the following. The PMMA pattern was covered by a Ni layer ≈80 nm. After that, the PMMA areas under the deposited Ni were removed in dimethyl formamide. A nickel mask thickness of ≈80 nm is sufficient to etch the GaAs layer to a depth of 2 μm and deeper. The reactive ion etching of GaAs was carried out as follows: the ratio between the gas components $Cl_2 : BCl_3 : Ar = 1 : 4 : 16$ sccm, the gas mixture pressure in the reactor 0.8 Pa, and the self-bias voltage 100 V. The average etch rate was ≈70 nm/min.

Stripes with a width of 200 nm and a depth of 1 μm (Fig. 4) have remarkably smooth and nearly vertical side walls. This is evidence that the bottom layers were almost entirely removed by etching in the oxygen plasmas and, therefore, the pattern on the substrate perfectly corresponds to that on the mould. Figure 4 presents the pictures of (a) the mould, (b) the bottom PMMA, and (c) the etched semiconductor.

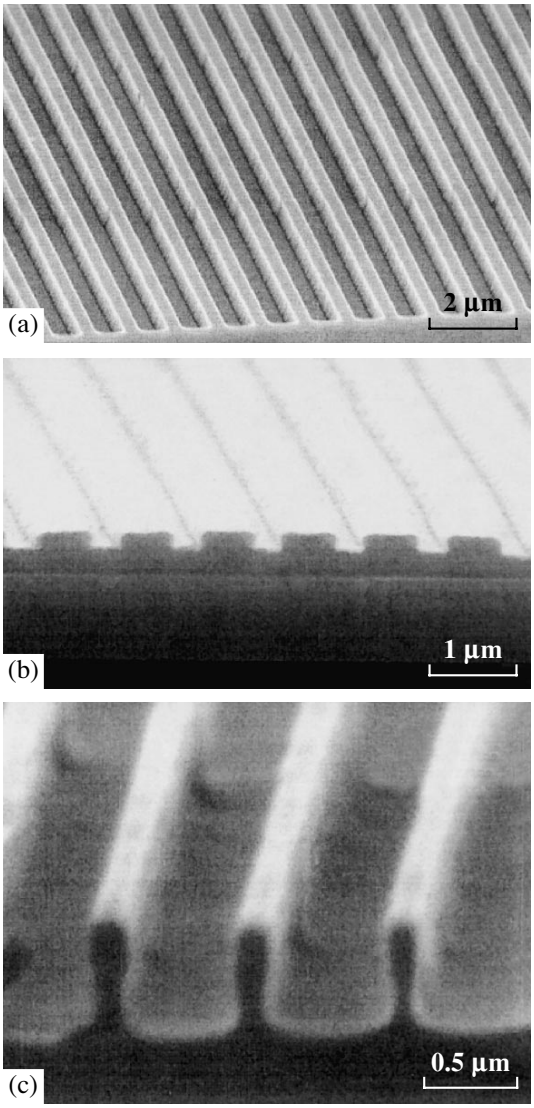


Fig. 4. (a) Silicon mould with an 800-nm period produced in Finland; (b) pressed-in PMMA on the SiO_2 substrate produced in Germany; and (c) one-dimensional photonic crystal obtained by the GaAs etching through the Ni mask produced at the Ioffe Physicotechnical Institute, Russia.

To simplify the technology of semiconductor photonic crystals, it appears reasonable to use polymers with a higher resistance to dry etching, which would allow one to etch the resist directly through the mask obtained by nanoimprinting. Of much considerable interest for improving the imprint precision is the possibility of a controllable diminishing of the mask feature size by etching in oxygen plasmas.

CONCLUSIONS

It is shown that the nanoimprint method provides the possibility of a direct fabrication of photonic crystals based on polymeric compounds; in addition, it is a rapid and efficient lithographic technique for the pro-

duction of masks intended for etching semiconductor structures. Thus, nanoimprinting seems to be a promising low-cost lithographic method for the industrial production of semiconductor devices and integrated circuits with high (up to 10-nm) resolution. There is good reason to believe that nanoimprinting will soon become the main method for the direct production of rapidly advancing micro- and optoelectronic devices on the basis of organic and polymeric compounds.

ACKNOWLEDGMENTS

This work was supported by the Department of General Physics and Astronomy, Russian Academy of Sciences, within the program "New Materials and Structures" and the Presidium of the Russian Academy of Sciences within the program "Low-Dimensional Quantum-Size Structures."

We thank V. M. Busov and S. I. Troshkov for carrying out the scanning electron microscopy studies.

REFERENCES

1. H. Kosaka, T. Kawashima, A. Tomita, *et al.*, J. Lightwave Technol. **17**, 2032 (1999).
2. H. Kosaka, T. Kawashima, A. Tomita, *et al.*, Appl. Phys. Lett. **74**, 1212 (1999).
3. M. Bayindir, E. Ozbay, B. Temelkuran, *et al.*, Phys. Rev. B **63**, 81107 (2001).
4. A. Chutinan, M. Mochizuki, M. Imada, *et al.*, Appl. Phys. Lett. **79**, 2690 (2001).
5. S. Rennon, F. Klopf, J. P. Reithmaier, *et al.*, Electron. Lett. **37**, 690 (2001).
6. M. V. Maximov, E. M. Ramushina, V. I. Skopina, *et al.*, Semicond. Sci. Technol. **17**, L69 (2002).
7. C. M. Sotomayor Torres, S. Zankovych, J. Seekamp, *et al.*, Mater. Sci. Eng. **23**, 23 (2003).
8. R. J. Shul and S. J. Pearton, *Handbook of Advanced Plasma Processing Techniques* (Springer, Berlin, 2000).

Translated by A. Sidorova

Technical Physics, Vol. 50, No. 8, 2005, pp. 1048–1052. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 85–89.

SOLID-STATE ELECTRONICS

MIS Structure with Polyamide Insulator under the Conditions of Water Vapor Sorption

E. A. Tutov, E. N. Bormontov, M. N. Pavlenko, G. A. Netesova, and E. E. Tutov

Voronezh State University, Universitetskaya pl. 1, Voronezh, 394693 Russia

e-mail: ssd126@phys.vsu.ru

Received October 26, 2004

Abstract—A silicon MIS structure with sulfonate-containing aromatic polyamide as a gate insulator is prepared with the aim of using it as a capacitive-type humidity detector. A detector made of this material offers a high sensitivity at an elevated relative humidity and a high speed of response. Measurements of the capacitance of the structure at different frequencies basically make it possible to determine the water distribution in the polymer matrix. The high-frequency (1 MHz) capacitance versus relative humidity dependence (i.e., the calibration curve of the detector) is typical of the isotherms of water vapor sorption by similar polymers. © *2005 Pleiades Publishing, Inc.*

INTRODUCTION

Impressive achievements in modern sold-state microelectronics are based on the mature solid-state physics and physics of semiconductors in combination with the progress in technologies of single crystals and related structures, of which Si technology occupies a prominent place.

Along with this main stream, study of novel materials differing from conventional ("ideal") semiconductors and insulators in structure and properties is of ever increasing fundamental and applied importance. These materials possess spatial and energy inhomogeneities (of different type and scale) and feature intricate profiles of dopants and localized states at nanocrystalline, amorphous, homogeneous, (micro)heterogeneous, and heterophase boundaries. Accordingly, their related structures frequently offer a number of unique properties missing from their single-crystal counterparts; hence, investigation in this field with emphasis on applications seems to be topical.

In the majority of microelectronic semiconductor devices, a thin layer of the semiconductor (the near-surface region or the interface between two materials) serves as an active region. The advances in planar technology have made it possible to create metal–insulator–semiconductor (MIS) structures. MIS structures are not only building blocks for many devices but also appropriate objects for gaining deeper insight into the mechanisms of electronic processes at the interfaces and inside of semiconductors and insulators.

Film materials based on aromatic polyamides (including those containing ionogen groups) are being widely used not only in the process of membrane separation [1] but also in hybrid technology as flexible low-permittivity ($\varepsilon \approx 3.5$) substrates and in chemical detectors as selectively permeable and sensitive layers [2].

Development of microelectronic integrated detectors is a promising line of inquiry aimed at producing a new component basis for advanced measuring–information systems. The possibility of using the achievements of Si technology is an important factor in choosing the design of primary transducers, including humidity detectors [3].

In detector structures, polymeric materials are applied as both passive [4] and active elements. In the latter case, the conductivity [5] or capacitance is usually used as a parameter being measured. Capacitive-type MIS detectors may directly measure not only the capacitance but also a shift of the *C–V* characteristic along the voltage axis upon sorption of a gas [6].

The sorptivity of and diffusion in polymers are, as a rule, subjects of independent research [7–9]. The feasibility of analyzing the response of heterostructure sensors to sorption—a promising approach in studying the properties of a material—has been demonstrated with porous silicon in [10, 11]. The high sensitivity of the electrophysical parameters of heterojunctions to external actions and the accessibility of the silicon–polyamide interface for sorbate molecules acting as specific "probes" make it possible to apply heterostructures in chemical (including humidity) detectors, as well as to study the structural–energy characteristics of materials under water vapor sorption and clarify attendant processes.

Since heterogeneous systems often exhibit the mixed electron–ion mechanism of conduction, investigation of both the active and reactive components of the conductivity in a wide range of variable field frequencies is of great importance. However, the measurements

1063-7842/05/5008-1048$26.00

taken in this way are difficult to treat and the results are usually ambiguous; therefore, the method of impedance spectroscopy has not as yet found wide application, specifically, because of its limited applicability.

The problems of impedance spectroscopy of heterogeneous materials stem from the fact that most effects are related not to the intrinsic properties of a composite's components but to the material structure. This extends the applied potential of the material but plagues the study of these effects.

In a number of cases, the high frequency of a measuring signal makes it possible to eliminate a great deal of "slow" processes. Therefore, a major part of the results presented below was obtained by taking high-frequency *C–V* characteristics.

The aim of this paper is to study the feasibility of aromatic polyamides as a sensitive layer in capacitive-type humidity detectors based on MIS structures. In addition, it is of interest to see whether such structures can be used in investigating the sorptivity of and diffusion in polyamides by electrophysical methods.

EXPERIMENTAL

As a gate insulator of the MIS structure, we tested a material from the class of aromatic polyamides (PAs): the copolymer of the sodium salt of 4.4'-diaminodiphenylamine-2-sulfonic acid and *M*-phenylenediamine with a different mole fraction of the α component containing the SO_3Na ionogen group. Figure 1 shows the structural formula of an elementary unit of the PA. Since the mole fraction of sulfogroup-containing α fragments directly influences the amount of water sorbed by the polymer [7], we took the polymer with a maximum value of α, α = 0.41. At α = 0, poly-*M*-phenylene isophthalamide (Phenylone) forms, whose sorption capacity with respect to water vapor is negligible. Therefore, Phenylone was studied as a reference material for developing a method of electrophysical measurements.

We studied the polymer films in the free state (20–80 μm in thickness) and also those prepared by covering a silicon substrate by a solution of the polymer (presynthesized in dimethyl formamide) with subsequent evaporation of the solvent at room temperature (in the latter case, the PA film thickness was several micrometers).

As substrates, we used standard KÉF-4.5 (phosphorus-doped, resistivity 4.5 Ω) single-crystal (100)Si wafers. Native oxide on the substrate surface was not removed, which provided a high quality of the semiconductor–insulator interface. Metallic contacts were made of aluminum (applied by magnetron sputtering) and In–Ga eutectic.

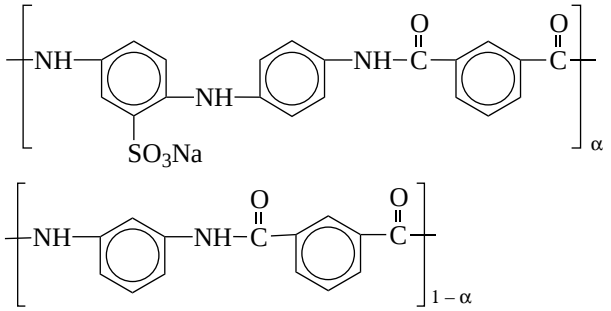


Fig. 1. Structural formula of the polyamide.

As in [11], high-frequency *C–V* characteristics were taken in the static regime using a computerized setup based on an E7-12 immitance meter (the frequency and amplitude of the test signal were 1 MHz and 20 mV, respectively). The frequency dependence of the dielectric losses in the polymer was recorded using a Goodwill 719 LCR meter in the 10 Hz–100 kHz range.

The structure was placed into a sealed measuring cell, which was desiccated with ShSM silica gel. It was assumed that the partial pressure of water vapor corresponds to the zero relative humidity ($p/p_0 = 0$). The desired values of p/p_0 in the range from 0 to 100% were set by the standard method using saturated solutions of various salts [12]. To study the kinetic characteristics of the detector, the structure was rapidly transferred in and out of the cells with $p/p_0 = 0$ into the cell with $p/p_0 =$ 100%. The measurements were carried out at 295 K.

RESULTS AND DISCUSSION

Note at once that the adhesion of the PA film to the $Si(SiO_2)$ surface was good: multiple cycling of the relative humidity had no effect even on the capacitance of the structure.

Low-frequency (1 kHz) measurements showed that the sorption of water vapor in the PA is accompanied by a sharp (by a factor of 1000) increase in the capacitance, which lasts about 1 h. Such a significant change in the capacitance cannot be explained by only the additive contribution of the water sorbed by the polymer. Contributing also are the structure-dependent capacitance of the near-electrode double layer at the PA–metal boundary and PA–Si boundary (the Helmholtz layer [13]), with respect to which the water-containing PA is an electrolyte. The Maxwell–Wagner effect, interlayer polarization due to the charge accumulated at the boundaries of layers with different conductivities when the current passes normally to the layers, is also a factor. In our case, these layers are internal interfaces between the polymeric matrix and sorbate.

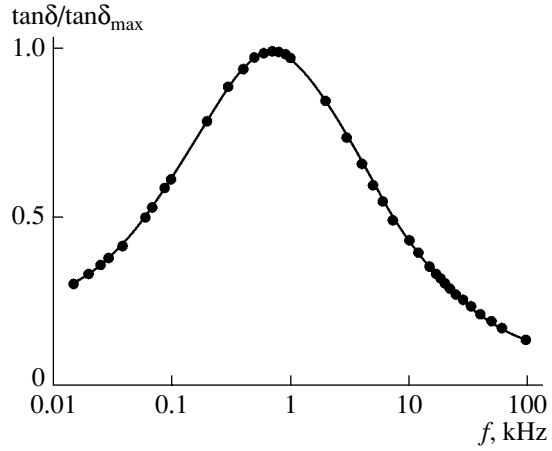


Fig. 2. Frequency dependence of the normalized dielectric loss tangent in the hydrated PA film (at $p/p_0 = 100\%$).

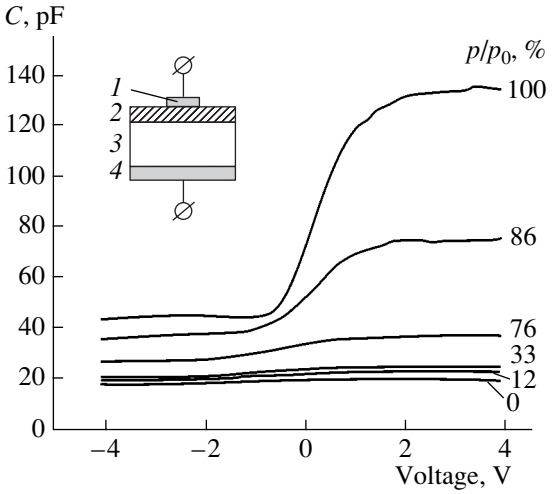


Fig. 3. High-frequency *C–V* characteristics of the Si/PA/Al structure vs. the relative air humidity. The detector structure is schematically shown in the inset. (*1*) Al, (*2*) PA, (*3*) *n*-Si, and (*4*) In–Ga eutectic.

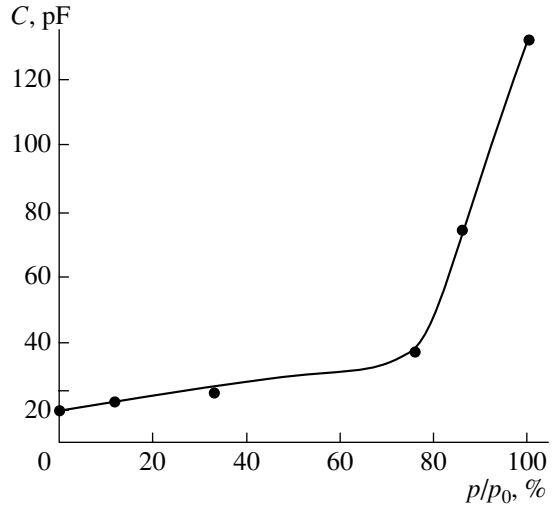


Fig. 4. Capacitance of the insulator vs. the relative humidity.

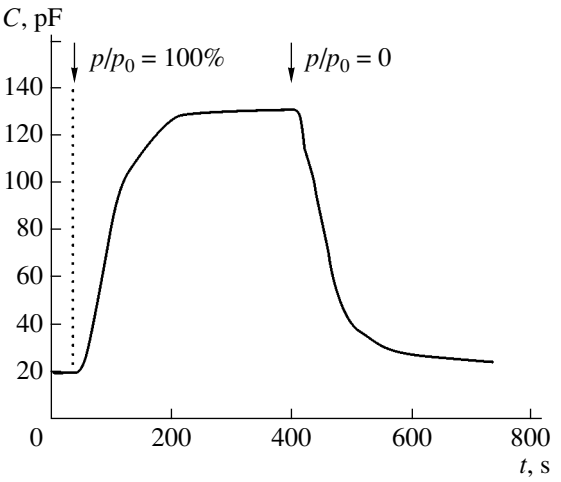


Fig. 5. Water sorption/desorption kinetics in the polyamide.

The shape of the dielectric loss dispersion curve for the hydrated PA films (Fig. 2) is typical of polymers of this kind [14]. The run of the curve, which is the envelope for relaxation oscillators in a wide frequency range and is usually explained by the Maxwell–Wagner effect, shows that, when the electric field frequency exceeds 1 MHz, the contribution from the structural effects decreases and influences the capacitive response of the humidity detector only slightly.

The capacitance of the capacitor structures with the PA insulator that is measured at frequencies below 1 MHz under the conditions of water vapor sorption is characterized by a large amount of the effect; however, the measuring procedure takes too much time: it is comparable to the time taken to establish sorption equilibrium, as estimated from gravimetric data. Moreover, the contribution to the capacitance of the polymer–sorbate composite has, in this case, both additive and structure-dependent components, which are problematic to discriminate.

In order to minimize the role of "slow" effects, we applied the method of high-frequency *C–V* characteristics, which was used earlier [11] to solve a similar problem (water vapor sorption in porous silicon). Note that the structure of porous silicon is simpler than that of polymers, which facilitated quantitative estimation.

The high-frequency *C–V* characteristics taken of the Si–PA–metal structure under the conditions of variable relative humidity (Fig. 3) are typical of MIS structures in which the density of surface states is not too high and the charge built in the insulator is minor. At positive biases, the capacitance of the structure depends on the permittivity of the polymer layer and its growth (with partial pressure of water vapor) reflects an increase in

the water concentration in the PA film. The dependence of the capacitance on the relative humidity, i.e., the calibration curve of the detector, is shown in Fig. 4. Its shape is in qualitative correspondence with the isotherm of water vapor sorption by similar polymers [14] (the Henry isotherm). This isotherm features a sharp increase when the relative humidity exceeds 80%.

At a frequency of 1 MHz, the capacitance of the structure under water sorption/desorption (Fig. 5) varies over much wider limits compared with lower frequencies and the time it takes for the capacitance to reach a steady-state value is about 2 min. Once the detector has been transferred from the cell with $p/p_0 = 0$ to the cell with $p/p_0 = 100\%$, water sorption naturally lasts more than 2 min. However, after a lapse of 2 min, the sorption, as well as the redistribution, if any, of the sorbed water in the PA layer, does not affect the capacitance. This fact, together with the frequency dependence of the dielectric losses (Fig. 2), allows us to argue that, at a frequency of 1 MHz, one can measure the contribution from the "free" water sorbed by the PA film [14], i.e., from the liquid water.

The free water in the polymeric matrix can be characterized by its thermodynamic parameters, in particular, relative permittivity ε. Nevertheless, the increase in the capacitance of the PA-containing structure by more than six times as the relative humidity varies from 0 to 100% is too high. If it is taken into account that the volume fraction of water in similar materials does not exceed 10% [14], such an increase cannot be predicted by the known estimators of the permittivity of composites (for example, by the Lichtenecker formulas applied to the structure with porous silicon [11]).

We believe that the observed increase in the capacitance upon polymer hydration is associated with an increase in the conductivity of the polymer and with a spread of the charge over the surface (hence, with an increase in the effective surface area of the metallic electrode, which seems to be impossible to estimate). To exclude the effect assumed, we performed measurements on the structure where the electrodes occupied the entire surface area. In this case, the maximal change in the capacitance was 30%, which corresponds to a volume fraction of the water phase in the PA matrix of about 4%, according to our estimates. This is in agreement with the published data for moisture absorption in polyamides.

In such a geometry of the experiment, a significant advantage of the electrophysical method of investigating water vapor sorption by polymers (high sensitivity) is almost lost. However, in this case, the time taken to establish the steady-state value of the capacitance is specified by the diffusion of water molecules through the PA film under the metallic electrode. Therefore, there appears a possibility to study diffusion in polymers by a simple convenient method.

CONCLUSIONS

Aromatic polyamides containing sulfogroups can be used as a sensitive layer in capacitive-type humidity metal–insulator–silicon detectors offering a high sensitivity and speed of response. Although water molecules can form hydrogen bonds with fragments of –C=O amide groups [7], the increase in the capacitance was observed only in the structures where the polymer film contains sulfonate groups.

Sorption/desorption of water vapor in the polymer is unrelated to its charge state, as indicated by the *C–V* characteristics of the structure. In water solutions of electrolytes, ion exchange reactions (the primary application of this material) change the charge of the polyamide film and so are bound to cause a characteristic shift of the *C–V* characteristics along the voltage axis. Our tentative experiments suggest that the MIS structure considered is promising for *p*Na-sensitive chemical detectors operating without a reference electrode.

The capacitance measurements performed in this work were complemented by the dependences of the conductivity of the PA-containing structure on the relative humidity that were recorded in a field varying with the same frequencies (to exclude polarization of the insulator). At 1 MHz, these dependences are in qualitative agreement with the capacitance measurements presented, showing even a better sensitivity. However, to establish a quantitative correlation between the conductivity of the PA film and the water concentration in it from the conductivity data is still a greater challenge than between the capacitance and water concentration from the capacitance measurements.

The fact that a polymer film can serve not only as a sensor of sorbate molecules but also as a transducer, the possibility of frequency separation of signals from variously associated sorbates, and easy control of the filler content in the polymer matrix make a Si/PA/metal structure promising not only for humidity detectors: it can be used to advantage in investigating sorption and diffusion in polymeric materials.

REFERENCES

1. S. F. Timashev, *Physical Chemistry of Membrane Processes* (Khimiya, Moscow, 1988) [in Russian].
2. G. Asch, *et al.*, *Le Capteur en Instrumentation Industrielle* (Dunod, 1991; Mir, Moscow, 1992), Vol. 2 [translated from French].
3. B. I. Podlepetskii and A. V. Simakov, Zarubezhn. Élektron. Tekh., No. 2, 64 (1987).
4. P. V. Yakovlev, A. V. Shaposhnik, V. S. Voishchev, *et al.*, Zh. Anal. Khim. **57**, 326 (2002).
5. A. K. Mamedov, Sensors, No. 4, 16 (2002).
6. A. A. Vasil'ev, E. K. Lyutikova, V. Morits, and V. I. Filippov, Sensors, No. 4, 8 (2002).

7. V. V. Valuev, O. Yu. Zemlyanova, N. V. Semina, *et al.*, Zh. Fiz. Khim. **68**, 1667 (1994).
8. A. T. Ponomarenko, V. G. Shevchenko, N. G. Ryvkina, *et al.*, Kondensirovannye Sredy i Mezhfaznye Granitsy **3** (1), 73 (2001).
9. V. V. Kotov, S. A. Sokolova, G. A. Netesova, and I. V. Kuznetsova, Zh. Fiz. Khim. **78**, 1678 (2004).
10. E. A. Tutov, A. Yu. Andryukov, and E. N. Bormontov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **35**, 850 (2001) [Semiconductors **35**, 816 (2001)].
11. E. A. Tutov, E. N. Bormontov, V. M. Kashkarov, *et al.*, Zh. Tekh. Fiz. **73** (11), 83 (2003) [Tech. Phys. **48**, 1442 (2003)].
12. G. Wiegleb, *Sensortechnik* (Franzis, Munchen, 1986; Mir, Moscow, 1990).
13. A. I. Kulak, *Electrochemistry of Semiconductor Heterostructures* (Universitetskoe, Minsk, 1986) [in Russian].
14. S. M. Rowland, *Water in Polymers* (American Chemical Society, Washington, 1980; Mir, Moscow, 1984).

Translated by Yu. Vishnyakov

Technical Physics, Vol. 50, No. 8, 2005, pp. 1053–1057. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 90–94.

OPTICS, QUANTUM ELECTRONICS

Simulation of Experiment on Total External Reflection of Electron Bremsstrahlung

F. K. Aliev, G. R. Alimov, A. T. Muminov, B. S. Osmanov, and V. V. Skvortsov

Research Institute of Applied Physics, National University of Uzbekistan, Tashkent, 700174 Uzbekistan

e-mail: gleb@iaph.tkt.uz

Received July 19, 2004

Abstract—Computer simulation of an experiment on scattering of a strongly collimated beam of 13-MeV bremsstrahlung electrons from a macroscopically smooth surface through ultrasmall angles is carried out by the Monte Carlo method taking into account the interaction of the beam with an extended air medium. A comparison of the results of calculations and experimental data indicates the possibility of total external reflection of γ quanta in the energy range ≤1 MeV. © *2005 Pleiades Publishing, Inc.*

INTRODUCTION

Total external reflection (TER) [1], which has been used in various X-ray systems [2], is observed when soft X rays ($E \le 10$ keV) are incident on the surface (interface between two media) with a roughness (whose height does not exceed a quarter of radiation wavelength λ) at an angle ψ smaller than the critical angle

$$\alpha_{cr} = (eh/E)(ZN_A\rho/\pi Am)^{1/2},$$

where e and m are the electron mass and charge; Z, A, and ρ are the charge, atomic mass, and density of the substance of the reflecting surface; h is the Planck constant; and N_A is the Avogadro number. The interest in small-angle scattering of hard γ quanta is stimulated to a considerable extent by the possibility of designing such systems in the range of harder electromagnetic radiation.

In contrast to soft X rays, the wavelength of hard X rays is comparable to the mean atomic spacing d in solids and liquids (in the γ range, it is even 2 or 3 orders of magnitude smaller than this quantity); any macroscopically smooth surface (consisting of individual scattering centers) is clearly discrete for this radiation. It was shown in [3–5] that the interaction of radiation with $\lambda \ll d$ with such a surface results in elastic as well as inelastic scattering of quanta. For any wavelength, there exists the so-called coherence cone θ_{coh} within which the coherent scattering amplitude is virtually constant, but rapidly decreases with increasing angle outside this cone. With increasing energy of photons, the angular size $\theta_{koh} \le 10^{-2}\lambda$ of the cone decreases together with the TER critical angle $\alpha_{cr} \sim \theta_{coh}/10$. The results of calculation of the TER critical angle [3] completely coincide with the corresponding expression obtained from the classical theory of X-ray dispersion, which was developed initially for soft X rays and repeatedly confirmed experimentally.

As regards the experimental verification of TER in the γ range, we are aware of only one report on its observation in experiments on scattering of 122-keV γ radiation from ^{57}Co on a 10-m path length [6]. Obviously, more convincing results can be obtained only for a much longer path length and a more powerful source of directional γ radiation. We began such experiments with an electron bremsstrahlung (EB) beam from the MT-22S microtron mounted in a bunker adjoining a 300-m-long transit channel (2.5 × 2.5 m in cross section) [7, 8]. Here, we report on the results of computer simulation of these experiments and compare them with individual preliminary experimental data.

EXPERIMENTAL SETUP

The experimental setup is shown in Fig. 1. An accelerated electron beam (E_e = 13 MeV, average current $I_a \le$ 30 μA, electron pulse duration τ = 2.5 μs, pulse repeti-

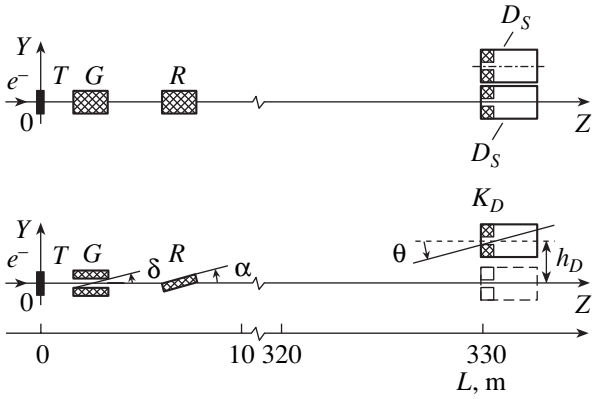


Fig. 1. Experimental setup.

1063-7842/05/5008-1053$26.00 © 2005 Pleiades Publishing, Inc.

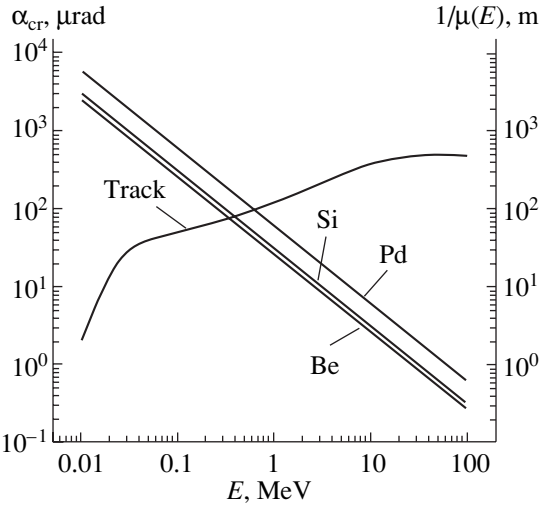


Fig. 2. Critical angles of TER for γ quanta, calculated by formula (1). The mean free path of γ quanta in air is taken from [10].

tion rate f = 386 Hz, beam diameter ∅ ~ 5 mm) extracted from a microtron is directed to a decelerating target T (horizontal stretched tungsten wire 1 mm in diameter). Electron bremsstrahlung generated in the target (with a natural divergence of ~35 mrad and a solid angle $\Omega \sim 3 \times 10^{-4}$ sr) is formed only along the vertical with the help of a slit collimator G (slit height $\Delta X_G \geq 10$ μm and vertical aperture $\delta \geq 15$ μrad). The lower part ($\theta < 0$) of the formed EB beam is absorbed at the end of reflector R (glass; ΔZ_R = 600 mm; ΔY_R = 300 mm; ΔX_R = 30 mm; surface finish of the 14th grade; and the grazing angle, i.e., the inclination of the surface to the beam axis, is controlled to within α = 0–1.5 mrad with a step $\Delta\alpha \geq 15$ μrad), while the upper part is scattered from the surface at $\theta = 0$–α or passes above it without interacting for $\theta > \alpha$. At a level of 330 m, γ quanta of the beam are detected by a fast counting D_f (Y = 100 mm) and a slow spectrometric D_S (Y_S = 0) detector. The detectors are rigidly fixed to each other and are displaced along the vertical with a step $h_D \geq$ 1 mm; the detection angle (formed by the beam axis OZ and TD line) may change in the limits $\theta = 0 \pm 30$ mrad ($\Delta\theta \geq 3$ μrad). Horizontal slit collimators K_D of the detector have a gap of height ΔX_K = 1 mm, which corresponds to the solid angle of detection $\Omega_D = 4 \times 10^{-11}$ sr; depending on the average current of extracted electrons, from a few to a hundred γ quanta pass through them during the acceleration cycle τ = 2.5 μs.

The error in adjustment of the test bench elements (decelerating target T, slit collimator G, reflector R, and detector D_S) relative to the EB beam axis is ~10 μrad. To reduce the background contribution associated with scattered γ quanta and charged particles to the events registered by the detectors, the primary (after target T), formed (after slit collimator G), and reflected (after reflector R) EB beam is cleaned by auxiliary collimators and permanent magnets (not shown on the diagram), respectively, and detection is blocked over a time interval of 2.5 μs between the end of an acceleration cycle and the beginning of the next cycle.

The experiment includes the measurement by detectors D_S and D_f of the spectra and intensity of the EB beam of γ quanta, respectively, as functions of the height ΔX_G of the gap of the slit collimator, angle of inclination α of reflector R, and detection angle θ.

The geometrical parameters of the test bench elements correspond to the conditions of observation of the expected effect of TER of hard X rays in the energy range $E \leq 13$ MeV; however, in the range of $E \leq 4$ MeV, the effect is slightly suppressed by the interaction of γ quanta with atmospheric air (Fig. 2). As regards the reflectivity of the surface of the reflector used in the test bench, the calculations and experiments on reflection of 122-keV γ rays emitted by ^{57}Co [6] from an analogous reflector (glass with a surface polished to the 14th grade with a roughness height of ~500 Å) showed that the reflection coefficient of γ quanta is close to unity for angles $\alpha < 264$ μrad.

For a fixed grazing angle α of the reflector, γ quanta from the low-energy part of the EB spectrum, for which $\psi = 0$–α_{cr}, must obviously experience TER, while this effect is not observed in the high-energy range, where $\psi > \alpha_{cr}$.

In spite of TER blurring due to edge effects at the collimators and reflector, interaction of γ quanta with the 300-m air medium, as well as the finite size of the decelerating target and gaps in the slit collimators, the high angular resolution of the test bench may ensure the observation of the TER in experiment.

SIMULATION OF EXPERIMENT

Since the problem being solved here is quite complicated, we developed a statistical model of the experiment using the Monte Carlo method to estimate the effect of the above-mentioned factors on the expected result. The computer model of the experiment was worked out with the help of the Mathematica software on computers of the Intel Pentium III and higher class. In this program, the energy of bremsstrahlung γ quanta is assumed to be distributed in accordance with the well-known Schiff formula [9]. The positions of γ quanta in the plane of the target are distributed in the horizontal and vertical projections in accordance with the normal law with the same standard deviation equal to 1.5 mm. The smallness of the solid angle of the γ quanta used in the experiment makes it possible to assume, to a high degree of accuracy, that their direction distribution is equiprobable and to simulate it with the help of two random quantities. Accordingly, the correlation between the angular and energy distributions of bremsstrahlung γ quanta was ignored.

The program approximately presumes that a γ quantum vanishes when its stimulated trajectory intersects with any material objects of the experimental setup (except the reflector and air). The effect of air on each γ quantum is simulated probabilistically. For this purpose, tabular data are used for the energy dependence of the total absorption coefficient of γ quanta in air [10]. We assumed that any interaction of a γ quantum with air leads to its annihilation. The case with the reflector is considered as follows. When the trajectory of the γ quantum intersects with the reflector, we first determine its grazing angle relative to the reflector surface and then compare it with the critical angle of scattering α_{cr} calculated by formula (1) for the preset energy E of the γ quantum. If this angle turns out to be larger than the critical angle, the γ quantum is assumed to disappear as a result of absorption and the program begins a new testing (i.e., the parameters of a new γ quantum are generated in the target). If, however, this angle is smaller than the critical angle, the program replaces the parameters of the trajectory of the γ quantum incident on the reflector by the parameters of the specularly reflected trajectory. The trajectory of the γ quantum behind the reflector is traced to the end of the channel, where the position of the detector is determined as well as the energy with which the γ quantum is detected. The program assumes 100% efficiency of its detection.

Using this program, we analyzed the effect of various values of the path length, the gap in the slit collimator, and the angle of inclination of the reflector on the energy distribution and vertical angular distribution of γ quanta.

RESULTS OF SIMULATION AND DISCUSSION

The following circumstances should be borne in mind while analyzing the results of simulation presented below.

The slit height of the collimator is set by three values (50, 20, and 10 μm), and the BE beam axis coincides with the center of the slit. The front edge of the reflecting surface is made to coincide with the lower plane of the slit collimator along the vertical; the distance from the edge to the slit exit along the horizontal was 30 cm, as in the experiment. The calculations were made for values of 20 and 40 μrad of the angle of inclination of the reflector to the beam axis. These values correspond to calculated values of the TER critical angle for γ quanta energies of 1.55 and 0.78 MeV, respectively. The BE beam incident in the reflector always contains a γ quanta satisfying and not satisfying the TER conditions.

We assumed that the energy of electrons incident on the target was 13 MeV, which corresponded to the first working mode of the MT-22S microtron. The simulated energy of bremsstrahlung quanta was bounded by values of 0.01 and 12.5 MeV from below and from above.

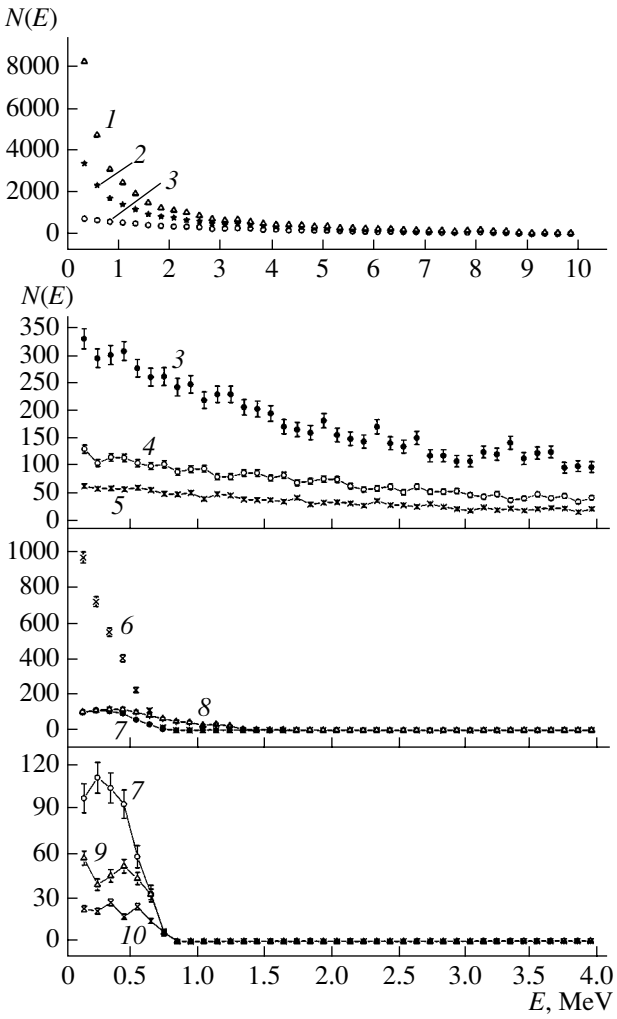


Fig. 3. Spectra of BE beam. Without reflector: (*1*) δ = 62 μrad, in vacuum, (*2*) δ = 62 μrad, L = 120 m; (*3*) δ = 62 μrad, L = 330 m; (*4*) δ = 25 μrad, L = 330 m; (5) δ = 12 μrad, L = 330 m. With reflector: (*6*) δ = 62 μrad, α = 40 μrad, L = 120 m; (*7*) δ = 62 μrad, α = 40 μrad, L = 330 m; (*8*) δ = 62 μrad, α = 20 μrad, L = 330 m; (*9*) δ = 25 μrad, α = 40 μrad, L = 330 m; (*10*) δ = 12 μrad, α = 40 μrad, L = 330 m.

To compare the data obtained for different simulation characteristics (corresponding to various experimental conditions), the data were recalculated taking into account the normalization factors obtained for a collimator slit height of 50 μm. The number of BE γ quantum formed by such a slit and sampled by the program exceeded 50 000. Figure 3 illustrates the transformation of the energy spectrum of γ quanta in a BE beam (integrated over all virtual positions of detector D_S) as a function of path length L, angular aperture δ of the slit collimator, and the angle of inclination α of the reflector; Fig. 4 shows the same dependences for the vertical angular distribution of the intensity of γ quanta with energy E < 12.5 MeV. The spectra of BE γ quanta scattered by the reflector through angles θ = 3, 40, and

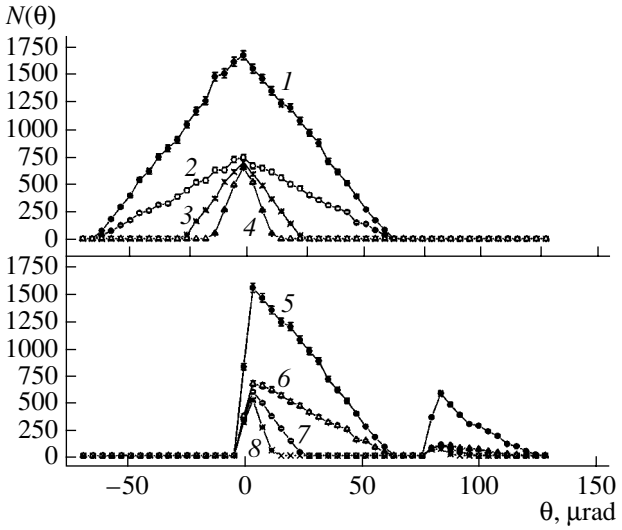


Fig. 4. Vertical distribution of the BE beam intensity. Without reflector: (*1*) $\delta = 62$ μrad, $L = 120$ m, (*2*) $\delta = 62$ μrad, $L = 330$ m; (*3*) $\delta = 25$ μrad, $L = 330$ m; (*4*) $\delta = 10$ μrad, $L = 330$ m. With reflector: (*5*) $\delta = 62$ μrad, $\alpha = 40$ μrad, $L = 120$ m; (*6*) $\delta = 62$ μrad, $\alpha = 40$ μrad, $L = 330$ m; (*7*) $\delta = 62$ μrad, $\alpha = 40$ μrad, $L = 330$ m; (*8*) $\delta = 12$ μrad, $\alpha = 40$ μrad, $L = 330$ m.

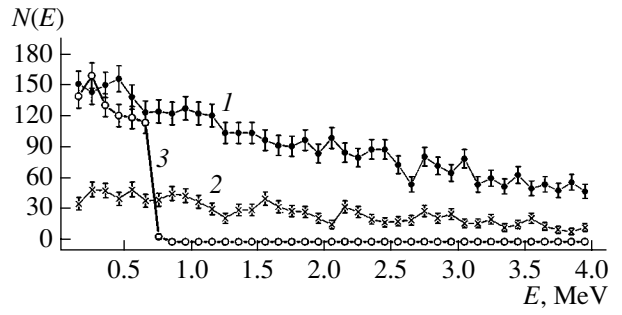


Fig. 5. Spectra of the BE beam for $\delta_G = 62$ μrad, $\alpha = 40$ μrad, $\theta = 3$ (*1*), 40 (*2*), and 80 μrad (*3*).

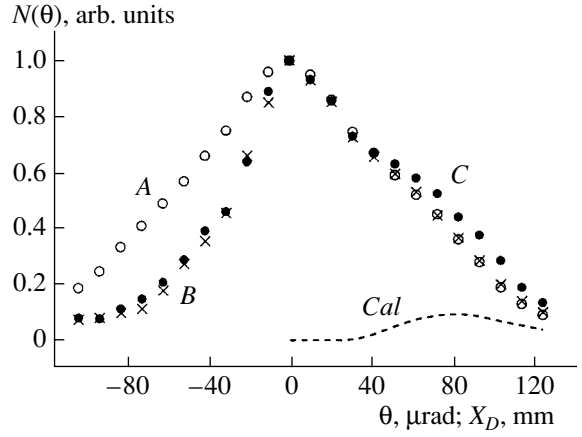


Fig. 6. Vertical distribution of the BE beam intensity for $\delta_G = 62$ μrad, $L = 330$ m without reflector (*A*) and with reflector (*2–4*) for $\alpha = 3$ (*B*) and 40 μrad (*C*); the calculated distribution with $\alpha = 40$ μrad (*Cal*).

80 μrad for fixed angles $\delta = 62$ μrad and $\alpha = 40$ μrad are shown in Fig. 5. Analysis of these data provided a quantitative estimate of the effect of experimental conditions on the intensity as well as the energy and angular distribution of detected γ quanta of the BE beam: (a) attenuation of BE γ quantum beam intensity in air increases upon a decrease in the γ quantum energy and reaches an order of magnitude for path length $L = 330$ m in the case of γ quanta with $E \le 4$ MeV (curves *1* and *3* in Fig. 3); (b) the intensity of the BE beam formed in experiment is proportional to the angular aperture δ of the slit collimator (curves *3–5* in Fig. 3 and *2–4* in Fig. 4); (c) with increasing angle of inclination α of the reflector introduced into the formed BE beam, the transformation of the spectrum to the low-energy region increases (curves *7* and *8* in Fig. 3) and the displacement of the vertical distribution to the range of larger angles θ are enhanced; in the region of $\theta \ge 2\alpha$, a TER component of γ quanta with an energy corresponding to critical angle $\alpha_{cr} \le \alpha$ appears (curves *4–6* in Fig. 4 and *1–3* in Fig. 5).

For a path length $L = 120$ m, the expected TER effect is manifested much more clearly; in this case, the angular resolution of the test bench is worse to the same extent as in the case of $L = 330$ m ($\Delta\theta/\Delta h_D = 8.3$ μrad/mm for $L = 120$ m and $\Delta\theta/\Delta h_D = 3$ μrad/mm for $L = 330$ m). Considering that the detection of the effect requires primarily a high angular resolution, we used path length $L = 330$ m in our experiments. Some of the preliminary results can be compared with the results of simulation. Figure 6 shows the experimental intensity distributions of γ quanta in the BE beam formed with a vertical aperture $\delta = 62$ μrad and measured without a reflector (*A*) and with a reflector inclined at angles $\alpha = 3$ μrad (*B*) and 40 μrad (*C*). It should be expected that the shape of distribution *A* in the angular interval $\theta = 0 \pm \delta$ must be close to normal, while distributions *B* and *C* must have components due to γ quanta that have not interacted with the reflector ($\psi = \alpha - \delta$) as well as due to γ quanta that have interacted with it ($\psi = 0 - \alpha$), since $\delta > \alpha$. If the TER effect is observed for γ quanta, the latter components must be manifested in the form of TER peaks for angles $\theta \ge 2\alpha$. This peak should not be resolved from the component due to noninteracting radiation in distribution *B* (since $2\alpha < \delta$) and should be resolved in distribution *C* (since $2\alpha > \delta$). These expectations with corrections for the finite size of decelerating target *T* and collimators *G* and K_D hold in the calculated distribution *C*'. As regards the experimental distribution *C*, it is strongly blurred by other disregarded factors (like beam aperture broadening due to the finite absorption coefficient of the material of collimators *G* and *K*, scattering in the air medium, and instability of the electron beam). Nevertheless, distribution *C* exhibits bulging in the region of $\theta \sim 2\alpha$. If the right slope of distribution *A* or *B* is taken as the component due to noninteracting radiation in distribution *C*, the difference between *C* and *A* or *B* in a

first approximation corresponds to the calculated TER peak of γ radiation (distribution *6* in Fig. 4).

Naturally, this fact should not be treated as a direct manifestation of the TER effect in the γ range; it is rather an indication of the existence of this effect and the possibility of its more exact resolution under improved experimental conditions (decrease in the vertical size of the formed BE beam and the noise level, use of remote control of the reflector tilting and position-sensitive detectors, increase in the temporal stability of the BE beam, shortening of the experimental time, etc.).

The comparison of the experimental and calculated distributions necessitates the development of a more perfect computational algorithm taking into account edge effects in the slit collimator and reflector and Compton scattering in air, as well as a more realistic dependence of the reflection coefficient of BE γ quanta on the tilt angle of the reflector.

REFERENCES

1. M. A. Blokhin, *Physics of X Rays* (GITTL, Moscow, 1957) [in Russian].
2. M. A. Kumakhov, Proc. SPIE **4765**, 20 (2002).
3. V. A. Arkad'ev, M. A. Kumakhov, and L. I. Ognev, Pis'ma Zh. Tekh. Fiz. **12** (21), 1307 (1986) [Sov. Tech. Phys. Lett. **12**, 540 (1986)].
4. J. H. Hubbel and I. Overbo, J. Phys. Chem. Ref. Data **8**, 69 (1979).
5. L. Kissel, R. H. Pratt, and S. C. Roy, Phys. Rev. A **22**, 1970 (1980).
6. V. I. Glebov, E. I. Denisov, N. K. Zhevago, *et al.*, Zh. Tekh. Fiz. **55**, 1785 (1985) [Sov. Phys. Tech. Phys. **30**, 1043 (1985)].
7. M. A. Kumarhov, T. M. Muminov, U. S. Salikhbaev, *et al.*, Proc. SPIE **4155**, 48 (2000).
8. F. K. Aliev, G. R. Alimov, A. K. Ikromov, *et al.*, Proc. SPIE **4765**, 145 (2002).
9. L. I. Schiff, Phys. Rev. **83**, 252 (1951).
10. S. V. Starodubtsev and A. M. Romanov, *Interaction of Gammas with Matter* (Nauka, Tashkent, 1964), Chap. 1 [in Russian].

Translated by N. Wadhwa

Technical Physics, Vol. 50, No. 8, 2005, pp. 1058–1061. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 95–98.

RADIOPHYSICS

Transition Radiation of Surface Electromagnetic Waves by Electron Bunches in a Cylindrical Waveguide

Yu. O. Averkov

Usikov Institute of Radiophysics and Electronics, National Academy of Sciences of Ukraine, ul. Akademika Proskura 12, Kharkov, 61085 Ukraine

e-mail: yuaver@online.kharkiv.com

Received May 7, 2004

Abstract—Transition surface electromagnetic radiation from electron bunches that cross a metal–insulator–semiconductor structure in a perfectly conducting semi-infinite cylindrical waveguide is studied. It is shown that, using a periodic sequence (train) of electron bunches, a particular surface waveguide eigenmode can be amplified by bringing its frequency to resonance with the bunch repetition frequency in the train. Those eigenmodes are amplified most efficiently whose frequency falls into the range occupied by the first maximum of the geometric factor of one bunch. © *2005 Pleiades Publishing, Inc.*

INTRODUCTION

It is well known that transition radiation, as well as Cherenkov radiation, is convenient to apply in diagnostics of pico- and femtosecond electron bunches and electrical performance of semiconductors. At the same time, together with the conventional approaches, such as the method of frustrated total internal reflection, method of diffraction gratings, and others [1], surface waves induced by the transition radiation from charges can be used for gaining insight into the structure and electrical properties of the surface. To this end, the energy of surface waves must be high enough for them to be detected in experiment (from 1 to 10 eV). In this paper, it is suggested that these waves be generated using coherent effects associated with the transition radiation from a periodic sequence (train) of electron bunches.

The properties of the transition radiation from one electron crossing the interface between two media have been studied at length in [2]. It is also well known that, due to coherent effects, the transition radiation produced by an extended bunch of charges may differ significantly from the transition radiation from one electron [3, 4]. In particular, the transition radiation from a charged bunch may produce wide-band electromagnetic pulses whose waveform copies that of the pulsed current of the bunch. The bunch kinetic energy-to-electromagnetic pulse energy conversion efficiency may be as high as several tens of percent.

The works cited above addressed the transition radiation of volume electromagnetic waves. It is also of interest to see how the coherent radiation from bunches influences the transition radiation of surface electromagnetic waves. The transition radiation of surface electromagnetic waves from a single nonrelativistic electron bunch crossing a vacuum–semiconductor interface has been studied in [5]. However, as was noted above, for these surface waves to be used in practice, their energy must be sufficiently high. Therefore, it is necessary to use a train of electron bunches. We will show that, by appropriately choosing the size and repetition period of bunches, the energy of a harmonic being amplified can be made proportional to the square of the product of the number of electrons in a bunch by the number of bunches in a train. Consequently, not only the radiation from all electrons in each bunch but also the radiation from all bunches taken together will be coherent. In this work, the system under study is a waveguide, which has a limited transverse dimension—a feature typical of real systems.

FORMULATION OF THE PROBLEM AND BASIC EQUATIONS

Consider a perfectly conducting semi-infinite cylindrical waveguide closed by a perfectly conducting plate at one end. The generatrix of the waveguide is parallel to the z axis. The origin of the z axis is on the inner surface of the end plate. The waveguide region $0 < z < d$ is filled with a dielectric of permittivity ε_d; the region $z > d$, with a semiconductor of permittivity ε_s. Monoenergetic electron bunches in the form of ellipsoids of revolution with their axes aligned with the waveguide axis are injected into the waveguide along its axis through an opening in the end plate. The transition radiation from the bunches induces surface and volume waves on the dielectric–semiconductor interface inside the waveguide. The surface (standing) waves are localized near the interface and have a discrete spectrum.

1063-7842/05/5008-1058$26.00 © 2005 Pleiades Publishing, Inc.

The charge density of the bunch is specified as follows:

$$q(\mathbf{r}, t) = \frac{en_\mathrm{b}V_\mathrm{b}}{2\pi^2R^2}\sum_{k=1}^{\infty}\frac{J_0(\kappa_k\rho)}{J_1^2(\mu_k^{(0)})} \times \int_{-\infty}^{\infty} dk_z f_k(\omega)\exp[i(k_z z - \omega t)], \quad (1)$$

where $J_k(x)$ is the Bessel function of order k, $\mu_k^{(0)}$ is a kth root of the Bessel function of order 0, $\kappa_k = \mu_k^{(0)}/R$ is the wavenumber of the surface wave in the plane of the interface, R is the waveguide radius, ρ is the magnitude of the radius vector in the waveguide's cross section, $k_z = \omega/v$, v is the bunch velocity, n_b is the electron density in a bunch, $V_\mathrm{b} = 4\pi a^2 b/3$ is the bunch volume, and

$$f_k(\omega) = \frac{1}{V_\mathrm{b}}\int_{(V_\mathrm{b})} J_0(\kappa_k\sqrt{x_0^2 + y_0^2}) \times \exp(-ik_z z_0)dx_0 dy_0 dz_0 \quad (2)$$

is the geometric factor of a bunch.

In (2), x_0, y_0, and z_0 are the coordinates of an electron in a bunch in the co-moving frame of reference.

Integration over the ellipsoid in formula (2) yields

$$f_k(\omega) = \frac{3}{\psi_k(\omega)^2}\left\{\frac{\sin[\psi_k(\omega)]}{\psi_k(\omega)} - \cos[\psi_k(\omega)]\right\}, \quad (3)$$

where $\psi_k(\omega) = \sqrt{(k_z b)^2 + (\kappa_k a)^2}$.

Next, we expand the electromagnetic fields of the bunch and radiation in the Fourier–Bessel series, apply the method developed in [2], and use the boundary conditions to obtain the following expressions for the surface wave fields in the dielectric:

$$E_{\rho,\mathrm{d}}^{\mathrm{sw},k}(\mathbf{r}, t) = \frac{\pi}{\beta}\frac{\lambda_{\mathrm{d}k}\lambda_{sk}}{\omega_k}\frac{\Delta(\omega_k, \kappa_k)}{\left(\frac{\partial\Delta_0(\omega_k, \kappa_k)}{\partial\omega_k}\right)} \times \sin(\lambda_{\mathrm{d}k}z)J_1(\kappa_k\rho)\exp(-i\omega_k t), \quad (4)$$

$$E_{z,\mathrm{d}}^{\mathrm{sw},k}(\mathbf{r}, t) = \frac{\pi}{\beta}\frac{\kappa_k\lambda_{sk}}{\omega_k}\frac{\Delta(\omega_k, \kappa_k)}{\left(\frac{\partial\Delta_0(\omega_k, \kappa_k)}{\partial\omega_k}\right)} \times \cos(\lambda_{\mathrm{d}k}z)J_0(\kappa_k\rho)\exp(-i\omega_k t), \quad (5)$$

$$H_{\varphi,\mathrm{d}}^{\mathrm{sw},k}(\mathbf{r}, t) = -\frac{i\pi\varepsilon_\mathrm{d}}{v}\lambda_{sk}\frac{\Delta(\omega_k, \kappa_k)}{\left(\frac{\partial\Delta_0(\omega_k, \kappa_k)}{\partial\omega_k}\right)} \times \cos(\lambda_{\mathrm{d}k}z)J_1(\kappa_k\rho)\exp(-i\omega_k t), \quad (6)$$

where

$$\Delta(\omega, \kappa_k) = \beta\left\{[\varepsilon_\mathrm{s}\tilde{E}_{\rho,\mathrm{s}}^{\mathrm{b}}(\omega, \kappa_k, d) - \varepsilon_\mathrm{d}\tilde{E}_{\rho,\mathrm{d}}^{\mathrm{b}}(\omega, \kappa_k, d)] - \frac{\kappa_z}{\lambda_{sk}}\varepsilon_\mathrm{s}[\tilde{E}_{\rho,\mathrm{s}}^{\mathrm{b}}(\omega, \kappa_k, d) - \tilde{E}_{\rho,\mathrm{d}}^{\mathrm{b}}(\omega, \kappa_k, d)]\right\} - \frac{\omega}{c}\left(\frac{\varepsilon_\mathrm{d}}{\lambda_{\mathrm{d}k}} + \frac{\varepsilon_\mathrm{s}}{\lambda_{sk}}\right)\tilde{E}_{\rho,\mathrm{d}}^{\mathrm{b}}(\omega, \kappa_k, 0)\exp(-i\lambda_{\mathrm{d}k}d), \quad (7)$$

$$\tilde{E}_{\rho,l}^{\mathrm{b}}(\omega, \kappa_k, z) = -\frac{2en_\mathrm{b}V_\mathrm{b}}{\pi R^2}\frac{\kappa_k}{J_1^2(\mu_k^{(0)})} \times \frac{f(\omega)}{\varepsilon_l(\lambda_{lk}^2 - k_z^2)}\exp(ik_z z), \quad (8)$$

$$\Delta_0(\omega, \kappa_k) = \varepsilon_\mathrm{d}\lambda_{sk}\cos(\lambda_{\mathrm{d}k}d) - i\varepsilon_\mathrm{s}\lambda_{\mathrm{d}k}\sin(\lambda_{\mathrm{d}k}d). \quad (9)$$

Here, $\beta = v/c$, c is the velocity of light in free space, and $\lambda_{lk}^2 = \omega_k^2\varepsilon_l/c^2 - \kappa_k^2$ is the axial wavenumber squared. Subscript l = d for the dielectric and s for the semiconductor, and frequency ω_k is a solution to the dispersion relation $\Delta_0(\omega, \kappa_k) = 0$ for the surface waves for κ_k found from the condition $J_0(\kappa_k R) = 0$.

For the surface wave fields to decay from the dielectric–semiconductor interface into either medium, it is necessary that $\mathrm{Im}\{\lambda_{sk}\} = \lambda_{sk}^* > 0$ and $\mathrm{Im}\{\lambda_{\mathrm{d}k}\} = \lambda_{\mathrm{d}k}^* < 0$.

Substituting formulas (4)–(6) into the standard expression for the energy density of a kth harmonic of the radiated field [2] and integrating the resulting expression over the region occupied by the dielectric, we obtain an expression for the energy radiated in the frequency range $\omega_k \ll \Omega_\mathrm{sp}$,

$$S_k^{\mathrm{sw}} \approx \beta^2\frac{e^2N_\mathrm{b}^2}{d\mu_k^{(0)2}J_1^2(\mu_k^{(0)})}f(\omega_k)^2\sin\left(\frac{\omega_k d}{2v}\right)^2, \quad (10)$$

where Ω_sp is the surface plasmon frequency at which $\varepsilon_\mathrm{s} = -\varepsilon_\mathrm{d}$ with $|\lambda_{\mathrm{d}k}|d \ll 1$ and $N_\mathrm{b} = n_\mathrm{b}V_\mathrm{b}$ is the number of electrons in the bunch.

When deriving formula (10), we assumed that $\varepsilon_\mathrm{s} = (1 - \Omega_0^2/\omega^2)$, where ε_0 is the permittivity of the semiconductor lattice, $\Omega_0 = \sqrt{4\pi e^2 n_\mathrm{s}/\varepsilon_0 m}$ is the plasma frequency, n_s is the electron concentration in the semiconductor, and m is the effective mass of electrons in the semiconductor.

DISCUSSION

Consider the transition radiation of surface electromagnetic waves from one electron bunch moving with velocity $v = 0.1c$. The waveguide radius is taken to be $R = 10^{-2}$ m. Let the dielectric be a vacuum and the semiconductor be GaAs, for which $\varepsilon = 12.53$ and $m =$

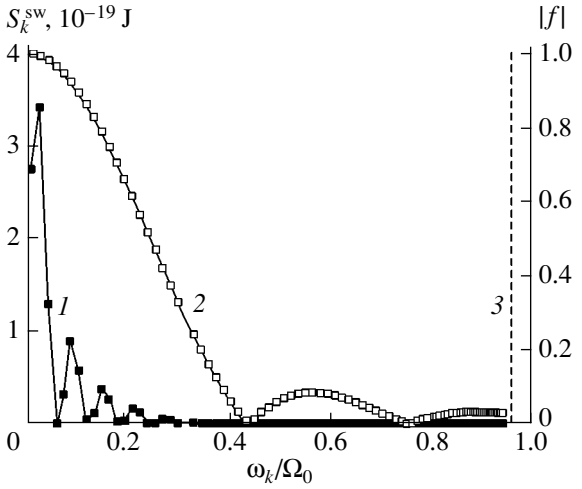


Fig. 1. (*1*) Energy $S^{sw}(\omega_k)$ of surface waves induced by one electron bunch and (*2*) geometric factor $|f(\omega_k)|$ of the bunch vs. frequency calculated by formulas (10) and (3), respectively. The bunch has the form of an ellipsoid of revolution (radius $a = 10^{-5}$ m, the axis of revolution $2b = 10^{-4}$ m long is aligned with the waveguide axis) and moves with velocity $v = 0.1c$ (where c is the velocity of light in free space) along the perfectly conducting waveguide with radius $R = 10^{-2}$ m.

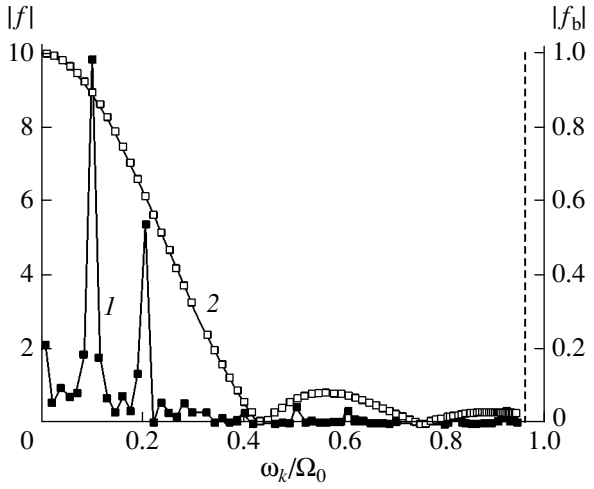


Fig. 2. (*1*) Resulting geometric factor $f(\omega_k)$ of the train of electron bunches and (*2*) geometric factor $|f_b(\omega_k)|$ of one bunch vs. frequency calculated by formulas (12) and (3), respectively. The bunches have the same dimensions: $a = 10^{-5}$ m and $b = 5 \times 10^{-5}$ m. The number of bunches in the train is $\sigma = 10$, and the bunch spacing $h = 3 \times 10^{-4}$ m is chosen such that the harmonic with $\omega_k \approx 0.1\,\Omega_0$ ($k = 7$) is amplified most efficiently. This harmonic corresponds to $M = 1$ in resonance condition (13).

$0.067m_0$ (where m_0 is the mass of a free electron). We also assume that $d = 5 \times 10^{-4}$ m and $n_s = 10^{22}$ m^{-3}. Figure 1 plots the frequency dependences of surface wave energy $S^{sw}(\omega_k)$ (curve *1*) and geometric factor $f(\omega_k)$ of the bunch (curve *2*) at $a = 10^{-5}$ m and $b = 5 \times 10^{-5}$ m. Each square data point in Fig. 1 is associated with a certain waveguide harmonic. Dashed line *3* shows the frequency of the surface plasmon. It is seen that the curve $S^{sw}(\omega_k)$ covers a discrete set of harmonics whose amplitudes decrease in an oscillatory manner with increasing k. The harmonics that fall into the frequency range of the first maximum of geometric factor $|f(\omega_k)|$ have the highest energies,

$$0 \le \sqrt{(\omega_k b/v)^2 + (a\mu_k^{(0)}/R)^2} \ll \frac{3}{2}\pi, \qquad (11)$$

where $\psi_k(\omega_k) \approx 3\pi/2$ roughly corresponds to the first zero of function $|f(\psi_k)|$.

The energy oscillations are due to the oscillating character of function $\sin(\omega_k d/2v)^2$ in expression (10). Energy maxima occur when $d = (M + 1/2)\lambda_{VK}$, where $\lambda_{VK} = 2\pi v/\omega_k$ is the van Kampen wavelength and $M = 1, 2, \ldots$, is an integer. In frequency range (11), the geometric factor the bunch takes is close to unity, which means that the vast majority of electrons in the bunch radiates coherently and the radiated energy is proportional to the squared number of particles in the bunch [2].

Consider the transition radiation of surface harmonics from a train of bunches with given spatial period h. As the period, we take the center distance between adjacent bunches. In this case, the expression for geometric factor $f(\omega_k)$ takes the form of the product of geometric factor $f_b(\omega_k)$ of one bunch by factor $f_s(\omega_k)$, which arises from summation over bunches in expression (1) for the charge density,

$$f(\omega_k) = f_b(\omega_k) f_s(\omega_k).$$

Here,

$$f_s(\omega_k) = \frac{\sin\left[\frac{\omega_k h}{2v}\sigma\right]}{\sin\left(\frac{\omega_k h}{2v}\right)} \exp\left[-i\frac{\omega_k}{v}\left(z_0 + \frac{(\sigma-1)h}{2}\right)\right], \qquad (12)$$

$f_b(\omega_k)$ is given by expression (3), σ is the number of bunches in the train, and z_0 is the distance between the first bunch and the vacuum–semiconductor interface.

As follows from formula (12), period h can be taken such that the resonance condition

$$\frac{\omega_k h}{2v} = \pi M \quad \text{and} \quad |f_s| = \sigma \qquad (13)$$

is met for one of the waveguide eigenmodes ($M = 1, 2, \ldots$, is an integer).

If this harmonic lies in frequency range (11), its energy is proportional to the squared product of the number of electrons in a bunch and the number of bunches in a train; accordingly, this harmonic will be amplified most efficiently. Figure 2 plots functions $|f(\omega_k)|$ (curve *1*) and $|f_b(\omega_k)|$ (curve *2*) versus frequency

at $a = 10^{-5}$ m, $b = 5 \times 10^{-5}$ m, $\sigma = 10$, and $h = 3 \times 10^{-4}$ m. It can be seen that the curve $|f_b(\omega_k)|$ cuts off only those oscillations of function $|f_s(\omega_k)|$ falling into frequency range (11). Therefore, resulting geometric factor $|f(\omega_k)|$ has only a few first maxima in this range. The higher the order of maximum (i.e., the greater integer M in formula (13)), the smaller its amplitude. Consequently, amplification of a kth waveguide harmonic is the most efficient when $M = 1$ and $h = 2\pi v/\omega_k$. In Fig. 2, the highest maximum of geometric factor $|f(\omega_k)|$ at frequency $\omega_k \approx 0.1\ \Omega_0$ ($k = 7$) corresponds to $M = 1$ in resonance condition (13). The next highest maximum of $|f(\omega_k)|$, which occurs at frequency $\omega_k \approx 0.2\Omega_0$ ($k = 14$), corresponds to $M = 2$ in resonance condition (13). The energy radiated in the first maximum of $|f(\omega_k)|$ in Fig. 2 is approximately equal to 10^{-17} J.

CONCLUSIONS

We considered surface electromagnetic waves excited by a train of electron bunches injected into a perfectly conducting semi-infinite cylindrical waveguide with a vacuum–semiconductor interface inside. The bunches, having the form of ellipsoids of revolution, are injected into the waveguide through an opening in the perfectly conducting end wall. Transition radiation from the bunches induces volume and surface waves at the interface. The surface waves have the form of standing waves with a discrete spectrum and are localized near the interface. It is shown that, by appropriately choosing the parameters of the train, a particular waveguide surface eigenmode (harmonic) can be amplified by bringing its frequency to resonance with the bunch repetition frequency in the train. In this case, the energy of this harmonic is proportional to the number of bunches in the train squared. Amplification will be the most efficient when the frequency of a harmonic being amplified falls into the range occupied by the first maximum of the geometric factor of one bunch, within which transition radiation from electrons is coherent and the radiated energy is proportional to the number of electrons squared.

ACKNOWLEDGMENTS

The author is grateful to V.M. Yakovenko and V.I. Karas' for valuable discussion.

REFERENCES

1. N. L. Dmitruk, V. G. Litovchenko, and V. L. Strizhevskiĭ, *Surface Polaritons in Semiconductors and Insulators* (Naukova Dumka, Kiev, 1989) [in Russian].
2. V. L. Ginzburg and V. N. Tsytovich, *Transient Radiation and Transient Scattering* (Nauka, Moscow, 1984; Hilger, Bristol, 1990).
3. V. A. Balakirev and G. L. Sidel'nikov, Zh. Tekh. Fiz. **69** (10), 90 (1999) [Tech. Phys. **44**, 1209 (1999)].
4. V. A. Balakirev, I. N. Onishchenko, D. Yu. Sidorenko, *et al.*, Zh. Tekh. Fiz. **72** (2), 88 (2002) [Tech. Phys. **47**, 227 (2002)].
5. Yu. O. Averkov and V. M. Yakovenko, Vopr. At. Nauki Tekh., Ser.: Plazm. Elektron. Novye Metody Uskor., No. 4, 45 (2003).

Translated by A. Khzmalyan

Technical Physics, Vol. 50, No. 8, 2005, pp. 1062–1068. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 99–105.

RADIOPHYSICS

Propagation of the H_{10} Mode in a Rectangular Waveguide with a Dielectric Discontinuity

Yu. V. Bludov

Usikov Institute of Radiophysics and Electronics, National Academy of Sciences of Ukraine, ul. Akademika Proskura 12, Kharkov, 61085 Ukraine

e-mail: bludov@ire.kharkov.ua

Received July 27, 2004

Abstract—The transmission of the H_{10} wave through a waveguide–capillary cavity in the form of a section of a rectangular waveguide with a square liquid-filled capillary inside is studied theoretically. The capillary passes through the center of the waveguide perpendicularly to its wide walls. The problem is solved by constructing the mirror images of the capillary on the narrow walls of the waveguide with subsequent consideration of the diffraction of the Brillouin components of the H_{10} waveguide mode by a periodic "array of capillaries." It is shown that a resonance minimum in the frequency dependence of the energy coefficient of transmission of the H_{10} wave through the waveguide–capillary cavity can be explained by the destructive interference of diffraction harmonics propagating in antiphase. The experimentally found existence of the capillary optimum dimension at which the resonance minimum is the most pronounced is corroborated analytically. *© 2005 Pleiades Publishing, Inc.*

INTRODUCTION

The structure and concentration of water solutions of various substances can be determined from the dielectric characteristics of these solutions measured in various ranges of electromagnetic wavelengths [1–7]. In particular, the coefficient of transmission of the waveguide H_{10} mode through a waveguide–capillary cavity representing a waveguide with a liquid-filled capillary placed between the centers of its wide walls was measured in the frequency ranges 27.5–37.5, 37–53, and 54–78 GHz [3–6]. It was found that the frequency dependence of the transmission coefficient of the H_{10} mode has a clear-cut resonance minimum in the frequency range where the waveguide is single-mode. The frequency of this minimum depends on the concentration [3–7] and temperature [8] of a solution under study. Based on these measurements, a method of measuring the concentration of water solutions of different substances was proposed [3–6] and patented [7].

Propagation of electromagnetic waves in a rectangular waveguide with an inductive dielectric cylinder (whose axis is parallel to the propagation direction of the H_{10} wave) was studied theoretically by various techniques [8–12]. In those works, theoretical frequency dependences of the transmission coefficient of the H_{10} wave exhibited the resonance minimum mentioned above, but a clear physical explanation of this effect was not given. Earlier [3], the resonance absorption was studied in terms of the model of surface wave excitation at the capillary and the resonance effect was explained by reflection from the wide walls of the waveguide. Later [13], it was found that resonance reflection of the surface waves from the wide walls cannot provide a high-Q resonance; therefore, it was conjectured that excitation of the H_{20} wave is also a factor. However, experiments did not reveal resonance absorption of the H_{20} wave energy.

In this paper, we theoretically study the propagation of the H_{10} wave through a waveguide–capillary cavity consisting of a square liquid-filled capillary placed in a waveguide perpendicularly to its wide walls. The problem is reduced to considering the diffraction of the H_{10} wave by a periodic array of capillaries that is constructed by multiply imaging the capillary on the narrow waveguide walls. The existence of a frequency at which the waveguide becomes nontransparent is physically justified.

BASIC EQUATIONS

Consider a rectangular waveguide of cross section $a \times b$ ($a > b$). Let the waves in the waveguide propagate in the z direction, the wide walls of the waveguide lie in the planes $y = 0$ and $y = b$, and the narrow walls lie in the planes $x = 0$ and $x = a$ (Fig. 1). A capillary is placed in the middle of the waveguide perpendicularly to its wide walls. It is made of a material with permittivity ε_c and is filled with a liquid of complex permittivity ε_l. For convenience, we assume that the capillary is a square tube with outer side D and inner side d and its axis is aligned with the y axis. Thus, the capillary with the liquid occupies the domain $0 < z < D$, $a/2 - D/2 < x < a/2 + D/2$, while the liquid itself occupies the domain $D/2 - d/2 < z < D/2 + d/2$, $a/2 - d/2 < x < a/2 + d/2$. An H_{10}

1063-7842/05/5008-1062$26.00 © 2005 Pleiades Publishing, Inc.

electromagnetic wave propagating in the positive z direction is incident on the capillary. The components of the electromagnetic field of the incident wave are given by

$$E_y(x, z, t) = E_0^{10}\sin(\pi x/a)\exp(i\Gamma z - i\omega t), \quad (1)$$

$$H_x(x, z, t) = -\frac{\Gamma}{k}E_0^{10}\sin(\pi x/a)\exp(i\Gamma z - i\omega t), \quad (2)$$

$$H_z(x, z, t) = -\frac{i}{k}\frac{\pi}{a}E_0^{10}\cos(\pi x/a)\exp(i\Gamma z - i\omega t), \quad (3)$$

where ω is the circular frequency of the wave, c is the speed of light in free space, $\Gamma = \sqrt{k^2 - (\pi/a)^2} > 0$, and $k \equiv \omega/c$.

To calculate the diffracted field, we will use the method of mirror images [14]. Since electromagnetic waves propagate in a waveguide by reflections from the metal walls, each event of reflection may be assigned an equivalent source of the reflected field, i.e., multiple imaging of the capillary on the waveguide walls takes place. These sources form an array of parallel square bars with period a (Fig. 1). Then, the incident wave may be represented as a superposition of two plane waves (Brillouin waves) as follows:

$$E_y(x, z) = E_i^+\exp(i\pi x/a + i\Gamma z) + E_i^-\exp(-i\pi x/a + i\Gamma z), \quad (4)$$

where

$$E_i^+ = E_0^{10}/(2i), \quad E_i^- = -E_0^{10}/(2i) \quad (5)$$

(in (5), factor $\exp(-i\omega t)$ is omitted for brevity). Components $H_x(x, z)$ and $H_z(x, z)$ of the incident electromagnetic wave can be represented as a superposition of two Brillouin waves in a similar fashion.

Eventually, the problem of propagation of the H_{10} wave in a waveguide with a dielectric capillary is reduced to the problem of diffraction of two plane *TE* waves by a periodic array of two-layer dielectric bars. The Maxwell equations for a *TE* wave can be written as follows:

$$\frac{\partial H_z^{(m)}}{\partial x} - \frac{\partial H_x^{(m)}}{\partial z} = ik\varepsilon_m(x)E_y^{(m)}, \quad (6)$$

$$\frac{\partial E_y^{(m)}}{\partial z} = -ikH_z^{(m)}, \quad (7)$$

$$\frac{\partial E_y^{(m)}}{\partial x} = ikH_z^{(m)}. \quad (8)$$

Here, superscripts m = 1–5 refer to the domains (Fig. 1) $z < z_1$, $z_1 < z < z_2$, $z_2 < z < z_3$, $z_3 < z < z_4$, and $z > z_4$, respectively, where $z_1 = 0$, $z_2 = (D - d)/2$, $z_3 = (D + d)/2$,

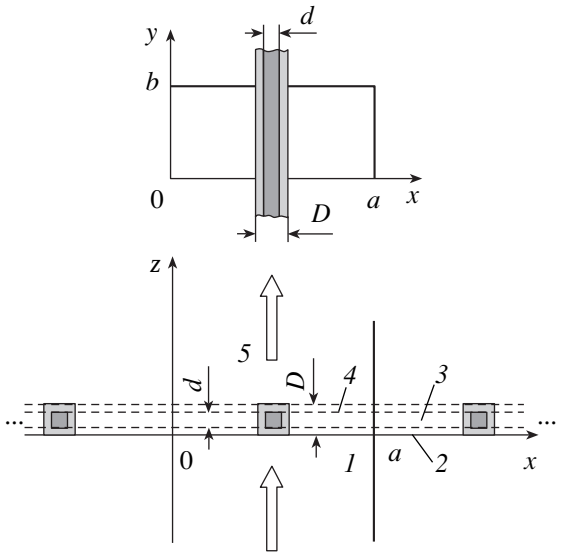


Fig. 1. Waveguide–capillary cavity consisting of a $a \times b$ rectangular waveguide with a square liquid-filled capillary inside. The outer and inner sides of the capillary are D and d, respectively.

and $z_4 = D$. In this case, $\varepsilon_1(x) = \varepsilon_5(x) = 1$ and

$$\varepsilon_2(x) = \varepsilon_4(x) = \begin{cases} 1, & na - (a - D)/2 < x < na + (a - D)/2 \\ \varepsilon_c, & na + (a - D)/2 < x < na + (a + D)/2, \end{cases} \quad (9)$$

$$\varepsilon_3(x) = \begin{cases} 1, & na - (a - D)/2 < x < na + (a - D)/2 \\ \varepsilon_c, & na + (a - D)/2 < x < na + (a - d)/2 \\ & na + (a + d)/2 < x < na + (a + D)/2 \\ \varepsilon_l, & na + (a - d)/2 < x < na + (a + d)/2, \end{cases} \quad (10)$$

where $n \in (-\infty, \infty)$.

It should be noted that periodic functions $\varepsilon_m(x)$ (m = 2–4) written in form (9) and (10) can be expanded into the Fourier–Floquet series in spatial harmonics,

$$\varepsilon_m(x) = \sum_{n=-\infty}^{\infty}\varepsilon_{m\|n}\exp(ingx). \quad (11)$$

Here, $g = 2\pi/a$ is the vector of the array, $\varepsilon_{2\|0} = \varepsilon_{4\|0} = 1 + (\varepsilon_c - 1)D/a$, $\varepsilon_{3\|0} = 1 + (\varepsilon_c - 1)D/a + (\varepsilon_l - \varepsilon_c)d/a$, $\varepsilon_{2\|n} = \varepsilon_{4\|n} = \frac{(-1)^n}{\pi n}(\varepsilon_c - 1)\sin(ngD/2)$, and $\varepsilon_{3\|n} = \frac{(-1)^n}{\pi n}[(\varepsilon_c - 1)\sin(ngD/2) + (\varepsilon_l - \varepsilon_c)\sin(ngd/2)]$ (at $n \neq 0$) are the amplitudes of the spatial harmonics.

Eliminating components $H_z^{(m)}$ and $H_x^{(m)}$ of the magnetic field from system of equations (6)–(8), we obtain

a wave equation in the form

$$\frac{\partial^2 E_y^{(m)}}{\partial x^2} + \frac{\partial^2 E_y^{(m)}}{\partial z^2} + k^2 \varepsilon_m(x) E_y^{(m)} = 0. \quad (12)$$

First, consider the diffraction of incident wave $E_y^+(x, z) = E_i^+ \exp(i\pi x/a + i\Gamma z)$ (the first term in expression (4)) by the array. Since a solution must be periodic in coordinate x, we represent the electric field of the *TE* wave as the Fourier–Floquet series in spatial harmonics:

$$E_y^{(m,+)}(x, z) = \sum_{s=-\infty}^{\infty} E_{y\|s}^{(m,+)}(z) \exp\left[i\left(s + \frac{1}{2}\right) g x\right].$$

Here, we used the property $\pi/a = g/2$. Substituting the series for functions $\varepsilon_m(x)$ and $E_y^{(m,+)}(x, z)$ into wave equation (12) and equating the coefficients multiplying the same exponential functions to zero, we come to the infinite system of differential equations

$$\frac{\partial^2 E_{y\|s}^{(m,+)}}{\partial z^2} - p_s^2 E_{y\|s}^{(m,+)} = 0; \quad m = 1, 5, \quad (13)$$

$$\frac{\partial^2 E_{y\|s}^{(m,+)}}{\partial z^2} - P_{m\|s}^2 E_{y\|s}^{(m,+)} + k^2 \sum_{n \neq s} \varepsilon_{m\|s-n} E_{y\|n}^{(m,+)} = 0; \quad m = 2, 3, 4. \quad (14)$$

Here, $p_s^2 = \left(s + \frac{1}{2}\right)^2 g^2 - k^2$ and $(P_s^{(m)})^2 = \left(s + \frac{1}{2}\right)^2 g^2 - k^2 \varepsilon_{m\|0}$. In addition, $\mathrm{Re}(p_s) - \mathrm{Im}(p_s) \geq 0$ and $\mathrm{Re}(P_s^{(m)}) - \mathrm{Im}(P_s^{(m)}) \geq 0$. A solution to system of equations (13) that meets the radiation condition at infinity can be obtained in the form

$$E_{y\|s}^{(1,+)}(z) = \delta_{s,0} E_i^+ \exp[-p_s z] + E_{y\|s}^{(1,+)}(0) \exp[p_s z], \quad (15)$$

$$E_{y\|s}^{(5,+)}(z) = E_{y\|s}^{(5,+)}(D) \exp[-p_s(z - D)]. \quad (16)$$

Note that the first term in expression (15) is the incident wave (with $i\Gamma = -p_0$). On rearrangement, a solution to system of linear differential equations (14) can be represented as

$$E_{y\|s}^{(m,+)}(z) = \sum_{n=-\infty}^{\infty} \varphi_{sn}^{(m)} \{B_{1\|n}^{(m)} \exp[\gamma_n^{(m)}(z - z_{m-1})] + B_{2\|n}^{(m)} \exp[-\gamma_n^{(m)}(z - z_{m-1})]\}. \quad (17)$$

Here, $\gamma_n^{(m)}$ is determined from the condition that the determinants of matrices $\widehat{F^{(m)}}$ ($m = 2, 3, 4$) equal zero; the elements of these matrices can be written as $F_{sn}^{(m)} = \delta_{sn}[(\gamma^{(m)})^2 - (P_s^{(m)})^2] + (1 - \delta_{sn}) k^2 \varepsilon_{m\|s-n}$ (δ_{sn} is the Kronecker delta). Also, in expression (17),

$$\varphi_{sn}^{(m)} = (-1)^{s+n} \frac{M_{ns}^{(m)}(\gamma_n^{(m)})}{M_{nn}^{(m)}(\gamma_n^{(m)})},$$

where $M_{ns}^{(m)}(\gamma_n^{(m)})$ is the algebraic complement of element $F_{ns}^{(m)}$ at $\gamma^{(m)} = \pm\gamma_n^{(m)}$.

As boundary conditions, we use the continuity conditions for the tangential components of the electric and magnetic fields at the boundaries $z = z_m$ ($m = 1$–4). Since the tangential components of the electric and magnetic fields are related through expression (7), the boundary conditions have the form

$$E_{y\|s}^{(m,+)}(z_m) = E_{y\|s}^{(m+1,+)}(z_m),$$
$$\frac{\partial E_{y\|s}^{(m,+)}(z_m)}{\partial z} = \frac{\partial E_{y\|s}^{(m+1,+)}(z_m)}{\partial z}. \quad (18)$$

Thus, by substituting expressions (15)–(17) into boundary conditions (18), we get a linear system of algebraic equations in the matrix form,

$$\sum_{n=-\infty}^{\infty} \{\widehat{H_{11\|sn}} X_{1\|n} + \widehat{H_{12\|sn}} X_{2\|n}\} = Q_{1\|s}, \quad (19)$$

$$\sum_{n=-\infty}^{\infty} \{\widehat{H_{21\|sn}} X_{1\|n} + \widehat{H_{22\|sn}} X_{2\|n}\} = Q_{2\|s}, \quad (20)$$

where

$$\widehat{H_{11\|sn}} = \begin{pmatrix} -\delta_{sn} & \varphi_{sn}^{(2)} & \varphi_{sn}^{(2)} & 0 \\ -\delta_{sn} p_n & \gamma_n^{(2)} \varphi_{sn}^{(2)} & -\gamma_n^{(2)} \varphi_{sn}^{(2)} & 0 \\ 0 & -\varphi_{sn}^{(2)} C_{n\|+}^{(2)} & -\varphi_{sn}^{(2)} C_{n\|-}^{(2)} & \varphi_{sn}^{(3)} \\ 0 & -\gamma_n^{(2)} \varphi_{sn}^{(2)} C_{n\|+}^{(2)} & \gamma_n^{(2)} \varphi_{sn}^{(2)} C_{n\|-}^{(2)} & \gamma_n^{(3)} \varphi_{sn}^{(3)} \end{pmatrix},$$

$$\widehat{H_{12\|sn}} = \begin{pmatrix} 0 & 0\,0\,0 \\ 0 & 0\,0\,0 \\ \varphi_{sn}^{(3)} & 0\,0\,0 \\ -\gamma_n^{(3)} \varphi_{sn}^{(3)} & 0\,0\,0 \end{pmatrix},$$

$$\widehat{H_{21\|sn}} = \begin{pmatrix} 0\,0\,0 & -\varphi_{sn}^{(3)} C_{n\|+}^{(3)} \\ 0\,0\,0 & -\gamma_n^{(3)} \varphi_{sn}^{(3)} C_{n\|+}^{(3)} \\ 0\,0\,0 & 0 \\ 0\,0\,0 & 0 \end{pmatrix},$$

$$\widehat{H_{22\|sn}} = \begin{pmatrix} -\varphi_{sn}^{(3)}C_{n\|-}^{(3)} & \varphi_{sn}^{(2)} & \varphi_{sn}^{(2)} & 0 \\ \gamma_n^{(3)}\varphi_{sn}^{(3)}C_{n\|-}^{(3)} & \gamma_n^{(2)}\varphi_{sn}^{(2)} & -\gamma_n^{(2)}\varphi_{sn}^{(2)} & 0 \\ 0 & -\varphi_{sn}^{(2)}C_{n\|+}^{(2)} & -\varphi_{sn}^{(2)}C_{n\|-}^{(2)} & \delta_{sn} \\ 0 & -\gamma_n^{(2)}\varphi_{sn}^{(2)}C_{n\|+}^{(2)} & \gamma_n^{(2)}\varphi_{sn}^{(2)}C_{n\|-}^{(2)} & -\delta_{sn}p_n \end{pmatrix},$$

$$X_{1\|n} = \begin{pmatrix} E_{y\|n}^{(1,+)}(0) \\ B_{1\|n}^{(2)} \\ B_{2\|n}^{(2)} \\ B_{1\|n}^{(3)} \end{pmatrix}, \quad X_{2\|n} = \begin{pmatrix} B_{2\|n}^{(3)} \\ B_{1\|n}^{(4)} \\ B_{2\|n}^{(4)} \\ E_{y\|n}^{(5,+)}(D) \end{pmatrix},$$

$$Q_{1\|s} = \begin{pmatrix} \delta_{s0}E_i^+ \\ -p_0\delta_{s0}E_i^+ \\ 0 \\ 0 \end{pmatrix}, \quad Q_{2\|s} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

$$C_{n\|+}^{(2)} = \exp(\gamma_n^{(2)}z_2), \quad C_{n\|-}^{(2)} = \exp(-\gamma_n^{(2)}z_2),$$

$$C_{n\|+}^{(3)} = \exp(\gamma_n^{(3)}d), \quad C_{n\|-}^{(3)} = \exp(-\gamma_n^{(3)}d).$$

When deriving expressions (19) and (20), we also used the relationships $z_2 - z_1 = z_2$, $z_3 - z_2 = d$, $z_4 - z_3 = z_2$, $\gamma_n^{(4)} = \gamma_n^{(2)}$, and $\varphi_{sn}^{(4)} = \varphi_{sn}^{(2)}$.

Since we are considering the single-mode waveguide, $g/2 < k < g$. Consequently, the waveguide supports the propagation of only those harmonics for which $n = 0$ and $n = -1$ (with $p_0 = p_{-1} = -i\Gamma$). The field of the transmitted propagating wave that arises as a result of diffraction of incident wave $E_y^+(x, z) = E_i^+ \exp(-i\pi x/a + i\Gamma z)$ by the array can be written in the form

$$E_y^{(t,+)}(x,z) = E_{y\|0}^{(5,+)}(D)\exp[igx/2 - p_0(z-D)] + E_{y\|-1}^{(5,+)}(D)\exp[-igx/2 - p_0(z-D)]. \quad (21)$$

From symmetry considerations, the relationships

$$\frac{E_{y\|0}^{(5,+)}(D)}{E_i^+} = \frac{E_{y\|0}^{(5,-)}(D)}{E_i^-} = t_0,$$
$$\frac{E_{y\|-1}^{(5,+)}(D)}{E_i^+} = \frac{E_{y\|1}^{(5,-)}(D)}{E_i^-} = t_1, \quad (22)$$

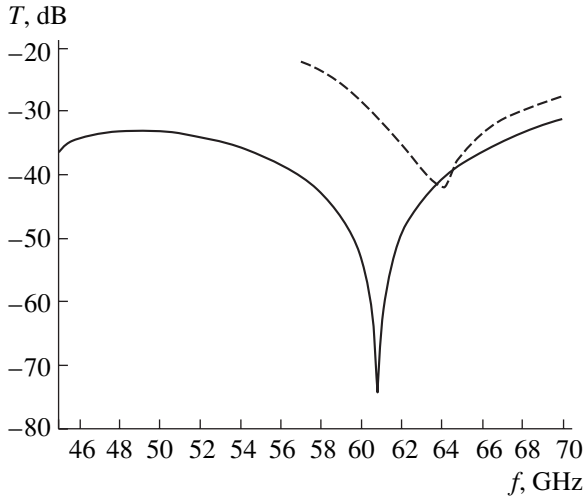


Fig. 2. Theoretical (solid line) and experimental (dashed line) frequency dependences of the energy transmission coefficient of the H_{10} wave propagating through the waveguide–capillary cavity filled with distilled water.

are valid for the diffraction of the wave $E_y^-(x, z) = E_i^- \exp(-i\pi x/a + i\Gamma z)$ (the second term in expression (4)) and the field of the propagating wave can be written as

$$E_y^{(t,-)}(x,z) = E_{y\|0}^{(5,-)}(D)\exp[-igx/2 - p_0(z-D)] + E_{y\|1}^{(5,-)}(D)\exp[igx/2 - p_0(z-D)]. \quad (23)$$

Using relationships (5) and (22), we can derive an expression for the total field of the transmitted H_{10} wave in the form

$$E_y^{(t)}(x,z) = E_y^{(t,+)}(x,z) + E_y^{(t,-)}(x,z) = (t_0 - t_1)E_0^{10}\sin(gx/2)\exp[-p_0(z-D)]. \quad (24)$$

Note that the term $(t_0 - t_1)$ in expression (24) has the meaning of the coefficient of transmission of the H_{10} wave through the waveguide–capillary cavity.

NUMERICAL RESULTS

Thus, propagation of waves in a waveguide with a discontinuity is completely defined by transmission coefficients t_0 and t_1 (expression (22)) of the propagating modes. Figure 2 plots the energy transmission coefficient $T = |t_0 - t_1|^2$ of the H_{10} wave versus its frequency. The parameters of the mathematical model are the same as those used in the experiment: we assumed that the wave propagates in a 3.4 × 1.8-mm rectangular waveguide and the capillary is made of Teflon with complex permittivity $\varepsilon_c = 2.04 + i3.468 \times 10^{-4}$, has dimensions D = 1.329 mm (outer side) and d = 0.975 mm (inner side), and is filled with distilled water

Table 1. Real (ε') and imaginary (ε") parts of the permittivity of distilled water found experimentally [15]

		f = 30 (GHz)	f = 37.5 (GHz)	f = 60 (GHz)	f = 100 (GHz)
t = 10°C	ε'	15.35	11.92	7.8	5.97
	ε"	26.78	22.49	14.86	9.13
t = 20°C	ε'	21.14	16.17	9.75	6.72
	ε"	30.92	26.82	18.45	11.55

at a temperature of 10°C.[1] The frequency and temperature dependences of the complex permittivity of distilled water, which are necessary for calculations, were obtained by polynomial fitting of the experimental data given in Table 1. The dashed line in Fig. 2 shows the experimental dependence for comparison.

As can be seen in Fig. 2, both the analytical and experimental frequency dependences of the coefficient of transmission of the H_{10} wave through the waveguide–capillary cavity exhibit a sharp minimum in the frequency range f = 60–63 GHz. The frequencies of this resonance minimum in the analytical and experimental curves differ insignificantly. This minimum stems from the fact that, at a certain frequency, $t_0 \approx t_1$ and, as a consequence, transmission coefficient $T = |t_0 - t_1|^2$ is close to zero at this frequency. Thus, the physical reason for the resonance minimum of the transmission coefficient for the H_{10} waveguide mode propagating in a waveguide–capillary cavity is that diffraction of either of the Brillouin components ($E_y^+(x, z)$ and $E_y^-(x, z)$, see (4)) of this mode generates two propagating waves whose wavevectors have the codirected z components and oppositely directed x components. In other words, the field diffracted by the capillary consists of four plane waves,

$$E_{y\|0}^+(x, z) = t_0\frac{E_0^{10}}{2i}\exp\left(i\frac{\pi}{a}x + i\Gamma z\right),$$

$$E_{y\|-1}^+(x, z) = t_1\frac{E_0^{10}}{2i}\exp\left(-i\frac{\pi}{a}x + i\Gamma z\right),$$

$$E_{y\|0}^-(x, z) = -t_0\frac{E_0^{10}}{2i}\exp\left(-i\frac{\pi}{a}x + i\Gamma z\right),$$

$$E_{y\|1}^-(x, z) = -t_1\frac{E_0^{10}}{2i}\exp\left(i\frac{\pi}{a}x + i\Gamma z\right).$$

If $t_0 = t_1$, waves $E_{y\|0}^+(x, z)$ and $E_{y\|1}^-(x, z)$, as well $E_{y\|0}^-(x, z)$ and $E_{y\|1}^+(x, z)$, cancel each other. Note that a small transmission coefficient in the low-frequency region (analytical curve) is explained by the fact that, at frequencies close to the waveguide cutoff frequency, the z components of the wavevectors of the incident Brillouin waves are much smaller than the x components. Thus, the Brillouin waves are incident on the capillary at a grazing angle, resulting in a small transmission coefficient [16].

Table 2. Outer and inner diameters of the capillaries used in the experiment

Capillary no.	D_c, mm	d_c, mm
1	1.0	0.7
2	1.5	1.1
3	2.3	1.8

Consider now how the transmission coefficient of the H_{10} wave varies with capillary dimensions. Figures 3–5 show the transmission coefficient as a function of the frequency and dimension of the outer wall of the capillary with the distilled water kept at 10, 15, and 20°C, respectively. The differently colored domains correspond to different values of the transmission coefficient: the lower T, the darker the domain. On the left of Figs. 3–5, the transmission coefficient is plotted versus the dimension of the capillary outer wall at fixed frequencies (the frequencies are shown on the right by vertical dashed lines). To make Figs. 3–5 more illustrative, dimension D of the capillary outer wall is converted to outer diameter D_c of the equivalent circular capillary. Here, it was assumed that diameter d of the capillary inner wall is a linear function of outer diameter D_c of the equivalent circular capillary,

$$d = \frac{\sqrt{\pi}}{2}\left[\frac{d_{c2} - d_{c1}}{D_{c2} - D_{c1}}(D_c - D_{c1}) + d_{c1}\right]. \qquad (25)$$

where d_{c2} = 1.1 mm, D_{c2} = 1.5 mm, d_{c1} = 0.7 mm, and D_{c1} = 1.0 mm. These values reflect the fact that the experiment was performed with capillaries of three dimensions (see Table 2). Thus, expression (25) describes a gradual transition from capillary no. 1 to capillary no. 2.

Figures 3–5 show that, as the outer dimension of the capillary increases, the minimum in the frequency dependence of the transmission coefficient shifts toward lower frequencies (because of the growth of zero-order spatial harmonics $\varepsilon_{2\|0}$, $\varepsilon_{3\|0}$, and $\varepsilon_{4\|0}$ of the

[1] When choosing the dimensions of the capillary, we recognized that, in the experiment, a circular capillary with outer diameter D_c = 1.5 mm and inner diameter d_c = 1.1 mm was used. Since our model uses the square capillary, its dimensions were taken such that the cross-sectional areas of the circular and square capillaries, as well as the cross-sectional areas of the liquid columns in them, are equal: $D^2 = \pi(D_c/2)^2$ and $d^2 = \pi(d_c/2)^2$.

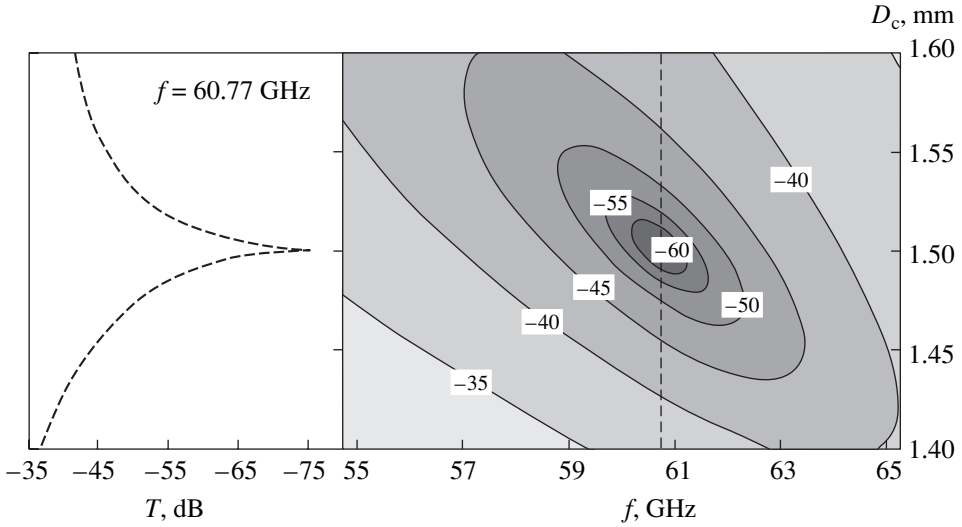


Fig. 3. Energy transmission coefficient of the H_{10} wave calculated as a function of frequency f and outer dimension D of the capillary at a temperature of distilled water of 10°C.

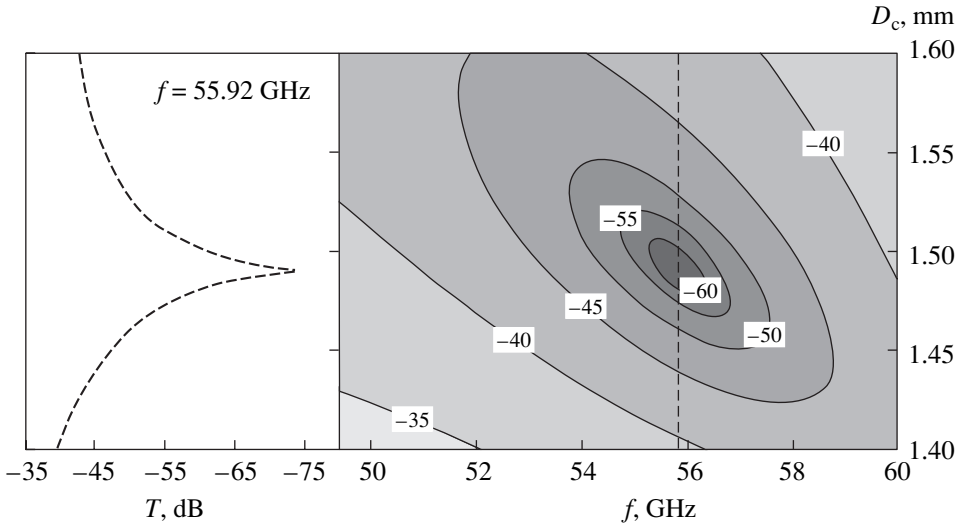


Fig. 4. Same as in Fig. 3 at 15°C.

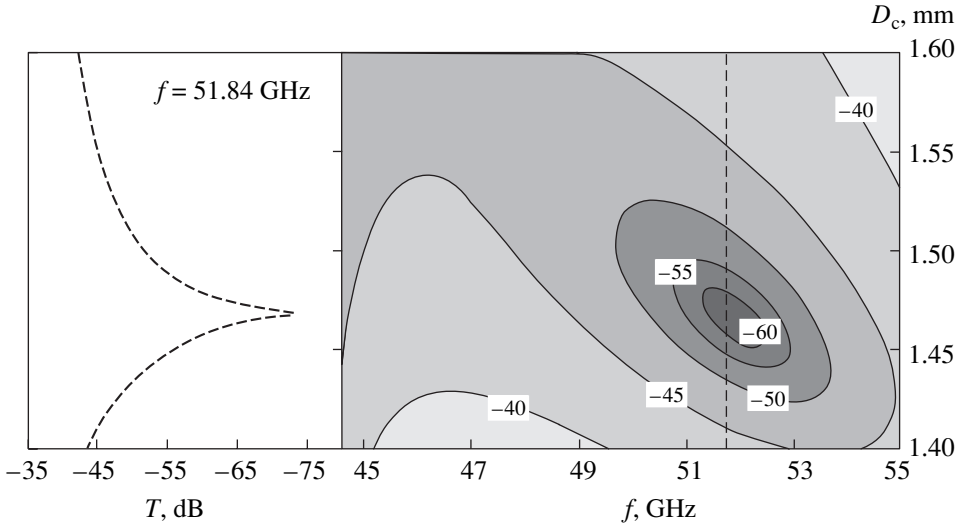


Fig. 5. Same as in Fig. 3 at 20°C.

permittivity). Also, the dependence of the transmission coefficient both on the frequency and outer dimension of the capillary has a clear-cut minimum (in other words, there exists an optimum dimension of the capillary at which the resonance minimum is the most pronounced). At a water temperature of 10°C, this optimum dimension is $D_c \approx 1.5$ mm (on the left of Fig. 3). Therefore, the transmission coefficient of the H_{10} wave is the smallest when the capillary dimensions roughly equal those of capillary no. 2 (Table. 2). This observation can explain the fact that the resonance minimum in the frequency dependence of the transmission coefficient in the experiments with capillaries no. 1 and 3 is indistinct. As the temperature of the distilled water rises, the frequency of the resonance minimum and the capillary optimum dimension decrease. In particular, at 15°C, the frequency of the minimum is $f \approx 55.9$ GHz and the optimal dimension is $D_c \approx 1.49$ mm (Fig. 4); at 20°C, $f \approx 51.9$ GHz and $D_c \approx 1.47$ mm (Fig. 5).

CONCLUSIONS

In this paper, we theoretically studied propagation of the waveguide H_{10} mode through a waveguide–capillary cavity consisting of a single-mode waveguide with a square capillary filled with a lossy liquid inside. The capillary runs parallel to the narrow walls of the waveguide through the middles of its wide walls (where the electric field of the H_{10} wave is maximal). The problem is solved by considering the diffraction of the H_{10} mode by a periodic array of capillary mirror images constructed on the narrow walls of the waveguide. The frequency dependence of the coefficient of transmission of the H_{10} mode through the system shows a resonance minimum. This minimum arises because, at a certain resonance frequency, the diffraction harmonics propagate in antiphase and cancel each other by interference. The experimental and theoretical frequency dependences of the transmission coefficient of the H_{10} mode are in satisfactory agreement. It is also corroborated theoretically that there exists an optimum dimension of the capillary at which the resonance minimum of the frequency dependence of the transmission coefficient is the most pronounced.

ACKNOWLEDGMENTS

The author is grateful to V.I. Lutsenko for submission of experimental data, A.I. Nosich for useful discussion, and A.Ya. Kirichenko for assistance.

This work was supported by the Science and Technology Center of Ukraine, project no. 2051.

REFERENCES

1. N. D. Devyatkov, Radiotekh. Élektron. (Moscow) **23**, 1882 (1978).
2. A. Yu. Akhadov, *Dielectric Properties of Binary Solutions* (Nauka, Moscow, 1977) [in Russian].
3. E. V. Belyakov, Elektron. Tekh. Ser. Elektronika SVCh, No. 9, 3 (1986).
4. G. F. Bakaushina, E. V. Belyakov, N. B. Zinov'eva, and A. M. Khrapko, Elektron. Tekh. Ser. Elektronika SVCh, No. 9, 3 (1986).
5. E. V. Belyakov and A. M. Khrapko, Elektron. Tekh. Ser. Elektronika SVCh, No. 10, 34 (1986).
6. E. V. Belyakov, Elektron. Tekh. Ser. Elektronika SVCh, No. 7, 51 (1987).
7. E. V. Belyakov and A. M. Khrapko, USSR Patent No. 1 307 315, Byull. Izobret., No. 16 (1987).
8. C. G. Hsu and H. A. Auda, IEEE Trans. Microwave Theory Tech. **34**, 883 (1986).
9. R. Gesche and N. Löchel, IEEE Trans. Microwave Theory Tech. **36**, 137 (1988).
10. R. Gesche and N. Löchel, IEEE Trans. Microwave Theory Tech. **37**, 962 (1989).
11. A. Valero and M. Ferrando, IEEE Trans. Microwave Theory Tech. **47**, 1997 (1999).
12. M. Polevski, R. Lech, and J. Mazur, IEEE Trans. Microwave Theory Tech. **52**, 1508 (2004).
13. E. V. Belyakov, Elektron. Tekh. Ser. Elektronika SVCh, No. 10, 59 (1989).
14. E. I. Veliev and V. P. Shestopalov, Zh. Tekh. Fiz. **49**, 1167 (1979) [Sov. Phys. Tech. Phys. **24**, 646 (1979)].
15. H. A. Hyatt, J. Quant. Spectrosc. Radiat. Transf. **10**, 217 (1970).
16. M. Born and E. Wolf, *Principles of Optics*, 4th ed. (Pergamon, Oxford, 1969; Nauka, Moscow 1973).

Translated by A. Khzmalyan

Technical Physics, Vol. 50, No. 8, 2005, pp. 1069–1075. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 106–112.

RADIOPHYSICS

On the Theory of Thomson Self-Oscillatory Systems

L. M. Lifshits

e-mail: l-m-lifshits@mtu-net.ru

Received October 26, 2004

Abstract—A nonlinear mechanism of the effect of broadband fluctuations on oscillations in a Thomson self-oscillator has been revealed. The correlation function for oscillations, which takes into account all appreciable effects in the second approximation, has been obtained for the first time. A new method has been developed for analyzing steady-state synchronous oscillations in a generator forming a part of the phase-synchronization system; linearized equations of steady-state synchronous oscillations, which take into account the effect of broadband noise, have been developed for the first time. An example of calculation of a complex phase-synchronization system is considered, in which the possibility of synthesizing highly effective systems with the help of the proposed methods of analysis is illustrated. The diffusion coefficients of the phase are estimated. © *2005 Pleiades Publishing, Inc.*

INTRODUCTION

Oscillations in a Thomson oscillator can be defined by the equation [1–6]

$$\ddot{u} + u = \varepsilon f(u) + n(t) + \mu q(t), \qquad (1)$$

where $u(t)$ is a normalized oscillatory process; ε and μ are small parameters ($\varepsilon \ll 1$, $\mu \ll \varepsilon$, the characteristic value of ε being much smaller than 0.1); $f(u)$ is a nonlinear function, $q(t)$ is an external stationary action; $n(t)$ is broadband (thermal, Schottky, or technical) noise, except in regions of substantial bursts $n(\sim)\varepsilon$; the symbols in parentheses indicate that the relation holds "almost everywhere"; the average value $\langle n \rangle = 0$; and the effective band of the noise spectrum $\omega_n \in (1 - \Delta_n, 1 + \Delta_n)$, $\Delta_n \gg \varepsilon$.

We assume that a soft excitation mode is implemented in system (1) and a steady-state mode of self-sustained oscillations sets in. The inclusion of the noise effect changes the nature of oscillations of the Thomson oscillator, increasing the number of degrees of freedom of the system. In particular, system (1) may acquire a type of steady-state oscillation such that the process evolving almost periodically in time changes randomly for relatively short accidental "bursts" (in the theory of random processes, this is associated with outbursts of the process, while a similar phenomenon in the theory of nonlinear dynamic systems is referred to as intermittence). It is well known [1–6] that many effects strongly influencing the oscillatory process can be studied only in the framework of the second approximation in ε. The closeness of system (1) to a linear one-dimensional oscillator and the difficulty of reliable measurement of many effects in the second approximation primarily necessitate the estimation of the correctness of the mathematical and physical models of the oscillatory system at each stage of theoretical analysis. The conventional asymptotic methods of analysis are well substantiated mathematically [1, 2, 7]; however, the additional conditions ensuring the uniqueness of the solution to Eq. (1) do not ensure physically realizable uniqueness in determining the parameters of oscillations. For $n(t) \neq 0$, such analytical methods may lead (in the second approximation in ε) to doubtful results from the standpoint of physics. In addition, operations with nonlinear functions are always nonlinear and incommutable with other operations. For example, in deriving the differential equation for a self-excited oscillator, Van der Pol adopted the cubic nonlinearity model $f(x) = (x^3/3)$; strictly speaking, this leads to the equation

$$\ddot{x} + \varepsilon \frac{d}{dt}\left(\frac{x^3}{3} - x\right) + x = 0$$

(naturally, this is immaterial in the case of a self-excited oscillator).

Combining the methods of statistical and harmonic linearization and assuming that harmonic linearization is a special case of statistical linearization [3–5] can ensure a correct analysis of Eq. (1). Such an approach makes it possible not only to refine the mechanism of oscillations in a Thomson self-excited oscillator taking into account second-approximation effects, but also to obtain new results.

ANALYSIS OF OSCILLATORY PROCESSES IN A GENERALIZED VAN DER POL OSCILLATOR

If we are dealing with cubic nonlinearity, the equation describing oscillations of a Thomson oscillator for $q(t) = 0$ can be written in the form

$$\ddot{u} + \varepsilon \frac{d}{dt}\left(\frac{4}{3}u^3 - u\right) + u = n_g, \qquad (2)$$

1063-7842/05/5008-1069$26.00

where noise $n_g(\sim)\varepsilon$ and $\langle n_g\rangle = 0$.

Let the spectral density of noise (in the spectral band $2\Delta_g$) be given by

$$S_g = 2N, \quad N \ll \varepsilon^2, \quad \omega_g \in (1-\Delta_g, 1+\Delta_g),$$

the noise dispersion being $\sigma_g^2 = 2N\Delta_g/\pi$. It can be seen that, being a source of energy of the oscillatory process, noise can be matched with the load only in a limited frequency band of the spectrum; consequently, the energy band of the noise participating in the formation of oscillations is finite (we assume that $\Delta_g \le \Delta_n$). In the case of a stationary low-intensity perturbation, the possibility of a steady-state mode of self-excited oscillations is known to be determined by the structural stability (roughness in Andronov's sense) of the Thomson oscillator [1–7]. To carry out a correct statistical linearization of nonlinearity in Eq. (2), the following obvious conditions must be satisfied (in the stationary oscillation mode) [3–5]:

$$\langle u\rangle = 0, \quad \langle u^2\rangle = \sigma_u^2 = \text{const}; \tag{3}$$

here, σ_u^2 is the dispersion of steady-state oscillation $u(t)$.

According to [3–5], we can derive the following expression for the criterion of the minimal mean square of the difference in processes at the exit of a nonlinear element and an equivalent linear element:

$$\frac{4}{3}u^3 = 2\sigma_u^2 u. \tag{4}$$

Relations (2) and (4) yield

$$\ddot{u} + \varepsilon(2\sigma_u^2 - 1)\dot{u} + u = n_g(t).$$

Assuming that $p = d/dt$, $p(=)i\omega$, and $i^2 = -1$, where ω is the spectral frequency, we define the dispersion of oscillations using the well-known formula [3]

$$\sigma_u^2 = \frac{N}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega}{\left|(i\omega)^2 + \varepsilon(2\sigma_u^2 - 1)i\omega + 1\right|^2}.$$

Here, we assume that $\varepsilon(2\sigma_u^2 - 1) \ll \Delta_g$; consequently, in the definition of σ_u^2, we assume that we are dealing with white noise. This is a tabulated integral [3]; its evaluation leads to the algebraic formula

$$2\sigma_u^2(2\sigma_u^2 - 1) - \frac{N}{\varepsilon} = 0,$$

whence

$$\sigma_u^2 = \frac{1}{2}\left(1 + \frac{N}{\varepsilon}\right) + O(\varepsilon^2). \tag{5}$$

Relations (2)–(5) lead to the statistically linearized equation for oscillations

$$\ddot{u} + N\dot{u} + u = n_g. \tag{6}$$

This equation can be treated as the initial-approximation equation for the generalized Van der Pol oscillator defined by Eq. (2).

From the standpoint of physics, the dynamic stability of self-excited oscillations (according to Kharkevich) is defined by the restoring (retrieving) force for the amplitude of steady-state oscillations [1]. Consequently, to take into account the dynamics of processes in the vicinity of steady-state oscillations, we carry out a change of variables in Eq. (6), assuming that $u = aX$, where a is the amplitude of oscillations. We also assume that the effect of higher harmonics has been taken into account in the value of the amplitude and that amplitude is independent of normalized oscillation X in the second approximation in ε. Qualitative analysis of Eq. (2) shows that

$$aX(=)ax + O(\varepsilon), \quad \sigma_u^2 = \sigma_{u_1}^2 + O(\varepsilon^2), \quad u_1 = ax,$$

where x is the first harmonic oscillation normalized in amplitude.

Taking into account the low intensity of noise, we can assume that parameters of oscillation are slow almost everywhere [1, 2, 4–6],

$$\begin{cases} u = aX, \quad \dot{a}(\sim)\varepsilon, \quad \langle\dot{a}\rangle = \langle\ddot{a}\rangle = 0 \\ (\langle(\dot{a})^2\rangle) = \sigma_{a'}^2 \gg \sigma_{a''}^2 \sim \varepsilon^3, \quad x(=)\cos\varphi \\ X(=)\cos\varphi + O(\varepsilon), \quad \dot{\varphi} = \Omega, \quad \ddot{\varphi}(\sim)\varepsilon^2, \end{cases} \tag{7}$$

where φ is the phase and Ω is the frequency of oscillations.

Steady-state conditions (3) should be apparently supplemented by analogous conditions for the steady-state oscillation parameters

$$\langle a^2\rangle = \sigma_a^2 = \text{const}, \quad \langle X\rangle = 0,$$
$$\langle X^2\rangle = \sigma_X^2 = \text{const}, \tag{3a}$$
$$\langle x\rangle = 0, \quad \langle x^2\rangle = \sigma_x^2 = \text{const}, \quad \langle\Omega\rangle = \text{const}.$$

It is clear from physical considerations that conditions (3) and (3a) can be regarded as sufficient. Using Eqs. (6) and (7), we obtain

$$\ddot{X} + V_x\dot{X} + \left[1 + \frac{\ddot{a}}{a} + O(\varepsilon^2)\right]X = \frac{n_g}{a},$$

where $V_x = N + 2\dot{a}/a$, $\langle V_x\rangle = N$, $\Delta_g \gg N \ll \varepsilon^2$.

Further, we assume [1–5] that $n_g \simeq n_{gk}X + n_{gi}\dot{X}$, where $\sigma_g^2 = \sigma_{gk}^2 = \sigma_{gi}^2$, $\langle n_g/a\rangle = 0$. In this case, we can easily determine the dispersion of the amplitude of steady-state oscillations from the conditions of energy

balance and normalization of steady-state oscillations [4, 5],

$$aV_x = n_{gi}, \quad \sigma_a^2 = 1 + O(\varepsilon^3),$$
$$\sigma_u^2 = \sigma_X^2 + O(\varepsilon^3) = \sigma_x^2 + O(\varepsilon^2), \tag{8}$$
$$2\sigma_u^2 - \sigma_a^2 = 2\sigma_{f_m}^2 + O(\varepsilon^2), \quad 2\sigma_{f_m}^2 = N/\varepsilon,$$

where $\sigma_{f_m}^2$ is the dispersion of the nonlinear frequency modulation by noise [4].

Since the amplitude of oscillations is independent of other oscillation parameters by definition (this is also in accordance with formulas (8)), we can easily derive the equation for the amplitude from Eq. (2) using the methods of harmonic linearization and assuming that

$$n_g = n_{gk}\cos\varphi - n_{gi}\sin\varphi, \quad \omega_{gi} \in (0, \Delta_g),$$
$$\sigma_a^2 = 1 + O(\varepsilon^3)\ [1, 2, 4], \tag{9}$$
$$2\dot{a} + \varepsilon a(a^2 - 1) = n_{gi}, \quad 2\ddot{a} + 2\varepsilon\dot{a} \simeq \dot{n}_{gi}.$$

In the vicinity of the steady-state oscillation mode, we assume that [1, 4, 5]

$$a^2 = 1 + 2\eta, \quad \langle\eta\rangle = 0, \quad \eta(\sim)\varepsilon,$$
$$\sigma_a^2 = (\langle a\rangle)^2 + \sigma_\eta^2, \tag{10}$$
$$a \simeq 1 + \eta - \eta^2/2, \quad [\langle\eta^2\rangle] = \sigma_\eta^2 < \varepsilon.$$

Using relations (9) and (10), we can easily derive the Rytov equation for amplitude fluctuations [1],

$$2\dot{\eta} + 2\varepsilon\eta = n_{gi}. \tag{11}$$

Obviously, taking into account formulas (9)–(11), we can write the expression for the correlation function of the amplitude in the form

$$K_a(\tau) = 1 - \frac{N}{2\varepsilon} + \frac{N}{2\varepsilon}\exp(-\varepsilon\tau), \tag{12}$$

where τ is the time shift.

Using the harmonic linearization methods, we can easily show [1, 2, 4] that, in the first approximation, Eq. (2) under conditions (7) can be presented in the form

$$u = u_1 - \frac{\varepsilon a^3}{8\Omega}\sin 3\varphi + O(\varepsilon^2), \quad u_1(=)a\cos\varphi. \tag{13}$$

Taking this relation into account, we can write the expression for the nonlinear component in Eq. (2) in the form [4]

$$\frac{4}{3}\varepsilon\frac{du^3}{dt} = \frac{4}{3}\varepsilon\frac{du_1^3}{dt} - \frac{\varepsilon^2 a^4}{8}u_1 + \varepsilon^2 u_j + O(\varepsilon^3). \tag{14}$$

Confining ourselves to analysis of the first harmonic of oscillation and assuming that

$$(4u_1^3/3) = 2a^3\sigma_x^2 x,$$

we obtain from Eq. (2) taking into account relations (5), (7), (8), and (14)

$$\begin{cases} \ddot{x} + \tilde{V}\dot{x} + \tilde{\Omega}^2 x = \dfrac{n_g}{a}, \quad \tilde{V} = \dfrac{2\dot{a}}{a} + \varepsilon(a^2 - 1) + N \\ \langle\tilde{V}\rangle = V = N \\ \tilde{\Omega}^2 = 1 - \dfrac{\varepsilon^2 a^4}{8} + \dfrac{\ddot{a}}{a} + \varepsilon\dfrac{\dot{a}}{a}(3a^2 - 1). \end{cases} \tag{15}$$

The value of $D_V = \langle[\tilde{V}^2]\rangle$ can be determined directly from the second formula in system (15) taking into account relations (7) and (9) [1, 4, 5].

In addition, we can easily obtain the values of V and D_V from the conditions of normalization and balance of average densities. This gives

$$\langle(n_{gi}/a)\rangle = V = N, \quad \langle(\tilde{V}^2)\rangle = D_V = \sigma_g^2 = 2N\Delta g/\pi,$$
$$n_g(=)n_{gk}\cos\varphi - n_{gi}\sin\varphi, \quad \langle(n_g/a)\rangle = 0.$$

In the vicinity of steady-state conditions of oscillations, parameter $\tilde{V}$ is small almost everywhere; frequency fluctuations are also small almost everywhere. Consequently, taking into account the structural stability (roughness) of the Thomson self-excited oscillator, we will determine the correlation function of the normalized process $x(t)$ from relations (15) using the method of freezing of parameters and then average the results [3–7],

$$K_x(\tau) = \sigma_x^2\exp(-N\tau/2)\cos\Omega_c\tau, \tag{16}$$

where $\langle(\tilde{\Omega}_c^2)\rangle = \Omega_c^2$, $\Omega_c = 1 - D_{\Omega_n} - \varepsilon^2/16$, and $D_{\Omega_n} = \sigma_g^2/8$.

A comparison of formula (16) with the known results [1, 3, 4] shows that D_{Ω_n} is the diffusion coefficient for the phase of oscillation and the dispersion of the phase diffusion is $D_\Phi(\tau) = D_{\Omega_n}\tau$. It follows from relation (16) that noise plays the role of drag force in this case [3, 4]. The formula defining $D_\Phi(\tau)$ can also be derived by the methods of harmonic linearization. Setting $u = ax$, $x = \cos\varphi$, $\dot{\varphi} = 1 + \dot{\Phi} + O(\varepsilon^2)$, and $\langle\dot{\Phi}\rangle = 0$ in formula (2), we can easily determine the dispersion of nonstationary phase fluctuations (dispersion of phase

diffusion) using the Rytov formula [1] for noise fluctuations of oscillation frequency:

$$D_\Phi(\tau) = \sigma_\Phi^2(\tau) = \frac{N}{8\pi}\int_0^\tau d\chi \int_0^\chi \frac{\sin\Delta_g\vartheta}{\vartheta}d\vartheta \simeq \frac{N}{16}\tau, \quad \chi\Delta_g \geq 1/\varepsilon. \tag{17}$$

Equating the values of diffusion coefficients of the phase, which are determined using different methods, we can estimate the energy band of noise participating in the formation of oscillation $u(t)$,

$$\Delta_g \simeq \pi/4, \quad \omega_u \in \left(\Omega_c - \frac{\pi}{4}, \Omega_c + \frac{\pi}{4}\right). \tag{18}$$

Equation (15) for the normalized oscillatory process and formula (18) show that the Thomson oscillator can be treated as a highly selective system only for averaged parameters of oscillations. The steady-state oscillation process is close to periodic almost everywhere except in the regions of relatively short outbursts. The correlation function of the first harmonic of oscillations is obviously given by

$$K_{u_1} = K_a(\tau)K_x(\tau) = \frac{1}{2}\left(1 + \frac{N}{2\varepsilon} - \frac{N}{2\varepsilon}\exp(-\varepsilon\tau)\right) \times\left(1 + \frac{N}{\varepsilon}\right)\exp(-N\tau/2)\cos\Omega_c\tau, \tag{19}$$

where

$$\Omega_c = 1 - \frac{\varepsilon^2}{16} - \frac{N}{16}.$$

Results (15) and (19) were obviously derived for the first time. The estimates of the effective band of noise close to the value given by formula (18) have been known for more than 20 years [4]; however, these results have not revealed the physical essence of this effect.

PHASE DIFFUSION IN A SELF-EXCITED THOMSON OSCILLATOR WITH LOCKING

We define oscillations in a Thomson oscillator by the equation

$$\ddot{u} + \varepsilon\frac{d}{dt}\left(\frac{4}{3}u^3 - u\right) + u = u_f + n_f, \tag{20}$$

where u_f is a periodic locking oscillation, $u_f = -b\sin\varphi_f$, $b \ll \varepsilon$, $\dot{\varphi}_f = \Omega_f = 1 + \nu_f$, $\nu_f = \text{const}$, $\nu_f \ll \varepsilon$, and $\nu_f^2 \sim \varepsilon^3$; the spectral density of oscillation u_f is $S_{u_f} = [\pi b^2/2]\delta(\omega - (1 + \nu_f))$, $\delta(\omega)$ is the Dirac delta function, and ω is the spectral frequency [1–3, 7]; and n_f is steady-state noise with a finite spectrum, the spectral density of noise is $S_{n_f} = 2N$, $2N < b$, the spectral noise band $\omega_{n_f} \in (\Omega_f - \Delta_n, \Omega_f + \Delta_n)$, $\Delta_n \approx 1$, $(1 - \Delta_n) > 0$, $n_f = n_{f_k}\cos\varphi_f + n_{f_i}\sin\varphi_f$, and $\langle n_f\rangle = 0$ (it should be noted that we are considering here the case of weak locking or, which is the same, locking due to a very weak external effect almost everywhere [1, 2, 8]); the remaining parameters of the system are defined in Eqs. (1) and (2).

We assume that steady-state conditions are satisfied in system (20) and the synchronous oscillation mode is realized almost everywhere. Obviously [1–3, 8], for $b \ll \varepsilon$, the amplitude stabilization rate is substantially higher than the phase stabilization rate and the amplitude dispersion differs from the value defined by the second formula in (8) only slightly. Consequently, in this problem, we can mainly confine our analysis to first-approximation effects and assume that

$$u(\simeq)\cos\varphi, \quad \dot{\varphi} = 1 + \dot{\Phi}, \quad \dot{\Phi}(\sim)\varepsilon, \quad \ddot{\Phi}(\sim)\varepsilon^2$$

in the reference frame connected with the oscillator; in addition, we assume that the oscillation phase in the reference frame of the oscillator is reduced to the interval $-\pi, \pi$. In this case, the steady-state conditions can be formulated as

$$\begin{cases}\sigma_u^2 = \langle u^2\rangle = \text{const}; \quad \langle u\rangle = 0\\ \langle\dot{\varphi}\rangle = \dot{\varphi}_f = \text{const}; \quad \varphi \in (-\pi, \pi).\end{cases} \tag{21}$$

Taking into account conditions (21) and formula (4), we obtain from Eq. (20)

$$\ddot{u} + C\dot{u} + u = u_f + n_f, \quad C = \varepsilon(2\sigma_u^2 - 1). \tag{22}$$

For an external frame of reference, steady-state oscillations of the oscillator can be represented in the form $u(t) = u_s(t) + u_g(t)$, where u_s is a synchronous harmonic oscillation and u_g is a steady-state quasi-harmonic (random) process. In this case, phase jumps causing phase diffusion are determined by the interaction of harmonic and random processes [1, 4, 6–9]. In the oscillator reference frame, phase jumps can be identified (for a very weak external interaction almost everywhere and for a large signal-to-noise ratio) with short-term (on the time scale bt) losses in controllability, which do not noticeably affect the characteristics of the oscillatory process. Linearized equation (22) can be simplified (truncated) in the first approximation in ε [1–5, 7]. We present oscillation $u(t)$ in the vicinity of the steady-state locking mode as follows:

$$u = U\cos t - R\sin t, \quad U(=)\cos\Phi_U, \quad R(=)\sin\Phi_R, \quad \ddot{U}(\sim)\varepsilon^2, \quad \dot{\Phi}_U = \dot{\Phi}_R = \dot{\Phi}, \quad \dot{\Phi}(\sim)\varepsilon, \quad \ddot{\Phi}(\sim)\varepsilon^2, \quad u(\simeq)\cos\varphi, \quad \dot{\varphi} = 1 + \dot{\Phi}, \quad \sigma_u^2 = \sigma_U^2 = \sigma_R^2. \tag{23}$$

Then we can easily derive from relations (22) and (23) the equation

$$2\dot{U} + CU = U_f + n_F, \quad (24)$$

where $U_f = b\cos\Phi_f$, $\dot{\Phi}_f = \nu_f$; the spectral density of the process $U_f(t)$ is $S_{U_f}(\omega) = \pi b^2\delta(\omega - \nu_f)/2$, noise $n_F = n_{F_i} + n_{F_k}$, $n_{F_i} = n_{f_i}\sin\Phi_f$, $n_{F_k} = n_{f_k}\cos\Phi_f$, spectral density of noise is $S_{n_F} = 2N$, $\omega_{n_F} \in (\nu_f - \Delta_n, \nu_f + \Delta_n)$.

For a more correct formulation of the problem, we represent oscillation U as the sum of the harmonic and quasi-harmonic components,

$$U(t) = U_s(t) + U_g(t).$$

Applying the spectrum translation operation [3, 7], we finally obtain

$$(2p + C)U_s = U_f, \quad (2(p - i\nu_f) + C)U_g = n_F, \quad (25)$$

where $p = d/dt$, $p(=)i\omega$, $i^2 = -1$, ω being the spectral frequency.

Using Eq. (25), we define the correlation function of oscillation,

$$K_U(\tau) = K_{U_s}(\tau) + K_{U_g}(\tau),$$

$$K_{U_s}(\tau) = \frac{A_s^2}{2}\cos\nu_f\tau = \frac{b^2\cos\nu_f\tau}{2(b^2\gamma^2 + C^2)}, \quad \gamma = \frac{2\nu_f}{b}, \quad (26)$$

$$K_{U_g}(\tau) = K_{U_{g0}}(\tau)\cos\nu_f\tau = \frac{N}{2C}\left[\exp\left(-\frac{C\tau}{2}\right)\right]\cos\nu_f\tau,$$

where $U_{g0}(t) = U_g(t)$ for $\nu_f = 0$.

Considering that $2\sigma_U^2 \approx 1$, we can easily derive from Eq. (26) the following equation defining the value of parameter C,

$$C^3 - NC^2 - b^2(1-\gamma^2)C - Nb^2\gamma^2 = 0. \quad (27)$$

The Descartes rule of signs [7] implies that Eq. (27) has only one positive root if $\gamma^2 < 1$. In this case,

$$C = b(1-\gamma^2)^{1/2} + \frac{N}{2(1-\gamma^2)}, \quad b > 2N, \quad \gamma^2 < 0.5.$$

For $\sigma_{u_s}^2 \gg \sigma_{u_g}^2$, small time intervals appear at random, in which envelope u_g of the process is larger than the amplitude of process u_s and the phases of these processes are opposite. This causes very fast (on the time scale bt) variations of the phase of the overall process; as a result, diffusion of the phase of oscillations is observed, the dispersion of nonstationary phase fluctuations increasing linearly during the observation time [1, 6, 9, p. 176).

The phase diffusion coefficient D is defined, taking into account relations (22) and (25)–(27), by the formula [9, p. 176]

$$D = \frac{\sqrt{\pi}}{A_s}\sigma_{\dot{U}_{g0}}\exp\left(-\frac{A_s^2}{2\sigma_{U_g}^2}\right), \quad (28)$$

where $\sigma_{U_g}^2 = \sigma_{U_{g0}}^2 = \sigma_{u_g}^2$ is dispersion of process $u_g(t)$, $\sigma_{\dot{U}_{g0}}$ is the standard deviation of process $\dot{U}_{g0}(t)$, $b\sqrt{(1-\gamma^2)} > 2N$, and A_s is the amplitude of processes $u_s(t)$ and $U_s(t)$.

Thus, the phase diffusion coefficient in an oscillator with locking can be written, taking into account relations (22) and (25)–(28), in the form

$$D \cong (\sqrt{N/2})\exp\left(-\frac{b\sqrt{1-\gamma^2}}{N}\right),$$

$$\gamma^2 < 1/2, \quad b\sqrt{1-\gamma^2} > 2N.$$

Let us estimate the statistically averaged duration of a transient process in an oscillator due to a phase jump (at instant $t = t_0$) of a locking oscillation (signal) for $b \gg 2N$, $\gamma^2 < 1/2$, assuming that steady-state synchronous oscillations have already been stabilized in the oscillator, processes in the vicinity of the steady-state mode being ergodic [1–3, 6, 8, 10]. In this case, without loss of generality, we can assume that $t_0 = 0$. This gives

$$U(\cong)U_0(t)\exp(i\Phi_f(t)), \quad U_0(0) = U_{e_0}(=)\exp(i\varphi_{e_\gamma}),$$

$$U_f = bU_\Phi\exp(i\Phi_f(t)), \quad U_\Phi = \exp(i\Phi_{0f} + i\varphi_{e_\gamma}),$$

where φ_{e_γ} is the steady-state phase error. $\varphi_{e_\gamma} \cong \gamma =$ const [8] and Φ_{0f} is the initial phase of the signal.

Since the random nature of the initial phase of the signal has been taken into account in Eqs. (20) and (24), the phase jump does not change the statistical linearization parameters. Thus, from Eq. (24), we obtain

$$U_0 \cong U_\Phi(1 - \exp(-Ct/2)) + U_{e_0}\exp(-Ct/2).$$

Finally, we obtain the following estimate of the duration of a transient process:

$$T_s \approx k_s T_1, \quad T_1 = 2/C, \quad (29)$$

here, T_s is the rated duration of the transient process and k_s is the proportionality factor.

A transient process can be treated as completed when the value of the phase of an oscillator being locked at instant T_s deviates from its steady-state value by less than 10^{-2}; in this case, $k_s = 2\ln 10$. It should be noted that estimate (29) virtually coincides with the known estimates [8].

CONDITIONS OF STATISTICAL EQUIVALENCE OF EQUATIONS FOR A LOCKED OSCILLATOR AND AN OSCILLATOR IN THE PHASE LOCKING SYSTEM

Oscillations in the Thomson oscillator in a phase-locking system (PLS) can be defined by the equations [1–3, 8]

$$\ddot{u} + \varepsilon\frac{d}{dt}\left(\frac{4}{3}u^3 - u\right) + (1 + 2\dot{\Phi})u + O(\varepsilon^2) = 0, \quad (30)$$

$$2\dot{\Phi} = b\sin\varphi_e + n_{f_i}\sin\varphi_e + n_{f_k}\cos\varphi_e, \quad (31)$$

where $\varphi_e = \Phi_f - \Phi$, $\dot{\Phi} = \nu_f - \dot{\varphi}_e$, and the remaining oscillation parameters are specified by formulas (20) and (21). It is well known [1–3, 8] that Eq. (31) describes the frequency balance in the first-order PLS (PLS-1). If the steady-state oscillation in the Thomson oscillator in a PLS is independent of amplitude fluctuations in the first approximation almost everywhere, the relations

$$u(=)\cos\varphi, \quad \varphi = t + \Phi, \quad \dot{U} = -(\sin\Phi)\dot{\Phi}, \quad (32)$$

satisfy Eq. (30), where U is the steady-state synchronous oscillation of the Thomson oscillator with a frequency shift.

Taking into account relations (21)–(23), we obtain from Eqs. (30)–(32)

$$\begin{aligned} 2\dot{U} + UC_\nu &= U_f + n_{f_i}\cos\Phi_f - n_{f_k}\sin\Phi_f, \\ C_\nu &= b\cos\varphi_e + n_{f_i}\cos\varphi_e - n_{f_k}\sin\varphi_e. \end{aligned} \quad (33)$$

Obviously, when physically substantiated conditions

$$\langle C_\nu\rangle = C \ll \varepsilon, \quad C > 0, \quad C^2 \sim \varepsilon^3 \quad (34)$$

are satisfied, Eq. (33) in the first approximation is statistically equivalent to Eqs. (24), (25). It should be emphasized that conditions (34) introduce only energy constraints ensuring dynamic stability of steady-state oscillations; the value of C can be determined from the energy balance conditions for steady-state synchronous oscillations (for Eq. (25), these conditions are obviously specified by formula (27)); consequently, conditions (34) can be regarded as general.

COMPLEX PHASE LOCKING SYSTEMS WITH IMPROVED DYNAMIC STABILITY

In a complex PLS, parameter $\dot{\Phi}$ appearing in Eq. (30) can be determined from the equation [1–3, 8]

$$\begin{aligned} 2\dot{\Phi} &= F(p)b\sin\varphi_e \\ &+ F(p)((\sin\varphi_e)n_{f_i} + (\cos\varphi_e)n_{f_k}), \end{aligned} \quad (35)$$

where $F(p)$ is the transmission function of the low-frequency filter ($F(p=0) = 1$, $b|F(p = i\omega)| \ll \varepsilon$, $i^2 = -1$, $\omega \in (0, \infty)$, and p is the differentiation operator).

The remaining parameters are defined in Eqs. (30), (31); the variables appearing in Eq. (35) permit multiple differentiation. We assume that Eq. (35) is stable in the small; the admissible initial detuning (locking band γ_m) is also assumed to be known [1–3, 7, 8]. Further, in analogy with relation (33), we obtain from Eqs. (32) and (35)

$$\begin{aligned} 2\dot{U} + UC_\nu = &-(\sin\Phi)F(p)((b + n_{f_i})\sin\varphi_e + n_{f_k}\cos\varphi_e) \\ &+ (\cos\Phi)F(p)((b + n_{f_i})\cos\varphi_e - n_{f_k}\sin\varphi_e). \end{aligned} \quad (36)$$

We assume that conditions (34) are satisfied in the vicinity of the steady-state synchronous oscillation mode. The possibility of representing the oscillatory process (36) in the form of the sum of processes analogous to (25) apparently imposes additional constraints on the fluctuation amplitude and the rate of variation of the phase error function in the vicinity of the steady-state oscillation mode:

$$\begin{aligned} b|F(p)\cos\varphi_e - \langle F(p)\cos\varphi_e\rangle|(\sim)\varepsilon^2, \quad b \ll \varepsilon, \\ |\cos\varphi_e - \langle\cos\varphi_e\rangle|(\sim)b, \\ |p(\langle F(p)\sin\varphi_e\rangle)| = o(\varepsilon^2). \end{aligned} \quad (37)$$

In transformations of nearly quasi-static functions for which conditions (37) hold, we can apply the method of frozen reactions in the right-hand side of Eq. (36) [3, 4, 7]. In this case, we obviously obtain the final result,

$$\begin{aligned} (2p + C)U_s &= F(p)U_f, \\ (2(p - i\nu_f) + C)U_g &= -(\sin\Phi_f)F(p)n_{f_k} + (\cos\Phi_f)F(p)n_{f_i}. \end{aligned} \quad (38)$$

The correlation function of oscillations determined from Eqs. (30) and (38) is obviously given by [1–7]

$$K_u(\tau) = \frac{A_s^2}{2}\cos(1 + \nu_f)\tau + K_{U_{g0}}(\tau)\cos(1 + \nu_f)\tau, \quad (39)$$

where

$$A_s^2 = \frac{b^2F(i\gamma)F(i\gamma)}{\gamma^2b^2 + C^2}, \quad \gamma = \frac{2\nu_f}{b}, \quad \sigma_{u_s}^2 \gg \sigma_{u_g}^2,$$

$$K_{U_{g0}}(\tau) = \frac{2N}{\pi}\int_0^{\Delta_n}\frac{F(i\omega)F(-i\omega)\cos\omega\tau d\omega}{4\omega^2 + C^2}.$$

The asymptotic value of the locking band γ_s can be determined from the inequality following from relation (39),

$$C^2 > b^2(F(i\gamma_s)F(-i\gamma_s) - \gamma_s^2) = 0, \quad \gamma_s^2 > \gamma^2, \quad (40)$$

where $C > 0$, C being the only real root of Eq. (39) for $\tau = 0$.

It can easily be verified that parameter $|\gamma_0| = 1$ for system (38) (which is usually referred to as the relative locking band [8]) characterizes only the static properties of the locking band.

The passage from Eq. (36) to Eq. (38) is obviously nontrivial since conditions (37), which ultimately change the characteristics of the energy balance, make the requirements imposed on the parameters of a complex PLS more stringent and in fact determine the class of phase locking systems with improved dynamic stability of synchronous oscillation mode. The PLSs defined by Eqs. (25) and (38) are characterized by short-term (on the time scale bt) outbursts of the oscillatory process.

Preliminary theoretical studies of PLSs defined by Eqs. (38) with various types of filters proved that the advantages of such systems are manifested when second- and higher-order filters are used. By way of an example, let us consider the results of analysis of PLS-3 with a low-frequency filter defined by the relation

$$F(p) = F_0\frac{p^2 + bp + F_0(b^2/4)}{(p + F_0b)(p + F_0b/4)},$$

where $0.1 < F_0 < 0.3$.

We can show that the locking band γ_m [3, 7, 8] and the locking retention band γ_s (formula (40)) in this case are given by $\gamma_m^2 \approx 3F_0 - 2F_0^2$ and $\gamma_s^2 \approx 2F_0 - 2F_0^2 + F_0^3$, respectively, and $\gamma_m^2 > \gamma_s^2$. Parameter $|\gamma_s|$ is smaller than the locking band $|\gamma_m|$, which is typical of locking systems defined by Eq. (38). It can easily be verified that

$$C^2 \approx b^2(1 - \gamma^2(1 + 1/(\gamma^2 + 4F_0^2))), \quad \gamma = 2\nu_f/b,$$

$$C^2 > 2Nb, \quad 0.1 < F_0 < 0.3, \quad 2\gamma^2 < \gamma_s^2.$$

The phase diffusion coefficient determined from formulas (28) and (39) is given by

$$D \approx (\sqrt{N/2})\exp\left(-\frac{C^2(1 + 2F_0)}{2F_0Nb}\right).$$

Since Eq. (30) takes into account far from all second-approximation effects and since we confined our analysis mainly to the first approximation, the estimates of phase diffusion in the locking system are authentic as long as $D > N$. For $b \cong C \gg 2N$, the statistically averaged duration of a transient process can be estimated, taking into account results (29), from the equation following from Eq. (38),

$$\left(p + \frac{C}{2}\right)U_0 \cong \frac{bF(p)U_\Phi}{2p} + U_{e_0}.$$

It can easily be verified that the transient process duration T_s in this case is $T_s \approx k_sT_F$, where $k_s = 2\ln 10$ and $T_F \approx (T_1/2F_0) = (1/F_0C)$.

It should be noted that we have considered a relatively simple example without touching upon the problems of PLS optimization.

CONCLUSIONS

We have presented a comprehensive theoretical analysis of steady-state oscillations in a free-running self-excited Thomson oscillator with allowance for second-approximation effects (the main results of analysis of flicker noise effects, which was carried out using an analogous approach, are given in [5]). For the first time, the nonlinear mechanism of influence of broadband fluctuations on oscillations in the Thomson oscillator is revealed. It is shown that the Thomson oscillator can be treated as a highly selective system only for average oscillation parameters. The process of steady-state oscillations is close to periodic almost everywhere; however, the process accidentally changes for relatively short random flares on the intervals characterized by noticeable outbursts.

A new method of analysis of steady-state oscillations in an oscillator forming a part of the phase locking system has been developed, and linearized equations for steady-state oscillations in a locking system with an elevated dynamic stability have been derived for the first time with regard to noise. An example of a complex PLS is considered, which confirms the possibility of synthesizing effective systems with high-order filters in the control circuit.

REFERENCES

1. S. M. Rytov, Yu. A. Kravtsov, and V. Tatarskiĭ, *Principles of Statistical Radiophysics* (Nauka, Moscow, 1976; Springer-Verlag, Berlin, 1987).
2. M. I. Rabinovich and D. I. Trubetskov, *Oscillations and Waves in Linear and Nonlinear Systems* (Nauka, Moscow, 1984; Kluwer, Dordrecht, 1989).
3. V. S. Pugachev, I. E. Kazakov, and L. G. Evlanov, *Statistical Theory of Automated Systems* (Mashinostroenie, Moscow, 1974) [in Russian].
4. L. M. Lifshits, Radiotekh. Élektron. (Moscow) **33**, 1899 (1988).
5. L. M. Lifshits, Radiotekh. Élektron. (Moscow) **37**, 1905 (1992).
6. H. Cramer and M. Leadbetter, *Stationary and Related Stochastic Processes* (Wiley, New York, 1967; Mir, Moscow, 1969).
7. G. A. Korn and T. M. Korn, *Mathematical Handbook for Scientists and Engineers* (McGraw-Hill, New York, 1968; Nauka, Moscow, 1984).
8. *Phase Locking*, Ed. by V. V. Shakhgil'dyan and L. N. Belyustina (Svyaz', Moscow, 1975) [in Russian].
9. V. I. Tikhonov, *Spikes in Random Processes* (Nauka, Moscow, 1970) [in Russian].
10. A. A. Koronovskiĭ, A. E. Khramov, and I. A. Khromova, Pis'ma Zh. Tekh. Fiz. **30** (6), 79 (2004) [Tech. Phys. Lett. **30**, 253 (2004)].

Translated by N. Wadhwa

Technical Physics, Vol. 50, No. 8, 2005, pp. 1076–1083. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 113–120.

RADIOPHYSICS

Diffraction of a Plane Electromagnetic Wave by a Slot in a Conducting Screen of Arbitrary Thickness

V. M. Serdyuk

Sevchenko Institute of Applied Physical Problems, Belarusian State University, Minsk, 220064 Belarus

e-mail: rhlab@bsu.by

Received November 4, 2004

Abstract—A 2D theoretical model of the diffraction of a plane electromagnetic wave by a slot in a perfectly conducting screen is constructed based on the partial domain method. The Tikhonov regularization is used to solve a system of algebraic equations for the slot-mode amplitudes. This makes it possible to extend the domain of applicability of the theory to conducting screens of arbitrary thickness and to significantly increase the accuracy of solution in the cases when the slot width and the screen thickness are comparable to the wavelength of the diffracted radiation. The absence of a continuous passage to the limit of an infinitesimal screen thickness from the case of an arbitrarily small finite thickness is demonstrated. The boundary conditions for the energy-flux vector are considered. A concept of the energy potential that is convenient for the computer calculations of the energy-flux lines of 2D diffraction fields is introduced. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The analysis of one of the most important diffraction elements of microwave, optical, and acoustical devices is the main reason for the interest in the classical problem of the plane-wave diffraction by a slot in an infinite conducting screen. A rigorous solution to the problem of diffraction by an additional screen of infinitesimal thickness representing a flat metal tape can be found in [1]. However, this solution is a limiting case of a more general solution for an elliptical cylinder and can hardly be used in practical calculations. Several attempts have been made to solve the above problems on the basis of approximate methods (see, for example, [2–7]). The method of partial domains, or the sewing method [8, 9], is the closest to the rigorous diffraction theory. As well as the rigorous theory [2, 3, 10], this method employs the division of the field-propagation volume into domains whose boundaries coincide with the coordinate surfaces. The fields in these domains are constructed as superpositions of elementary modes that satisfy all of the conditions except for the conditions on interfaces. Unknown amplitudes of the modes are found using the total-field sewing at the interfaces. This method was used in [6] to find a solution to the diffraction problem for the finite-thickness screens (the thickness is comparable to the wavelength of the diffracted radiation). In this work, we employ the partial-domain method to construct a solution to the problem of plane electromagnetic wave diffraction by a slot in a perfectly conducting screen of arbitrary thickness.

2. FORMULATION OF THE PROBLEM AND ITS SOLUTION FOR *H* POLARIZATION

We assume that a plane wave,

$$u_0 = \exp(i\mathbf{k}_0\mathbf{r} + i\varphi_0) = \exp(i[\alpha_0(x + d) + \beta_0 z]), \quad (1)$$

is incident on a screen with a slot (Fig. 1). Here, $\varphi_0 = \alpha_0 d$ is the initial phase; $\mathbf{k}_0$ is the wave vector of the incident wave; $\alpha_0 = k\cos\theta$ and $\beta_0 = k\sin\theta$ are its projections along onto the x and z coordinate axes, respectively ($\alpha_0^2 + \beta_0^2 = k^2$); $k = \omega/c$ is the wave number; and θ is the angle of incidence.

We need to determine the spatial components of the electric and magnetic vectors of the stationary field that results from the diffraction of the wave given by expression (1) by the aforementioned slot structure. We assume that the field is monochromatic and that its time dependence is given by a factor of $\exp(-i\omega t)$ (below, we omit this factor). Evidently, the field-propagation space can be divided into domains in the following way: the first domain ($x \le -d$) lies to the left-hand side of the screen, the second one is inside the slot ($-d \le x \le d$, $-l \le z \le l$), and the third domain lies to the right-hand side ($x \ge d$) of the screen (Fig. 1).

The solution to the diffraction problem under consideration depends on the polarization of the incident wave. First, we consider the *H* polarization, when the electric vector of the incident wave is perpendicular to the propagation plane. In this case, various components

1063-7842/05/5008-1076$26.00

of the electric and magnetic vectors can be represented in terms of single scalar function u [2, 3, 10]:

$$E_y = u, \quad H_x = \frac{i}{k}\frac{\partial u}{\partial z}, \quad H_z = -\frac{i}{k}\frac{\partial u}{\partial x}. \tag{2}$$

The remaining components appear to be equal to zero. Note that scalar function u should satisfy the Helmholtz equation:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial z^2} + k^2 u = 0. \tag{3}$$

We formulate the boundary conditions for this function based on the known conditions for the electric and magnetic field vectors. At a conducting surface, the tangential components of the electric field should vanish. At the domain boundaries, we impose the condition for the continuity of the tangential components of electric and magnetic fields. Then, in accordance with expression (2), function u should satisfy the following conditions:

$$(u)_{x=-d-0} = \begin{cases} 0 & (|z| > l) \\ (u)_{x=-d+0} & (|z| < l), \end{cases} \tag{4a}$$

$$(\partial u/\partial x)_{x=-d-0} = (\partial u/\partial x)_{x=-d+0} \quad (|z| < l), \tag{5a}$$

$$(u)_{x=d+0} = \begin{cases} 0 & (|z| > l), \\ (u)_{x=d-0} & (|z| < l), \end{cases} \tag{4b}$$

$$\begin{aligned} (\partial u/\partial x)_{x=d+0} &= (\partial u/\partial x)_{x=d-0} \ (|z| < l), \\ u(x, \pm l) &= 0 \quad (-d \le x \le d), \end{aligned} \tag{5b}$$

where zero in subscripts denotes an infinitesimal positive quantity.

In addition, we impose a condition for the finiteness of the field at $x \longrightarrow \pm\infty$ and $z \longrightarrow \pm\infty$.

In each of the above domains, we represent field function u as a superposition of sine waves (plane waves) that satisfy Helmholtz equation (3) and the aforementioned boundary conditions except for the conditions on the planes $x = \pm d$, which separate the domains. We assume that, inside the slot ($-d \le x \le d$ and $-l \le z \le l$), function u is given by

$$\begin{aligned} u = \sum_{n=1}^{+\infty} \{ & [a_{sn}\exp(i\sigma_{sn}(d+x)) \\ & + b_{sn}\exp(i\sigma_{sn}(d-x))]\cos\xi_{sn}z \\ & + i[a_{an}\exp(i\sigma_{an}(d+x)) \\ & + b_{an}\exp(i\sigma_{an}(d-x))]\sin\xi_{an}z \}, \end{aligned} \tag{6}$$

where $a_{sn,\,an}$ and $b_{sn,\,an}$ are the amplitudes of symmetric (subscript s) and antisymmetric (subscript a) modes along the z coordinate whose propagation parameters form infinite discrete series $\xi_{sn} = (\pi/l)(n - 1/2)$ and $\xi_{an} = \pi n/l$. This provides the zero field of each mode at the slot boundaries ($z = \pm l$) and their mutual orthogonality:

$$\int_{-l}^{+l} \cos\xi_{sn}z\cos\xi_{sm}z\,dz = \int_{-l}^{+l} \sin\xi_{an}z\sin\xi_{am}z\,dz = l\delta_{nm}.$$

Here, δ_{nm} is the Kronecker delta.

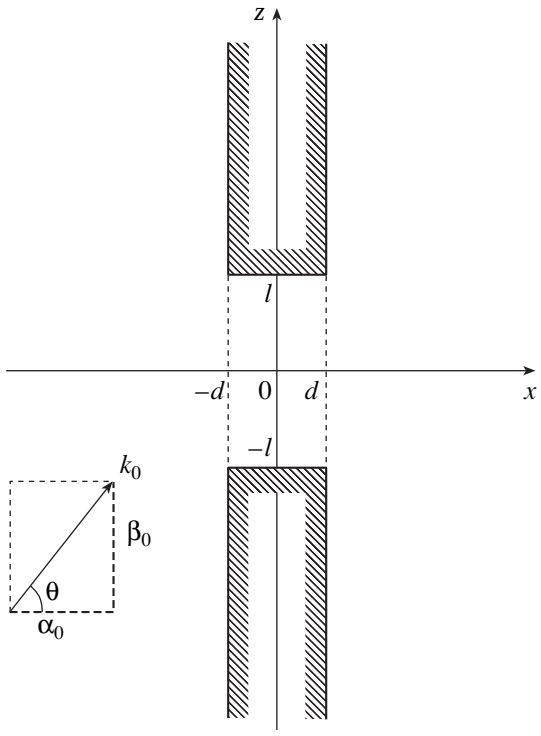


Fig. 1. Scheme of the plane-wave diffraction by a conducting screen with a slot: l is the slot halfwidth, d is the screen half-thickness, and θ is the angle of incidence of the wave.

Helmholtz equations (3) for modes (6) are fulfilled if the following conditions are satisfied:

$$\sigma_{sn,\,an} = \sqrt{k^2 - \xi^2_{sn,\,an}}. \tag{7}$$

The domains lying on both sides of the screen do not contain conducting surfaces that confine the field propagation along the z axis. Therefore, the field is determined by a continuous spectrum of the modes of free space rather than a discrete spectrum. To the left of the screen ($x \le -d$), we have

$$\begin{aligned} u = & (\cos\beta_0 z + i\sin\beta_0 z) \\ & \times \{\exp(i\alpha_0(x+d)) - \exp(-i\alpha_0(x+d))\} \\ & + \int_0^{+\infty} [A_s(\beta)\cos\beta z + iA_a(\beta)\sin\beta z]\exp(-i\alpha(x+d))d\beta. \end{aligned} \tag{8}$$

Here, the first term on the right-hand side represents an explicit sum of incident plane wave (1) and the plane

wave reflected from the screen. The integral term represents an expansion of the diffraction field into the Fourier integral. As well as in expression (6), we represent each mode as the sum of the symmetric and antisymmetric components with the amplitudes $A_s(\beta)$ and $A_a(\beta)$, respectively. For the right-hand-side domain ($x \geq d$), we employ the following representation:

$$u = \int_0^{+\infty} [B_s(\beta)\cos\beta z + iB_a(\beta)\sin\beta z] \times \exp(i\alpha(x-d))d\beta. \quad (9)$$

The fields given by expressions (8) and (9) should have finite amplitudes at $x \longrightarrow -\infty$ and $x \longrightarrow +\infty$. Therefore, for the parameters

$$\alpha = \sqrt{k^2 - \beta^2}, \quad (10)$$

we need to choose the root with the nonnegative imaginary part ($\mathrm{Im}\alpha \geq 0$). A necessary condition for the boundedness of these fields at $z \longrightarrow \pm\infty$ is an exact coincidence of the path of integration with the real axis β.

The modes of the infinite space for fields (8) and (9) also exhibit orthogonality,

$$\int_{-\infty}^{+\infty} \cos\beta z\cos\tilde{\beta}z dz = \int_{-\infty}^{+\infty} \sin\beta z\sin\tilde{\beta}z dz = \pi\delta(\beta - \tilde{\beta}),$$

where δ is the Dirac delta function.

The overlap integrals involving these functions and slot modes (6) are given by

$$\int_{-l}^{+l} \cos\beta z\cos\xi_{sn}z dz = lQ_n^{(s)}(\beta),$$

$$\int_{-l}^{+l} \sin\beta z\sin\xi_{an}z dz = lQ_n^{(a)}(\beta),$$

where we use the following notation:

$$Q_n^{(s)}(\beta) = \mathrm{sinc}((\beta - \xi_{sn})l) + \mathrm{sinc}((\beta + \xi_{sn})l),$$
$$Q_n^{(a)}(\beta) = \mathrm{sinc}((\beta - \xi_{an}l)) - \mathrm{sinc}((\beta + \xi_{an})l). \quad (11)$$

Here, sinc is the conventional notation for the function $\mathrm{sinc}x = \sin x/x$.

We find unknown mode amplitudes of the diffraction fields using the field sewing at the domain boundaries ($x = \pm d$). We substitute expressions (6), (8), and (9) into the boundary conditions (4) and expand the resulting equations with respect to the orthogonal set of functions $\cos\beta z$ and $\sin\beta z$ on the entire z axis (from $-\infty$ to $+\infty$). The resulting expressions make it possible to determine the amplitudes of the diffraction modes of free space in terms of the slot-mode amplitudes:

$$A_s(\beta) = \frac{l}{\pi}\sum_{n=1}^{\infty}(a_{sn} + b_{sn}\exp(2i\sigma_{sn}d))Q_n^{(s)}(\beta),$$

$$A_a(\beta) = \frac{l}{\pi}\sum_{n=1}^{\infty}(a_{an} + b_{an}\exp(2i\sigma_{an}d))Q_n^{(a)}(\beta),$$

$$B_s(\beta) = \frac{l}{\pi}\sum_{n=1}^{\infty}(a_{sn}\exp(2i\sigma_{sn}d) + b_{sn})Q_n^{(s)}(\beta), \quad (12)$$

$$B_a(\beta) = \frac{l}{\pi}\sum_{n=1}^{\infty}(a_{an}\exp(2i\sigma_{an}d) + b_{an})Q_n^{(a)}(\beta).$$

Equations for the latter are derived from boundary conditions (5) after substituting expressions (6), (8), (9), and (12) and using an expansion in orthogonal functions $\cos\xi_{sm}z$ and $\sin\xi_{am}z$ ($m = 1, 2, 3, \ldots$) for z ranging from $-l$ to l. The term-by-term summation and subtraction of the resulting equations yield four independent systems of infinite-order linear algebraic equations (two systems for the amplitudes of the symmetric slot modes and two systems for the antisymmetric modes). The expressions resulting from the truncation of the number of equations and unknown variables to a finite value of N are written as

$$\sum_{n=1}^{N} A_{mn}c_n = f_m, \quad (13)$$

where

$$A_{mn} = W_{nm}^{(s,a)}\Gamma_{sn,an}^{\pm} + \sigma_{sn,an}\Gamma_{sn,an}^{\mp}\delta_{nm},$$

$$c_n = c_{sn,an}^{\pm}, \quad f_m = \alpha_0 Q_m^{(s,a)}(\beta_0), \quad (14)$$

$$W_{nm}^{(s,a)} = \frac{l}{\pi}\int_0^{+\infty} \alpha Q_n^{(s,a)}(\beta)Q_m^{(s,a)}(\beta)d\beta,$$

$$\Gamma_{sn,an}^{\pm} = 1 \pm \exp(2i\sigma_{sn,an}d),$$
$$c_{sn,an}^{\pm} = (a_{sn,an} \pm b_{sn,an})/2. \quad (15)$$

Each of these systems can directly be solved using one of the conventional methods from [11, 12] (as was done in [6]). However, at a relatively small thickness $2d$ of the conducting screen, system (13) becomes ill-conditioned [11]: its determinant becomes so small that the conventional methods do not yield stable solutions. To obtain a desired solution to system (13) that is valid for any screen thickness, we employ the Tikhonov regularization method [13]. We search for the minimum of

functional

$$\sum_{m=1}^{N}\left|\sum_{n=1}^{N}A_{mn}c_n - f_m\right|^2 + \mu\sum_{m=1}^{N}|c_m|^2\xi_m^2 = \min, \quad (16)$$

rather than for a direct solution to the original system of equations. In expression (16), μ is the regularization parameter (a small positive quantity) that we choose based on the condition that the minimum of this parameter corresponds to a relatively well-conditioned system.

The additional term in on the left-hand side of expression (16) corresponds to the regularization of the desired solution with respect to the second derivative (i.e., to the application of the criterion for choosing the solution based on the minimum rate of variation along the slot). Differentiating condition (16) with respect to all of c_m^* (asterisk denotes complex conjugation), we derive a new well-conditioned system of equations,

$$\sum_{n=1}^{N}\left(\sum_{k=1}^{N}A_{km}^*A_{kn} + \mu\xi_n^2\delta_{mn}\right)c_n = \sum_{k=1}^{N}A_{km}^*f_k; \quad (17)$$

$$m = 1, 2, 3, \ldots, N.$$

We can now solve system (17) using any conventional approach from [11, 12] (e.g., the square-root method).

After solving system (13) (or, strictly speaking, regularized systems (17)), we may use expressions (12) to calculate the amplitudes of the field modes inside the slot and to the left-hand and right-hand sides of the screen:

$$a_{sn,an} = c^+_{sn,an} + c^-_{sn,an}, \quad b_{sn,an} = c^+_{sn,an} - c^-_{sn,an},$$

$$A_{s,a}(\beta) = \frac{l}{\pi}\sum_{n=1}^{N}(c^+_{sn,an}\Gamma^+_{sn,an} + c^-_{sn,an}\Gamma^-_{sn,an})Q_n^{(s,a)}(\beta), \quad (18)$$

$$B_{s,a}(\beta) = \frac{l}{\pi}\sum_{n=1}^{N}(c^+_{sn,an}\Gamma^+_{sn,an} - c^-_{sn,an}\Gamma^-_{sn,an})Q_n^{(s,a)}(\beta).$$

Then, we may calculate the spatial components of the electric and magnetic fields in all three domains by substituting expressions (6), (8), and (9) in Eqs. (2). For the numerical calculation of integrals (8) and (9) and for the calculation of the integrated elements of matrices (14), we may introduce a discrete mesh of argument β and apply approximate quadrature formulas from [12] that employ the values of integrands at the nodes or inside the cells. From the physical point of view, this means that the continuous spectrum of the field modes outside the slot (expressions (8) and (9)) is approximated using a discrete spectrum with the corresponding weighting coefficients. In this case, the number of modes that we take into account should be limited (e.g., the number should be no greater than certain integer M). For a uniform mesh, this corresponds to substituting finite number ΔM (Δ is the mesh step) for the infinite upper limit of integration in formulas (8), (9), and (14). Here (as well as in the case of limiting the dimension of system (13)), the verification of the boundary conditions of the diffraction problem after all calculations may serve as a criterion for choosing the correct values of number M and step Δ. At a uniform mesh of argument β, it is expedient to calculate the integrated elements of matrices (14) using a simple quadrature formula for the mean values from [12]. However, in the integrals entering into expressions (8) and (9), one should take into account the possibility of fast oscillations of the exponential functions at relatively large values of the x and z coordinates. This can easily be realized if the integrals are calculated in the linear approximation with respect to argument β of the exponential function and in the approximation of the mean values with respect to the remaining integrands.

3. SOLUTION FOR *E* POLARIZATION

If the electric vector of the incident wave is polarized parallel to the propagation plane (*E* polarization), we also represent various spatial components of the electric and magnetic vectors in terms of a single scalar function [2, 3, 10]:

$$E_x = -\frac{i}{k}\frac{\partial\bar{u}}{\partial z}, \quad E_z = \frac{i}{k}\frac{\partial\bar{u}}{\partial x}, \quad H_y = \bar{u}. \quad (19)$$

The remaining components equal zero.

This function $\bar{u}$ should satisfy Helmholtz equation (3). The boundary conditions for this function are established using expressions (19) and the above conditions for the vectors of electric and magnetic fields at the domain interfaces:

$$(\partial\bar{u}/\partial x)_{x=-d-0} = \begin{cases}0 & (|z|>l)\\ (\partial\bar{u}/\partial x)_{x=-d+0} & (|z|<l),\end{cases} \quad (20a)$$

$$(\bar{u})_{x=-d-0} = (\bar{u})_{x=-d+0} \quad (|z|<l), \quad (21a)$$

$$(\partial\bar{u}/\partial x)_{x=d+0} = \begin{cases}0 & (|z|>l)\\ (\partial\bar{u}/\partial x)_{x=d-0} & (|z|<l),\end{cases} \quad (20b)$$

$$(\bar{u})_{x=d+0} = (\bar{u})_{x=d-0} \quad (|z|<l),$$
$$(\partial\bar{u}/\partial x)_{z=\pm l} = 0 \quad (-d\le x\le d). \quad (21b)$$

Note that the condition for the boundedness of the fields at $x \longrightarrow \pm\infty$ and $z \longrightarrow \pm\infty$ remains valid.

As in the case of the *H* polarization, we construct field function $\bar{u}$ as a superposition of the sine modes satisfying Helmholtz equation (3) in each of the three domains. We assume that, inside the slot ($-d \le x \le d$ and

$-l \le z \le l$), the function is written as

$$\begin{aligned}\bar{u} &= k\sum_{n=1}^{+\infty}\{\bar{\sigma}_{sn}^{-1}[\bar{a}_{sn}\exp(i\bar{\sigma}_{sn}(d+x))\\ &- \bar{b}_{sn}\exp(i\bar{\sigma}_{sn}(d-x))]\cos\bar{\xi}_{sn}z\\ &+ i\bar{\sigma}_{an}^{-1}[\bar{a}_{an}\exp(i\bar{\sigma}(d+x))\\ &- \bar{b}_{an}\exp(i\bar{\sigma}_{an}(d-x))]\sin\bar{\xi}_{an}z\},\end{aligned} \tag{22}$$

where

$$\bar{\sigma}_{sn,\,an} = \sqrt{k^2 - \bar{\xi}^2_{sn,\,an}}, \tag{23}$$

and the propagation parameters of the symmetric and antisymmetric modes form infinite discrete series $\bar{\xi}_{sn} = (\pi/l)(n-1)$ and $\bar{\xi}_{an} = (\pi/l)(n-1/2)$, respectively. This provides for a zero value of the normal derivative of the field at the slot boundaries ($z = \pm l$) for each mode and their mutual orthogonality:

$$\int_{-l}^{+l}\cos\bar{\xi}_{sn}z\cos\bar{\xi}_{sm}z\,dz = l\delta_{nm}(1+\delta_{n1}),$$

$$\int_{-l}^{+l}\sin\bar{\xi}_{an}z\sin\bar{\xi}_{am}z\,dz = l\delta_{nm}.$$

We search for the field in the left-hand-side domain ($x \le -d$) represented as

$$\begin{aligned}\bar{u} &= (\cos\beta_0 z + i\sin\beta_0 z)\\ &\times[\exp(i\alpha_0(x+d)) + \exp(-i\alpha_0(x+d))]\\ &- k\int_0^{+\infty}\alpha^{-1}[\bar{A}_s(\beta)\cos\beta z + i\bar{A}_a(\beta)\sin\beta z]\\ &\times\exp(-i\alpha(x+d))d\beta.\end{aligned} \tag{24}$$

In the right-hand-side domain ($x \ge d$), the desired field is written as

$$\begin{aligned}\bar{u} &= k\int_0^{+\infty}\alpha^{-1}[\bar{B}_s(\beta)\cos\beta z + i\bar{B}_a(\beta)\sin\beta z]\\ &\times\exp(i\alpha(x-d))d\beta.\end{aligned} \tag{25}$$

For parameter α given by expression (10), we choose the branch of the root with the nonnegative imaginary part. The first term ion the right-hand side of expression (24) represents an explicit sum of incident plane wave (1) and the reflected plane wave.

We determine unknown amplitudes of the diffraction field modes in terms of the parameters of incident wave (1) substituting expressions (22), (24), and (25) into boundary conditions (20). We derive the corresponding expressions in the same way as in the case of the *H* polarization. Expressions obtained for the amplitudes of the free-space modes are identical to expressions (12) and contain the same overlap integrals (11). Then, we substitute these expressions into formulas (24) and (25) and boundary conditions (21) and expand in orthogonal functions $\cos\bar{\xi}_{sm}z$ and $\sin\bar{\xi}_{am}z$ ($m = 1, 2, 3, \ldots$) at the slot. The resulting systems are identical to the systems of linear equations (13) for the amplitudes of the symmetric and antisymmetric slot modes, in which

$$A_{mn} = \bar{W}^{(s,a)}_{nm}\bar{\Gamma}^{\pm}_{sn,\,an} + \chi_m\bar{\sigma}^{-1}_{sn,\,an}\bar{\Gamma}^{\mp}_{sn,\,an}\delta_{nm},$$

$$c_n = \bar{c}^{\pm}_{sn,\,an}, \quad f_m = k^{-1}Q^{(s,a)}_m(\beta_0),$$

$\chi_m = 2$ for the symmetric mode with $m = 1$, and $\chi_m = 1$ for the remaining modes,

$$\bar{W}^{(s,a)}_{nm} = \frac{l}{\pi}\int_0^{+\infty}\alpha^{-1}Q^{(s,a)}_n(\beta)Q^{(s,a)}_m(\beta)d\beta \tag{26}$$

and

$$\begin{aligned}\bar{\Gamma}^{\pm}_{sn,\,an} &= 1 \pm \exp(2i\bar{\sigma}_{sn,\,an}d),\\ \bar{c}^{\pm}_{sn,\,an} &= (\bar{a}_{sn,\,an} \pm \bar{b}_{sn,\,an})/2.\end{aligned} \tag{27}$$

The Tikhonov regularization yields new systems that are similar to systems (17). We employ the square-root method to derive the solution. Then, we find the mode amplitudes outside the slot using formulas (18), in which we substitute the parameters of *E* polarization given by expression (27) for the quantities *c* and Γ given by formula (15). Finally, we substitute expressions (22), (24), and (25) with the calculated mode amplitudes into expression (19) and calculate various components of the electric and magnetic fields in the entire space. Numerical calculation of integrals (24) and (25), as well as of integrated matrix elements (26), is performed using a discrete uniform mesh of argument β and approximate quadrature formulas similar to the formulas derived for the *H* polarization.

The theoretical model constructed can be applied for a numerical calculation of electromagnetic fields resulting from the diffraction of *H*- and *E*-polarized plane waves by perfectly conducting screens of arbitrary thickness at various slot widths. By way of example, Figs. 2 and 3 demonstrate (on the same scale) the results of such calculations on coordinate mesh *xQz* for the angle of incidence $\theta = 30°$, the slot halfwidth is $l = 1.4k^{-1}$, and the screen half-thickness is $d = 0.5l$. In the calculations, we employ the following parameters of the model: the step of the uniform mesh of the field-integration argument outside the slot, $\Delta = 0.1k$ ($0.01k$) at $kl \le 4$ ($kl > 4$); the regularization parameter for the system of equations (17), $\mu = 10^{-5}$; the number of the slot modes that are taken into account in expressions (6) and (22), $N = 19 + [2kl/\pi]$; and the number of the out-

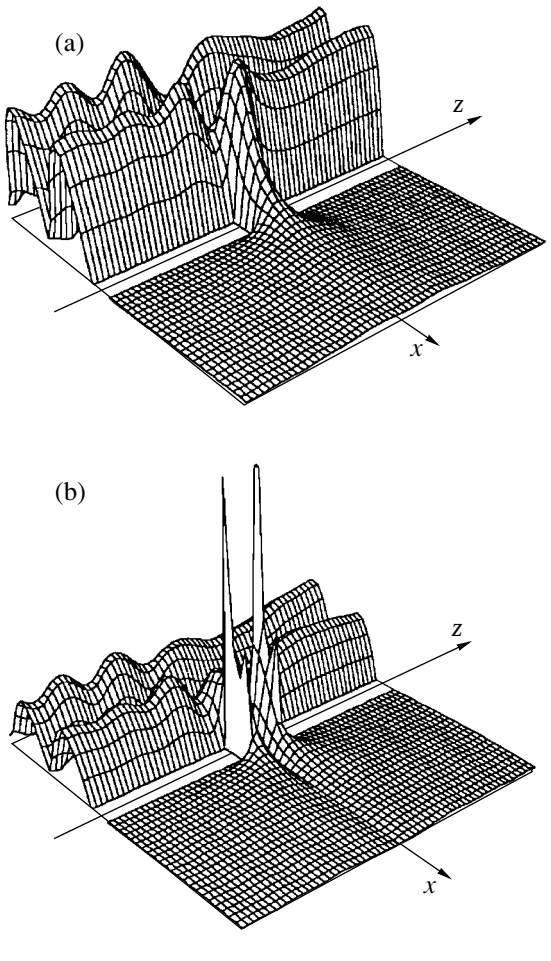


Fig. 2. Spatial distribution of the components of (a) electric E_y and (b) magnetic H_x fields for the diffraction of an *H*-polarized plane wave by a slot in a conducting screen.

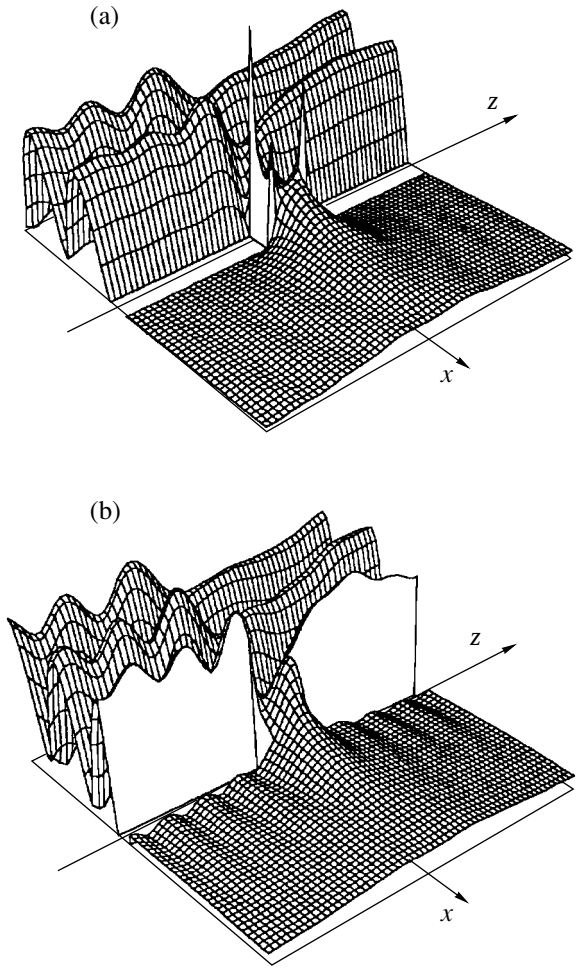


Fig. 3. Spatial distribution of the components of (a) electric E_z and (b) magnetic H_y fields for the diffraction of an *E*-polarized plane wave by a slot in a conducting screen.

of-slot modes of the approximating discrete spectrum in expressions (8), (9), (24), and (25), $M = 600 + N[\pi/\Delta l]$. Here, the brackets denote the integer part of a real number.

It is seen from Figs. 2b and 3a that H_x and E_z components exhibit discontinuities at the slot edges (such a discontinuity is also observed for H_z and E_x components). The reason for this lies in the fact that each slot edge represents a rectangular wedge. On one side of this wedge, each of the above field components is orthogonal to the conducting surface. On the other side, the same component appears tangential. This component of the electric field vanishes on the conductor. The normal component should be proportional to $\rho^{-1/3}$ (ρ is the distance from the edge) in the vicinity of the wedge edge [8, 14]. An inverse scenario is realized for the magnetic field. The normal component vanishes on the conducting surface, whereas the tangential component increases indefinitely when the distance from the edge decreases [8, 14].

4. CALCULATION OF ENERGY FLUXES FOR THE DIFFRACTION FIELD

For a comprehensive analysis, we should additionally calculate the energy characteristics of the diffraction field. One of the most important characteristics of the electromagnetic field is the vector of the mean energy flux (Poynting vector), which is written in terms of the electric and magnetic field strengths at a given point [14]:

$$\mathbf{S} = (c/8\pi)\mathrm{Re}(\mathbf{E} \times \mathbf{H}^*). \qquad (28)$$

In theory, this vector is normally interpreted as the radiation field intensity [10]. The value and direction of vector **S** at each point can directly be calculated using formula (28), provided that the spatial distributions of the fields **E** and **H** are known. This approach can easily be used for the calculation of scalar field $|\mathbf{S}|$. However, it is not convenient for the construction of the field of directions. Below, we propose an alternative method to construct the field of directions of the energy-flux vector that can easily be realized in computer calculations.

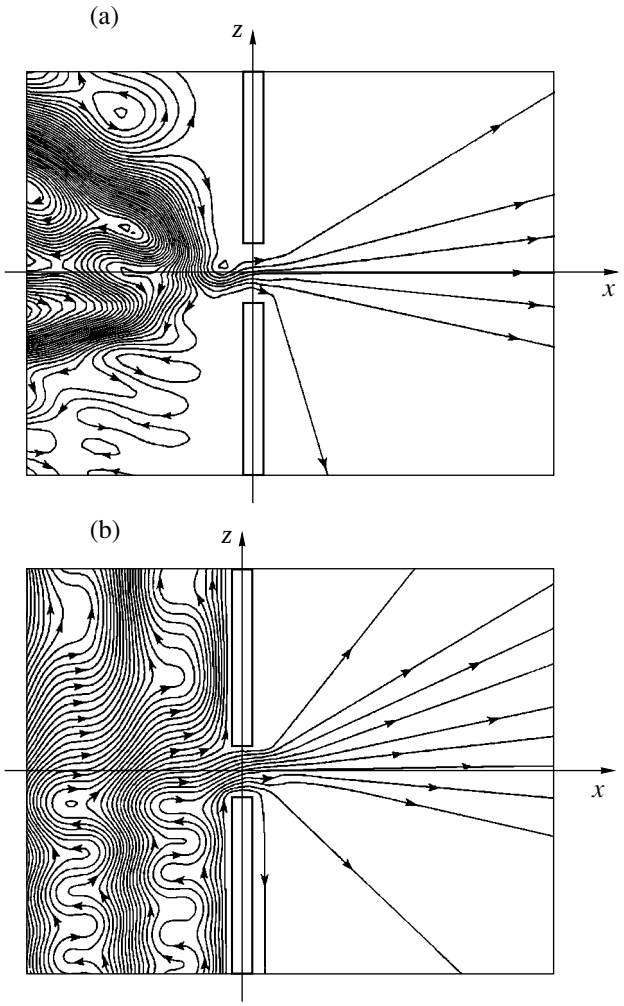


Fig. 4. Energy-flux lines for the diffraction of (a) *H*-polarized and (b) *E*-polarized plane waves by the slot in a conducting screen.

Consider vector $\mathbf{V} = (\mathbf{S} \times \mathbf{e}_y)$, where $\mathbf{e}_y$ is a unit vector along the *y* axis of the Cartesian coordinates. For a plane field completely belonging to the *x*0*z* plane and a nonabsorbing medium (currents and free charges are absent), we have $(\nabla \cdot \mathbf{S}) = 0$ [14]. Then, $(\nabla \times \mathbf{V}) = 0$ and the field of vectors **V** can be represented as a gradient of certain scalar function *U* which we call the energy potential:

$$U = U_0 - \int_{\mathbf{r}_0}^{\mathbf{r}} \mathbf{V} d\mathbf{r} = U_0 + \frac{c}{8\pi}\int_{\mathbf{r}_0}^{\mathbf{r}} \mathrm{Re}(\mathbf{e}_y \times (\mathbf{E} \times \mathbf{H}^*)) d\mathbf{r}. \quad (29)$$

Here $U_0 = U(\mathbf{r}_0)$ and $\mathbf{r}_0$ is the radius vector of an arbitrarily chosen fixed point in space.

Equipotential lines of the scalar field *U* are the flux lines of the solenoidal field **S** owing to the orthogonality of vectors **S** and **V**. Thus, it suffices to construct the equipotential lines of energy potential (29) for the construction of the field of directions of the energy-flux vector for a plane field in a nonabsorbing medium. In this case, that value of vector $\mathbf{V} = -\nabla U$ which is equal to the value of the energy flux of electromagnetic field is proportional to the rate of variation of the energy potential. Then, this value can be estimated locally using the density of equipotential lines in a given field region.

For an *H*-polarized plane field, the energy potential is defined as

$$U = U_0 - \frac{c}{8\pi}\left(\int_{x_0}^{x} \mathrm{Re}(E_y H_x^*) dx + \int_{z_0}^{z} \mathrm{Re}(E_y H_z^*) dz\right). \quad (30)$$

For an *E*-polarized field, we have

$$\bar{U} = \bar{U}_0 + \frac{c}{8\pi}\left(\int_{x_0}^{x} \mathrm{Re}(E_x H_y^*) dx + \int_{z_0}^{z} \mathrm{Re}(E_z H_y^*) dz\right). \quad (31)$$

Figure 4 demonstrates the equipotential lines of the energy potential (energy-flux lines) calculated with formulas (30) and (31) for the electromagnetic fields of the plane-wave diffraction by a slot shown in Figs. 2 and 3. It is seen that, for the case of *H* (*E*) polarization, the flux lines are repulsed (attracted) by the conducting surface. The reason for such a behavior of the energy-flux lines lies in the fact that the flux vector decreases down to zero for *H* polarization and increases to a maximum value for *E* polarization with a decrease in the distance from the conducting surface. This effect results from boundary conditions (28) for the energy-flux vector, which are derived from the boundary conditions for the fields **E** and **H** (expressions (2) and (19), respectively). Indeed, for *H* polarization, the value of this vector on the conductor surface vanishes, since all of the components of the electric field (including the tangential component E_y) equal zero on this surface. For *E* polarization, the tangential component of vector **S** (28) is proportional to $|H_y|^2 \mathrm{Im}(H_y^{-1} \partial H_y/\partial \xi)$, whereas the normal component is proportional to $|H_y|^2 \mathrm{Im}(H_y^{-1} \partial H_y/\partial \eta)$, where ξ is the tangential coordinate of the conducting surface (*z* or *x*) and η is the normal coordinate (*x* or *z*). The tangential component of the electric field equals zero on the conductor: $\partial H_y/\partial \eta = 0$ and $\partial |H_y|^2/\partial \eta = 0$. Hence, the tangential component of the energy-flux vector and $|H_y|^2$ reach a maximum in terms of argument η on the conducting surface.

CONCLUSIONS

The theoretical model proposed can easily be generalized to the case when the slot in a conducting screen is filled with an insulator (permittivity ε). For this purpose, we need to substitute $k^2\varepsilon$ for k^2 in formulas (7) and (23). Note that this model employs only computational approximations; in particular, infinite-dimensional systems are reduced to finite-dimensional systems and the integrals with respect to the free-space modes are

approximated with quadrature formulas. We do not use any additional approximations. Therefore, the model can be characterized as rigorous. Note that, in the rigorous classical theory [2, 3, 10], the solution for the fields is derived as integrals or infinite power series that are calculated using similar approximations.

In the case of an infinitesimal screen thickness ($d \longrightarrow 0$), the domain lying between the screen boundaries ($-d \le x \le d$) degenerates into the aperture plane. In this case, the model proposed yields relatively large values of the field components on at the slot edges (Figs. 2b and 3a). However, the model does not provide for the Sommerfeld-type singularities of type $\rho^{-1/2}$, which should emerge on the edge of a perfectly conducting half-plane [2, 3, 10]. This difference is related to different mathematical approaches to the construction of the field in the aperture plane. In the rigorous theory (in particular, in the Sommerfeld model [2, 3, 10]), the diffraction field is expanded in terms of the sine basis functions (Fourier components) at the entire aperture plane including the conductor. In the framework of the partial-domain method, the diffraction field is calculated only in the slot domain, where the metal is absent and the field differs from zero. Hence, the behavior of the field in the vicinity of the screen edges obtained for a screen with an arbitrarily small finite thickness differs from that obtained for the screen with an infinitesimal thickness. Therefore, to solve the problems of diffraction by screens with extremely small thickness, it is expedient to concentrate on the selection of a theoretical model. In particular, the question arises as to whether these problems can be solved using the methods and results of the rigorous diffraction theory, based on the assumption regarding infinitesimally thin screens.

Real conductors always have a finite thickness. Therefore, in our opinion, the preference should be given to the methods that take into account the finiteness of the thickness, in particular, the partial-domain method (see, for example, [15]).

The Tikhonov regularization used to solve system of equations (13) for the amplitudes of slot modes is helpful not only in the case of a thin conducting screen. This method makes it possible to significantly increase the accuracy of solution of the diffraction problem for finite-thickness screens, since it diminishes the error related to the difference between resulting fields (2) and (19) on both sides of the slot boundaries. The difference between the regularized and unregularized solutions decreases with an increase in the slot width and the screen thickness. However, this difference is significant when the slot width is approximately the wavelength of incident radiation. If the slot width is substantially greater than the wavelength, the theoretical model under consideration remains valid but the amount of computations significantly increases. The reason for this lies in an increase in the number of the slot modes that are taken into account and of the mesh nodes used to calculate the field (the mesh step should be a few times smaller than the wavelength to adequately describe the diffraction pattern).

REFERENCES

1. N. W. McLachlan, *Theory and Application of Mathieu Functions* (Clarendon, Oxford, 1947; Inostrannaya Literatura, Moscow, 1953).
2. H. Hönl, A. Maue, and K. Westpfahl, *Theorie der Beugung* (Springer-Verlag, Berlin, 1961; Mir, Moscow, 1964) [translated from German].
3. L. A. Vainshtein, *The Theory of Diffraction and the Factorization Method* (Sov. Radio, Moscow, 1966; Golem, Boulder, 1969).
4. L. N. Litvinenko and S. L. Prosvirnin, *Spectral Operators of Scattering in Problem of Wave Diffraction by Plane Screens* (Naukova Dumka, Kiev, 1984), pp. 100–174 [in Russian].
5. S. C. Kashyap and M. A. Hamid, IEEE Trans. Antennas Propag. **19**, 499 (1971).
6. L. N. Litvinenko, S. L. Prosvirnin, and V. P. Shestopalov, Radiotekh. Élektron. (Moscow) **22**, 474 (1977).
7. A. P. Anyutin and A. G. Kyurkchan, Radiotekh. Élektron. (Moscow) **49**, 15 (2004).
8. R. Mittra and S. W. Lee, *Analytical Techniques in the Theory of Guided Waves* (Macmillan, New York, 1971; Mir, Moscow, 1974).
9. A. D. Grigor'ev and V. B. Yankevich, *Microwave Cavities and Slow-Wave Systems: Numerical Methods of Analysis and Design* (Radio i Svyaz', Moscow, 1984) [in Russian].
10. M. Born and E. Wolf, *Principles of Optics*, 4th ed. (Pergamon, Oxford, 1969; Nauka, Moscow 1973).
11. D. K. Faddeev and V. N. Faddeeva, *Computational Methods of Linear Algebra* (Lan', St. Petersburg, 2002; Dover, New York, 1959).
12. N. N. Kalitkin, *Numerical Methods* (Nauka, Moscow, 1978) [in Russian].
13. A. N. Tikhonov and V. Ya. Arsenin, *Solutions of Ill-Posed Problems* (Nauka, Moscow, 1979; Halsted, New York, 1977), pp. 110–127.
14. L. A. Vaĭnshteĭn, *Electromagnetic Waves* (Radio i Svyaz', Moscow, 1988) [in Russian].
15. P. D. Kukharchik, V. M. Serdyuk, and I. A. Titovitskiĭ, Radiotekh. Élektron. (Moscow) **49**, 551 (2004).

Translated by A. Chikishev

Technical Physics, Vol. 50, No. 8, 2005, pp. 1084–1088. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 121–125.

RADIOPHYSICS

Energy Characteristics of Radiation from an Oscillating Dipole Moving in a Dielectric Medium with Resonant Dispersion

A. V. Tyukhtin

Radiophysics Research Institute, St. Petersburg State University, St. Petersburg, 198504 Russia
e-mail: tyukhtin@niirf.spbu.ru

Received December 9, 2004

Abstract—Radiation produced by oscillations of an electric dipole moving along its dipole moment through an insulator with the resonance-type dispersion is considered. The total power of radiation and the power density spectrum are studied both analytically and numerically. It is shown that the radiation spectrum consists of either two separate frequency ranges or a single frequency range depending on the parameters of the problem. The dependences of the radiation power on the velocity of the source are revealed for various values of the resonant and Langmuir frequencies. © *2005 Pleiades Publishing, Inc.*

INTRODUCTION

By now, the radiation from oscillating dipole sources that move relative to the ambient homogeneous medium has been studied over a long period of time (see, e.g., [1–6]). The most thorough consideration was given to the motion of various oscillators in dispersion-free media and in cold plasmas [2, 4, 6]; in these cases, the problem was analyzed not only in the coordinate system related to the medium, but also in the system related to the source [7–11]. This study deals with a medium with resonant frequency dispersion. It should be emphasized that the dispersion of this type is typical of media with bound charges in a very wide frequency range with its lower edge lying from several gigaherz to several thousands of gigaherz, depending on the chemical composition of the substance.

It is known that moving oscillating sources may be both micro- and macroscopic (radiating atoms, antennas, etc.). To macroscopic sources, one could assign small polarized domains of the medium itself (for example, a running polarized domain produced by laser radiation in recent experiments [12]). It is noteworthy that a number of recent experiments were aimed at producing media where the velocity of light is extremely low (see, e.g., [13, 14]). Therefore, even slow motion of a source in such media significantly affects the electromagnetic field structure. In these conditions, the motion of conventional macroscopic sources (such as antennas), which is usually characterized by velocities much lower than the velocity of light in free space, has an appreciable effect on the emission process.

In this study, we consider the radiation of an oscillating electric dipole that moves uniformly in a homogeneous nonmagnetic isotropic nonabsorptive medium with permittivity in the form

$$\varepsilon(\omega) = 1 + \frac{\omega_L^2}{\omega_r^2 - \omega^2} = \frac{\omega_r^2\varepsilon_0 - \omega^2}{\omega_r^2 - \omega^2}, \qquad (1)$$

where ω_r and ω_L are the resonance and Langmuir (plasma) frequencies, respectively, and $\varepsilon_0 = 1 + \omega_L^2/\omega_r^2$ is the permittivity of the medium with respect to a constant field.

Note that a detailed study even in the simplest case of a point charge has been accomplished only recently. In particular, the radiation of a charge moving in a boundless dispersive resonance dielectric medium was analyzed in [15, 16], and radiation in a waveguide filled with a dielectric was considered in [17].

1. GENERAL EXPRESSIONS FOR THE FIELD AND THE POWER OF RADIATION

Let an electric dipole in the intrinsic coordinate system be characterized by dipole moment density $\mathbf{P}' = \mathbf{p}'\delta^3(x', y', z')$, where $\mathbf{p}' = p_0'\mathbf{e}_z'\exp(-i\omega_0' t')$ and ω_0' is the frequency in this coordinate system (primed values relate to the intrinsic coordinate system of the source). We assume that the dipole moves in the direction of its moment with the velocity $\mathbf{v} = v\mathbf{e}_z$ ($v > 0$). In the lab coordinates, the source also possesses only electric dipole moment $p = p'\sqrt{1-\beta^2}$. General expressions for describing the field of a longitudinal oscillating dipole in an isotropic nonabsorptive medium free from spatial dispersion were derived in [4]. Recall that the nonzero field components can be represented in the form [4, 6]

$$E_z = \frac{i}{c}\int_{-\infty}^{\infty} a\omega\left[1 - \left(\frac{\omega-\omega_0}{\omega\beta n}\right)^2\right]e^{i\psi}d\omega,$$

1063-7842/05/5008-1084$26.00

$$E_r = -\int_{-\infty}^{\infty} \frac{\partial a}{\partial r} \frac{\omega - \omega_0}{\omega \beta n^2} e^{i\psi} d\omega, \quad (2)$$

$$H_\varphi = -\int_{-\infty}^{\infty} \frac{1}{\mu} \frac{\partial a}{\partial r} e^{i\psi} d\omega,$$

where

$$\psi(\omega) = \frac{\omega - \omega_0}{v} z - \omega t,$$

$$a(\omega, r) = \frac{p_0' \sqrt{1-\beta^2}\, \omega \mu(\omega)}{2cv} H_0^{(1)}(rs(\omega)),$$

$$s(\omega) = \begin{cases} \operatorname{sgn}(\omega) v^{-1} \sqrt{(\omega n \beta)^2 - (\omega - \omega_0)^2} \\ \text{at} \quad (\omega n \beta)^2 > (\omega - \omega_0)^2 \\ i v^{-1} \sqrt{(\omega - \omega_0)^2 - (\omega n \beta)^2} \\ \text{at} \quad (\omega - \omega_0)^2 > (\omega n \beta)^2, \end{cases} \quad (3)$$

$$n^2(\omega) = \varepsilon(\omega)\mu(\omega), \quad \omega_0 = \omega_0' \sqrt{1-\beta^2}$$

and the radicals in the expressions for $s(\omega)$ are assumed to be positive.

Note that, in the case $(\omega - \omega_0)^2 > (\omega n \beta)^2$, it appears convenient to express Hankel function $H_0^{(1)}(rs)$ in the expression for $a(\omega, r)$ in terms of the Macdonald function with the use of the formula $H_0^{(1)}(i\xi) = -2i\pi^{-1} K_0(\xi)$.

In what follows, we investigate radiation power Σ and its spectral distribution $\sigma(\omega)$. The radiation power can be expressed in the form [4, 6]

$$\Sigma = \int_{\omega > 0} \sigma(\omega) d\omega, \quad \sigma(\omega) = \tilde{\sigma}(\omega) + \tilde{\sigma}(-\omega), \quad (4)$$

$$\tilde{\sigma}(\omega) = \frac{p_0'^2 (1-\beta^2)}{4c^2 v} \mu |\omega|^3 \left[1 - \left(\frac{\omega - \omega_0}{\omega n \beta}\right)^2\right] \times 1((\omega n \beta)^2 - (\omega - \omega_0)^2), \quad (5)$$

where $1(\xi)$ is the unit step function

$$1(\xi) = \begin{cases} 1 & \text{for} \quad \xi > 0 \\ 0 & \text{for} \quad \xi < 0. \end{cases}$$

The value of $\sigma(\omega)$, which presents the spectral density of radiation, is seen to differ from zero in the frequency range defined by

$$\omega^2 \beta^2 \varepsilon(\omega)\mu(\omega) > (\omega - \omega_0)^2. \quad (6)$$

Within this range, the value of $s(\omega)$ is real, which seems natural, since only propagating waves can contribute to the radiation field. Note that, in the case of the more rigid condition

$$\omega^2 \beta^2 \varepsilon(\omega)\mu(\omega) > (\omega + \omega_0)^2, \quad (7)$$

both integrand terms in (4), $(\tilde{\sigma}(\omega))$ and $(\tilde{\sigma}(-\omega))$, are nonzero. If, however, condition (6) is met and condition (7) is violated, then $\tilde{\sigma}\omega \neq 0$ and $\tilde{\sigma}(-\omega) = 0$.

Let us apply the results obtained above to a nonmagnetic medium with a permittivity expressed in the form (1). This leads to the expression

$$\tilde{\sigma}(\omega) = \frac{p_0'^2 (1-\beta^2)}{4v^3} \frac{|\omega| f(\omega)}{\omega_r^2 + \omega_L^2 - \omega^2}, \quad (8)$$

where

$$f(\omega) = (1-\beta^2)\omega^4 - 2\omega_0 \omega^3 + [\omega_0^2 - \omega_r^2(1-\beta^2) + \omega_L^2 \beta^2]\omega^2 + 2\omega_0 \omega_r^2 \omega - \omega_0^2 \omega_r^2. \quad (9)$$

Condition (6), which defines the range of the generated frequencies, can be reduced to the following requirements:

$$f(\omega) > 0 \quad \text{for} \quad \omega^2 \leq \omega_r^2,$$
$$f(\omega) < 0 \quad \text{for} \quad \omega^2 \geq \omega_r^2 + \omega_L^2. \quad (10)$$

Note that no radiation can exist in the frequency region $\omega_r^2 < \omega^2 < \omega_r^2 + \omega_L^2$, since condition (6) is violated because of negative value of ε.

2. THE CASE OF LOW RESONANT FREQUENCY

An approximate analytical description of the most essential features of radiation is possible in a number of specific cases. Assume first that the resonance frequency of the medium is much lower than the oscillator frequency in the lab coordinates,

$$\omega_r \ll \omega_0 = \omega_0' \sqrt{1-\beta^2}. \quad (11)$$

It is not difficult to show that the radiation spectrum includes two frequency ranges in this case. The first range is defined by the inequalities

$$\omega_1 < \omega < \omega_r, \quad \omega_1 \approx \omega_r (1 + \beta^2 \omega_L^2 \omega_0^{-2})^{-1/2}. \quad (12)$$

This frequency range can be referred to as "resonant" since it lies near the resonance frequency ω_r.

The second range of the generated frequencies is defined by the inequalities

$$\omega_2 < \omega < \omega_3, \quad \omega_{2,3} \approx \frac{\omega_0' \mp \beta \sqrt{\omega_0'^2 - \omega_L^2}}{\sqrt{1-\beta^2}}. \quad (13)$$

This range can be called "intrinsic" with a certain degree of conditionality, since it includes the oscillator

frequency ω_0' as long as the velocity is comparatively low (at a sufficiently high β its lower edge rises above ω_0'). It should be emphasized that this frequency range exists in the radiation spectrum only on the condition $\omega_0' > \omega_L$. Note also that, if condition (11) is met, the resonance and intrinsic frequency ranges are spaced well away from each other.

With constraint (11), it is easy to estimate the total radiation power of the oscillator, which comprises two terms: $\Sigma = \Sigma_r + \Sigma_o$. For resonant radiation, the approximate calculation yields the following result:

$$\Sigma_r \approx \frac{p_0'^2\omega_0'^4}{8c^3}\frac{\Omega_L^2\Omega_r^4\beta(1-\beta^2)}{1-\beta^2+\beta^2\Omega_L^2}, \tag{14}$$

where $\Omega_L = \omega_L/\omega_0'$ and $\Omega_r = \omega_r/\omega_0'$.

The intrinsic radiation power of the oscillator can be approximately calculated as

$$\Sigma_o \approx \frac{p_0'^2\omega_0'^4}{4v^3}\left\{2\beta\sqrt{1-\Omega_L^2}\left[\Omega_L^2+\frac{\beta^2}{3}(2-5\Omega_L^2)\right]\right.$$
$$-\frac{\Omega_L^2}{2}(1-\beta^2)\left[(\Omega_L+\sqrt{1-\beta^2})^2\ln\frac{\Omega_3+\Omega_L}{\Omega_2+\Omega_L}\right. \tag{15}$$
$$\left.\left.+(\Omega_L-\sqrt{1-\beta^2})^2\ln\frac{\Omega_3-\Omega_L}{\Omega_2-\Omega_L}\right]\right\},$$

where $\Omega_{2,3} = \omega_{2,3}/\omega_0'$.

Note that formula (15) presents the exact expression for the total radiation power of a dipole oscillating in a cold plasma with permittivity $\varepsilon(\omega) = 1 - \omega_L^2/\omega^2$ [6]. This fact is explained by the remoteness of the intrinsic frequency range from the resonance frequency and the resulting closeness of the permittivity in this region to the permittivity of the plasma.

It is readily seen that, subject to condition (11), resonant radiation is much weaker than intrinsic radiation (certainly, if the latter exists, i.e., $\Omega_L < 1$). This is due to the smallness of term Ω_r^4 entering (14). The dependences of Σ_r and Σ_o on the dipole velocity appear to be different. The intrinsic radiation power monotonically decreases with β from the value

$$\Sigma_o\big|_{\beta=0} \approx \frac{p_0'^2\omega_0'^4}{3c^3}(1-\Omega_L^2)^{1/2}$$

to the value

$$\Sigma_o\big|_{\beta\to 1} \approx \frac{p_0'^2\omega_0'^4}{3c^3}(1-\Omega_L^2)^{3/2}.$$

The resonant radiation power turns to zero both at $\beta = 0$ and at $\beta \longrightarrow 1$ and peaks at a certain intermediate value. It seems important that, in the case $\Omega_L > 1$, intrinsic radiation vanishes, but resonant radiation takes place at any arbitrary relationship between the Langmuir frequency and the intrinsic frequency of the source.

3. THE CASE OF HIGH RESONANCE FREQUENCY

Let us now suppose that $\omega_0 \ll \omega_r$. This gives rise to two qualitatively different possibilities. The first one is realized if $\beta\sqrt{\varepsilon_0} < 1$, i.e., the dipole velocity v is lower than the phase velocity $c/\sqrt{\varepsilon_0}$ of the low-frequency radiation. In this case, the radiation spectrum includes both the intrinsic and resonant frequency range, the latter extending to frequency ω_r.

To conveniently estimate the boundary frequencies and radiation energies requires a number of additional assumptions. Thus, putting $\omega_0\sqrt{1-\beta^2} \ll \omega_r\sqrt{1-\beta^2\varepsilon_0}$ in the case of the resonant radiation, we obtain

$$\omega_1 < \omega < \omega_r, \quad \omega_1 \approx \omega_r\sqrt{\frac{1-\beta^2\varepsilon_0}{1-\beta^2}},$$
$$\Sigma_r \approx \frac{p_0'^2}{4v^3}(1-\beta^2)\omega_L^2\left\{(\omega_r^2+\omega_L^2)\right. \tag{16}$$
$$\left.\times[\ln(1-\beta^2)^{-1}-\beta^2]-\frac{\omega_L^2\beta^4}{2(1-\beta^2)}\right\}.$$

And assuming $\omega_0 \ll \omega_r(1-\beta\sqrt{\varepsilon_0})$ for the intrinsic radiation, we have

$$\omega_2 < \omega < \omega_3, \quad \omega_{2,3} \approx \frac{\omega_0}{1\pm\beta\sqrt{\varepsilon_0}},$$
$$\Sigma_o \approx \frac{p_0'^2\omega_0'^4}{3c^3}\sqrt{\varepsilon_0}\left(\frac{1-\beta^2}{1-\varepsilon_0\beta^2}\right)^3. \tag{17}$$

This expression for Σ_o is identical to the exact expression for the total radiation power of a dipole moving in a nondispersive medium with a refractive index ε_0 [2, 6]. It is naturally explained by the fact that the intrinsic range falls into the low-frequency region, where the permittivity is close to ε_0.

Note that, if the motion of the source is nonrelativistic ($\beta^2 \ll 1$), the resonant radiation power can be

approximated as

$$\Sigma_r \approx \frac{p_0'^2 \omega_r^2 \omega_L^2 \beta}{8c^3},$$

and the power ratio is defined as

$$\frac{\Sigma_r}{\Sigma_o} \approx \frac{3\omega_r^2 \omega_L^2 \beta}{8\omega_0'^4 \sqrt{\varepsilon_0}}. \quad (18)$$

Because of condition $\omega_r \gg \omega_0'$, resonant radiation may dominate over intrinsic radiation even in spite of the presence of small parameter β.

Now consider the case $\omega_0 \ll \omega_r$ and $\beta\sqrt{\varepsilon_0} > 1$, i.e., the oscillator velocity does not exceed the phase velocity of low-frequency waves. Then, the radiation spectrum consists of only one frequency range,

$$\omega_2 < \omega < \omega_r, \quad \omega_2 \approx \frac{\omega_0}{1+\beta\sqrt{\varepsilon_0}}. \quad (19)$$

The total radiation power can be approximated under additional condition $\omega_0 \ll \omega_r(1-\beta\sqrt{\varepsilon_0})$ as

$$\Sigma \approx \frac{p_0'^2}{4v^3}(1-\beta^2)\left[\omega_L^2(\omega_r^2+\omega_L^2)\ln\frac{\omega_r^2+\omega_L^2}{\omega_L^2} - \frac{\omega_r^4}{2}(1-\beta^2) - \omega_r^2\omega_L^2\right]. \quad (20)$$

As is evident from the above considerations, the radiation power is almost independent of the oscillation frequency under the restrictions imposed.

4. NUMERICAL RESULTS

Figures 1a–1d show the spectral energy density of radiation versus dimensionless frequency $\Omega = \omega/\omega_0'$ for various values of the resonance and Langmuir frequencies and different velocities of the source. No small parameter was implied: all calculations were carried out using exact formulas. All curves in each figure share the same resonance and Langmuir frequencies (given in the figure caption), the oscillator velocity is indicated by numbers by the curves.

Figure 1a shows the typical arrangement of curves in the case when both the resonance and Langmuir frequencies are lower than the oscillator frequency $\omega_r < \omega_0'$, $\omega_L < \omega_0'$. Two separate parts of the spectrum are seen: intrinsic radiation dominates and resonant radiation is only insignificant.

Figure 1b illustrates the situation when $\omega_r < \omega_0'$ and $\omega_L > \omega_0'$. It is only resonant radiation that exists in this case.

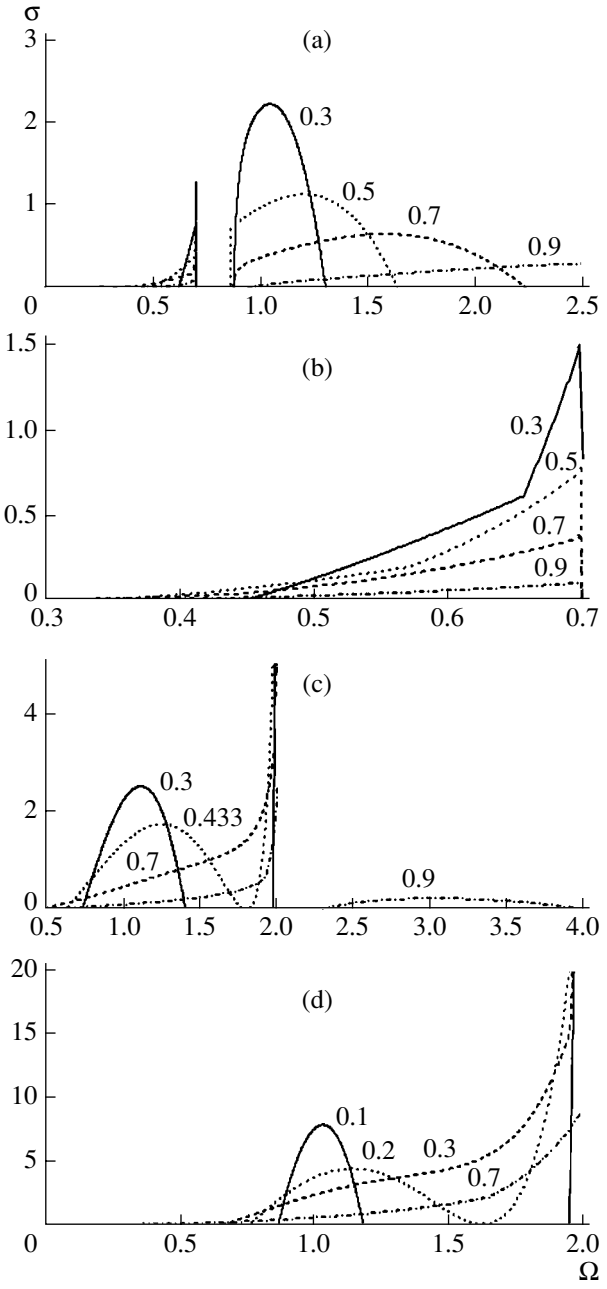


Fig. 1. Spectral power density (in the units of $\omega_0'^3 p_0'^2/3c^3$) vs. dimensionless frequency $\Omega = \omega/\omega_0'$ for Ω_r and Ω_L equal to (a) 0.7 and 0.5, (b) 0.7 and 2, (c) 2 and 0.5, and (d) 2 and 2, respectively. The number on each curve is the value of the oscillator velocity.

Figure 1c is typical of the case when $\omega_r > \omega_0'$ and $\omega_L < \omega_0'$. There are three possibilities here. If the dipole velocity is less than a certain β_* (for the parameters used in Fig. 1c, $\beta_* \approx 0.433$), the spectrum includes both the resonant and intrinsic (comparatively low-frequency) ranges. If $\beta_* < \beta < \beta_{**}$ (in our case, $\beta_{**} \approx$

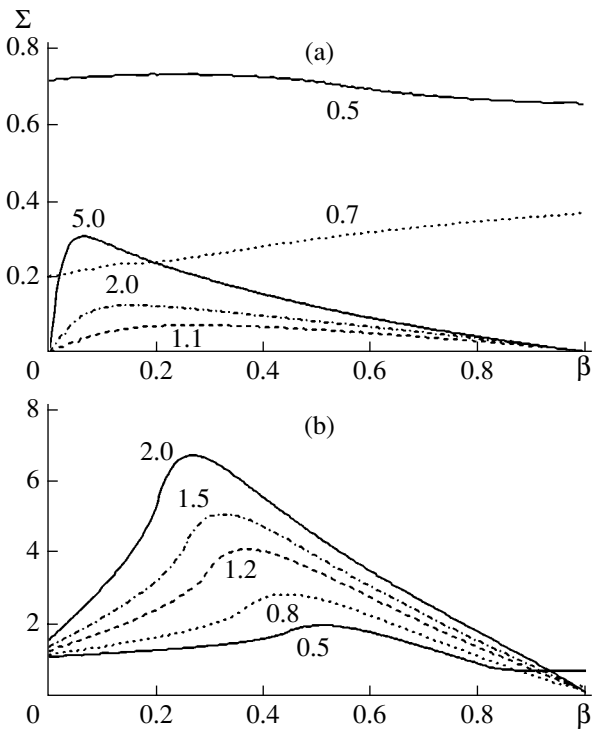


Fig. 2. Spectral power density (in the units of $\omega_0'^4 p_0'^2/3c^3$) vs. dipole velocity β for Ω_r = (a) 0.7 and (b) 2. The number on each curve is the value of the Langmuir frequency Ω_L.

0.82), the radiation exists only in one range covering the oscillator frequency and adjacent to the resonance frequency. If, finally, the oscillator velocity is so high that $\beta > \beta_{**}$, radiation appears in one more frequency range lying above the resonance frequency.

Figure 1d illustrates the case when $\omega_r > \omega_0'$ and $\omega_L > \omega_0'$. Two possible variants can be realized: with either two spectral ranges for relatively low velocity or only one spectral range otherwise.

Figures 2a and 2b present the total radiation power versus the dipole velocity for various values of the resonance and Langmuir frequencies. Figure 2a corresponds to the case when the resonance frequency is lower than the oscillator frequency ($\Omega_r < 1$). If the Langmuir frequency is also below the oscillator frequency ($\Omega_L < 1$), both the intrinsic and resonant parts of radiation exist and the former one dominates. In this situation, the dependence of power on the velocity appears to be insignificant. In contrast, at $\Omega_L > 1$, only relatively weak resonant radiation is generated and the velocity dependence with a peak at a certain value is fairly well pronounced.

Figure 2b corresponds to the case when the resonance frequency exceeds the frequency of the oscillator. The radiation power as a function of velocity shows a peak, which grows with the plasma frequency.

It should be emphasized in conclusion that the results of this study give evidence of a number of uncommon features that characterize radiation of an oscillator moving in a medium with resonant dispersion. Whether the radiation spectrum comprises two separate frequency ranges or they merge into a single range depends on the parameters of the problem. The dependence of the radiation power on the velocity of the source appears different for different ratios between the resonance frequency, Langmuir frequency, and the oscillator frequency. Comparing the results to their analogues obtained in the cases of nondispersive medium [2, 6] and cold plasmas [6], we find that the presence of resonant dispersion precludes the use of the same dependences but calls for the employment of some other substantially more complicated relationships to describe the radiation of an oscillator moving in such media.

REFERENCES

1. I. M. Frank, *Cherenkov Radiation* (Nauka, Moscow, 1988) [in Russian].
2. I. M. Frank, Izv. Akad. Nauk SSSR **6**, 3 (1942).
3. V. L. Ginzburg and I. M. Frank, Dokl. Akad. Nauk SSSR **56**, 583 (1947).
4. V. L. Ginzburg and I. M. Frank, Dokl. Akad. Nauk SSSR **56**, 699 (1947).
5. G. M. Garibyan and F. A. Kostanyan, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **14**, 1857 (1971).
6. A. V. Tyukhtin, Zh. Tekh. Fiz. **74** (8), 67 (2004) [Tech. Phys. **49**, 1021 (2004)].
7. P. Daly, K. S. H. Le, and C. H. Papas, IEEE Trans. Antennas Propag. **13**, 583 (1965).
8. B. M. Bolotovskiĭ and S. N. Stolyarov, in *Einstein Studies 1978–1979* (Nauka, Moscow, 1983), pp. 173–277 [in Russian].
9. E. G. Doil'nitsina, Yu. B. Zhuravlev, and A. V. Tyukhtin, Vestn. St. Petersb. Univ., Ser. 4: Fiz., Khim., No. 4, 44 (2001).
10. E. G. Doil'nitsina and A. V. Tyukhtin, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **46**, 21 (2003).
11. E. G. Doil'nitsina and A. V. Tyukhtin, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **47**, 301 (2004).
12. T. E. Stevens, J. K. Wahlstrand, J. Kuhl, *et al.*, Science **291**, 627 (2001).
13. I. Carusotto, M. Artoni, G. C. La Rocca, *et al.*, Phys. Rev. Lett. **87**, 064801 (2001).
14. M. M. Kash, V. A. Sautenkov, A. S. Zibrov, *et al.*, Phys. Rev. Lett. **82**, 5229 (1999).
15. G. N. Afanasiev and V. G. Kartavenko, J. Phys. D **31**, 2760 (1998).
16. G. N. Afanasiev, V. G. Kartavenko, and E. N. Magar, Physica B **269**, 95 (1999).
17. A. V. Tyukhtin, Pis'ma Zh. Tekh. Fiz. **30** (14), 68 (2004) [Tech. Phys. Lett. **30**, 605 (2004)].

Translated by A. Sidorova

Technical Physics, Vol. 50, No. 8, 2005, pp. 1089–1094. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 126–131.

SURFACE, ELECTRON AND ION EMISSION

Dynamics of Polarization Reversal in Thin PZT Films

V. M. Mukhortov, V. V. Kolesnikov, S. V. Biryukov, Yu. I. Golovko, and A. I. Maschenko

Prokhorov Institute of General Physics, Russian Academy of Sciences, ul. Vavilova 38, Moscow, 117333 Russia

e-mail: muhortov@aaanet.ru

Received June 30, 2004

Abstract—New experimental data on the dynamics of polarization reversal in thin PZT films prepared by rf sputtering are reported. The frequency and amplitude dependences of the hysteresis loop area and coercive field are investigated. It is shown that, at infralow frequencies, there exists a critical frequency at which the shape of the hysteresis loop changes stepwise. Theoretical analysis of the related problems is given. © *2005 Pleiades Publishing, Inc.*

INTRODUCTION

Recent extensive investigations into polarization reversal in thin ferroelectric films have been aimed at applying them in nonvolatile memory devices [1, 2], microelectromechanical systems [3, 4], high-frequency optical modulators [5], and controlled microwave devices [6–8]. The stringent requirements for the parameters of these devices stimulate, in turn, efforts to gain a deeper insight into the processes of polarization reversal and develop adequate theoretical models and approaches. From the practical standpoint, of primary importance is the frequency dependences of reversal, on which investigators have focused their attention in recent years. As a result, a number of fundamental physical effects have been found, in particular, those related to the frequency dependences of dielectric hysteresis loop area A and coercive field E_c in the range of low and infralow frequencies [9–13]. It has been shown by numerical calculations that quantities A and E_c as functions of frequency f of an external electric field have a power-type singularity. Such behavior is bound to take place in both magnets and ferroelectrics, although to date these features have been investigated at length only in magnets [9–13]. With an increase in the frequency, the area of hysteresis loops in thin magnetic films passes through a maximum, while at high frequencies, the loops collapse. In particular, this effect was observed in ferroelectric films [14]. Explanation of its physical nature is not only of applied significance. This effect is a key issue in the problem of polarization reversal and hysteresis. In this study, we analyze new experimental data concerning the amplitude–frequency dependence of hysteresis loops in ferroelectric films.

EXPERIMENTAL

Thin (~3-μm) PZT films ($Pb(Zr_{0.53}Ti_{0.47})O_3$) were produced by reactive rf sputtering of a stoichiometric target (for details of specimen preparation, see [15, 16]). Here, we only note that synthesis and crystallization of the films were performed during deposition by using a high oxygen pressure (0.7 Torr) in the chamber and a dense plasma with electron concentration $n_e > 10^{12}$ cm^{-3}. The 40-μm-thick metallic substrate (stainless steel mark 1Kh18N19T) was kept at 650°C, which provides the polycrystalline state the film crystallized. The top electrodes (1 mm^2 in area) were made from aluminum with a vanadium sublayer. X-ray analysis data indicated that foreign phases in the films were absent and the lattice parameter was c = 0.4057 nm. The dynamics of the reversal process was studied at different amplitudes (from 0.5 to 300 kV/cm) of a sawtooth voltage in the frequency range 10^{-3}–10^4 Hz with a modified Sawyer–Tower method. To provide a high input resistance necessary for infralow-frequency measurements, the signal from a reference capacitor (C_0 = 1.0 μF) was applied to a high-input-resistance (250 MΩ) unity-gain noninverting matching amplifier based on an AD743 op amp.

The exciting signal from a generator and the polarization signal picked up from the standard capacitor were applied (through a voltage divider and the amplifier, respectively) to two independent analog-to-digital converting channels of a data-acquisition (DAQ) board. At an exciting field frequency <5 Hz, a PCL-812 Advantech board was used; at a field frequency >5 Hz, an LA-1.5PCI-1.4 board. The conversion times of the signals were 10 μs and 1 ns, respectively. To ensure independent entry of the signals to the inputs of the ADCs and independent analog-to-digital conversion, two separate program flows in the software of the DAQ board driver were used. This made it possible to independently measure both the response to the field action and the polarization signal as functions of time.

1063-7842/05/5008-1089$26.00

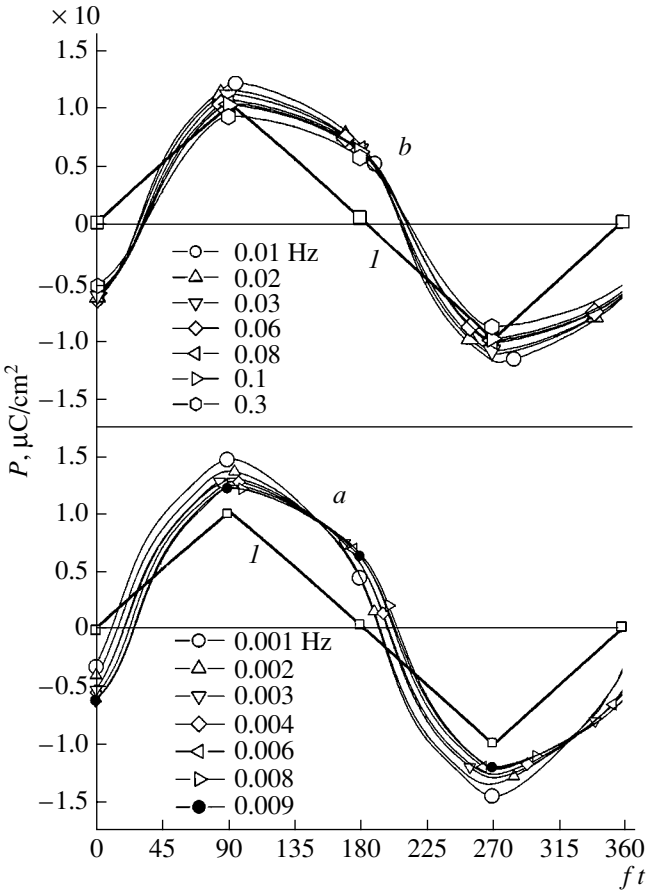


Fig. 1. Variation of the shape of polarization in the *t* representation for different frequencies: (*a*) low-frequency range, (*b*) high-frequency region, and (*1*) the applied field.

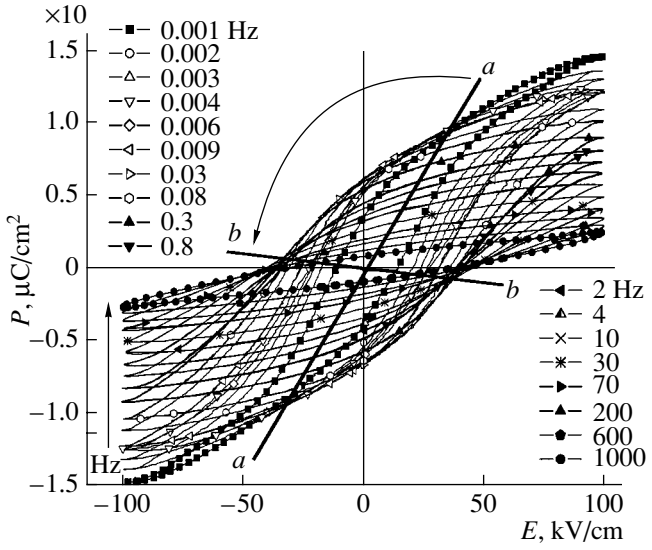


Fig. 2. Loops of dielectric hysteresis for different frequencies.

Such a modification of the Sawyer–Tower method made it possible to reveal an interesting feature in the frequency dependence of phase shift $\Delta\varphi$ of polarization response P on external field E. This feature is illustrated in Fig. 1, where the polarization response of the PZT film subjected to the external sawtooth voltage is plotted against time (in Fig. 1, the dimensionless time scale $\theta = ft$ is used). For frequencies lower or higher than critical frequency $f_{cr} = 0.0085$ Hz, a fixed point (*a* or *b*, respectively, in Fig. 1) exists on the P–θ plane. Going through the critical frequency is accompanied by a stepwise change in the position of the fixed point within range $\Delta f = 0.0005$ Hz.

The presence of the critical frequency in polarization reversal also shows up in the frequency dependence of the shape of the dielectric hysteresis loops (Fig. 2), which is similar to that in Fig. 1. Here, in either frequency range ($f < f_{cr}$ or $f > f_{cr}$), the shape of the loops changes in such a way that two fixed points, which are symmetric about the origin, appear on the (P–E) plane (in Fig. 2, they are connected by straight lines *a*–*a* and *b*–*b*). In going from one frequency range into the other, the position of these lines changes stepwise.

Figure 3 shows the variation of the hysteresis loop with external field amplitude at frequencies $f = 0.005$ Hz $< f_{cr}$ and $f = 8.4$ Hz $> f_{cr}$. In the frequency range studied, area $A = (f, E)$ of the loops increases monotonically with external field amplitude E and is approximated well by a power relationship of type $A = bE^s$ (Fig. 3). In passing through f_{cr}, exponent s changes jumpwise from $s = 2.15$ (at $f < f_{cr}$) to $s = 1.49$ ($f > f_{cr}$), remaining constant in either of these ranges within the experimental error.

The loop area and coercive field E_c as functions of frequency are plotted in Fig. 4 for two values of the external field amplitude. At certain frequency f_m, the curves $A(f)$ have a maximum. A similar situation also takes place in magnets [10–13]. When the external field amplitude decreases, frequency f_m shifts toward the low-frequency range, approaching f_{cr}. As the frequency drops, the remanent polarization increases, whereas the coercive field declines, tending to $E_c = 5$ kV/cm. The frequency dependence of E_c is characterized by a smooth maximum at a frequency slightly exceeding f_m. As the external field decreases, the maximum of E_c also shifts toward lower frequencies.

ANALYSIS AND DISCUSSION

Prior to analyzing the features described above, let us first consider the physics of the point at issue. According to the present-day concepts, polarization reversal in ferroelectrics is categorized as the first-order phase transition induced by an external electric field, as a result of which the volumes of two phases corresponding to different polarization directions change [17–21]. External field E is responsible for overcooling

or supersaturation in ordinary first-order phase transitions. In the case of polarization reversal, the critical parameter is coercive field E_c. Therefore, to describe hysteresis effects in ferroelectrics, theoretical models based on the classical nucleation concept [20] (such as the Kolmogorov–Avrami theory [1, 17–19] and its contemporary modifications [17–19]) are routinely used. However, application of these approaches to analyzing reversal processes at $|E| > E_c$ has not been validated in full measure. Under cyclic reversal with frequency f, parameter E falls into two ranges: $|E| < E_c$ (a two-phase range, where a new phase nucleates and grows) and $|E| > E_c$ (stable nuclei cannot exist). In the latter case, the phase with the initial polarization direction is absolutely unstable and the evolution of nuclei emerging at the first stage ($|E| < E_c$) carries the system to a complicated spatially nonuniform state. This state is bound to depend on a relationship between frequency f of the external field and the characteristic relaxation times of polarization in the ferroelectric. Because of this, a more general approach to describing hysteresis phenomena in magnets was proposed [10–14].

The numerical data obtained in [10–14] were treated in terms of the theory of dynamic scale invariance [21], which equally applies to magnets and ferroelectrics. According to this theory, the dependences of the physical characteristics of a system undergoing a phase transition on the frequency and amplitude of the external field must exhibit power-type singularities in the neighborhood of critical points. The corresponding exponents are called critical exponents. In our case, an important consequence following from this model is slowing-down of the phase transition near critical point E_c, which makes it possible to estimate the characteristic relaxation time of polarization versus the field magnitude. Let us demonstrate this effect with the purely dynamic approach (in the absence of fluctuations) proposed in [21] to describe the reversal process in ferroelectrics. In the case of the spatially uniform state and constant electric field $E > E_c$, characteristic relaxation time $\tau_r \sim \tau_0(E - E_c)^{-\beta}$, where $\beta = 0.5$. In the two-phase range ($E < E_c$), the characteristic time depends on the evolution time of a nucleus with a new polarization direction. In this case, $\tau_r \sim \tau_0(E_c - E)^{-\beta}$ with the same β ($\beta = 0.5$). In the combined form,

$$\tau_r \sim \tau_0|E - E_c|^{-\beta}, \quad (1)$$

where τ_0 is a constant independent of field E.

The presence of a singularity in this expression at $E = E_c$ (at $1/\tau_r$, it has the form of a bird's beak) leads to a sharp increase in relaxation time τ_r of polarization in the vicinity of E_c. There lies the above-mentioned effect of critical slowing-down of polarization reversal. Polarization reversal in this case is characterized by dimensionless parameter $\eta = f\tau_r(E)$ depending on both frequency f and field E. In view of the fact that parameter τ_0 in expression (1) is independent of field E, it can be found from independent experiments on linear response detection (for example, from the data on the frequency dispersion of the permittivity). The effect of critical slowing-down results in a power-type singularity in the frequency dependence of the coercive field at low frequencies, $\eta = f\tau_r \ll 1$,

$$\Delta E_c \sim f^{\gamma}. \quad (2)$$

In this expression, ΔE_c is a change in the coercive field with respect to its static value E_c specified by the equation of state of a ferroelectric [21]. In the absence of fluctuations, $\gamma = 2/3$; the same exponent also

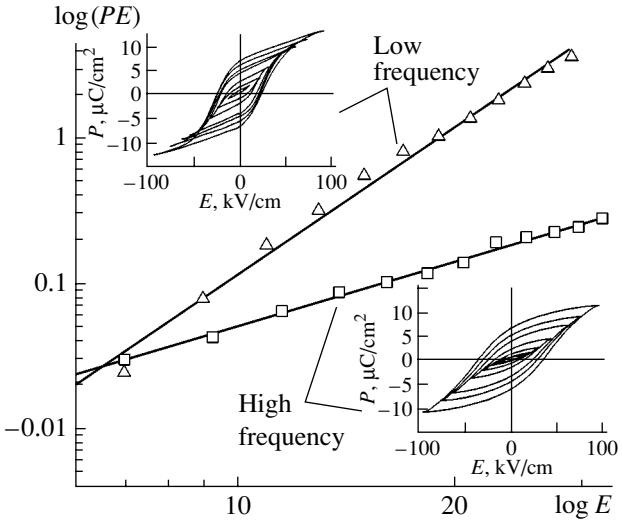


Fig. 3. Variation of the area and shape of dielectric hysteresis vs. the external electric field in the low- (0.005 Hz) and high-frequency (8.4 Hz) ranges.

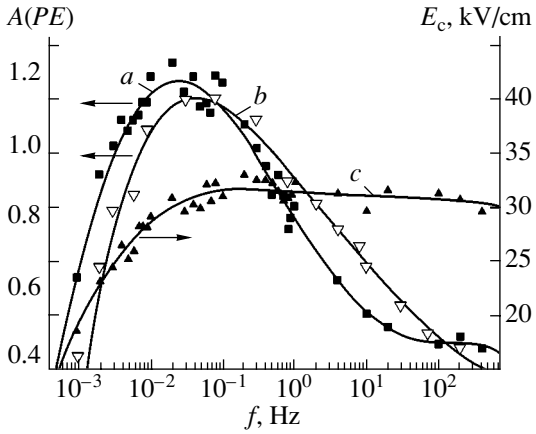


Fig. 4. Frequency dependence of the dielectric hysteresis loop area for an external field of (*a*) 54 and (*b*) 96 kV/cm. (*c*) Frequency dependence of the coercive field for an external field of 96 kV/cm.

describes loop area A as a function of the frequency [22]. It follows from expression (2) that, at low frequencies, the hysteresis loop in the neighborhood of E_c has a near-vertical portion, which is responsible for a sharp maximum of the switching current and rectangular hysteresis loops, as was observed, for example, in barium titanate single crystals [23]).

However, it is known that, along with nonlinearity, fluctuations (temperature-induced and static spatial nonuniformities of the electric field and polarization, which are observed, for example, in polycrystals) play an important role near the critical point. Fluctuations flatten out the singularity of type (1) and change the values of critical indices. For example, from numerical calculations performed for magnets [10–14] where fluctuations were taken into account, it follows that characteristic reversal time τ_r is also described by expression (1) but with $\beta = 1$, as in the experimental dependence for the rate of spin relaxation.

In [12–14], calculated frequency dependences of hysteresis loop area A were processed with scaling function $W(\eta)$ related to A as

$$A = A_0\tau_0/\tau_r W(\eta), \quad (3)$$

where

$$\eta = f\tau_r, \quad \tau_r = \tau_0 C/E^\beta. \quad (4)$$

Here, A_0 and C are constants that have the asymptote $W \sim \eta^\alpha$ with $\alpha = 1/3$ and $\beta = 1$ in the range of low ($\eta \ll 1$) frequencies and $W \sim \eta^{-1}$ in the range of high frequencies ($\eta \gg 1$).

In expressions (3) and (4), E is the external field amplitude and $\tau_r = \tau_0(C/E)$ specifies the characteristic relaxation time of polarization in ferroelectrics. It meets expression (1) with $\beta = 1$ at fields far exceeding E_c. Numerical calculations show that constant C is very large, $C = 10^4$–10^5 (amplitude E is normalized to coercive field E_c), and $\tau_r \gg \tau_0$, which reflects the critical slowing-down of the phase transition. From the asymptote of scaling function $W(\eta)$, it follows that the amplitude–frequency dependence of the hysteresis loop area is given by expressions

$$A \sim E^{2/3} f^{1/3}, \quad (5)$$

$$A \sim E^2/f \quad (6)$$

at low and high frequencies, respectively. At low frequencies, the exponent in the frequency dependence of the area is $\alpha = 0.33$. Processing of our experimental dependences $A(f)$ for PZT films yields α in the range from 0.246 to 0.417. Its mean value is in good agreement with the analytical data. It has already been noted that, in this frequency range, one can expect a near-vertical portion of the hysteresis loop near the coercive field and a value of exponent γ (see expression (2)) that is close to exponent $\alpha \approx 0.33$. It should be remembered that these exponents appear in the frequency dependences of the coercive field and loop area, respectively. Similar behavior has been observed in a number of experiments. According to [24], for strontium tantalate–bismuth tantalate solid solutions, $\gamma = 0.3$–0.4. However, the experimental results for the same solid solution presented in [1] give a much smaller value of γ, $\gamma \approx 0.12$. Note that, in those experiments, the range of "low-frequency" power-type growth of the coercive field is observed in the range 10–100 MHz. This can be explained by very small values of parameter τ_0 in the works cited. According to [21], the ultimate value of τ_0 for an "ideal" electric circuit can be estimated from the expression $\tau_0 \sim \varepsilon/f_{lat}$, where ε is the permittivity and f_{lat} is the characteristic frequency of lattice vibrations in a ferroelectric. Putting $\varepsilon = 10^2$–10^4 and $1/f_{lat} = 10^{-12}$–10^{-13} s, we find that the ultimate value of τ_0 ranges from 10^{-11} to 10^{-8} s. In our experiments, the value of $1/\tau_0$ was not too large (1–10 kHz). This, as well as the effect of critical slowing-down, leads to a shift of the maximum in the frequency dependence of the hysteresis loop area, $A(f)$, toward infralow frequencies. It appears that the condition $\gamma \approx \alpha \approx 0.33$ is violated in our experiments for the same reason: exponent γ in the frequency dependence of the coercive field is much smaller, $\gamma \approx 0.17$ at $f < 0.02$ Hz. In other words, because of large values of parameter τ_0, the condition of low frequencies ($\eta = f\tau_r \ll 1$), strictly speaking, is not fulfilled. This conclusion is confirmed, in particular, by the absence of a vertical portion in the hysteresis loops near E_c, although the role of field spatial nonuniformity in thin polycrystalline films should not be ruled out here. Studies are now under way to clarify this point.

According to relationships (5) and (6), a sharp increase in the hysteresis loop area at low frequencies changes to a decrease in this area at high frequencies by the law $A \sim E^2/f$. Quantity A exhibits a maximum, its position depending on field amplitude E. It should be noted that a maximum in the frequency dependence of the loss ellipse area is also observed in linear insulators. However, the run of the curves $A(f)$ in these two cases differs radically: at low frequencies, $A(f)$ for ferroelectrics and ferromagnets has a power-type singularity, $A(f) \sim f^\alpha$ ($\alpha < 1$), while for linear insulators, $A(f)$ linearly increases with the frequency, $A(f) \sim f$. Physically, this is because effects like critical slowing-down are absent in linear systems. At the same time, in nonlinear systems, the transition from low frequencies (the range of critical slowing-down) to high frequencies (the range of the quasi-linear behavior) is accompanied by a radically new effect. This effect consists in a qualitative change in the shape of limiting cycles on the P–E plane (experimentally, these cycles are hysteresis loops). At low frequencies, there is one stable cycle enclosing the origin. As the frequency grows (above a certain critical value), there appear two cycles lying in the upper and lower half-planes and enclosing the stable points of

spontaneous polarization $+P_s$ and $-P_s$. A similar effect observed experimentally in magnetic Co films on copper substrates [25, 26] was speculatively related to the dynamic phase transition [27, 28]. Although the physics of this effect has been studied largely in magnets, it is expected to take place in ferroelectrics as well, as indicated, e.g., by the variation of limiting cycles with frequency in simple dynamic models of ferroelectrics [21] and magnets [27, 28] in which only the nonlinearity of the equation of state of a solid (in the absence of fluctuations) is taken into account. The stepwise variation in the shape of the hysteresis loops (the fixed points in Figs. 1 and 3) also suggests that the mechanism of polarization reversal in ferroelectrics changes radically when the frequency exceeds a certain critical value. However, further experimental and theoretical investigations are needed to gain better insight into the physical nature of the effects described above and clarify their relation to nonlinearity in the equations of motion and fluctuations.

Comparing the theoretical and experimental dependences of the hysteresis loop area on the field amplitude leads us to the same conclusion. Expressions (3) and (4) imply that these are power-type dependences of type $A \sim E^s$ with strongly differing exponents: $s \approx 0.67$ and 2 at high and low frequencies, respectively. Our experimental data also indicate that s changes from 2.15 (at $f < f_{cr}$) to 1.49 (at $f > f_{cr}$) (see Fig.2). However, in the experiments, exponent s varies in a radically different manner: the lower the frequency, the larger s. Of even greater importance is the fact that the exponent exhibits a step in going through the critical value of the frequency, $f_{cr} = 0.0085$ Hz, remaining virtually constant on the right and on the left of f_{cr}. The hypothesis for dynamic phase transition described above basically predicts such behavior of the physical characteristics of polarization reversal. However, a number of issues still remain unclear and tackling them call for detailed insight into the physics of hysteresis phenomena.

CONCLUSIONS

The basic results of our study are as follows.

(1) The frequency dependence of hysteresis loop area A (losses per cycle) is characterized by a sharp rise at infralow frequencies, $f < 10^{-2}$ Hz. As the frequency increases, it changes to a smooth decay. Irrespective of external field amplitudes, the maximum of the curve $A(f)$ is at $f_m \approx 0.08$ Hz.

(2) At certain critical frequency $f_{cr} \approx f_m$, the frequency dependences of the hysteresis loop shape exhibit steps.

(3) Comparative analysis of hysteresis effects in magnets and ferroelectrics in terms of the dynamic scale theory indicates that the sharp growth of the loop area at infralow frequencies is due to the critical slowing-down of the process of polarization reversal, which causes a power-type singularity in the dependence $A(f) \sim f^{\alpha}$. The experimental value of the exponent agrees qualitatively with its theoretical value, $\alpha \approx 0.33$. The transition from low frequencies to high frequencies is expected to be accompanied by a qualitative change in the shape of hysteresis loops and, eventually, by their collapse. In a number of works, this effect is associated with the dynamic phase transition, which may also be responsible for the observed steps in the frequency dependences of the hysteresis loop shape near critical frequency f_{cr}. However, relevant investigation is still in its infancy.

REFERENCES

1. J. F. Scott, Ferroelectr. Rev. **1**, 1 (1998).
2. Y. Shimada, K. Nakao, A. Inoue, *et al.*, Appl. Phys. Lett. **71**, 2538 (1997).
3. C. A.-P. De Araujo, J. D. Cuchiaro, L. D. McMillan, *et al.*, Nature **374**, 627 (1995).
4. J. F. Scott and C. A. Araujo, Science **246**, 1400 (1989).
5. D. M. Gill, C. W. Conrad, and G. Ford, Appl. Phys. Lett. **71**, 1783 (1997).
6. P. Padmini, T. R. Taylor, M. J. Lefevre, *et al.*, Appl. Phys. Lett. **75**, 3186 (1999).
7. O. G. Vendik, E. K. Hollmann, A. B. Kozyrev, *et al.*, J. Supercond. **12**, 145 (1999).
8. O. G. Vendik, I. G. Mironenko, and L. T. Ter-Martirosyan, J. Phys. **33**, 277 (1972).
9. J. F. Scott, *Thin Film Ferroelectric Materials and Devices* (Kluwer, Boston, 1997), pp. 115, 145–151.
10. M. Rao, H. R. Krishnamurthy, and R. Pandit, Phys. Rev. B **42**, 856 (1990).
11. M. Rao and R. Pandit, Phys. Rev. B **43**, 3373 (1991).
12. J.-M. Liu, H. P. Li, C. K. Ong, *et al.*, J. Appl. Phys. **86**, 5198 (1999).
13. J.-M. Liu, H. L. Chan, *et al.*, Phys. Rev. B **65**, 1 (2001).
14. J.-M. Liu, Q. C. Li, W. M. Wang, *et al.*, J. Phys.: Condens. Matter **13**, L153 (2001).
15. V. M. Mukhortov, Yu. I. Golovko, G. N. Tolmachev, *et al.*, Zh. Tekh. Fiz. **69** (12), 87 (1999) [Tech. Phys. **44**, 1477 (1999)].
16. V. M. Mukhortov, Yu. I. Golovko, G. N. Tolmachev, *et al.*, Ferroelectrics **247**, 75 (2000).
17. S. A. Kukushkin and A. V. Osipov, Fiz. Tverd. Tela (St. Petersburg) **43**, 80 (2001) [Phys. Solid State **43**, 82 (2001)].
18. S. A. Kukushkin and A. V. Osipov, Fiz. Tverd. Tela (St. Petersburg) **43**, 88 (2001) [Phys. Solid State **43**, 90 (2001)].
19. M. A. Zakharov, S. A. Kukushkin, and A. V. Osipov, Fiz. Tverd. Tela (St. Petersburg) **46**, 1238 (2004) [Phys. Solid State **46**, 1277 (2004)].
20. E. M. Lifshitz and L. P. Pitaevskii, *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon, Oxford, 1981).

21. V. V. Kolesnikov, A. T. Kozakov, and A. V. Nikol'skii, Fiz. Tverd. Tela (St. Petersburg) **42**, 141 (2000) [Phys. Solid State **42**, 146 (2000)].
22. G. H. Goldsztein, F. Broner, and S. H. Strogatz, J. Appl. Math. **57**, 1163 (1997).
23. G. A. Smolenskiĭ, V. A. Bokov, V. A. Isupov, N. N. Kraĭnik, R. E. Pasynkov, and M. S. Shur, *Ferroelectrics and Antiferroelectrics* (Nauka, Moscow, 1971), p. 476 [in Russian].
24. C. Kuhn, H. Hoingschmid, O. Kowarik, *et al.*, in *Proceedings of the International Symposium on the Application of Ferroelectrics, Honolulu (USA), 2000.*
25. Q. Jiang, H.-N. Yang, and G.-C. Wang, Phys. Rev. B **52**, 14911 (1995).
26. Q. Jiang, H.-N. Yang, and G.-C. Wang, J. Appl. Phys. **79**, 5122 (1996).
27. S. W. Sides, P. A. Rikvold, and M. A. Novotny, Phys. Rev. E **59**, 2710 (1999).
28. T. Tome and M. J. de Oliveira, Phys. Rev. A **41**, 4251 (1990).

Translated by Yu. Vishnyakov

Technical Physics, Vol. 50, No. 8, 2005, pp. 1095–1097. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 132–135.

SHORT COMMUNICATIONS

Phase Transitions in the CN_x–Co System Induced by Changes in the Film Growth Temperature

Z. A. Samoilenko, E. I. Pushenko, N. N. Ivakhnenko, V. N. Varyukhin, and E. I. Shemchenko

Galkin Physicotechnical Institute, National Academy of Sciences of Ukraine, Donetsk, 83114 Ukraine

Received October 8, 2004

Abstract—The structures of amorphous CN_x–Co films grown at different temperatures (T_s = 200–365°C) are studied by X-ray diffraction. As the growth temperature increases above T_s = 200°C, a concentration phase transition is found to occur in the amorphous state; this transition is related to a change in the major portion of carbon or cobalt in the structure of the cluster films. At T_s = 365°C, a disorder–order phase transition, which is accompanied by the transition from the amorphous to crystalline state, occurs in the films. © *2005 Pleiades Publishing, Inc.*

INTRODUCTION

Interest in composite materials consisting of a semiconductor matrix with embedded conducting nanoparticles has grown in the past few years. Nitrogen alloying of a carbon film not only transforms it into a semiconductor but also preserves all the advantages of diamond-like carbon [1, 2]. Moreover, the nitrogen concentration affects the fundamental properties of the film, such as the bandgap width and the impurity activation energy, and induces local conduction zones [3, 4]. Thus, depending on the crystal structure of a carbon-nitride-based composite, one can produce materials combining the semiconductor properties of carbon nitride and the properties of diamond-like carbon [4, 5]. Therefore, the purpose of this work is to study the crystal structure of CN_x–Co films.

EXPERIMENTAL

Diamond-like carbon nitride films with a metallic impurity were deposited in a VUP-5M vacuum device by ion-plasma sputtering of a graphite target in a nitrogen plasma. We used a dc magnetron with the induction stabilization of a glow discharge. As the target, we used a plate of pure graphite having a staggered impurity consisting of a powdered metal oxide. The substrate (thin glass) was placed onto an electrically insulated radiation heater, whose temperature was controlled by a chromel–alumel thermocouple to an accuracy of 5°C. Thus, we prepared homogeneous CN_x–Co films grown at different temperatures and at the same impurity concentration.

Using diffuse X-ray scattering, we studied the structure of the CN_x–Co films synthesized at T_s = 200–365°C. To this end, we applied long-wavelength CrK_α radiation with a V filter and a photodiode sensitive to weak diffuse scattering of low-intensity rays. X-ray diffraction patterns were recorded using the reflection scheme at two angles of inclination of the samples (25° and 155°). The exposure time was 20 min, and the X-ray diffraction patterns were subjected to microphotometering in 0.5° steps (Fig. 1).

RESULTS AND DISCUSSION

Amorphous CN_x–Co films are characterized by a substantially inhomogeneous structure: the medium of random atoms contains mesoscopic clusters with an incipient long-range order [6]. A similar situation is typical of many materials [7]. The introduction of impurities in the films is accompanied by the formation of impurity clusters in the matrix. Cluster transformations detected by us in materials of different chemical compositions suggest that, apart from the formation of impurity clusters [8, 9], the alloying of amorphous materials can also be accompanied by structural phase transitions. We detected such a structural transition in the amorphous CN_x–Co films when studying the effect of the growth temperature T_s on their structure. Figure 1 shows X-ray diffraction patterns taken from CN_x–Co films grown at different temperatures T_s. Two diffusion maxima (halo I and II) are visible in the X-ray diffraction patterns of the films; they are mainly caused by coherent X-ray scattering by clusters of two types. The appearance of a halo as the main diffraction element means that the films have an amorphous structure formed by mesoscopic atomic groups (clusters) with an incipient long-range order and by a random intercluster medium (which manifests itself in an X-ray diffraction pattern as a background whose intensity is proportional to $\sim\sin^2\theta$). As T_s increases, the film structure changes substantially. As follows from Fig. 1, this increase is accompanied by changes in the shape and ratio of halo intensities I_1 and I_2 (Fig. 2a), which indicates continu-

1063-7842/05/5008-1095$26.00

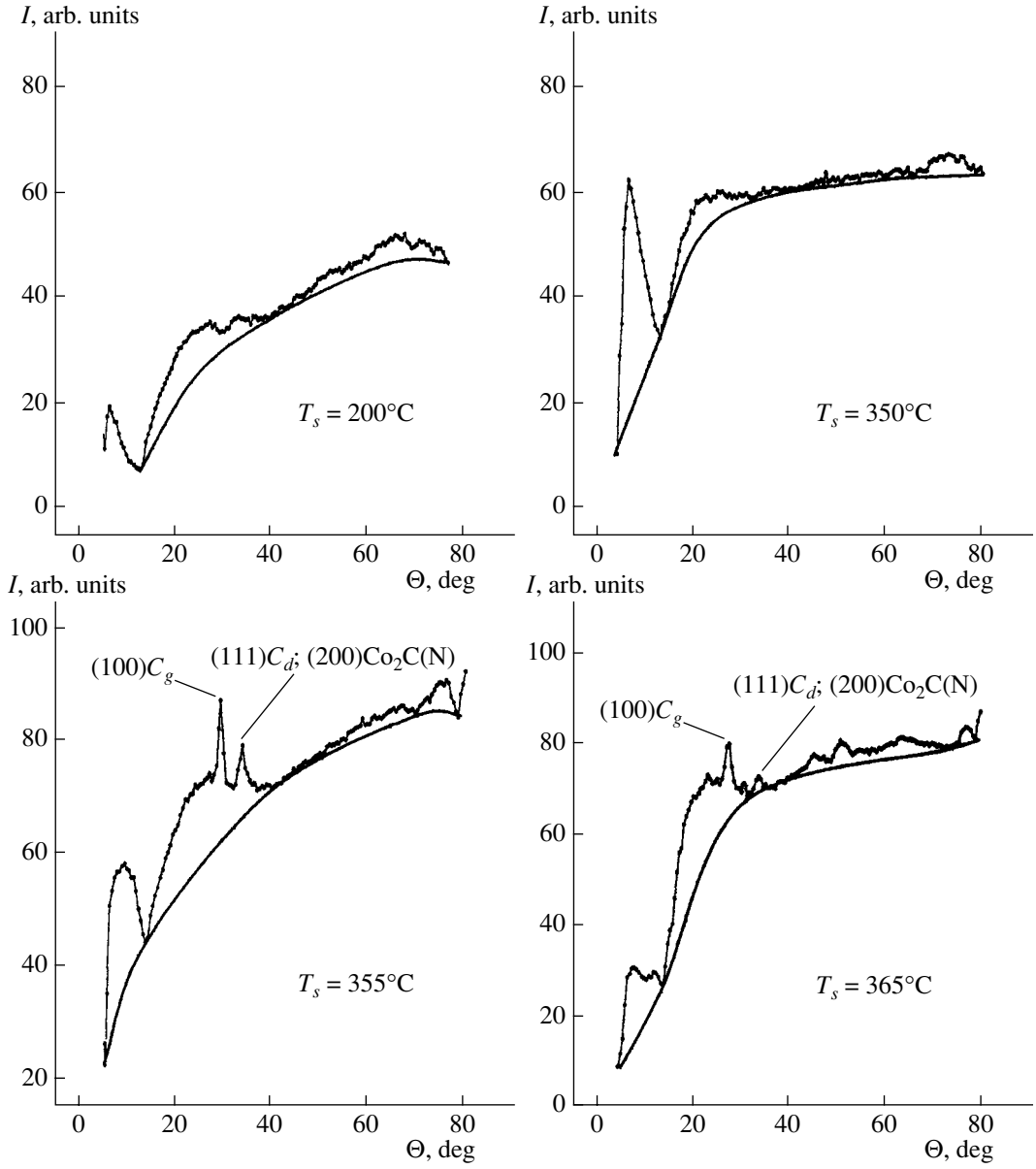


Fig. 1. X-ray diffraction patterns of CN_x–Co films grown at different temperatures T_s: (100)C_g belongs to the graphite phase, and (111)C_d is a diamond-like phase. Θ is the scattering angle.

ous evolution of the structure with increasing T_s. Based on the structures of semiconductor materials doped with metallic atoms [10], we attributed the first halo to a matrix carbon-based phase and the second halo to an impurity cobalt-rich phase. As the growth temperature T_s changes, these mesostructural phase states compete with each other, which is indicated by a change in the intensity ratio of halos I and II (Figs. 1, 2a). This competition results in a concentration phase transition for the major portion of carbon (the base phase, which mainly generates halo I) or cobalt (the impurity phase, which mainly generates halo II) in the cluster structure:

impurity phase ⟶ base phase ⟶ impurity phase
200°C 300°C 355°C

Figure 2b shows the temperature dependences of the relative intensities of coherent (I_{coh}/I) and incoherent (I_{incoh}/I) scattering. Before the amorphous–crystalline phase transition at T_s = 360°C, the incoherent scattering intensity increases (Fig. 2b, curve *1*) and the coherent scattering intensity decreases thereby (Fig. 2b, curve *2*). This behavior indicates a decrease in the order parameter in the system during the disorder–order transformation in the structure, namely, during the formation of a crystalline phase, which is supported by the emergence of lines against the background of the halo at T_s = 350–365°C (Fig. 1).

By measuring the halfwidths β of the fluctuation maxima and the angles Θ from Fig. 1 and knowing the

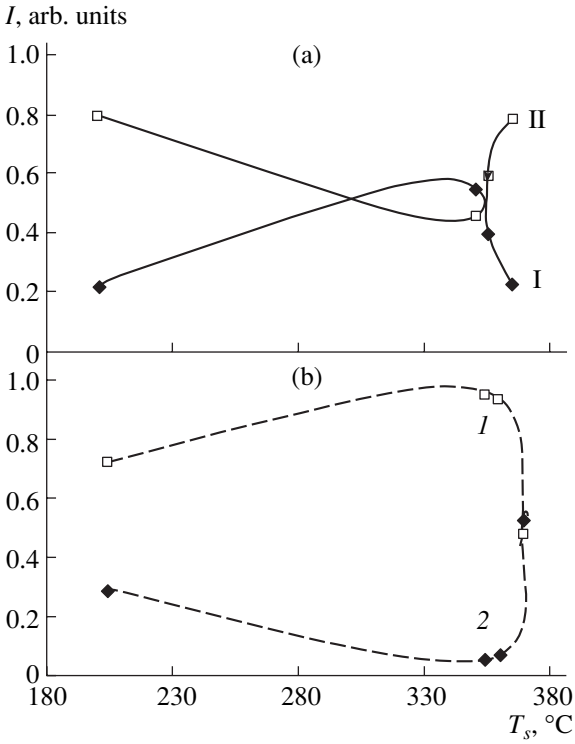


Fig. 2. Intensity as a function on the growth temperature: (a) curve I is I_1/I_{1+2} for halo I; curve II is I_2/I_{1+2} for halo II; (b) (*1*) I_{incoh}/I and (*2*) I_{coh}/I.

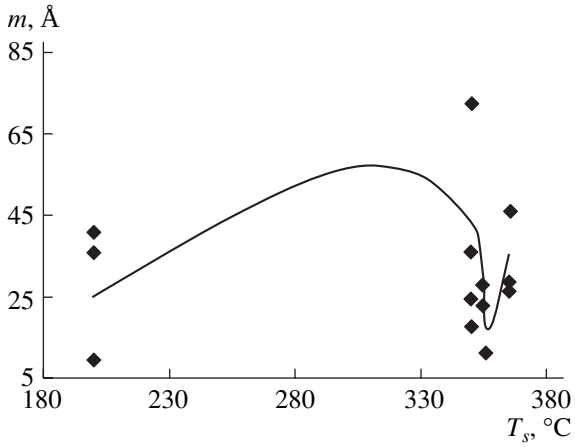


Fig. 3. Dependence of the cluster size on the growth temperature.

wavelength λ for CrK_α radiation, we used the Kitaĭgorodskiĭ technique [11] to estimate the linear dimensions of clusters in Å: $m = \lambda/4\pi\eta k$. Here, η is a dimensionless coefficient that relates the cluster dimensions to the width of a diffusion maximum with allowance for the geometrical parameters of the X-ray camera used [11]; η is determined from the formula $B/D = 19\eta/(0.004 + 0.084\cos\theta) + 0.0064\theta$, where B is the width of the diffusion maximum (deg) and D is the diameter of the focus spot on the sample (mm). To plot the dependence of the cluster size on the growth temperature, we used averaged cluster sizes for each temperature (Fig. 3). A nonmonotonic character of the $m = f(T_s)$ dependence indicates structural evolution. The minimum on the curve confirms the presence of a phase transition in the structure [11]. Upon growth at $T_s = 200$°C, the sample has an amorphous structure, and, as the temperature increases to $T_s = 365$°C, the structure becomes crystalline with nanocrystals of graphite and diamond-like phases (the latter phase is alloyed by cobalt).

CONCLUSIONS

X-ray diffraction studies of CN_x–Co films have shown that they exhibit two comparable halos and diffraction lines appearing against the background of the second halo at a growth temperature $T_s > 360$°C. These results indicate the formation of a two-phase nanostructure in the films with an ordered arrangement of chemical elements (which is similar to heterophase separation) in the form of 20–50 Å clusters of base ($CN + Co_y$) and impurity ($Co + CN_y$) phases.

We have revealed concentration and crystallographic phase transitions that occur in the CN_x–Co films at a growth temperature $T_s > 200$°C.

REFERENCES

1. V. I. Ivanov-Omskiĭ, S. G. Yastrebov, A. O. Golubok, *et al.*, Pis'ma Zh. Tekh. Fiz. **24** (20), 186 (1998) [Tech. Phys. Lett. **24**, 800 (1998)].
2. E. A. Smorogonskaya, T. K. Zvonareva, E. I. Ivanova, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **45**, 1579 (2003) [Phys. Solid State **45**, 1658 (2003)].
3. I. Hu, P. Yang, and C. M. Lieber, Phys. Rev. B **57**, 3185 (1998).
4. D. Li, Y. Chung, and M. J. Wong, Appl. Phys. **74**, 219 (1993).
5. H. Sjostrom and S. Stafstrom, Phys. Rev. Lett. **75**, 1336 (1995).
6. V. D. Okunev and Z. A. Samoĭlenko, Pis'ma Zh. Éksp. Teor. Fiz. **43**, 24 (1986) [JETP Lett. **43**, 28 (1986)].
7. A. S. Bakaĭ, *Polycrystalline Amorphous Solids* (Énergoizdat, Moscow, 1988) [in Russian].
8. V. D. Okunev and Z. A. Samoĭlenko, Fiz. Tverd. Tela (Leningrad) **13**, 257 (1989) [Sov. Phys. Solid State **31**, 1623 (1989)].
9. V. D. Okunev and Z. A. Samoĭlenko, Pis'ma Zh. Éksp. Teor. Fiz. **53**, 42 (1991) [JETP Lett. **53**, 44 (1991)].
10. M. A. Krivoglaz, *Electronic Structure and Electronic Properties of Metals and Alloys* (Naukova Dumka, Kiev, 1988), pp. 3–39 [in Russian].
11. L. I. Kitaigorodskii, *X-ray Diffraction Analysis of Fine-Crystalline and Amorphous Solids* (Nauka, Moscow, 1952) [in Russian].

Translated by K. Shakhlevich

Technical Physics, Vol. 50, No. 8, 2005, pp. 1098–1101. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 136–139.

SHORT COMMUNICATIONS

Mesoscopy of Cluster Gases

L. I. Kurlapov

e-mail: lkurlapov@yandex.ru

Received October 29, 2004

Abstract—The properties of gases in the near-critical region acquire the features of the liquid state; in other words, mesoscopy of a dense gas is manifested. In the cluster model [3], mesoscopy is explained by the effect of large clusters. The results of calculation of the temperature dependence of viscosity agree with experimental data in the prediction of mesoscopy, which was also observed in the pressure dependence of the thermodiffusion baric effect (the direction of this effect at high pressures corresponds to the liquid state). The phenomena associated with a change in the properties of a gas without the emergence of an interface should be referred to as a mesoscopic phase transition. © *2005 Pleiades Publishing, Inc.*

Under various conditions in gases, formations consisting of several molecules [1, 2] may exist, which are treated as subcomponents of a cluster mixture in the cluster model [3]. An essential reason for deviation of the properties of real gases from ideality is the changes in the cluster composition upon a change in macroparameters, since the cluster formation or disintegration corresponds to a change in the number of moles and the molar mass of this mixture. Each cluster subcomponent makes a contribution to transport processes, which is determined by the mass and collision cross section of clusters. In this communication, it is shown that large clusters existing in the near-critical region are mesoscopic particles and that the entire cluster mixture exhibits its intermediate (mesoscopic) properties between a gas and a liquid.

Under the local thermodynamic equilibrium conditions for each locally equilibrium domain in a gas, a certain size distribution of clusters (determined by the number of molecules constituting the clusters) emerging in the gas can be described by an exponential function [3],

$$n_g = n_1 \exp[-G(g-1)], \tag{1}$$

where n_g is the number density of clusters containing g molecules, n_1 is the number density of molecules as monomers, and G is the normalization factor.

To describe transport processes in a cluster mixture, it is convenient to characterize the composition of a cluster mixture by the numerical fraction relative to the total number density of all clusters,

$$C_g^{(c)} = C_1^{(c)} \exp[-G(g-1)],$$
$$C_g^{(c)} = \frac{n_g}{n^{(c)}}, \quad n^{(c)} = \sum_{g=1}^{r} n_g, \tag{2}$$

where $C_1^{(c)}$ is the concentration of molecules (monomers) and r is the size of the largest cluster, which is taken into account in calculations.

The application of conventional relations for a cluster mixture leads to an algorithm for calculating concentration $C_g^{(c)}$ using tabulated data on specific volume [4] and effective diameters of particle collisions. In this communication, we use the effective diameters of collisions determined by approximating the viscosity of a rarefied gas [4] by a fifth-degree polynomial and by calculating effective diameters by the formula for the viscosity of a rarefied gas in the kinetic theory. The computational algorithm implies that clusters beginning with dimers up to clusters containing 45 molecules should be taken into account.

Figures 1 and 2 represent the results of calculations for argon and carbon dioxide; it can be seen that quite large clusters can exist in a gas at low temperatures under high pressures. Large clusters noticeably affect the transport properties of gases. This is demonstrated in Fig. 3 showing the results of calculations of the viscosity of these gases based on a formula stemming from the application of the formula from the kinetic theory of multicomponent mixtures to a cluster to a cluster gas [2, 3, 5].

It can be seen from the curves that the temperature dependence of viscosity in the high-temperature range corresponds to a gas, while this dependence in the range of lower temperatures corresponds to the liquid state of aggregation (viscosity increases upon cooling). This effect is also observed in the experiments [4], which demonstrate that the cluster model correctly reflects the behavior of the temperature dependence of the viscosity of gases in the entire temperature range. The results correspond to temperatures above the critical value for each gas; consequently, the substance exists in the gaseous state (there is no visible interface between two phases). This feature indicates that heavy

1063-7842/05/5008-1098$26.00

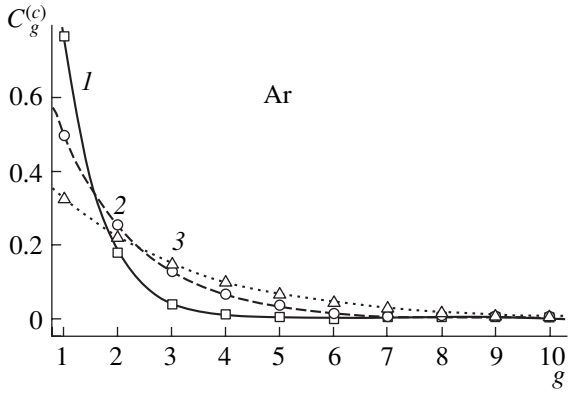


Fig. 1. Concentration $C_g^{(c)}$ of cluster subcomponents as a function of cluster size g (in the number of molecules) for argon at different temperatures T, K: (*1*) 280, (*2*) 210, and (*3*) 190. The curves are calculated by the formula $C_g^{(c)} = C_1^{(c)} \exp[-G(g-1)]$.

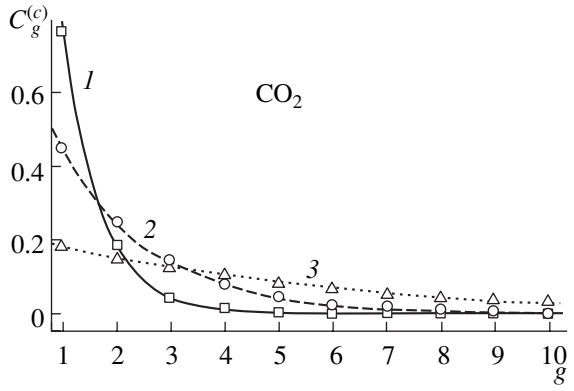


Fig. 2. The same as in Fig. 1 for carbon dioxide at different temperatures T, K: (*1*) 280, (*2*) 330, and (*3*) 310. The curves are calculated by the formula $C_g^{(c)} = C_1^{(c)} \exp[-G(g-1)]$.

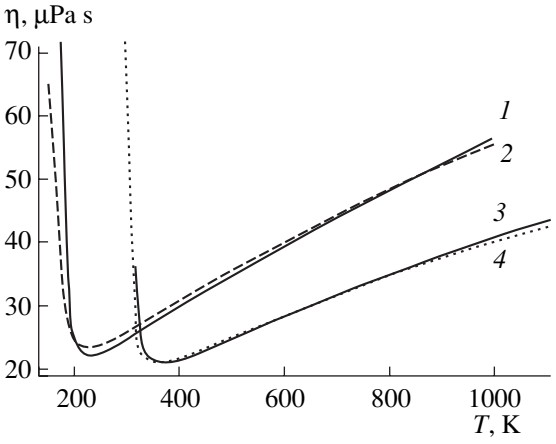


Fig. 3. Temperature dependence of the viscosity coefficient η (μPa s). The solid curves are calculated using the cluster model, and the dashed curves are smoothed experimental data [4]: (*1*, *2*) argon at a pressure of 10 MPa; (*3*, *4*) carbon dioxide under a pressure of 8 MPa.

clusters impose the properties of the second phase on the gas since these clusters are intermediate (mesoscopic) particles and a cluster gas exhibits mesoscopic properties.

Analyzing the contribution from each cluster subcomponent to the internal friction of the gas can reveal the physical mechanism leading to the observed features of viscosity. Figures 4 and 5 show the results of calculations of the partial viscosity coefficient for each cluster subcomponent and partial viscosity.

It can be seen from Fig. 4 that the partial viscosity coefficient increases with the cluster size. This is due to the velocity persistence effect after collisions (heavy particles colliding with light particles preserve the direction of the initial velocity to a larger extent) [6]. It can also be seen that this increase slows down upon an increase in pressure since the fraction of collisions between heavy particles increases in this case. The contribution to the viscosity of the cluster mixture substantially depends on the concentration of the corresponding clusters (Fig. 5). Consequently, at relatively high temperatures, the contribution from large clusters is small due to their small fraction and, hence, their mesoscopic properties do not noticeably affect the entire gas. At low temperatures under a given pressure, mesoscopy does not affect the properties of the entire gas any longer. Viscous properties of the liquid state appear partly due to enhancement of the momentum transfer mechanism over a distance equal to the effective diameter of collisions, which is taken into account in the formula for viscosity [3] in the *n*th type correction.

Mesoscopy of cluster gases is also manifested in other properties of high-density gases. For example, the concentration dependence of the mutual diffusion coefficient in dense gases corresponds to the dependence observed in liquids and solids [7]. Calculations based on the cluster model correspond to the experimentally observed passage of the ratio of tracing diffusion coefficients through unity [8], which is due to mesoscopy of a cluster gas.

Another example of mesoscopy is illustrated in Fig. 6, according to which the thermodiffusion baric effect in gases changes sign upon an increase in pressure: an increase in pressure is observed not in the heated region as in rarefied gases [9, 10], but in the cold region as in liquids [11]. These data were obtained by calculating the difference in pressure (baric effect) existing in a closed device consisting of two vessels connected by a capillary tube or via a porous tablet; the gas in the vessels is kept at different temperatures [9, 10]. The kinetic theory of this effect is constructed on the basis of the assumption that the particle flow observed in the general case must contain two components, reversible and irreversible. In a closed device under steady-state conditions, the observed particle flux must be zero, which is ensured by the flow caused by the baric effect associated with the irreversible part of the flow. As applied to the cluster model, this physi-

cal mechanism is preserved, but clusters of various sizes also participate in the process in addition to molecules. The variability of the cluster composition and, hence, the number of structural elements complicate the process, but the main pattern remains unchanged. As applied to clusters containing g molecules in a gas consisting of identical molecules, the expression derived for the particle flow in the kinetic theory of inhomogeneous multicomponent mixtures [2, 12] has the form

$$\begin{aligned}\mathbf{\Gamma}_g &= -n_g D_g \nabla \ln C_g^{(c)} - n_g D_g \nabla \ln p \\ &\quad + n_g D_g^T \nabla \ln T + n_g \mathbf{W},\end{aligned} \tag{3}$$

where $\mathbf{\Gamma}_g$ is the flux density for g-dimensional clusters, D_g and D_g^T are the true diffusion coefficient and the thermodiffusion coefficient for g-dimensional clusters, p is the pressure, $\mathbf{W}$ is the velocity of the flow, and T is the temperature.

In experiments on baric effects, pressure diffusion can be disregarded; in this case, expression (3) for the molecular flux gives

$$g\mathbf{\Gamma}_g = g n_g D_g^T \left(1 - \frac{D_g}{D_g^T}\frac{T n^{(c)}}{n_g}\frac{\partial C_g^{(c)}}{\partial T}\right)\nabla \ln T + g n_g \mathbf{W}. \tag{4}$$

The expression for the pressure effect is usually derived by using the Poiseuille formula for velocity W averaged over the cross section of the tube, which is written precisely in terms of pressure difference [9, 10]. This approach leads to the formula for the thermodiffusion pressure effect,

$$\Delta p = \frac{8\langle\eta\rangle}{r_c^2 T}\sum_{g=1}^{r} d C_g^{(n)} D_g^T \left(1 - \frac{D_g T}{D_g^T}\left(\frac{\partial \ln C_g^{(c)}}{\partial T}\right)\right), \tag{5}$$

where $\langle\eta\rangle$ is the average viscosity coefficient and r_c is the radius of the capillary or the effective pore radius of the porous tablet.

In deriving formula (5), it is convenient to use the concentration $C_g^{(n)}$ introduced in terms of the total number density $n^{(n)}$ of molecules [3],

$$C_g^{(n)} = \frac{n_g}{n^{(n)}} = \frac{n_g}{\sum_{g=1}^{r} g n_g} \tag{6}$$

and which can be expressed in terms of $C_g^{(c)}$ by the following relation [3]:

$$C_g^{(n)} = C_g^{(c)} \frac{1}{\sum_{g=1}^{r} g C_g^{(c)}}. \tag{7}$$

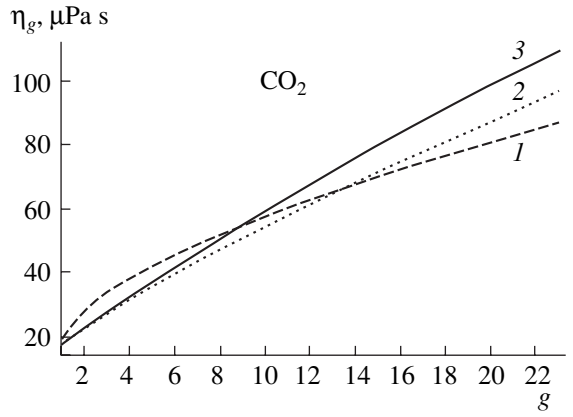


Fig. 4. Partial viscosity coefficient for a cluster component under a pressure of 8 MPa as a function of cluster size in carbon dioxide at different temperatures T, K: (*1*) 310, (*2*) 330, and (*3*) 350.

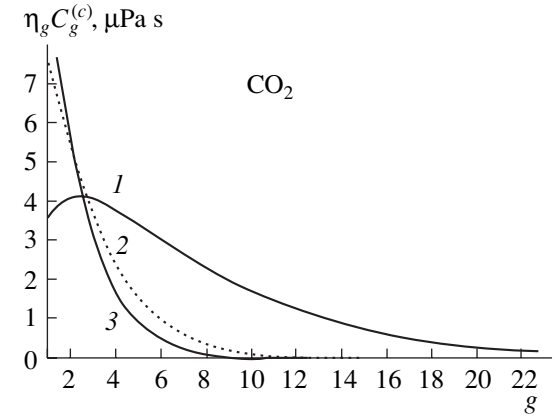


Fig. 5. Partial viscosity for a cluster component as a function of cluster size in carbon dioxide at different temperatures. Notation is the same as in Fig. 4.

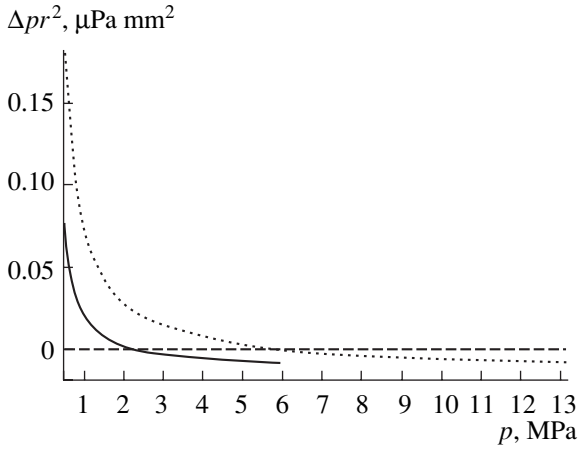


Fig. 6. Thermodiffusion baric effect for a unit (in mm) radius of the capillary tube and a unit temperature difference for an average temperature T = 300 K. The solid and dashed curves correspond to CO_2 and Ar, respectively.

For low pressures and high temperatures, when the effect of clusters can be ignored, formula (5) is transformed to the well-known formula for a rarefied gas, which leads to the results matching the experimental data [9, 10]. The results of calculations for argon and carbon dioxide shown in Fig. 6 indicate that the baric effect differs from that for a rarefied gas upon an increase in pressure (the sign of the effect is reversed). This means that an elevated pressure in a cluster mixture is observed in the cold region as in liquids [11]. This is a manifestation of mesoscopy of the cluster gas.

Thus, the experiment and theory show that the properties of gases at low temperatures and high pressures correspond to the properties of liquids; in the cluster model, this is attributed to the effect of large clusters and is a manifestation of cluster gas mesoscopy. Such an intermediate region corresponds to a homogeneous system, and the observed transition can be rightfully referred to as a mesoscopic phase transition.

REFERENCES

1. S. Ya. Khmel' and R. G. Sharafutdinov, Zh. Tekh. Fiz. **67** (3), 43 (1997) [Tech. Phys. **42**, 291 (1997)].
2. L. I. Kurlapov, *Kinetic Theory of Irreversible Processes in Gases* (Kazakhsk. Nats. Univ. im. Al'Farabi, Almaty, 2000) [in Russian].
3. L. I. Kurlapov, Zh. Tekh. Fiz. **73** (2), 51 (2003) [Tech. Phys. **48**, 185 (2003)].
4. N. B. Vargaftik, *Tables of the Thermophysical Properties of Liquids and Gases* (Nauka, Moscow, 1972; Halsted Press, New York, 1975).
5. L. I. Kurlapov, Zh. Tekh. Fiz. **48**, 1302 (1978) [Sov. Phys. Tech. Phys. **23**, 731 (1978)].
6. J. H. Jeans, *The Dynamical Theory of Gases*, 4th ed. (Dover, New York, 1925).
7. R. Reid, J. Praustniz, and T. Sherwood, *The Properties of Gases and Liquids*, 3rd ed. (McGraw-Hill, New York, 1977; Khimiya, Leningrad, 1982).
8. L. I. Kurlapov, Inzh.-Fiz. Zh. **54**, 438 (1988).
9. N. D. Kosov, A. F. Bogatyrev, and L. I. Kurlapov, in *Heat and Mass Transfer* (Nauka i Tekhnika, Minsk, 1968), Vol. 7, pp. 507–517 [in Russian].
10. N. D. Kosov, A. F. Bogatyrev, and L. I. Kurlapov, Zh. Tekh. Fiz. **39**, 1119 (1969) [Sov. Phys. Tech. Phys. **14**, 841 (1969)].
11. B. Deryagin and G. Sidorenko, Dokl. Akad. Nauk SSSR **32**, 622 (1941).
12. L. I. Kurlapov, Zh. Tekh. Fiz. **56**, 386 (1986) [Sov. Phys. Tech. Phys. **31**, 233 (1986)].

Translated by N. Wadhwa

Technical Physics, Vol. 50, No. 8, 2005, pp. 1102–1104. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 140–142.

SHORT COMMUNICATIONS

Formation of Higher Manganese Silicide Films on Silicon

T. S. Kamilov*, D. K. Kabilov*, I. S. Samiev*, Kh. Kh. Khusnutdinova*, R. A. Muminov, V. V. Klechkovskaya*****

** Tashkent State Aviation Institute, Tashkent, 700047 Uzbekistan*

e-mail: tdai_kts@mail.ru

*** Physicotechnical Institute, Physics–Sun Research and Production Corporation, Academy of Sciences of Uzbekistan, Tashkent, Uzbekistan*

**** Shubnikov Institute of Crystallography, Russian Academy of Sciences, Leninskiĭ pr. 59, Moscow, 117333 Russia*

Received January 18, 2005

Abstract—The formation of higher manganese silicide (HMS) films on silicon and the properties of the silicide–silicon interface are studied. Morphology analysis of the surface and thin transition layer at the HMS–Si interface suggests that the growth of HMS films by the method of Mn reactive diffusion follows the vapor–liquid–crystal mechanism. © *2005 Pleiades Publishing, Inc.*

INTRODUCTION

Comprehensive study of HMS films is of great importance for thermoelectric and optoelectronic applications [1, 2], since these films are characterized by a high thermal stability in the temperature range 300–1000 K [3]. It is worth noting that the physical properties of transition metal silicide films may differ substantially from those of the bulk silicides because of the difference in their structure. When designing film thermoelements and thermoelectric microbatteries based on manganese silicide films, one should take into account a number of issues, such as the direction of silicide growth, the silicide–silicon interface condition, and the influence of the silicon substrate on the electrophysical parameters of the films.

Investigation into HMS film formation and the properties of the HMS–silicon interface is a challenging problem, which this paper is devoted to.

EXPERIMENTAL RESULTS

In [4], we showed that growth of HMS films on a silicon substrate is accompanied by the formation of a 5- to 7-μm-thick transition layer at the HMS–Si interface. It was conjectured that this layer is amorphous or polycrystalline. However, in that work, manganese silicide films were grown on mechanically polished silicon substrates, which are known to have an 8- to 12-μm-thick damaged layer on their surface; therefore, the transition layer at the interface was most likely to consist mainly of the remaining (lower) part of the damaged silicon layer. The lattice mismatch between HMS and Si is ≈1.8% (a_{Si} = 0.543 nm, a_M = 0.551 nm) [2]. According to the published data [5], when the silicide–silicon lattice mismatch is no more than 3%, the growth of silicide films on a silicon substrate may be accompanied by the formation of a thin intermediate layer, so that transition metal silicide films will grow epitaxially.

In this work, silicon substrates were prepared with the technique detailed in [6]. This technique provides a surface finish satisfying the stringent requirements of planar technology (class ∇14), where the thickness of the damaged layer must not exceed 50 nm [7]. Electron microscopy studies showed that the transition layer at the HMS–Si interface is 0.1–0.3 μm thick. Optical and electron micrographs demonstrate a sharp transition from the Si substrate to the silicide film (Figs. 1a and 1b).

The surface morphology of HMS films grown on (111)Si and (100)Si substrates is shown in Fig. 2. It is seen that the habit and arrangement of HMS crystallites depend on the substrate orientation. While in the HMS-on-(111)Si films the crystallites form a texture, i.e., are misoriented in the substrate plane but cooriented in the perpendicular direction, the HMS-on-(100)Si films grow nearly epitaxially.

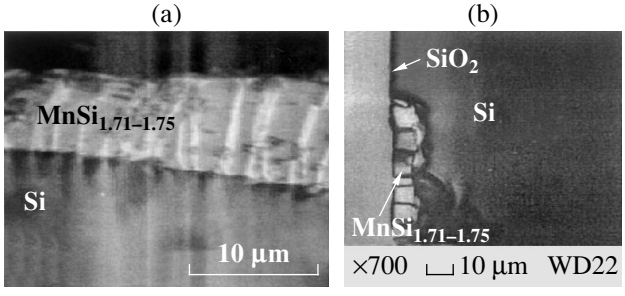


Fig. 1. (a) Cross-sectional image of the HMS films that is obtained under an optical microscope and (b) electron micrograph taken from the cross section of the HMS films grown on silicon in the presence of a SiO_2 mask.

1063-7842/05/5008-1102$26.00 © 2005 Pleiades Publishing, Inc.

DISCUSSION

In accordance with [8], if elemental crystalline substance M is in contact with a vapor phase containing atoms of element X that can combine with M to form solid compound M_aX_b, crystal M will be covered by a film of this compound. The reaction proceeds as a result of a change in the free energy of the system when the compound forms,

$$aM + bX = M_aX_b + \Delta G.$$

A film of the resulting compound will be continuous (free of voids and cracks) if the molecular volume of compound M_aX_b is larger than, or equal to, that of element M. Once a continuous fine-grained film of the compound has formed, the reaction is controlled by the diffusion of the reagents through the crystal lattice of compound M_aX_b and along grain boundaries.

During reactive diffusion, new phases arise when atoms of the matrix and diffusate have a limited mutual solubility or can form compounds [9, 10]. As is known [9], the solubility of silicon in metals is low: in particular, 14% in manganese. The solubility of all metals in silicon is extremely low, because the energy of vacancy formation in Si is high.

In general, films on a substrate form in two stages [10]: nucleation and growth of nuclei. In turn, nucleation involves the following processes: (i) mass transfer of a initial phase (e.g., by means of diffusion), i.e., delivery of the material to the substrate surface and growing layer; (ii) adsorption of initial phase atoms on the surface, which results in formation of critical nuclei; and (iii) surface diffusion (migration) of atoms adsorbed, as a result of which particles are delivered to critical nuclei, making them crystallization centers.

In crystallization from the vapor (gas) phase, structural ordering in the films depends largely on the adatom mobility on the substrate: a high mobility of adatoms favors growth of oriented nuclei [10]. Films grown may be single-crystalline epitaxial or polycrystalline textured depending on the orienting influence of the substrate. In the solid-phase reaction mentioned above, silicide films grow via Mn–Si interdiffusion [11].

However, in our case, manganese diffuses from the gas phase and solid manganese layers on the surface into which silicon atoms could penetrate are absent. Therefore, the growth of a manganese silicide film is basically controlled by the inward diffusion of Mn atoms.

Actually, once early layers of a solid solution or compound have formed, the outward diffusion of substrate atoms through the growing film is observed. Depending on the atomic structure of the film grown by reactive diffusion, the diffusion rate of substrate atoms may be both higher and lower than that of the diffusate. The ratio between these rates specifies the direction of subsequent growth of the film and the direction of the diffusate concentration gradient in it. Calculations of the coefficients of diffusion of manganese into silicon and silicon into manganese silicide at high temperatures showed that the former exceeds the latter by almost six orders of magnitude ($D_{Mn} = 2 \times 10^{-7}$ cm^2/s, $D_{Si} = 2.2 \times 10^{-13}$ cm^2/s) [10].

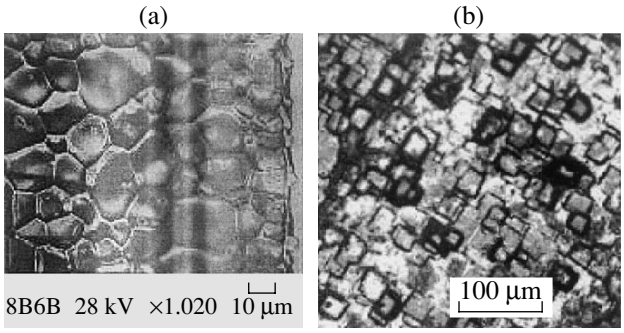


Fig. 2. Surface morphology of the HMS films grown on (a) (111)Si and (b) (100)Si.

Thus, it seems that, in our case, a thin low-concentration surface layer of the Mn–Si solid solution first forms during Mn diffusion from the gas phase into the single-crystalline silicon and then compound $MnSi_{2-x}$ nucleates as a result of reactive diffusion. Here, surface diffusion plays an important role. The silicon content in this layer varies, increasing with distance from the substrate surface. Therefore, the rate of outward diffusion of the silicon through $MnSi_{2-x}$ is much lower than the rate of manganese diffusion and the $MnSi_{2-x}$ film thickens at the $MnSi_{2-x}$–Si interface rather than at the Mn–$MnSi_{2-x}$ one. It is seen from Fig. 1b that the silicide film grows inward to the silicon substrate. Thus, during the formation of the silicide, the HMS–Si interface shifts inward to the silicon; i.e., the manganese silicide film grows from the surface of the silicon substrate into its volume.

It is known [10] that films grow by the vapor–liquid–crystal mechanism at relatively high temperatures such that $T > 2/3T_{mel}$ (T_{mel} is the melting point of the condensing metal). At the initial stage of condensation, the phase that is in equilibrium with vapor is liquid: its islands coalesce and solidify, producing a polycrystalline film [10]. Film growth and, accordingly, melting of small nuclei of manganese silicide take place at a temperature of 1000–1040°C, which is below the melting point of bulk HMS (T = 1145°C [11]) but approaches it as the nucleus size increases. The liberation of the heat of condensation raises the process temperature and, thus, favors the vapor–liquid–crystal mechanism.

The data for the morphology of the surface and thin transition layer at the HMS–Si interface are further evidence that the growth of manganese silicide films follows this mechanism.

CONCLUSIONS

It is found that a thin (0.1–0.3 μm) transition layer exists at the HMS–Si interface. The cross-sectional structure of the manganese silicide films grown on the silicon substrate shows that they grow from the surface of the substrate into its volume. The surface morphology of the HMS-on-(111)Si and HMS-on-(100)Si films indicates that the orientation of grains in the HMS film depends on the substrate orientation, the HMS-on-(100)Si films growing nearly epitaxially. The morphology of the surface of the film and thin transition layer at the HMS–Si interface suggests that the growth of HMS films via reactive diffusion follows the vapor–liquid–crystal mechanism.

ACKNOWLEDGMENTS

The work was supported by the State Committee for Science and Technology of Uzbekistan, grant no. GNTP-10.20.

REFERENCES

1. Q. R. Hou, Z. M. Wang, and Y. J. He, Appl. Phys. A: Mater. Sci. Process. **79**, 1 (2004).
2. J. E. Mahan, Thin Solid Films **461**, 152 (2004).
3. V. K. Zaitsev, in *CRC Handbook of Thermoelectrics*, Ed. by D. M. Rome (CRC, New York, 1995), pp. 299–309.
4. T. S. Kamilov, A. Zh. Khusanov, M. K. Bakhadyrkhanov, *et al.*, Poverkhnost: Rentgenovskie, Sinkhrotronnye, Neĭtronnye Issled., No. 6, 100 (2002).
5. S. P. Murarka, *Silicides for VLSI Applications* (Academic, New York, 1983; Mir, Moscow, 1986).
6. V. V. Klechkovskaya, T. S. Kamilov, S. I. Adasheva, *et al.*, Kristallografiya **39**, 894 (1994) [Crystallogr. Rep. **39**, 815 (1994)].
7. A. A. Chernyshev, *Reliability of Semiconductor Devices and Integrated Circuits* (Radio i Svyaz', Moscow, 1988) [in Russian].
8. S. A. Medvedev, *Introduction to the Semiconductor Materials Technology* (Vysshaya Shkola, Moscow, 1970).
9. *Thin Films: Interdiffusion and Reactions*, Ed. by J. M. Poate, K. Tu, and J. Meier (Wiley, New York, 1978; Mir, Moscow, 1982).
10. A. A. Barybin and V. G. Sidorov, *Physicotechnical Basics of Electronics* (Lan', St. Petersburg, 2001) [in Russian].
11. G. V. Samsonov, L. A. Dvorina, and B. M. Rud', *Silicides* (Metallurgiya, Moscow, 1979) [in Russian].

Translated by M. Astrov

Technical Physics, Vol. 50, No. 8, 2005, pp. 961–966. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 1–6.

THEORETICAL AND MATHEMATICAL PHYSICS

On the Simplest Equivalent Currents of an Ellipsoidal Body

R. Z. Muratov and V. L. Shkuratnik

Moscow State Mining University, Moscow, 117935 Russia

e-mail: rodeszmu@mtu-net.ru

Received March 9, 2004

Abstract—The solution is obtained to the Frenkel problem of determining the surface current on the boundary of an ellipsoidal region of space, equivalent (i.e., producing the same external magnetic field) to the simplest steady-state current specified in the bulk of this region. Simultaneously, multipole representations of pseudoscalar magnetic potentials of the volume and surface currents are determined in the framework of this particular problem. It is shown that these representations are completely analogous to the corresponding multipole representations of the scalar potentials of the ellipsoid associated with the volume or surface distributions of scalar sources (charges or masses). © *2005 Pleiades Publishing, Inc.*

INTRODUCTION

This study is devoted to the problem of the magnetic field of an ellipsoid, which may have an exact analytic solution.

Historically, the first problem that was solved in macroscopic magnetostatics was the boundary-value problem of a homogeneous ellipsoid (with $\mu \neq 1$) in a uniform magnetic field. This solution, which was obtained by Poisson, was included by Maxwell in his famous *Treatise* [1] and referred to as "rather ingenious." In Poisson's work, the ellipsoidal body was an insulating or conducting (in the absence of conduction currents) nonferromagnetic magnet. In this case (see, for example, [2]) the system of magnetostatics equations and the boundary conditions to these equations formally coincide with the system of equations defining the electrostatic field in insulators (in the absence of free charges) and differs from these equations only in the substitution

$$\mathbf{E} \Rightarrow \mathbf{H}, \quad \mathbf{D} \Rightarrow \mathbf{B}. \tag{1}$$

Among other things, this means that the pseudoscalar magnetic potential and the interpretation based on the concept of bound "magnetic charges" [3] can be rightfully used in the entire space in analysis of such magnetostatic problems. It should be noted, bearing in mind the substitution (1), that generalization of Poisson's results can be found in the works [4, 5] devoted to a layered inhomogeneous dielectric ellipsoid in a nonuniform external electrostatic field.

Here, we consider the following problem. In the bulk of an ellipsoid bounded by the surface

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, \tag{2}$$

the simplest current distribution characterized by the density

$$j_x = 0, \quad j_y = B_{001}\frac{z}{c}, \quad j_z = C_{010}\frac{y}{b}, \tag{3}$$

which is a linear function of coordinates, is specified. We must find on surface (2) the density $\mathbf{i}(\mathbf{r})$ of a current, which is equivalent to the volume current (3).[1]

Since the magnetic field sources (currents) in the given problem are distributed directly in the bulk of the ellipsoid, the introduction of the pseudoscalar potential in the inner region is ruled out (in contrast to the Poisson-type problems). The method of analytic solution which is effective for a sphere [8] and based on the multipole expansion of potential is also inapplicable since the number of nonzero magnetic multipole moments is always infinitely large in the case of an ellipsoid. This necessitates that vector potentials and magnetic fields be calculated preliminarily and the pseudoscalar potentials produced by the preset and sought currents be determined. It will be proved, however, that the results of this rather cumbersome solution indicate the existence of the same fundamental relation between magnetostatic problems of an ellipsoid (in the presence and absence of currents in it), which was established for it in analogous electrostatic problems (in the presence and absence of charges).

1. POTENTIALS AND MAGNETIC MULTIPOLES OF BULK CURRENT

Any constant bulk current must satisfy the steady-state condition

$$\mathrm{div}\mathbf{j} = 0 \tag{4}$$

[1] A class of problems of equivalent (i.e., producing the same external fields) sources of the scalar (charges) or vector (currents) origin was formulated by Frenkel [6, p. 103; 7, p. 524]. The solution to the problem of equivalent currents for concentric spherical spatial regions was given in [8].

1063-7842/05/5008-0961$26.00 © 2005 Pleiades Publishing, Inc.

and the boundary condition

$$j_n|_S = 0 \tag{5}$$

at the boundary S of the ellipsoid. In the case of the ellipsoidal boundary (1) under consideration, we can write the latter condition in the form

$$\frac{1}{p}\langle n_x j_x\rangle\Big|_S = \left\langle \frac{x}{a^2} j_x\right\rangle\Big|_S = 0. \tag{6}$$

Here, $p = \langle x^2/a^4\rangle^{-1/2}$ is the length of the perpendicular dropped from the center of the ellipsoid on the plane tangential to its surface at the point characterized by the unit normal vector **n**; angle brackets $\langle\ldots\rangle$ denote everywhere the sum of three terms of a cyclic permutation. Current (3) satisfies condition (4) automatically, while condition (6) implies that

$$cB_{001} = -bC_{010}. \tag{7}$$

Current (3) corresponds to the vector potential[2]

$$\mathbf{A} = \frac{1}{c}\int \frac{\mathbf{j}}{R} dV \tag{8}$$

of the magnetic field. Volume integrals contained in expression (8) for the ellipsoid in question are considered in detail in [9]. In particular, for points of observation outside the ellipsoid, we have

$$\int x\frac{dV}{R} = 2\pi a^2 x\left(\mathcal{M}_{100} - \frac{1}{3}\mathcal{M}_{200}x^2 - \mathcal{M}_{110}y^2 - \mathcal{M}_{101}z^2\right); \tag{9}$$

the external potential factors $\mathcal{M}_{lmn}(\xi)$ appearing in this relation are defined as

$$\mathcal{M}_{lmn}(\xi) = \Pi_{lmn}\frac{abc}{2}\int_\xi^\infty \frac{du}{(a^2+u)^{l+\frac{1}{2}}(b^2+u)^{m+\frac{1}{2}}(c^2+u)^{n+\frac{1}{2}}}. \tag{10}$$

Here, $\Pi_{lmn} = (2l-1)!!(2m-1)!!(2n-1)!!$ and ξ is the ellipsoidal coordinate of the point of observation (x, y, z), which is a nonnegative root of the equation $x^2/(a^2+\xi) + y^2/(b^2+\xi) + z^2/(c^2+\xi) = 1$. On the surface of ellipsoid (2), the ellipsoidal coordinate ξ is equal to zero. The expressions for the integrals differing from expression (9) in that x in the integrand is replaced by y or z can be obtained from expression (9) with the help of cyclic permutation.

[2] Semiaxis c of the ellipsoid should not be confused with the velocity of light c.

Taking into account relation (7), we obtain the vector potential components outside the ellipsoid:

$$A_x^{(e)} = 0, \quad A_y^{(e)} = -\frac{V}{5c}bC_{010}\varphi_{001}, \quad A_z^{(e)} = \frac{V}{5c}bC_{010}\varphi_{010}. \tag{11}$$

Here, V is the volume of the ellipsoid,

$$\varphi_{100} = \frac{15}{2abc}\left(\mathcal{M}_{100} - \frac{1}{3}\mathcal{M}_{200}x^2 - \mathcal{M}_{110}y^2 - \mathcal{M}_{101}z^2\right)x, \tag{12}$$

and expressions for φ_{010} and φ_{001} can be obtained from relation (12) by cyclic permutation. Quantities $\varphi_{\alpha\beta\gamma}$ introduced in [10] form the so-called tensor potential of rank $\nu = \alpha + \beta + \gamma$ of the ellipsoid. The transition from the three-index form of tensor $\varphi_{\alpha\beta\gamma}$ which is used here and is valid only for symmetric tensors, as well as tensor $\psi_{\alpha\beta\gamma}$ appearing in the next section, to the conventional tensor notation is explained by the following example:

$$\varphi_{\alpha\beta\gamma} \equiv \varphi_{\underbrace{x\ldots x}_{\alpha\text{ times}}\,\underbrace{y\ldots y}_{\beta\text{ times}}\,\underbrace{z\ldots z}_{\gamma\text{ times}}}.$$

In particular, components $\varphi_{100} \equiv \varphi_x$, $\varphi_{010} \equiv \varphi_y$, and $\varphi_{001} \equiv \varphi_z$ are a vector representative of this family of tensors.

Vector potential (11) corresponds to the components of magnetic induction $\mathbf{B} = \mathrm{curl}\,\mathbf{A}$, which are described by the formulas

$$B_x^{(e)} = \frac{V}{5c}bC_{010}\left(\frac{\partial\varphi_{010}}{\partial y} + \frac{\partial\varphi_{001}}{\partial z}\right),$$

$$B_y^{(e)} = -\frac{V}{5c}bC_{010}\frac{\partial\varphi_{010}}{\partial x}, \quad B_z^{(e)} = -\frac{V}{5c}bC_{010}\frac{\partial\varphi_{001}}{\partial x}.$$

Using relations

$$\frac{\partial\varphi_{100}}{\partial x} + \frac{\partial\varphi_{010}}{\partial y} + \frac{\partial\varphi_{001}}{\partial z} = 0, \tag{13}$$

$$\frac{\partial\varphi_{010}}{\partial x} = \frac{\partial\varphi_{100}}{\partial y}, \quad \frac{\partial\varphi_{001}}{\partial y} = \frac{\partial\varphi_{010}}{\partial z}, \quad \frac{\partial\varphi_{100}}{\partial z} = \frac{\partial\varphi_{001}}{\partial x}, \tag{14}$$

which are proved in [11], we arrive at the final expression,

$$\mathbf{B}^{(e)} = -\mathrm{grad}\Phi^{(e)}, \tag{15}$$

where

$$\Phi^{(e)} = \frac{V}{5c}bC_{010}\varphi_{100}. \tag{16}$$

Thus, we have derived the expression for the pseudoscalar magnetic field potential $\Phi^{(e)}$ produced by bulk current (3).

Using the Lagrange formula

$$\int\left(\frac{x}{a}\right)^{2l}\left(\frac{y}{b}\right)^{2m}\left(\frac{z}{c}\right)^{2n}dV = 3\frac{(2l-1)!!(2m-1)!!(2n-1)!!}{(2l+2m+2n+3)!!}V, \tag{17}$$

we can now evaluate the magnetic dipole $\mathbf{m}^{(\mathbf{j})}$ and octupole $m^{(\mathbf{j})}_{klm}$ moments of the ellipsoid[3] produced by current (3).

This gives

$$m^{(\mathbf{j})}_x = \frac{1}{2c}\int[\mathbf{rj}]_x dV = \frac{V}{5c}bC_{010}, \qquad m^{(\mathbf{j})}_y = m^{(\mathbf{j})}_z = 0, \tag{18}$$

$$m^{(\mathbf{j})}_{xyy} = \frac{3}{4c}\int\{10xyzj_x + (x^2+z^2-4y^2)zj_y + (4y^2-11x^2-z^2)yj_z\}dV = -\frac{9V}{7!!c}bC_{010}(3a^2-4b^2+c^2), \tag{19}$$

$$m^{(\mathbf{j})}_{xzz} = -\frac{9V}{7!!c}bC_{010}(3a^2+b^2-4c^2), \tag{20}$$

$$m^{(\mathbf{j})}_{xxx} = -m^{(\mathbf{j})}_{xyy} - m^{(\mathbf{j})}_{xzz}. \tag{21}$$

All the remaining components of the octupole magnetic moment of current (3) are equal to zero.

The octupole moment of bulk current will be required at a later stage. Formulas (18) for the dipole moment make it possible to write relation (16) in the form of the scalar product,

$$\Phi^{(e)} = m^{(\mathbf{j})}_x\varphi_x. \tag{22}$$

2. MAGNETIC MULTIPOLES AND POTENTIALS OF SURFACE CURRENT

It is well known (see, for example, [12]) that the volume charge distributed in an ellipsoid with a density which is a polynomial (of degree L) function of coordinates is equivalent to a charge of density σ on the surface of the ellipsoid, where σ/p is a polynomial of degree $L+2$. This implies that the Cartesian components of the surface current density $\mathbf{i}$ at boundary (2), which is equivalent to preset current (3), should be sought in the form of cubic polynomials. In addition, it follows from symmetry considerations that i_x must be an odd function of all Cartesian coordinates, i_y must be an even function of x and y and an odd function of z, and i_z must be an even function of x and z and odd function of y. Consequently, the sought surface current can be written in the form

$$\left.\begin{aligned}\frac{i_x}{p} &= \alpha_{111}\frac{xyz}{abc}\\ \frac{i_y}{p} &= \left(\beta_{201}\frac{x^2}{a^2}+\beta_{021}\frac{y^2}{b^2}+\beta_{003}\frac{z^2}{c^2}\right)\frac{z}{c}\\ \frac{i_z}{p} &= \left(\gamma_{210}\frac{x^2}{a^2}+\gamma_{030}\frac{y^2}{b^2}+\gamma_{012}\frac{z^2}{c^2}\right)\frac{y}{b}.\end{aligned}\right\} \tag{23}$$

For the surface current, the boundary condition (6) assumes the form

$$\left\langle\frac{x}{a^2}i_x\right\rangle\Big|_S = 0. \tag{24}$$

Substituting expressions (23) into this relation and equating the coefficients of like powers of coordinates, we arrive the equalities

$$\frac{\alpha_{111}}{a}+\frac{\beta_{201}}{b}+\frac{\gamma_{210}}{c} = 0, \quad \frac{\beta_{021}}{b} = -\frac{\gamma_{030}}{c}, \quad \frac{\beta_{003}}{b} = -\frac{\gamma_{012}}{c}. \tag{25}$$

If we assume that the entire ellipsoid (2) is split by ellipsoidal surfaces (which are similar to it, arranged similarly, and concentric to it) into infinitely thin ellipsoidal layers,[4] it can easily be proved that equalities (25) for a bulk current, which differ from (23) only in that they do not contain factor p, do not allow such a current to cross the boundaries of the layers. Consequently, the steady-state condition (4) for surface current $\mathbf{i}$ assumes the form

$$\mathrm{div}\left(\frac{\mathbf{i}}{p}\right) = \left\langle\frac{\partial}{\partial x}\frac{i_x}{p}\right\rangle = 0, \tag{26}$$

supplementing relations (25) by a relation that follows from (26):

$$\frac{\alpha_{111}}{a}+2\frac{\beta_{021}}{b}+2\frac{\gamma_{012}}{c} = 0. \tag{27}$$

It should be noted that, on account of four relations (25) and (27), only three of seven sought coefficients appearing in polynomials (23) remain independent. In terms of these three coefficients (we take γ_{030}, γ_{012}, and γ_{210} for this purpose), we can express, in particular, the multipole moments of surface current (23).

It is well known that, in the simply connected region of space external relative to steady-state electric cur-

[3] The quadrupole and higher moments of the same parity are obviously equal to zero.

[4] Such layers are referred to as homeoids.

rents, the multipole expansion of pseudoscalar potential of their magnetic field has the form [13]

$$\Phi(\mathbf{r}) = \sum_{l=1}^{\infty} \frac{(-2)^l}{(2l)!} m_{i_1 \ldots i_l} \nabla_{i_1} \ldots \nabla_{i_l} \frac{1}{r}, \tag{28}$$

where ∇ is the Hamilton operator.

It follows from the universality of the multipole expansion that all multipole moments corresponding to equivalent current systems coincide. In particular, the dipole and octupole magnetic moments of bulk (3) and surface (23) currents in the problem considered here must coincide.

Using the Lagrange formula

$$\oint \left(\frac{x}{a}\right)^{2l} \left(\frac{y}{b}\right)^{2m} \left(\frac{z}{c}\right)^{2n} p\,dS = 3\frac{(2l-1)!!(2m-1)!!(2n-1)!!}{(2l+2m+2n+1)!!} V, \tag{29}$$

we can calculate the magnetic dipole $\mathbf{m}^{(\mathbf{i})}$ and octupole $m^{(\mathbf{i})}_{klm}$ moments of the ellipsoid caused by current (23). Taking into account relations (25), we obtain

$$m_x^{(\mathbf{i})} = \frac{1}{2c}\oint [\mathbf{ri}]_x dS = \frac{Vb}{5c}(\gamma_{012} + \gamma_{210} + 3\gamma_{030}), \quad m_y^{(\mathbf{i})} = m_z^{(\mathbf{i})} = 0; \tag{30}$$

$$m_{xyy}^{(\mathbf{i})} = \frac{3}{4c}\oint \{10xyzi_x + (x^2 + z^2 - 4y^2) z i_y + (4y^2 - 11x^2 - z^2) y i_z\} dS = -\frac{3Vb}{35c}\{(5a^2 - 20b^2 + 3c^2)\gamma_{030} + (7a^2 - 4b^2 + 3c^2)\gamma_{012} + (9a^2 - 4b^2 + c^2)\gamma_{210}\}, \tag{31}$$

$$m_{xzz}^{(\mathbf{i})} = -\frac{3Vb}{35c}\{(25a^2 + 5b^2 - 12c^2)\gamma_{030} + (13a^2 + b^2 + 12c^2)\gamma_{012} + (9a^2 + b^2 - 4c^2)\gamma_{210}\}, \tag{32}$$

$$m_{xxx}^{(\mathbf{i})} = -m_{xyy}^{(\mathbf{i})} - m_{xzz}^{(\mathbf{i})}. \tag{33}$$

All remaining components of the octupole moment of the surface current (23) are equal to zero.

For subsequent analysis, we also need inverse formulas expressing coefficients α, β, and γ in terms of magnetic multipoles,

$$-c\beta_{021} = b\gamma_{030} = \frac{c}{6V\delta_a}\left\{15(14a^2 + 7a^2c^2 - \delta_a)m_x^{(\mathbf{i})} + \frac{7}{3}(11a^2 + 4c^2)m_{xyy}^{(\mathbf{i})} - \frac{7}{3}(a^2 - c^2)m_{xzz}^{(\mathbf{i})}\right\}, \tag{34}$$

$$-b\gamma_{210} = \frac{c}{6V\delta_a}\left\{15(\delta_a - 14a^2b^2 - 7b^2c^2)m_x^{(\mathbf{i})} + \frac{7}{3}(32a^2 + b^2 + 12c^2)m_{xyy}^{(\mathbf{i})} + \frac{7}{3}(8a^2 + 4b^2 + 3c^2)m_{xzz}^{(\mathbf{i})}\right\}, \tag{35}$$

$$-\frac{bc}{a}\alpha_{111} = \frac{7c}{9V\delta_a}\{45a^2(b^2 - c^2)m_x^{(\mathbf{i})} - (12a^2 - b^2 + 4c^2)m_{xyy}^{(\mathbf{i})} + (12a^2 + 4b^2 - c^2)m_{xzz}^{(\mathbf{i})}\}, \tag{36}$$

where

$$\delta_a \equiv 8a^4 + 3a^2b^2 + 3a^2c^2 + b^2c^2. \tag{37}$$

As regards the expressions for $b\gamma_{012} = -c\beta_{003}$ and $c\beta_{201}$, they can be obtained from formulas (34) and (35), respectively, with the help of mutual substitution $y \longleftrightarrow z$. For all pseudoquantities (including multipole magnetic moments), we must reverse their signs.

From the nonzero components of the dipole and octupole moments presented in relations (30)–(33), three components are independent (by virtue of relation (33)). This coincides with the number of independent coefficients of polynomials (23). Consequently, equating $m_x^{(\mathbf{i})} = m_x^{(\mathbf{j})}$, $m_{xyy}^{(\mathbf{i})} = m_{xyy}^{(\mathbf{j})}$, and $m_{xzz}^{(\mathbf{i})} = m_{xzz}^{(\mathbf{j})}$, we obtain a system of algebraic equations leading to the required coefficients. The solution to this system has the form

$$\left.\begin{aligned} \frac{\gamma_{030}}{c} &= -\frac{\beta_{021}}{b} = \frac{a^2(2a^2 + c^2)}{\delta_a}\frac{C_{010}}{c} \\ \frac{\gamma_{012}}{c} &= -\frac{\beta_{003}}{b} = \frac{a^2(2a^2 + b^2)}{\delta_a}\frac{C_{010}}{c} \\ \gamma_{210} &= \frac{b^2(2a^2 + c^2)}{\delta_a} C_{010} \\ \frac{\beta_{201}}{b} &= -\frac{c^2(2a^2 + b^2)}{\delta_a}\frac{C_{010}}{c} \\ \frac{\alpha_{111}}{a} &= -\frac{2a^2(b^2 - c^2)}{\delta_a}\frac{C_{010}}{c}. \end{aligned}\right\} \tag{38}$$

It should be emphasized that, having ensured the coincidence of the dipole and octupole moments of the bulk (3) and surface (23) currents, we have not proved the coincidence of all higher nonzero magnetic multipoles of these currents. Consequently, we must verify that the magnetic field produced by current (23) with coefficient (38) outside the ellipsoid coincides with the field of current (3).

The vector potential

$$\mathbf{A} = \frac{1}{c}\oint \frac{\mathbf{i}}{R} dS \tag{39}$$

of currents (23) contains surface integrals considered in detail in [9]. In particular, for points of observation outside the ellipsoid, the formulas

$$\oint\left(\frac{x}{a}\right)^2\frac{y}{b}\frac{pdS}{R} = Vb\left(\frac{1}{5}\varphi_{010}+\frac{18a^2}{7}\psi_{210}\right),$$

$$\oint\left(\frac{z}{c}\right)^3\frac{pdS}{R} = Vc\left(\frac{1}{15}\varphi_{001}+\frac{18c^2}{7}\psi_{003}\right),$$

$$\oint\frac{xyz}{abc}\frac{pdS}{R} = \frac{18}{7}Vabc\psi_{111}$$

hold, as well as those obtained from the first two formulas by cyclic permutation or mutual substitution of the coordinates. Here,

$$\psi_{210} = -\frac{7}{12abc}\left(\mathcal{M}_{110}-\mathcal{M}_{210}x^2-\frac{1}{3}\mathcal{M}_{120}y^2-\mathcal{M}_{111}z^2\right)y, \quad (40)$$

$$\psi_{003} = -\frac{7}{12abc}\left(\mathcal{M}_{002}-\mathcal{M}_{102}x^2-\mathcal{M}_{012}y^2-\frac{1}{3}\mathcal{M}_{003}z^2\right)z, \quad (41)$$

$$\psi_{111} = \frac{7}{6abc}\mathcal{M}_{111}xyz. \quad (42)$$

Quantities $\psi_{\alpha\beta\gamma}$ introduced in [10] form the so-called tensor potential of a homeoid of rank $\nu = \alpha + \beta + \gamma$. It is shown in [11] that the components of tensor $\psi_{\alpha\beta\gamma}$ satisfy the relations

$$\frac{\partial\psi_{\alpha+1,\beta,\gamma}}{\partial x}+\frac{\partial\psi_{\alpha,\beta+1,\gamma}}{\partial y}+\frac{\partial\psi_{\alpha,\beta,\gamma+1}}{\partial z} = 0, \quad (43)$$

$$\frac{\partial\psi_{\alpha,\beta+1,\gamma}}{\partial x} = \frac{\partial\psi_{\alpha+1,\beta,\gamma}}{\partial y}, \quad \frac{\partial\psi_{\alpha,\beta,\gamma+1}}{\partial y} = \frac{\partial\psi_{\alpha,\beta+1,\gamma}}{\partial z}, \quad \frac{\partial\psi_{\alpha+1,\beta,\gamma}}{\partial z} = \frac{\partial\psi_{\alpha,\beta,\gamma+1}}{\partial x}, \quad (44)$$

which are analogous to relations (13) and (14).

As a result, we obtain the following relations for the vector potential components outside the ellipsoid:

$$A_x^{(e)} = \frac{18Vabc}{7c}\alpha_{111}\psi_{111}, \quad (45)$$

$$A_y^{(e)} = \frac{Vc}{5c}(\beta_{201}+\beta_{021}+3\beta_{003})\varphi_{001} + \frac{18Vc}{7c}(a^2\beta_{201}\psi_{201}+b^2\beta_{021}\psi_{021}+c^2\beta_{003}\psi_{003}), \quad (46)$$

while the expression for $A_z^{(e)}$ can be obtained from Eq. (46) by the substitution $y \longleftrightarrow z$.

Accordingly, the magnetic field of the surface current is defined by the formulas

$$\begin{aligned}B_x^{(e)} = \frac{V}{c}\Bigg\{&\frac{b}{5}(\gamma_{210}+\gamma_{012}+3\gamma_{030})\frac{\partial\varphi_{010}}{\partial y}\\ &-\frac{18b}{7}(b^2\gamma_{030}-a^2\gamma_{210})\frac{\partial\psi_{210}}{\partial y}\\ &-\frac{18b}{7}(b^2\gamma_{030}-c^2\gamma_{012})\frac{\partial\psi_{012}}{\partial y}\\ &-\frac{c}{5}(\beta_{201}+\beta_{021}+3\beta_{003})\frac{\partial\varphi_{001}}{\partial z}\\ &+\frac{18c}{7}(c^2\beta_{003}-a^2\beta_{201})\frac{\partial\psi_{201}}{\partial z}\\ &+\frac{18c}{7}(c^2\beta_{003}-a^2\beta_{021})\frac{\partial\psi_{021}}{\partial z}\Bigg\},\end{aligned} \quad (47)$$

$$\begin{aligned}B_y^{(e)} = &\frac{18Vabc}{7c}\alpha_{111}\frac{\partial\psi_{111}}{\partial z}\\ -\frac{V}{c}\Bigg\{&\frac{b}{5}(\gamma_{210}+\gamma_{012}+3\gamma_{030})\frac{\partial\varphi_{010}}{\partial x}\\ &-\frac{18b}{7}(b^2\gamma_{030}-a^2\gamma_{210})\frac{\partial\psi_{210}}{\partial x}\\ &-\frac{18b}{7}(b^2\gamma_{030}-c^2\gamma_{012})\frac{\partial\psi_{012}}{\partial x}\Bigg\}.\end{aligned} \quad (48)$$

The formula for $B_z^{(e)}$ can be obtained from Eq. (48) with the help of mutual substitution $y \longleftrightarrow z$.

Instead of substituting formula (38) into these expressions, we represent (47) and (48) in terms of magnetic multipoles of the surface current analogously to writing relations (15) and (22). Using simple transformations based on equalities (30) and (34)–(36), relations (13) and (44), the irreducibility (i.e., vanishing of the convolution in any two indices) of the octupole magnetic moment tensor and the tensor potential of the homeoid, as well as the easily verifiable identity

$$\varphi_{100} = \psi_{100}+\frac{18}{7}(a^2-b^2)\psi_{120}+\frac{18}{7}(a^2-c^2)\psi_{102}, \quad (49)$$

where

$$\psi_{100}\equiv\psi_x = \frac{3}{abc}\mathcal{M}_{100}x,$$

we can transform expressions (47) and (48) to the universal vectorial form

$$\mathbf{B}^{(e)} = -\mathrm{grad}\Psi^{(e)}. \quad (50)$$

The pseudoscalar magnetic field potential appearing in this relation in tensor form can be described by the

formula

$$\Psi^{(e)} = m_x^{(\mathbf{i})}\psi_x + (m_{xxx}^{(\mathbf{i})} - 9m_x^{(\mathbf{i})}a^2)\psi_{xxx} + 3(m_{xyy}^{(\mathbf{i})} - 3m_x^{(\mathbf{i})}b^2)\psi_{xyy} + 3(m_{xzz}^{(\mathbf{i})} - 3m_x^{(\mathbf{i})}c^2)\psi_{xzz}. \quad (51)$$

The following remark is appropriate here. If we considered not currents but scalar sources (e.g., charges), the electrostatic potential outside ellipsoid (2) carrying a bulk charge of density $\rho \sim x$ would be given, according to [14], by any of formulas (22) or (51), in which magnetic multipoles are replaced by the corresponding electrical multipoles. A formula analogous to (51) also describes the electrostatic potential outside the ellipsoid, which is produced by its surface charge of density $\sigma = P_3(x, y, z)p$, where P_3 is a cubic polynomial whose coordinate dependence is characterized by the same symmetry as for the bulk charge.[5]

It is clear from the above arguments that, if the pseudoscalar potential possesses the same multipole representations as the electrostatic potential, the substitution of expressions (18)–(21) for multipole moments of bulk currents into formula (51) for surface currents must lead to formula (22). An elementary verification also based on identity (49) and on irreducibility of the tensor potential of the homeoid confirms this statement. Thus, currents (3) and (23) produce identical fields outside the ellipsoid; this fact was established without using formulas (38) explicitly. As regards these formulas, their substitution into relation (23) provides the solution to our problem.

If ellipsoid (2) degenerates into a sphere ($c = b = a$), the density of bulk current (3) in view of (7) is transformed into

$$j_x = 0, \quad j_y = -C_{010}\frac{z}{a}, \quad j_z = C_{010}\frac{y}{a},$$

while the equivalent surface currents (23) is simplified, as expected [8], to the form

$$i_x = 0, \quad i_y = -\frac{1}{5}C_{010}z, \quad j_z = \frac{1}{5}C_{010}y.$$

CONCLUSIONS

Summing up, we confine ourselves to the following remarks.

(1) The fact that multipole representations (22) and (51) of pseudoscalar magnetic potentials of ellipsoids for currents (3) and (23) also exist after the replacement of magnetic multipoles by electric ones (these representations are transformed into the corresponding exact formulas for multipole representations of electrostatic potentials)[6] is an indication of the existence of multipole representations of magnetic potentials of the ellipsoid in the general case of currents with a polynomial coordinate dependence.

(2) It is shown, using the specific example of simplest equivalent bulk and surface currents, that the multipole representations of their pseudoscalar potentials, which can be expressed in terms of tensor potentials of the homeoid, are completely identical.[7]

(3) Let us suppose that a superconductor filling the entire space has a cavity which is bounded by ellipsoidal surface (2) and carries a preset current (3). In this case, a current with a surface density $\mathbf{i}'(\mathbf{r}) = -\mathbf{i}(\mathbf{r})$ will be induced on the boundary surface of the cavity, where $\mathbf{i}(\mathbf{r})$ is the solution to the problem considered here, which is described by formulas (23) and (38). Naturally, an infinite superconductor can be replaced by a finite-size grounded superconducting body.

REFERENCES

1. J. C. Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed. (Clarendon, Oxford, 1892; Nauka, Moscow, 1989), Vol. 2.
2. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1992; Pergamon, New York, 1984).
3. J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill, New York, 1941; Gostekhizdat, Moscow, 1948).
4. M. L. Levin and R. Z. Muratov, Zh. Tekh. Fiz. **47**, 2464 (1977) [Sov. Phys. Tech. Phys. **22**, 1430 (1977)].
5. R. Z. Muratov, Zh. Tekh. Fiz. **57**, 2097 (1987) [Sov. Phys. Tech. Phys. **32**, 1267 (1987)].
6. Ya. I. Frenkel, *Electrodynamics (General Theory of Electricity): Collection of Selected Works* (Akad. Nauk SSSR, Moscow, 1959), Vol. 1 [in Russian].
7. Ya. I. Frenkel, *Lehrbuch der Elektrodynamik* (Springer, Berlin, 1926; ONTI, Moscow, 1935), Vol. 2.
8. R. Z. Muratov, Zh. Tekh. Fiz. **72** (4), 6 (2002) [Tech. Phys. **47**, 380 (2002)].
9. R. Z. Muratov, *Ellipsoid Potential* (Atomizdat, Moscow, 1976) [in Russian].
10. S. P. Efimov and R. Z. Muratov, Astron. Zh. **67**, 302 (1990) [Sov. Astron. **34**, 152 (1990)].
11. R. Z. Muratov, Astron. Zh. **70**, 1271 (1993) [Astron. Rep. **37**, 641 (1993)].
12. R. Z. Muratov, Zh. Tekh. Fiz. **67** (4), 1 (1997) [Tech. Phys. **42**, 325 (1997)].
13. B. V. Medvedev, *Principles of Theoretical Physics* (Nauka, Moscow, 1977) [in Russian].
14. S. P. Efimov and R. Z. Muratov, Astron. Zh. **67**, 314 (1990) [Sov. Astron. **34**, 157 (1990)].

Translated by N. Wadhwa

[5] For example, the volume charge $\rho \sim x$ corresponds to P_3 containing only monomials x^3, xy^2, and xz^2.

[6] Naturally, the total charge of the ellipsoid must be set equal to zero in these formulas in view of the absence of magnetic charges.

[7] It should be noted that the vector potential of the magnetic field outside the ellipsoid does not possess a universal property of this type. This follows, for example, from the fact that the number of nonzero vector potential components is not the same for equivalent currents (3) and (23).

Technical Physics, Vol. 50, No. 8, 2005, pp. 967–975. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 7–14.

THEORETICAL AND MATHEMATICAL PHYSICS

Possible Spacing between Linear Vortices in a 3D Ordered Josephson Medium

M. A. Zelikman

St. Petersburg State Polytechnical University, St. Petersburg, 195251 Russia

Received December 17, 2004

Abstract—A method is proposed for solving the nonlinear system of equations of fluxoid quantization for two interacting linear vortices. It is shown that the centers of the vortices may lie in adjacent cells only if the pinning parameter $I > 0.91$, in alternate cells if $I > 0.44$, and in each third cell if $I > 0.25$. These critical values are substantially lower than analogous values for planar vortices. It is shown that, as the value of I tends to zero, the minimal spacing between linear vortices does not increase indefinitely, but attains a certain finite value and then remains unchanged. This means that pinning of linear vortices cannot be ignored even for values of I quite close to zero. It is shown that two linear vortices with centers in the neighboring cells along a diagonal may coexist for indefinitely small values of I. *© 2005 Pleiades Publishing, Inc.*

INTRODUCTION

One of the main problems in the physics of high-temperature superconductors (HTSCs) is analysis of the structure, motion, and pinning of vortices emerging in the sample placed in an external magnetic field. A large number of publications [1–7] are devoted to analysis of vortices. The behavior of 1D vortices in a long Josephson junction is analyzed in [3, 4]. However, in this analysis, a vortex is assumed to be a spatially continuous phase distribution and its pinning is explained by the interaction with discretely distributed pinning centers. In fact, a Josephson medium is a cellular structure that ensures pinning, which is determined by the energy required for displacement of the vortex center from one cell to another.

The behavior of a vortex in a linear SQUID chain was analyzed in [5]. However, analysis is carried out in this case in the 2D configuration; i.e., the magnetic field of an individual loop is taken into account only in the magnetic flux piercing the loop itself. However, in the 3D case, a vortex is a system of coaxial "solenoids"; consequently, the magnetic flux through the loop is produced not only by the loop itself, but also by other current-carrying subcircuits (including remote ones). As the critical current in the junction decreases, the vortex size (i.e., the number of loops participating in the formation of the magnetic flux through the central cell of the vortex) increases, which compensates for the decrease in the contribution from each loop to the magnetic flux. A system of equations of fluxoid quantization in cells was obtained in [6, 7] for a 3D ordered Josephson medium.

The equilibrium configurations of a laminar (planar) vortex in a 3D ordered Josephson medium were calculated in [8], where the minimal spacings between two planar vortices were determined. Various configurations of a linear vortex in such a medium were analyzed in [9]. This study is aimed at calculating minimal possible spacing between two linear vortices in a 3D Josephson medium for various values of the pinning parameter. This problem is of considerable practical interest, since such vortices are encountered in most HTSC materials.

We will carry out our analysis using a model in the form of a cubic grid with period a, formed by superconducting wires of diameter δ, in which each link contains a Josephson junction; all junctions are small in size and have the same value of critical current J_c. The current distribution has a planar structure (i.e., currents are distributed identically in all parallel planes perpendicular to the vortex axis and spaced by distance a). In spite of its simplicity, this model makes it possible to draw quite definite conclusions concerning the structure of vortices, their pinning, and dynamics. These results are qualitatively preserved in more complex models.

LINEAR VORTICES WITH CENTERS IN ADJACENT CELLS

Let us consider the possibility of the existence of two isolated linear vortices of the same orientation with centers lying in adjoining cells far away from the sample boundary.

Figure 1 shows a quarter of the section of this pattern by a plane perpendicular to its axis. Central cells of vortices are marked by bold dots. Since the vortices are oriented identically, a symmetry plane exists, in which all phase jumps at the junctions are equal to zero. A complete pattern can be obtained by symmetric extrapolation of Fig. 1 to the left and downwards.

1063-7842/05/5008-0967$26.00

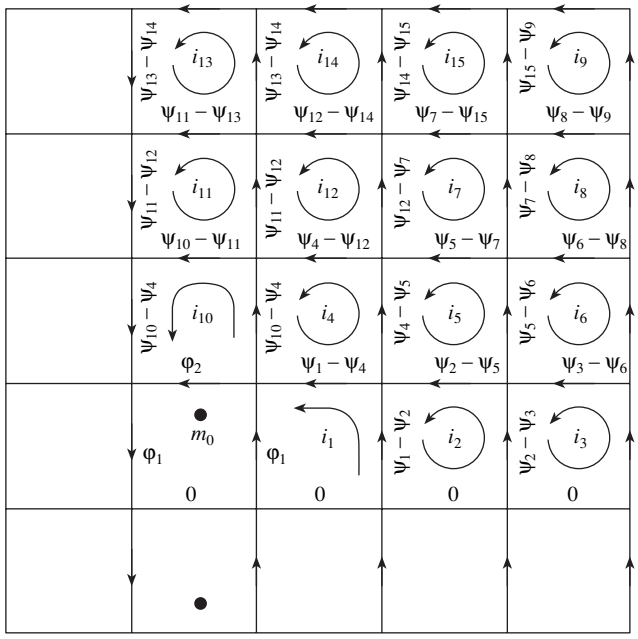


Fig. 1. Distribution of currents and phase jumps at the junctions for two linear vortices with the centers in adjoining cells in the plane perpendicular to their axes; i_m and ψ_m are the loop current in the mth loop and the corresponding "loop" phase jump. The phase jumps expressed in terms of ψ_m are indicated above each junction. Jumps φ_1 and φ_2 are not regarded as small. The cells marked by bold dots contain a flux quantum Φ_0, while the remaining cells do not contain a flux quantum. The pattern can be symmetrically continued to the left and downwards.

In each cell, the condition of fluxoid quantization [6] is satisfied,

$$2\pi\Phi_m/\Phi_0 + \sum_k \varphi_k^{(m)} = 2\pi K_m, \qquad (1)$$

where $\sum_k \varphi_k^{(m)}$ is the sum of phase jumps at the Josephson elements of the mth cell, Φ_m is the total magnetic flux through the mth cell, Φ_0 is the magnetic flux quantum, and K_m is an integer equal to unity for central cells of the vortices (with a current i_0) and zero for all remaining cells.

With increasing distance from the centers, the Josephson currents $J_k = J_c\sin\varphi_k$ decrease, the rate of the decrease increasing with the critical current J_c. In this case, we consider only the configurations in which $\varphi_k \ll 1$; i.e., $\sin\varphi_k \approx \varphi_k$ for all φ_k except for the largest jumps in phases φ_1 and φ_2 in the cells nearest to the center (see Fig. 1). The validity of this assumption will be confirmed by the result. To avoid writing the balance conditions for currents in junctions, it is convenient to use the loop current method for the cells. Let us suppose that a loop current $i_m = J_c\psi_m$, where ψ_m is the loop phase jump, flows in the mth cell ($m \neq 0$). Then the phase jumps φ_k at junctions (except φ_1 and φ_2) are defined as the differences in the corresponding loop values (see Fig. 1). The magnetic flux through the mth cell can be written in the form [6]

$$\Phi_m = \mu_0 S/a\left(i_m + b\sum_k J_k^{(m)}\right) \quad (m \neq 0), \qquad (2a)$$

$$\Phi_m\big|_{m=0} = \mu_0 S/a(i_0 + i_1 + 4bi_0) \quad (m = 0), \qquad (2b)$$

where b is the coefficient of field nonuniformity [6] due to discreteness of the current distribution along the vortex axis, $\sum_k J_k^{(m)}$ is the algebraic sum of currents in the junctions of the mth cell, and S is the cell area.

For a wire thickness $\delta \ll a$, parameter b is defined as

$$b = -\frac{1}{2\pi}\ln\left(2\sinh\frac{\pi\delta}{a}\right).$$

For simplicity, we will henceforth consider the case of $Ib \ll 1$, where I is the pinning parameter defined by formula (4). It should be noted that such an analysis is also valid for a structure formed by superconducting filaments glued to one another over the entire length so that the surfaces of contact between filaments play the role of long Josephson contacts. The cross section of the structure must be in the form of a square lattice, in which the shape of a cell is not necessarily a square (in particular, it can be formed by filaments of a circular cross section).

Substituting relations (2) into formula (1) and assuming that $i_0 = J_c\sin\varphi_1$ and loop currents $i_k = J_c\psi_k$ ($k \neq 0$), we can write (for $Ib \ll 1$) a system of equations for fluxoid quantization (m is the cell number in Fig. 1),

$$\begin{gathered}
I(\sin\varphi_1 + \psi_1) + 2\varphi_1 + \psi_2 = 2\pi \ (m = 0),\\
(I+2)\psi_1 - \psi_4 - \psi_2 - \varphi_1 = 0 \ (m = 1),\\
(I+3)\psi_{10} - \psi_2 - 2\psi_4 - \psi_{11} = 0 \ (m = 10),\\
(I+4)\psi_4 - \psi_1 - \psi_5 - \psi_{10} - \psi_{12} = 0 \ (m = 4),\\
(I+3)\psi_2 - \psi_1 - \psi_5 - \psi_3 = 0 \ (m = 2),\\
(I+4)\psi_5 - \psi_2 - \psi_4 - \psi_6 - \psi_7 = 0 \ (m = 5),\\
(I+4)\psi_{11} - \psi_{10} - 2\psi_{12} - \psi_{13} = 0 \ (m = 11),\\
(I+4)\psi_{12} - \psi_4 - \psi_7 - \psi_{11} - \psi_{14} = 0 \ (m = 12),\\
(I+4)\psi_7 - \psi_5 - \psi_8 - \psi_{12} - \psi_{15} = 0, \text{ etc. } (m = 7),\\
\sin\varphi_1 + \psi_1 = \sin\varphi_2 + \psi_{10},
\end{gathered} \qquad (3)$$

where

$$I \equiv 2\pi\mu_0 J_c a/\Phi_0. \qquad (4)$$

The last equation in system (3) was derived from the additional condition of current continuity in the upper

right junction of the cell with number 0, which is denoted by a bold dot in Fig. 1.

A transition from infinite system (3) to a finite system can be carried out by disregarding currents at large distances from the center (Fig. 1). We assume that the loop currents in the extreme upper and extreme right rows of cells (Fig. 1) and farther are zero. It is clear from symmetry considerations that the situation with zero currents in the extreme upper junctions in Fig. 1 approximately corresponds to the presence of another configuration, which is symmetric to that shown in Fig. 1 and is centered at a cell separated by seven cells from the cell with $m = 0$ (strictly speaking, this is not completely true, since the currents in the vertical subcircuits of the upper row of cells would not be equal to zero in this case). The effect of this remote configuration can be ignored as compared to the interaction of the two vortices in question with the centers in adjoining cells. The same is true of zero currents in the extreme right row. These remarks justify the analysis of the problem in such geometry.

In this case, system (3) contains only ten equations written explicitly above, in which we must set $\psi_3 = \psi_6 = \psi_8 = \psi_9 = \psi_{13} = \psi_{14} = \psi_{15} = 0$. This system will be solved by expressing all values of ψ_m in terms of φ_1 and φ_2 from linear equations of system (3) and substituting them into the first and last equations of system (3). This gives a system of two nonlinear equations,

$$\varphi_2 = (2\pi - I\sin\varphi_1 - A(I)\varphi_1)/B(I), \quad (5a)$$

$$\varphi_1 = (2\pi - I\sin\varphi_2 - C(I)\varphi_2)/D(I), \quad (5b)$$

where A, B, C, and D are polynomials in I.

For values of pinning parameter I larger than a certain critical value $I_0 = 0.9093$ (the subscript corresponds to the number of cells between the centers), curves (5) have two points of intersection (Fig. 2) corresponding to different possible configurations of vortices (both *a* or both *b* [9]). The values of φ_1 and φ_2 of the points of intersection can be determined numerically. Knowing these values, we can also find all values of ψ_m. For $I < 0.9093$, curves (5) have no intersections at all. This means that the centers of vortices cannot lie in neighboring cells for such values of I. Figure 3 shows the calculated distribution of the phase jump over the junctions of a vortex for $I = 0.9093$. It can be seen that the assumption concerning the smallness of all values of φ_k except φ_1 and φ_2 is valid.

Another argument in favor of the geometry chosen here is the fact that analogous calculation gives the value of $I_0 = 0.9758$ under the assumption of nonzero currents only in cells with numbers 0, 1, 10, and 4 (and symmetric to them). This means that a further increase in the size of the configuration in question does not appreciable affect the value of I_0. Consequently, we can state with confidence that the centers of linear vortices may lie in neighboring cells only for $I > 0.9$.

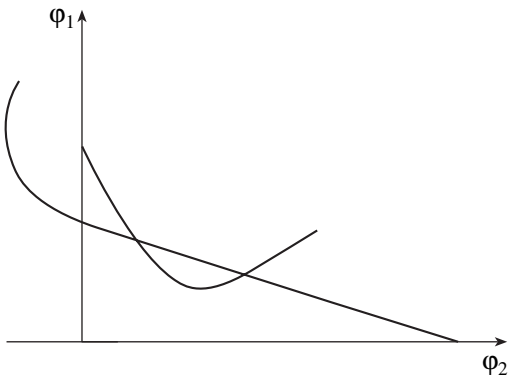


Fig. 2. Intersection of the curves described by Eqs. (7).

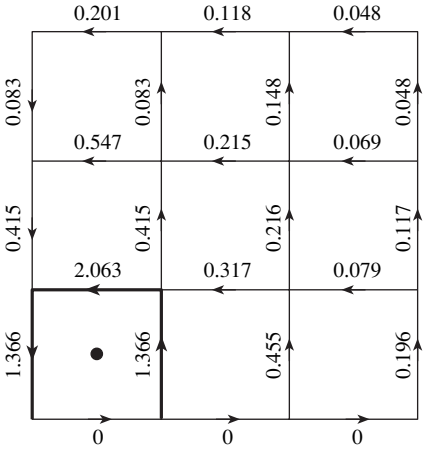


Fig. 3. Distribution of phase jumps for two linear vortices with the centers in the neighboring cells for $I = 0.9093$. Bold lines show the segments on which phase jumps are not regarded as small. The pattern can be symmetrically continued to the left and downwards.

For planar vortices, the value of I_0 is higher (2.907, according to [8]). This can be explained as follows. If the superposition principle were satisfied, i.e., if we could assume that a planar vortex is a continuous row of linear vortices (which is certainly not true since the equations are nonlinear in the central part of the vortex), all linear vortices of another plane vortex would exert forces on each vortex from this row. The resultant force would be stronger than the force of interaction of two linear vortices; i.e., it is more difficult for planar vortices to be confined in the neighboring cells than for linear vortices. For this reason, the critical value of I_0 for planar vortices is larger than for linear ones.

Figure 4 shows the phase jump distribution over junctions for two linear vortices located in adjacent rows of chains with the centers located in alternate cells in each chain. According to calculations, the critical value of I for this situation is 1.64. It should be noted that if we confine our calculation not to three vertical rows from the central cell of a vortex as in Fig. 4, but

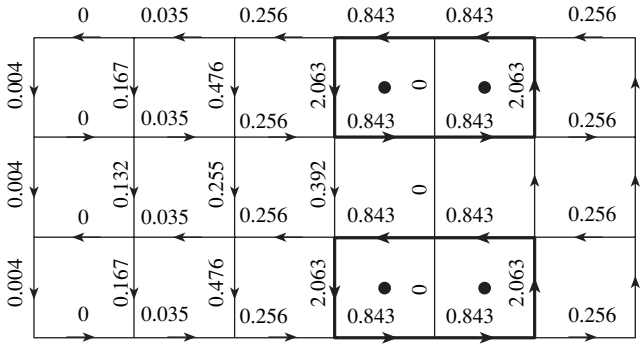


Fig. 4. Distribution of phase jumps over junctions for two linear vortices located in neighboring vertical rows of infinite chains, whose centers in each chain are in alternate cells. Bold lines show the segments on which phase jumps are not regarded as small. The pattern can be continued periodically upwards and downwards and symmetrically to the right. The cells marked by bold dots contain a flux quantum, while all remaining cells do not contain flux quanta.

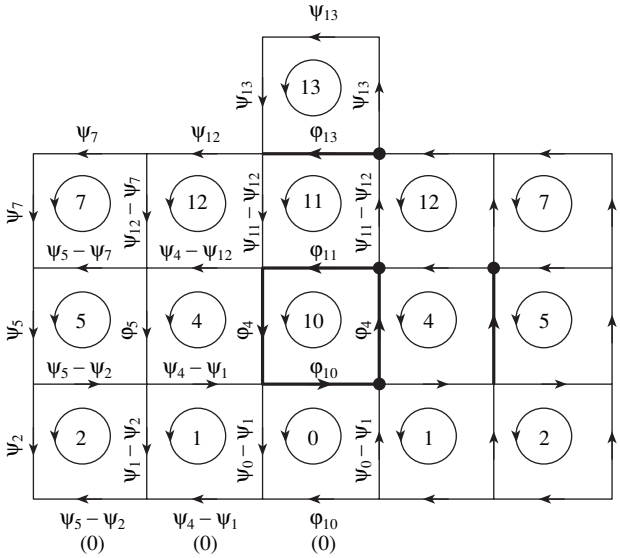


Fig. 5. Distribution of phase jumps over junctions for two linear vortices with the centers in every second or every third cell. The phase jump expressed in terms of "loop" phase jumps ψ_m is indicated over each junction. The upper values in the lower row are obtained for the centers of the vortices located in alternate cells, while the lower values (zeros in the parentheses) are for the centers lying in every third cell. Bold lines mark the segments on which the phase jumps are not regarded as small. The pattern can be symmetrically continued to the right and downwards.

only to two rows, the critical value of I will be 1.65. The closeness of these values indicates that they are close the exact value. The configuration considered here is intermediate between the previous case of two linear vortices with the centers in neighboring cells and two planar vortices with the centers in adjacent rows. The force acting in this case on each linear vortex is stronger than in the case of two linear vortices, but weaker than for planar vortices. The same should be expected for the critical values of the pinning parameter.

CENTERS OF VORTICES LOCATED IN ALTERNATE CELL

It follows from symmetry considerations that half the section of this pattern by a plane perpendicular to its axis has the form shown in Fig. 5. A transition from an infinite system of equations to a finite system will be carried out in the same way as in the case considered above (i.e., by ignoring currents far away from the centers). We confine our analysis to the cells shown in Fig. 5. The central cell of the vortex is denoted by a bold dot. The complete pattern can be obtained by symmetric extrapolation of Fig. 5 in the downward direction.

It can be seen as a result of calculations that phase jumps at all junctions except subcircuits shown by bold lines in Fig. 5 (namely, φ_4, φ_5, φ_{10}, φ_{11}, and φ_{13}) are small (such that $\sin\varphi_k \approx \varphi_k$) for all values of I for which the given configuration can exist. We introduce the loop currents $i_m = J_c\psi_m$ for $m \neq 10$. This allows us to automatically satisfy the current balance conditions in all junctions except those shown by bold dots in Fig. 5.

The current balance conditions at these junctions have the form

$$\begin{aligned}
\sin\varphi_{13} &= \psi_{11} - \psi_{13},\\
\psi_4 + \sin\varphi_4 &= \psi_{11} + \sin\varphi_{11},\\
\psi_0 + \sin\varphi_{10} &= \psi_4 + \sin\varphi_4,\\
\psi_5 + \sin\varphi_5 &= \psi_4 + \sin\varphi_4.
\end{aligned} \tag{6}$$

Analogously to relation (3), we derive the following system of equations of fluxoid quantization in cells for $b = 0$ (m is the cell number in Fig. 5):

$$\begin{aligned}
I(\sin\varphi_4 + \psi_4) + 2\varphi_4 + \varphi_{10} + \varphi_{11} &= 2\pi \ (m = 10),\\
(I+2)\psi_{11} + \psi_{13} - 2\psi_{12} - \psi_{11} &= 0 \ (m = 11),\\
(I+2)\psi_0 - 2\psi_1 - 2\varphi_{10} &= 0 \ (m = 0),\\
(I+2)\psi_4 - \psi_1 - \psi_{12} + \varphi_5 - \varphi_4 &= 0 \ (m = 4),\\
(I+3)\psi_{13} - 2\psi_{14} - \varphi_{13} &= 0 \ (m = 13),\\
(I+3)\psi_5 - \psi_2 - \psi_7 - \varphi_5 &= 0 \ (m = 5),\\
(I+4)\psi_1 - \psi_0 - \psi_2 - 2\psi_4 &= 0 \ (m = 1),\\
(I+4)\psi_2 - \psi_1 - 2\psi_5 &= 0 \ (m = 2),\\
(I+4)\psi_{12} - \psi_4 - \psi_7 - \psi_{11} - \psi_{14} &= 0 \ (m = 12),\\
(I+4)\psi_7 - \psi_5 - \psi_{12} &= 0 \ (m = 7).
\end{aligned} \tag{7}$$

System of equations (6), (7) differs from system (3) in that it is nonlinear not in two but in more (five in our case) unknowns. Consequently, this system cannot be

reduced to a system of two nonlinear equations, which can be solved numerically. The problem can be solved using the method of successive approximations. We present φ_5, φ_{10}, and φ_{13} as well as their sines in the form $\varphi_5 = \varphi_5^0 + \delta_5$, $\sin\varphi_5 = \sin\varphi_5^0 + \cos\varphi_5^0\delta_5$, etc., where φ_5^0, φ_{10}^0, and φ_{13}^0 are assumed to be known quantities (equal to zero in the first iteration), while δ_5, δ_{10}, and δ_{13} are new unknowns (instead of φ_5, φ_{10}, and φ_{13}), relative to which the system is linear.

We express the values of all ψ_m, as well as δ_5, δ_{10}, and δ_{13}, in terms of φ_4 and φ_{11} from the linear equations of the system; substituting these values into the remaining two nonlinear equations, we obtain a system of the form

$$\varphi_4 = [2\pi - I\sin\varphi_{11} - E(I)\varphi_{11} - D(I)]/F(I), \quad (8)$$

$$\varphi_{11} = [2\pi - I\sin\varphi_4 - G(I)\varphi_4 - H(I)]/M(I), \quad (9)$$

where E, D, F, G, H, and M are rational fractional functions of I, whose coefficients depend on values of φ_5^0, φ_{10}^0, and φ_{13}^0.

As in the previous case, for pinning parameters I exceeding a certain critical value I_1, curves (8) and (9) have two points of intersection (as in Fig. 2) corresponding to various possible configurations of vortices (both of type *a* or both of type *b* [9]). The values of φ_4 and φ_{11} at the points of intersection can be found numerically. Knowing the values of φ_4 and φ_{11}, we obtain δ_5, δ_{10}, δ_{13}, and all values of ψ_m. For the next iteration, we take new values of φ_5^0, φ_{10}^0, and φ_{13}^0; add them to the previous values of δ_5, δ_{10}, and δ_{13}; and solve the system again. After each such iteration, we must find the corrected value of I_1 and carry out the next iteration for the new value of I_1. The results of calculations show that the iterative procedure converges (i.e., each next iteration gives values of δ_5, δ_{10}, δ_{13}, which are smaller by several orders of magnitude). In this way, the initial system of equations can be solved to the required degree of accuracy in just a few steps. The authenticity of the resulting solutions was verified by their substitution into all initial equations of system (6), (7).

Calculations show that the critical value of I_1 is equal to 0.441. (It should be noted that the value of I_1 before the iterative procedure was equal to 0.422.)

Figure 6 shows the calculated distributions of phase jumps over the vortex junctions for $I = I_1 = 0.450$. It can be seen that the condition of the correctness of calculation ($\sin\varphi_k \approx \varphi_k$ for all values of φ_k except φ_4, φ_5, φ_{10}, φ_{11}, and φ_{13}) is satisfied.

It should be noted that $I_1 = 1.947$ for planar vortices [8].

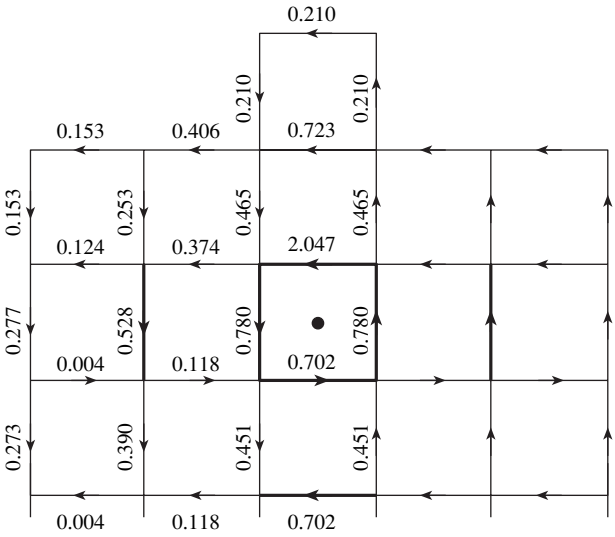


Fig. 6. Distribution of phase jumps for two linear vortices with the centers in alternate cells for $I = 0.441$. Bold lines show the segments on which phase jumps are not regarded as small. The pattern can be symmetrically continued to the right and downwards.

CENTERS OF VORTICES LOCATED IN EVERY THIRD CELL

It follows from symmetry considerations that half the section of this pattern by a plane perpendicular to its axis has the same form as in the previous case (see Fig. 5), the only difference being that phase jumps on the segments of the lowermost line are equal to zero. As before, we confine our analysis to the cells shown in Fig. 5. The central cell of the vortex is marked by a bold dot. The complete pattern can be obtained by symmetric extrapolation of Fig. 5 in the downward direction.

The equations for cells with numbers 0, 1, and 2 in system (6) and (7) should be replaced by the following equations:

$$(I+2)\psi_0 - 2\psi_1 - \varphi_{10} = 0 \ (m = 0),$$
$$(I+3)\psi_1 - \psi_0 - \psi_2 - \psi_4 = 0 \ (m = 1), \quad (10)$$
$$(I+3)\psi_2 - \psi_1 - \psi_5 = 0 \ (m = 2).$$

All remaining equations of system (6) and (7) are unchanged.

We will assume that phase jumps at the same contacts φ_4, φ_5, φ_{10}, φ_{11}, and φ_{13} are not small. All the remarks concerning the previous case remain valid in this case also.

Calculations analogous to those presented above show that the critical value I_2 for vortices whose centers are spaced by two cells from one another is equal to 0.246.

Figure 7 shows the phase jump distribution over the junctions of the vortex calculated for $I = I_2 = 0.246$. It can be seen that the condition for the validity of calcu-

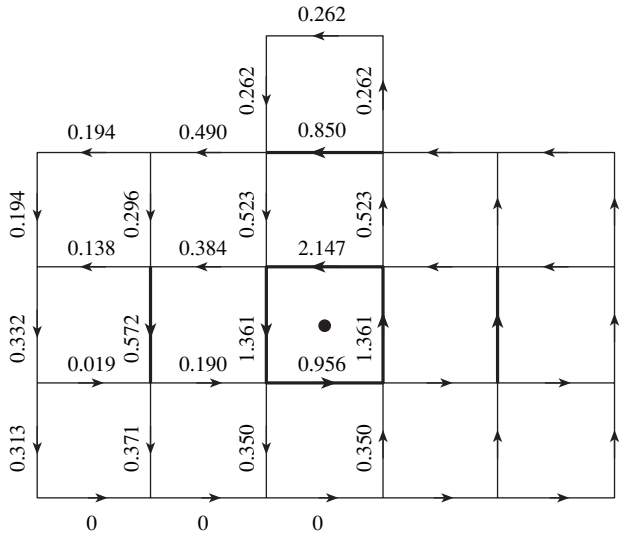


Fig. 7. Distribution of phase jumps for two linear vortices with the centers in every third cell for $I = 0.246$. Bold lines show the segments on which phase jumps are not regarded as small. The pattern can be symmetrically continued to the right and downwards.

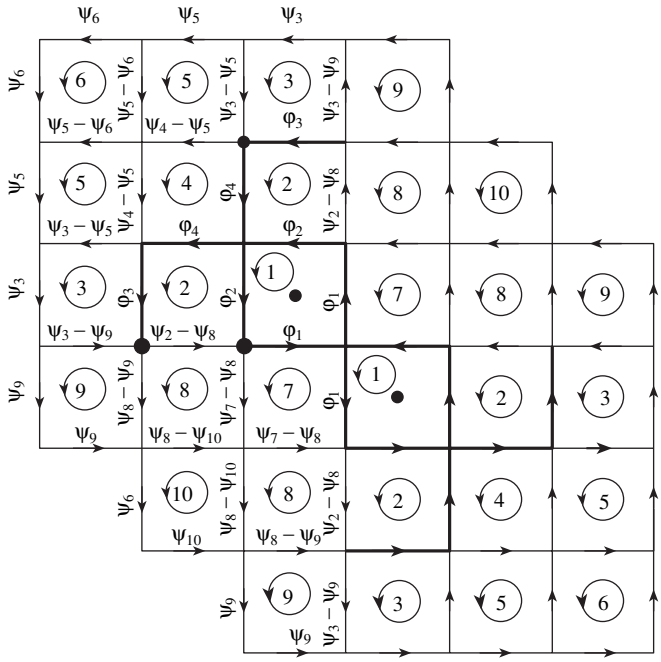


Fig. 8. Distribution of phase jumps over junctions for two linear vortices with the centers in neighboring cells on the diagonal. The phase jump expressed in terms of "loop" phase jumps ψ_m is indicated over each junction. Bold lines mark the segments on which the phase jumps are not regarded as small. The cells marked by bold dots contain a flux quantum, while all remaining cells do not contain flux quanta. The pattern is symmetric relative to diagonals.

lation ($\sin\varphi_k \approx \varphi_k$ for all values of φ_k except φ_4, φ_5, φ_{10}, φ_{11}, and φ_{13}) is satisfied.

It should be noted that $I_2 = 1.428$ for planar vortices [8].

LINEAR VORTICES WITH CENTERS LOCATED IN ADJACENT CELLS ALONG THE DIAGONAL

The section of this pattern by a plane perpendicular to its axis is shown in Fig. 8. The central cells of the vortices are marked by bold dots. The segments on which the phase jumps are not regarded a priori as small (φ_1, φ_2, φ_3, and φ_4) are shown by bold lines in Fig. 8.

We introduce the loop currents $i_m = J_c\psi_m$ for $m \neq 1$. This allows us to automatically satisfy the current balance conditions at all junctions except those marked by bold dots in Fig. 8.

The current balance conditions at these junctions have the form

$$\begin{aligned}\psi_3 + \sin\varphi_3 &= \psi_4 + \sin\varphi_4,\\ \psi_3 + \sin\varphi_3 &= \psi_2,\\ \psi_2 + \sin\varphi_2 &= \psi_7 + \sin\varphi_1.\end{aligned} \tag{11}$$

Analogously to system (3), we obtain the following system of equations for fluxoid quantization in cells (m is the number of a cell in Fig. 8):

$$\begin{aligned}I(\sin\varphi_1 + \psi_7) + 2\varphi_1 + 2\varphi_2 &= 2\pi \ (m = 1),\\ (I+1)\psi_2 + \varphi_3 + \varphi_4 - \varphi_2 - \psi_8 &= 0 \ (m = 2),\\ (I+3)\psi_3 - \psi_5 - \varphi_3 - \psi_9 &= 0 \ (m = 3),\\ (I+2)\psi_4 - 2\psi_5 - 2\varphi_4 &= 0 \ (m = 4),\end{aligned}$$

$$\begin{aligned}(I+4)\psi_5 - \psi_3 - \psi_4 - \psi_6 &= 0 \ (m = 5),\\ (I+4)\psi_6 - 2\psi_5 &= 0 \ (m = 6),\end{aligned} \tag{12}$$

$$\begin{aligned}(I+2)\psi_7 - 2\psi_8 - 2\varphi_1 &= 0 \ (m = 7),\\ (I+4)\psi_8 - \psi_2 - \psi_7 - \psi_9 - \varphi_{10} &= 0 \ (m = 8),\\ (I+4)\psi_9 - \psi_3 - \psi_8 &= 0 \ (m = 9),\\ (I+4)\psi_{10} - 2\psi_8 &= 0 \ (m = 10).\end{aligned}$$

System of equations (11), (12) can be solved by the same method of successive approximations as system (6), (7). Let us linearize these equations in φ_3 and φ_4, as it was done in the previous section in φ_5, φ_{10}, and φ_{13}.

Having expressed the values of all ψ_m as well as δ_3 and δ_4 in terms of φ_1 and φ_2 from the linear equations of the system and substituting these values into the two remaining nonlinear equations, we again arrive at a system of form (5). However, in contrast to the cases considered above, Eqs. (5) have only one point of intersection. On the one hand, this should be expected since both vortices in such a diagonal arrangement can have only type *a* configuration [8] and the other point of intersection cannot appear from anywhere. On the other hand, the solution to the system cannot disappear in the manner shown in Fig. 2 upon a decrease in I. Thus, we

can expect that the system will have a solution for any (indefinitely small) values of I.

The same iterative procedure as the one used before can be employed for solving the system to the required degree of accuracy in just a few steps. Calculations confirm the existence of a solution for any value of I. Figure 9 shows the calculated distribution of phase jumps over junctions for $I = 0.0001$. It can be seen that the condition for the validity of calculation ($\sin\varphi_k \approx \varphi_k$ for all values of φ_k except φ_1, φ_2, φ_3, and φ_4) is satisfied, although with a lower degree of accuracy as compared to the previous cases.

The above arguments lead to a nontrivial conclusion: two linear vortices with the centers in the neighboring cells on the diagonal can coexist for indefinitely small values of the pinning parameter.

LINEAR VORTICES AT LARGE DISTANCES FROM EACH OTHER

For small values of the pinning parameter ($I \ll 1$), the pinning force is weak, while the force of interaction between the vortices is strong; consequently, the vortices can be in equilibrium only at large distances from each other.

For small values of I, the vortices can be regarded as quasi-continuous. The energies of solitary vortices at distances a and b can be described by the expressions [9, 10]

$$E_a = E_0\frac{\pi}{2}I\left(\ln\frac{1}{I} + 2.77\right),$$
$$E_b = E_0\frac{\pi}{2}I\left(\ln\frac{1}{I} + 2.81\right), \quad (13)$$

where $E_0 = \Phi_0^2/4\pi^2\mu_0 a^2$, Φ_0 being the magnetic flux quantum and μ_0 the magnetic constant.

The approximate average value of the pinning force is equal to the ratio of the pinning energy, i.e., energy difference (13), to the distance between the positions of the center of the vortex in states a and b, i.e., to half the cell size $a/2$. The maximal pinning force is larger than the mean value. Introducing the corresponding coefficient k, we obtain

$$F_p^{\max} = k0.04E_0\frac{\pi}{a}I. \quad (14)$$

Since vortices for small values of I are described by the same equation [7] as the Abrikosov vortices, we will use the expression for the energy of interaction of two Abrikosov vortices [10] per vortex,

$$U_{\text{int}} = E_0\pi I K_0(x), \quad (15)$$

where K_0 is the zero-order Bessel (Hankel) function of the imaginary argument, $x = r/\lambda$, r is the distance between the centers of the vortices, and $\lambda = a/\sqrt{I}$ is the characteristic vortex size.

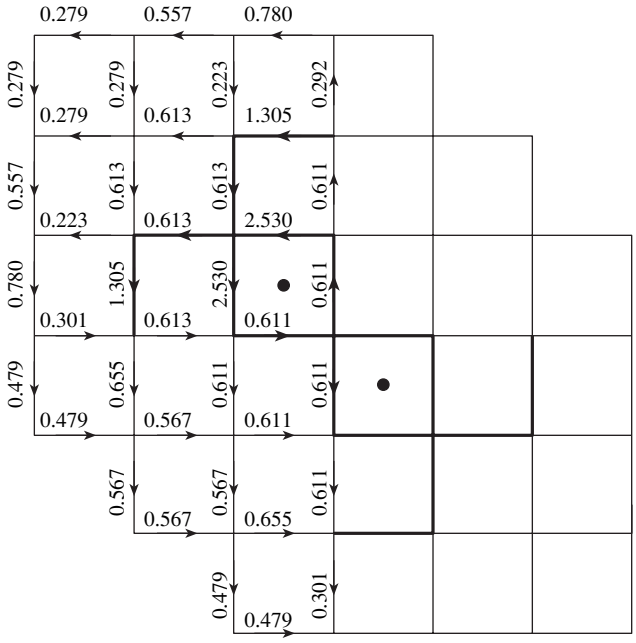


Fig. 9. Distribution of phase jumps over junctions for two linear vortices with the centers in neighboring cells on the diagonal for $I = 0.0001$. Bold lines mark the segments on which the phase jumps are not regarded as small. The cells marked by bold dots contain a flux quantum, while all remaining cells do not contain flux quanta. The pattern is symmetric relative to diagonals.

Equating the force of interaction $F_{\text{int}} = -\partial U_{\text{int}}/\partial r$ to maximal pinning force (14), we arrive at the equation for determining the maximal spacing between the centers of the vortices,

$$\frac{dK_0(x)}{dx} = -k0.04I^{-0.5}. \quad (16)$$

Using the fact that $dK_0/dx = -K_1(x)$, we can write Eq. (16) in the form

$$K_1(x) = k0.04I^{-0.5}. \quad (17)$$

We choose the value of k so that relation (17) leads exactly to the result obtained in one of the previous sections: for $I = 0.25$, the minimal spacing between the central cells of the vortices is equal to two cells; i.e., the distance between the centers of these cells is $3a$. Substituting $\sqrt{I} = 0.5$, $r/a = 3$, and $x = 1.5$ into relation (17), we obtain $k = K_1(1.5)/0.08 = 3.5$, and relation (17) assumes the form

$$K_1(x) = 0.14I^{-0.5}. \quad (18)$$

We analyze Eq. (17) as follows. Choosing a certain value of I, we calculate the value of K_1 corresponding to it by formula (17); using the table of Bessel func-

Calculation of minimal distances between vortices

I	$\sqrt{I}$	$K_1 = 0.14/\sqrt{I}$	x	$r/a = x/\sqrt{I}$
0.0001	0.01	14	0.071	7.1
0.0049	0.07	2	0.43	6.1
0.01	0.1	1.4	0.57	5.7
0.0625	0.25	0.56	1.05	4.2
0.25	0.50	0.28	1.5	3.0
0.45	0.67	0.21	1.7	2.5
1.00	1.00	0.14	2.0	2.0

tions, we obtain the corresponding value of x and then calculate $r/a = x/\sqrt{I}$. The results of such calculations are presented in the table. Strictly speaking, the approximation of quasi-continuous vortices is satisfied only for $\sqrt{I} \ll 1$ (i.e., for the first three lines of the table). However, it follows from the table that the results can be regarded as admissible for large values of I. For example, for $I = 0.45$, the exact calculations carried out in the previous sections give a spacing of one cell between the central cells of the vortices (i.e., the distance between the centers of these cells is equal to two cells. Calculations based on formula (18) give a value of $r/a = 2.5$. Consequently, determining coefficient k in the derivation of relation (18) by applying relation (17) to the case $I = 0.25$ can be regarded as substantiated. The table shows, however, that the use of formula (18) for large values of I gives values of r/a higher than the exactly calculated values. Consequently, the value of r/a obtained from relation (18) for $I = 0.25$ must also be higher than the exact value equal to 3. Thus, the value of k must be smaller than the value calculated above. It should be noted that the obtained ratio of the maximal pinning force to the mean value ($k = 3.5$) appears as quite plausible on account of the approximate derivation of expressions (13) in [9].

For small values of x, function $K_0(x)$ is approximately equal to $\ln(1/x)$, and relation (16) implies that the minimal spacing between the vortices is $r/a = 1/0.04k = 7.1$ cells irrespective of the value of I. In this case, $x = 7.1\sqrt{I}$; in other words, the condition of smallness of x assumes the form $\sqrt{I} \ll 0.14$ (i.e., $I \ll 0.02$).

This means that, as the value of I tends to zero, the minimal spacing between linear vortices in a discrete medium does not increase indefinitely (as in the case of planar vortices [8] as well as Abrikosov vortices in a continuous medium), but attains a certain finite value and then remains unchanged. This fact indicates that, as I tends to zero, we cannot disregard pinning of linear vortices, while pinning of planar vortices can be ignored.

This is due to the fact that the width of a vortex tends to infinity upon a decrease in I for planar vortices. Consequently, when a vortex is displaced by a distance smaller than its size, the change in the vortex energy (and, hence, the pinning force) decreases at a much higher rate than the force of interaction between vortices as the value of I tends to zero [8]. On the other hand, a linear vortex for indefinitely small values of I has the central part of a few cells in size [9], where phase jumps are much larger than in the remaining region of the vortex, and the shape of the central part does not change as I tends to zero. The pinning force of such a vortex decreases for I tending to zero at a much lower rate than for planar vortices; consequently, for very small values of I, the maximal pinning force and the mutual repulsion force of linear vortices exhibit the same dependence on I.

CONCLUSIONS

The method proposed for solving a nonlinear system of equations of fluxoid quantization in cells of a 3D ordered Josephson medium makes it possible to determine the configuration of currents for two interacting linear vortices, as well as the minimal possible spacing between them for a given value of the pinning parameter. The method is based on approximate linearization of the most nonlinear equations and on the application of the procedure of successive iterations, which leads to the exact solution after a small number of steps. The convergence of the method in all cases is confirmed by verifying the fulfillment of the initial equations of the system.

It is shown that the centers of the vortices can be in adjoining cells only for $I > 0.91$, in alternate cells for $I > 0.44$, and in every third cell for $I > 0.25$. These critical values are substantially lower than their counterparts for planar vortices.

In the case of small values of I, a comparison of the maximal pinning force and the mutual repulsion force between the vortices has led to the nontrivial conclusion that the minimal spacing between linear vortices in a discrete medium with I tending to zero does not increase indefinitely, but attains a certain finite value and then remains unchanged. This means that pinning of linear vortices cannot be ignored even for values of I very close to zero, while pinning of planar vortices can be ignored.

Another nontrivial result is the fact that two linear vortices with the centers in the neighboring cells on the diagonal may coexist for indefinitely small values of the pinning parameter.

The results may be of interest for solving problems concerning the penetration of a magnetic field in a 3D Josephson medium and in determining the magnetic field profile at the surface boundary.

REFERENCES

1. K.-H. Muller, J. C. Macfarlane, and R. Driver, Physica C **158**, 69 (1989).
2. M. S. Rzchowski, S. P. Benz, M. Tinkham, and C. J. Lobb, Phys. Rev. B **42**, 2041 (1990).
3. Y. S. Kivshar and B. A. Malomed, Rev. Mod. Phys. **61**, 763 (1989).
4. V. V. Bryksin and S. N. Dorogovtsev, Zh. Éksp. Teor. Fiz. **102**, 1025 (1992) [Sov. Phys. JETP **75**, 558 (1992)].
5. F. Parodi and R. Vaccarone, Physica C **173**, 56 (1991).
6. M. A. Zelikman, Supercond. Sci. Technol. **10**, 469 (1997).
7. M. A. Zelikman, Supercond. Sci. Technol. **10**, 795 (1997).
8. M. A. Zelikman, Zh. Tekh. Fiz. **71** (7), 44 (2001) [Tech. Phys. **46**, 831 (2001)].
9. M. A. Zelikman, Zh. Tekh. Fiz. **75** (1), 37 (2005) [Tech. Phys. **50**, 36 (2005)].
10. P. De Gennes, *Superconductivity of Metals and Alloys* (Benjamin, New York, 1966; Mir, Moscow, 1968).

Translated by N. Wadhwa

Technical Physics, Vol. 50, No. 8, 2005, pp. 976–986. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 15–25.

THEORETICAL AND MATHEMATICAL PHYSICS

Numerical Study of Converging Shock Waves in Porous Media

A. A. Charakhchyan*, K. V. Khishchenko, V. V. Milyavskiy**, V. E. Fortov**, A. A. Frolova*, I. V. Lomonosov**, and L. V. Shurshalov***

** Dorodnitsyn Computation Center, Russian Academy of Sciences, Moscow, 119991 Russia*

*** Institute for High Energy Densities, Associated Institute for High Temperatures, Russian Academy of Sciences, Moscow, 125412 Russia*

e-mail: chara@ccas.ru

Received January 12, 2005

Abstract—The dependences of the solutions to the hydrodynamic equations of compressed media that describe converging shock waves on the density of a substance ahead of a wave front are studied. The properties of Hugoniot adiabats that can explain the qualitatively different characters of these dependences for the equations of state of perfect gas and condensed matter are analyzed. The one-dimensional problems of converging shock waves in graphite and aluminum are considered, and the two-dimensional problem of the compression of graphite in a steel target with a conical cavity is solved. The latter problem is also investigated in terms of a simple model for a deformable solid that takes into account shear stresses. © *2005 Pleiades Publishing, Inc.*

INTRODUCTION

A continuation of the simulation of the shock compression of graphite placed in conical solid targets [1] led to the necessity of taking into account the initial porosity of the samples. Here, the authors encountered unusual properties of converging shock waves in porous media, which consisted in a significant increase in the pressure with decreasing initial porosity of graphite as compared to the crystalline density, all other things being equal. This stimulated the study of converging one-dimensional spherical and cylindrical shock waves. Earlier [2], we published a brief communication about the preliminary results.

To demonstrate that an increase in the pressure with decreasing initial density is unusual, we give two problems where this situation does not occur. The first problem is the problem of a converging spherical or cylindrical shock wave in a medium with the equation of perfect-gas state, and it has a self-similar solution (e.g., see monograph [3]). In this solution, the pressure at a wave front as a function of the radius is determined by the exponent of self-similarity, which is independent of the initial density, to within a multiplicative constant. The constant in the self-similar solution is specified by the shock-wave intensity far from the center. As will be shown below, when a shock wave passes through an initially porous medium, the pressure at its front decreases with decreasing initial density of the sample. Therefore, one should expect that the pressure at the front of a wave with a given radius also decreases with decreasing initial density. The calculation of the problem with the equation of perfect-gas state supports this conclusion. However, the numerical solution of the problem with a real equation of state for solids has a qualitatively different character: the pressure at the front of a wave with a given radius increases with decreasing initial density. As will be shown in this work, this change in the character of solution is related to the degeneration of the Hugoniot adiabat of a perfect gas in the strong shock wave approximation.

Another example that demonstrates an unusual increase in the pressure when a continuous substance is changed into a porous one is the one-dimensional problem of the entrance of a shock wave into a porous substance, which is a particular case of the problem of the decay of an arbitrary discontinuity for the hydrodynamic equations of compressible media. All the unknown functions depend on one spatial variable x. The equation of state is assumed to be defined in the form $\varepsilon = \varepsilon(p, v)$, where ε is the specific internal energy, p is the pressure, and v is the specific volume. In the initial instant, the half-space with $x > 0$ is occupied by an immovable (with a velocity $u = 0$) porous substance having a pressure p_0, a specific volume v_{00}, a specific volume of continuous substance $v_0 < v_{00}$, and a specific internal energy $\varepsilon_0 = \varepsilon(p_0, v_0)$. A shock wave in this substance is taken to be a moving boundary that has no pores behind it [3]. The wave velocity and the values of the unknown functions behind the wave are related by conventional laws of conservation. In particular, the pressure p and specific volume v behind the wave front are related by the equation of Hugoniot adiabat

$$\varepsilon(p, v) - \varepsilon_0 = (p + p_0)(v_{00} - v)/2. \qquad (1)$$

We assume that the initial substance that has thermodynamic parameters p_1, v_1, and a velocity u_1 and is

1063-7842/05/5008-0976$26.00

compressed by a shock wave exists at $x < 0$. The solution to this problem (e.g., see monograph [4]) has the form of two waves, which can be either a shock wave or a centered rarefaction wave, that pass to the right and to the left from a discontinuity and a contact discontinuity where pressure and velocity are continuous. For the considered case of a shock wave going to the right, the pressure behind the wave fronts can be found from the solution to the algebraic equation

$$U(p, p_1, v_1) + U_s(p, p_0, v_{00}) + u_1 = 0, \qquad (2)$$

which specifies the function $p = p(p_0, v_{00}, p_1, v_1, u_1)$.

The velocity of a substance flux to the left of the front is $U = U_s(p, p_1, v_1)$ at $p > p_1$ or $U = U_r(p, p_1, v_1)$ at $p < p_1$:

$$U_s(p, p_*, v_*) = \sqrt{(p - p_*)(v_* - v(p, v_*, p_*))},$$

$$U_r(p, p_*, v_*) = \int_{p_*}^{p} \frac{dp'}{a(p', v(p', v_*, p_*))}.$$

The functions U_s and U_r have the meaning of a particle velocity behind the shock wave and the rarefaction wave, respectively, with a pressure p that passes through the immovable substance with parameters p_* and v_* and $a(p, v)$ is the Lagrangian sound velocity. In terms of the U_s and U_r functions, the $v(p, v_*, p_*)$ dependence is specified by Hugoniot adiabat (1) and the isentrope $d\varepsilon + pdv = 0$, respectively.

By differentiating Eq. (2), we obtain

$$\frac{\partial p}{\partial v_{00}} = -\frac{\partial U_s(p, p_0, v_{00})/\partial v_{00}}{\partial U_s(p, p_0, v_{00})/\partial p + \partial U(p, p_1, v_1)/\partial p}. \qquad (3)$$

For the derivative in the numerator of Eq. (3), we have

$$\frac{\partial U_s(p, p_0, v_{00})}{\partial v_{00}} = \frac{p - p_0}{2U_s}\left(1 - \frac{\partial v}{\partial v_{00}}\right). \qquad (4)$$

By differentiating Hugoniot adiabat (1), we find

$$\frac{\partial v}{\partial v_{00}} = \frac{1}{1 + \xi}, \quad \xi = \frac{2}{p + p_0}\left(\frac{\partial \varepsilon}{\partial v}\right)_p.$$

Substituting this expression into Eq. (4), we obtain

$$\frac{\partial U_s(p, p_0, v_{00})}{\partial v_{00}} = \frac{(p - p_0)}{2U_s}\frac{\xi}{(1 + \xi)}.$$

If a rarefaction wave goes to the left, the second term in the denominator of Eq. (3) is $\partial U_r/\partial p = 1/a > 0$. As a result, we have

$$\frac{\partial p}{\partial v_{00}} < 0 \qquad (5)$$

when two conditions are satisfied; namely, if the derivative of the velocity behind the shock wave with respect to pressure along the Hugoniot adiabat is $\partial U_s/\partial p > 0$ and if the thermodynamic derivative is $(\partial \varepsilon/\partial v)_p > 0$. Both conditions are likely to be always satisfied. Therefore, an increase in the porosity, i.e., an increase in v_{00}, virtually always results in a decrease in the intensity of a shock wave when it enters into a substance. However, as will be shown in our work, the further motion of a converging shock wave into a porous substance can be accompanied by the formation of a high pressure behind its front; this pressure can be significantly higher than the pressure that would appear in the case of a less porous or even continuous substance.

At present, there is no theory that could analytically represent the properties of the solutions to the equations of fluid dynamics for compressible media that were detected upon a numerical study. Nevertheless, a comparison of the solutions obtained with the properties of Hugoniot adiabats revealed one such property; it can explain the dependence of the pressure at a wave front on the initial density in converging shock waves on the assumption that the density at the wave front only weakly depends on the initial density. This property of Hugoniot adiabats is described in the next section of this work. Then, we present the results of calculation of spherical and cylindrical converging shock waves for the equations of state for graphite, aluminum, and perfect gas. In particular, these results support the fact that the density at a wave front only weakly depends on the initial density for condensed matter. In the last section, we consider the two-dimensional problem of shock compression of graphite placed in a conical steel target in terms of both hydrodynamic equations and the simplest model of dynamics of a deformable solid that takes into account shear stresses.

ON ONE PROPERTY OF HUGONIOT ADIABATS IN SOLIDS

For shock waves in solids, the strong shock wave approximation, where the pressure and internal energy ahead of a wave front may be neglected as compared to their values behind the front, holds true with a high degree of accuracy. Then, Hugoniot adiabat (1) transforms into the following relation between three quantities, namely, between p, v, and v_{00}:

$$\varepsilon(p, v) = p(v_{00} - v)/2. \qquad (6)$$

If an equation of state has the form $\varepsilon(p, v) = p\varphi(v)$, pressure p is eliminated from Eq. (6) and Eq. (6) specifies the function $v = v(v_{00})$. In particular, for the equation of perfect-gas state, we have $\varphi(v) = v/(\gamma - 1)$, where $\gamma = \text{const} > 1$. For the equation of condensed-matter state, such degeneration does not occur, and Eq. (6) specifies the function $p = p(v, v_{00})$. This is illustrated with Fig. 1, which shows Hugoniot adiabats in the form of the dependences of the compression $\sigma = \rho/\rho_0$ on the pressure p at an initial porosity $m_p = \rho_0/\rho_{00} = 1$ and 2 ahead of a wave for the equations of state for

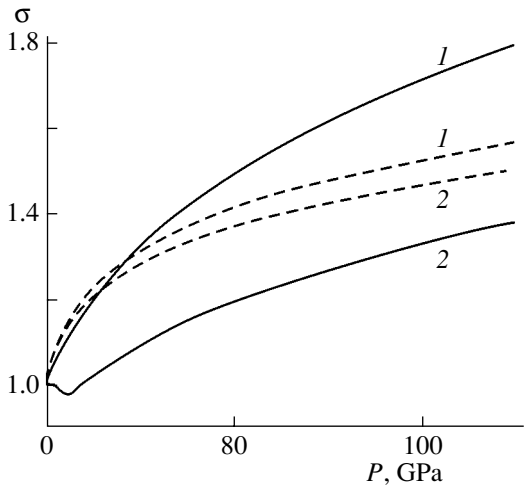


Fig. 1. Hugoniot adiabats of (solid lines) aluminum and (dashed lines) graphite with an initial density m_p = (*1*) 1 and (*2*) 2.

graphite [5, 6] and aluminum [7]. Here, $\rho = v^{-1}$ is the density, $\rho_0 = v_0^{-1}$ is the density of a continuous substance under normal conditions, and $\rho_{00} = v_{00}^{-1}$ is the initial sample density. For graphite, the curves $\sigma(p, \rho_0)$ and $\sigma(p, \rho_0/2)$ are close to each other, which means a weak dependence of ρ on ρ_{00} in the Hugoniot adiabat. For aluminum, these curves differ significantly from each other. Nevertheless, we assume that, when a shock wave moves, the density behind its front ρ, which is a function of the density ρ_{00} and any other parameters characterizing the shock-wave position, depends weakly on ρ_{00}. For example, for one-dimensional shock waves, we assume that the function $\rho = \rho(r, \rho_{00})$, where r is the wave radius, depends weakly on ρ_{00}. As a preliminary basis for this hypothesis, we refer to works [8, 9], where the authors assumed that ρ was independent of ρ_{00} when simulating explosions in pore media. Then, depending on the change in the density ρ_{00} ahead of a wave, an increase or decrease in the pressure behind the wave front p is determined by the sign of the derivative $(\partial p/\partial \rho_{00})_\rho$ or, in terms of specific volume, the sign of the derivative $(\partial p/\partial v_{00})_v$. Let us consider a few examples of Hugoniot adiabats for which this derivative can be found analytically.

By taking the differential of Eq. (6) and eliminating the difference $v_{00} - v$ from the resulting expression using Eq. (6), we obtain

$$\left(\frac{\partial p}{\partial v_{00}}\right)_v = \frac{p^2}{2[p(\partial\varepsilon/\partial p)_v - \varepsilon]}. \qquad (7)$$

For the equation of perfect-gas state, $(\partial\varepsilon/\partial p)_v = \varepsilon/p$ and derivative (7) becomes infinite. This behavior should be expected, since pressure p is eliminated from Eq. (6) in this case.

When simulating some shock-wave flows in solids, researchers used the so-called two-term equation of state [10]

$$\varepsilon = \frac{p - (\rho - \rho_*)c_*^2}{(\gamma - 1)\rho},$$

where ρ_* and c_* are the density and sound velocity of the substance in the normal state, respectively.

From Eq. (7), we have

$$\left(\frac{\partial p}{\partial v_{00}}\right)_v = \frac{(\gamma - 1)\rho p^2}{2(\rho - \rho_*)c_*^2} > 0$$

at $\rho > \rho_*$, i.e., for a substance compressed with respect to the normal state.

Another example is a Hugoniot adiabat with a linear relation between the shock-wave velocity D and the particle velocity u behind the wave:

$$D = a + bu, \quad a > 0, \quad b > 1.$$

A relation between p, v, and v_{00} for such a Hugoniot adiabat has the form (e.g., see monograph [11])

$$p = \frac{a^2(v_{00} - v)}{[v_{00} - b(v_{00} - v)]^2}. \qquad (8)$$

The field of application of Eq. (8) is restricted to the range $v_{00} \geq v > v_{00}(b - 1)/b$. By differentiating Eq. (8), one can make sure that $(\partial p/\partial v_{00})_v > 0$ in this range.

Thus, for both considered approximations of a Hugoniot adiabat in solids, the pressure behind a wave front increases with decreasing initial density ρ_{00} if the density ρ behind the wave does not change with ρ_{00}. The pressure is also expected to increase for a weak dependence of ρ on ρ_{00}, e.g., for one-dimensional shock waves when the function $\rho = \rho(r, \rho_{00})$ depends weakly on ρ_{00}. This situation is impossible for the equation of perfect-gas state, where the relation between ρ and ρ_{00} is unambiguously determined by a Hugoniot adiabat. We may assume that the weak dependence of ρ on ρ_{00} is also absent in the case of very strong shock waves, where the density ρ is close to its limiting value depending on ρ_{00}. In this case, the character of the ρ_{00} dependence of the shock-wave intensity is likely to be qualitatively identical to that typical of the equation of perfect-gas state.

For flows with converging shock waves, it is of interest to study not only the pressure at a converging-wave front but also the pressure appearing behind the front of the reflected shock wave. Therefore, as a characteristic of the converging-wave intensity, we introduce a pressure p_r that appears upon the interaction of a shock wave with a rigid immovable wall and that is the solution to the equation

$$\sqrt{(p_r - p)(v - v_r(p_r))} = u, \qquad (9)$$

where u is the particle velocity behind the wave front that is normal to this front and the $V_r(p_r)$ function is specified by the Hugoniot adiabat constructed from the state p, V:

$$\varepsilon(p_r, V_r) - \varepsilon(p, V) = (p + p_r)(V - V_r)/2.$$

The pressure p_r increases with both p and u. For example, in the acoustic approximation $V - V_r \approx (p_r - p)/\rho^2 c^2$ (where c is the velocity of sound), we have $p_r \approx p + \rho c u$ from condition (9).

In the strong shock wave approximation, we obtain $u = \sqrt{p(V_{00} - V)}$. Therefore, if $(\partial p/\partial V_{00})_V > 0$, then $(\partial u/\partial V_{00})_V > 0$. If we now assume that the $\rho = \rho(r, \rho_{00})$ function depends weakly on ρ_{00} for one-dimensional shock waves, then not only the pressure p but also the velocity u and pressure p_r increase with decreasing ρ_{00}.

ONE-DIMENSIONAL CONVERGING SHOCK WAVES

The problem is formulated in terms of the equations of nondissipative hydrodynamics of a compressible medium:

$$\frac{d\rho}{dt} + \rho \operatorname{div} u = 0, \quad \rho\frac{du}{dt} + \frac{\partial p}{\partial r} = 0,$$
$$\rho\frac{d(\varepsilon + u^2/2)}{dt} + \operatorname{div} pu = 0, \tag{10}$$

which are completed by the equations of state

$$p = p(\rho, T), \quad \varepsilon = \varepsilon(\rho, T). \tag{11}$$

Here, r is the radius; u is the velocity; T is the temperature; and $d/dt = \partial/\partial t + u\partial/\partial r$ is the Lagrangian derivative with respect to time,

$$\operatorname{div}(\cdot) = \frac{1}{r^\nu}\frac{\partial r^\nu(\cdot)}{\partial r},$$

where superscript $\nu = 1$ corresponds to cylindrical symmetry and superscript $\nu = 2$ corresponds to spherical symmetry. Equation (11) can be inverted to determine, e.g., $\varepsilon = \varepsilon(\rho, p)$ and $T = T(\rho, p)$ or $p = p(\rho, \varepsilon)$ and $T = T(\rho, \varepsilon)$. At the initial instant, we specify the density ρ_{00}, pressure p_0, and, in the case of a porous sample, the normal density in the absence of voids $\rho_0 > \rho_{00}$. The motion of a porous medium is assumed to begin when a bow shock wave passes through it. The boundary condition for the wave is a Hugoniot adiabat with parameters ρ_{00}, $p_0 = p(\rho_0, T_0)$, and $\varepsilon_0 = \varepsilon(\rho_0, T_0)$ ahead of its front.

We used two numerical methods, which are based on different approaches to the calculation of a converging shock wave. In the first method, a wave front is explicitly separated. We initially introduce a small calculation region with a small number of mesh points, and the values of the desired functions at these points are taken from the solution to the problem of discontinuity decay, which determines the parameters of a shock wave entering into a porous substance. Then, at each time step, we calculate the values of the desired functions in the upper time layer using a scheme providing second-order accuracy. From the pressure at the mesh point nearest to the wave front and the Hugoniot adiabat, we determine the wave velocity and, thus, a new front position. As the calculation region extends, new mesh points are added to it until a given number of points is achieved. When a mesh point is added, new points are distributed uniformly in r and the values of the desired functions are calculated at them with interpolation by fourth-order accuracy local splines based on the Hermitian polynomial. The wave front is separated explicitly until it reaches a rigid wall located at sufficiently small r. Then, the initial porous substance does not affect the solution to the problem, and the reflected wave is calculated using a standard through-calculation scheme for Eqs. (10) and (11). In the stage of explicit separation of a bow converging wave front, shock waves inside the calculation region are absent, which allows us to use the isentrope equation $d\varepsilon = pd\rho/\rho^2$ instead of the third equation in (10). Along with Eq. (11), the former equation can be integrated at a rather small ρ step for each point of a Lagrangian mesh.

The second method is based on calculation through a bow-wave front and on a simple model of pore kinetics (e.g., see [12]). We can also formally introduce a pore volume V_p in the formulation of the problem and define it at the initial instant as follows: $V_p = \rho_{00}^{-1} - \rho_0^{-1}$ ahead of the front and $V_p = 0$ behind the front. In calculations, the V_p function is assumed to be a new unknown function satisfying the equation

$$\frac{dV_p}{dt} = -AV_p(p - p_0). \tag{12}$$

In equation of state (11), the density ρ is replaced by the density of the continuous substance $\rho_s = (\rho^{-1} - V_p)^{-1}$. Equation (12) is written at each point of the difference mesh, and the constant A is chosen so that the V_p function decreases by a few orders of magnitude behind the bow shock wave spread over the mesh.

Test calculations performed using both methods described above agreed well with each other.

As a test problem for studying the properties of converging shock waves, we consider a ball or cylinder of radius $r_0 = 2$ mm that consists of a substance with a normal density ρ_0 and an initial porosity falling in the range $1 \le m_p \le 2$ and collide with a poreless ball or cylinder that can be made of another substance and move at a velocity u_0. The initial pressure in the pore substance and the flyer body is $p_0 = 0.1$ MPa. The quantities related to the wave front are assumed to be functions of the wave radius r and $\rho_{00} = \rho_0/m_p$.

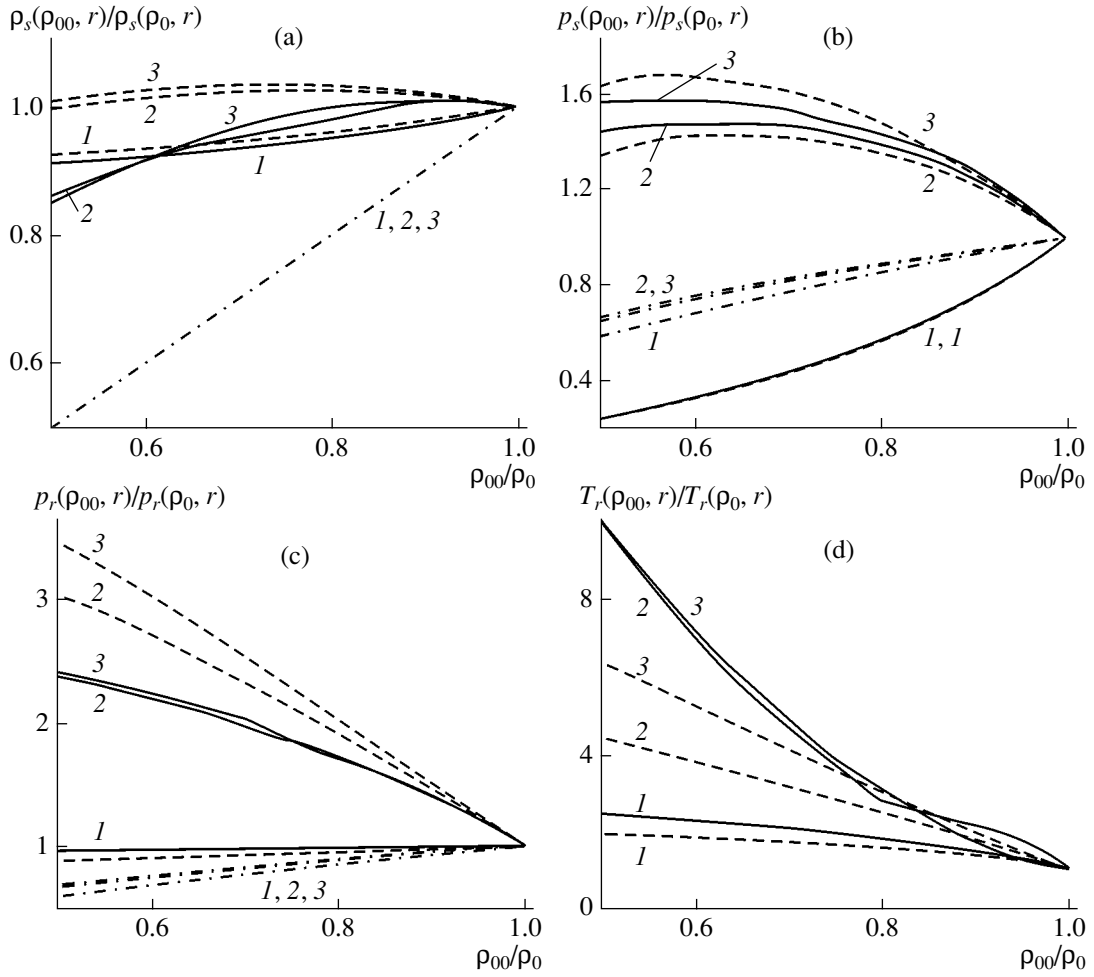


Fig. 2. Relative changes in (a) the density ρ_s and (b) the pressure p_s in a converging shock wave and relative changes in (c) the pressure p_r and (d) the temperature T_r in the wave reflected by a rigid wall at $r = (1)\ r_0$, $(2)\ r_0/5$, and $(3)\ r_0/10$ for (solid lines) aluminum, (dashed lines) graphite, and (dot-and-dash lines) perfect gas. Aluminum flyer body moving at $u_0 = 1$ km/s.

For the spherical case and an aluminum flyer ball moving at $u_0 = 1$ km/s and at $r = r_0$, $r_0/5$, and $r_0/10$, Fig. 2 shows the ratio $f(\rho_{00}, r)/f(\rho_0, r)$ as a function of the relative initial density $\rho_{00}/\rho_0 = m_p^{-1}$, where f is the density ρ_s and pressure p_s in the converging wave and the pressure p_r and temperature T_r in the wave reflected from a rigid wall. The problem was solved for a ball with the equations of state for aluminum, graphite, and perfect gas at $\gamma = 3$. In the latter case, the density ρ_0 coincided with the graphite density.

For a perfect gas, ρ_s is a linear r-independent function of ρ_{00}. The pressure p_s decreases with decreasing ρ_{00}, as follows from the properties of the corresponding self-similar solution and condition (5) at $r = r_0$. The pressure p_r in the reflected wave also decreases.

For the equation of condensed-matter state, the ρ_{00} dependence of the solution has a radically different character. For graphite, the density ρ_s is almost independent of ρ_{00}. For aluminum, this dependence is stronger, especially at low ρ_{00}. However, even at a range boundary $m_p = 2$, the density ρ_s decreases by only 15% as compared to the solution to the problem at $m_p = 1$. At $r = r_0$, i.e., at the instant of entering the shock wave into the porous substance, the pressure p_s decreases with decreasing ρ_{00}, as it follows from analysis of the problem of discontinuity decay. However, at low r, $p_s(\rho_{00}, r)$, in contrast, increases as ρ_{00} decreases from ρ_0 to $\sim 0.7\rho_0$.

The ρ_{00} dependences of the parameters of the reflected shock wave at low r are of particular interest. The pressure p_r in graphite increases almost linearly with decreasing ρ_{00}. For $\rho_{00} = 0.5\rho_0$, p_r increases by a factor of 3–3.5 as compared to the case of $\rho_{00} = \rho_0$. In aluminum, the increment in p_r is less significant: it increases by a factor of ~2.5. A striking increase in the temperature T_r at low r in aluminum is noteworthy: this temperature increases by about an order of magnitude at $\rho_{00} = 0.5\rho_0$ as compared to the case with $\rho_{00} = \rho_0$.

However, at $r = r_0$, this temperature increases by a factor of only 2.5.

In Fig. 3, these ρ_{00} dependences of the pressure p_r and temperature T_r for aluminum and graphite and a spherical wave are compared with such dependences for a cylindrical wave induced by a flyer cylinder moving at the same velocity. In the cylindrical wave, both parameters also increase with decreasing ρ_{00}. In both substances, the pressure p_r in the cylindrical wave increases weaker than in the spherical wave. The temperature T_r in aluminum increases approximately identically for the cylindrical and spherical waves, and that in graphite increases more strongly for the cylindrical wave.

The results shown in Fig. 4 answer the question of the processes that occur as the shock-wave intensity increases. As a porous material, we only considered aluminum, since graphite undergoes phase transformations accompanied by a large density jump at high temperatures and pressures (which can exert the main effect on the entire process). For the case of a spherical wave, we present calculation results for three flyer balls, namely, for an aluminum ball with $u_0 = 1$ km/s and for steel balls moving at $u_0 = 1$ and 2 km/s. For the first steel ball with $r = r_0/10$ and $\rho_{00} = \rho_0$, the pressure in the converging wave increases approximately by a factor of 1.5 as compared to the aluminum ball. For the steel ball with $u_0 = 2$ km/s, the pressure increases by another factor of 2.4. As is seen from Fig. 4, as the converging-wave intensity increases, the rate of increase of the pressure and temperature in the reflected wave decreases noticeably with decreasing ρ_{00}. In particular, for the steel ball with $u_0 = 2$ km/s, the pressure in the reflected wave does not increase with decreasing ρ_{00}.

Above, we have considered the flow characteristics related to a shock-wave front. Let us describe the characteristics of the entire compressed volume. Figure 5 shows a typical example of the time dependence of the radius of the flyer ball–compressed ball interface ($r_b(t)$) that corresponds to the average density $\rho_b(t) = m/v_b(t)$, where $v_b = 4\pi r_b^3/3$ is the ball volume and m is the ball weight

$$m = 4\pi\int_0^{r_b}\rho r^2 dr = \text{const},$$

and the total internal energy

$$E(t) = 4\pi\int_0^{r_b}\varepsilon\rho r^2 dr.$$

It is seen that the functions $\rho_b(t)$ and $E(t)$ have a maximum. As integrated characteristics, we chose $\rho_{b\,\max}$ and $\varepsilon_{b\,\max} = E_{\max}/m$. The latter quantity is the maximum average internal energy of a mass unit in time.

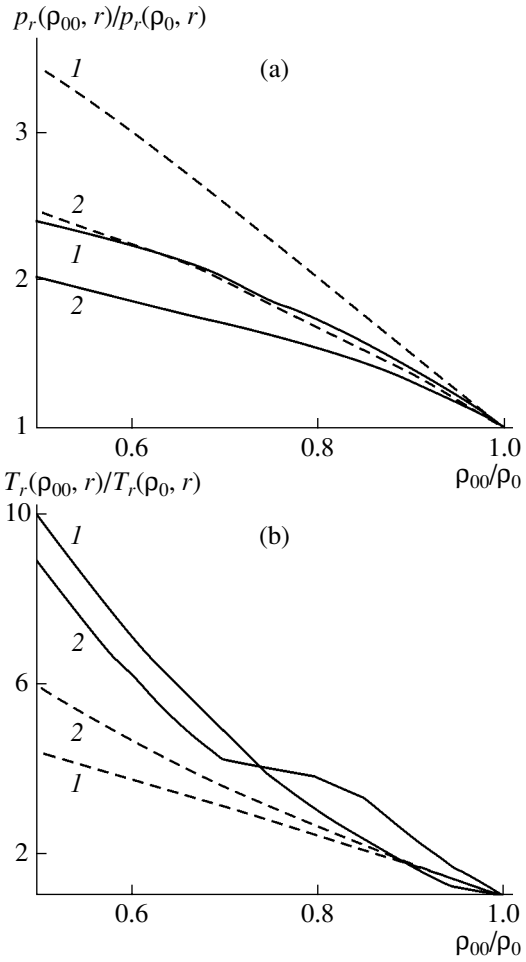


Fig. 3. Relative changes in (a) the pressure p_r and (b) the temperature T_r in the reflected wave at $r = r_0/10$ in (solid lines) aluminum and (dashed lines) graphite for a (*1*) spherical or (*2*) cylindrical wave.

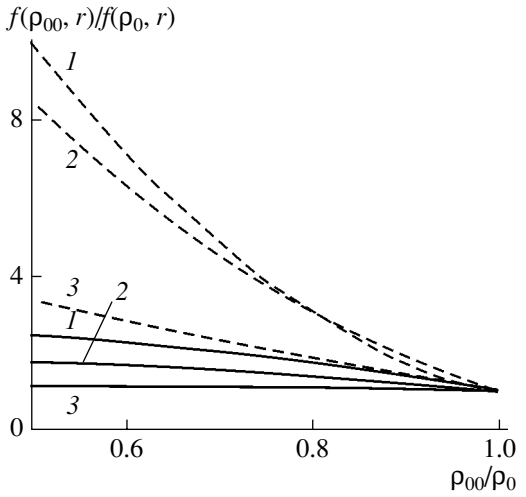


Fig. 4. Spherical shock waves of different intensities in aluminum: (*1*) aluminum flyer body, $u_0 = 1$ km/s, (*2*, *3*) steel flyer body, $u_0 = 1$ and 2 km/s, respectively. (Solid lines) $f = p_r$ and (dashed lines) $f = T_r$; $r = r_0/10$.

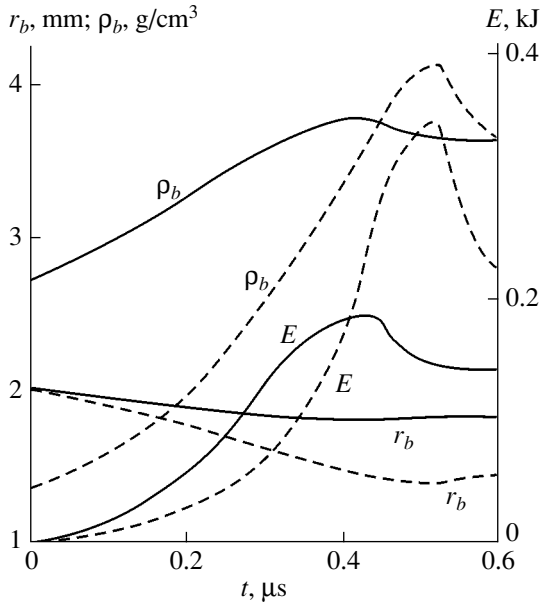


Fig. 5. Radius of an aluminum ball r_b, its average density ρ_b, and its total internal energy E as functions of time for (solid lines) $m_p = 1$ and (dashed lines) $m_p = 2$. An aluminum flyer body, $u_0 = 1$ km/s.

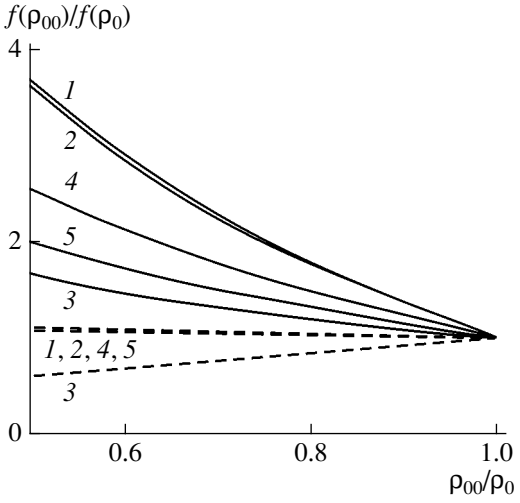


Fig. 6. Relative changes in the average quantities that are maximum in time: (solid lines) $f = \varepsilon_{b\,\max}$ and (dashed lines) $f = \rho_{b\,\max}$ in (*1*) graphite, (*2*) aluminum, and (*3*) perfect gas for an aluminum flyer body with $u_0 = 1$ km/s and in aluminum for a steel flyer body with $u_0 =$ (*4*) 1 and (*5*) 2 km/s.

Figure 6 shows relative changes in these characteristics with ρ_{00} in a perfect gas, graphite, and aluminum at different shock-wave intensities. For graphite and aluminum, $\rho_{b\,\max}$ very weakly depends on ρ_{00} in all cases. For a perfect gas, $\rho_{b\,\max}$ decreases with decreasing ρ_{00} similarly to the density at the converging-wave front. $\varepsilon_{b\,\max}$ increases most intensely in aluminum and graphite in the case of the aluminum flyer ball moving at $u_0 = 1$ km/s. As the converging-wave intensity increases, the rate of increase of $\varepsilon_{b\,\max}$ decreases substantially and approaches that in the perfect gas.

In the conclusion of this section, we propose a possible mechanism for the density behind the converging-wave front in a porous substance to be as high as that in the poreless substance. As is seen from Fig. 2, at $r_b = r_0$ (i.e., at the instant of entering a shock wave into a porous substance), the density behind the front ρ decreases with decreasing ρ_{00} and the pressure p_r is almost independent of ρ_{00}. The latter fact means that the substance velocity at $r_b = r_0$ increases with decreasing ρ_{00}, whereas the pressure p decreases. An increase in the velocity in the converging flow results in a more pronounced cumulative effect and, correspondingly, in a higher relative increase in the density as compared to the case of a crystalline initial density. As was shown by the calculations given above, this increase compensates for a pore-induced decrease in the density in many cases.

CONVERGING SHOCK WAVE IN A CONICAL SOLID TARGET

We now solve the following problem (Fig. 7). An aluminum flyer plate moving at a velocity of 2.5 km/s collides with a steel target. The target contains a cavity in the form of a truncated cone filled by graphite. The opening of the cone is 75°, the radius of the smaller base is 1 mm, and the cavity height is 7 mm. The flow in the flyer plate is considered within the framework of hydrodynamic equations. To describe the flow in the target and graphite, we applied both hydrodynamic equations and a more general model of a hypoelastic, ideally plastic medium. The initial graphite density ρ_{00} changed in the range $\rho_0 = 2.265$ g/cm^3 $\geq \rho_{00} \geq$ 1.1 g/cm^3. Since the purpose of our work was to study the effect of porosity on the pressure, we did not take into account a phase transition in graphite, which also affects the pressure.

Let us present equations that describe the axisymmetric flow of a hypoelastic, ideally plastic medium (e.g., see [13, 14]). In the cylindrical coordinates r, z, and ϑ, all functions are supposed to be independent of ϑ. A velocity vector $\mathbf{u}$ has two components, namely, u_r and u_z. Nonzero stress-tensor components are σ_{rr}, σ_{zz}, σ_{rz}, and $\sigma_{\vartheta\vartheta}$; for the sake of brevity, they will be referred to as σ_r, σ_z, τ, and σ_ϑ, respectively. Equations for the laws of conservation of mass, momentum, and energy have the form

$$\frac{d\rho}{dt} + \rho \mathrm{div}\mathbf{u} = 0, \quad \mathrm{div}\mathbf{u} = \frac{1}{r}\frac{\partial r u_r}{\partial r} + \frac{\partial u_z}{\partial z},$$

$$\rho\frac{du_r}{dt} = \frac{\partial\sigma_r}{\partial r} + \frac{\partial\tau}{\partial z} + \frac{\sigma_r - \sigma_\vartheta}{r},$$
$$\rho\frac{\partial u_z}{dt} = \frac{\partial\sigma_z}{\partial z} + \frac{1}{r}\frac{\partial r\tau}{\partial r}, \quad (13)$$
$$\rho\frac{d(\varepsilon + |\mathbf{u}|^2/2)}{dt} = \frac{1}{r}\frac{\partial r(\sigma_r u_r + \tau u_z)}{\partial r} + \frac{\partial(\tau u_r + \sigma_z u_z)}{\partial z},$$

where $d/dt = \partial/\partial t + u_r\partial/\partial r + u_z\partial/\partial z$ is the Lagrangian derivative with respect to time.

From the stress tensor, we separate a deviator part: $\sigma_r = -p + s_r$, $\sigma_z = -p + s_z$, and $\sigma_\vartheta = -p + s_\vartheta$, where p is the pressure and s_r, s_z, and s_ϑ are the diagonal elements of the stress-tensor deviator that satisfy the condition

$$s_r + s_z + s_\vartheta = 0. \quad (14)$$

Equations (13) are completed by equations of state (11) and the equations

$$\frac{ds_r}{dt} - \tau\psi = 2G\varphi_r - \frac{2}{3}G\mathrm{div}\mathbf{u} - \lambda s_r, \quad \psi = \frac{\partial u_r}{\partial z} - \frac{\partial u_z}{\partial r},$$
$$\frac{ds_z}{dt} + \tau\psi = 2G\varphi_z - \frac{2}{3}G\mathrm{div}\mathbf{u} - \lambda s_z,$$
$$\frac{ds_\vartheta}{dt} = 2G\varphi_\vartheta - \frac{2}{3}G\mathrm{div}\mathbf{u} - \lambda s_\vartheta, \quad (15)$$
$$\frac{d\tau}{dt} + \frac{s_r - s_z}{2}\psi = G\varphi_{rz} - \lambda\tau,$$

which describe the evolution of the deviator components. Here, the left-hand side contains the Jaumann derivative of the tensor components with respect to time; G is the given shear modulus of the medium; and $\varphi_r = \partial u_r/\partial r$, $\varphi_z = \partial u_z/\partial z$, $\varphi_\vartheta = u_r/r$, and $\varphi_{rz} = \partial u_r/\partial z + \partial u_z/\partial r$ are the strain rate tensor components.

To determine s_ϑ, we can also use Eq. (14), which follows from Eq. (15) as a consequence of the equality $\varphi_r + \varphi_z + \varphi_\vartheta = \mathrm{div}\mathbf{u}$. The function λ is chosen so that the Mises condition for plastic flow

$$Q = s_r^2 + s_z^2 + s_\vartheta^2 + 2\tau^2 \le \frac{2}{3}Y^2, \quad (16)$$

where Y is the given yield strength of the medium, to be satisfied.

Using Eq. (15), we can obtain

$$\frac{d}{dt}\frac{Q}{2} = W - \lambda Q,$$
$$W = 2G(\varphi_r s_r + \varphi_z s_z + \varphi_\vartheta s_\vartheta + \varphi_{rz}\tau). \quad (17)$$

Therefore, assuming

$$\lambda = 0 \text{ at } Q < 2Y^2/3,$$
$$\lambda = 3W/2Y^2 \text{ at } Q = 2Y^2/3, \quad (18)$$

we find that $dQ/dt = 0$ from Eq. (17) at $Q = 2Y^2/3$.

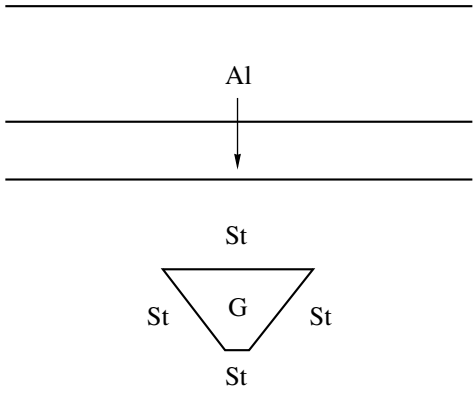


Fig. 7. Schematic diagram for the shock loading of a steel target (St) having a conical cavity filled by graphite (G) via an aluminum flyer plate (Al).

We omit the description of the numerical method used and dwell on its certain specific features. If the yield strength is $Y = 0$, the stress-tensor deviator becomes zero and the equations written above transform into hydrodynamic equations. The numerical method has a similar property. At $Y = 0$, it transforms into a method for hydrodynamic equations providing second-order accuracy for mobile regular meshes where interfaces are explicitly separated in the form of mesh lines. This method was used in our previous studies [1].

Following [13], we approximate Eqs. (15) without using the last terms, which are responsible for flow plasticity. Instead, when Mises condition (16) is violated, we multiply the deviator by a coefficient that renders condition (16) valid at each time step. As was shown by Grigoryan in his remark to [13] (also see [15]), this algorithm gives a difference scheme that can approximate Eqs. (15) and (18).

We realized two variants of calculation through the front of a bow shock wave passing across a porous substance. The first method is based on a simple model of pore kinetics and was described in the previous section. The second method is based on the model [1] of a mixture of two phases, namely, graphite and a perfect gas, whose initial density ρ_g is determined from the initial pressure and temperature. The gas weight concentration in the mixture β is constant and determined from the condition $\rho_{00}^{-1} = \beta\rho_g^{-1} + (1 - \beta)\rho_0^{-1}$. Very good agreement between the results of test calculations carried out by both methods indicates that through calculation does not introduce noticeable distortions in the solution to the problem.

The main calculations were performed using a 25×50 mesh in the graphite, where the first number is the number of intervals along the base of cone and the second number is the number of intervals along the axis of symmetry. For test calculations, we also used a mesh with a double scale in both directions.

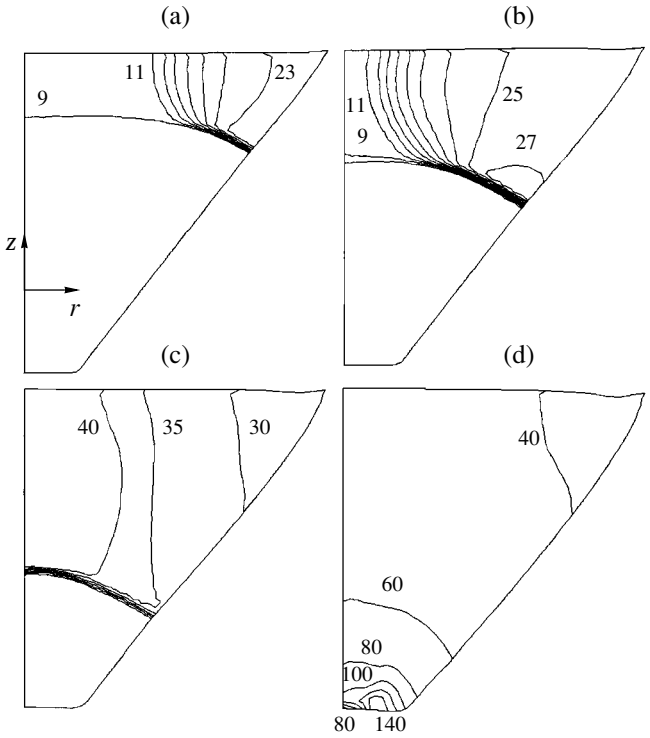


Fig. 8. Isobars in a conical target of a porous graphite at t = (a) 2, (b) 2.25, (c) 2.5, and (d) 2.75 μs. Numerals at curves give pressure in GPa.

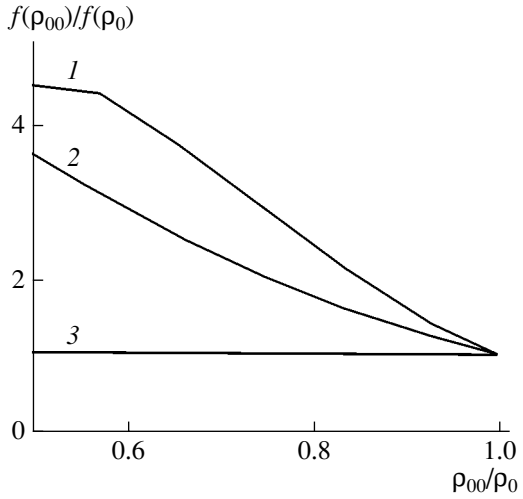


Fig. 9. Characteristics of a flow in the conical target: f = (*1*) p_{max}, (*2*) ε_{max}, and (*3*) ρ_{max}.

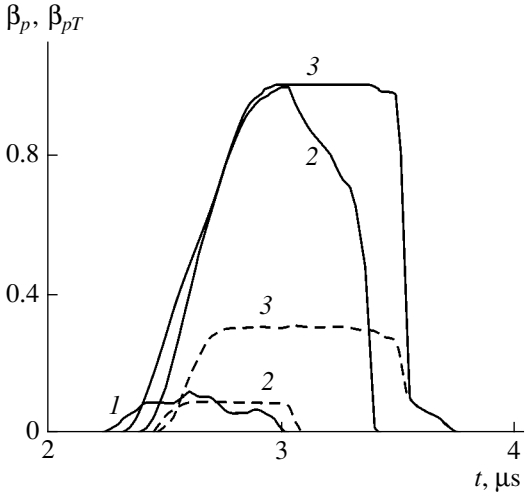


Fig. 10. Weight fractions of (solid lines) high-pressure graphite (β_p) and (dashed lines) high-temperature, high-pressure graphite (β_{pT}) as functions of time at ρ_{00} = (*1*) 2.265, (*2*) 1.9, and (*3*) 1.7 g/cm^3.

First, we present the calculation results for the purely hydrodynamic problem with $Y = 0$. Figure 8 shows isobars in the initial porous graphite with ρ_{00} = 1.7 g/cm^3 at four sequential instants. In Fig. 8a, which corresponds to the first instant, one can see a bow plane shock wave near the axis of symmetry and a lateral compression pulse coming from the cone base angle. The exact calculation of the parameters of the bow plane shock wave based on conservation laws demonstrates that the pressure behind the front is ≈9 GPa. The numerical method gives approximately the same value, which indicates its good accuracy. Note that, in the case of continuous graphite, the pressure behind the bow-wave front increases to ≈17 GPa.

Then, as is seen from Figs. 8b and 8c, a near-spherical converging shock wave appears. A mechanism that limits cumulation in this wave is the deformation of the cone walls. Figure 8d corresponds to the instant after the wave has collapsed.

As an analog of the pressure p_r in the one-dimensional reflected wave, we take the pressure p_{max} that is maximum in both time and space. As in the case of one-dimensional flows, we can introduce the average density ρ_{max} and the average internal energy ε_{max} that are maximum in time. Relative changes in these quantities with increasing ρ_{00} are shown in Fig. 9. As in the case of one-dimensional converging waves, a decrease in ρ_{00} leads to a substantial increase in p_{max} and ε_{max} at almost constant ρ_{max}.

The increase in the pressure and temperature of graphite with decreasing ρ_{00} is not local. To demonstrate this behavior, we take a certain, rather high pressure ($p_* = 40$ GPa) and temperature ($T_* = 1500$ K) and introduce the functions

$$\beta_p(t) = \int\limits_{p > p_*} \rho dV/m, \quad \beta_{pT}(t) = \int\limits_{p > p_*,\, T > T_*} \rho dV/m,$$

where m is the total graphite mass.

The function $\beta_p(t)$ is taken to be the weight fraction of high-pressure graphite, and the function $\beta_{pT}(t)$ is

taken to be the weight fraction of high-temperature, high-pressure graphite. The latter function approximately corresponds to the weight fraction of graphite compressed to a high pressure in a converging shock wave, since the temperature increases only insignificantly upon repeated compression in the reflected wave.

Figure 10 shows the $\beta_p(t)$ and $\beta_{pT}(t)$ functions for continuous and pore graphite with ρ_{00} = 2.265, 1.9, and 1.7 g/cm^3. In the case of the continuous sample, ≈10% of graphite has a pressure $p > p_*$ for ~0.5 μs and graphite with $p > p_*$ and $T > T_*$ is virtually absent (less than 0.1%). For the porous sample with ρ_{00} = 1.9 g/cm^3, the fraction of graphite with $p > p_*$ increases significantly and approaches unity in a small time interval. Moreover, ~10% graphite has a pressure $p > p_*$ and a temperature $T > T_*$ for ~0.5 μs. As ρ_{00} decreases to 1.7 g/cm^3, the entire graphite has a pressure $p > p_*$ for ~0.5 μs. The fraction of graphite with $p > p_*$ and $T > T_*$ increases approximately to 30%, and the lifetime of this graphite increases to ~1 μs.

In conclusion, we present data calculated by the model of a hypoelastic, ideally plastic medium. The yield strength Y of the same material can change significantly in different experiments. For example, for Kh18N10T steel, the authors of [16] give a value Y = 0.26 GPa, which was obtained from certain shock-wave experiments. However, Makarov [17] asserts that the shear stresses in strong shock waves in metals are compared with the theoretical shear strength ($k = Y/\sqrt{3}$), which is taken to be $G/30$ [18]. A typical value for steels is G = 80 GPa [18, 19]. As a result, we obtain $Y \approx$ 4.6 GPa, which is likely to be the upper estimate of possible Y values in steels.

For the case of ρ_{00} = 1.7 g/cm^3, Fig. 11 shows the graphite-cavity boundary at t = 5 μs and the $\beta_{pT}(t)$ function for yield strengths Y = 0, 0.26, and 4.6 GPa in a steel. The calculation results for Y = 0 and 0.26 GPa are seen to virtually coincide with each other. For Y = 4.6 GPa, shear stresses play a significant role: the deformation of the graphite cavity changes, the weight fraction of β_{pT} decreases from ~30 to ~25%, and the lifetime of the high-temperature, high-pressure graphite decreases to ~0.5 μs. This decrease is related to the weakening of the lateral compression pulse that forms a converging shock wave. Nevertheless, when passing from the continuous to porous graphite, the effect remains very strong, since β_{pT} = 0 for the continuous graphite.

The elastic modulus of graphite (G ~ 0.1 GPa) is well below that of a steel. Correspondingly, the yield strength Y is lower. Our calculations at G = 0.1 GPa and Y = 0.05 GPa in graphite show that shear stresses in graphite exert only a weak effect on the solution to the problem.

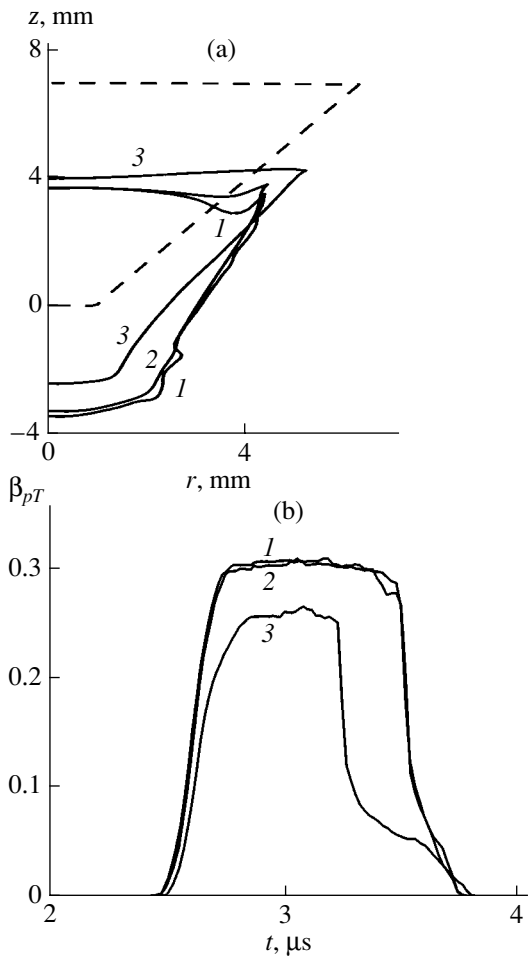


Fig. 11. Comparison of solutions for Y = (*1*) 0, (*2*) 0.26, and (*3*) 4.6 GPa in a steel; G = 80 GPa. (a) The graphite-cavity boundary at (dashed line) t = 0 and (solid lines) t = 5 μs and (b) the weight fraction of high-temperature, high-pressure graphite as a function of time.

CONCLUSIONS

The solutions to the hydrodynamic equations of compressed media that describe one-dimensional converging shock waves of a moderate intensity in solids have the following properties. As the initial density of a solid decreases, the pressure and temperature of the reflected wave decrease and the average, maximum-in-time internal energy of the whole mass of the compressed substance increases substantially. At a given radius of a converging-wave front, both the density at the front and the average, maximum-in-time density of the whole mass of the compressed substance only weakly depend on the initial porosity.

For the equation of perfect-gas state, the dependences of the flow characteristics on the initial density have a qualitatively different character, which is related

to the degeneration of the Hugoniot adiabat of the perfect gas in the strong wave approximation, where the density behind the wave front is unambiguously controlled by the initial density.

At a high intensity of a converging shock wave, the pressure behind the reflected-wave front almost ceases to depend on the initial sample porosity, and the temperature and the average internal energy in the reflected wave increase with decreasing initial density to a lesser degree as compared to a low intensity of the converging wave.

As the initial density decreases, the pressure in the reflected wave also increases in the two-dimensional problem of the shock compression of graphite in a steel target with a conical cavity, where the converging shock wave is not spherical and its cumulation is restricted to the deformation of the target walls. As the initial density decreases, both the weight fraction of graphite compressed to a high pressure and the time of maintenance of a high pressure in the target increase. This effect is retained when possible shear stresses in the target are taken into account within the framework of the model of a hypoelastic, ideally plastic medium.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project nos. 04-01-00051, 04-02-16471), the Russian Academy of Sciences (project no. 3 OMN), and the Council on Grants from the President of the Russian Federation (project nos. MK-2103.2003.08, MK-2572.2004.8).

We also thank the Russian Foundation for Support of Science.

REFERENCES

1. I. V. Lomonosov, V. E. Fortov, A. A. Frolova, *et al.*, Zh. Tekh. Fiz. **73** (6), 66 (2003) [Tech. Phys. **48**, 727 (2003)].
2. A. A. Charakhch'yan, I. V. Lomonosov, V. V. Milyavskiĭ, *et al.*, Pis'ma Zh. Tekh. Fiz. **30** (1), 72 (2004) [Tech. Phys. Lett. **30**, 33 (2004)].
3. Ya. B. Zel'dovich and Yu. P. Raizer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena*, 2nd ed. (Nauka, Moscow, 1966; Academic, New York, 1967).
4. B. L. Rozhdestvenskii and N. N. Yanenko, *Systems of Quasilinear Equations and Their Applications to Gas Dynamics* (Nauka, Moscow, 1978; Am. Math. Soc., Providence, 1983).
5. K. V. Khishchenko, V. E. Fortov, I. V. Lomonosov, *et al.*, in *Proceedings of the Conference of the American Physical Society Topical Group on Shock Compression of Condensed Matter, Atlanta, 2001*, Ed. by M. D. Furnish, N. N. Thadhani, and Y. Horie (AIP, New York, 2002), pp. 759–762.
6. I. V. Lomonosov, V. E. Fortov, A. A. Frolova, *et al.*, Teplofiz. Vys. Temp. **41**, 515 (2003).
7. A. V. Bushman, V. E. Fortov, G. I. Kanel', and A. L. Ni, *Intense Dynamic Loading of Condensed Matter* (Inst. Khim. Fiz. AN SSSR, Chernogolovka, 1988; Taylor and Francis, Washington, 1993).
8. A. S. Kolomeets, Dokl. Akad. Nauk SSSR **109**, 49 (1956).
9. E. I. Andriankin and V. P. Koryavov, Dokl. Akad. Nauk SSSR **128**, 257 (1959) [Sov. Phys. Dokl. **4**, 966 (1959)].
10. S. K. Godunov, A. V. Zabrodin, M. Ya. Ivanov, G. P. Prokopov, and A. N. Kraĭko, *Numerical Solution of Multidimensional Problems in Gas Dynamics* (Nauka, Moscow, 1976) [in Russian].
11. A. V. Bushman, I. V. Lomonosov, and V. E. Fortov, *Equations of State for Metals at High Energy Densities* (IKhFCh RAN, Chernogolovka, 1992) [in Russian].
12. A. Z. Zhuk, A. V. Ivanov, and G. I. Kanel', Teplofiz. Vys. Temp. **29**, 486 (1991).
13. M. L. Wilkins, in *Methods of Computational Physics*, Ed. by B. Alder, S. Fernbach, and M. Rotenberg (Academic, New York, 1964; Mir, Moscow, 1967), pp. 212–263.
14. A. G. Kulikovskii, N. V. Pogorelov, and A. Yu. Semenov, *Mathematical Aspects of Numerical Solution of Hyperbolic Systems* (Fizmatlit, Moscow, 2001; Chapman and Hall, Boca Raton, 2001).
15. V. V. Kukudzhanov, Izv. Ross. Akad. Nauk, Mekh. Tverd. Tela, No. 1, 98 (2004).
16. G. I. Kanel', S. V. Razorenov, A. V. Utkin, and V. E. Fortov, *Shock-Wave Phenomena in Condensed Media* (Yanus-K, Moscow, 1996) [in Russian].
17. P. V. Makarov, in *Shock Waves and Extreme States of Media* (Nauka, Moscow, 2000), pp. 219–254 [in Russian].
18. C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1976; Nauka, Moscow, 1978).
19. G. P. Sleptsova, in *Physics Encyclopedia: A Dictionary* (Sov. Éntsiklopediya, Moscow, 1963), Vol. 3, pp. 273–274 [in Russian].

Translated by K. Shakhlevich

Technical Physics, Vol. 50, No. 8, 2005, pp. 987–992. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 26–31.

ATOMS, SPECTRA, RADIATION

Effect of Electrons of the Multiply Charged Ion Core on Single-Electron Capture

V. V. Afrosimov, A. A. Basalaev, and M. N. Panov

Ioffe Physicotechnical Institute, Russian Academy of Sciences, St. Petersburg, 194021 Russia

e-mail: a.basalaev@mail.ioffe.ru

Received December 7, 2004

Abstract—The population of various electronic states of particles that arise during the capture of a single electron in hydrogen and helium atoms, as well as hydrogen molecules, by Ar^{3+} and Ne^{3+} ions with an energy of several kiloelectronvolts was studied by collision spectroscopy, viz., precision analysis of kinetic energy variation for ions formed as a result of interaction between ions and atoms. It is shown that single-electron capture in many cases is a multielectron process accompanied by the rearrangement of a multiply charged ion core. It is found that the triply charged Ne^{3+} ions formed as a result of ionization of Ne atoms by electron impacts are formed mainly in metastable states. The population of excited states of particles during their multiple ionization should be taken into account in determining the characteristics of various particles by the appearance potential method. Collision spectroscopy can be used for analyzing the metastable ion impurities in ionic beams. © *2005 Pleiades Publishing, Inc.*

INTRODUCTION

The interaction of slow multiply charged ions with atoms has been studied intensely since 1970s. This research was stimulated primarily by the need for information of the cross section of electron capture by impurity ions in the plasma of controlled nuclear fusion devices for analyzing their effect on the energy balance and the charge composition of the plasma. The interaction of multiply charged ions with atomic hydrogen is of special practical importance. Analysis of single-electron systems revealed that single-electron capture in slow ion–atom collisions is correctly described by a two-level quasi-molecular model, in which the electron capture is determined by the interaction of particles in the nuclear spacing range corresponding to pseudocrossing of the energy levels of the initial and final states of a quasi-molecule. It was also shown that transition probabilities can be calculated to a high degree of accuracy by using the Landau–Zeener model [1, 2] in view of the exothermal nature of the process involving multiply charged ions and occurring at a large nuclear spacing. The capture of a single electron from an atom by an ion having several electrons may occur as a two-electron process known as transfer excitation, in which the general regularities of the process inherent in a one-electron system are preserved [3]. Two types of such two-electron processes are obviously possible in this case. Electron capture can be accompanied by either the excitation of the core of a multiply charged ion [3–7], or by electron excitation of a target particle [5, 8].

Theoretical analysis of electron correlations during the interaction of multiply charged ions with atoms are usually associated with analysis of the mutual effect of electrons in the target atom [9]. In most cases, it was assumed that the electrons in the core of a multiply charged ion do not participate in the process and only determine, together with the electrons of the target, the structure of energy levels of the quasi-molecular system formed as a result of slow ion–atom collisions. However, experiment shows that single-electron capture may occur as a multielectron process accompanied by excitation of the electrons in the core of a multiply charged ion, and the contribution of this transfer excitation to the capture cross section of a single electron may be quite large even for a high multiplicity of the impinging ion. A significant role of transfer excitation was observed in [3] in the interaction of lithium-like ions $Ne^{7+}(1s^22s)$ and $F^{6+}(1s^22s)$ with He and Ne atoms.

This study is devoted to analysis of the role of multielectron channels in single-electron capture by Ar^{3+} and Ne^{3+} ions from various targets on the basis of recording of the electronic states of collision partners. The results were also used for analyzing the initial electronic states of impinging A^{z+} ions.

1. EXPERIMENTAL TECHNIQUE

We studied the electronic states of the ions formed during single-electron capture by using the method of collision spectroscopy based on precision analysis of the kinetic energy of a multiply charged ion after the collision. The experimental technique is described in [8, 10] and is based on the fact that the change in the kinetic energies (ΔE) of interacting particles as a result of inelastic single-electron capture

$$A^{z+}(L) + B \longrightarrow A^{(z-1)+}(L', n1) + B^{+} + Q, \qquad (1)$$

1063-7842/05/5008-0987$26.00 © 2005 Pleiades Publishing, Inc.

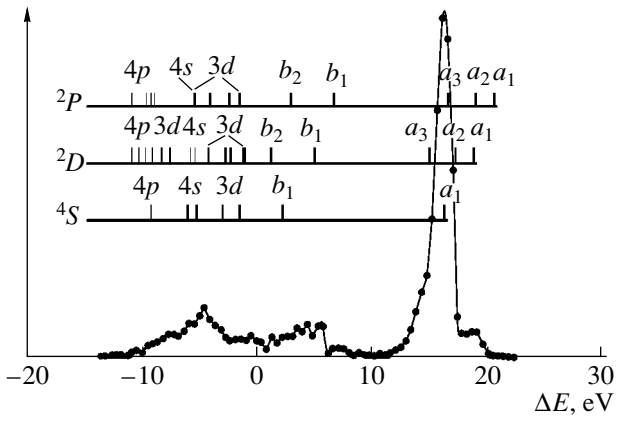


Fig. 1. Spectrum of kinetic energy variation ΔE for ions in the case of single-electron capture by Ar^{3+} ion from He atoms. Horizontal lines marked by ^4S, ^2D, and ^2P denote the Q scales for three possible initial states of the impinging ion $Ar^{3+}(3s^23p^3)$. The calculated positions of peaks for various finite electron states of collision partners are marked by vertical segments.

is determined by the difference Q in their potential energies before and after the collision. Here, $A^{z+}(L)$ is a multiply charged ion with a charge z, whose core is in the initial electron state L; $A^{\{z-1\}+}(L', n1)$ is a multiply charged ions that has captured an electron in a state with quantum numbers n, l, whose core is in the quantum state L'; B and B^+ are the target atom in the initial (ground) state and the ion that has lost an electron; and Q is the change in the potential energy of the system of colliding particles, which is calculated from the data on the energy of atomic levels [11, 12].

For small scattering angles, the kinetic energy transferred to a target particle is negligibly small; consequently, the measured change ΔE in the kinetic energy of an impinging particle as a result of an inelastic collision virtually coincides with the change Q in the potential energy of the system. This makes it possible to use quantity ΔE to identify excited states of particles after collisions from the data on the energy of their levels. For appreciable scattering angles, the kinetic energy of the recoil ion can also be calculated and taken into account.

In the present work, we used a source of multiply charged ions with an electron impact; the relatively low intensity of the electron beam and the low density of the gas in the source ensured the formation of ions in single collisions. The energy of ions crossing the collision chamber was analyzed with the help of an electrostatic analyzer of the "plane mirror" type. The possibility of separating the peaks in the experimentally measured kinetic energy spectrum of ions is ensured by the instrument function of the energy analyzer and the monochromaticity of the primary beam. Here, the total half-amplitude width of the instrument function of the experimental setup was $0.4z$ eV for a collision energy of z keV. According to the results of multiple measurements, the error in determining the position of the peaks in the energy spectra of $A^{(z-1)+}$ ions was smaller than 1 eV.

The gas target was produced using a chamber with differential evacuation. To obtain a target of atomic hydrogen, a collision chamber made of tungsten was heated by an electron beam to 2500 K. At this temperature, the degree of dissociation of molecular hydrogen in the collision chamber was 78%.

2. EXPERIMENTAL RESULTS

a. Interaction of Ar^{3+} ions with He atoms. Figure 1 shows the spectruum of kinetic energies of Ar^{2+} ions formed as a result of single-electron capture by Ar^{3+} ions with an initial kinetic energy of 5.25 keV from He atoms. The highest peak in the spectrum corresponds to ions that have increased their kinetic energy by ΔE = 16 eV as a result of electron capture. The Ar^{2+} ions with such a kinetic energy correspond to the process

$$Ar^{3+}(3s^23p^{3\,4}S) + He(1s^2) \longrightarrow Ar^{2+}(3s^23p^{4\,3}P) + He^+(1s) + 16.16\ \text{eV}, \quad (2)$$

in which an electron is captured by ions in the ground state $Ar^{3+}(3s^23p^3S)$ to form the ground state of the $Ar^{2+}(3s^2\,3p^{4\,3}P)$ ion. The calculated position of the peak associated with a change Q = 16.16 eV in the potential energy of collision partners is denoted in Fig. 1 by a_1 on the scale of the ground state ^4S.

Primary-beam ions $Ar^{3+}(3s^2\,3p^3)$ can also be in two metastable electronic states ^2D and ^2P with excitation energies E_{MS} = 2.62 eV and E_{MS} = 4.3 eV, respectively. When a metastable ion captures an electron to the $3p$ subshell, the latter is rearranged and the potential energy of the metastable state liberated in the process of capture leads to an additional (as compared to process (2)) increase in the kinetic energy of the ion that has captured an electron. If an electron is captured to the $3p$ subshell by an ion in the ground quartet state ^4S, the Ar^{2+} ion formed in this case can be only in the triplet ground state ^3P, while the ions formed from the metastable ions $Ar^{3+}(^2D)$ and $Ar^{3+}(^2P)$ can be in three electron states: ^3P, ^1P, and ^1S. The calculated positions of the peaks corresponding to these final states are marked in Fig. 1 by a_1, a_2, and a_3 on the scales corresponding to two metastable initial states of Ar^{3+} ions. It can be seen in Fig. 1 that a small peak with energy $\Delta E \approx$ 18.5 eV can be explained only by the presence of an impurity of metastable ions in states ^2D and/or ^2P in the primary beam.

The spectral region with ΔE approximately from –4 to –11 eV is associated with single-electron endothermic processes, in which an electron is captured to the excited state of the $Ar^{2+}(3s^23p^3\,n1)$ ion ($n1 = 3d, 4s, 4p$) without a change in the initial state of the core of the multiply charged ion. The contribution of such pro-

cesses amounts to approximately 15% of the total capture cross section. The spectral region with $\Delta E \approx 1$–7 eV corresponds to ions formed in the two-electron process

$$\mathrm{Ar}^{3+}(3s^2 3p^{2}{}^{2s+1}L) + \mathrm{He}(1s^2) \longrightarrow \mathrm{Ar}^{2+}(3s3p^{5}{}^{3}P, {}^{1}P) + \mathrm{He}^{+}(1s), \quad (3)$$

which occurs with excitation of the core of the multiply charged ions. As a result, the 3*s* electron passes to the 3*p* level. The calculated positions of the peaks for this process are denoted in Fig. 1 by letters b_1 for the final state $\mathrm{Ar}^{2+}(3s3p^{5}\,{}^{3}P)$ and b_2 for $\mathrm{Ar}^{2+}(3s3p^{5}\,{}^{1}P)$.

An analysis of the spectrum shown in Fig. 1 leads to the conclusion that the major part of Ar^{3+} ions of the primary beam are in the ground state. Some of the channels of single-electron capture (about 10% of the total cross section) are in fact two-electron processes, which are accompanied by excitation of one of the 3*s* electrons of the core to a higher 3*p* level of the doubly charged ion formed in the process. The excitation of a 3*s* electron of the core of the Ar^{3+} ion during the capture of an electron from the CF_4 molecule was also observed in [13]. The electron capture by Ar^{3+} ions, which is accompanied by excitation of He^+ target ions formed in the process ($n = 2$), has a much smaller cross section, since statistically significant peaks in the ΔE spectra of the Ar^{2+} ions, corresponding to $\Delta E \approx -25$ eV, were not detected.

b. Interaction of Ne^{3+} ions with He atoms. The energy spectra of fast Ne^{2+} ions formed as a result of capture of a single electron by Ne^{3+} ions with initial energies $E = 2.7$ and 5.25 keV from helium atoms are shown in Fig. 2. Here and below in the figures, the spectra for various collision energies are normalized to the maximum of the main peak. Since the outer electron subshell of impinging Ne^{3+} ions consists of three *p* electrons as in the case of Ar^{3+}, the Ne^{3+} ions formed in the primary beam can also be in three electron states: ^4S, ^2D ($E_{MS} = 5.08$ eV), and ^2P ($E_{MS} = 7.7$ eV), and all final electron states of the reaction product are analogous to the states of Ar^{2+}.

In experiments, ions that have increased their kinetic energy by ΔE ranging from 0 to 25 eV are detected. This means that the electron capture by a $Ne^{3+}(^4S)$ ion in the ground state to the ground state of the $Ne^{2+}(^3P)$ ion is not observed to within the sensitivity of the experiment for the collision energies studied here (the process

$$\mathrm{Ne}^{3+}(2s^2 2p^{3}{}^{4}S) + \mathrm{He}(1s^2) \longrightarrow \mathrm{Ne}^{2+}(2s^2 2p^{4}{}^{3}P) + \mathrm{He}^{+}(1s) + 38.88\ \mathrm{eV}, \quad (4)$$

which is an analog of process (2) for the impinging $Ar^{3+}(^4S)$ ion in Fig. 1).

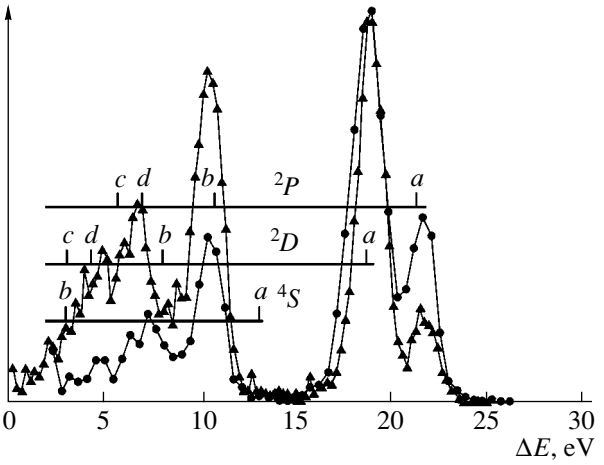


Fig. 2. Spectrum of kinetic energy variation ΔE for ions in the case of single-electron capture by Ne^{3+} ion from He atoms. Initial energies *E* of ions are 2.7 (▲) and 5.25 keV (●). Notation is the same as in Fig. 1 (notation of processes is given in the table).

The processes of formation of Ne^{2+} ions in states with a configuration $2s^2 2p^{4}{}^{2s+1}L$ from impinging ions in the metastable states $Ne^{3+}(^2D)$ and $Ne^{3+}(^2P)$ are not observed either, since the corresponding peaks should have occupied the region of $\Delta E = 37$–47 eV. The low probability of the processes listed above can be explained by their large resonance defects [11, 12]. As a result, the quasi-crossings of the energy levels of the initial and final states of the system of colliding particles $\{\mathrm{NeHe}\}^{3+}$ corresponding to these electronic transitions must lie at a small nuclear spacing. This in turn leads to considerable splitting of energy levels, which makes electron transitions very unlikely in the range of low collision rates.

When an electron is captured by Ne^{3+} ions to excited levels with $n > 3$, all channels of the process are endothermic with $\Delta E < 0$ and improbable according to the experimental results, since no statistically significant peaks are discovered in this spectral region. Thus, the observed electron capture accompanied by a change in the kinetic energy of impinging ions for system $(\mathrm{NeHe})^{3+}$ in the range of $\Delta E = 0$–25 eV must correspond to channels of electron capture to the 2*p* level, for which the final electron state of the projectile ion or the target is excited. The list of such channels is given in the table.

For ions in the ground state ($2s^2 2p^{3}{}^{4}S$), in the range of kinetic energy variation ΔE where the capture of a single electron is observed, only one capture channel accompanied by simultaneous electron transition from level 2*s* to level 2*p* and liberation of energy $Q = 13.55$ eV can exist (channel a–^4S in the table). However, a peak corresponding to the realization of this channel is not observed. The clearly manifested doublet in Fig. 2 denoted by a–^2P and a–2^D corresponds to elec-

Possible channels of single electron capture, corresponding to experimentally observed variation of the kinetic energy of fast particles

Reaction participants	Reaction products	Q, eV	Notation of the position of Q in Fig. 2	ΔE, eV
$Ne^{3+}(^2P) + He(1s^2)$	$Ne^{2+}(2s2p^5{}^3P) + He^+(1s)$	21.25	a–^2P	21.3 ± 0.5
$Ne^{3+}(^2D) + He(1s^2)$		18.63	a–^2D	18.7 ± 0.5
$Ne^{3+}(^4S) + He(1s^2)$		13.55	a–^4S	No
$Ne^{3+}(^2P) + He(1s^2)$	$Ne^{2+}(2s2p^5{}^1P) + He^+(1s)$	10.69	b–^2P	10.6 ± 0.5
$Ne^{3+}(^2D) + He(1s^2)$		8.07	b–^2D	8.07 ± 0.5
$Ne^{3+}(^2P) + He(1s^2)$	$Ne^{2+}([^4S]3s^3S) + He^+(1s)$	6.97	d–^2P	6.8 ± 0.5
$Ne^{3+}(^2D) + He(1s^2)$		4.35	d–^2D	4.8 ± 0.5
$Ne^{3+}(^4S) + He(1s^2)$		–0.73	d–^4S	No
$Ne^{3+}(^2P) + He(1s^2)$	$Ne^{2+}(2s^2p^4{}^3P) + He^+(n = 2)$	5.78	c–^2P	5
$Ne^{3+}(^2D) + He(1s^2)$		3.07	c–^2D	3
$Ne^{3+}(^4S) + He(1s^2)$		–1.92	c–^4S	No

tron capture by metastable ions ^2P and ^2D, accompanied by the formation of Ne^{2+} ions in the triplet electron state $2s2p^5{}^3P$. The next (in magnitude) contribution to the single-electron capture cross section corresponds to the channel of electron capture by metastable ions $Ne^{3+}(^2P)$ with the formation of the $Ne^{2+}(2s2p^5{}^1P)$ ion in the singlet state (b–^2P).

The range $1 < \Delta E < 8$ eV of energy spectrum corresponds to capture occurring as a complex multielectron process (third group of processes in the table denoted by d–^{2s+1}L in Fig. 2). First, the d–^2P peak at ΔE = 6.8 eV, which corresponds to the formation of $Ne^{2+}([^4S]3s^3S)$ ions from the $Ne^{3+}(^2P)$ state, can be clearly seen in the spectrum. Second, the capture channel associated with the formation of the $Ne^{2+}([^4S]3s^3S)$ ion from the $Ne^{3+}(^2D)$ state, to which we can attribute the experimentally observed d–^2D peak in the spectrum at ΔE = 4.8 eV, falls into this energy interval. These processes change the multiplicity of the core of a multiply charged ion and can occur only as the electron capture to the 2p subshell of the impinging ion with simultaneous excitation of a 2p electron to state 3s and quenching of the metastable state. Simple capture with an electron transition to the 3s level does not cause quenching of the metastable state by virtue of the spin conservation law; this will be confirmed experimentally in the subsequent analysis of the interaction of $Ne^{3+}(^{2s+1}L)$ ions with H atoms and H_2 molecules. The ions formed as a result of electron capture to the ground state of the $Ne^{3+}(2s^22p^4{}^3P)$ ion with simultaneous excitation of the He^+ target ion (n = 2) fall approximately in the same range of variation of kinetic energy (third group of processes in the table, denoted by c–^2P and c–^2D in Fig. 2).

The absence of the peak corresponding to electron capture by ions in the ground state (a^4–S in Fig. 2) just in the middle of the spectrum of kinetic energy variation ΔE of the Ne^{2+} ions formed in the process suggests that Ne^{3+} ion in the ground state are practically not formed in the ion source during the production of triply charged neon ions by an electron impact under the conditions of single collisions. This also implies that the lifetime of metastable Ne^{3+} ions is much longer than the transit time ($t < 1.5$ μs) to the collision chamber since the decay of metastable ions does not result in the formation of ions in the ground state in the beam used in experiments.

The change in the electron energy in an ion source in the range 150–1500 eV has not led to an appreciable change in the ΔE spectrum being measured for the electron capture by Ne^{3+} ions and, hence, to a change in the ion beam composition. Thus, the mechanism of formation of the multiply charged Ne^{3+} ion under the conditions of single electron–atom collisions is mainly associated with electrons of the outer shell $2s^2\ 2p^6$ of the neon atom. In all probability, the formation of the 2s vacancy in the neon atom is accompanied by excitation of 2p electrons with subsequent decay of excited states. Conversely, the role of K vacancies in the neon atom is insignificant even for the energy of the electron beam in the ion source, which is approximately twice as high as the binding energy of a 1s electron (867 eV [14]).

The presence of only metastable Ne^{3+} ions in the ion beam is in qualitative agreement with the results obtained in [15, 16]. The measured spectra of the Ne^{2+} ions formed during collisions of Ne^{3+} ions with energies of 600 eV [15] and 150 eV [16] with He atoms show that the electron capture by the metastable $Ne^{3+}(^2P)$ ion is the main process accompanied by excitation of a 2s electron of the multiply charged ion core, which is denoted by b–^2P in Fig. 2. The contribution from processes leading to the formation of the triplet state $Ne^{2+}(^3P)$ (denoted by a–^2P and a–^2D) for collision energies studied in [15, 16] is insignificant. A compari-

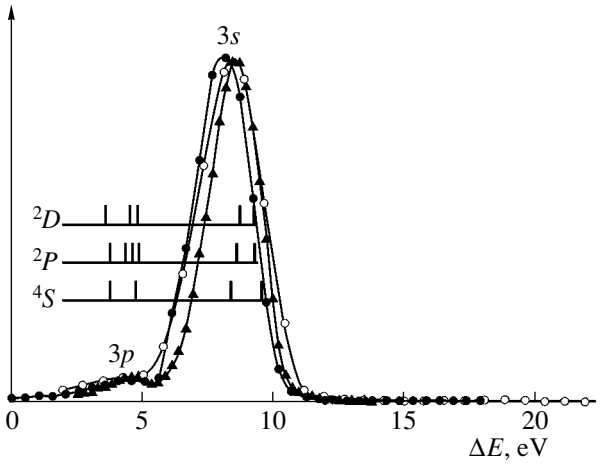


Fig. 3. Spectrum of kinetic energy variation ΔE for ions in the case of single-electron capture by Ne^{3+} ion from H_2 molecules. Initial energies E of ions are 3 (▲), 6 (■), and 9 keV (●). Notation is the same as in Fig. 1.

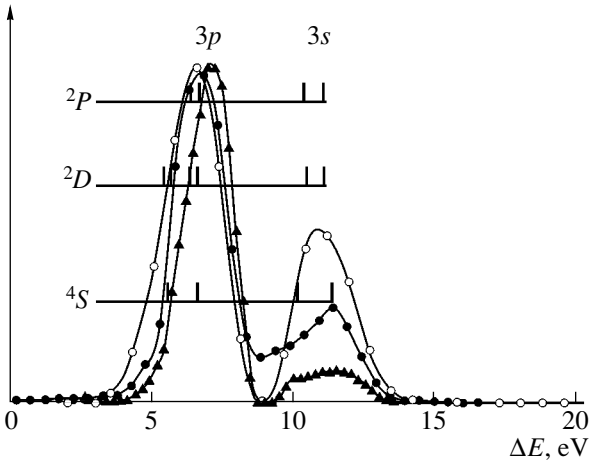


Fig. 4. Spectrum of kinetic energy variation ΔE for ions in the case of single-electron capture by Ne^{3+} ions from H atoms. Notation is the same as in Fig. 3.

son of the spectra obtained by us at E = 2.70 and 5.25 keV shows that the relative role of the latter processes increases indeed with the collision energy.

Thus, electrons from the 2*s* and 2*p* subshells of neon and 1*s* subshell of helium are involved in the simple (from the standpoint of change in the charge state of the particles) process of single-electron capture by Ne^{3+} ions from He atoms; consequently, it is a complex multielectron process.

(c) Interaction of Ne^{3+} ions with H atoms and H_2 molecules. Electron capture by a metastable ion may lead to a rearrangement of the electron configuration of the multiply charged ion core and to quenching of the metastable state during the interaction. This occurs during the interaction of triply charged Ne^{3+} ions with helium atoms, when an electron is captured directly to an incompletely filled shell in the core of the multiply charged ion. The answer to the question about quenching of the metastable state at the instant of interaction during electron capture to higher levels of the impinging ion, which do not belong to the core of the multiply charged ion can be obtained either by reducing the collision energy, or by using the target atoms with an ionization potential lower than the ionization potential of the helium atom.

Figure 3 shows the results of measurements of the energy of Ne^{3+} ions with initial energies of 3, 6, and 9 keV during the electron capture from hydrogen molecules. The higher peak in the spectrum at $\Delta E \sim 9$ eV corresponds to electron capture to the 3*s* level of the Ne^{2+} ion. Capture to the 3*p* level is also observed. These channels of the process are not associated with exchange of electrons between the 2*p* and 3*s* subshells as in the case of electron capture for the Ne^{3+}–He pair and do not cause a rearrangement of the core of metastable ions. As a results, the 3*s* and 3*p* peaks can be attributed to electron capture by the Ne^{3+} ion in any of three states. Only one electron is active during capture, and the process can be described by the single-electron approximation. In this case, the hydrogen molecule can be treated as an indivisible particle with an ionization potential equal to 15.43 eV. The possibility of such a treatment of the molecule is based on the fact that the role of processes accompanied by dissociation is relatively small for the system studied here. This follows, first, from analysis of the obtained energy spectra and, second, from direct measurements by the method of coincidence of the electron capture cross sections, viz., capture with dissociation and capture with ionization, which were performed by us for the Ar^{z+}–He_2 pairs (z = 3, 6). Our measurements showed that the sum of the cross sections of processes accompanied by dissociation of molecules in a wide range of collision rates does not exceed 10% (for z = 3) and 20% (for z = 6) of the total single-electron capture cross section. Insignificant role of dissociation processes even for multielectron molecules (such as N_2 and O_2) in the electron capture by slow Ne^{3+} ions (E = 75 eV) was demonstrated by Kamber and Ferguson [7]. They also observed peaks in the energy loss spectrum, which are associated with electron capture by metastable $Ne^{3+}(^2D)$ ions with the formation of the excited $Ne^{2+}([^2P]3s^3P, {}^1P)$ ion.

Thus, the electron capture during a collision with the hydrogen molecule whose electron has a binding energy lower than that in the helium atom does not cause a rearrangement of the electron L shell of the Ne^{3+} ion in the collision energy range under investigation.

Figure 4 shows the spectrum of kinetic energy variation for Ne^{2+} ions formed during the capture of a single electron from hydrogen atoms by Ne^{3+} ions with energy 3.0, 5.25, and 9.0 keV. Since the binding energy of the

electron in a hydrogen atom is lower than in a molecule, the main channel of the process is population of the higher-lying 3*p* level of the Ne^{2+} ion as compared to the 3*s* level. Such a transition to the 3*p* subshell is carried out for still longer nuclear spacings corresponding to intersection of the energy levels of the initial and final states for a nuclear spacing of about 8.5 a.u., while these distances for populating the 3*s* subshell are ~5 a.u. Owing to such an arrangement of the regions of interaction, the relative probability of capture to the 3*s* subshell increases with collision rate (see Fig. 4). All conclusions concerning the effect of electron states and the participation of electrons from the core of multiply charged ions on the single-electron capture, which were drawn for collisions with hydrogen molecules, are also confirmed for the interaction with hydrogen atoms.

CONCLUSIONS

Population of the electron states of ions formed as a result of capture of a single electron during the interaction of multiply charged ions having an electron core with atoms is analyzed. It is shown that electron capture can occur, with a high probability, as a two-electron process and can be accompanied by excitation of the multiply charged ion core. The cross sections of exothermal channels of this capture are effective in the range of potential energy variation $Q \cong +5 + 25$ eV in the system of colliding particles with transfer of the corresponding kinetic energy to particles participating in a collision. The cross sections of such processes may exceed by many times the cross sections of competing single-electron processes, either endothermic, or with $Q > 30$ eV. In the presence of multiply charged ions in metastable states in the beam, the metastable state is effectively quenched upon the capture of an electron to the excited state of the ion core. The method of collision spectroscopy used by us can be employed for diagnostics of electron states of particles in a beam of multiply charged ions. Choosing atoms with different ionization potentials as target particles at which electron capture takes place and changing the collision energy, it is possible to cause effective quenching of metastable states of ions with different multiplicity.

The formation of multiply charged ions in metastable states during ionization by an electron impact indicates that multiple ionization of an atom by an electron impact is associated with excitation of the outer atomic shell and its subsequent relaxation; the multiply charged ion formed in this process may remain in the excited state, which is obviously not necessarily metastable. The possibility of a high probability of populating long-lived excited states during multiple ionization of atoms by an electron impact is demonstrated. This effect must be taken into account in interpreting the variation of the multiple ionization potentials of heavy atoms, molecules, and especially clusters if measurements are made under conditions of single collisions of heavy particles with electrons.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project no. 02-02-17590) and the program "Scientific Schools" (grant NSh-2215.2003.2).

REFERENCES

1. A. Salop and R. E. Olson, Phys. Rev. A **13**, 1312 (1976).
2. Yu. Sato and J. H. Moore, Phys. Rev. A **19**, 495 (1979).
3. M. N. Gaboriaud, M. Barat, P. Roncin, *et al.*, J. Phys. B **27**, 4595 (1994).
4. E. Bloemen, H. Winter, F. J. de Heer, *et al.*, J. Phys. B **11**, 4207 (1978).
5. V. V. Afrosimov, A. A. Basalaev, A. V. Samoĭlov, *et al.*, Zh. Éksp. Teor. Fiz. **91**, 465 (1986) [Sov. Phys. JETP **64**, 273 (1986)].
6. C. Schmeissner, C. L. Cocke, R. Mann, *et al.*, Phys. Rev. A **30**, 1661 (1984).
7. E. Y. Kamber and S. M. Ferguson, Phys. Rev. A **63**, 022701 (2001).
8. V. V. Afrosimov, A. A. Basalaev, G. A. Leĭko, *et al.*, Zh. Éksp. Teor. Fiz. **74**, 1605 (1978) [Sov. Phys. JETP **47**, 837 (1978)].
9. V. K. Nikulin, D. Dijkkam, Yu. S. Gordeev, *et al.*, J. Phys. B **17**, L721 (1984).
10. V. V. Afrosimov, G. A. Leiko, and M. N. Panov, in *Proceedings of the 9th International Conference on Photonic, Electronic and Atomic Collisions (ICPEAC), Seattle, 1975*, pp. 183–184.
11. C. E. Moore, *Atomic Energy Levels: NSRDS–NBS No. 35* (US Government Printing Office, Washington DC, 1971).
12. *NIST Atomic Spectra Database*, http://physics.nist.gov/cgi-bin/AtData/main_asd.
13. K. Motohashi and S. Tsurubuchi, J. Phys. B **36**, 1811 (2003).
14. K. Siegbahn, C. Nordling, A. Fahlman, H. Hamrin, J. Hedman, G. Johansson, T. Bergmark, S. E. Karlsson, J. Lindgren, and B. Lindberg, *Electron Spectroscopy for Chemical Analysis: Atomic, Molecular and Solid State Structure Studies by Means of Electron Spectroscopy* (Almqvist and Wiksells, Uppsala, 1967; Mir, Moscow, 1971).
15. H. Lebius, H. R. Koslowski, and B. A. Huber, Z. Phys. D **11**, 53 (1989).
16. R. Said, E. Y. Kamber, S. Yaltkaya, *et al.*, J. Phys. B **27**, 3993 (1994).

Translated by N. Wadhwa

Technical Physics, Vol. 50, No. 8, 2005, pp. 993–998. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 32–36.

GASES AND LIQUIDS

Wave Instability of a Molten Metal Layer Formed by Intense Laser Irradiation

F. Kh. Mirzade

Institute of Problems of Laser and Information Technologies, Russian Academy of Sciences, Moscow Oblast, 140700 Russia

e-mail: fmirzade@rambler.ru

Received May 20, 2004

Abstract—Wave structure of a molten metal layer flowing over the walls of a vapor–gas cavern that appears as an intense laser radiation penetrates deep into condensed media is studied theoretically taking into account surface tension, gravitation, thermocapillary effect, and nonuniform evaporation from the free surface of the melt. A long-wavelength evolution equation describing the evolution of nonlinear waves on the free surface of a plane molten layer is derived. The spatially periodic running solution to this equation is obtained, and the main characteristics (amplitude and period) of the nonlinear wave structures are determined. © *2005 Pleiades Publishing, Inc.*

The evolution of perturbations on the free surface of a molten metal layer that is formed under the action of highly intense external energy fluxes (laser or electron beams) on condensed media is governed by the capillary and thermocapillary forces, as well as by the vapor recoil pressure. Another factor that plays a significant role in the wave perturbation dynamics is the process of direct mass efflux, which accompanies spatially nonuniform evaporation. The viscous liquid layer moving under the action of various forces presents an example of an active–dissipative condensed medium, where the energy input is provided by external forces (for example, gravitation or thermocapillary forces) and the dissipation is associated with the viscous friction forces [1]. The instability of the interface significantly affects the behavior of the heat- and mass-transfer processes, primarily, due to wave formation, turbulization, and droplet detachment.

There is a large number of theoretical studies devoted to the wave processes that accompany the flow of a thin viscous liquid layer in both the linear and nonlinear statements of the problem in the case of free flowing (see, e.g., [1, 2]). The channeled penetration of intense laser beams into metals gives rise to the formation of vapor–gas caverns in the material. Studying the hydrodynamic instabilities that are responsible for the formation of nonlinear wave structures (for example, nonlinear periodic waves) on the molten walls of such caverns is complicated by their relation to the heat transfer and evaporation processes, as well as the interaction of the metal vapor with the external energy flux [3]. In this study, we make a pioneering attempt to investigate, as a first approximation, the evolution of nonlinear surface waves appearing in the molten metal layer flowing over the melt front on the front wall of the cavern (with a conic configuration) with taking into account the surface tension, gravitation, thermocapillary effect, and nonuniform evaporation from the free surface of the melt.

We consider the motion of a layer of viscous incompressible liquid evaporating under the laser irradiation; the layer is bounded by the vapor–gas flow on one side and a semi-infinite surface inclined by angle ϕ to the horizontal on the other. The flow is assumed to be planar and periodic with wavelength Λ_0. The schematic diagram of the flow is shown in the figure. The system of coordinates is chosen so that the x and y axes are parallel and perpendicular to the unpertubed interface, respectively. The free surface equation is taken in the form $y = h(x, t)$, where h is the melt thickness, t is time, and $y = 0$ is the solid–liquid interface. The process takes place in vacuum conditions. Absorption of radiation in the vapor phase is ignored.

Subject to the assumptions noted above, the equations for the melt dynamics and the heat transfer equation have the form [3, 4]

$$\nabla \cdot \mathbf{v} = 0, \tag{1}$$

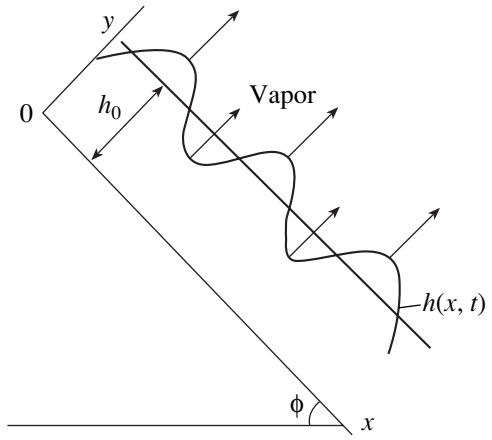
$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} = \mathbf{g} - \frac{1}{\rho}\nabla p + \nu \Delta \mathbf{v}, \tag{2}$$

$$\frac{\partial T}{\partial t} + \mathbf{v}\nabla T = \chi \Delta T, \tag{3}$$

where $\mathbf{v} = (u, v)$ is the velocity vector of the melt with components u and v directed along the x and y axes, respectively; p is the pressure; ρ is the density; $\mathbf{g}$ is the acceleration of gravity; ν is the kinematic viscosity; T is the temperature; and $\Delta = \partial^2/\partial x^2 + \partial^2/\partial y^2$ is the Laplace operator.

In the above system, Eq. (1) is the continuity equation, Eq. (2) is the Navier–Stokes equation for two

1063-7842/05/5008-0993$26.00



Wave flow of molten metal motion of over the melt front; $h(x, t)$ is the melt thickness and ϕ is the angle of inclination.

velocity components, and Eq. (3) is the heat conduction equation. A solution to system (1)–(3) should meet the following boundary conditions:

$$q \equiv \rho(\mathbf{v} - \mathbf{v}_{(i)}) \cdot \mathbf{n} = \rho^{(v)}(\mathbf{v}^{(v)} - \mathbf{v}_{(i)}) \cdot \mathbf{n}, \quad (4)$$

$$q\lambda_v + k\nabla T \cdot \mathbf{n} - k^{(v)}\nabla T^{(v)} \cdot \mathbf{n} - q_L = 0, \quad (5)$$

$$p - p^{(v)} + q_v(\mathbf{v} - \mathbf{v}^{(v)}) \cdot \mathbf{n} - (s - s^{(v)})\mathbf{n} \cdot \mathbf{n} = 2\sigma(T)H, \quad (6)$$

$$q_v(\mathbf{v} - \mathbf{v}^{(v)}) \cdot \boldsymbol{\tau} - (s - s^{(v)})\mathbf{n} \cdot \boldsymbol{\tau} = -\nabla\sigma \cdot \boldsymbol{\tau}, \quad (7)$$

$$(\mathbf{v} - \mathbf{v}^{(v)}) \cdot \boldsymbol{\tau} = 0, \quad (8)$$

$$\mathbf{n} = (-h_x, 1)/Y, \quad \boldsymbol{\tau} = (1, h_x)/Y, \quad Y = (1 + h_x^2)^{1/2},$$

on the free surface ($y = h(x, t)$) [5] and

$$\mathbf{v} = 0, \quad T = T_m$$

at the solid–liquid interface ($y = 0$).

In the above expressions, q is the density of the mass flow due to the evaporation of fluid liquid; $\mathbf{v}_{(i)}$ is the velocity of particles at the liquid–vapor interface; the quantities with and without the superscript correspond to the vapor and liquid, respectively; λ_v is the latent heat of evaporation; k and $k^{(v)}$ are the thermal conductivity coefficients of fluid liquid and vapor, respectively; s is the viscous stress tensor; q_L is the absorbed laser energy; T_m is the melting point; $2H = \nabla \cdot \mathbf{n}$ is the mean curvature of free surface; and $\mathbf{n}$ and $\boldsymbol{\tau}$ are the normal and tangent unit vectors, respectively.

Equation (4) presents the definition and the conservation law of the mass flux (q), and Eqs. (5)–(7) describe the conservation of energy and normal and tangential momentum components. The energy balance equation takes into account the absorption of laser energy and the energy loss through evaporation and heat conduction. The heat balance is assumed to be independent of the viscous dissipation and of the radiation loss on the free surface of the melt. Condition (8) expresses the tangential velocity continuity across the interface. Surface tension σ of the melt is related to temperature as $\sigma(T) = \sigma_v - \sigma_T(T - T_v)$, where $\sigma_T = -d\sigma/dT$ is the temperature coefficient of the surface tension. For most liquids, $\sigma_T > 0$. The density, viscosity, and thermal conductivity of the melt are considered as temperature independent.

For the stress tensor, we have $s = \mu_v(\partial u/\partial y + \partial v/\partial x)$, were μ_v is the kinematic viscosity. The boundary conditions expressed by Eqs. (4)–(8) should be supplemented with the formula for the mass flux density as a function of the local temperature on the free surface [5],

$$q = \left(\frac{k_0\rho_v\lambda_v}{T_v^{3/2}}\right)\left(\frac{M_w}{2\pi r_g}\right)(T - T_v), \quad (9)$$

where k_0 is the accommodation coefficient, r_g is the universal gas constant, M_w is the molecular weight of the liquid, and T_v is the temperature of the evaporation front.

Let a plane layer of an evaporating viscous liquid flow over an inclined surface. For arbitrary Reynolds numbers (Re), the system of Navier–Stokes and heat conduction equations has a stationary spatially inhomogeneous solution, which defines film thickness h_0, temperature distribution T_0, and the field of velocities V_0. This flow mode is unstable with respect to infinitesimal perturbations even at very small Re numbers. The growth of long-wave perturbations is limited by nonlinear effects and leads to the formation of nonlinear wave modes with finite amplitudes.

Let v_0 denote a characteristic velocity equal to the maximum x component of the melt flow velocity attained at the upper (free) boundary of the melt layer in the case of a laminar flow. Using h_0/v_0, v_0, ρv_0^2, $\Delta T = T_v - T_m$, and $k\Delta T/h_0\lambda_v$ as the scales for time, velocity, pressure, temperature, and the mass flux density, we pass to dimensionless variables in the following way:

$$(x', y', t') = \left(\frac{\varepsilon x}{h_0}, \frac{y}{h_0}, \frac{\varepsilon v_0 t}{h_0}\right),$$

$$h' = \frac{h}{h_0}, \quad (u', v') = \left(\frac{u}{v_0}, \frac{v}{\varepsilon v_0}\right),$$

$$p' = \frac{p}{\rho v_0^2}, \quad v_0 = \frac{g(1 - \Lambda_p)}{2\nu}h_0^2, \quad q' = \frac{h_0\lambda_v q}{k\Delta T},$$

$$\Delta T = \frac{q_L h_0}{k}, \quad T' = \frac{T - T_m}{\Delta T}.$$

In the above expressions, the longitudinal scale Λ_0 is of the order of the perturbation wavelength at the melt surface; h_0 is the scale of the melt layer thickness (for

example, the average thickness), which depends solely on longitudinal coordinate x; and $\Lambda_\rho = \rho^{(v)}/\rho$. To simplify the writing, we below omit the primes in the superscripts of the dimensionless variables and imply that $x = x'$, $y = y'$, $h = h'$, etc. Then, system of equations (1)–(3) with boundary conditions (4)–(8) appears as

$$u_x + v_y = 0, \tag{10}$$

$$u_t + vu_y + uu_x = (\varepsilon \mathrm{Fr})^{-1} - p_x + (\varepsilon \mathrm{Re})^{-1}\Delta u, \tag{11}$$

$$\varepsilon^2(v_t + vv_y + uv_x) = -\cot\phi \mathrm{Fr}^{-1} - p_y + \varepsilon \mathrm{Re}^{-1}\Delta v, \tag{12}$$

$$T_t + vT_y + uT_x = (\varepsilon \mathrm{Pe})^{-1}\Delta T. \tag{13}$$

On the free surface $y = h(x, t)$, we have:

$$q = \varepsilon K(v - h_t - uh_x)Y^{-1}, \tag{14}$$

$$\varepsilon K(v - h_t - uh_x)Y^{-1} + (T_y - \varepsilon^2 h_x T_x)Y^{-1} + q_L = 0, \tag{15}$$

$$\begin{aligned} p + \varepsilon^2 W h_{xx} Y^{-3} - K^{-2}[(T_y - \varepsilon^2 T_x h_x)Y^{-1} - q_L]^2 \\ + 2\varepsilon \mathrm{Re}^{-1} u_x (1 + \varepsilon^2 h_x^2)(1 - \varepsilon^2 h_x^2)^{-1} = 0, \end{aligned} \tag{16}$$

$$\begin{aligned} 4\varepsilon^2 u_x h_x - (u_y + \varepsilon^2 v_x)(1 - \varepsilon^2 h_x^2) \\ = 2\varepsilon \mathrm{MaPr}^{-1}(T_x + T_y h_x)Y, \end{aligned} \tag{17}$$

and at the solid–liquid interface ($y = 0$), $u = v = 0$ and $T = 0$.

The parameters used in Eqs. (10)–(17) are the following: the Reynolds number $\mathrm{Re} = v_0 h_0/\nu$; the Prandtl number $\mathrm{Pr} = \nu/\chi$; the Peclet number $\mathrm{Pe} = \mathrm{Re}\,\mathrm{Pr}$; the Weber number $W = \sigma/\rho v_0^2 h_0$; the Kutateladze number $\mathrm{Ku} = \lambda_v/C_p\Delta T$; the Froude number $\mathrm{Fr} = v_0^2/gh_0\sin\phi$; the Marangoni number $\mathrm{Ma} = \sigma_T \Delta T h_0/2\rho\nu k$; $K = \mathrm{RePrKu}$, $F = \mathrm{Re}/\mathrm{Fr}$; $\Omega = 2\lambda_v/v_0^2$; and $\varepsilon = h_0/\Lambda_0$, where C_p is the heat capacity, χ is the thermal diffusivity; and $q_L = \alpha_L I$ (I is the intensity of laser radiation and α_L is the absorption coefficient). Subscript $i = (x, y, t)$ designates the partial derivative with respect to the corresponding variable.

Note that parameter K^{-1} presents the ratio between the viscous time scale $t_\nu = h_0^2/\nu$ and vaporization time scale $t_E = h_0^2 \lambda_v/k\Delta T$ (t_E is the time taken to complete evaporation of the liquid): $K^{-1} = E = k\Delta T/\rho\nu\lambda_v$.

In dimensionless form, Eq. (9) becomes

$$q = N^{-1}(T - 1), \tag{18}$$

where $N = (kT_v^{3/2}/k_0 h_0 \rho^{(v)} \lambda_v^{(2)})(2\pi r_g/M_w)^{1/2}$.

Parameter N characterizes the degree of nonequilibrium of the evaporating liquid surface. The condition $N = 0$ corresponds to the quasi-equilibrium approximation; $N^{-1} = 0$, to the absence of the generated vapor flow (the case of nonvolatile liquids).

Let a solution to system of equations (10)–(18) be expanded in terms of small parameter $\varepsilon \ll 1$ (long-wavelength approximation)

$$Z = \sum_{l=0}^{\infty} \varepsilon^1 Z_1, \tag{19}$$

where $Z = \{U, v, p, q, T\}$.

Substituting expansion (19) into Eqs. (10)–(18) and setting the coefficients with the same powers of ε equal to zero, we obtain a system of differential equations. Having solved the system, we express the velocity and temperature fields as polynomials of the transverse coordinate with the coefficients depending on thickness $h(x, t)$ and its derivatives. Using condition (15) and leaving terms up to ε^2 inclusively, we arrive at the following evolutionary equation for the variable melt thickness:

$$\begin{aligned} h_t + a(h) + b(h)h_x + c(h)h_{xx} + d(h)h_{xxxx} + n_1(h)h_{tx} \\ + n_2(h)h_t h_x + n_3(h)h_x^2 + n_4(h)h_x h_{xxx} = 0, \end{aligned} \tag{20}$$

where a, b, c, d, n_1, n_2, n_3, and n_4 are defined by formulas

$$a(h) = (q_L h - 1)/\varepsilon K(h + N),$$

$$b(h) = F[(5\mathrm{Re}/8K - 1)h^2 - ((5\mathrm{Re}/6\mathrm{K})q_L - \cot\phi/3)h^3],$$

$$c(h) = \varepsilon \mathrm{Re}[2(1 - q_L h)h^3/3\Lambda_\rho K^2(h + N)^3 + 2F^2 h^6/15 + \mathrm{MaPr}^{-1} N(h/(n + N))^2],$$

$$d(n) = \mathrm{Re}W\varepsilon^3 h^3/3, \quad n_1(h) = 5\varepsilon \mathrm{Re}Fh^4/2,$$

$$n_2(h) = 5\varepsilon \mathrm{Re}Fh^4/6,$$

$$n_3(h) = \varepsilon(9\mathrm{Re}F^2 h^5/20 + h^2\cot\phi/2),$$

$$n_4(h) = \varepsilon \mathrm{Re}Wh^2.$$

Thus, within the limitations considered, the investigation of the motion of evaporating liquid over an inclined plane surface is reduced to the analysis of the solutions to the nonlinear partial differential equation for variable thickness $h(x, t)$. The second term on the left-hand side of Eq. (20) describes the mass lost by the liquid via direct evaporation and due to the action of the external energy flux. The next two terms account for the influence of the vapor reactive force exerted on the melt surface and the thermocapillary effect, respectively; the fifth term stands for the influence of the surface tension. The following terms characterize the change in the liquid layer thickness due to inertial forces.

In the stationary state ($h_t = 0$), Eq. (20) has the spatially inhomogeneous solution $h = h_0(x)$, which describes the waveless flow of the molten layer over the inclined surface. Accurate to the first coordinate derivative of the thickness, the solution can be numerically found from the equation $b(h_0)h_{0x} + a(h_0) = 0$.

We now explore the linear stability of the solution $h = h_0(x)$ and possible nonlinear modes that may stem from the development of unstable perturbations. Using Eq. (20) with the change $h = h_0 + \xi$ (ξ is the free surface perturbation) and making appropriate transformations, we arrive at a nonlinear partial differential equation (which is omitted here because of its complexity) with x-dependent coefficients. However, an exact analytic solution to this equation that describes both the temporal and spatial behavior of perturbations (ξ) cannot be obtained. The analysis is substantially simplified when restricted to the case $K \gg 1$ (or small $\Delta T \ll \Delta T_* = \lambda_\nu \mathrm{RePr}/C_p$). For example, for molten iron, $\Delta T_* = 10^3$ K. Considering a slightly varying basic flow as a background, we proceed with analyzing the linear stability and possible nonlinear spatially periodic (along the x coordinate) solutions. This approximation is valid on the condition that $|d\xi/dx| \approx \varepsilon \gg |dh_0/dx|$, i.e., when the wavelength (λ) of the periodic perturbations considered here is much smaller than the characteristic scale of changes in the main flow. Note that the approach used (quasi-parallel approximation) is widely applied for solving problems with spatial inhomogeneity [1].

Taking into consideration the aforesaid, we represent the solution to Eq. (20) in the form $h = 1 + \xi$, where 1 is the equilibrium value of the dimensionless thickness. As a result, we have

$$\widehat{L}\xi = -\frac{1}{2}a_1''\xi^2 - \left(b_1'\xi + \frac{1}{2}b_1''\xi^2\right)\xi_x - \left(c_1'\xi + \frac{1}{2}c_1''\xi^2\right)\xi_{xx} - \left(d_1'\xi + \frac{1}{2}d_1''\xi^2\right)\xi_{xxxx} + O(\xi^4), \tag{21}$$

where

$$\widehat{L} = \frac{\partial}{\partial t} + a_1' + b_1\frac{\partial}{\partial x} + c_1\frac{\partial^2}{\partial x^2} + d_1\frac{\partial^4}{\partial x^4}$$

(b_1, c_1, d_1, a_1', b_1', b_1'', c_1', c_1'', d_1', d_1'' are the values of a, b, c, d and their derivatives at $h = 1$; in what follows, the subscript 1 on these quantities is omitted).

Let us study the stability of the trivial solution $\xi \equiv 0$ with respect to the perturbations in the form

$$\xi = \eta \exp[i\alpha(x - \gamma t)] + \text{c.c.}, \tag{22}$$

where $\gamma = \gamma_r + i\gamma_i$ is the dimensionless complex increment of the instability, γ_r is the instability frequency, γ_i is the instability build-up increment, η is the perturbation amplitude, and $\alpha = 2\pi h_0/\lambda_0$ is the dimensionless wave number.

The substitution of formula (22) into the linearized equation $\widehat{L}\xi = 0$ yields

$$\gamma = -i(a' - c + d) + b,$$

whence

$$\gamma_r = F\left(1 + \frac{5}{6}\frac{\mathrm{Re}}{K}\left(\frac{3}{4} - q_L\right) + \frac{\cot\phi}{3}\right),$$

$$\gamma_i = \frac{2\alpha\mathrm{Re}}{3}[F^2/5 + (1 - q_L)/\Lambda_\rho K^2(1 + N)^3 + 3\mathrm{Ma}N/2\mathrm{Pr}(N + 1)^2] - \mathrm{Re}W\alpha^3/3 + (q_L N + 1)/\alpha K(N + 1)^2. \tag{23}$$

If $\gamma_i > 0$, the perturbation grows with time and the waveless flow is unstable. On the contrary, if $\gamma_i < 0$, the waveless flow is stable. The condition $\gamma_i = 0$ defines the boundary of the neutral stability.

For the maximum increment (γ_{i_m}) and wave period (d_L), expressions (23) yield

$$\gamma_{im} = \beta^2\mathrm{Re}/3W + E(q_L N + 1)/(1 + N)^2, \quad d_L = 2\pi/\alpha_* = 2\pi\sqrt{W\beta^{-1}}, \tag{24}$$

where

$$\beta = F^2/5 + E^2(1 - q_L)/\Lambda_\rho(1 + N)^3 + 3\mathrm{Ma}N/2\mathrm{Pr}(1 + N)^2.$$

The analysis of formulas (23) and (24) suggests that the loss of stability occurs at a certain critical value of the Reynolds number $\mathrm{Re}_* = 3$ (the physical characteristics correspond to iron: Pr = 4, W = 10^5, Ku = 6, and $\Lambda_\rho = 0.003$). The inertial parameters (Re number) and the thermocapillary effect (Ma number) play a destabilizing role in the instability development, while the surface tension (W number) acts as a stabilization factor. In view of the fact that the transverse heat flux in the ridges is lower than that in the valleys, the melt thickness increases in the valleys, which precludes the development of small perturbations of the free surface. For high values of Ku, the destabilizing role of the number E^2 and, therefore, of the reaction forces caused by an intense flow of the evaporating mass (as well as the role of the inertial effects) becomes dominating.

Let a set of spatial and time scales be introduced at the nonlinear stage of the instability evolution:

$$x_n = \varepsilon^n x, \quad t_n = \varepsilon^n t \ (n = 0, 1, 2, \ldots).$$

In all the expressions for differential operators in (21), we make the following change:

$$\frac{\partial}{\partial t} \longrightarrow \frac{\partial}{\partial t_0} + \varepsilon\frac{\partial}{\partial t_1} + \ldots, \quad \frac{\partial}{\partial x} \longrightarrow \frac{\partial}{\partial x_0} + \varepsilon\frac{\partial}{\partial x_1} + \ldots,$$

$$\frac{\partial^2}{\partial x^2} \longrightarrow \frac{\partial^2}{\partial x_0^2} + 2\varepsilon\frac{\partial^2}{\partial x_0 \partial x_1} + \varepsilon^2\left(\frac{\partial^2}{\partial x_1^2} + 2\frac{\partial^2}{\partial x_2 \partial x_0}\right) + \ldots,$$

$$\frac{\partial^4}{\partial x^4} \longrightarrow \frac{\partial^4}{\partial x_0^4} + 4\varepsilon\frac{\partial^4}{\partial x_1 \partial x_0^3} + \varepsilon^2\left(6\frac{\partial^4}{\partial x_0^2 \partial x_1^2} + 4\frac{\partial^4}{\partial x_2 \partial x_0^3}\right) + \ldots.$$

Let us expand the solution to Eq. (21) into a series in small parameter ε and leave the terms $O(\varepsilon^3)$,

$$\begin{aligned}(\widehat{L}_0 + \varepsilon\widehat{L}_1 + \varepsilon^2\widehat{L}_2)(\varepsilon\xi_1 + \varepsilon^2\xi_2 + \varepsilon^3\xi_3) \\ = -\varepsilon^2 Nl_2 - \varepsilon^3 Nl_3,\end{aligned} \tag{25}$$

where

$$\widehat{L}_0 = \widehat{L}, \quad \widehat{L}_1 = \frac{\partial}{\partial t_1} + b\frac{\partial}{\partial x_1} + 2c\frac{\partial^2}{\partial x \partial x_1} + 4d\frac{\partial^4}{\partial x^3 \partial x_1},$$

$$\widehat{L}_2 = \frac{\partial}{\partial t_2} + c\frac{\partial^2}{\partial x_1^2} + 6d\frac{\partial^4}{\partial x^2 \partial x_1^2},$$

$$Nl_2 = \frac{1}{2}a''\xi_1^2 + b'\xi_1\xi_{1x} + c'\xi_1\xi_{1xx} + d'\xi_1\xi_{1xxxx} + n_1'\xi_{1x}\xi_{1xxx_1},$$

$$Nl_3 = \frac{1}{2}a''\xi_1\xi_2 + b'(\xi_1\xi_{2x} + \xi_1\xi_{1x} + \xi_2\xi_{1x}) + \frac{1}{2}b''\xi_1^2\xi_{1x} + c'(\xi_1\xi_{2xx} + 2\xi_1\xi_{1xx_1} + \xi_2\xi_{1xx}) + \frac{1}{2}c''\xi_1^2\xi_{1xx} + d'(\xi_1\xi_{2xxxx} + \xi_2\xi_{1xxxx}) + \frac{1}{2}d''\xi_1^2\xi_{1xxxx}.$$

Consider now the spatially periodic solutions with a finite amplitude and low modulation over the time interval of the order of $O(\varepsilon^{-m})$. Parameter ε is taken to be of the order of the harmonic amplitude in (22).

Equating the coefficients with the same powers of ε, we obtain

$$\widehat{L}_0\xi_1 = 0, \tag{26}$$

$$\widehat{L}_1\xi_1 + \widehat{L}_0\xi_2 = Nl_2, \tag{27}$$

$$(\widehat{L}_0\xi_3 + \widehat{L}_1\xi_2 + \widehat{L}_2\xi_1) = Nl_3. \tag{28}$$

Solution (26), (27) can be represented in the form of various sums comprised of terms in the form

$$\xi_1 = \eta(x_1, t_1, t_2)e^{i\alpha x} + \text{c.c.} \tag{29}$$

Using Eq. (27) and the condition of solvability $\partial\eta/\partial t_1 + 2i\alpha(c - 4d)\partial\eta/\partial x_1 = 0$, we derive an equation for ξ_2,

$$c\xi_{2xx} + d\xi_{2xxxx} = -(a''/2 + ib' - c' + d')\eta^2 e^{2i\alpha x}.$$

From whence it follows

$$\begin{aligned}\xi_2 = R\eta^2 e^{2i\alpha x}, \quad R = R_r + iR_i, \\ R_r = \frac{a'' + 2(c' - d')}{16\alpha(4d - c)}, \quad R_i = \frac{b'}{16\alpha(4d - c)}.\end{aligned} \tag{30}$$

Combining solutions (30) and the secular condition for $O(\varepsilon^3)$, we have

$$\frac{\partial\eta}{\partial t_2} + i\varepsilon^{-1}Q\frac{\partial\eta}{\partial x_1} - \varepsilon^{-2}\gamma_i\eta - D\frac{\partial^2\eta}{\partial x_1^2} + U\eta^2\bar{\eta} = 0, \tag{31}$$

where

$$D = 6d - c, \quad Q = 2(c - 4d), \quad U = U_r + iU_i,$$

$$U_r = -3b'R_i + (17d' - 5c')R_r + \frac{1}{2}(d'' - c''),$$

$$U_i = -3b'R_r + (17d' - 5c')R_i + \frac{1}{2}b''.$$

After the change $\eta = \psi\exp(i\phi_0 x_1)$, $\phi_0 = \varepsilon^{-1}Q/2D$, Eq. (31) takes the form

$$\frac{\partial\psi}{\partial t_2} - \varepsilon^{-2}\gamma_{1i}\psi - D\frac{\partial^2\psi}{\partial x_1^2} + U\psi^2\bar{\psi} = 0, \tag{32}$$

where $\gamma_{1i} = \gamma_i + Q/4D$.

Equation (32) is an amplitude equation (similar to the Landau–Ginzburg equation) describing the wave flow of a viscous molten metal layer over an inclined surface with account taken for the phase transition, thermocapillary effect, capillary forces, and gravitation. The equation allows one to investigate the nonlinear wave behavior of the melt in relation to the parameters of external energy fluxes and the physical properties of the medium.

Disregarding spatial variations, we can present the solution to Eq. (32) in the form

$$\psi = \psi_0\exp(-i\Gamma_0 t_2).$$

Substituting this expression into Eq. (32) and neglecting the diffusion terms, we obtain the nonlinear wave amplitude and the increment in the form

$$\psi_0 = (\gamma_i/\varepsilon^2 U)^{1/2}, \quad \varepsilon^2\Gamma_0 = U_i(\gamma_i/U_r).$$

Thus, we have studied the linear and nonlinear stages in the wave instability that evolves on the surface of a molten metal layer evaporating under the laser action and gives rise to the formation of stationary traveling nonlinear waves with a finite amplitude on the walls of the cavern. The corresponding nonlinear evolutionary equation for the variable thickness of the mol-

ten metal layer is derived from the system of the Navier–Stokes hydrodynamic equations and the heat conduction equation with the phase transition (evaporation) and thermocapillary effect being taken into account. An equation for the traveling wave amplitude similar to the Ginzburg–Landau equation is obtained in the third order of the perturbation theory. The main physical factors that govern the wave formation on the free surface of a plane melt layer and define the period and amplitude of the steady nonlinear wave structures are revealed.

ACKNOWLEDGMENTS

The author thanks L.A. Shelepin and V.S. Golubev for their participation in discussion of the results.

REFERENCES

1. L. P. Kholpanov and V. Ya. Shkadov, *Hydrodynamics and Heat-and-Mass Exchange with the Interface* (Nauka, Moscow, 1990) [in Russian].
2. Yu. Ya. Trifonov and O. Yu. Tsvelodub, J. Fluid Mech. **229**, 531 (1991).
3. F. Kh. Mirzoev, V. Ya. Panchenko, and L. A. Shelepin, Usp. Fiz. Nauk **166**, 3 (1996) [Phys. Usp. **38**, 1 (1996)].
4. S. W. Joo, S. H. Davis, and S. G. Bankoff, J. Fluid Mech. **230**, 117 (1991).
5. J. P. Burelbach, S. G. Bankoff, and S. H. Davis, J. Fluid Mech. **195**, 463 (1988).

Translated by A. Sidorova

Technical Physics, Vol. 50, No. 8, 2005, pp. 999–1005. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 75, No. 8, 2005, pp. 37–43.

GASES AND LIQUIDS

Completely Two-Dimensional Model for Examining the Characteristics of a Linear Cylindrical Induction Pump

I. R. Kirillov and D. M. Obukhov

Efremov Research Institute of Electrophysical Apparatus, Metallostroy, St. Petersburg, 196641 Russia

e-mail: kirillir@sintez.niiefa.spb.su

Received September 30, 2003

Abstract—A completely two-dimensional mathematical model for calculating the characteristics of induction magnetichydrodynamic (MHD) machines with a cylindrical channel is proposed. The flow pattern of a liquid metal in an electromagnetic pump under MHD instability is obtained by numeric analysis. This pattern is characterized by the formation of large-scale vortices traveling longitudinally and azimuthally. The calculated basic characteristics of the pump are in good qualitative and in satisfactory quantitative agreement with the experiment. © *2005 Pleiades Publishing, Inc.*

INTRODUCTION

In recent years, the instability of a uniform flow of a liquid metal in induction MHD machines has been intensively studied theoretically and by experiments [1–6]. It has been demonstrated that, under certain conditions (the induced magnetic field in the liquid metal is comparable to or stronger than the applied field, whereas electromagnetic forces far exceed the friction forces), such machines exhibit MHD instability of the flow. Under these conditions, an insignificant disturbance of the electromagnetic field or flow hydrodynamic characteristics results in developing perturbations and causes the instability of the uniform flow. It has been experimentally established [4–6] that, in induction electromagnetic pumps (EMB) with a cylindrical channel, the MHD instability of the flow is manifested in azimuthally nonuniform distribution of the liquid metal velocity and magnetic field, low-frequency pulsations of hydrodynamic (pressure, liquid metal flow rate) and electric (current, voltage) parameters of the pump, and in channel pulsation.

Theoretical investigation of the stability of induction MHD flows in the framework of the one-dimensional jet model [1, 2] confirms that, as soon as the uniform flow becomes unstable, secondary flows are formed and allow qualitative estimation of the possible velocity profiles of the secondary flows and the asymptotic value of the developed pressure. EMP characteristics estimated with the use of the aforementioned model [7] and its modification [8] are in good qualitative agreement with experimental data.

The one-dimensional turbulent model has shown the theoretical possibility of secondary flows. However, being too simple, it could not describe the experimentally observed real flow pattern. Investigation of the instability of a uniform flow on the basis of a two-dimensional model [3] made it possible to detect vortex secondary flows such as rotating stalls in high-pressure compressors.

In this paper, we numerically solve the electromagnetic field equations and Navier–Stokes equations for a two-dimensional model in order to obtain the secondary flow pattern in an induction EMP and to reveal the nature of the low-frequency pulsations of the parameters.

Along with the problem of MHD instability, which is of profound importance for the design of induction MHD machines with metal working substance, of undeniable theoretical interest is the investigation of complex flows of conducting liquids in a magnetic field.

1. CALCULATION MODEL

A schematic view of the model in the coordinates x and z is presented in Fig. 1. The liquid metal flows along the x-axis through a $2b$-high channel between

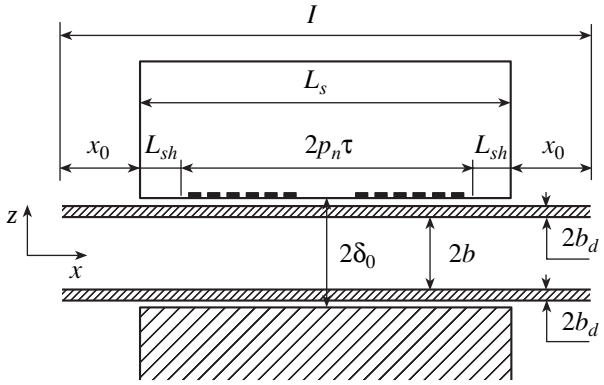


Fig. 1. The calculation domain scheme. I—calculation domain.

1063-7842/05/5008-0999$26.00 © 2005 Pleiades Publishing, Inc.

two ferromagnetic surfaces. On one of the surfaces, an inductor winding is mounted which produces an x-traveling magnetic field and is really located inside the inductor slots. The winding is replaced with thin current sheets; each current sheet carries the total current of a single slot and has the same x-size as the slot. The active (wound) zone of the inductor has the length $2p_n\tau$, where $2p_n$ is the number of poles of the inductor and τ is the pole pitch; L_{sh} is the length of the shunt zones of the inductor with due regard for the magnetic flux leakage at the ends of the inductor; and x_0 is the length of the calculation zone outside the inductor. The cylindrical geometry of the pump is replaced with a $2\pi R$-wide plane one, where R is the average radius of the channel. The error introduced by such a replacement is insignificant, because the average radius of the channel is usually much greater than its height $2b$. In this model, the external magnetic field and the field induced in the metal have only z-components $\mathbf{B}_0 = (0, 0, B_0(x, y))$ and $\mathbf{b} = (0, 0, b_z(x, y))$ depending on the longitudinal and azimuth coordinates and averaged with respect to the nonmagnetic gap height $2\delta_0$. The velocity components $U = (u, v, 0)$ and the pressure also depend on the longitudinal and azimuth coordinates and are averaged with respect to the channel height.

In this case, the equation for the magnetic field induced in the metal, which is derived from the Maxwell equation system, has the form

$$-\frac{R_{me}}{\omega}\frac{\partial(b_z + B_0)}{\partial t} + \frac{\partial^2 b_z}{\partial \bar{x}^2} + \frac{\partial^2 b_z}{\partial \bar{y}^2} - R'_m u^* \frac{\partial(b_z + B_0)}{\partial \bar{x}} - R'_m v^* \frac{\partial(b_z + B_0)}{\partial \bar{y}} = 0, \quad (1)$$

where $R_{me} = R'_m(1 + \sigma_k b_k/\sigma b)$; σ_k and σ are the conductivity of the channel walls and of the liquid metal; $2b_k$ and $2b$ are the thickness/height of the channel walls and of the liquid metal layer; $R'_m = R_{m0}(2b/2\delta')$; $R_{m0} = \mu_0\sigma\omega/\alpha^2$ is the magnetic Reynolds number, μ_0 is the magnetic permeability, $\omega = 2\pi f$, f is the current frequency in the winding; $\alpha = \pi/\tau$; $2\delta' = 2\delta_0 k_\delta k_\mu$ is the equivalent nonmagnetic gap, k_δ is the tooth-slot coefficient of the inductor, k_μ is the magnetic circuit saturation coefficient; and $u^* = u(\alpha/\omega)$, $v^* = v(\alpha/\omega)$, $\bar{x} = x\alpha$, and $\bar{y} = y\alpha$ are the dimensionless velocities and coordinates.

The boundary conditions for the induced field are

$$b_z|_{x=0} = 0, \quad b_z|_{x=\alpha L} = 0, \quad b_z|_{y=0} = b_z|_{y=2\pi R\alpha}, \quad \left.\frac{\partial b_z}{\partial y}\right|_{y=0} = \left.\frac{\partial b_z}{\partial y}\right|_{y=2\pi R\alpha}, \quad (2)$$

where $L = L_s + 2x_0$.

The induced field can be presented in the form $\dot{b}_z = b_z^{Re} + ib_z^{Im}$, while the external field is given as $\dot{B}_0 = \dot{B}_{0m}\exp(i\omega t)$, where $\dot{B}_{0m}(x, y)$ is the complex amplitude.

In this case, Eq. (1) takes the form

$$-\frac{R_{me}}{\omega}\frac{\partial(\dot{b}_z)}{\partial t} + \frac{\partial^2 \dot{b}_z}{\partial \bar{x}^2} + \frac{\partial^2 \dot{b}_z}{\partial \bar{y}^2} - R'_m u^* \frac{\partial(\dot{b}_z)}{\partial \bar{x}} - R'_m v^* \frac{\partial(\dot{b}_z)}{\partial \bar{y}} - iR_{me}\dot{B}_{0m}\exp(i\omega t) - R'_m u^* \frac{\partial(\dot{B}_{0m})}{\partial \bar{x}}\exp(i\omega t) - R'_m v^* \frac{\partial(\dot{B}_{0m})}{\partial \bar{y}}\exp(i\omega t) = 0. \quad (3)$$

The external field amplitude $\dot{B}_{0m}$ can be found from the equation

$$\mathrm{curl}\mathbf{B} = \mu_0 \mathbf{J}_0$$

as the superposition of the fields produced by the currents in the inductor slots. The field produced by a single slot current is calculated by the formula

$$B_{0n}(x) = \begin{cases} c_{1n}, & x \in [x_0, x_n] \\ c_{1n} + \dfrac{(c_{2n} - c_{1n})}{L_{slot}(x - x_n)}, & x \in [x_n, x_n + L_{slot}] \\ c_{2n}, & x \in [x_n + L_{slot}, x_0 + L_s], \end{cases}$$

where

$$c_{1n} = \frac{\mu_0 J_n\left(L_s - x_n - \dfrac{L_{slot}}{2}\right)}{L_s}, \quad c_{2n} = -\frac{\mu_0 J_n\left(x_n + \dfrac{L_{slot}}{2}\right)}{L_s}.$$

Here, $J_n = \sqrt{2}\, IU_n/2\delta'$ is the amplitude of the linear current density in the nth slot, U_n is the number of conductors in the nth slot, I is the effective current in the nth slot, x_n is the coordinate of the beginning of the nth slot, and L_{slot} is the x-width of the slot. Thus, the complex amplitude of the external field is determined by the formula

$$\dot{B}_{0m}(x) = \sum_{n=1}^{N} B_{0n}(x)\exp(i\varphi_n),$$

where φ_n is the current phase in the nth slot.

Equations (1) and (3) take into account the azimuth dependence of the external field, which is due to the constructive features of MHD machines,

$$\dot{B}_{0m}(x, y) = \dot{B}_{0m}(x) f(y).$$

There is a possibility of specifying the experimentally observed azimuth dependence of the external field, which is determined by the location of the induc-

tor packages. In the calculations, our concern was with two model types of the external field, namely, the azimuthally uniform one and the weakly nonuniform field with the azimuth dependence of the form

$$f(y) = 1 + 0.1\cos(y/R), \quad y \in [0, 2\pi R].$$

The induced field obtained from Eq. (3) and the corresponding induced current have components in the channel walls (I) and in the liquid metal (II). Calculation of the electromagnetic pressure requires only the current components in the liquid metal (II), whose complex amplitudes can be obtained by the following formulas:

$$j_x^{\mathrm{II}} = -\frac{\partial \dot{b}_z/\partial y 2\delta'/\mu_0 + 2b_d\sigma_d v(\dot{b}_z + \dot{B}_0)}{2b(1 + \sigma_d b_d/\sigma b)},$$

$$j_y^{\mathrm{II}} = -\frac{\partial \dot{b}_z/\partial x 2\delta'/\mu_0 + 2b_d\sigma_d u(\dot{b}_z + \dot{B}_0)}{2b(1 + \sigma_d b_d/\sigma b)}.$$

To determine the electromagnetic force components, one should take the product of the real parts of the currents' complex amplitudes and the magnetic fields:

$$f_x = \frac{1}{2}\{\mathrm{Re}\, j_y^{\mathrm{II}} \mathrm{Re}(\dot{b}_z + \dot{B}_0)\},$$

$$f_y = -\frac{1}{2}\{\mathrm{Re}\, j_x^{\mathrm{II}} \mathrm{Re}(\dot{b}_z + \dot{B}_0)\}.$$

In the general case, the two-dimensional system of Navier–Stokes equations for turbulent flow conditions contains pulsating terms. Applying the Reynolds averaging procedure, expressing the fluctuating velocities in terms of the averaged velocities, and averaging the equations with respect to the channel height, we put the system of Navier–Stokes equations in the following form:

$$\begin{aligned}
\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} &= -\frac{1}{\rho}\frac{\partial p}{\partial x} + \nu_{\mathrm{eff}}\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) \\
&\quad + \frac{f_x}{\rho} - \frac{c_f}{2b}u\sqrt{u^2 + v^2}, \\
\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} &= -\frac{1}{\rho}\frac{\partial p}{\partial x} + \nu_{\mathrm{eff}}\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right) \\
&\quad + \frac{f_y}{\rho} - \frac{c_f}{2b}v\sqrt{u^2 + v^2},
\end{aligned} \tag{4}$$

where $u(x, y)$ and $v(x, y)$ are the longitudinal and azimuth averaged velocity components, p is the pressure, ν_{eff} is the viscosity effective coefficient (the sum of the laminar ν and turbulent ν_t viscosity coefficients), ρ is the liquid metal density, and the last term is the channel wall friction (c_f is the local coefficient of friction).

The third component of the velocity is set equal to zero, while the components u and v are related by the continuity equation

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0. \tag{5}$$

The boundary conditions for the velocities have the form

$$u|_{x=0} = u_0(y), \quad u|_{x=L} = u_s, \quad u|_{y=0} = u|_{y=2\pi R},$$

$$\left.\frac{\partial u}{\partial y}\right|_{y=0} = \left.\frac{\partial u}{\partial y}\right|_{y=2\pi R},$$

$$v|_{x=0} = 0, \quad v|_{x=L} = 0, \quad v|_{y=0} = v|_{y=2\pi R},$$

$$\left.\frac{\partial v}{\partial y}\right|_{y=0} = \left.\frac{\partial v}{\partial y}\right|_{y=2\pi R}.$$

Similarly to the case of an applied magnetic field, velocity initial profiles of two types were considered, namely, the uniform one $u_0 = u_s$ (u_s is the average flow rate) and the nonuniform profile of the form $u_0 = u_s(1 + 0.1\cos(y/R))$.

With the use of Eqs. (4) and (5), one can obtain the Poisson equation for the pressure. Thus, we arrive at a closed equation system composed of the system of equations for the real and imaginary components of induced field (3), Navier–Stokes system (4), and the Poisson equation. This equation system was numerically solved by the finite-difference method on a biased uniform mesh (velocity calculation points were displaced from the points at which the pressure and magnetic field were calculated). Equations (3) and (4) were split by the method of stabilizing correction. Convective terms were approximated by upstream differences. The stationary Poisson equation was solved by the successive overrelaxation method.

At each time step, calculations include the following operations: (1) determining the components of the field induced in the metal for the initial values of the velocity components and calculating the electromagnetic force instantaneous components; (2) solving the system of Navier–Stokes equations, i.e., finding the provisional values of the velocities u' and v'; (3) solving the Poisson equation for the pressure correction by the iterative method; and (4) refining the values of the velocities for their use at the next step.

Our model involves the Poisson equation for the pressure correction p' [9] of the form

$$\frac{\partial^2 p'}{\partial x^2} + \frac{\partial^2 p'}{\partial y^2} = \frac{\rho}{\Delta t}\left(\frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y}\right), \tag{6}$$

where Δt is the time step of the calculation scheme, with the boundary conditions

$$\left.\frac{\partial p'}{\partial y}\right|_{x=0} = 0, \quad \left.\frac{\partial p'}{\partial x}\right|_{x=L} = 0, \quad p'|_{y=0} = p'|_{y=2\pi R},$$

$$\left.\frac{\partial p'}{\partial y}\right|_{y=0} = \left.\frac{\partial p'}{\partial y}\right|_{y=2\pi R}.$$

Here, the pressure at the current time step is the sum of the pressure at the previous step and the correction

$$p^n = p^{n-1} + p'.$$

At the current time step, the calculation terminates with the refinement of the velocities obtained from the system of Navier–Stokes equations,

$$u = u' - \frac{\Delta t}{\rho}\frac{\partial p'}{\partial x},$$

$$v = v' - \frac{\Delta t}{\rho}\frac{\partial p'}{\partial y}.$$

In this model, the pressure Δp developed by the pump means the channel width average pressure drop between the lengthwise first and last mesh nodes.

The choice of the initial conditions has a little effect on the final results and determines only the value of the initial/nonstationary period of flow development. The essential feature required of the initial conditions is their physical adequacy. In this model, we use the following initial conditions: at $t = 0$, the induced field is zero and, for all x, the velocities are set equal to the velocity at the input (when $x = 0$).

2. INITIAL DATA FOR THE CALCULATION

The above mathematical model (EMP-MHD2D) was employed for analyzing the electromagnetic and hydrodynamic characteristics of the liquid metal (sodium) flow in an electromagnetic pump with a cylindrical linear channel (CLIP-B). The flow MHD instability present in this pump under certain values of $R'_m s$, which is manifested in azimuthally nonuniform distribution of the magnetic field and liquid metal velocity and in the low-frequency pulsations of the pressure, is experimentally observed [6]. Although the liquid metal velocity is not directly measured in the experiment, the length and azimuth distribution of its x-component can be inferred by the distribution of the resulting magnetic field, as has been shown in [5].

In the calculations, the main parameters of the pump are taken as follows: the inductor length is $L_s = 0.85$ m, the number of poles is $2p_n = 6$, the pole pitch is $\tau = 0.1398$ m, the inner tooth length is $L_{\text{tooth }1} = 6.3$ mm, the extreme tooth length is $L_{\text{tooth }2} = 8$ mm, the slot length is $L_{\text{slot}} = 17$ mm, the channel average diameter is $D = 0.296$ m, the total height of the channel walls is $2b_d = 4.5$ mm, the channel height is $2b = 12$ mm, the nonmagnetic gap height is $2\delta_0 = 20$ mm, the metal conductivity is $\sigma = 6.6 \times 10^6\ (\Omega\ \text{m})^{-1}$, the conductivity of the channel walls is $\sigma_d = 1.15 \times 10^6\ (\Omega\ \text{m})^{-1}$, the supply voltage frequency is $f = 50$ Hz, and the magnetic Reynolds number $R_{m0} = 5.44$.

3. RESULTS

Some parameters of the model must be predetermined. Among them is the length of the shunt zone, which is made up of the real zone and the equivalent one, the latter taking into account the magnetic field attenuation outside the inductor. The length of the equivalent shunt zone is here set equal to $2\delta'$, which corresponds to the exponential attenuation of the magnetic field at the ends of the inductor (see [8] for details on the choice of L_{sh}). Outside the inductor, the zones must be sufficiently extended for the boundary conditions to exert little influence upon the flow parameters in the inductor region. After a number of optimizing estimations, in the calculations, we took the value $x_0 = 4\tau$.

The value of the local coefficient of friction c_f for turbulent flow conditions in Eq. (4), which is a function of the velocity, could not be exactly determined. For this reason, c_f was set equal to the resistance coefficient λ of a plain pipe under turbulent flow conditions. There is a number of semiempirical dependences for λ (logarithmic resistance law). In the calculations, we used the formula [10]

$$c_f = 1/4/(1.81\log(|V|4b/\nu) - 1.64)^2,$$

where $|V|$ is the local velocity.

The value of the turbulent viscosity coefficient ν_t in Eq. (4) is several orders greater than the laminar viscosity ν [11] and, in the general case, is a function of the velocity coordinate derivative. In our calculations, we took the viscosity effective coefficient to be constant over the whole domain and two orders greater than ν. The impact of the turbulent viscosity coefficient on the calculation results is to be studied in more detail.

Calculations with the model described above demonstrate that, after a certain nonstationary period, two regimes of the liquid metal steady flow are possible. For $R'_m s$ below some critical value $(R'_m s)_{\text{cr}} \approx 1.5$, the stationary velocity field is realized. The velocity field is completely uniform ($v = 0$ over the whole calculation domain) if the external magnetic field and the initial velocity are free of nonuniformities; it is weakly nonuniform if they are not. The nonuniformity degree of the velocity field has the same order as that of the external magnetic field or the velocity initial profile. Therefor, along with the time-independent component, the pressure developed by the pump contains a component that pulsates at twice the supply voltage frequency $2f$.

For R_m exceeding the critical value, the flow regime is also completely uniform within the framework of the present model in the absence of external nonuniformity. However, nonuniformity of the external magnetic field and/or the initial velocity brings about a strongly nonuniform flow regime with the formation of large-scale vortices in the inductor region which travel from the pump input to the output and move azimuthally. An example of the vortex evolution is given in Fig. 2. The

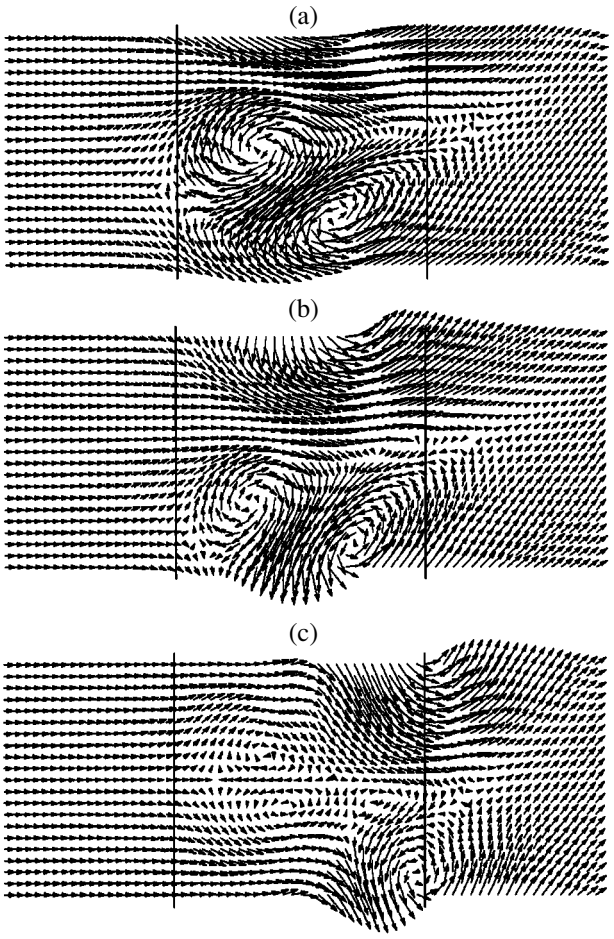


Fig. 2. Velocity field development for a strongly nonuniform flow regime, $s = 0.782$. The channel input is on the left. The vertical lines mark the inductor ends. Time instant: (a) 0, (b) 0.1 and (c) 0.2 s.

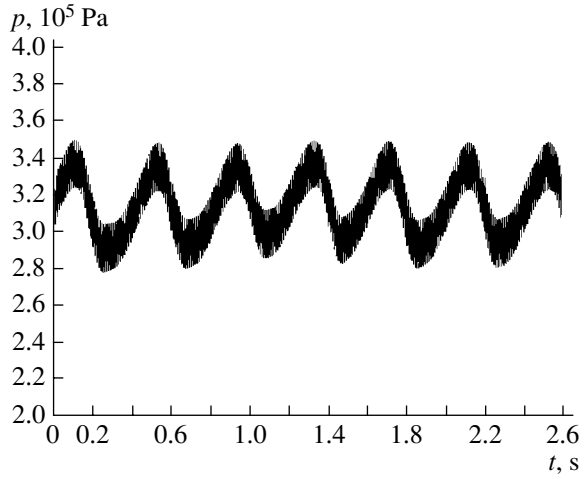


Fig. 3. The pressure developed by the pump, $s = 0.782$.

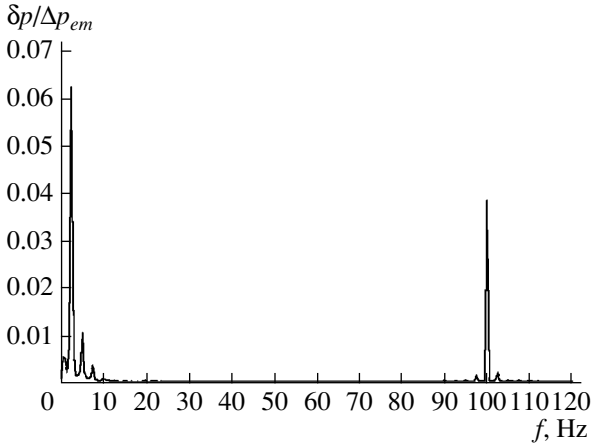


Fig. 4. The frequency spectrum of the pressure developed by the pump, $s = 0.782$.

vortices are formed in the inductor region at the pole pitch distance from the input and travel to the inductor output at the longitudinal velocity close to the average flow rate. The azimuth displacement of the vortices increases as they approach the inductor output, where the vortex azimuth velocity can exceed its longitudinal velocity. The larger R_m, the more developed is the flow nonuniformity exhibited by the regime. This is manifested in the fact that the vortex azimuth velocities become comparable to the longitudinal velocities right after the vortices are formed. In such a flow regime, the azimuth displacements of the vortices attain the value of 360°. In a nonuniform flow regime, the pressure developed by the pump exhibits both the high-frequency and low-frequency pulsations (see Fig. 3). For a strongly nonuniform regime, the frequency spectrum of the pressure is presented in Fig. 4. It qualitatively tallies with the one obtained experimentally in [6]. The flow pattern is also in good agreement with the calculation results in [6], though the mathematical models are essentially different.

The most significant low frequency of the pressure pulsation corresponds to the vortex longitudinal displacement time determined by the average flow rate ($f_l = u_s/L_{vor}$, where L_{vor} is the area occupied by the vortices). The calculation data presented in Table 1 confirm that this estimate for f_l is sufficiently accurate (for

Table 1. The value of the basic frequency of the pressure pulsation (in the low-frequency region) for two flow regimes

Slip, s	Average flow rate u_s, m/s	Basic low frequency of the pressure pulsation, Hz		
		experiment	$f_l = u_s/L_{vor}$	calculated by the two-dimensional model
0.715	3.99	3.6	4	4
0.782	3.05	2.9	3	2.5

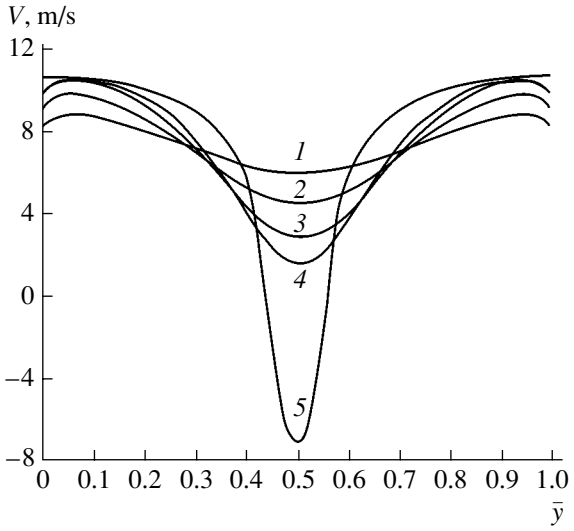


Fig. 5. Instantaneous profiles of the liquid metal velocity, $s = 0.466$. *1*, Quarter the inductor length; *2*, half the inductor length; *3*, three quarters of the inductor; *4*, end of the inductor; and *5*, calculated by the one-dimensional jet model.

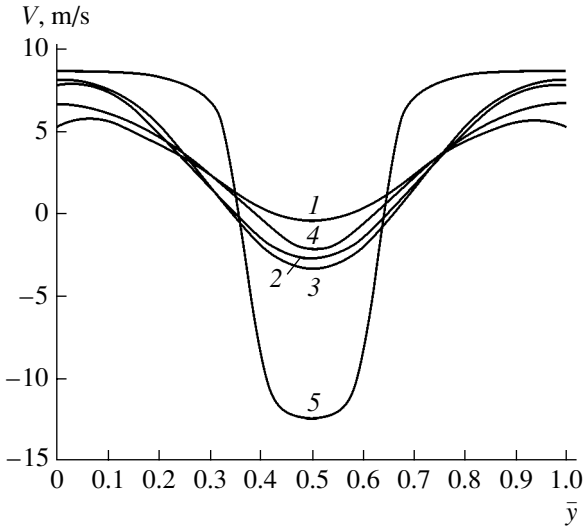


Fig. 6. Time averaged profiles of the liquid metal velocity, $s = 0.782$. (*1*–*5*) The same as in Fig. 5.

CLIP-B, $L_{vor} \approx 1$ m and slightly exceeds the inductor length L_s). An increase in the slip, i.e., a decrease in the average flow rate, involves a decrease in the basic frequency of the pressure pulsations. The subsidiary peaks in the low-frequency region observed in Fig. 4 along with the basic low frequency seem to be attributed to the azimuth displacement of the vortices.

Figures 5 and 6 show the time averaged profiles of the liquid metal velocity calculated by the present model at various points along the channel. For reference, the calculation by the one-dimensional jet model (EMP2D-JET [8]) is also presented. For the two-dimensional model, in the regime $s = 0.466$ ($R'_m s = 1.4$), the averaged velocities coincide with the instantaneous velocities. For $s = 0.782$, the averaging period was set equal to four periods of the pressure low-frequency pulsations in this regime. Calculation shows that, in the regime $s = 0.466$, the nonuniformity degree of the velocity profile uniformly increases from the inductor input to the output. In the regime $s = 0.782$ ($R'_m s = 2.3$), the velocity profile nonuniformity is stronger; it increases up to approximately 3/4 of the inductor length and then reduces to some extent. The velocity values calculated by the two-dimensional model are physically adequate, especially as compared to the one-dimensional model, which gives overvalued velocities in the reversed-flow region. Both models have a reversed-flow region of approximately the same width. However, for the two-dimensional model, the reverse-flow region appears at higher slip values than for the one-dimensional model.

The results of calculating the time averaged velocity profiles by the two-dimensional model are in good qualitative agreement with the measured distribution of the resulting magnetic field [6], which can be considered as an indicator of the liquid metal velocity distribution, as has been shown in [5].

The completely two-dimensional model EMP-MHD2D demonstrates the following quantitative results. The pressure developed by the pump (the time-independent component) is 3–9% above the experimental values and is much closer to the results of calculations by the one-dimensional model EMP2D-JET (see Table 2).

Table 2. Pressure developed by the pump

Slip, s	Δp, 10^5 Pa		
	experiment	calculated by the one-dimensional model	calculated by the two-dimensional model
0.466	2.12	2.42	2.19
0.715	2.78	2.99	2.94
0.782	2.87	3.14	3.12

Table 3. Ratio between the pressure pulsation amplitudes of the basic low frequency and the electromagnetic pressure

Slip, s	$p_{low}/\Delta p_{em}$	
	experiment	calculated by the two-dimensional model
0.715	0.05	0.012
0.782	0.052	0.059

Table 4. Ratio between the pressure pulsation amplitudes of the double frequency and the electromagnetic pressure

Slip, s	$p_{2f}/\Delta p_{em}$		
	experiment	calculated by the one-dimensional model	calculated by the two-dimensional model
0.466	0.026	0.125	0.076
0.715	0.015	0.127	0.050
0.782	0.013	0.123	0.036

The amplitudes of the basic low frequency p_{low} differ from the experimental data. This is easy to explain, since the model nonuniformity of the external magnetic field taken in the calculations differs from the real one (see Table 3). The one-dimensional model does not allow calculation of the pressure low-frequency pulsations.

The pressure pulsation amplitudes of the double frequency p_{2f} calculated by the two-dimensional model are three to four times higher than the experimental data (see Table 4). However, the two-dimensional model provides much more accurate calculation results than the one-dimensional jet model.

The pump electromagnetic pressure Δp_{em} is calculated by integrating the x-component of the electromagnetic force over the volume of the calculation domain.

CONCLUSIONS

The flow pattern of a liquid metal in an electromagnetic pump with a cylindrical channel under the MHD instability is obtained on the basis of a completely two-dimensional mathematical model. This flow is characterized by the presence of two-dimensional vortices which travel from the pump input to the output and move azimuthally. These vortices give rise to low-frequency pulsations of the pump pressure.

The calculation analysis of the flow regimes in the CLIP-B pump demonstrates that the weakly nonuniform regime turns into the strongly nonuniform one at the value $R'_m s \approx 1.5$. In the experiment, the nonuniform flow regime is attained at $R'_m s \approx 1.3$–1.4, which is evidenced by the substantially increasing amplitudes of the pressure low-frequency pulsation.

Most of the calculated characteristics except for the pressure pulsation amplitudes are in good qualitative and satisfactory quantitative agreement with the experiment.

REFERENCES

1. Yu. A. Polovko, E. P. Romanova, and E. A. Tropp, Zh. Tekh. Fiz. **66** (4), 36 (1996) [Tech. Phys. **41**, 315 (1996)].
2. Yu. A. Polovko, E. P. Romanova, and E. A. Tropp, Zh. Tekh. Fiz. **67** (6), 5 (1997) [Tech. Phys. **42**, 591 (1997)].
3. Yu. A. Polovko, E. P. Romanova, and E. A. Tropp, in *Proceedings of the 3rd International Conference on Transfer Phenomena in Magnetohydrodynamic and Electroconducting Flows, Aussois, 1997,* Vol. 2, pp. 451–456.
4. I. R. Kirillov *et al.*, Magn. Gidrodin., No. 2, 107 (1980).
5. I. R. Kirillov and V. P. Ostapenko, Magn. Gidrodin., No. 2, 95 (1987).
6. H. Araseki *et al.*, Nucl. Eng. Des. **227**, 29 (2004).
7. R. A. Valdamane *et al.*, Magn. Gidrodin., No. 3, 98 (1982).
8. I. R. Kirillov and D. M. Obukhov, Energy Convers. Manage. **44**, 2687 (2003).
9. D. A. Anderson, J. C. Tannehill, and R. H. Pletcher, *Computational Fluid Mechanics and Heat Transfer* (Hemisphere, New York, 1984; Mir, Moscow, 1990), Vol. 2.
10. I. E. Idel'chik, *Handbook of Hydraulic Resistance* (Gosénergoizdat, Moscow, 1960; Hemisphere, Washington, 1986).
11. L. G. Loitsyanskii, *Mechanics of Liquids and Gases* (Nauka, Moscow, 1978; Elsevier, Amsterdam, 1966).

Translated by A. Pankrat'ev