Technical Physics, Vol. 47, No. 7, 2002, pp. 795–802. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 1–8.

THEORETICAL AND MATHEMATICAL PHYSICS

Complex Dynamics of a Simple Distributed Self-Oscillatory Model System with Delay

N. M. Ryskin and A. M. Shigaev

Saratov State University, ul. Moskovskaya 155, Saratov, 410026 Russia

e-mail: RyskinNM@info.sgu.ru

Received July 6, 2001

Abstract—A simple model for a distributed self-oscillatory system with cubic nonlinearity and delay is presented. Conditions for oscillation self-excitation and stationary oscillation conditions, as well as the stability of the oscillations, are analyzed. Nonstationary self-modulation regimes (including conditions of complex dynamics and chaos) are simulated numerically over a wide range of control parameters. As the factor of nonequilibrium grows, regular and chaotic regimes alternate in a complex manner. The transitions to chaos may follow all scenarios known for finite-dimensional systems. The model suggested is somewhat akin to a number of earlier finite-dimensional models aimed at studying mode competition in resonance electron masers. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Distributed self-oscillatory systems (DSOSs) with delayed feedback (DF) are of great importance in radiophysics, nonlinear optics, biophysics, etc. [1–4]. It is well known that such systems can generate complex, including chaotic, oscillations. Therefore, their study may provide a better insight into the dynamics of systems with a large number of degrees of freedom. Specifically, the simulation of some properties of fully developed turbulence by means of self-excited oscillators with delay has been discussed [2, 5]. Note also that this problem is of great practical significance, since self-excited oscillators with DF are the basis for many devices of vacuum and plasma microwave electronics. Typical examples are traveling-wave tubes (TWTs) with DF [6–8], DF klystrons [9–12], free-electron lasers (FELs) [13–15], electron linacs with external feedback [16], etc. In backward-wave oscillators (BWOs), the self-modulation and complex dynamics are also associated with delayed internal feedback. Note that recent attempts to develop controllable sources of noise-like signals have sharpened interest in chaotic oscillation conditions.

The complex dynamics of distributed systems is extremely difficult to study because of an infinitely large number of degrees of freedom. The presence of several control parameters is an additional aggravating factor. It would be therefore tempting to consider a simple model system that demonstrates all dynamics features of DSOSs with DF and still can be investigated numerically and, if possible, analytically. For this purpose, we suggest the equation

$$\dot{A} + \gamma A = \alpha \exp(i\Theta)(1 - |A(t-1)|^2)A(t-1). \quad (1)$$

Here, A is the slowly varying complex amplitude of oscillations; α and γ are the factors of nonequilibrium and dissipation, respectively; and Θ is the phase shift in the feedback loop. The right-hand side of Eq. (1) depends on the amplitude value at the delayed time instant $t - 1$ (the variables can always be normalized so that the delay time is unity). Equation (1) describes, for example, the dynamics of a van der Pol triode oscillator with a delay line in the anode circuit in the case when the grid–plate characteristic of the oscillator is approximated by a cubic polynomial [17]. Without delay, Eq. (1) becomes the reduced van der Pol equation

$$\dot{A} + \gamma A = \alpha \exp(i\Theta)(1 - |A|^2)A,$$

which describes the conditions of periodic one-frequency self-oscillations near the self-excitation threshold for a wide class of self-oscillatory systems [1, 4, 18]. It is therefore natural to use Eq. (1) as a universal model describing the dynamics of self-oscillatory systems with DF. Note that the equation

$$\dot{A} + \gamma A = \alpha \exp(i\Theta) J_1(|A(t-1)|)\frac{A(t-1)}{|A(t-1)|}$$

(J_1 is the first-order Bessel function of the first kind), which describes the operation of a self-excited DF klystron oscillator [9–11], is also reduced to form (1) in the case of weak nonlinearity, where one can expand the function J_1 into a series and leave the first two terms.

Different problems associated with the complex dynamics of a self-excited oscillator with cubic nonlinearity and delay have been considered in [19–21]. It has been shown [19] that the stationary solution becomes unstable and self-modulation appears as the factor of nonequilibrium grows. With a further increase in α, the

1063-7842/02/4707-0795$22.00

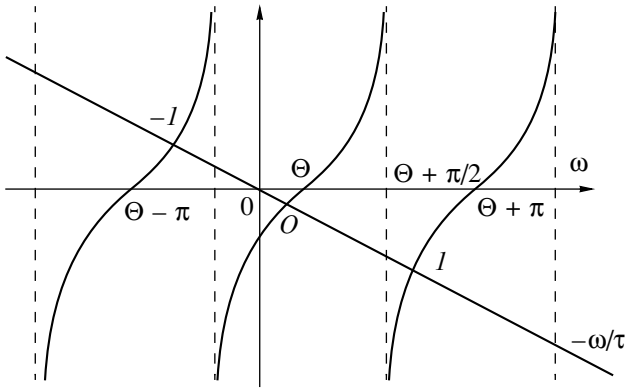


Fig. 1. On determination of the eigenmode frequencies of an oscillator with delay.

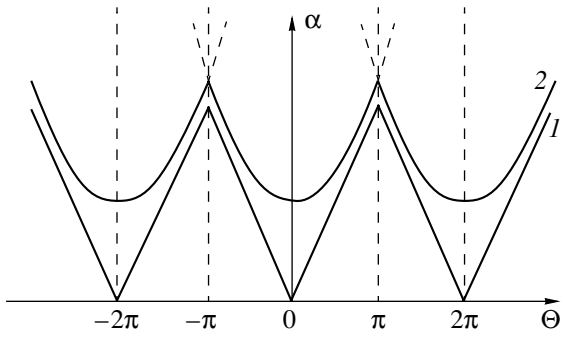


Fig. 2. Boundaries of the self-excitation regions on the parameter plane (α, Θ). γ = 0 (*1*) and >0 (*2*).

situation becomes chaotic, following the Feigenbaum scenario; that is, the system experiences a series of self-modulation period doubling bifurcations [20]. Also, it has been found [21] that so-called multimode chaos sets in the system far above the threshold of the chaotic regime. This is because the attractors related to different modes combine.

In this work, we study at length the reduced version of van der Pol equation with delay (1) over a wide range of control parameters. Conditions for the self-excitation of oscillations are analyzed, solutions meeting stationary oscillation conditions are found, and stability conditions of the oscillations are determined. Since this article elaborates upon works [4, 17, 19], we will present only those results uncovered previously.

1. THEORETICAL ANALYSIS

1.1. Conditions for oscillation self-excitation. Linearizing Eq. (1) and finding a solution in the form $A \sim \exp(pt)$, we come to the characteristic equation

$$p + \gamma = \alpha\exp(i\Theta - p). \qquad (2)$$

Transcendent equation (2) has an infinite number of complex roots because the system is distributed. Taking into account that p is purely imaginary ($p = i\omega$, where ω is the oscillation frequency) near the instability threshold, we find from (2)

$$\gamma = \alpha\cos(\omega - \Theta), \quad \omega = -\alpha\sin(\omega - \Theta). \qquad (3)$$

Eliminating α from Eqs. (3), one can find the equation for oscillation frequencies:

$$\frac{\omega}{\gamma} = -\tan(\omega - \Theta). \qquad (4)$$

Equation (4) can be conveniently solved graphically. Let its roots be numbered as shown in Fig. 1. From the physical meaning of the problem, the parameters α and γ are considered to be positive; hence, only those for which $\cos(\omega - \Theta) > 0$ (i.e., even-number roots) should be left. The eigenmode frequencies depend on γ as follows (Fig. 1). For $\gamma \longrightarrow 0$,

$$\omega_{\pm n} \approx \Theta \pm \left(\pi n - \frac{\pi}{2}\right),$$

where $n = 2, 4, \ldots$.

For the fundamental mode ($n = 0$), $\omega_0 \longrightarrow 0$ regardless of the value of Θ. At $\gamma \longrightarrow \infty$, $\omega_{\pm n} \approx \Theta \pm \pi n$.

One can also find the starting value of the parameter α at which the nth mode is excited ($\alpha_{st}^{(n)}$). Squaring both equations in (3) and adding them yields

$$\alpha_{st}^{(n)} = \sqrt{\gamma^2 + \omega_n^2(\gamma, \Theta)}, \qquad (5)$$

where $\omega_n(\gamma, \Theta)$ is the nth-mode frequency defined by (4).

It follows that the value $\Theta = 0$ is optimal for the excitation of the fundamental mode $n = 0$. Then, $\omega_0 = 0$ and $\alpha_0 = \gamma$. For $\Theta = \pm\pi$, self-excitation is difficult. In this case, the frequencies ω_0 and $\omega_{\mp 2}$ become roughly equal. As Θ increases further, the second mode becomes fundamental. Thus, the self-excitation boundary in the plane (α, Θ) has the form depicted in Fig. 2. The regions $(2m - 1)\pi < \Theta < (2m + 1)\pi$ will be called oscillation generation regions. For $\Theta = 2\pi m$, the parameter α takes the minimal value; for $\Theta = 2\pi m + \pi$, the maximal value.

1.2. Stationary oscillation conditions. Under stationary oscillation conditions,

$$A = A_0\exp[i(\omega t + \varphi)], \qquad (6)$$

where the amplitude A_0 and the phase φ are constants.

Substituting (6) into (1) and separating the real and imaginary parts, we have

$$\begin{aligned} \gamma &= \alpha(1 - A_0^2)\cos(\omega - \Theta), \\ \omega &= -\alpha(1 - A_0^2)\sin(\omega - \Theta). \end{aligned} \qquad (7)$$

Dividing one of the relationship in (7) by the other, we arrive at Eq. (4). This means that the stationary oscillation frequency equals one of the natural frequen-

cies ω_n and does not vary with oscillation amplitude. From (7), one also readily finds the amplitude A_0:

$$A_0 = \sqrt{1 - \frac{(\gamma^2 + \omega_n^2)^{1/2}}{\alpha}} = \sqrt{1 - \frac{\alpha_{st}^{(n)}}{\alpha}}. \quad (8)$$

It is seen that if the supercriticality is weak, i.e., when $\alpha = \alpha_{st}^{(n)}(1 + \varepsilon)$, where ε is small, the amplitude grows by the law $A_0 \sim \sqrt{\varepsilon}$, which is typical of Andronov–Hopf supercritical bifurcation (i.e., of the soft excitation of self-oscillations). If α greatly exceeds its starting value, A_0 tends to unity.

1.3. Conditions for self-modulation. Let us test the solutions for stability. To do this, we set a small disturbance of the stationary regime as a pair of side-band (satellite) modes equidistant from the fundamental frequency:

$$A = [A_0 + A_+(t)\exp(i\Omega t) + A_-(t)\exp(-i\Omega t)] \times \exp[i(\omega_n t + \varphi)]. \quad (9)$$

Here, A_0 satisfies Eq. (8) ($|A_\pm| \ll |A_0|$), ω_n is one of the natural frequencies, Ω is the self-modulation frequency, and φ is an arbitrary phase. One should substitute Eq. (9) into starting equation (1), linearize the resulting expression in terms of small disturbances, and analyze the behavior of the satellite amplitudes $A_\pm$. We will consider only the case $\Theta = 0$ (the center of the oscillation generation region), for which an analytical solution in explicit form can be found. At the stability boundary, we have $A_\pm = \text{const}$; that is, the small disturbances neither grow nor decay. After rearrangement, we obtain the following equations:

$$\begin{aligned} [\gamma + i\Omega + (\alpha - 2\gamma)\exp(-i\Omega)]A_+ &= -(\alpha - \gamma)A_-^*\exp(-i\Omega), \\ [\gamma - i\Omega + (\alpha - 2\gamma)\exp(i\Omega)]A_- &= -(\alpha - \gamma)A_+^*\exp(i\Omega). \end{aligned} \quad (10)$$

System (10) has two kinds of solutions: $A_+ = A_-^*$ or $A_+ = -A_-^*$. It can be shown that the former corresponds to purely amplitude disturbances; the latter, to purely phase ones [13]. Thus, the amplitude and phase disturbances split in this case and can be analyzed separately.

Consider the amplitude disturbances. The characteristic equation for them takes the form

$$\gamma + i\Omega = -(2\alpha - 3\gamma)\exp(-i\Omega). \quad (11)$$

Separating the real and imaginary parts in (11), we find

$$\gamma = -(2\alpha - 3\gamma)\cos\Omega, \quad \Omega = (2\alpha - 3\gamma)\sin\Omega. \quad (12)$$

It can be shown that the self-modulation frequency Ω also satisfies Eq. (4) for the eigenmode frequencies; now, however, one has to take odd-number roots. Recall that we consider the case $\Theta = 0$ and the parameters α and γ are assumed to be positive ($\alpha > \gamma$, as follows from the self-excitation conditions; see Subsection 1.1). Thus, the following conditions must be met:

$$2\alpha - 3\gamma > 0, \quad (13)$$

$$\cos\Omega < 0. \quad (14)$$

From (14), it follows that odd roots of (4) should be taken (see Fig. 1).

Since our system represents an amplifier with an external DF loop, it is convenient to recast inequality (13) in view of (8) in the form

$$1 - 3A_0^2 = dF/dA_0 < 0,$$

where $F(A_0) = A_0 - A_0^3$ is the amplitude characteristic of the amplifier.

Thus, the physical meaning of condition (13) is that the stationary amplitude must lie in the descending branch of the amplitude characteristic.

From system (12), one can also find the critical value of α, $\alpha_{sm}^{(n)}$, at which the instability arises:

$$\alpha_{sm}^{(n)} = \frac{1}{2}(3\gamma + \sqrt{\gamma^2 + \Omega_n^2}). \quad (15)$$

Here, Ω_n is one root of Eq. (4). Note that $n = \pm 1, \pm 3, \ldots$ in this case. Side-band modes that are the closest to the fundamental frequency have the lowest instability threshold. Since they are excited only when the fundamental mode is excited with a sufficiently large amplitude, the associated disturbances will be referred to as self-modulation modes.

Virtually, condition (15) means that the slope of the descending branch must be sufficiently large. The self-modulation arises in a soft manner through the appearance of a pair of satellites equidistant from the fundamental frequency. Their frequencies are $\omega_0 \pm \Omega_n$. In the theory of distributed self-oscillatory systems (see, e.g., [8]), such a self-modulation mechanism is called the amplitude mechanism.

We will not perform in-depth analysis of phase disturbances. Note only that the stationary conditions are stable against a constant phase shift; that is, when the disturbance neither grows nor decays. This is an obvious consequence of the substitution $A \longrightarrow A\exp(i\varphi_0)$, where φ_0 is an arbitrary constant.

In the general case $\Theta \neq 0$, the equations for the amplitude and phase disturbances do not split and the instability causes both amplitude and phase oscillations. Exact analytical conditions for self-modulation are impossible to derive. However, on the whole, the results obtained for the specific case $\Theta = 0$ remain valid. The self-modulation arises by the amplitude mechanism and is accompanied by the soft excitation of a pair of adjacent self-modulation modes. The self-modulation frequency varies with Θ insignificantly. This is

supported by the results of numerical simulation which follow.

2. RESULTS OF NUMERICAL SIMULATION

The dynamics of the system described by the reduced version of van der Pol equation with delay (1) was numerically simulated with the fourth-order Runge–Kutta method. The data were processed by constructing temporal realizations and bifurcation diagrams, reconstructing the attractors by the Packard–Takens technique, and calculating the Fourier spectra. The dissipation factor γ was taken to be rather small ($\gamma < 1$), because the self-modulation oscillations are quasi-harmonic in this case. At large γ, they are of a relaxation character and the validity of the approximation of slowly varying amplitudes becomes questionable. Looking ahead, we note that the self-excitation threshold, the amplitude and frequency of stationary oscillations, and the threshold and frequency of self-modulation obtained by the numerical simulation totally agree with the theoretical predictions (Subsection 1.1).

We studied in detail the transition to chaos over a wide range of control parameters. It turned out that the dynamics of the system exhibits a much wider variety of forms than was earlier surmised [20, 21]. For instance, the Feigenbaum scenario is observed only at the center of the oscillation generation region. A typical sequence of bifurcations for this case is illustrated in Fig. 3, which shows temporal realizations, phase portraits, and spectra for different values of the factor of nonequilibrium α ($\gamma = 0.3$, $\Theta = 0.1\pi$). For the interval $\alpha_{st}^{(0)} < \alpha < \alpha_{sm}^{(1)}$, the transient ends with the establishment of the single-frequency operation (Fig. 3a). Here, $\alpha_{st}^{(0)}$ is the threshold for the fundamental mode and $\alpha_{sm}^{(1)}$ is the threshold for the excitation of the first self-modulation mode (see Section 1). For $\alpha > \alpha_{sm}^{(1)}$, this regime looses stability against the first self-modulation mode and the periodic self-modulation regime (Fig. 3b) to which the limiting cycle in the phase space corresponds sets in. The phase trajectory segments associated with the transient are left in the phase portraits (Figs. 3a, 3b) for clarity.

First, the self-modulation is quasi-harmonic. As α grows, the self-modulation amplitude and the number of harmonics in the spectrum rise. Simultaneously, the shape of the limiting cycle becomes distorted: the cycle increases in size, approaching the saddle equilibrium state at the origin. This state repulses the phase trajectory, which is wound on the unstable saddle manifold. Because of this, loops arise in the limiting cycle (Fig. 3c). A similar situation was described in [11]. Accordingly, the temporal realization of the output signal becomes more complicated: several local maxima appear during one oscillation period. In addition, the self-modulation frequency diminishes, since the phase trajectory spends more time near the unstable equilibrium. There may appear a large number of variously shaped cycles (i.e., cycles with a different number of loops) depending on the parameter values. For $\gamma \ll 1$, the transitions from one cycle to another are soft. With an increase in γ, hard transitions, which exhibit hysteresis, start to prevail. Such behavior is typical of many self-oscillatory systems with delay (for example, of that considered in [22]), as well as of microwave oscillators based on a drift klystron, which were investigated theoretically and experimentally in [12].

As α increases, the transition to chaos proceeds through a series of self-modulation period doubling bifurcations. In this case, universal quantitative rules characteristic of the Feigenbaum scenario [1, 18] are corroborated. Figure 3d shows the temporal realization, spectrum, and phase portrait under the chaotic oscillation conditions. The general picture is well illustrated by the bifurcation diagram, which demonstrates the positions of the amplitude maxima at various values of α (Fig. 4). A hard transition between two cycles at $\alpha \approx 2.33$, regions of period doubling bifurcations, and windows of periodic motion in the chaos are clearly seen. As α grows further, the regular and chaotic conditions alternate, which is a characteristic property of such systems [11, 12]. The transitions to chaos proceed either through period doubling or softly.

Near the boundaries of the oscillation generation region, the dynamics is much more complex. Here, one of the two adjacent modes may be excited depending on the initial conditions; in other words, we observe a typical case of mode competition at strong coupling [18]. For either mode, the transition to chaos follows a particular scenario. By way of example, let us consider the case when Θ is close to π, so that modes with $n = 0$ (fundamental) and $n = 2$ (side-band) can be excited (see Subsection 1.1). The bifurcation diagrams for both modes at $\gamma = 0.3$ and $\Theta = 0.9\pi$ are depicted in Figs. 5a and 5b. As follows from the results of numerical simulation, the Feigenbaum scenario is first realized, but then ($\alpha = 3.722$) the hard transition to the periodic self-modulation conditions takes place. Subsequently, the self-modulation becomes quasi-periodic ($\alpha = 3.76$) and one more transition to chaos, which breaks the three-frequency quasi-periodic motion (Ruelle–Takens scenario), is observed. Near the oscillation generation region, the doubling cascade terminates and the hard transition to the periodic regime takes place after a finite number of bifurcations, the number of bifurcations decreasing as Θ approaches the boundary of this region.

For the side-band modes, the transition to chaos occurs only by breaking the quasi-periodicity (Fig. 5b). In this case, many windows of periodic behavior are observed. Within some of them, period doubling bifurcation can be noticed. These bifurcations correspond to the doubling of the resonance cycle on a torus.

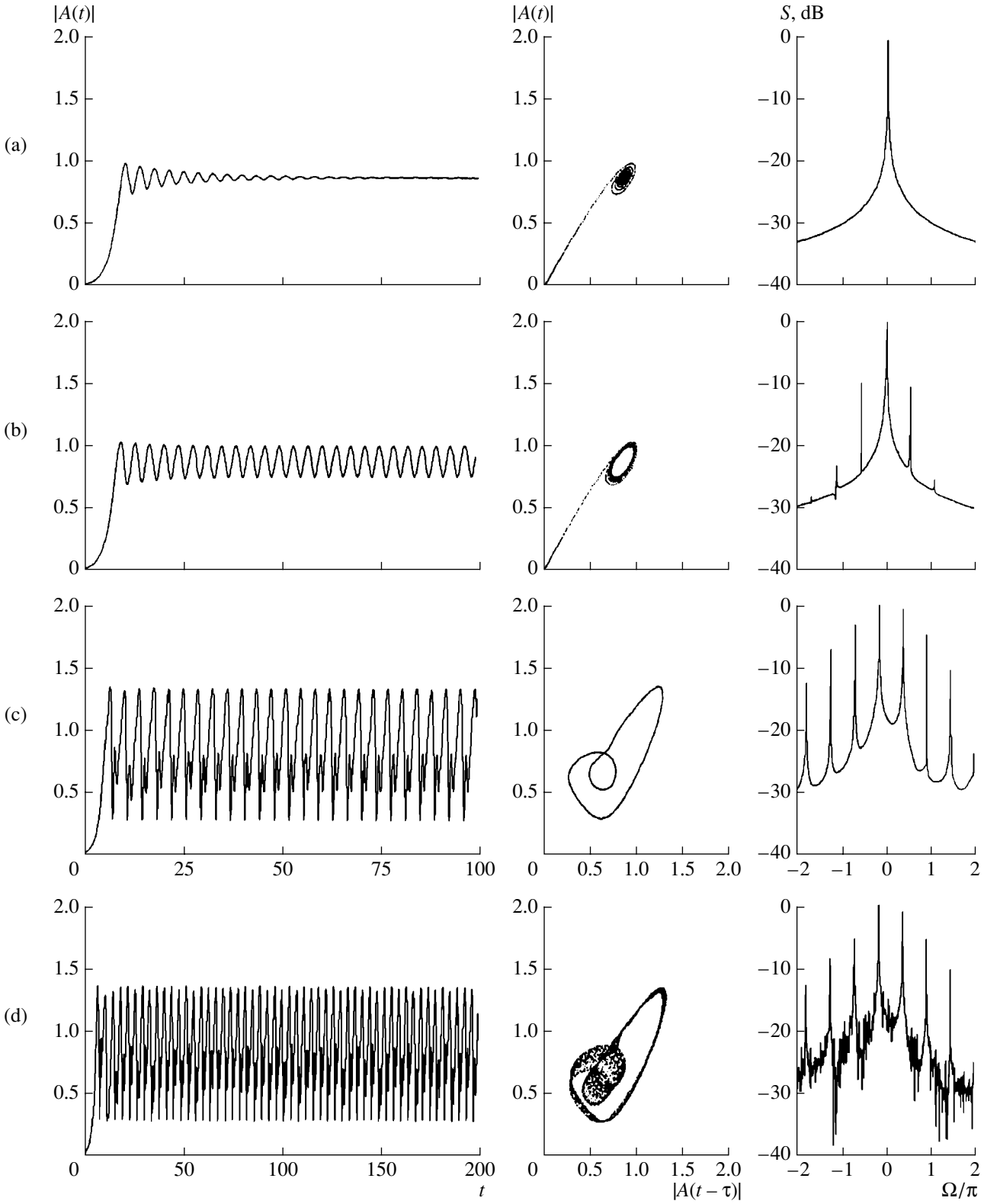


Fig. 3. Time series, phase portraits, and spectra for various dynamic conditions near the center of the oscillation generation region at $\gamma = 0.3$ and $\Theta = 0.1\pi$. $\alpha = 1.3$ (a), 1.4 (b), 2.35 (c), and 2.38 (d).

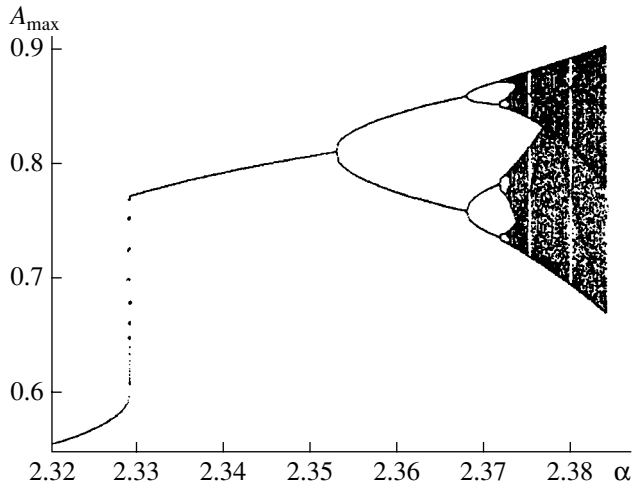


Fig. 4. Bifurcation diagram near the transition to chaos through a sequence of period doubling bifurcations for $\gamma = 0.3$ and $\Theta = 0.1\pi$.

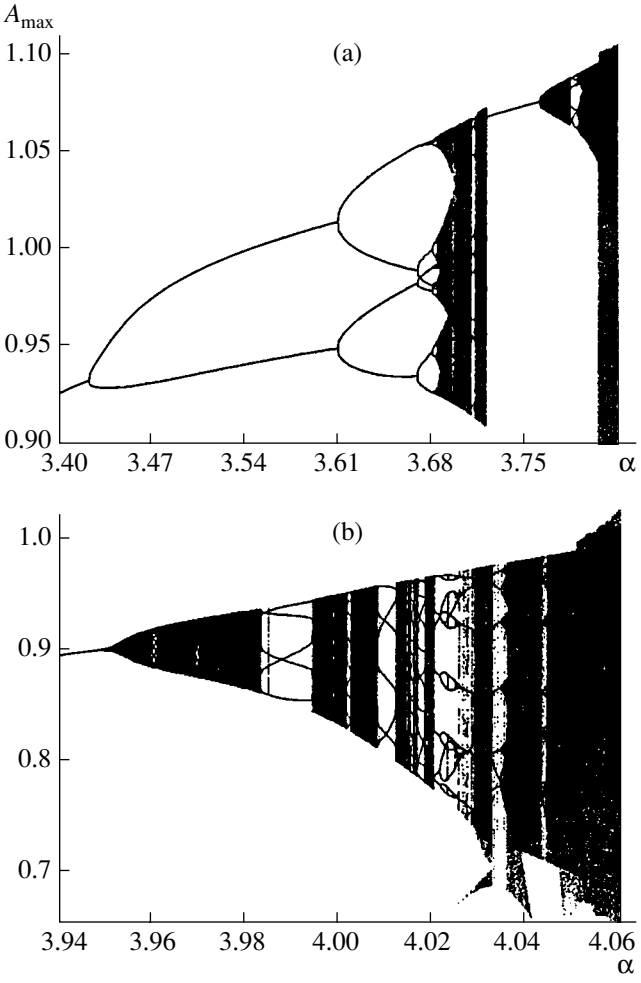


Fig. 5. Bifurcation diagrams for the (a) fundamental and (b) side-band modes at $\gamma = 0.3$ and $\Theta = 0.9\pi$.

It should be emphasized that our results are similar to the experimental data obtained in [7], where a microwave oscillator built around a TWT with DF was studied. The feedback loop of the oscillator contained a tunable narrow-band filter. When the resonance frequency of the filter was close to that of one of the eigenmodes, the transition to chaos followed the Feigenbaum scenario. If the resonance frequency lay roughly in the middle between the frequencies of two modes, the quasi-periodic route to chaos occurs.

When the factor of nonequilibrium α far exceeds the threshold of chaos initiation, the attractors corresponding to the fundamental and side-band modes combine. This effect was also observed earlier [21]. The dynamics of the system with the combined attractor is shown in Fig. 6. The temporal realization is highly irregular and virtually exhibits random switchings between the two attractors, i.e., chaotic intermittency. The spectrum, being rather noisy, has two distinct peaks at the frequencies of the modes. Any well-defined large-scale structure can no longer be revealed in the phase portrait. Such conditions have been called high-dimensional or fully developed chaos and are a feature of distributed systems, in particular, of many electron–wave DSOSs (see, e.g., [7, 23, 24]).

3. THREE-MODE APPROXIMATION

Investigators of distributed systems often invoke various simplifying finite-dimensional models, which consider a small number of interacting modes. The best-known example is the Lorentz system in the problem of thermal convection [1, 4, 18]. Let us construct a model of system (1) where only the fundamental and two self-modulation modes are involved. The solution to this three-mode system is found in the form of (9), where A_0 and $A_\pm$ are slowly varying functions. We assume that they change insignificantly within the delay time (which equals unity in our normalization). Certainly, the amplitudes of the satellites are no longer small. Substituting (9) into (1) and applying the method of averaging [17, 18], we easily come to

$$\begin{aligned}
&\dot{A}_0 + (\gamma + i\omega)A_0 \\
&= \alpha e^{i(\Theta-\omega)}[(1 - |A_0|^2 - 2|A_+|^2 - 2|A_-|^2)A_0 - 2A_0^* A_+ A_-], \\
&\dot{A}_+ + (\gamma + i\omega + i\Omega)A_+ \\
&= \alpha e^{i(\Theta-\omega-\Omega)}[(1 - |A_+|^2 - 2|A_0|^2 - 2|A_-|^2)A_+ - A_0^2 A_-^*], \\
&\dot{A}_- + (\gamma + i\omega - i\Omega)A_- \\
&= \alpha e^{i(\Theta-\omega+\Omega)}[(1 - |A_-|^2 - 2|A_0|^2 - 2|A_+|^2)A_- - A_0^2 A_+^*].
\end{aligned} \tag{16}$$

By passing to the real amplitudes and phases in (16) provided that $A_0 = a_0\exp(i\varphi_0)$ and $A_\pm = a_\pm\exp(i\varphi_\pm)$, it is easy to check that the phase space of the system is four-dimensional (three real amplitudes and the phase Φ =

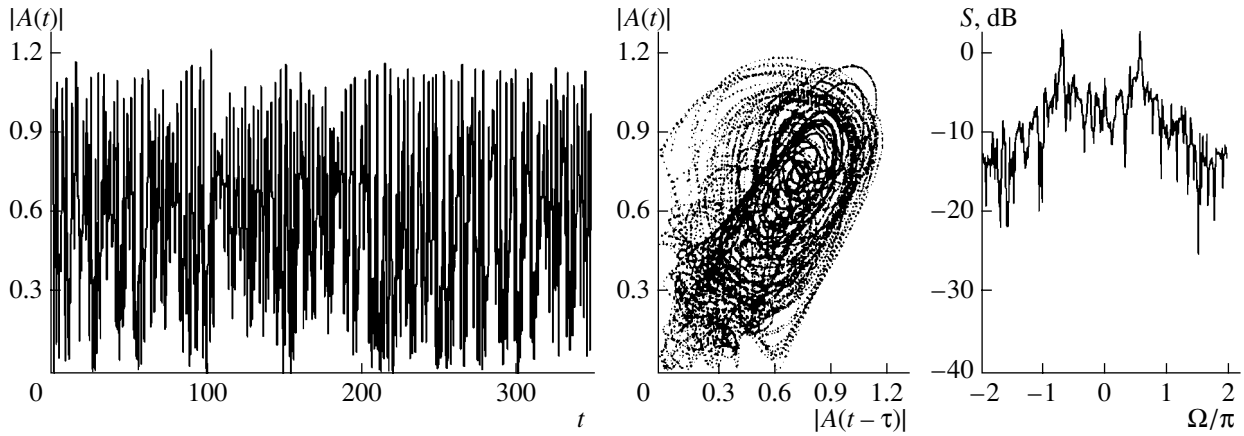


Fig. 6. Temporal realization, phase portrait, and spectrum after the combination of the attractors for $\gamma = 1.0$, $\Theta = 0.9\pi$, and $\alpha = 5.68$.

$2\varphi_0 - \varphi_+ - \varphi_-$). In the single-mode case ($A_\pm = 0$), Eq. (16) becomes

$$\dot{A}_0 + (\gamma + i\omega)A_0 = ae^{i(\Theta - \omega)}(1 - |A_0|^2)A_0.$$

This equation describes the conditions of single-frequency oscillations at the fundamental mode, with the self-excitation conditions, oscillation amplitude, and oscillation frequency coincident with those found in Section 1 [formulas (4), (5), and (8)]. In addition, from (16), one can also derive stability conditions for the single-frequency regime that are totally coincident with those in Subsection 1.3.

It is important that system (16) coincides (up to factors) with finite-dimensional models widely used in the analysis of mode competition in resonant electron masers (free-electron lasers, gyros, etc.) [25, 26]. However, these models fail to describe the complex dynamics inherent in such devices. For them, only constant-amplitude (single-mode or three-mode) conditions are stable. Therefore, we can argue that distributed model with delay (1), providing a variety of dynamic conditions, closer approximates actual systems.

CONCLUSION

Thus, the simple model of a self-excited oscillator with cubic nonlinearity and delay exhibits a wide diversity of self-oscillation conditions in the parameter space. As the factor of nonequilibrium grows, a complicated sequence of alternating regular and random self-modulation conditions is observed prior to chaos sets in the system. It is basically incorrect to consider a single scenario of transition to chaos, since almost all scenarios typical of finite-dimensional dynamic systems are realized. A sequence of bifurcations ends with fully developed chaos, which is characterized by a much more uniform continuous spectrum and the absence of any well-defined large-scale structure on the phase portrait projection. Such a situation is typical of electron DSOSs, such as TWTs with DF [7] and BWOs [23, 24]. It should also be noted that in the three-mode approximation, model (1) may be transformed into model (16), which was used earlier to analyze mode competition in electron masers [25, 26]. The results of the stability analysis for the single-mode conditions coincide for both models; however, the distributed system, being more advanced, allows one to describe the complex dynamics and chaos. Thus, the reduced version of van der Pol equation with delay (1) can be used as a reference model of oscillators with delayed feedback.

ACKNOWLEDGMENTS

This work was supported by the CRDF (grant Award REC-006) and the Federal Program “Integration” (project no. A0057/2001).

REFERENCES

1. Yu. I. Neimark and P. S. Landa, *Stochastic and Chaotic Oscillations* (Nauka, Moscow, 1987; Kluwer, Dordrecht, 1992).
2. A. S. Dmitriev and V. Ya. Kislov, *Stochastic Oscillations in Radiophysics and Electronics* (Nauka, Moscow, 1989).
3. S. P. Kuznetsov, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **25**, 1410 (1982).
4. P. S. Landa, *Nonlinear Oscillations and Waves in Dynamic Systems* (Kluwer, Dordrecht, 1996; Nauka, Moscow, 1997).
5. A. S. Dmitriev and S. O. Starkov, Radiotekh. Élektron. (Moscow) **33**, 1427 (1988).
6. V. Ya. Kislov, E. A. Myasin, and N. N. Zalogin, Radiotekh. Élektron. (Moscow) **25**, 2160 (1980).
7. V. A. Kats, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **28**, 161 (1985).
8. Yu. P. Bliokh, A. V. Borodkin, M. G. Lyubarskiĭ, *et al.*, Izv. Vyssh. Uchebn. Zaved., Prikl. Nelineĭnaya Din. **1** (1–2), 34 (1993).
9. V. S. Ergakov and M. A. Moiseev, Radiotekh. Élektron. (Moscow) **31**, 962 (1986).

10. V. V. Afanas'eva and A. G. Lazerson, Izv. Vyssh. Uchebn. Zaved., Prikl. Nelineĭnaya Din. **3** (4), 88 (1995).
11. T. V. Dmitrieva, N. M. Ryskin, V. N. Titov, and A. M. Shigaev, Izv. Vyssh. Uchebn. Zaved., Prikl. Nelineĭnaya Din. **7** (6), 66 (1999).
12. B. S. Dmitriev, Yu. D. Zharkov, N. M. Ryskin, and A. M. Shigaev, Radiotekh. Élektron. (Moscow) **46**, 561 (2001).
13. T. M. Antonsen and B. Levush, Phys. Fluids B **1**, 1097 (1989).
14. N. S. Ginzburg and M. I. Petelin, Izv. Vyssh. Uchebn. Zaved., Prikl. Nelineĭnaya Din. **2** (6), 3 (1994).
15. N. S. Ginzburg and A. S. Sergeev, Zh. Tekh. Fiz. **61** (6), 133 (1991) [Sov. Phys. Tech. Phys. **36**, 665 (1991)].
16. N. I. Aĭzatskiĭ, Zh. Tekh. Fiz. **57**, 1532 (1987) [Sov. Phys. Tech. Phys. **32**, 914 (1987)].
17. V. P. Rubanik, *Oscillations of Quasilinear Systems with Delay* (Nauka, Moscow, 1969).
18. M. I. Rabinovich and D. I. Trubetskov, *Oscillations and Waves in Linear and Nonlinear Systems* (Nauka, Moscow, 1984; Kluwer, Dordrecht, 1989).
19. V. A. Solntsev and T. M. Andreevskaya, Radiotekh. Élektron. (Moscow) **28**, 561 (1983).
20. T. M. Andreevskaya and V. A. Solntsev, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **32** (1), 34 (1989).
21. V. A. Kats and S. P. Kuznetsov, Pis'ma Zh. Tekh. Fiz. **13**, 727 (1987) [Sov. Tech. Phys. Lett. **13**, 302 (1987)].
22. D. V. R. Reddy, A. Sen, and G. L. Johnston, Physica D (Amsterdam) **144**, 335 (2000).
23. N. M. Ryskin and V. N. Titov, Izv. Vyssh. Uchebn. Zaved., Prikl. Nelineĭnaya Din. **6** (1), 75 (1998).
24. N. M. Ryskin, V. N. Titov, and D. I. Trubetskov, Dokl. Akad. Nauk **358**, 620 (1998) [Dokl. Phys. **43**, 90 (1998)].
25. G. S. Nusinovich, Pis'ma Zh. Tekh. Fiz. **6**, 848 (1980) [Sov. Tech. Phys. Lett. **6**, 364 (1980)].
26. L. G. Blyakhman and G. S. Nusinovich, Radiotekh. Élektron. (Moscow) **27**, 996 (1982).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 7, 2002, pp. 803–806. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 9–12.

THEORETICAL AND MATHEMATICAL PHYSICS

Nonlinear Friction as a Mechanism of Directed Motion

Yu. L. Bolotin*, A. V. Tur, and V. V. Yanovsky*****

** Institute of Theoretical Physics, National Scientific Center Kharkov Institute of Physics and Technology, ul. Akademicheskaya 1, Kharkov, 61108 Ukraine*

e-mail: bolotin@kipt.kharkov.ua

*** Center D'etude Spatiale Des Rayonnements, 9, Av. du colonel Roche, 31028 Toulouse, France*

**** Institute of Single Crystals, National Academy of Sciences of Ukraine, pr. Lenina 60, Kharkov, 61001 Ukraine*

Received July 16, 2001

Abstract—A simple model (ratchet model) of occurrence of directed motion under the action of a zero-mean fluctuating force is proposed. The motion arises when the symmetry in the velocity space is violated by nonlinear friction. The mechanism of the directed motion is discussed qualitatively. Existence conditions of the motion are derived. The efficiency of conversion of the fluctuating random force to the directed motion is estimated. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Over the past decade, interest has been rekindled in the problem of rectification of fluctuations, i.e., in producing a macroscopic flux of particles in the absence of applied forces or pressure, temperature, concentration, etc. gradients [1]. Not only conditions under which related devices could function have been formulated but also direct observations of the particle flux have been accomplished [2–5]. The majority of such devices (models) are called ratchets because interest in this problem has been largely inspired by Feynman *et al*. [6], who discussed the ratchet-and-pawl model (first used by Smoluchowski [7]). Emphasis has been focused on the problem of transport in spatially periodic systems in the absence of external forces (i.e., when space-, time-, and ensemble-averaged driving forces are zero). The most attention has been placed on small-scale systems (the so-called Brownian world [8]), in which thermal fluctuations cannot be neglected. If transients are ignored, directed transport in these systems, being in contact with a single source of dissipation and fluctuations (a thermal reservoir), is forbidden by the second principle of thermodynamics. The question arises as to what the conditions are for ordered Brownian transport to exist in the absence of directed forces? These are nonequilibrium of the fluctuations and a violated reflection symmetry. The first condition is obvious in the context of the second principle of thermodynamics, while the second one is necessary for selecting the preferred transport direction. By violating the symmetry of detailed balance (i.e., by making fluctuations nonequilibrium) and the reflection symmetry of the spatially periodic potential, we eliminate all forbidding symmetries and thereby (in accordance with the Curie principle [9]) allow the flux to exist in the absence of macroscopic gradients. One of the best-studied techniques for creating nonequilibrium is to make the temperature space- or time-dependent [10]. Such models have been called diffusion ratchets. In terms of the Einstein relation for motion in a potential field with violated reflection symmetry, nonequilibrium can be achieved by introducing a space or time dependence of the friction factor (model of a friction ratchet). In this paper, we consider another approach; namely, we introduce the dependence of the friction factor on the direction of particle motion. In the simplest case, this dependence can be induced by the shape of a particle or by its conformations. In our model, conformation transitions are independent of intrinsic transitions in the particle but are induced by the environment. Further, we use the term "umbrella" to refer to particles that undergo such conformations. For an umbrella subjected to zero-mean forces, the question of existence of a nonzero flux even at the zero potential may be posed, because the reflection symmetry is violated even at the friction factor level. The associated model can be called the model of nonpotential friction ratchet. The main purpose of this paper is to study the dynamics of this model. We will derive conditions for the directed motion of an umbrella under the action of zero-mean random forces. The efficiency of such a mechanism is discussed. The possibility of overcoming the constant field of the forces is demonstrated and conditions imposed on the noise amplitude for such a motion are derived in terms of the kinetic approach.

MECHANISM OF THE DIRECTED MOTION

Consider the motion of an object (umbrella) in a medium with friction under the action of a random force $F(t)$ (Fig. 1). The statistical characteristics of the random force are assumed to be known. The main prop-

1063-7842/02/4707-0803$22.00

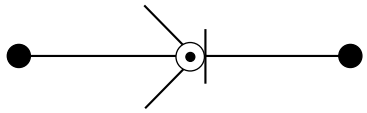


Fig. 1. Object consisting of an axis, two movable blades, and a stopper.

erty of this force is $\langle F(t)\rangle = 0$. The equation of motion of this object can be written as

$$m\frac{d^2x}{dt^2} = F(t) - \alpha(\dot{x})\frac{dx}{dt}, \quad (1)$$

where α is the friction factor dependent on the motion direction.

The reason for this dependence is clarified from Fig. 1. Let us consider it in more detail. The simplest reasoning is as follows. When the umbrella moves in the positive x direction, it closes and the friction factor is α_1. When the umbrella moves in the opposite direction, it opens up and the friction factor increases to α_2. Such a variation is easily described as

$$\alpha = \frac{1}{2}(\alpha_2 + \alpha_2) + \frac{1}{2}(\alpha_1 - \alpha_2)\frac{\dot{x}}{|\dot{x}|}. \quad (2)$$

Now, let us write the starting equation in terms of velocity $\dot{x} = v$:

$$m\dot{v} = F(t) - \frac{1}{2}((\alpha_1 + \alpha_2)v + (\alpha_1 - \alpha_2)|v|). \quad (3)$$

For simplicity, this model assumes that the umbrella opens up and closes instantaneously. Obviously, this assumption is reasonable if the characteristic times T_c of the random force and the characteristic times τ_c of the umbrella satisfy the inequality $T_c \gg \tau_c$. More sophisticated models can be constructed by introducing, for example, continuous α vs. velocity dependences

$$\alpha_e(v) = \alpha_1 + \frac{\alpha_2 - \alpha_1}{1 + \exp(\beta v)}. \quad (4)$$

Here, β is the characteristic scale of the transition between the two values of the friction factor. The model of the umbrella and its characteristics are specified by this function. Another example is

$$\alpha_p(v) = \frac{1}{2}\left((\alpha_1 + \alpha_2) + \frac{(\alpha_1 - \alpha_2)v}{\sqrt{\varepsilon^2 + v^2}}\right), \quad (5)$$

where ε^2 is the characteristic parameter of the transition.

Now, let us discuss the mechanism behind the directed motion under the action of a fluctuating random force. The reason for this effect is that the symmetry of the equation of motion with respect to the change of variables $v \rightleftharpoons -v$ is violated. To estimate the directed velocity, we reason as follows. Let us average Eq. (3) over the random force. The steady state is given by the equality

$$\langle v\rangle = \frac{\alpha_2 - \alpha_1}{\alpha_2 + \alpha_1}\langle|v|\rangle. \quad (6)$$

Now the average value of the velocity magnitude is expressed through the quadratic moment:

$$\langle|v|\rangle \simeq \langle v^2\rangle^{\frac{1}{2}}.$$

The velocity pair correlator can be estimated from the energy balance condition for the steady state:

$$\langle vF\rangle = \frac{1}{2}(\alpha_1 + \alpha_2)\langle v^2\rangle.$$

The correlator on the left side is expressed in terms of correlation characteristics of the force. This is easy to do if we bear in mind the estimate

$$v \simeq \frac{1}{m}\int F(\tau)d\tau.$$

Then,

$$\langle Fv\rangle = \frac{1}{m}\int\langle F(t)F(\tau)\rangle d\tau \approx \frac{1}{m}\langle F^2\rangle\tau_c.$$

By substituting this expression into Eq. (6), we finally obtain

$$\langle v\rangle \simeq \frac{(\alpha_2 - \alpha_1)}{(\alpha_2 + \alpha_1)}\left(\frac{2\tau_c\langle F^2\rangle}{m(\alpha_2 + \alpha_1)}\right)^{\frac{1}{2}}. \quad (7)$$

Equation (7) specifies the velocity of the directed motion due to the fluctuating random force with zero mean. The magnitude and direction of the velocity is defined by the difference $(\alpha_2 - \alpha_1)$. Clearly, these simple estimates demonstrate only the possibility of the directed motion and need more rigorous substantiation. To this end, we will derive a kinetic equation for the umbrella velocity distribution.

KINETIC EQUATION

To derive the kinetic equation, we use the usual definition of the distribution function:

$$f(V, t) = \langle\delta(V - v(t))\rangle.$$

Here, $v(t)$ is a solution to the dynamic equation with random force over which averaging is performed. By taking the time derivative of $f(V, t)$ and using the equation of motion, we obtain

$$\frac{\partial f}{\partial t} = -\frac{1}{m}\frac{\partial}{\partial V}(\langle F(t)\delta(V - v(t))\rangle - \alpha(V)Vf) \quad (8)$$

or, in a more convenient form,

$$\frac{\partial f}{\partial t} - \frac{\partial}{\partial V}\left(\frac{\alpha(V)}{m}Vf\right) = -\frac{1}{m}\frac{\partial}{\partial V}\langle F(t)\delta(V - v(t))\rangle. \quad (9)$$

Thus, to solve the problem, we must average the term on the right-hand side and to express it through the distribution function. Using the conventional methods (see, e.g., [11–13]), one arrives at a closed equation for the distribution function when the random force is Gaussian and δ-correlated:

$$\frac{\partial f}{\partial t} - \frac{\partial}{\partial V}\left(\frac{\alpha(V)}{m}Vf\right) = \frac{\langle F^2\rangle\tau_c}{2m^2}\frac{\partial^2 f}{\partial V^2}. \quad (10)$$

Here, we use the notation $\int_{-\infty}^{\infty}\langle F(t)F(\tau)\rangle\, d\tau \equiv \langle F^2\rangle\tau_c$, which has clear physical meaning. The steady-state solution to this equation (the zero flux in the velocity space) can easily be found for two-level model (2) of $\alpha(V)$:

$$f(V) = C\exp\left\{-\frac{2m}{\langle F^2\rangle\tau_c}\int V\alpha(V)dV\right\}$$
$$= C\exp\left\{-\frac{mV^2}{\langle F^2\rangle\tau_c}(\alpha_2 + (\alpha_1 - \alpha_2)\Theta(V))\right\}. \quad (11)$$

The constant factor C is defined by the normalization condition:

$$C = \left(\frac{m\alpha_1\alpha_2}{\pi\langle F^2\rangle\tau_c}\right)^{\frac{1}{2}}\frac{1}{\sqrt{\alpha_2} + \sqrt{\alpha_1}}.$$

Using the above equilibrium distribution function, we obtain the umbrella average velocity:

$$\langle V\rangle = \int_{-\infty}^{\infty} Vf(V)dV = \left(\frac{\langle F^2\rangle\tau_c}{\pi m\alpha_1\alpha_2}\right)^{1/2}\frac{\alpha_2 - \alpha_1}{\sqrt{\alpha_2} + \sqrt{\alpha_1}}. \quad (12)$$

This velocity differs from estimate (7) not only by the constant factor but also by its more complex dependence on the two-level friction factor. Now, let us estimate the efficiency of such an umbrella motion in terms of its average energy:

$$\zeta = \frac{\langle V\rangle^2}{\langle V^2\rangle}$$
$$= \frac{2(\alpha_2 - \alpha_1)^2}{\pi(\sqrt{\alpha_1} + \sqrt{\alpha_2})(\alpha_1^{3/2} + \alpha_2^{3/2})} \xrightarrow{\alpha_2 \gg \alpha_1} \frac{2}{\pi}. \quad (13)$$

In a sense, this way of motion is sufficiently effective when the difference in the friction factors is large ($\alpha_2 \gg \alpha_1$). The standard efficiency is given by

$$\eta = \frac{\langle V\rangle^2}{\langle V\rangle^2 + \langle V^2\rangle} = \frac{\zeta}{1+\zeta} \xrightarrow{\alpha_2 \gg \alpha_1} \frac{2}{\pi + 2}.$$

This value is the maximum achievable efficiency for this way of motion.

ACTION OF CONSTANT FORCE

The efficiency of the resultant directed motion is sufficiently high. To corroborate this statement, we will consider the motion of an umbrella in a constant gravitational field. This case is very close to the way of motion of jellyfishes. Let a constant gravitational force be directed in the negative x direction. Then, the kinetic equation that describes the umbrella motion has the form

$$\frac{\partial f}{\partial t} - \frac{\partial}{\partial V}\left(gf + \frac{\alpha(V)}{m}Vf\right) = \frac{\langle F^2\rangle\tau_c}{2m^2}\frac{\partial^2 f}{\partial V^2}. \quad (14)$$

A solution to this equation can easily be found. As in the previous case,

$$f(V) = C\exp\left\{-\frac{2m}{\langle F^2\rangle\tau_c}\int(gm + V\alpha(V))dV\right\}$$
$$= C\exp\left\{-\frac{2gm^2V}{\langle F^2\rangle\tau_c} - \frac{mV^2}{\langle F^2\rangle\tau_c}(\alpha_2 + (\alpha_1 - \alpha_2)\Theta(V))\right\}. \quad (15)$$

Here, C is the normalization constant. This constant is found from the normalization condition for the function f:

$$\frac{C}{q}(I_1 + I_2) = 1,$$

where $q = 2gm^2/\langle F^2\rangle\tau_c > 0$ and $I_i\,(i = 1, 2)$ are defined as

$$I_i = \sqrt{\pi}\beta_i e^{\beta_i^2}\operatorname{erfc}((-1)^{i+1}\beta_i).$$

The parameters $\beta_i\,(i = 1, 2)$ are

$$\beta_i = \frac{m^{3/2}g}{\sqrt{\alpha_i\langle F^2\rangle\tau_c}}. \quad (16)$$

Then, the condition for the directed motion in the positive x direction is written as

$$\langle V\rangle = \frac{-2}{q(I_1 + I_2)}(\beta_1^2 I_1 + \beta_2^2 I_2 + \beta_2^2 - \beta_1^2) > 0. \quad (17)$$

It is easy to check that there exists a range of parameters where this condition is satisfied. This condition can be qualitatively understood from simple physical considerations. The Langevin equation that is used to derive the kinetic equation is equivalent to the overdamped case of the usual Langevin equation. This can readily be shown by applying the substitutions $V \longrightarrow x$ and $m \longrightarrow \alpha$. This means that the behavior of this system can be described in terms of the overdamped Langevin equation, which essentially considers the particle to execute finite movements in the field of the effective potential

$$U(x) = mgx + \frac{x^2}{2}(\alpha_2 + (\alpha_1 - \alpha_2)\Theta(x)).$$

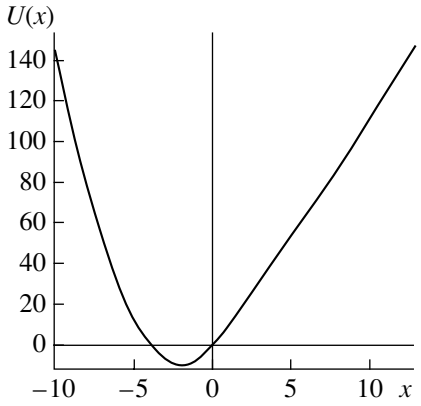


Fig. 2. Effective potential vs. coordinate.

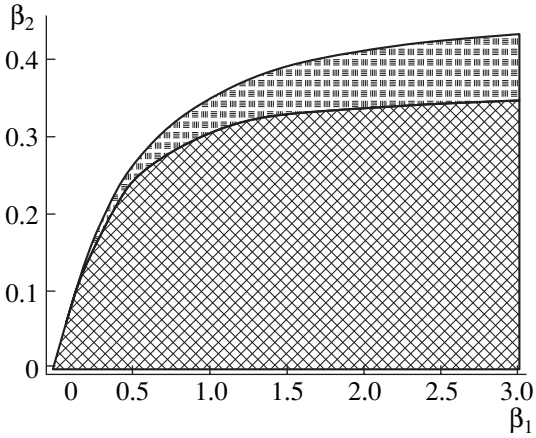


Fig. 3. Parameter ranges in which the upward motion of the umbrella is predicted by the crude (heavy shade) and exact (light shade) conditions.

It is clear that the global minimum of this potential lies in the range of negative x (Fig. 2).

The foregoing implies that the average particle coordinate is negative when the particle energy is low. Turning back to the initial variables, we find that the umbrella has a negative average velocity and falls under gravity. However, the situation changes with increasing particle energy. The average coordinate of the particle may become positive because of the different asymptotic behavior of the potential ($\alpha_1 < \alpha_2$). It can be proved that the average coordinate is positive if the energy satisfies the condition

$$E \geq \frac{4(mg)^2(\alpha_1 + \alpha_2)}{(\alpha_2 - \alpha_1)^2}.$$

In the overdamped case, the energy is defined by the pair correlator of the random external force. This means that, when

$$\frac{\langle F^2 \rangle \tau_c}{2m} \geq \frac{4(mg)^2(\alpha_1 + \alpha_2)}{(\alpha_2 - \alpha_1)^2}, \qquad (18)$$

the umbrella will move in the positive x direction (opposite to the gravity). Physically, this condition is rather crude because it ignores the difference between the motion characteristic times in the regions $x > 0$ and $x < 0$. Taking this difference into account will predict the antigravitational motion at β_2 to be higher than those given by (18). This also follows from exact condition (17). Figure 3 illustrates the difference between the exact and crude conditions. Thus, the efficiency of this way of motion is sufficient to overcome the counteracting constant forces. In a sense, this mechanism of directed motion may be observed in micro- and macrofauna. In macrofauna, the role of random forces may be played by periodic forces due to the suction and ejection of a medium (liquid). Such a mechanism is used by many aquatic organisms, for example, by jellyfishes. It should be noted that here, the ejection force does not necessarily exceed the suction force. The period-averaged values of these forces may and must be zero. Molecules of special asymmetric shape can use a similar mechanism to move in random external fields.

REFERENCES

1. P. Reimann, Phys. Rep. **361**, 57 (2002); cond-mat/0010237.
2. J. Rousselet, L. Salome, A. Ajdari, *et al.*, Nature **370**, 446 (1994).
3. L. P. Faucheux, L. S. Bourdieu, R. D. Kaplan, *et al.*, Phys. Rev. Lett. **74**, 1504 (1995).
4. C. Mennerat-Robillard, D. Lucas, S. Guibal, *et al.*, Phys. Rev. Lett. **82**, 851 (1999).
5. V. B. Belinicher and B. I. Sturman, Usp. Fiz. Nauk **130**, 415 (1980) [Sov. Phys. Usp. **23**, 199 (1980)].
6. R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics* (Addison-Wesley, Reading, 1963; Mir, Moscow, 1965).
7. M. Smoluchowski, Phys. Z. **13**, 1069 (1912).
8. M. O. Magnasco, Phys. Rev. Lett. **71**, 1477 (1993).
9. P. Curie, J. Phys. (Paris), No. 3, 393 (1894).
10. P. Reinmann, R. Bartussek, P. Haussler, and P. Hanggi, Phys. Lett. A **215**, 26 (1996).
11. Yu. L. Klimontovich, *Statistical Physics* (Nauka, Moscow, 1982; Harwood Academic, New York, 1986).
12. V. I. Klyatskin, *Stochastic Equations and Waves in Randomly Inhomogeneous Media* (Nauka, Moscow, 1980).
13. H. Riskin, *The Fokker–Planck Equation* (Springer-Verlag, Berlin, 1989).

Translated by A. Khzmalyan

Technical Physics, Vol. 47, No. 7, 2002, pp. 807–813. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 13–20.

THEORETICAL AND MATHEMATICAL PHYSICS

Diffusion of Nanoparticles in a Rarefied Gas

V. Ya. Rudyak and S. L. Krasnolutskiĭ

Novosibirsk State University of Architecture and Civil Engineering, Novosibirsk, 630008 Russia

e-mail: rudyak@ngasu.nsk.su

Received August 1, 2001; in final form, December 7, 2001

Abstract—It is suggested to describe the diffusion of nanoparticles in rarefied gases in terms of the kinetic theory. For this purpose, the potential of interaction between a carrier gas molecule and a dispersed particle is constructed by summing the interactions of the given gas molecule with all atoms (molecules) of the dispersed particle. With this potential, a formula for the diffusion coefficient of the dispersed nanoparticle is derived. The dependence of the diffusion coefficient on the radius and temperature is studied. Analytical results are compared with experimental data. The well-known experimental Cunningham–Millikan correlation is shown to apply only in the range of near-room temperatures, for which the parameters of this correlation were determined. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The rapid advance in nanotechnologies necessitates the study of particle transport in gases and liquids. The size of nanoparticles ranges from several tens to several hundred of angströms. By nanoparticles, both individual dispersed particles and their clusters are meant. In studying the transport processes, the diffusion of nanoparticles and the calculation of their diffusion coefficient are key points. However, there is no general agreement among researchers as to how to treat the diffusion process. Some argue that the diffusion of nanoparticles can be described in the same way as that of usual Brownian particles, that is, with the Einstein diffusion coefficient

$$D_E = kT/\gamma_s, \quad \gamma_s = 6\pi\eta R, \tag{1}$$

where η is the viscosity of the medium, R is the characteristic radius of a Brownian particle, and T is the temperature.

Other researchers, however, try to describe the diffusion of nanoparticles in terms of the kinetic theory of rarefied or dense gases, considering the particles as molecules. Strictly speaking, it is unjustified to treat the diffusion of nanoparticles by either of the two methods. In fact, the nanoparticle sizes are on the order of the hydrodynamic infinitesimal scale of a carrier medium. Therefore, the use of the hydrodynamic approach to characterize the drag to particle motion from the carrying medium is incorrect; however, exactly this type of approach is implied in formula (1), where the Stokes force appears as the drag force. If the kinetic theory is applied, an appropriate potential of interaction between gas molecules and a nanoparticle should be chosen. The use of the conventional intermolecular potential in this case is also unjustified.

By means of the molecular dynamics method, it was shown [1, 2] that the mechanism of diffusion of nanoparticles in liquids and gases differs radically from both the Einstein mechanism (for Brownian particles) and the molecular mechanism. If, however, the particles are sufficiently small, their dynamics in rarefied gases can be well described (at least in the first approximation) by a set of Boltzmann kinetic equations [3–5]. In this case, an appropriate interaction potential should be available, as was noted above. Such a potential was constructed in our earlier works [6, 7]. In this article, this potential of interaction between gas molecules and dispersed particles is applied to calculate the diffusion coefficient of nanoparticles in a rarefied gas.

MOLECULE–DISPERSED PARTICLE INTERACTION POTENTIAL

The interaction of a carrier gas molecule with a dispersed particle is an extremely complex process. In the general case, its description requires the solution of a quantum-mechanical problem. However, in a wide range of dispersed system parameters, the interaction can be described classically in terms of model potential functions. Interactions between atoms and a solid surface are usually treated in this way [8–10]. The models have proven to be adequate in various aerodynamic applications, microelectronics, reactor materials science, and thermal nuclear materials science. The idea behind such a model is fairly simple, which makes it very attractive. A solid surface is considered a set of atoms (or molecules), and the potential of interaction between a carrier gas molecule and the surface is found as the sum of the potentials of interaction of a given molecule with all surface atoms. The same approach is adopted in this article to simulate the molecule–solid particle interaction potential.

Let a solid dispersed particle be simulated by a set of atoms of a solid that are enclosed in a sphere of

1063-7842/02/4707-0807$22.00

radius R. The interaction potential of a gas molecule colliding with the ith atom of the particle is described by the pair potential $\Phi_i(|\mathbf{r}_m - \mathbf{r}_i|)$, where $\mathbf{r}_m$ and $\mathbf{r}_i$ are the radius vectors of the striking molecule and the atom, respectively. Then, if it is assumed that the gas molecule–atom interaction is additive, the molecule–particle interaction potential $\Phi(|\mathbf{r}|)$ has the form

$$\Phi(|\mathbf{r}|) = \sum_{i=1}^{N} \Phi_i(|\mathbf{r}_m - \mathbf{r}_i|), \quad (2)$$

where N is the number of atoms (molecules) of the particle.

There exist many models of pair interaction potentials, among which the Lennard-Jones, Kihara, and Morse potentials are the most popular. In numerical simulations and direct calculations, the first-named potential is usually used. We also will apply the Lennard-Jones potential to simulate the interaction of a gas molecule with an atom of the solid surface. The Lennard-Jones potential has the form

$$\Phi_i(|\mathbf{r}|) = 4\varepsilon_{ii}\left[\left(\frac{\sigma_{ii}}{r}\right)^{12} - \left(\frac{\sigma_{ii}}{r}\right)^{6}\right], \quad (3)$$

where ε_{ii} is the potential well depth and $\sigma_{ii}/2$ is the effective repulsion radius.

It is usually impossible to take the sum of series (2). Therefore, summation in (2) is replaced by integration in practice [8]. Physically, this means that the solid is approximated by a continuous model.

We will also assume that the particle is spherical and homogeneous and has a radius R. Then, series (2) to be summed has a finite number of terms. Within the continuous model, the potential is calculated by taking the triple integral. In this case, the potential is averaged over current coordinates of integration. These are radius a and two, polar (Θ) and azimuth, angles. Moreover, if the interaction potential does not depend on the mutual arrangement of the gas molecule and the atom (which is strictly valid only for monoatomic and nonpolar molecules), formula (2) is reduced to the integral [6]

$$\Phi(x) = \int_0^R\int_0^\pi \frac{2\pi a^2 da \sin\Theta d\Theta}{V}\Phi_{LJ} \times \left(\sqrt{(x - \arccos\Theta)^2 + a^2\sin^2\Theta}\right), \quad (4)$$

where x is the distance from the molecule to the center of the sphere (where the origin is placed); Φ_{LJ} is the Lennard-Jones potential for the interaction of a gas molecule with an atom (molecule) of the dispersed particle; V is the effective volume per molecule, $1/V = \rho_* N_A/\mu$; ρ_* is the density of the dispersed particle material, N_A is the Avogadro number, and μ is the molecular weight of molecules (atoms) constituting the dispersed particle; a and Θ are current coordinates of integration, which characterize the position and the radius of the circle of integration; and $dV = 2\pi a^2 da \sin\Theta d\Theta$ is the elementary volume of integration.

If Lennard-Jones potential 12–6 (3) is taken as the molecule–atom interaction potential, potential (4) is calculable and has the form

$$\Phi(r) = \Phi_9(r) - \Phi_3(r), \quad (5)$$

where

$$\Phi_9(r) = C_9\left\{\left[\frac{1}{(r-R)^9} - \frac{1}{(r+R)^9}\right] - \frac{9}{8r}\left[\frac{1}{(r-R)^8} - \frac{1}{(r+R)^8}\right]\right\},$$

$$\Phi_3(r) = C_3\left\{\left[\frac{1}{(r-R)^3} - \frac{1}{(r+R)^3}\right] - \frac{3}{2r}\left[\frac{1}{(r-R)^2} - \frac{1}{(r+R)^2}\right]\right\},$$

$$C_9 = \frac{4\pi\varepsilon_{ij}\sigma_{ij}^{1/2}}{45V}, \quad C_3 = \frac{2\pi\varepsilon_{ij}\sigma_{ij}^2}{3V}.$$

Potential (5) depends on the parameters ε_{ij} and σ_{ij} of the pair interaction potential between gas molecules and the dispersed particle and also on the size of the particle. If the particle is coarse (with a characteristic size of more than 10^{-5} cm), one can use, instead of formula (5), the much simpler solid surface–gas molecule interaction potential [8, 9]

$$\Phi(r) = C_9/z^9 - C_3/z^3, \quad (6)$$

where $z = r - R$. This follows from (5) for large R.

CALCULATION OF THE DIFFUSION COEFFICIENT OF RAREFIED GAS SUSPENSIONS

It has been shown [1–3] that the dynamics of even rarefied suspensions in gas (gas suspensions) can be described by a set of kinetic equations that include many-body collision integrals. However, rarefied ultradisperse (and in some cases finely disperse) gas suspensions can be well described by a set of Boltzmann kinetic equations. This, in particular, applies to gas suspensions and aerosols wherein nanoparticles are suspended. In the one-liquid approximation, the states of the gas suspension are characterized by unified mass-averaged macroscopic variables. The solution of the set of Boltzmann equations yields Navier–Stokes equations in which the transfer coefficients are defined by well-known expressions. Specifically, for the diffu-

sion coefficient, we have [11]

$$D = \frac{3}{16}\frac{\sqrt{2\pi\mu kT}}{n\mu\pi R^2\Omega_{ij}^{(1,1)*}(T^*, \sigma_{ij}/R)}, \quad (7)$$

where $\mu = mM/(m + M)$; m and M are the masses of a carrier gas molecule and a dispersed particle, respectively; $T^* = T/\varepsilon_{ij}$; and $\Omega_{ij}^{(l,r)*}$ are the so-called Ω integrals [11].

1. Calculation of Ω Integrals

Thus, to calculate diffusion coefficient (7) in our gas suspension, it is necessary first to calculate the Ω integrals for potential (5). To do this, one needs to know the constants of the pair potential of interaction [specifically, Lennard-Jones potential (3)] between a gas molecule and an atom (molecule) of the particle. Force constants are usually evaluated from experimental data for transport coefficients (viscosity, thermal conductivity, and diffusion coefficient) or virial coefficients. This can be done reliably, rigorously, and fairly accurately only for molecules of the same material. With these constants known, one usually calculates the parameters of the interaction potentials for dissimilar molecules using so-called combination relationships. They allow determination of the force constants of the potentials for molecular mixtures from the interaction constants for pure (single-component) substances. Naturally, these relationships can be derived only if the same potential model is used for components and for the mixture as a whole. The applicability of the conventional and original combination relationships was studied in [6, 7]. In this work, we took the simplest ones:

$$\sigma_{ij} = \sqrt{\sigma_{ii}\sigma_{jj}}, \quad \varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}}, \quad (8)$$

$$\sigma_{ij} = \frac{1}{2}(\sigma_{ii} + \sigma_{jj}), \quad \varepsilon_{ij}\sigma_{ij}^6 = \sqrt{\varepsilon_{ii}\varepsilon_{jj}}\sigma_{ii}^3\sigma_{jj}^3. \quad (9)$$

The Ω integral for potential (5) was calculated with an original computer program. It was tested with available data for mixtures of rarefied gases and proved to be highly accurate.

Ω integrals (and the transport coefficients as a consequence) depend strongly on the parameters of the molecular interaction potential in a rarefied gas (see, e.g., [11, 12]). Therefore, one must choose the proper combination relationships to calculate the integrals. Extensive calculations showed that, in contrast to the case of a rarefied gas, the Ω integrals for potential (5) depend on the choice of combination relationships insignificantly. Their dependence on the parameters of the intermolecular potential, the published data for which are very contradictory, is also weak. As an example, Table 1 lists the calculated $\Omega^{(1,1)*}$ for condensation nuclei in air. The parameters of potential (5) were determined with combination relationships (8) and (9) (see the first column in Table 1) and five pairs of ε_{ii} and σ_{ii} for water [13] (Table 2). The parameters of the interaction potential for air molecules were taken from the monograph [12]: $\sigma_{ij} = 3.617$ Å and $\varepsilon_{jj} = 97$ K. The spread in the values of $\Omega^{(1,1)*}$ did not exceed 0.5% in all cases.

Table 1. Interaction potential parameters, Ω integrals, and diffusion coefficients ($\times 10^{-5}$ cm^2/s) for condensation nuclei

Formula	σ_{ij}, Å	ε_{ij}/k, K	$\Omega^{(1,1)*}$	D	ΔD, %
(8)	3.13	222	1.040	1.997	10.9
(8)	3.02	274	1.040	1.997	10.9
(8)	3.10	279	1.043	1.992	10.7
(8)	3.18	159	1.038	2.001	11.2
(8)	3.10	192	1.040	1.996	10.9
(9)	3.16	208	1.041	1.995	10.8
(9)	3.07	249	1.040	1.997	11.0
(9)	3.13	259	1.042	1.993	10.7
(9)	3.21	151	1.038	2.001	11.2
(9)	3.13	179	1.039	1.999	11.1

Table 2. Interaction potential parameters for water molecules [13]

σ_{ii}, Å	ε_{ii}/k, K
2.71	506
2.52	755
2.65	800
2.80	260
2.65	380

The weak dependence of the Ω integrals on the parameters of the potentials of interaction between gas molecules and the dispersed particle is explained by the fact that potential (5) was obtained by averaging over a large ensemble of atoms in the particle. Such a weak sensitivity of the transport factors of nanoparticles to the parameters of the intermolecular potential is a very important finding, because experimental values of these parameters are often unreliable. Cases are known where the parameters of intermolecular potentials determined for the same material under similar conditions greatly diverge.

2. Dependence of the Diffusion Coefficient on the Radius of the Dispersed Particle

The dependence of the diffusion coefficient of dispersed particles on their radius is a very important characteristic. If the particles are coarse ($R \gg l$, where l is the free path of gas molecules), this dependence is defined by formula (1); that is, the diffusion coefficient is inversely proportional to the radius. In the other limiting case $R \ll l$, the diffusion coefficient is defined by

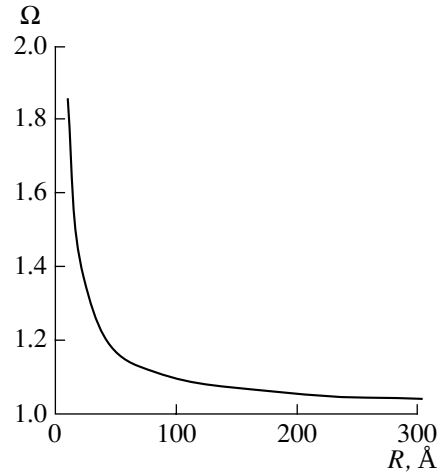


Fig. 1. $\Omega_{ij}^{(1,1)*}$ vs. particle radius R.

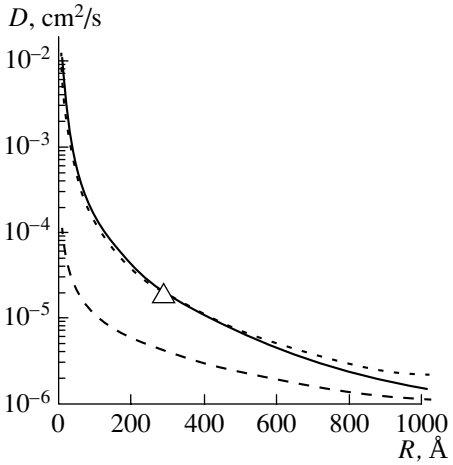


Fig. 2. Diffusion coefficient D vs. particle radius R. Solid curve, potential (5); dotted curve, Cunningham–Millikan formula; dashed line, Einstein formula; and Δ, data point [17].

the Epstein formula [14]

$$D_{\text{fm}} = \frac{3\sqrt{mkT/2\pi}}{8\rho R^2(1+\alpha\pi/8)}, \tag{10}$$

from which D_{fm} is inversely proportional to the particle radius squared. Here, m is the mass of a carrier gas molecule, ρ is the gas density, and α is the accommodation coefficient.

Formula (7) is much more complicated than (10), since the value of the Ω integral also varies with particle radius. This is shown in Fig. 1 for some fixed temperature. This dependence is well approximated by the relationship

$$\Omega_{ij}^{(1,1)*} \sim 1 + \frac{a_1}{\sqrt{R}} + \frac{a_2}{R}, \tag{11}$$

where a_1 and a_2 are temperature-dependent constants. For $R > 5 \times 10^2$ Å, the Ω integrals become virtually independent of the radius and their values come close to unity.

Thus, neither Einstein formula (1) nor formula (10) give correct values for the diffusion coefficient of nanoparticles. By way of example, the dependence of this coefficient on the radius at $T = 288$ K is shown in Fig. 2 for condensation nuclei in air under atmospheric pressure. Here, the solid curve meets formula (7) and the dashed curve reflects the Einstein formula (1). As was expected, the discrepancy is especially noticeable for fine dispersed particles. For example, for clusters with a characteristic size of about 10 Å, the diffusion coefficient given by (1) is two orders of magnitude lower than that defined by (7).

3. Temperature Dependence of the Diffusion Coefficient

It should be noted that the diffusion coefficient of nanoparticles, like that of molecules and Brownian particles, depends considerably on the temperature. Under free molecular conditions, it depends on the square root of the temperature, as in (10). Taking the viscosity coefficient in Einstein formula (1) in the Chapman–Enskog form [11], we arrive at

$$D_{\text{E}} = \frac{8\sigma^2}{15R}\sqrt{\frac{kT}{m\pi}}\Omega_i^{2,2*}(T^*), \tag{12}$$

where σ is the effective diameter of a molecule. Formula (12) involves the Ω integral for Lennard-Jones potential 12–6. Its temperature dependence can be approximated by the power relationship[1]

$$\Omega_i^{(2,2)*}(T^*, 12\text{–}6) = 1.16145/T^{*0.14874}.$$

Then, the Einstein diffusion coefficient takes the form

$$D_{\text{E}} = 0.61944\frac{\sigma^2}{R}\sqrt{\frac{k}{m\pi}}T^{0.35}. \tag{13}$$

It is easy to check from (7) and (13) that

$$\frac{D}{D_{\text{E}}} \sim \frac{1}{R\Omega_{ij}^{(1,1)*}(9-3)\Omega_{ii}^{(2,2)*}(12-6)} \tag{14}$$

at a given density of the carrier gas. This ratio depends on the temperature nonlinearly. The discrepancy is the most pronounced for fine particles when the Ω integrals are far from unity.

Figure 3 demonstrates the temperature dependence of the diffusion coefficient for condensation nuclei of a fixed radius ($R = 300$ Å) in air at atmospheric pressure. Again, formula (7) is depicted by the solid curve; Einstein formula (1), by the dashed curve. In the latter case,

[1] This approximation is satisfactory when temperatures are not too low ($T^* > 1$). The exact relationship is more complex [15].

the temperature dependence of the viscosity for air was determined from tabulated data.

COMPARISON WITH EXPERIMENTAL DATA

Comparison of calculated and experimental data is a necessary test for the applicability for any theory. Our analytical results were compared with available experimental data for the diffusion of condensation nuclei in air [17]. The experiment was carried out at 288 K and atmospheric pressure (p = 101325 Pa). It was found that R = 285 Å and $D_e = 1.8 \times 10^{-5}$ cm^2/s, where R is the size of the aerosol particle and D_e is the experimental value of the diffusion coefficient.

Diffusion coefficient (7) with potential (5) was calculated with combination relationships (8) and (9) and five pairs of σ_{ii} and ε_{ii} for water [13] (Table 2). The parameters of the interaction potential for air molecules were taken from monograph [12]: σ_{jj} = 3.617 Å, ε_{jj} = 97 K, and m_a = 29 g/mol. The density of water, its molecular weight κ_w, and the effective volume of a water molecule V_w were set equal to 1, 18.0153 g/mol, and 29.916 Å^3, respectively. The parameters of the interaction potentials determined with combination relationships (8) and (9), Ω integrals, and diffusion coefficients for the condensation nuclei are listed in Table 1. The last column of Table 1 displays the relative discrepancy between our analytical results and the experimental data [17], $\Delta D = (D - D_e)100\%/D_e$.

The aggregate discrepancy, which is the sum of the error of measurement and the processing error, is about 10% in all cases. In Figs. 2 and 3, the experimental value of the diffusion coefficient (triangle) is applied on the analytical curve. The agreement between the calculation and the experiment is seen to be good.

Formula (1) for the diffusion coefficient can be applied directly only to liquids and sufficiently dense gases. To extend the applicability of formula (1), Cunningham suggested the following modification:

$$D_k = kT/\gamma_k,$$
$$\gamma_k = 6\pi\eta R\left(1 + A\frac{l}{R} + Q\frac{l}{R}e^{-bR/l}\right)^{-1}. \quad (15)$$

The constants appearing in (15) were found in the Millikan experiments [16]: A = 0.864, Q = 0.29, and b = 1.25. Below, formula (15) will be called the Cunningham–Millikan formula. This formula is an experimental correlation. In this sense, the comparison of our results with those obtained by this formula is effectively a comparison with experimental findings. Under our conditions, the Knudsen number is relatively high; thus it could be hoped that this formula would give adequate results. Figure 2 shows the results obtained with this formula under the conditions of our experiment (dotted line). The Cunningham–Millikan formula does correct formula (1) considerably and fits our data well nearly throughout the range considered. The discrepancy (5–7%) is observed only at large and very small radii.

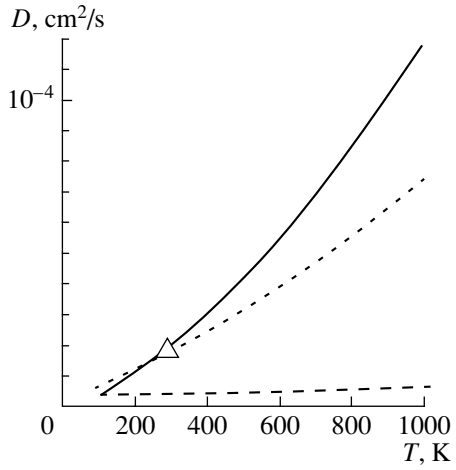


Fig. 3. Diffusion coefficient D vs. temperature T. Designations are the same as in Fig. 2.

From physical considerations, it follows that the diffusion coefficient must depend substantially on the gas temperature. In Fig. 3, the diffusion coefficient of the aerosol particle varies several tens of times as the temperature grows from 100 to 1000 K. On the other hand, according to the Einstein formula, the diffusion coefficient varies with temperature insignificantly (Fig. 3, dashed curve). This should also be taken into account in constructing experimental correlations. In the general case, the coefficients A, Q, and b are temperature-dependent. Unfortunately, Millikan, in his classical experiments, determined these parameters in a very narrow temperature range (19–24°C). Accordingly, the temperature range where formula (15) applies is narrow. The dotted line in Fig. 3 is obtained with this formula. It agrees with our data at temperatures between 100 and 300 K. At higher temperatures, the Cunningham–Millikan formula strongly underestimates the diffusion coefficient.

DISCUSSION

Our model of the interaction potential is straightforward and is, therefore, very attractive. Its basic assumptions are as follows.

(1) Molecule–surface interaction is potential and adiabatic and can be described in classical terms.

(2) The potential of gas molecule–particle atom interaction is pair and additive.

(3) The effect of the surface structure is ignored.

(4) Thermal oscillations of the particle surface are not taken into account.

The question arises as to how serious these limitations are.

The motion of electrons and nuclei is of quantum nature; hence, the problem of finding the intermolecular potential is reduced, strictly speaking, to solving the Schrödinger equation for interacting molecules. For most systems, this problem can be solved only approximately. It can be simplified by separating the nuclear and electron motions and introducing the concept of adiabatic potentials. This so-called adiabatic approximation is based on a great difference in the masses of electrons and nuclei and allows researchers to study the motion of electrons when the nuclei are at rest. With such an approach, the Schrödinger equation does not contain the operator of nucleus kinetic energy and the nucleus coordinates are used as fixed parameters. As a result, one can find the system energy as a function of nucleus spacing. This energy is taken as the potential energy of nuclei and is often referred to as the adiabatic potential. Knowing this potential, one can study the behavior of an ensemble of interacting molecules.

The principle of separation of the electron and nuclear motions is not entirely incorrect when the electron terms come closer together and completely fails when the electron states become degenerate. Therefore, one should be especially careful when applying the adiabatic approximation to excited electron states of a system. Moreover, the adiabatic approximation may be invalid in high-energy collisions, where the operator of nucleus kinetic energy can no longer be neglected: at high velocities of atoms, the electrons have no time to accommodate the instantaneous nuclear configuration adiabatically.

The applicability of the adiabatic approximation is specified by the Massey condition

$$\xi_{ij} = \omega_{ij}\frac{a}{v} \gg 1, \qquad (16)$$

where ω_{ij} is the frequency of electron–electron transitions, v is the atom velocity, and a is the length over which the adiabatic wave function considerably changes. With the condition (16) met, the collision time a/v far exceeds the time of electron–electron transitions ω_{mn}^{-1}. This condition is usually satisfied in dispersed liquids at moderate temperatures. At high collision energies, the adiabatic approximation must fail. However, available experimental data suggest that the molecular interaction is well described in terms of the adiabatic approximation up to energies of several kiloelectronvolts [18].

The model potential derived in this work can be applied in both classical and quantum-mechanical calculations. The limitations of the classical approach refer not to the potential as such but to its applicability within the classical transport theories. The classical approach is valid if the de Broglie wavelength $\lambda = 2\pi\hbar/p$ is small in comparison with the characteristic spatial scale of the problem. It is easy to check that this condition is fulfilled if the temperatures are not too low.

Since the gas molecule–surface atom interaction potential is assumed to be pair and additive, a carrier gas, generally speaking, must be sufficiently dilute. However, it should be noted in this respect that the Lennard-Jones potential used in the calculations is effective; hence, the nonadditivity of the interaction is also taken into account (at least partially).

With the effect of thermal lattice vibrations and surface structure taken into consideration, the quantitative results may differ substantially. The potential model suggested is thus only a zero approximation to what we would like to have. However, the results obtained with this model may fit the experimental data well. This is because the potential is effective in essence owing to the specific definition of its constants.

The comparison between the calculated diffusion coefficient of the nanoparticles and the experimental data suggests that the potential can successively be used for studying the transport of nanoparticles in dilute gases. In this work, the temperature dependence of the diffusion coefficient of ultradispersed particles in a dilute gas was investigated virtually for the first time. Neither the Einstein formula (1) nor Cunningham–Millikan correlation (15) describes this dependence correctly. To be more specific, the latter applies in a very narrow temperature range near room temperature where Millikan determined the constants in formula (15). Formula (10), using several constants found experimentally, also gives incorrect values.

It should be noted that the majority of today's studies of ultradisperse aerosols apply indirect methods to find the dispersed particle size. As a rule, the diffusion coefficient (or the mobility) of dispersed particles is first found and then their radius is calculated with formula (15). It can be concluded that such an approach to finding the particle size yields error in the results. The error grows sharply with carrier gas temperature. Direct measuring methods, for example, with an electron microscope, are therefore necessary.

Some comments on the applicability domain of formula (7) for potential (5) follow. When studying transport processes in gases, one should consider two similarity parameters: the Knudsen number within the particle radius, $Kn_R = l/R$, and the Knudsen number away from a particle, Kn_∞. In the latter case, the Knudsen number is defined as the ratio of the molecule free path to the characteristic linear size of the system (such as the transverse size of a flowed-around body, thickness of the boundary layer, characteristic scale of the macroscopic parameter gradient, etc.). The Einstein formula is valid for a continuous medium, where $Kn_\infty \ll 1$ and $Kn_R \ll 1$.

The theory put forward in this work is strictly applicable if $Kn_\infty \ll 1$ or at least <1. In this case, the Boltzmann equations can be solved by the Chapman–Enskog method, the result of which is formula (7). In addition, a particle must be finely dispersed such that the condi-

tion

$$R \sim \sqrt{\epsilon} l \ll l \tag{17}$$

is fulfilled. Here, ϵ is the van der Waals parameter for a rarefied gas. Otherwise, the dynamics of particle–molecule collisions cannot be treated in terms of binary collisions [1–3]. In this case, it is necessary to consider many-body correlations and formula (7) may be modified to the form

$$D = \frac{3}{16}\frac{\sqrt{2\pi\mu kT}}{n\mu\pi R^2 g\Omega_{ij}^{(1,1)^*}(T^*, \sigma_{ij}/R)}, \tag{18}$$

where g is some correlation function.

As the rarefaction of a carrier gas increases ($Kn_\infty \gtrsim 1$ or $\gg 1$), formula (7) becomes still more invalid. However, its applicability can be extended by introducing the correction

$$\mathcal{D} = D[1 + B(1 - \exp\{-\beta Kn_\infty\})Kn_\infty], \tag{19}$$

where B and β are some temperature-dependent functions and D is the diffusion coefficient defined by formula (7) or formula (18). It is believed that formula (7) for potential (5), rather than Einstein formula (1), should be taken to describe the diffusion of nanoparticles in rarefied gases. The applicability domain for formulas (18) and (19) will be discussed later.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grants no. 00-15-96164 and 01-01-0045).

REFERENCES

1. V. Ya. Rudyak, G. V. Kharlamov, and A. A. Belkin, Pis'ma Zh. Tekh. Fiz. **26** (13), 29 (2000) [Tech. Phys. Lett. **26**, 553 (2000)].
2. V. Ya. Rudyak, G. V. Kharlamov, and A. A. Belkin, Teplofiz. Vys. Temp. **31**, 283 (2001).
3. V. Ya. Rudyak, Pis'ma Zh. Tekh. Fiz. **18** (20), 77 (1992) [Sov. Tech. Phys. Lett. **18**, 681 (1992)].
4. M. Yu. Gladkov and V. Ya. Rudyak, Zh. Tekh. Fiz. **64** (4), 170 (1994) [Tech. Phys. **39**, 441 (1994)].
5. M. Yu. Gladkov and V. Ya. Rudyak, Izv. Akad. Nauk, Mekh. Zhidk. Gaza, No. 2, 165 (1994).
6. V. Ya. Rudyak and S. L. Krasnolutskiĭ, Preprint No. 3(13)-98, NGASU (Novosibirsk, 1998).
7. V. Y. Rudyak and S. L. Krasnolutskii, in *Proceedings of the 21st International Symposium on Rarefied Gas Dynamics, Marseilles, 1998*, Vol. 1, p. 263.
8. F. O. Goodman and H. Y. Wachman, *Dynamics of Gas-Surface Scattering* (Academic, New York, 1976; Mir, Moscow, 1980).
9. R. G. Barantsev, *Interaction of Rarefied Gases with Streamline Surfaces* (Nauka, Moscow, 1975).
10. W. Eckstein, *Computer Simulation of Ion-Solid Interactions* (Springer-Verlag, Berlin, 1991; Mir, Moscow, 1995).
11. J. H. Ferziger and H. G. Kaper, *Mathematical Theory of Transport in Gases* (North-Holland, Amsterdam, 1972; Mir, Moscow, 1976).
12. J. O. Hirschfelder, C. F. Curtiss, and R. B. Bird, *Molecular Theory of Gases and Liquids* (Wiley, New York, 1954; Inostrannaya Literatura, Moscow, 1961).
13. L. Monchick and A. Mason, J. Chem. Phys. **35**, 1676 (1961).
14. L. S. Epstein, Phys. Rev. **23**, 710 (1924).
15. P. D. Neufeld, A. R. Janzen, and R. A. Aziz, J. Chem. Phys. **57**, 1100 (1972).
16. R. A. Millikan, Phys. Rev. **22**, 1 (1923).
17. J. J. Nolan and V. H. Guerrini, Trans. Faraday Soc. **32**, 1175 (1936).
18. I. G. Kaplan, *Theory of Intermolecular Interactions* (Nauka, Moscow, 1982; Elsevier, New York, 1986).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 7, 2002, pp. 814–820. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 21–27.

THEORETICAL AND MATHEMATICAL PHYSICS

Electromagnetic Field Induced by Shock Compression of a Current-Carrying Conductor

S. D. Gilev* and T. Yu. Mikhaĭlova**

** Lavrent'ev Institute of Hydrodynamics, Siberian Division, Russian Academy of Sciences, pr. Akad. Lavrent'eva 15, Novosibirsk, 630090 Russia*

e-mail: Gilev@hydro.nsc.ru

*** Novosibirsk State University, ul. Pirogova 2, Novosibirsk, 630090 Russia*

Received November 8, 2001

Abstract—The electromagnetic field induced by shock compression of a current-carrying conductor is shown to consist of two current waves. One propagates in the uncompressed material at the shock-wave velocity. The other is due to current inward diffusion. As the shock wave propagates, the current passes from the first wave to the second. At large observation periods, the situation resembles conventional current diffusion into a conducting half-space. Control parameters for electrodynamic problems with shock waves are found. Their physical meaning is the ratio between the times of field convection and diffusion in different regions. In specific cases, the problem is reduced to the motion of the surface of a current-carrying half-space and to shock metallization of an insulator. *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Shock compression of a current-carrying conductor is a basic principle of operation of pulsed electromagnetic systems (explosive magnetic cumulation generators, MHD generators, and current switches) [1]. Also, this effect is used in shock-wave experiments to measure electrical conductivity, magnetization, and mass velocity [2, 3]. The study of electrodynamic processes accompanying shock compression of conductors would help to improve the efficiency of related energy devices and gain new information on the material state in extreme conditions.

Upon shock compression of a material in a magnetic field, the resulting electromagnetic pattern is governed by two, diffusion and convective, mechanisms of magnetic field modification. When moving, the shock front induces currents (specific current waves) in the conductor. The structure of the current waves depends on the electrodynamic conditions and shock compression parameters. The electrodynamics problem aimed at finding the electromagnetic field due to material shock compression has been known since the late 1950s [4–7]. It has become of still greater importance in recent years owing to the development of novel dynamic methods for diagnosing the material state at high energy densities [8–17]. With these methods, closed electrodynamic models of shock compression are constructed to discover unknown material properties. Experimental data are used as an input for solving inverse boundary-value problems of shock compression electrodynamics. Here, the exact solutions of direct electrodynamic problems are of special value. In the quasi-stationary approximation, the direct problems are reduced to the equation of diffusion with nonstationary boundaries, which is hard to solve. One of the boundaries of the conducting region is the shock-wave front; therefore, correct statement of the conditions at the shock front is needed.

In this work, we analyze the structure of the electromagnetic field resulting from shock compression of a current-carrying conductor by extending a number of known problems. The solution is based on a method that is rather general and can be helpful in solving other physical problems.

STATEMENT OF THE PROBLEM AND ELECTROMAGNETIC MODEL

In our consideration, we assume that (1) the shock wave is nonrelativistic, i.e., the mass velocity u of the material is much less than the velocity of light c; (2) the shock wave is stationary and the mass velocity u and the wave velocity D remain constant during wave propagation; (3) the magnetic field is sufficiently weak and has no effect on shock-wave propagation; (4) the shock front width is negligibly small; (5) the conductivity of the metal experiences a step at the shock front from σ_1 to σ_2; and (6) the conductor is made of a nonmagnetic material.

These assumptions reasonably describe many shock-wave experiments [2, 3]. Indeed, in metals, the typical velocity of shock waves generated by chemical explosives is about 5 km/s, which is much smaller than the velocity of light. The stationarity condition is readily fulfilled if the appropriate scale of the laboratory experiments is chosen. The shock-wave pressure in metals under these conditions is above 10 GPa, which

1063-7842/02/4707-0814$22.00

usually far exceeds the magnetic field pressure. The shock wave front width in metals is on the order of 10 μm. Finally, it has been shown [17] that the step-conductivity model fits the experimental data adequately.

Consider the simple 1D geometry. A plane shock wave enters into the conducting space with a given current passing over its surface at the time instant $t = 0$. The coordinate system is related to the surface of the half-space (Fig. 1). Figures 1 and 2 refer to the uncompressed and compressed material regions, respectively. Our goal is to find the magnetic field $B(x, t)$ in the half-space $x \geq 0$ when the shock wave propagates in the material ($t > 0$).

In the quasi-stationary approximation ($u \ll c$), the 1D Maxwell equations and generalized Ohm's law give the following set of equations:

$$\frac{\partial B_1}{\partial t} - u\frac{\partial B_1}{\partial x} - \frac{1}{\mu_0\sigma_1}\frac{\partial^2 B_1}{\partial x^2} = 0, \quad (D-u)t \leq x < \infty, \quad (1)$$

$$\frac{\partial B_2}{\partial t} - \frac{1}{\mu_0\sigma_2}\frac{\partial^2 B_2}{\partial x^2} = 0, \quad 0 \leq x \leq (D-u)t, \quad (2)$$

$$B_1(x, 0) = 0, \quad x \geq 0, \quad (3)$$

$$B_2(0, t) = B_0, \quad (4)$$

$$B_1((D-u)t, t) = B_2((D-u)t, t), \quad (5)$$

$$\frac{1}{\mu_0\sigma_1}\frac{\partial B_1}{\partial x} + uB_1\Big|_{x=(D-u)t} = \frac{1}{\mu_0\sigma_2}\frac{\partial B_2}{\partial x}\Big|_{x=(D-u)t}. \quad (6)$$

Here, Eqs. (1) and (2) describe the magnetic field diffusion in the uncompressed and compressed regions, respectively. Equation (1) includes the convective term describing the motion of the material with the velocity u. Equation (3) represents the initial condition, and Eq. (4) specifies the boundary condition. Equations (5) and (6) reflect the continuity of the magnetic and electric fields at the shock-wave front, which follows from the boundary conditions for the shear component of the related vectors [18]. In fact, the shock front is assumed to be indefinitely thin; hence, it does not contain the current. Equation (6), of electric-field continuity, combines the differential Ampere law and generalized Ohm's law.

The set of Eqs. (1)–(6) specifies the boundary-value problem for the magnetic field $B(x, t)$. In the absence of the shock wave, the problem is reduced to conventional diffusion of a field into a half-space [1]. With the shock wave present, the situation becomes more complicated. Behind the shock front, the material is in the compressed state and the size of the compressed region grows monotonically. Moreover, compression changes the conductivity of the material (the limit cases are insulator–metal and metal–insulator transitions). The problem stated by (1)–(6) is of a rather general character. Its solution must describe the electromagnetic pattern in different physical situations. For this reason, it is hoped that general parameters controlling the electrodynamics of material shock compression will be found.

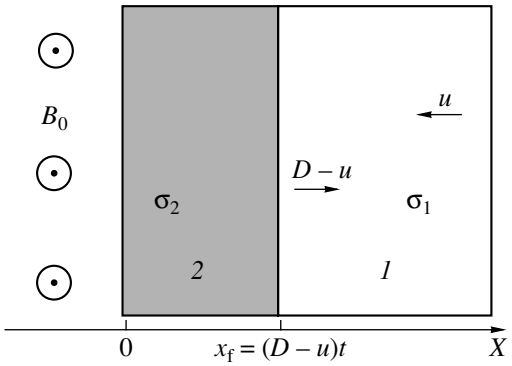


Fig. 1. Entering of a plane shock wave into a conducting half-space with surface current. The reference system is related to the free boundary of the half-space. (*1*) Uncompressed material; (*2*) compressed material. The moving position of the shock front is $x_f = (D-u)t$.

SOLUTION METHOD

The theory of solving boundary-value problems of heat conduction with nonstationary boundaries has a rich history. It dates back to the classical work [19] and its subsequent criticism [20, 21]. Later, the special analytical methods developed in [22–27] were extended in [28, 29]. The Green function method, the method of thermal potentials, the method of series, and the method of functional transformation have become the most popular. A basically new approach to solving the boundary-value problems with nonstationary boundaries, which can be called the method of extension, was suggested in [13]. Its essence is in that an initial range of variables (x, t), which is defined by the shock front kinematics, is transformed into an extended range of a simpler form. As an extended range, the quadrant $x \geq 0$, $t \geq 0$ has turned out to be the most convenient. For this quadrant with $t = 0$, a solution to the boundary-value problem can be constructed with the Green function. The unknown function at $t = 0$ is determined from the condition at the moving boundary. The solution thus obtained has physical meaning in the initial range of variables (x, t). Once work [13] had been published, associated solutions for a linearly extending range with given edge fields were found by the Laplace–Carson method of functional transformation [22] and by the Green function method [26]. A comparison between the three different solutions of this problem for the case of stationary edge fields [13, 22, 26] shows that the method of extension offers an advantage as applied to similar physical problems. Subsequently, the method of extension has been used to analyze electromagnetic effects in double-line systems [14] and the behavior of a shock wave entering into a half-space [16]. This method is partly employed in the solution given below [13, 14, 16].

SOLUTION TO THE PROBLEM

Let us reduce problem (1)–(6) to the problem with zero boundary conditions. To do this, we introduce the new functions B_1' and B_2' ($B_1 = B_0 B_1'$, $B_2 = B_0(1 + B_2')$) and put $B_0 = 1$ without loss of generality. Then, we arrive at two boundary-value problems for the functions $B_1'(x, t)$ and $B_2'(x, t)$:

$$\frac{\partial B_1'}{\partial t} - u\frac{\partial B_1'}{\partial x} - \frac{1}{\mu_0\sigma_1}\frac{\partial^2 B_1'}{\partial x^2} = 0, \quad (D-u)t \le x < \infty,$$

$$B_1'(x, 0) = 0, \quad x \ge 0, \tag{7}$$

$$\frac{1}{\mu_0\sigma_1}\frac{\partial B_1'}{\partial x} + uB_1'\big|_{x=(D-u)t} = \mu(t),$$

$$\frac{\partial B_2'}{\partial t} - \frac{1}{\mu_0\sigma_2}\frac{\partial^2 B_2'}{\partial x^2} = 0, \quad 0 \le x \le (D-u)t,$$

$$B_2(0, t) = 0, \tag{8}$$

$$\frac{1}{\mu_0\sigma_2}\frac{\partial B_2'}{\partial x}\bigg|_{x=(D-u)t} = \mu(t),$$

which are related by the conditions

$$B_1'((D-u)t, t) - B_2'((D-u)t, t) = 1 \tag{9}$$

at the mobile boundary. Here, $\mu(t)$ is an auxiliary function. Problems (7) and (8) will be referred to as problems I and II, respectively. The solution to problem I is given by

$$B_1'(x', t) = \exp\left(-\frac{\mu_0\sigma_1 D^2 t}{4} - \frac{\mu_0\sigma_1 D x'}{2}\right)$$

$$\times\left[-\frac{\sqrt{\mu_0\sigma_1}}{\sqrt{\pi}}\int_0^t \frac{\mu(\tau)}{\sqrt{t-\tau}}\exp\left(\frac{\mu_0\sigma_1 D^2\tau}{4} - \frac{\mu_0\sigma_1 x'^2}{4(t-\tau)}\right)d\tau\right. \tag{10}$$

$$+ h\int_0^t \mu(\tau)\exp\left(\frac{\mu_0\sigma_1 D^2\tau}{4}\right)\exp\left(hx' + \frac{h^2(t-\tau)}{\mu_0\sigma_1}\right)$$

$$\left.\times\operatorname{erfc}\left(\frac{\sqrt{\mu_0\sigma_1}x'}{2\sqrt{t-\tau}} + h\sqrt{\frac{t-\tau}{\mu_0\sigma_1}}\right)d\tau\right],$$

where $x' = x - (D-u)t$, $h = (D/2 - u)\mu_0\sigma_1$, $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x)$, and $\operatorname{erf}(x)$ is the error function.

Based on the method of extension, we seek a solution to problem II in the form

$$B_2'(x, t) = \frac{1}{\sqrt{\pi t}}\exp\left(-\frac{\mu_0\sigma_2 x^2}{4t}\right)\int_0^\infty f\left(\frac{\sqrt{\mu_0\sigma_2}(D-u)\tau}{2}\right) \tag{11}$$

$$\times\sinh\left(\frac{\sqrt{\mu_0\sigma_2}x\tau}{2t}\right)\exp\left(-\frac{\tau^2}{4t}\right)d\tau.$$

In view of (8) and (9), we come to the following set of equations for the functions $\mu(t)$ and $f(t)$:

$$-\frac{D-u}{2\sqrt{\pi t}}\exp\left(-\frac{\mu_0\sigma_2(D-u)^2 t}{4}\right)\int_0^\infty f\left(\frac{\sqrt{\mu_0\sigma_2}(D-u)\tau}{2}\right)$$

$$\times\sinh\left(\frac{\sqrt{\mu_0\sigma_2}(D-u)\tau}{2}\right)\exp\left(-\frac{\tau^2}{4t}\right)d\tau + \frac{1}{2\sqrt{\pi\mu_0\sigma_2}t^{3/2}} \tag{12}$$

$$\times\exp\left(-\frac{\mu_0\sigma_2(D-u)^2 t}{4}\right)\int_0^\infty f\left(\frac{\sqrt{\mu_0\sigma_2}(D-u)\tau}{2}\right)$$

$$\times\cosh\left(\frac{\sqrt{\mu_0\sigma_2}(D-u)\tau}{2}\right)\tau\exp\left(-\frac{\tau^2}{4t}\right)d\tau = \mu(t),$$

$$\exp\left(-\frac{\mu_0\sigma_1 D^2 t}{4}\right)\left[-\frac{\sqrt{\mu_0\sigma_1}}{\sqrt{\pi}}\int_0^t \frac{\mu(\tau)}{\sqrt{t-\tau}}\exp\left(\frac{\mu_0\sigma_1 D^2\tau}{4}\right)d\tau\right.$$

$$+ h\int_0^t \mu(\tau)\exp\left(\frac{\mu_0\sigma_1 D^2\tau}{4} + \frac{h^2(t-\tau)}{\mu_0\sigma_1}\right)$$

$$\left.\times\operatorname{erfc}\left(h\sqrt{\frac{t-\tau}{\mu_0\sigma_1}}\right)d\tau\right] = \frac{1}{\sqrt{\pi t}}\exp\left(-\frac{\mu_0\sigma_2(D-u)^2 t}{4}\right) \tag{13}$$

$$\times\int_0^\infty f\left(\frac{\sqrt{\mu_0\sigma_2}(D-u)\tau}{2}\right)\sinh\left(\frac{\sqrt{\mu_0\sigma_2}(D-u)\tau}{2}\right)$$

$$\times\exp\left(-\frac{\tau^2}{4t}\right)d\tau + 1.$$

By applying the Laplace transformation to set (12)–(13) and eliminating the Laplace transform from the function $\mu(t)$, we obtain

$$-\frac{u}{D-u}\frac{F_1(p)}{p} + F_2(p) + \sqrt{s}\frac{\sqrt{p^2+k^2}-p}{p}F_1(p)$$

$$+\sqrt{s}F_1(p) = \frac{2u-D}{D-u}\frac{1}{p^2-1} - \frac{\sqrt{p^2+k^2}-p}{p^2-1}\sqrt{s} \tag{14}$$

$$-\frac{p}{p^2-1}\sqrt{s}.$$

Here, $F_1(p) = L[f(t)\sinh t]$ and $F_2(p) = L[f(t)\cosh t]$ are the Laplace transforms and

$$s = \frac{\sigma_2}{\sigma_1}, \quad k^2 = \frac{D^2}{(D-u)^2}\frac{\sigma_1}{\sigma_2} - 1.$$

Taking the inverse Laplace transform from (14), we arrive at the integro-differential equation for the function $f(t)$:

$$\frac{d}{dt}[f(t)(\cosh t + \sqrt{s}\sinh t)] - \frac{u}{D-u}f(t)\sinh t$$

$$+ \sqrt{s}k\int_0^t f(\tau)\sinh\tau\frac{J_1(k(t-\tau))}{t-\tau}d\tau = \frac{2u-D}{D-u}\cosh t \quad (15)$$

$$-\sqrt{s}\sinh t - \sqrt{s}k\int_0^t \cosh\tau\frac{J_1(k(t-\tau))}{t-\tau}d\tau$$

for $f(0) = -\sqrt{s}$.

With the function $f(t)$ found, the magnetic field $B_2(x, t)$ is given by

$$B_2(x,t) = B_0 + B_0\frac{2}{\sqrt{\pi}}\exp\left(-\frac{R_2\varphi^2}{4}\right) \times \int_0^\infty f(\sqrt{R_2}\tau)\sinh(\sqrt{R_2}\varphi\tau)\exp(-\tau^2)d\tau, \quad (16)$$

where

$$\varphi = \frac{x}{(D-u)t}, \quad R_2 = \mu_0\sigma_2(D-u)^2t.$$

Assuming that the magnetic field at the moving boundary is found, we can state the problem for the uncompressed material. It involves Eqs. (1), (3), and (5). The last equation assumes that the edge field is already known.

The magnetic field in the uncompressed material has the form

$$B_1(\xi,t) = B_0\frac{2}{\sqrt{\pi}}\exp\left(-\frac{R_2}{4}\right)\exp(-\sqrt{R_1}\xi)\int_0^\infty f(\sqrt{R_2}\tau) \times \sinh(\sqrt{R_2}\tau)\exp(-(\tau+\xi)^2)G(\xi,t,\tau)d\tau + \frac{B_0}{2}\exp(-2\sqrt{R_1}\xi)\operatorname{erfc}\left(\xi - \frac{\sqrt{R_1}}{2}\right) + \frac{B_0}{2}\operatorname{erfc}\left(\xi + \frac{\sqrt{R_1}}{2}\right), \quad (17)$$

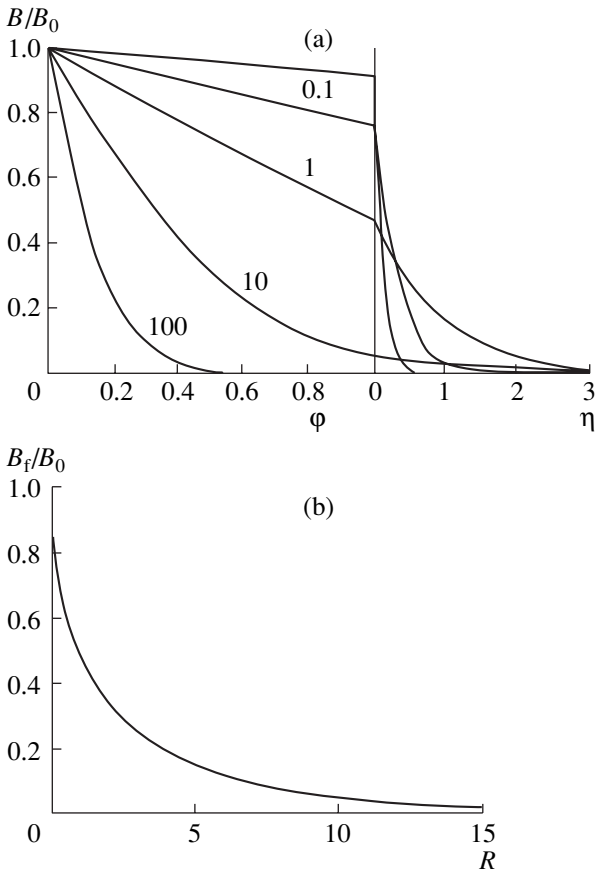


Fig. 2. Magnetic field distribution when the shock wave enters into the conducting half-space ($u/D = 0.5$, $s = 4$, $R_1 = R_2 = R$). (a) Magnetic field distribution in the conductor: $0 \le \varphi \le 1$, compressed material; $\eta \ge 0$, uncompressed material (figures on the curves indicate the values of R); the shock front position corresponds to $\varphi = 1$ and $\eta = 0$. (b) Magnetic field at the shock front vs. R.

where

$$G(\xi,t,\tau) = \frac{2}{\sqrt{\pi}}\int_0^\infty \exp\left(-\left(\Theta - \frac{\xi\tau}{\Theta}\right)^2\right) \times \exp\left(-\frac{(R_1-R_2)\xi^2}{4(\xi^2+\Theta^2)}\right)d\Theta,$$

$$\xi = \frac{x'}{2}\sqrt{\frac{\mu_0\sigma_1}{t}}, \quad R_1 = \mu_0\sigma_1D^2t. \quad (18)$$

Formulas (16)–(18) and also Eq. (15) for the function $f(t)$ yield a solution to the problem stated by (1)–(6).

Figure 2 shows the distributions of the magnetic field for $R_1 = R_2 = R$ and $u/D = 0.5$. For the uncompressed material, the spatial variable $\eta = \mu_0\sigma_1Dx'$ was used. Figure 2b demonstrates the dependence of the magnetic field at the shock-wave front on the parameter R.

The value of the field is descriptive of the ratio of the currents in the compressed and uncompressed material.

If the time of shock propagation is small ($R \ll 1$), the thickness of the compressed region is also small and the magnetic field at the front differs from that on the surface only slightly. In this case, almost the entire current passes in the uncompressed region before the front. As the shock wave propagates, the compressed region thickness grows and the magnetic field at the front decreases. The current passes from the uncompressed region to the compressed one. At large observation times ($R \gg 1$), the field at the front tends to zero and the current occupies only a part of the compressed region.

Such a current pattern may be treated as the current redistribution between two current waves. One propagates through the uncompressed material at the shock-front velocity. The maximal thickness of this wave is given by the condition $\eta \approx 1$. The other wave is a diffusion wave in the compressed material. Its thickness grows continuously. For short times, the current is uniformly distributed over the compressed material. For long times, the shock front breaks away from the second current wave, so that there is no current at the front. It is essential that the value and the spatial distribution of the resulting currents strongly vary with time.

The electromagnetic field pattern is readily illustrated by the asymptotes to the solution and specific cases of the problem.

ASYMPTOTES TO THE SOLUTION

The asymptotes to the solution obtained are easy to find. We will take advantage of the properties of the function $f(t)$, which follows from (15):

$$f(t) = -\sqrt{\frac{\sigma_2}{\sigma_1}} + \left(\frac{2u - D}{D - u} + \frac{\sigma_2}{\sigma_1}\right)t + o(t) \quad \text{at} \quad t \longrightarrow 0,$$

$$\lim_{t \to \infty} f(t) = -1.$$

For short times ($R_1 \longrightarrow 0$, $R_2 \longrightarrow 0$), the field in the uncompressed material is

$$B_1(x', t) \approx B_0\left[\operatorname{erfc}\left(\frac{\eta}{2\sqrt{R_1}}\right) - \sqrt{\frac{\sigma_2}{\sigma_1}}\frac{\sqrt{R_2}}{\sqrt{\pi}} \times \exp\left(-\frac{\eta}{2} - \frac{\eta^2}{4R_1}\right) + \sqrt{\frac{\sigma_2}{\sigma_1}}\frac{\sqrt{R_2}}{\sqrt{R_1}}\exp\left(-\frac{\eta}{2}\right)\frac{\eta}{2} \times \operatorname{erfc}\left(\frac{\eta}{2\sqrt{R_1}}\right)\right]. \quad (19)$$

With $\eta \gg \sqrt{R_1}$, this yields

$$B_1(\eta, R_1) \approx B_0\frac{2\sqrt{R_1}}{\sqrt{\pi}\eta}\exp\left(-\frac{\eta^2}{4R_1}\right);$$

that is, the field rapidly drops with the distance from the shock front. In the compressed material, the field is given by

$$B_2(x, t) \approx B_0\left(1 - \sqrt{\frac{\sigma_2}{\sigma_1}}\frac{\sqrt{R_2}}{\sqrt{\pi}}\varphi\right). \quad (20)$$

From (19) and (20), one finds the magnetic field B_f at the shock front:

$$B_f(t) \approx B_0\left(1 - \sqrt{\frac{\sigma_2}{\sigma_1}}\frac{\sqrt{R_2}}{\sqrt{\pi}}\right).$$

From the last two formulas, it follows that the field in the compressed material varies with the coordinate linearly and the field at the front diminishes with time rather slowly: $\approx 1 - \alpha\sqrt{t}$, where $\alpha = \text{const}$.

For long times ($R_1 \longrightarrow \infty$, $R_2 \longrightarrow \infty$), the asymptotes have the form

$$B_1(\eta, R_1) \approx \frac{B_0}{2}\operatorname{erfc}\left(\frac{\sqrt{R_1}}{2} + \frac{\eta}{2\sqrt{R_1}}\right) - \frac{B_0}{2}\exp(-\eta)\operatorname{erfc}\left(\frac{\sqrt{R_1}}{2} - \frac{\eta}{2\sqrt{R_1}}\right),$$

$$B_2(x, t) \approx B_0\left(1 - \operatorname{erf}\left(\frac{x}{2}\sqrt{\frac{\mu_0\sigma_2}{t}}\right)\right).$$

It is easy to see that the field $B_1(\eta, t)$ at the front tends to zero. The expression for the field $B_2(x, t)$ in the compressed material is the solution to the classical problem of diffusion into a conducting half-space [1].

SPECIFIC CASES

Let $R_1 = R_2$; then, $k = 0$ and the function $f(t)$ can be found from (15) in explicit form:

$$f(t) = \frac{1}{\cosh t + \sqrt{s}\sinh t} \times \left(-\sqrt{s}q(t) + q(t)\int_0^t \frac{(\sqrt{s} - 2)\cosh\tau - \sqrt{s}\sinh\tau}{q(\tau)}d\tau\right),$$

$$q(t) = \exp\left(\frac{\sqrt{s}t}{\sqrt{s} + 1}\right)(\cosh t + \sqrt{s}\sinh t)^{-\frac{1}{\sqrt{s} + 1}}.$$

In this case, $G(\xi, t, \tau) = 1$, which simplifies the use of (17).

Let $\sigma_1 \longrightarrow 0$ and σ_2 be finite. Then, $R_1 \longrightarrow 0$, the function $f(t)$ can be found, and the solution to the problem is

$$B_1(x, t) = 0,$$

$$B_2(x, t) = B_0\frac{2}{\sqrt{\pi}}\exp\left(-\frac{\varphi^2 R_2}{4}\right)$$
$$\times\int_0^\infty \frac{\sinh[(1-\varphi)\sqrt{R_2}\eta]}{\sinh(\sqrt{R_2}\eta)}\exp(-\eta^2)d\eta.$$

The latter formula, derived in [13], illustrates the structure of the magnetic field for the insulator–metal transition in a shock wave.

Now let $\sigma_2 \longrightarrow 0$ and σ_1 be finite. Then, $R_2 \longrightarrow 0$ and the solution to the problem is

$$B_1(\eta, R_1) = \frac{B_0}{2}\exp(-\eta)\operatorname{erfc}\left(\frac{\eta}{2\sqrt{R_1}} - \frac{\sqrt{R_1}}{2}\right)$$
$$+\frac{B_0}{2}\operatorname{erfc}\left(\frac{\eta}{2\sqrt{R_1}} + \frac{\sqrt{R_1}}{2}\right), \quad (21)$$
$$B_2(x, t) = B_0.$$

Formula (21) is the solution to the problem of motion of the current-carrying half-space surface. This problem is akin to the well-known problem of motion in a medium with a heat source [19–21]. A correct solution to the thermal problem is given in [21]. Formula (21) differs in that it uses the variables η and R_1 and, thus, has clear physical meaning. It is easy to check that formula (21) is reduced to

$$B_1(\eta) \approx B_0\exp(-\eta) \quad (22)$$

at long times ($t \longrightarrow \infty$, $R_1 \longrightarrow \infty$, $\eta/2\sqrt{R_1} \longrightarrow 0$).

Thus, for $t \longrightarrow \infty$, the magnetic field distribution before the moving surface becomes stationary. In this case, the magnetic field varies (and the associated current concentrates) largely in a conductor layer of thickness $x_* = 1/\mu_0\sigma_1 D$. For short times, the effective thickness of the current layer is less than x_*. The magnetic field distribution (22) becomes stationary at $R_1 \approx 10$.

Note in conclusion that the solution to (16)–(18) offers sufficient generality and can be helpful in solving well-known problems [13, 19–21] in specific cases.

PHYSICAL MEANING OF CONTROL PARAMETERS

The solution found depends on the parameters R_1 and R_2. The former can be represented as the ratio of two times: $R_1 = t/\mu_0\sigma_1 x_*^2$. The time t characterizes convective motion. The quantity $\mu_0\sigma_1 x_*^2$ is the time of electromagnetic relaxation in a conducting layer of thickness x_*. The thickness x_* and the related relaxation time are characteristic parameters in electrodynamic problems with a moving (nonstationary) boundary. When a shock wave propagates in an infinite conductor in a field, the current concentrates within a layer of thickness x_* before the front [6, 7]. When the wave enters into the conductor, the thickness x_* is the limiting thickness of the current-carrying layer [16]. When the surface of a half-space moves, the size of the region with current is also x_* for $t \longrightarrow \infty$ [see formula (22)].

Let the magnetic field before the front be kept constant. Then, for $R_1 \ll 1$, the magnetic field distribution in the uncompressed material is nonstationary. If $R_1 \gg 1$, the distribution becomes stationary in the front-related reference system. The distribution becomes stationary for $t_1 \approx 10\mu_0\sigma_1 x_*^2 \approx 10/\mu_0\sigma_1 D^2$. If the magnetic field at the wave front changes, it will take a time on the order of t_1 for the magnetic field distribution to "accommodate" this change.

The parameter R_2 can be represented as $R_2 = \mu_0\sigma_1 x_c^2/t$. Here, the numerator is the diffusion time in the compressed material ($x_c = (D-u)t$) and the denominator is the time of convective motion. Thus, the parameter R_2 is also the ratio of two times, the times of diffusion and convection. For $R_2 \ll 1$, diffusion prevails in the compressed material and the magnetic field distribution is near-equilibrium. If $R_2 \gg 1$, the effect of diffusion is significant only in the region adjacent to the free boundary of the half-space and the currents near the front are absent. The condition for the separation of the shock front from the diffusion wave ($R_2 \gg 1$) gives an estimate for the time t_2: $t_2 \approx 10/\mu_0\sigma_2(D-u)^2$. Within this time, the magnetic field at the shock front diminishes considerably in our statement.

For a continuous metal (without phase transitions), t_1 and t_2 are of the same order of magnitude. In the shock-wave experiment [2, 3], the typical value of t_1 was $\simeq$10 ns for copper and $\simeq$300 ns for constantan. Two time ranges can be distinguished. For short times ($t < t_1$, $t < t_2$), the magnetic field distribution is nonequilibrium in the uncompressed material but is equilibrium in the compressed state. For long times ($t > t_1$, $t > t_2$), the situation is reversed. In the latter case, the magnetic field at the shock front is close to zero and the entire current passes in the compressed region (in our statement).

The importance of the parameters R_1 and R_2 is that they have clear physical meaning and allow for a simpler representation of the nonstationary problem. Owing to their generality, these parameters characterize solutions to the class of electrodynamic problems involving shock waves. The parameter R_2 has appeared in the problem with one conducting region [13]; both parameters were present in the two-region problem [16]. The solution to specific electrodynamic problems depends on the initial and boundary conditions. At the same, these problems have mutual control parameters and characteristic times specified by them.

It should be noted that our approach can be extended for the processes of heat transfer and molecular diffusion, where nonstationary-boundary problems are frequently encountered. In these cases, one can also determine the (thermal or molecular) skin depth and the related diffusion time. The use of the associated dissipation factor, instead of the conductivity, will then suffice to find the control parameters R_1 and R_2. In the general case, the parameters R_1 and R_2 can be written in the form $R_1 = V_1^2 t/\chi_1$ and $R_2 = (V_1 - V_2)^2 t/\chi_2$, where V_1 and V_2 are the velocities of the front and back boundaries, respectively, and χ_i is the diffusion coefficient in the associated region. In this case, the meaning of the parameters R_1 and R_2 remains the same (the convection-to-diffusion time ratio). The use of these parameters in nonstationary-boundary problems makes it possible to simplify the solution of nonstationary heat conduction problems and make them physically more meaningful.

CONCLUSION

The method of extension of the operating domain allows one to find the electromagnetic field upon shock compression of a current-carrying conductor in the sufficiently general statement of the problem. The specific cases of our analysis are the problem of motion of the half-space surface [19–21], the problem of shock-induced insulator–metal transition [13], and the problem of classical current diffusion into the half-space. It is important that the analysis reveals mutual parameters and characteristic times in the class of electrodynamic problems with shock waves. This is essential for developing new experimental techniques aimed at studying the material state at high energy densities and in high-energy electromagnetic systems. Our approach can be helpful in treating diffusion and other thermal problems with nonstationary boundary.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 99-02–16807).

REFERENCES

1. H. Knoepfel, *Pulsed High Magnetic Fields* (North-Holland, Amsterdam, 1970; Mir, Moscow, 1972).
2. L. V. Al'tshuler, Usp. Fiz. Nauk **85**, 197 (1965) [Sov. Phys. Usp. **8**, 52 (1965)].
3. *The Physics of High Energy Density*, Ed. by P. Caldirola and H. Knoepfel (Academic, New York, 1971; Mir, Moscow, 1974).
4. W. Marshall, Proc. R. Soc. London, Ser. A **233**, 367 (1955).
5. H. Sen, Phys. Rev. **102**, 5 (1956).
6. J. M. Burgers, in *Magnetohydrodynamics: A Symposium*, Ed. by R. K. M. Landshoff (Stanford Univ. Press, Stanford, 1957; Atomizdat, Moscow, 1958).
7. E. I. Zababakhin and M. N. Nechaev, Zh. Éksp. Teor. Fiz. **33**, 442 (1957) [Sov. Phys. JETP **6**, 345 (1958)].
8. J. N. Fritz and J. A. Morgan, Rev. Sci. Instrum. **44**, 215 (1973).
9. Yu. N. Zhugin and K. K. Krupnikov, Prikl. Mekh. Tekh. Fiz. **20** (1), 102 (1983).
10. V. V. Paĭ, I. V. Yakovlev, and G. E. Kuz'min, Fiz. Goreniya Vzryva **32** (2), 124 (1996).
11. Yu. N. Zhugin and Yu. L. Levakova, Prikl. Mekh. Tekh. Fiz. **41** (6), 199 (2000).
12. V. V. Paĭ, Ya. L. Luk'yanov, I. V. Yakovlev, and G. E. Kuz'min, Fiz. Goreniya Vzryva **36** (6), 164 (2000).
13. S. D. Gilev and T. Yu. Mikhaĭlova, Zh. Tekh. Fiz. **66** (5), 1 (1996) [Tech. Phys. **41**, 407 (1996)].
14. S. D. Gilev and T. Yu. Mikhaĭlova, Zh. Tekh. Fiz. **66** (10), 109 (1996) [Tech. Phys. **41**, 1029 (1996)].
15. S. D. Gilev, Fiz. Goreniya Vzryva **31** (4), 109 (1995).
16. S. D. Gilev and T. Yu. Mikhaĭlova, Fiz. Goreniya Vzryva **36** (6), 153 (2000).
17. S. D. Gilev, Fiz. Goreniya Vzryva **37** (2), 121 (2001).
18. I. E. Tamm, *The Principles of Electricity Theory* (Nauka, Moscow, 1989).
19. V. L. Shevel'kov, Zh. Tekh. Fiz. **16**, 207 (1946).
20. B. Ya. Lyubov, Zh. Tekh. Fiz. **18**, 713 (1948).
21. G. A. Grinberg, Zh. Tekh. Fiz. **21**, 382 (1951).
22. B. Ya. Lyubov, Dokl. Akad. Nauk SSSR **57**, 551 (1947).
23. D. V. Redozubov, Zh. Tekh. Fiz. **30**, 606 (1960) [Sov. Phys. Tech. Phys. **5**, 570 (1960)].
24. V. I. Kval'vasser and Ya. F. Rutner, Dokl. Akad. Nauk SSSR **156**, 1273 (1964).
25. É. M. Kartashov and G. M. Bartenev, Izv. Vyssh. Uchebn. Zaved., Fiz., No. 2, 70 (1969).
26. É. M. Kartashov and G. M. Bartenev, Izv. Vyssh. Uchebn. Zaved., Fiz., No. 3 (82), 20 (1969).
27. É. M. Kartashov and G. M. Bartenev, Izv. Vyssh. Uchebn. Zaved., Fiz., No. 12, 97 (1970).
28. É. M. Kartashov and B. Ya. Lyubov, Izv. Akad. Nauk SSSR, Énerg. Transp., No. 6, 83 (1974).
29. É. M. Kartashov, *Analytical Methods in the Theory of Heat Conductivity of Solids* (Vysshaya Shkola, Moscow, 1985).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 7, 2002, pp. 821–828. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 28–35.

THEORETICAL AND MATHEMATICAL PHYSICS

Different Methods of Calculating the Pinning Energy of Plane Vortices in a 3D Josephson Medium: A Comparative Study

M. A. Zelikman

St. Petersburg State Technical University,
ul. Politekhnicheskaya 29, St. Petersburg, 195251 Russia

Received December 4, 2001

Abstract—The pinning energy of plane (laminar) vortices in a 3D Josephson medium is calculated within a continuous vortex model considering functions of two types: $V = 1 - \cos\varphi$ and $V = 2/\pi^4\varphi^2(2\pi - \varphi)^2$. The shape and energy of the stable and unstable vortices are found with an algorithm for the exact numerical solution of a set of difference equations. The vortex magnetic and Josephson energies diverge. The magnetic and Josephson components of the pinning energy are close in magnitude but differ in sign; as a result, the total pinning energy is smaller than its components by one order of magnitude. This result is confirmed analytically. An analytical computing method within the continuous vortex model is suggested. This method preserves the difference terms in the energy expression. The magnetic energy found by this method differs from the Josephson energy in magnitude, and the magnetic component of the pinning energy is opposite in sign to the Josephson component. Comparative analysis of the approximate approaches to energy calculation within the continuous vortex model when the difference terms are retained and when they are replaced by derivatives is performed. It is shown that the continuous vortex model gives incorrect values of the Josephson and magnetic components of the pinning energy. The actual values are several tens or several hundreds of times higher than those obtained with the continuous vortex model. Yet, since the Josephson and magnetic components of the pinning energy have different signs, the exact value of the total pinning energy and the approximate value obtained within the continuous vortex model differ insignificantly. *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Recent studies of high-temperature superconductors have shown that vortex structures are to a great extent responsible for processes arising in the material. In particular, vortex pinning and interaction specify the processes taking place when a high-temperature superconductor is placed in a magnetic field. The motion of vortices results in non-Joule heat evolution in the sample. The critical currents and fields can be increased (a fundamental problem in physics and technology) if the vortex behavior, structure, and interaction, as well as pinning mechanisms and forces, are adequately understood.

In [1], plane (laminar) vortices in a 3D-ordered Josephson medium representing a cubic lattice consisting of superconducting wires were considered. Each connection of the lattice contains a Josephson contact. In such a structure, the energy of the current configuration per unit length (1 m) is given by

$$E = \varepsilon_0 \sum_{k=-\infty}^{\infty} \left[\frac{1}{2}(\varphi_{k+1} - \varphi_k)^2 + I(1 - \cos\varphi_k)\right], \quad (1)$$

where φ_k are the phase jumps at the contacts, I is the pinning parameter, and ε_0 is the normalizing constant. Hereafter, the energies are expressed in terms of ε_0. The first term in (1) describes the vortex magnetic energy; the second, the energy of Josephson contacts.

The vortex equilibrium states, which correspond to the energy extrema, are described by a set of difference equations that follows from the condition $\dfrac{\partial E}{\partial \varphi_m} = 0$:

$$\varphi_{m+1} - 2\varphi_m + \varphi_{m-1} = I\sin\varphi_m. \quad (2)$$

A system of equations similar to (2) also describes vortices in a long periodically modulated Josephson contact [2, 3].

Nonlinear system (2) characterizes the behavior of one type of soliton. Therefore, great interest has been displayed in this set by physicists and mathematicians and many interesting results have been found. This set was first obtained by Frenkel' and Kontorova when considering the behavior of edge dislocations in a crystal [4]. Because of its discrete nature, the system allows researchers to analyze the soliton motion with allowance for pinning. The pinning force was defined as the finite energy needed for the vortex center to be displaced from one cell to another.

The exact solution to nonlinear finite-difference system (2) is impossible to find analytically. When the parameter I is small ($I < 1$), the distribution becomes

quasi-continuous and system (2) transforms into the differential equation

$$\frac{d^2\varphi}{\partial x^2} = I\sin\varphi. \quad (3)$$

The vortex Josephson, E_J, and magnetic, E_H, energies obtained within the quasi-continuous approximation (when I tends to zero) approach their exact values. However, the use of such an approximation for finding the pinning energy is questionable. The point is that the Josephson and magnetic components of the pinning energy, which equal the differences between the associated energies in different vortex positions, have a much higher order of smallness in I and may be beyond the approximation error. It will be shown below that the components of the pinning energy vary as $\exp(-1/I)$, which makes our doubts still more justified.

In this work, Eq. (3) is used to compare different approximate approaches to evaluating the energy values by comparing results obtained in various approximations with the exact solution to the problem stated.

Mathematically, the discrete nature of the problem follows from (i) discrete summation in energy expression (1); (ii) the presence of difference terms in expressions under the summation sign, e.g., $(\varphi_{k+1} - \varphi_k)^2$ in (1); and (iii) finding the vortex shape with finite-difference system (2).

The replacement of summation by integration, difference terms by derivatives, and finite-difference system (2) by differential equation (3) would mean the total abandonment of discreteness. Then, the Josephson, E_J, and magnetic, E_H, energies would equal each other and both components of the pinning energy would vanish.

In [5], another method of calculating the vortex pinning energy was proposed. Here, summation in (1) is retained but the term $1/2(\varphi_{k+1} - \varphi_k)^2$ is replaced by $1/2(\varphi')^2$ and the vortex shape is found from differential equation (3). With this method, the pinning energy was calculated from Eq. (1) in work [3]. The values of E_J and E_H turned out to equal each other and were given by the same expressions as in the continuous approach. As for the pinning energy components, they were exactly equal to each other, being other than zero.

In this work, we suggest an algorithm for exact energy computation that meets all points (i)–(iii). The magnetic and Josephson energies differ. The magnetic and Josephson components of the pinning energy are close in magnitude and differ in sign; thereby, the total pinning energy is one order of magnitude smaller than its components. In addition, we use an analytical computational method that retains the first two discretenesses (summation and differences terms) but assumes the vortex shape is found from a differential equation. The magnetic energy E_H is smaller than the Josephson energy E_J. The magnetic and Josephson components of the pinning energy are opposite in sign.

VORTEX PINNING EVALUATION

1. Analytical Calculation Based on the Continuous Vortex Model with Summation and Differences Terms Taken into Account

Let the stationary vortex energy per unit length and height (1 m) be expressed as [5]

$$E = E_0 \sum_{k=-\infty}^{\infty} \left[\frac{1}{2}(\varphi_{k+1} - \varphi_k)^2 + IV(\varphi_k)\right], \quad (4)$$

where E_0 is the normalizing constant. Hereafter, the energies are expressed in terms of E_0.

The equilibrium states of the vortex, which correspond to its energy extrema, are given by the set of finite-difference equations

$$\varphi_{m+1} - 2\varphi_m + \varphi_{m-1} = IV'(\varphi_m). \quad (5)$$

We will seek the vortex pinning energy at $I \ll 1$ by expanding the δ functions into a Fourier series [5]:

$$\sum_{k=-\infty}^{\infty} \delta(x-k) = \sum_{n=-\infty}^{\infty} \exp(i2\pi nx) = \mathrm{Re}\sum_{n=-\infty}^{\infty} \exp(i2\pi nx). \quad (6)$$

For this purpose, we represent Eq. (4) in the form

$$E = \sum_{k=-\infty}^{\infty} f(\varphi_k) = \int_{-\infty}^{\infty} f(\varphi(x+\alpha)) \sum_{k=-\infty}^{\infty} \delta(x-k)dx = \mathrm{Re}\int_{-\infty}^{\infty} f(\varphi(x)) \sum_{n=-\infty}^{\infty} \exp(i2\pi n(x-\alpha))dx, \quad (7)$$

where α is the coordinate of the vortex center relative to the cell edge. If, for example, the vortex is centered on the boundary between two cells, $\alpha = 0$; if the vortex is at the center of a cell, $\alpha = 0.5$.

If the function $f(x)$ is even, expression (7) takes the form

$$E = 2\sum_{n=-\infty}^{\infty} \cos(2\pi n\alpha)\mathrm{Re}\int_0^{\infty} f(\varphi(x))\exp(i2\pi nx)dx. \quad (8)$$

In (8), a solution to the differential equation

$$\frac{d^2\varphi}{dx^2} = IV'(\varphi) \quad (9)$$

is used as $\varphi(x)$. Equation (9) replaces discrete equation (5). For the boundary conditions $\varphi \xrightarrow[x\to-\infty]{} 0$, $\varphi \xrightarrow[x\to+\infty]{} 2\pi$,

and $\varphi' \xrightarrow[x \to \pm\infty]{} 0$ (which correspond to a single vortex), Eq. (9) yields

$$(\varphi')^2 = 2IV(\varphi), \tag{10}$$

from whence the solution is found.

This solution corresponds to a vortex in a continuous medium. Below, it will be shown that the use of the approximate solution that ignores the discreteness of the medium in calculating the pinning energy is an assumption that somewhat corrupts the real situation. In [5], however, one more assumption is made (which is omitted in this work). Namely, it is suggested to retain summation and, at the same time, replace the term $1/2(\varphi_{k+1} - \varphi_k)^2$ by $1/2(\varphi')^2$. In this case, as follows from (10), the terms in (4) equal each other and the calculation by (8) at $f = 2IV$ would suffice to find the total energy. As was already noted, this assumption is omitted in this work, which greatly improves the accuracy of the results, as comparison with the computer calculation shows (Tables 1, 2). In this case, the first and second terms in (4) make different contributions to the vortex total energy, as well as to the pinning energy.

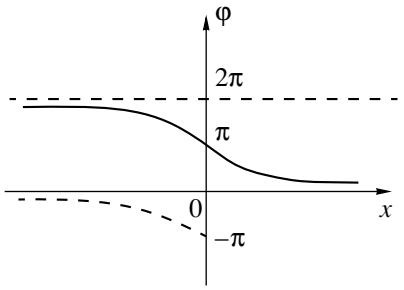
Two forms of the function $V(\varphi)$ will be considered: $V = 1 - \cos\varphi$ and $V = 2/\pi^4\varphi^2(2\pi - \varphi)^2$. The former was analyzed in [6]. It corresponds to the behavior of plane vortices in a 3D Josephson medium [1] and in a long periodically modulated Josephson contact [2, 3]. Therefore, the contributions from the first and second terms in (4) to the total energy will be called the magnetic energy and Josephson energy, respectively.

The latter form is of independent interest, since it also has been analyzed in a number of works. On the other hand, it is reasonable to check the generality of results obtained as applied to different forms of the function $V(\varphi)$ and to estimate the effect of the function type on the energy values, especially on the values of the pinning energy components. The function $V = 2/\pi^4\varphi^2(2\pi - \varphi)^2$ approximates the function $V = 1 - \cos\varphi$ in the range $0 \le \varphi \le 2\pi$, since φ_k for a single vortex varies precisely in this interval. Both functions are symmetric about $\varphi = \pi$ and have equal roots ($\varphi = 0$ and $\varphi = 2\pi$) and maxima ($V = 2$ at $\varphi = \pi$). However, in spite of their similarity, the functions give greatly differing vortex total energies and, particularly, pinning energies. The generality of rules found for both functions is therefore all the more convincing.

The case $V = 1 - \cos\varphi$. The solution to Eq. (10) for a single vortex has the form

$$\varphi(x) = 4\arctan(\exp(-x\sqrt{I})). \tag{11}$$

The plot of function (11) is depicted in the figure.



Solution (11) to Eq. (10) for a single vortex.

Substituting (11) into (8) and leaving the terms with $n = 0, \pm 1$ yields

$$E = \frac{2}{\sqrt{I}}\int_0^\infty f(4\arctan(\exp(-\xi)))d\xi + \frac{4}{\sqrt{I}}\cos(2\pi\alpha)\,\mathrm{Re}\int_0^\infty f(4\arctan(\exp(-\xi)))\exp\left(\frac{i2\pi\xi}{\sqrt{I}}\right)d\xi, \tag{12}$$

where $\xi = x\sqrt{I}$.

Consider the Josephson energy $E = I\sum_{k=-\infty}^{\infty}(1 - \cos\varphi_k)$. In this case,

$$f_1 = I(1 - \cos(4\arctan(e^{-\xi}))) = \frac{2I}{\cosh^2\xi}. \tag{13}$$

Taking into account that f_1 is even with respect to ξ and using formula (3.985) from [7], we find

$$E_{J0} = 4\sqrt{I}\int_0^\infty \frac{d\xi}{\cosh^2\xi} = 4\sqrt{I}\lim_{b\to 0}\int_0^\infty \frac{\cos(b\xi)d\xi}{\cosh^2\xi} = 4\sqrt{I}\lim_{b\to 0}\frac{b\pi}{2\sinh(b\pi/2)} = 4\sqrt{I}, \tag{14}$$

$$E_{J1} = 8\sqrt{I}\cos(2\pi\alpha)\int_0^\infty \frac{\cos(2\pi\xi/\sqrt{I})d\xi}{\cosh^2\xi} = \frac{8\pi^2}{\sinh(\pi^2/\sqrt{I})} \times \cos(2\pi\alpha) \approx 16\pi^2\exp\left(-\frac{\pi^2}{\sqrt{I}}\right)\cos(2\pi\alpha). \tag{15}$$

The energy E_{J1} is maximal when $\alpha = 0$, that is, when the vortex is centered on the boundary between two cells.

Now, consider the magnetic energy $E_H = 0.5\sum_{k=-\infty}^{\infty}(\varphi_k - \varphi_{k+1})^2$. In this case,

$$f_2 = 8(\arctan(e^{-(x+\alpha)\sqrt{I}}) - \arctan(e^{-(x+\alpha+1)\sqrt{I}}))^2 = 8\arctan^2\frac{e^{-\alpha\sqrt{I}} - e^{-(\alpha+1)\sqrt{I}}}{e^{-x\sqrt{I}} - e^{-(x+2\alpha+1)\sqrt{I}}}. \tag{16}$$

Table 1. Magnetic and Josephson energies for $V = 1 - \cos\varphi$ found by different ways

I	Theory (13)	Theory (19)	Numerical model (25) and (26)		Exact solution	
	E_{J0}	E_{H0}	E_{H0}	E_{J0}	E_H	E_J
0.15	1.5492	1.5428	1.5428	1.5492	1.5525	1.5393
0.40	2.5298	2.5022	2.5022	2.5298	2.5452	2.4860
0.70	3.347	3.2836	3.2837	3.347	3.3853	3.2393
1.00	4.000	3.894	3.894	4.000	4.073	3.806

At $\alpha_1 = -0.5$, the function

$$f_2 = 8\arctan^2\frac{\sinh(\sqrt{I}/2)}{\cosh(x\sqrt{I})} \tag{17}$$

is even. Expanding the function arctan$x$ into a series in x, squaring it, and integrating termwise yields

$$E_{H0} = \frac{8}{\sqrt{I}}\int_{-\infty}^{\infty}\arctan^2\frac{\sinh(\sqrt{I}/2)}{\cosh\xi}d\xi = \frac{16\eta}{\sqrt{I}}\left[1 - \frac{4}{9}\eta + \frac{184}{675}\eta^2 - \frac{704}{3675}\eta^3 + \ldots\right], \tag{18}$$

where $\eta = \sinh^2\dfrac{\sqrt{I}}{2}$.

Going from the series in η to that in $\sqrt{I}$, we obtain

$$E_{H0} = 4\sqrt{I} - \frac{1}{9}I^{3/2} + \frac{7}{1350}I^{5/2} + \ldots. \tag{19}$$

To calculate E_{H1}, we will use the methods of the theory of complex variables [6]:

$$E_{H1} = -16\exp(-\pi^2/\sqrt{I})\cos(2\pi\alpha)(2.43766 - I/12). \tag{20}$$

The vortex total energy is

$$E = E_{j0} + E_{H0} + (E_{j1} + E_{H1})\cos 2\pi\alpha. \tag{21}$$

By the pinning energy E_p, we mean the difference between the maximal and minimal vortex energies. The energy is maximal when the vortex is centered on the boundary between two cells ($\alpha = 0$) and minimal when the it is concentric with a cell ($\alpha = 0.5$):

$$E_p = E(0) - E(0.5) = \Delta E_J + \Delta E_H. \tag{22}$$

According to the analysis performed, the quantities ΔE_J and ΔE_H are given by

$$\Delta E_J = 32\pi^2\exp\left(-\frac{\pi^2}{\sqrt{I}}\right), \tag{23}$$

$$\Delta E_H = -32\exp\left(-\frac{\pi^2}{\sqrt{I}}\right)(2.43766 - I/12). \tag{24}$$

Note that expressions (23) and (24) in their analytical form cannot be expanded in powers of I; hence, they could not be found by the method of successive approximations.

To check the formulas obtained, we performed an exact computer calculation of the quantities

$$E_J(\alpha) = I\sum_{k=-\infty}^{\infty}(1 - \cos\varphi_k), \qquad E_H(\alpha) = \frac{1}{2}\sum_{k=-\infty}^{\infty}(\varphi_{k+1} - \varphi_k)^2 \tag{25}$$

for $\alpha = 0$ and $\alpha = 0.5$. Here, φ_k is taken in the form of (11), which corresponds to a vortex in a continuous medium:

$$\varphi_k(\alpha) = 4\arctan(\exp(-k - \alpha)\sqrt{I}). \tag{26}$$

According to (22), the difference between the associated values for various α gives ΔE_J and ΔE_H, while their half-sum, according to (21), equals E_{J0} and E_{H0}.

The results thus obtained for several values of I, as well as the energies E_{J0}, E_{H0}, ΔE_J, and ΔE_H calculated, respectively, from (14), (19), (23), and (24), are listed in Tables 1 and 2.

The case $V = 2/\pi^4\varphi^2(2\pi - \varphi)^2$. To calculate the energies, one must substitute the solution to (10) for a single vortex into (8). The solution in this case has the form

$$\varphi(x) = \frac{2\pi}{1 + \exp(4x\sqrt{I}/\pi)}. \tag{27}$$

The plot of function (27) is shown in figure. Consider the energy $E = 2I\sum_{k=-\infty}^{\infty}\psi^2(2 - \psi)^2$, where $\psi = \varphi/\pi$. Then,

$$f_1 = 2I\psi^2(2 - \psi)^2 = \frac{2I}{\cosh^4(2x\sqrt{I}/\pi)}. \tag{28}$$

Taking into account the evenness of f_1 with respect to x and using formula (3.985) from [7], we arrive at

$$E_{J0} = 2\pi\sqrt{I}\int_0^{\infty}\frac{d\xi}{\cosh^4\xi} = 2\pi\sqrt{I}\lim_{b\to 0}\int_0^{\infty}\frac{\cos(b\xi)d\xi}{\cosh^4\xi} = 2\pi\sqrt{I}\lim_{b\to 0}\frac{b\pi(b^2 + 4)}{12\sinh(b\pi/2)} = \frac{4\pi\sqrt{I}}{3}, \tag{29}$$

$$E_{J1} = 8\sqrt{I}\cos(2\pi\alpha)\int_0^{\infty}\frac{\cos(2\pi\xi/\sqrt{I})d\xi}{\cosh^4(2\xi/\pi)} = \frac{\pi^4(\pi^4/I + 4)}{3\sinh(\pi^3/2\sqrt{I})}\cos(2\pi\alpha) \approx \frac{2\pi^8}{3I}\exp\left(-\frac{\pi^3}{2\sqrt{I}}\right)\cos(2\pi\alpha). \tag{30}$$

The energy E_{J1} is maximal at $\alpha = 0$, i.e., when the vortex is centered on the boundary between two cells.

Table 2. Pinning energy components for $V = 1 - \cos\varphi$ found by different ways

I	Theory (27)	Theory (24)	Numerical model (25) and (26)		Exact solution			Theory (22)–(24)
	ΔE_J	ΔE_H	ΔE_H	ΔE_J	ΔE_H	ΔE_J	E_p	E_p
0.15	2.705×10^{-9}	-6.6477×10^{-10}	-6.6479×10^{-10}	2.705×10^{-9}	-6.095×10^{-8}	6.617×10^{-8}	5.22×10^{-9}	2.04×10^{-9}
0.40	5.27×10^{-5}	-1.2849×10^{-5}	1.2852×10^{-5}	5.27×10^{-5}	-6.21×10^{-4}	7.14×10^{-4}	9.29×10^{-5}	3.99×10^{-5}
0.70	2.38×10^{-3}	-5.734×10^{-4}	-5.738×10^{-4}	2.38×10^{-3}	-1.86×10^{-2}	2.24×10^{-2}	3.90×10^{-3}	1.81×10^{-3}
1.00	1.63×10^{-2}	-3.897×10^{-3}	-3.902×10^{-3}	1.62×10^{-2}	-9.5×10^{-2}	1.21×10^{-1}	2.6×10^{-2}	1.24×10^{-2}

Consider the energy $E_H = 0.5\sum_{k=-\infty}^{\infty}(\varphi_k - \varphi_{k+1})^2$. In this case,

$$f_2 = 2\pi^2\left[\frac{1}{1+\exp[(x+\alpha)4\sqrt{I}/\pi]} - \frac{1}{1+\exp[(x+\alpha+1)4\sqrt{I}/\pi]}\right]^2. \quad (31)$$

The function $f_2(x)$ is even at $\alpha = -0.5$:

$$f_2 = 2\pi^2\frac{\sinh^2(2\sqrt{I}/\pi)}{(\cosh(4x\sqrt{I}/\pi)+\cosh(2\sqrt{I}/\pi))^2}. \quad (32)$$

Substituting (32) into (8), we obtain for $n = 0$ [7]

$$E_{H0} = \frac{\pi^3}{\sqrt{I}}\frac{\beta\cosh\beta - \sinh\beta}{\sinh\beta}, \quad (33)$$

where $\beta = \dfrac{2\sqrt{I}}{\pi}$.

Leaving the first terms in $\sqrt{I}$, we find

$$E_{H0} = \frac{4\pi\sqrt{I}}{3}\left(1 - \frac{4I}{15\pi^2} + \ldots\right). \quad (34)$$

To calculate E_{H1}, we will use methods of the theory of complex variable:

$$E_{H1} = \pi^2\cos(2\pi\alpha)\sinh^2\frac{2\sqrt{I}}{\pi} \times \mathrm{Re}\int_{-\infty}^{\infty}\frac{\exp(i2\pi x)}{\cosh^2((x+0.5)2\sqrt{I}/\pi)\cosh^2((x-0.5)2\sqrt{I}/\pi)}dx. \quad (35)$$

To obtain E_{H1} numerically, we will take the integral over the contour bounded by the real axis and the upper semicircle of radius R and then approach R to infinity. Expression (35) satisfies the moderate Jordan lemma; therefore, the integral over the larger semicircle tends to zero at $R \longrightarrow \infty$. The integrand inside the contour of integration has second-order poles at points

$$z = \pm0.5 + \frac{\pi^2}{4\sqrt{I}}(2k+1)i, \quad k = 0, 1, 2, 3\ldots. \quad (36)$$

The residues at the points $z = \pm0.5 + \dfrac{\pi^2}{4\sqrt{I}}i$ (at $k = 0$) play a dominant role in taking integral (35). Calculating the integral and taking into account the initial shift $\alpha = -0.5$ for reducing the resulting expression to a single frame of reference for α, we eventually obtain

$$E_{H1} = -\frac{2\pi^6}{I}\exp(-\pi^3/2\sqrt{I})\cos(2\pi\alpha), \quad (37)$$

where $\alpha = 0.5$ corresponds to the position of the vortex center at the center of the cell. The vortex total energy is

$$E = E_{j0} + E_{H0} + (E_{j1} + E_{H1})\cos2\pi\alpha. \quad (38)$$

The pinning energy, that is, the difference between the maximal ($\alpha = 0$, the vortex is centered on the boundary) and minimal ($\alpha = 0.5$, the vortex is concentric with the cell) energies of the vortex, is given by

$$E_\mathrm{p} = E(0) - E(0.5) = \Delta E_J + \Delta E_H. \quad (39)$$

According to the analysis performed,

$$\Delta E_{J1} = \frac{2\pi^4(\pi^4/I+4)}{3\sinh(\pi^3/2\sqrt{I})} \approx \frac{4\pi^8}{3I}\exp\left(-\frac{\pi^3}{2\sqrt{I}}\right), \quad (40)$$

$$\Delta E_{H1} = -\frac{4\pi^6}{I}\exp(-\pi^3/2\sqrt{I}). \quad (41)$$

To check the formulas obtained, we carried out an exact computer calculation of the quantities

$$E_J(\alpha) = \frac{2I}{\pi^4}\sum_{k=-\infty}^{\infty}\varphi_k^2(2\pi-\varphi_k)^2,$$
$$E_H(\alpha) = \frac{1}{2}\sum_{k=-\infty}^{\infty}(\varphi_{k+1}-\varphi_k)^2 \quad (42)$$

for $\alpha = 0$ and $\alpha = 0.5$. Here, φ_k is taken in form (31), which corresponds to a vortex in a continuous medium:

$$\varphi_k = \frac{2\pi}{1+\exp(4(k+\alpha)\sqrt{I}/\pi)}. \quad (43)$$

According to (39), the differences between the associated values obtained for various α give ΔE_J and ΔE_H,

Table 3. Magnetic and Josephson energies for $V = 2/\pi^4\varphi^2 \times (2\pi - \varphi)^2$ found using different methods

I	Theory (29)	Theory (34)	Numerical model (42) and (43)		Exact solution	
	E_{J0}	E_{H0}	E_{H0}	E_{J0}	E_H	E_J
0.15	1.6223	1.6158	1.6158	1.6223	1.6256	1.6123
0.40	2.6492	2.6210	2.6210	2.6492	2.6645	2.6046
0.70	3.5046	3.4400	3.4400	3.5046	3.5422	3.3978
1.00	4.1887	4.0798	4.0798	4.1888	4.258	3.999

while their half-sums, according to (38), equal E_{J0} and E_{H0}.

The results thus obtained for several values of I, as well as the energies E_{J0}, E_{H0}, ΔE_J, and ΔE_H calculated, respectively, from (29), (34), (40), and (41), are summarized in Tables 3 and 4.

2. Analytical Calculation Based on the Solution of Difference Equations

Let us evaluate the vortex total energy when it is in different positions relatively to the center of the cell, that is, at various α. To do this, we substitute the expression

$$f(\varphi(x)) = \frac{1}{2}((\varphi(x+1) - \varphi(x))^2 + IV(\varphi(x)) \quad (44)$$

into Eq. (8).

Taking the integrals by parts, we find for the terms of the sum in (8) ($n \neq 0$)

$$E_n(\alpha) = \frac{\cos(2\pi n\alpha)}{i2\pi n}\mathrm{Im}\int_{-\infty}^{\infty}((\varphi(x+1) - \varphi(x)) \quad (45)$$

$$\times(\varphi'(x+1) - \varphi'(x)) + I\varphi'(x)V'(\varphi(x)))\exp(i2\pi nx)dx.$$

Substituting $IV'(\varphi(x)) = \varphi(x+1) - 2\varphi(x) + \varphi(x-1)$ into this expression yields

$$E_n(\alpha) = \frac{\cos(2\pi n\alpha)}{i2\pi n}\mathrm{Im}\int_{-\infty}^{\infty}[\varphi'(x+1)(\varphi(x+1) - \varphi(x)) \quad (46)$$

$$-\varphi'(x)(\varphi(x) - \varphi(x-1))]\exp(i2\pi nx)dx.$$

Dividing (46) by two parts and changing the variables ($x \longrightarrow x+1$) in one of them, we come to

$$E_1(\alpha) = \frac{\cos(2\pi n\alpha)}{i2\pi n}\mathrm{Im}\left[(1 - \exp(i2\pi n)) \right. \quad (47)$$

$$\left. \times \int_{-\infty}^{\infty}[\varphi'(x+1)(\varphi(x+1) - \varphi(x))]dx\right] = 0.$$

Thus, the total energy of a vortex having the shape described by the exact solution of finite-difference system (5) does not depend on its position relative to the cells. In other words, the pinning energy equals zero. This rather unexpected result is due to the fact that expression (8) was derived under the assumption that the shape of the vortex is independent of its position. On the other hand, the solution to system (2) depends on the position of the vortex relative to the center of the cell. With this fact taken into consideration, the pinning energy would be other than zero. However, such a calculation is an independent mathematical problem, an approach to which has not been found so far.

From the aforesaid, the following important conclusion can be drawn: the total pinning energy is much smaller than its magnetic and Josephson components; that is, E_J and E_H are nearly equal in magnitude and opposite in sign. This conclusion will be supported by the exact computer solution to the problem.

3. Computer Calculation of the Pinning Energy

The foregoing consideration was based on functions (26) and (43), which correspond to the solutions to differential equation (10), approximating finite-difference system (5) at small I. To find the applicability limit for such an approach, the results obtained should be compared with the exact solution of system (5).

To obtain a numerical solution to system (5), we write it in the form of the recurrent relationship

$$\varphi_{m+1} = 2\varphi_m - \varphi_{m-1} + IV'. \quad (48)$$

Then, knowing the values of the phase discontinuity at two adjacent contacts, one can find the whole configuration. Let us consider a single (isolated) vortex. In the stable equilibrium state [1], the phase discontinuities at the contacts of the central cell are related as $\varphi_{-1} = 2\pi - \varphi_1$. Choosing a certain value of φ_1 and substituting $2\pi - \varphi_1$ and φ_1 into (48) as two successive values of φ_k, we start calculating the subsequent values of the phase discontinuity. From the vortex shape (see the figure), it follows that, if any of the next values is less than zero or larger than the preceding one, the starting value of φ_1 is, respectively, increased or decreased and the calculation starts anew. In this way, the whole set of φ_k values for a single vortex can be found, after which the energies can be calculated by formula (1).

A similar calculation can be carried out for the unstable vortex configuration. In this case, $\varphi_{-1} = \pi$ [1]; therefore, the values of π and φ_1 are taken as two successive values of φ_k and then the scheme of computation reverts the same as that previous.

Tables 1–4 list the computer-calculated Josephson and magnetic energies (the half-sums of the values for the stable and unstable configurations), as well as the pinning energies for both forms of $V(\varphi)$.

Table 4. Pinning energy components for $V = 2/\pi^4\varphi^2(2\pi - \varphi)^2$ found using different methods

I	Theory (40)	Theory (41)	Numerical model (42) and (43)		Exact solution			Theory (39)–(41)
	ΔE_J	ΔE_H	ΔE_H	ΔE_J	ΔE_H	ΔE_J	ΔE_p	ΔE_p
0.15	3.502×10^{-13}	-1.058×10^{-13}	-1.064×10^{-13}	3.508×10^{-13}	-2.08×10^{-11}	2.2×10^{-11}	1.16×10^{-12}	2.4×10^{-13}
0.40	7.269×10^{-7}	-2.174×10^{-7}	-2.174×10^{-7}	7.269×10^{-7}	-1.96×10^{-5}	2.15×10^{-5}	1.96×10^{-6}	5.10×10^{-7}
0.70	1.667×10^{-4}	-4.926×10^{-5}	-4.926×10^{-5}	1.667×10^{-4}	-2.69×10^{-3}	3.07×10^{-3}	3.85×10^{-4}	1.17×10^{-4}
1.00	2.436×10^{-3}	-7.113×10^{-4}	-7113×10^{-4}	2.436×10^{-3}	-2.76×10^{-2}	3.26×10^{-2}	5.04×10^{-3}	1.72×10^{-3}

4. Analysis of the Results

First of all, it should be noted that the magnetic and Josephson energies and also the pinning energy components obtained by the numerical computation based on expressions (25), (26), (42), and (43) (Tables 1–4, columns 4 and 5) agree well with those found with analytical formulas (Tables 1–4, columns 2 and 3). This supports the validity of the method for finding the sum of the series and the absence of calculation errors, indicating the high reliability of the results obtained.

Now, we pass to the analysis of the results. First, we will see how the preservation of the difference terms in the expression for the magnetic energy in the continuous vortex model influences the results. (i) In the case $V = I(1 - \cos\varphi)$, expressions (18) and (28) for the magnetic energies differ from the relationships $E_H = 4\sqrt{I}$ and $\Delta E_H = 32\pi^2\exp\left(-\frac{\pi^2}{\sqrt{I}}\right)$ obtained in [3]. While in the formula for E_H, small corrections are added, the value of ΔE_H is four times smaller and opposite in sign. At the same time, expressions (13) and (27) for the Josephson energies coincide with the formulas in [3]. (ii) In the case $V = 2/\pi^4\varphi^2(2\pi - \varphi)^2$, the slight discrepancy between the energies E_J and E_H and the great difference between the pinning energy components persist. The magnetic component turns out to be nearly 3.5 times smaller than the Josephson component in magnitude and also opposite in sign. (iii) The comparison of the results obtained for the two forms of $V(\varphi)$ shows that even such an insignificant (at first glance) change in the form of the function may considerably (by several thousands of times) alter the pinning energy, affecting the energies E_J and E_H only slightly.

Now, let us contrast the results of the approximate approaches with the true values of the energies obtained in the computer calculation.

(1) The exact computer calculation of the vortex energy (Tables 1–4, columns 6 and 7) showed that the magnetic and Josephson energies differ and the magnetic and Josephson components of the pinning energy are close in magnitude but opposite in sign. As a result, the total pinning energy is one order of magnitude smaller than its components.

The theoretical analysis [expression (47)] indicates that the total energy of the vortex having the shape described by finite-difference system (5) (on the assumption that the shape is invariant under shift) is independent of its position relative to the cell. In other words, the pinning energy equals zero. This means that the Josephson and magnetic components of the pinning energy are equal in magnitude and opposite in sign. The total pinning energy is other than zero, being much smaller than both its components, if the shape of the vortex depends on its position. Such a result is yielded by the exact computer calculation.

Thus, the difference between the Josephson and magnetic components of the vortex energy, as well as the opposite signs of the Josephson and magnetic components of the pinning energy and their rough equality, can be considered specific features of the given problem.

(2) The comparison of the two approximate approaches using the continuous vortex model leads us to the following conclusions (in view of the above features).

When the difference term $1/2(\varphi_{k+1} - \varphi_k)^2$ in the expression for the magnetic energy is replaced by $1/2(\varphi')^2$, as was done in [5], the magnetic and Josephson energies, as well as the magnetic and Josephson components of the pinning energy, turn out to equal each other.

With the difference terms in the expression for the magnetic energy retained, the magnetic and Josephson energies of the vortex differ in magnitude, while the magnetic and Josephson components of the pinning energy are opposite in sign. Thus, this approach reflects both features mentioned above and is therefore a step forward to the exact solution.

However, the comparison between the energies obtained with this approach and those obtained by the exact numerical solution of finite-difference system (5) sheds light on a number of significant differences.

(i) The magnetic and Josephson energies obtained by these two approaches (Tables 1, 3; columns 4 and 5–7) differ: in the approximate model, the magnetic energy E_H is lower than the Josephson energy at any I; in the exact solution, the situation is reverse, since the

exact values of E_H and E_J are higher and lower than the approximate values, respectively.

(ii) The exact values of the Josephson and magnetic components of the pinning energy, ΔE_J and ΔE_H (Tables 2, 4; columns 6, 7), are many times higher than those based on the approximate approach (Tables 2, 4; columns 4, 5). In spite of this, since the Josephson and magnetic components of the pinning energy are opposite in sign, the exact values of the total pinning energy $E_p = \Delta E_J + \Delta E_H$ (Tables 2, 4; columns 8) do not greatly differ from the approximate results (Tables 2, 4; columns 9).

Thus, the exact value of the total pinning energy is well approximated by the continuous vortex model with the difference term replaced by the derivative. This fact is rather intriguing, because all other results based on this approach are far from being an exact solution even qualitatively. Specifically, the pinning energy components have equal signs, the magnetic and Josephson energies of the vortex are exactly equal to each other, etc.

The good agreement between the approximate and exact results for the total pinning energy and the sharp discrepancy between the magnetic and Josephson components individually are unlikely to be accidental. It seems that this effect has some intrinsic reason. The use of the approximate approach for finding the total pinning energy remains to be substantiated. This is a separate mathematical problem to which an approach has not yet been found.

CONCLUSION

An algorithm for the exact numerical solution of a set of difference equations describing the shape of a stable and an unstable vortex is suggested. It does not require application of the continuous vortex model. The calculation is based on two forms of the function $V(\varphi)$: $V = 1 - \cos\varphi$ and $V = 2/\pi^3\varphi^2(2\pi - \varphi)^2$. The vortex magnetic and Josephson energies diverge. The magnetic and Josephson components of the pinning energy are close in magnitude but differ in sign; as a result, the total pinning energy is smaller than its components by one order of magnitude. This result is confirmed theoretically.

An analytical computing method within the continuous vortex model is suggested. This method preserves the difference terms in the energy expression. The magnetic energy found by this method differs from the Josephson energy in magnitude, and the magnetic component of the pinning energy is opposite in sign to the Josephson component.

A comparative analysis of the approximate approaches to energy calculation within the continuous vortex model when the difference terms are retained and when they are replaced by the derivatives is performed.

A comparison of the results obtained in various approximations with each other and also with the exact solution to the problem stated shows that the continuous vortex model gives incorrect values of the Josephson and magnetic components of the pinning energy. The actual (exact) values of the Josephson and magnetic components of the pinning energy are several tens or several hundreds of times higher than those obtained within the continuous vortex model. However, since the Josephson and magnetic components of the pinning energy have different signs, the exact value of the total pinning energy and the approximate value obtained within the continuous vortex model differ insignificantly. It appears that this fact is not accidental and has some intrinsic reason. The use of the continuous vortex model for finding the total pinning energy remains to be substantiated. This is a separate mathematical problem to which an approach has not yet been found.

REFERENCES

1. M. A. Zelikman, Supercond. Sci. Technol. **12**, 1 (1999).
2. S. N. Dorogovtsev and A. N. Samukhin, Europhys. Lett. **25**, 693 (1994).
3. V. V. Bryksin, A. V. Gol'tsev, S. N. Dorogovtsev, *et al.*, Zh. Éksp. Teor. Fiz. **100**, 1281 (1991) [Sov. Phys. JETP **73**, 708 (1991)].
4. Ya. I. Frenkel' and T. M. Kontorova, Fiz. Zh. **1**, 137 (1939).
5. S. Aubry, *The Rieman Problem* (Springer-Verlag, Berlin, 1982), Lect. Notes Math., Vol. 925, p. 240.
6. M. A. Zelikman, Zh. Tekh. Fiz. **71** (7), 44 (2001) [Tech. Phys. **46**, 831 (2001)].
7. I. M. Ryzhik and I. S. Gradshteyn, *Table of Integrals, Series, and Products* (Nauka, Moscow, 1963; Academic, New York, 1980).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 7, 2002, pp. 829–833. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 36–40.

GASES AND LIQUIDS

A Numerical Simulation of Nucleation and Dynamics of Bubbles Formed under Drastic Depressurization of a Liquid

E. Yu. Kumzerova and A. A. Schmidt

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

e-mail: alexander.schmidt@pop.ioffe.rssi.ru

Received September 17, 2001.

Abstract—The formation and development of bubbles in a liquid under a drastic fall in pressure is numerically studied. The mathematical model is based on the Lagrangian–Eulerian approach to the description of a two-phase mixture and includes the conservation equations for the liquid, an equation describing the nucleation process, and equations describing the development of a probe bubble. For numerical solutions of equations for the liquid, we apply a high-resolution Godunov-type scheme, and a rigid set of ordinary differential equations describing the bubble development is solved using Adam's implicit method. The numerical simulation allowed us to obtain fields of the gasdynamic functions for the liquid, the size and concentration of the bubbles, the vapor temperature and pressure, and the range of parameters corresponding to a metastable liquid state. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Investigation of the processes of bubble formation and development under a drastic change in the liquid pressure is of great interest both for a wide range of applications [1–3] and for developing the theory of nonequilibrium multiphase media [4].

A convenient approach to solving such problems is numerical simulation. The mathematical model should include conservation equations for the mass, momentum, and energy of the liquid and equations describing the formation of nuclei of the vapor phase and their development due to changes in the parameters of the ambient liquid and the interphase mass transfer. Substantially different spatial and temporal scales of the processes in the liquid and vapor phases and drastic gradients of the gasdynamic parameters in the processes under study necessitate the use of numerical schemes of high resolution for integration of the equations describing the liquid behavior, on the one hand, and solution of rigid sets of ordinary differential equations describing the bubble dynamics, on the other hand.

The objective of this paper is a numerical investigation of the ebullition of water at a high pressure and temperature in the wake of rarefaction wave propagation through it. To solve this problem, we have formulated a mathematical model of the process in the framework of the Lagrangian–Eulerian description of a multiphase medium that allows us to study the formation of nuclei, interphase mass transfer, and bubble dynamics. This model has been used for investigating a one-dimensional nonsteady flow of a liquid coming to a boil.

MATHEMATICAL MODEL

In the framework of the Eulerian approach to describing liquid behavior on the assumption of equal velocities of the liquid and vapor phases, the conservation equations can be written in the following forms:

the equation of conservation of mass,

$$\frac{\partial(1-\alpha)\rho_l}{\partial t} + \frac{\partial u(1-\alpha)\rho_l}{\partial x} = -\Gamma; \quad (1)$$

the equation of conservation of momentum,

$$\frac{\partial u(1-\alpha)\rho_l}{\partial t} + \frac{\partial u^2(1-\alpha)\rho_l}{\partial x} = -(1-\alpha)\frac{\partial p_l}{\partial x} - u\Gamma; \quad (2)$$

and the equation of conservation of energy,

$$\frac{\partial \rho_l(1-\alpha)E_l}{\partial t} + \frac{\partial \rho_l(1-\alpha)E_l u}{\partial x} = -\frac{\partial(1-\alpha)p_l u}{\partial x} - U_{lv} - HW_{cl}. \quad (3)$$

Here, u is the liquid–vapor mixture velocity; ρ_l and p_l are the liquid density and pressure, respectively; α is the volume fraction of the vapor; $E_l = c_l T_l + u^2/2$ is the specific total energy of the liquid; and Γ is the rate of the interphase mass transfer, which can be written in the following form [4]:

$$\Gamma = 4\pi R^2 N_b \frac{\eta_{ac}}{\sqrt{2\pi R_v}}\left(\frac{p^{sat}(T_l)}{\sqrt{T_l}} - \frac{p_v}{\sqrt{T_v}}\right), \quad (4)$$

where N_b is the concentration of bubbles in the liquid; T_l and T_v are the temperatures of the liquid and vapor

phases, respectively; c_l is the specific heat of the liquid; R_v is the gas constant for the vapor; p^{sat} is the saturated vapor pressure; p_v is the actual vapor pressure; R is the bubble radius; η_{ac} is the accommodation coefficient (in the case considered, $\eta_{ac} = 0.04$); H is the nucleation rate; W_{cl} is the work needed for the formation of a critical vapor nucleus;

$$U_{lv} = 4\pi R^2 N_b \frac{\eta_{ac}}{\gamma_v - 1}\sqrt{\frac{R_v}{2\pi}}(p^{sat}(T_l)\sqrt{T_l} - p_v\sqrt{T_v}) \quad (5)$$

is the interphase energy transfer resulting from the phase transition; and γ_v is the polytropic exponent.

The set of Eulerian equations also includes an equation describing the formation and convective transfer of bubbles,

$$\frac{\partial N_b}{\partial t} + \frac{\partial u N_b}{\partial x} = H, \quad (6)$$

and an equation of the water state in Tait's form,

$$p_l = p_a K\left[\left(\frac{\rho}{\rho_a}\right)^{\beta} - 1\right] + p_a, \quad (7)$$

where p_a and ρ_a are the water pressure and temperature under normal atmospheric conditions, respectively; $K = 3045$; and $\beta = 7.15$.

The equations of the Lagrangian approach expressing conservation of mass and energy and describing the interface motion for a size-averaged probe bubble can be transformed into the following forms:

the equation of conservation of the vapor mass inside the bubble,

$$\frac{dp_v}{dt} = p_v\left[\frac{1}{T_v}\frac{dT_v}{dt} - \frac{3}{R}\left\{\frac{dR}{dt} - \frac{\eta_{ac}T_v}{p_v}\sqrt{\frac{R_v}{2\pi}}\left(\frac{p^{sat}(T_l)}{\sqrt{T_l}} - \frac{p_v}{\sqrt{T_v}}\right)\right\}\right]; \quad (8)$$

the conservation equation for the bubble vapor energy,

$$\frac{dT_v}{dt} = -3\frac{T_v}{R_v p_v}\left\{(\gamma_v - 1)p_v\frac{dR}{dt} + \eta_{ac}p^{sat}(T_l)(T_v - T_l)\sqrt{\frac{R_v}{2\pi T_l}}\right\}; \quad (9)$$

and the Rayleigh–Lamb equation,

$$R\frac{d^2R}{dt^2} + \frac{3}{2}\left(\frac{dR}{dt}\right)^2 = \frac{1}{\rho_l}\left\{p_v - p_l - \frac{2\sigma}{R} - \frac{4\mu}{R}\frac{dR}{dt}\right\}, \quad (10)$$

where μ is the liquid viscosity.

A relation between the volume fraction α of the vapor phase, concentration N_b of the bubbles, and their mean radius R is given by the obvious relationship

$$\alpha = (4/3)\pi R^3 N_b. \quad (11)$$

One of the key points of the simulation is identification of the formation mechanism of critical nuclei of vapor bubbles. For extreme superheating of the liquid, the density of nucleation centers calculated by the known theory of homogeneous formation of nuclei [5] is in good agreement with the experiment [6]. In this case, the nucleation rate can be calculated as follows:

$$H = \rho_l\left(\frac{N_a}{m}\right)^{\frac{3}{2}}\sqrt{\frac{2\sigma}{\pi}}\exp\left\{-\frac{W_{cl}}{k_b T_l}\right\}, \quad (12)$$

where m is the mass of a water molecule; σ is the surface tension; N_a is Avogadro's number; k_b is the Boltzmann constant; and W_{cl} is the work performed in the formation of a bubble with a critical radius of $R_{cr} = 2\sigma/(p^{sat} - p_l)$, which is defined as

$$W_{cl} = \frac{16\pi\sigma^3}{3(p^{sat}(T_l) - p_l)^2}. \quad (13)$$

In practice, the formation and growth of nuclei start at a considerably lower liquid superheating than the theory [5] predicts because of the predominant role of heterogeneous nucleation. Theoretical calculations of the number of sites of heterogeneous nucleation fail because, generally, there is no a priori data on its nature. However, when considering small sites of heterogeneous nucleation, one can use a modified theory of homogeneous nucleation by introducing a heterogeneity factor $G(G \in (0, 1])$, which characterizes reduction of the work W_{cl} needed for nucleation on the existing sites. Thus, one can allow for the heterogeneous character of the nucleation by using expression (12) for the rate of homogeneous nucleation and substituting $W_{cl}G$ for W_{cl}.

In [3], on the basis of experimental data, the dependence of the heterogeneous factor on the initial liquid temperature and the rate of liquid depressurization are obtained in the following form:

$$G = (1 + 1.4\times 10^{-10}V_p^{0.8})\left(\frac{T_{l0}}{T_{cr}}\right)^{28.46}, \quad (14)$$

where T_{cr} is the critical liquid temperature and V_p is the depressurization rate in units of Pa/s.

INITIAL AND BOUNDARY CONDITIONS

We assume that at the initial time instant, a homogeneous quiescent liquid with no vapor bubbles exists at a given pressure and temperature of

$$t = 0: p_l = p_{l0}, \quad T_l = T_{l0}, \quad N_b = 0, \quad u = 0. \quad (15)$$

Next, in the cross section $x = 0$ of a tube containing the liquid, we impose a pressure drop rate of V_p and a "soft" boundary condition for the liquid velocity,

$$x = 0: p_l = p_{l0} - V_p t, \quad \frac{\partial u}{\partial x} = 0. \qquad (16)$$

After a bubble has arisen in the liquid, we postulate the following initial conditions for the ordinary differential equations describing its development:

$$R = R_{cr}, \quad p_v = p^{sat}, \quad T_v = T_l.$$

NUMERICAL APPROACH

The conservation equations for the liquid are solved simultaneously with the nucleation equations with the use of Godunov's method of second-order accuracy in time and space [7]. Fluxes of the mass and momentum through the faces of the calculation cell are found by solving Riemann's problem for water. The ordinary differential equations describing the development of bubbles are solved simultaneously with the equations for the liquid. Because of the appreciably different characteristic times of the processes in the liquid and the bubbles, we apply Adams' implicit method for solving rigid sets of ordinary differential equations [8].

For the calculations, we used a uniform grid with cell dimensions sufficient for appropriate description of the regions of large gradients of the parameters. It should be noted that limitations imposed on the time integration step by the depressurization rate are more severe than the Courant–Friedrichs–Levi conditions.

DISCUSSION OF THE RESULTS

In the following, some results of the numerical simulation of a rarefaction wave propagating through a tube filled with water at a high pressure and temperature are presented. A diagram of the flow of the boiling water from the tube is shown in Fig. 1.

Under the boundary and initial conditions given above, a rarefaction wave propagates inside the tube with an intensive bubble nucleation in the wake of its front followed by subsequent growth of the bubbles. We assume that the opposite end of the tube is too far away to influence the flow pattern. We analyzed the influence of the pressure fall rate V_p and the initial temperature T_{l0} and pressure p_{l0} on the formation and development of the vapor bubbles.

To test the algorithm developed and the assumptions made, we compared the simulation results with the known experimental data [3, 9] (see Figs. 2, 3).

Figure 2 shows a typical variation of the pressure with time in a tube cross section at a distance of $x = 0.43$ m from the open end of the tube at an initial liquid temperature of $T_{l0} = 495$ K and pressure falling at a rate of $V_p = 22$ GPa/s. It is seen that the pressure first drops to a magnitude that is considerably less than the saturation pressure and the liquid is found in a metastable state. Then, intensive nucleation and growth of critical nuclei begin. Ultimately, this process leads to an increase in the pressure and its stabilization near the saturation level. The difference between the saturation pressure and the minimum liquid pressure is conventionally called the pressure undershoot, the determination of which is one of the most important problems in practical applications. In Fig. 3, the calculated pressure undershoot as a function of the liquid temperature at various pressure drop rates (circles) and approxima-

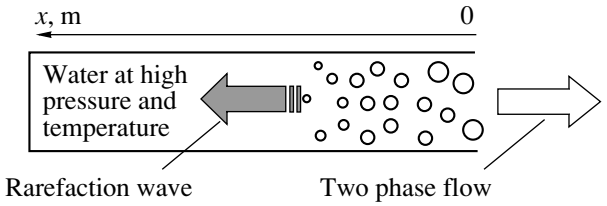


Fig. 1.

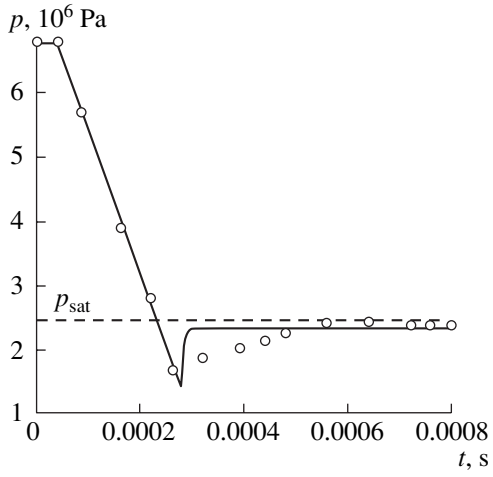


Fig. 2.

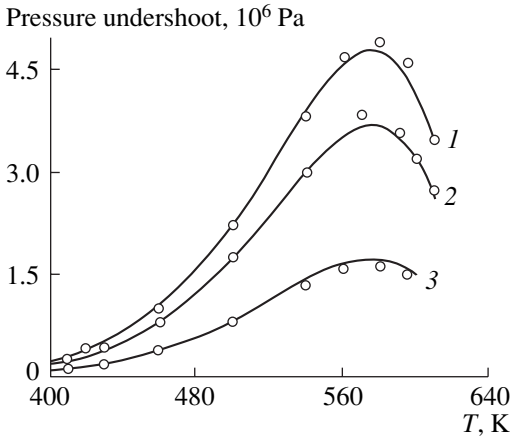


Fig. 3.

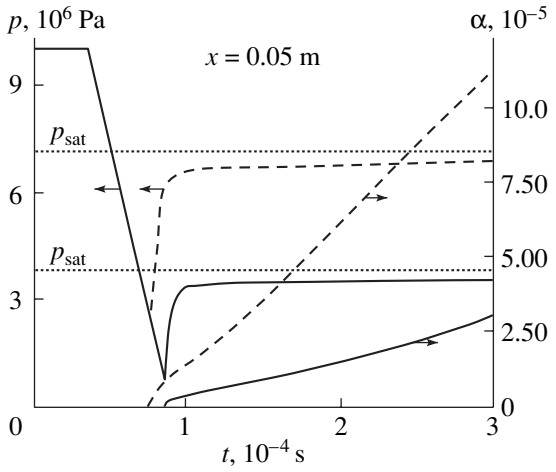


Fig. 4.

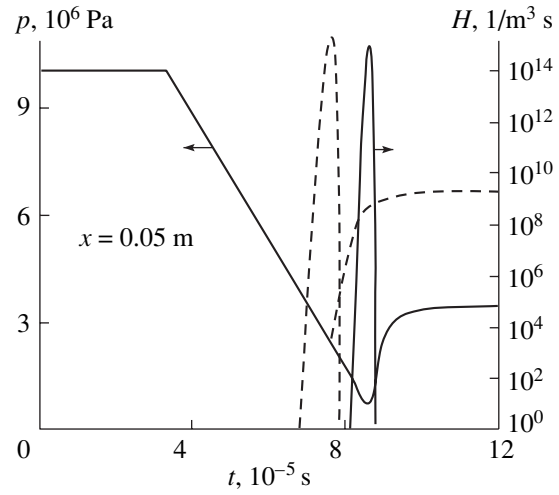


Fig. 5.

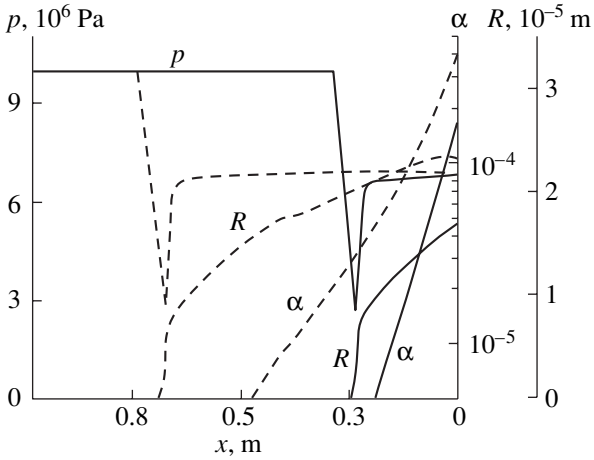


Fig. 6.

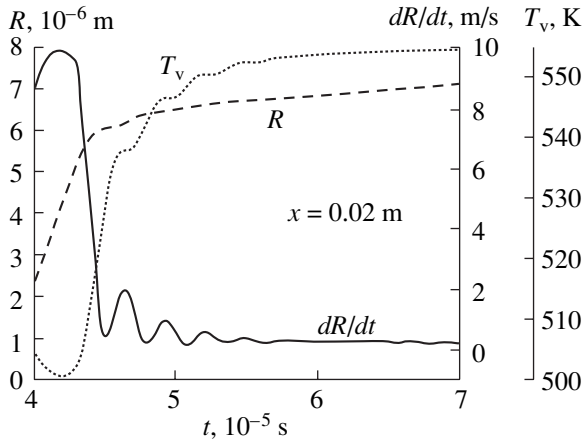


Fig. 7.

tions of experimental data [3] (solid lines) are shown. Here, V_p = (*1*) 200, (*2*) 100, and (*3*) 10 GPa/s.

It is seen that the adopted flow model of a boiling liquid with the use of the modified nucleation model has allowed us to obtain good agreement between the calculated and experimental data and to demonstrate the applicability of the approach considered to the problem arising when the pressure of a hot liquid falls rapidly.

Figures 4 and 5 illustrate the influence of the initial liquid temperature on variation with time of the liquid pressure p_l, the volume fraction of the vapor phase α, and the nucleation rate H. Solid lines in the plots correspond to a liquid temperature of T_{l0} = 520 K; dashed lines, to 560 K at a pressure drop rate of V_p = 180 GPa/s.

As seen in Fig. 4, the initial stage of decreasing the liquid pressure in the rarefaction wave does not depend on the liquid temperature, while the difference in the pressure undershoots at the initial temperatures considered is about 10^6 Pa. Thus, an increase in the initial temperature leads to a widening of the region of the metastable liquid state. A volume fraction of the vapor formed also rises appreciably with the initial liquid temperature.

On the basis of the plot in Fig. 5, one can draw a conclusion that the "wave of boiling" propagating through the liquid and corresponding to the maximum nucleation rate and minimum liquid pressure is appreciably shorter than the rarefaction wave.

Distributions of the water pressure, the vapor volume fraction, and the mean bubble radius along the tube for the two instants t = 0.0002 (solid lines) and 0.0005 s (dashed lines) at T_{l0} = 560 K and V_p = 180 GPa/s are presented in Fig. 6. The fast initial growth of the bubbles can be explained by dynamic processes initiated by the high initial pressure and temperature of the vapor inside the bubbles formed. The subsequent growth of the bubbles and the increase in the vapor phase fraction are governed by the predominant influence of the interphase mass transfer. The pressure undershoot and the

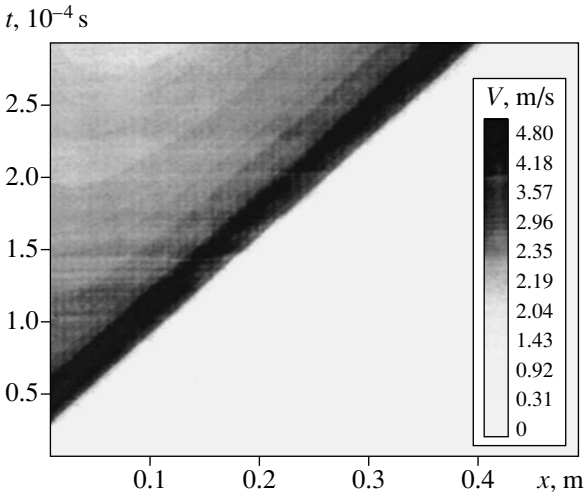


Fig. 8.

saturation pressure, as seen in the figure, are nearly constant along the tube. This indicates that there are relatively small changes in the liquid temperature.

Further development of the processes in the vapor phase of the two-phase flow considered is determined by the evolution of the bubbles formed. Figure 7 demonstrates variations with the time of the bubble radius, the interface velocity, and the vapor temperature at initial conditions of T_{l0} = 560 K and V_p = 180 GPa/s.

The nonmonotonic interface velocity at the initial process stage can be explained by irregularities in the interphase mass transfer between the bubbles being formed and the ambient liquid.

Figure 8 shows a mixture velocity distribution corresponding to initial conditions of T_{l0} = 560 K and V_p = 180 GPa/s. In particular, one can see that in a certain tube cross section, the mixture accelerates rapidly behind the rarefaction wave front up to the maximum velocity corresponding to the minimum liquid pressure (see Fig. 4) and then decelerates with time due to a growth in pressure caused by nucleation. The mixture deceleration caused by the nucleation weakens near the open end of the tube because of the assumed drop in pressure in the tube.

CONCLUSION

A mathematical model has been formulated that allows one to study the nucleation, interphase mass transfer, and development of bubbles as a result of variation in the liquid pressure. The numerical simulation of the distribution of a rarefaction wave through a tube filled with water allowed us to describe time-dependent variation along the tube length of the liquid phase parameters, volume fraction, temperature, and pressure of the vapor phase and the concentration and sizes of the vapor bubbles.

In the process under study, the following flow pattern takes place: the liquid pressure in the rarefaction wave propagating from the open end of the tube drops below the saturation pressure (the liquid is found in a metastable state); then, in the region of minimum pressure, an intensive nucleation starts and the bubbles grow quickly due to the interphase mass transfer, causing a growth in pressure and, ultimately, the liquid pressure to stabilize at a level close to the saturation pressure.

It has been demonstrated that in the case considered, the introduction of the heterogeneity factor in the expression for the formation energy of a critical nucleus allows one to obtain correct values of the pressure undershoot and to describe the behavior of a liquid in a metastable state.

REFERENCES

1. E. Hahne and G. Barthou, Int. J. Multiphase Flow **26**, 531 (2000).
2. J. Bartak, Int. J. Multiphase Flow **5**, 789 (1990).
3. Md. Alamgir and J. H. Lienhard, J. Heat Transfer **103**, 52 (1981).
4. R. I. Nigmatulin, *Dynamics of Multiphase Media* (Nauka, Moscow, 1987), Vol. 2.
5. Ya. I. Frenkel', *Kinetic Theory of Liquids* (Nauka, Leningrad, 1975).
6. V. P. Skripov *et al.*, *Thermophysical Properties of Liquids in Metastable State* (Atomizdat, Moscow, 1980).
7. A. V. Rodionov, Zh. Vychisl. Mat. Mat. Fiz. **27**, 1853 (1987).
8. E. S. Oran and J. P. Boris, *Numerical Simulation of Reactive Flows* (Elsevier, New York, 1987).
9. Md. Alamgir, C. Y. Kan, and J. H. Lienhard, J. Heat Transfer **102**, 433 (1980).

Translated by N. Mende

Technical Physics, Vol. 47, No. 7, 2002, pp. 834–840. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 41–47.

GASES AND LIQUIDS

Equilibrium Shape and Stability of a Charged Drop in an Air Flow under an Electric Field

A. I. Grigor'ev

Demidov State University, Sovetskaya ul. 14, Yaroslavl, 150007 Russia

Received November 12, 2001

Abstract—The pressure balance on the surface of a charged liquid drop moving along a uniform electrostatic field is analyzed. The liquid is assumed to be nonviscous and incompressible. In the approximation linear in deformation amplitude, the equilibrium shape of the drop as a function of the charge, field strength, and velocity of travel can be both a prolate and an oblate spheroid. Critical conditions for the surface instability of such a drop are obtained analytically in the form of a relationship between the charge, field strength, and velocity of travel. An instability criterion is found by extrapolating to large Reynolds numbers. This makes it possible to fit the earlier model of a corona-initiated lightning in the vicinity of large charged water drops or hailstones to the charges of the drops, field strengths, and velocities of travel (relative to the medium) typical of thunderclouds. © *2002 MAIK "Nauka/Interperiodica".*

(1) The instability of the surface of a drop against self-charge or induced charge is of interest for many applications in geophysics, technical physics, technology, and scientific instrument making (see, e.g., [1, 2] and Refs. therein). However, associated research has been concentrated mostly on elementary events in a thundercloud. According to modern concepts, streak lightning is initiated by a corona ignited in the vicinity of a large drop or a flooded hailstone (with the resulting instability of the charged surface of a water drop) [3, 4]. However, these ideas have not been corroborated by full-scale experiments with thunderclouds. The maximum values of the self-charge of the drops and of intracloud electric fields measured experimentally are much less [5] than those necessary for the surface of the drop to be unstable against self-charge or induced charge [6]. It is likely that the available models of lightning initiation overlook some important factors, such as the aerodynamic pressure near a falling drop, which loosens critical conditions for the instability of the free surface of the drop [7, 8]. Of significance also is the equilibrium shape of a charged drop moving in an external electric field, since analysis of the surface stability of the drop is impossible without knowing its equilibrium shape.

(2) Let a laminar gas stream of density ρ and velocity $\mathbf{U}$ aligned with an external electrostatic field $\mathbf{E}_0$ ($\mathbf{U} \parallel \mathbf{E}_0$) flow around a drop of a nonviscous incompressible liquid of charge Q. Our goal is to find the equilibrium shape of the drop. We assume that the flow velocity U is much less than the velocity of sound in the gas and approximate the gas by a nonviscous incompressible fluid.

At $Q = 0$, $E_0 = 0$, and $U = 0$, the spherical shape of the drop and its radius R are easily found from the equation of pressure balance on its surface,

$$\frac{2\sigma}{R} = \Delta\rho,$$

where σ is the surface tension coefficient and $\Delta\rho$ is the pressure drop between the drop and the medium.

Now, let $Q \neq 0$, $E_0 \neq 0$, and $U \neq 0$. In this case, the shape of the drop is other than spherical. Its new equilibrium shape can be represented by

$$r(\Theta) = R + h(\Theta) \equiv R + \sum_{n=0}^{\infty} a_n P_n(\mu). \qquad (1)$$

Here, $P_n(\mu)$ are the nth-order Legendre polynomials, a_n are the amplitudes of disturbance individual modes, $\mu \equiv \cos(\Theta)$, and $h(\Theta)$ is the virtual disturbance of the spherical surface of the drop. We will seek for the disturbance $h(\Theta)$ (or for the amplitudes a_n of the disturbance modes) again using the condition for the pressure balance on the surface,

$$p_\sigma = \Delta p + p_E + p_U. \qquad (2)$$

The terms on the right of (2) will be found on the initial spherical surface; the bubble pressure p_σ on the left of (2), on the virtually disturbed spherical surface. Accordingly, p_E is the electrostatic pressure of the field of the self-charge and polarization charge on the spherical surface and p_σ is the aerodynamic pressure on the spherical surface from the laminar gas flow.

Our aim is to find the amplitudes a_n of modes that are excited when the virtual disturbance $h(\Theta)$ interacts with the electric and aerodynamic fields in the vicinity

of a conducting spherical drop. Note first of all that [9, 10]

$$p_\sigma = \frac{2\sigma}{R} - \frac{\sigma}{R^2}\sum_{n=0}^{\infty}[2-n(n+1)]a_nP_n(\mu), \quad (3)$$

$$p_E = \frac{1}{8\pi}\left\{9E_0^2\left[\frac{1}{3}P_0(\mu)+\frac{2}{3}P_2(\mu)\right] + 6E_0\frac{Q}{R^2}P_1(\mu)+\frac{Q^2}{R^4}P_0(\mu)\right\}, \quad (4)$$

$$p_U = \frac{9}{8}\rho U^2\left[\frac{2}{3}P_0(\mu)-\frac{2}{3}P_2(\mu)\right]. \quad (5)$$

Substituting (3)–(5) into (2) and equating the coefficients of the Legendre polynomials of different order, we determine the amplitudes of the modes excited. From (4) and (5), it is easy to see that only the first three modes can be excited: $\sim P_0(\mu)$, $\sim P_1(\mu)$, and $\sim P_2(\mu)$. The mode $\sim P_1(\mu)$ corresponds to the translational motion of the drop and does not affect its shape. The modes $\sim P_0(\mu)$ and $\sim P_2(\mu)$ are easy to calculate, but the former contains the pressure drop Δp, which is unknown. Therefore, a_0 is more conveniently calculated through a_2 by using the fact that the volume of an incompressible liquid drop is constant. For a_2, we then have

$$\frac{a_2}{R} = \frac{1}{3}\left(\frac{9}{16\pi}w-\frac{9}{16}We\right), \quad w \equiv E_0^2R\sigma^{-1}, \quad We \equiv \rho U^2R\sigma^{-1}. \quad (6)$$

The amplitude a_0 calculated from the constant volume condition depends quadratically on the dimensionless amplitude of the fundamental mode $n = 2$, $(a_0/R) \sim (a_2/R)^2$, and the associated expression is ignored in the linear approximation.

Thus, only the fundamental mode grows in the linear approximation. For $(w/\pi) > We$, the drop becomes a prolate spheroid [6, 9]; for $(w/\pi) < We$, an oblate spheroid. Since we are interested in fine charged drops (with small We) in strong (large w) electric fields of thunderclouds, further considerations will be restricted to the case $(w/\pi) > We$. Therefore, we will subsequently require that pressure balance condition (2) on the surface of a prolate spheroid be fulfilled.

(3) We assume that $U = 0$ and Q and E_0 are nonzero at the initial time instant. According to [6], in the linear approximation of the surface distortion amplitude, the surface of the drop is a prolate spheroid with an eccentricity e given by

$$e^2 = \frac{9}{16\pi}\frac{w}{1-W}, \quad W \equiv \frac{Q^2}{16\pi\sigma R^3}. \quad (7)$$

The parameter W is a measure of the drop stability against its self-charge. The critical value of W at which the drop loses stability equals unity [11]. We will increase the velocity U of the flow along the axis of symmetry of the drop to see how the aerodynamic laminar flow around the drop affects its initially spherical shape. We again require that pressure balance condition (2) on the surface of the prolate spheroid be fulfilled. The surface of the prolate spheroid is given by the equation

$$\eta(\Theta) = \frac{r(\Theta)}{a} = (1-e^2)^{1/2}(1-e^2\cos^2\Theta)^{-1/2}, \quad e^2 = 1-b^2/a^2, \quad r(\Theta) = R\left[P_0(\mu)+\frac{1}{3}e^2P_2(\mu)+O(e^4)\right], \quad e^2 \ll 1, \quad (8)$$

where $b = R(1-e^2)^{1/6}$ and $a = R(1-e^2)^{-1/3}$ are the minor and major semiaxes of the spheroid and the angle Θ is measured from the direction of $\mathbf{E}_0$ (or $\mathbf{U}$, which is the same).

Instead of (1), the equation for the desired equilibrium surface will now be sought in the form

$$r(\Theta) = r(\Theta) + h(\Theta) \equiv r(\Theta) + \sum_{n=0}^{\infty}a_nP_n(\mu). \quad (9)$$

The pressure balance condition will be recast as

$$p_\sigma^s + \delta p_\sigma^s = \Delta p^s + p_E^s + p_U^s. \quad (10)$$

Here, p_σ^s and p_E^s are the bubble and electrostatic pressures under the spheroidal surface. In the approximation linear in e^2, these pressures take the form [6]

$$p_\sigma^s = \frac{2\sigma}{R}\left[1+e^2\frac{2}{3}P_2(\mu)+O(e^4)\right],$$

$$p_E^s = \frac{1}{\varepsilon\pi}\left\{9E_0^2\left[\frac{1}{3}P_0(\mu)+\left(\frac{2}{3}+\frac{12}{35}e^2\right)P_2(\mu) + \frac{16}{35}e^2P_4(\mu)\right] + 6E_0\frac{Q}{R^2}\left[\left(1+\frac{2}{15}e^2\right)P_1(\mu) + \frac{3}{5}e^2P_3(\mu)\right] + \frac{Q^2}{R^4}\left[P_0(\mu)+\frac{2}{3}e^2P_2(\mu)\right]+O(e^4)\right\}.$$

The addition δp_σ^s to the bubble pressure is associated with the virtual deformation $h(\Theta)$ of the spheroidal drop surface. In the approximation linear in e^2, it is given by [6]

$$\delta p_\sigma^s = -\frac{\sigma}{R^2}\sum_{n=0}^{\infty}\left\{[4+n(n+1)]\frac{(n+1)(n+2)}{(2n+1)(2n+3)}e^2P_{n+2}(\mu)\right.$$

$$+\left[2\left(1-\frac{2}{3}e^2\right)-\left(1+\frac{e^2}{3}\right)n(n+1)\right]P_n(\mu)$$

$$+e^2[4+n(n+1)]\frac{n(n-1)}{(2n-1)(2n+1)}P_{n-2}(\mu)$$

$$+e^2[4+n(n+1)]\frac{[2n(n+1)-1]}{(2n-1)(2n+3)}P_n(\mu)+O(e^4)\Big\}a_n.$$

In the same linear approximation, the aerodynamic pressure on the surface of a spheroid prolate in the direction of its motion on the side of the laminar flow of a nonviscous fluid is easily found as (see Appendix)

$$p_U^s = \frac{9}{8}\rho U^2\left\{\frac{2}{3}P_0(\mu)-\left(\frac{2}{3}-\frac{16}{35}e^2\right)P_2(\mu)\right.$$

$$\left.-\frac{16}{35}e^2P_4(\mu)+O(e^4)\right\}.$$

In (10), the sum of the terms p_σ^s, Δp^s, and p_E^s vanishes on the initial spheroidal surface with the eccentricity e^2. Eventually, we have

$$\sigma p_\sigma^s = p_U^s.$$

We substitute the above expressions for δp_σ^s and p_U^s into this equality to find the additions to the amplitudes of the modes that set in as a result of the interaction between the aerodynamic pressure and the virtual addition to the bubble pressure. In the approximation linear in e^2, we find that only the addition to the fundamental mode is nonzero:

$$\frac{a_2}{R} = -\frac{3}{16}We.$$

This corresponds to a prolate spheroid for which the eccentricity squared is decreased by $\delta e^2 = 9We/16$ [6]. Thus, in the linear (in e^2) approximation, the aerodynamic laminar flow around the prolate spheroidal drop along its axis of symmetry merely decreases the square of eccentricity:

$$e_*^2 = e^2-\frac{9}{16}We \equiv \frac{9}{16\pi}\frac{w}{1-W}-\frac{9}{16}We. \qquad (11)$$

It follows from (11) that in the space of the parameters *W*, *w*, and *We*, there exists a locus of functionally related points forming a surface where the drop has a spherical shape at $W = 0$, $w = 0$, and $We = 0$.

(4) Let us examine the effect of a laminar flow around a charged drop placed in a uniform electrostatic field on the stability of the drop. According to [12, 13], the drop loses stability against its self-charge or the polarization charge starting from the fundamental mode ($n = 2$); that is, the fundamental mode causes the instability. The instability of higher ($n > 2$) modes appears sequentially in order of increasing mode number. As they become unstable, the amplitude of the fundamental mode grows because of an increase in the surface charge density at the vertices of the extending drop. Therefore, we will study the stability of a spheroidal drop with the eccentricity defined by (11) against the virtual increment $\kappa_0 P_2(\mu)$ of the fundamental mode amplitude (here, $P_2(\mu)$ is the Legendre polynomial, $\mu \equiv \cos(\Theta)$). In the linear approximation, such a disturbance of the fundamental mode amplitude entails growth of the square of the eccentricity by δe^2:

$$\delta e^2 = \frac{3\kappa_0}{R}\left(1-\frac{7}{6}e^2\right).$$

As a result, the pressures p_E^s and p_U^s grow by δp_E^s and δp_U^s, respectively [6, 10]:

$$\delta p_E^s = \frac{9}{4\pi}\frac{\kappa_0}{R}\left\{E_0^2\left[\frac{237}{175}e^2P_0(\mu)+\frac{18}{35}\left(1+\frac{553}{90}e^2\right)P_2(\mu)\right.\right.$$

$$\left.+\frac{24}{35}\left(1+\frac{1663}{880}e^2\right)P_4(\mu)+\frac{66}{77}e^2P_6(\mu)\right]$$

$$+\frac{Q^2}{R^4}\left[\frac{8}{135}e^2P_0(\mu)+\frac{1}{9}\left(1+\frac{67}{42}e^2\right)P_2(\mu)\right.$$

$$\left.+\frac{16}{105}e^2P_4(\mu)\right]+\frac{QE_0}{R^2}\left[\frac{2}{15}\left(1-\frac{193}{210}e^2\right)P_1(\mu)\right.$$

$$\left.\left.+\frac{3}{5}\left(1+\frac{421}{270}e^2\right)P_3(\mu)+\frac{43}{63}e^2P_6(\mu)\right]+O(e^4)\right\},$$

$$\delta p_U^s = \frac{27}{4}\rho U^2\frac{\kappa_0}{R}\left[-\frac{8}{175}e^2P_0(\mu)\right.$$

$$+\frac{8}{35}\left(1-\frac{1}{10}e^2\right)P_2(\mu)-\frac{8}{35}\left(1-\frac{315}{154}e^2\right)P_4(\mu)$$

$$\left.-\frac{64}{231}e^2P_6(\mu)+O(e^4)\right].$$

These expressions correspond to the terms $\sim\delta e^2$ in the Taylor series for p_E^s and p_0^s.

As a result of the interaction between the virtual disturbance $\kappa_0 P_2(\mu)$ of the surface shape and the electrical, aerodynamic, and Laplace (bubble) forces, the shape of the drop may change if its initial spheroidal shape is ultimate in terms of stability against deformation. A new shape of the drop is found from the pressure bal-

ance condition [6]

$$p_\sigma^s + \delta p_\sigma^s = \Delta p^s + p_E^s + \delta p_E^s + p_U^s + \delta p_U^s. \quad (12)$$

Taking into account that the initial spheroidal shape of the drop satisfies Eq. (10) and its eccentricity is defined by (11), we find from (12) the relationship between the additions to the pressures:

$$\delta p_\sigma^s = \delta p_E^s + \delta p_U^s.$$

Substituting the expressions for δp_σ^s, δp_E^s, and δp_U^s into this relationship and equating the coefficients before the equal Legendre polynomials, we arrive at the expression for the addition to the fundamental mode amplitude:

$$\frac{a_2}{R} = \frac{\kappa_0}{R}\left\{\frac{27}{70}We\left(1 + \frac{79}{210}e^2\right) + W\left(1 + \frac{87}{42}e^2\right) + w\frac{81}{280\pi}\left(1 + \frac{4171}{630}e^2\right)\right\}. \quad (13)$$

It is easy to see that if the expression in the braces in (13) is less than unity, the virtual disturbance diminishes with time. This is evident from the following considerations. Let us assume that the expression for the addition to the fundamental mode amplitude is the initial virtual disturbance κ_1 and repeat our calculation from the beginning. According to (13), a new addition to the fundamental mode amplitude will be smaller than κ_1; that is, the disturbance will decrease. If the expression in the braces in (13) exceeds unity, the addition to the fundamental mode amplitude will grow with time; that is, the fundamental mode, along with all higher modes, will become unstable and the drop as a whole will suffer from instability following the scheme described in [14, 15].

Eventually, when a laminar gas flow passes around a charged drop subjected to a uniform electrostatic field, the critical condition for the instability of the drop will have the form

$$\frac{27}{70}We\left(1 + \frac{79}{210}e^2\right) + W\left(1 + \frac{87}{42}e^2\right) + w\frac{81}{280\pi}\left(1 + \frac{4171}{630}e^2\right) > 1. \quad (14)$$

For $U = 0$, this conditions is reduced to the instability condition for a charged drop in a uniform electrostatic field [6].

The surface defined by (11) and (14) in the space of the parameters *We*, *W*, and *w* is depicted in the figure; it separates the stable (below the surface) and unstable (above the surface) states of charged drops in the laminar flow of a nonviscous fluid parallel to the electrostatic field.

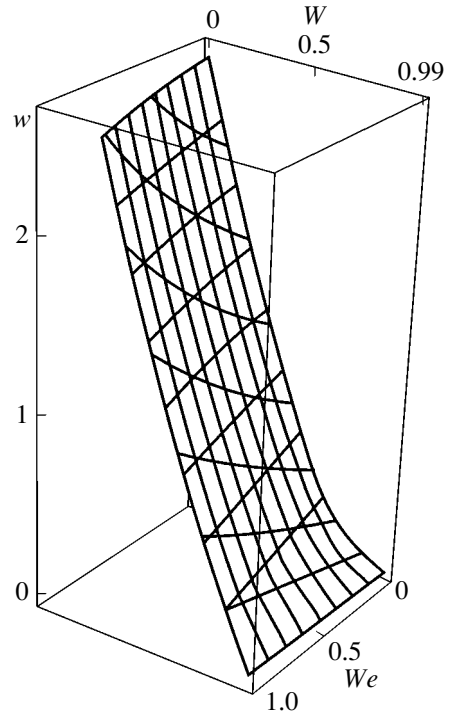


Figure.

(5) Let us verify the validity of condition (14) by using maximum values of the charges of the drops and electric field strengths observed in thunderclouds.

Assume that a drop has the radius $R = 100$ μm; then, it falls in a cloud at a velocity $U = 72$ cm/s and the Reynolds number is Re = 9.61 [16]. The flow around the drop can be considered laminar; thereby, all the above results totally apply to it (and to finer drops). The self-charge of the drop is small, $Q \le 3 \times 10^{-12}$ C [5]; hence, the parameter *W* is low. The parameters *We* and *w* will also be small (the former because of the small flow velocity and the latter because of the small radius of the drop). Eventually, we will find that condition (14) does not hold; hence, this drop (and finer drops) is stable against the electrostatic and aerodynamic forces.

Let a drop in a thundercloud have a radius $R = 1$ mm. According to observations [5], its charge may reach $Q = 2 \times 10^{-10}$ C. The maximum electrostatic field strength in thunderclouds was measured to be $E = 2000$ kV/m [5]. We will, however, take a much lower value, $E = 1200$ kV/m. This drop will fall in the gravitational field at a constant velocity $U = 6.49$ m/s [16]. The flow around this drop is characterized by the Reynolds number Re = 866 [13]. With so high Re, the flow is turbulent; hence, criterion (14), which was derived for laminar conditions, fails. Note however that the fulfillment or nonfulfillment of condition (14) indicates whether our model is basically adequate or in-adequate for analysis of the lightning initiation mechanism. With the above parameter values substituted into (11) and

(14), the criterion of drop instability is met. Thus, the stability analysis of the charged drop surface must include aerodynamic flow conditions.

Consider a feature of the behavior of the drop at large flow velocities, where the flow is no longer laminar. According to the analysis performed in [8] and experimental observations of the shape of moving drops [17], as the flow velocity rises (the Reynolds number increases), lower modes of capillary oscillations become unstable (starting from the fundamental one) owing to the unstable tangential discontinuity of the velocity field at the boundary of the drop. The fundamental mode becomes unstable first, and the eccentricity of the drop grows by a periodic law. This, according to [18], may cause the instability of the charged surface even if condition (14) fails.

(6) Relationships (11) and (14) can be used to advantage for experimental verification of the condition for stability against self-charge. In the experiments [19], heavily charged drops were suspended in an electrostatic field with an aerodynamic flow parallel to the field. The radii of the drops were small, $R \le 100$ μm, and the Reynolds number was $\mathrm{Re} \le 6$. Under these conditions, the charge of the drops was close to the limiting value in terms of stability against self-charge ($W \approx 1$). The spread in experimental results was significant: the deviation from the Rayleigh criterion for the stability of a conducting nonviscous incompressible liquid spherical drop against its self-charge, $W = 1$ [11], was as high as 17%. Such a large discrepancy occurs because the shape of the drop was taken to be spherical, rather than spheroidal, as (11), and the effect of the electrostatic field and aerodynamic forces on stability condition (14) was ignored in the experimental data processing.

(7) Of interest is the situation where the external electrostatic field $\mathbf{E}_0$ and the velocity $\mathbf{U}$ are normal to each other. This situation is typical of thunderclouds and is common in experiments [20, 21] where a charged drop falls in the gravitational field normal to the electric field of a capacitor.

Assuming that the effect of the electric field on the surface shape is stronger than that of the aerodynamic forces and that the drop is almost spheroidal (see [20, 21]), we come to the problem of flow around a prolate spheroid with the axis of symmetry aligned with the *OY* axis (with the external electric field), which runs normally to the flow (the flow is directed along the *OX* axis in opposition to the free-fall acceleration). Using the well-known solution [10, 22] to the problem of potential flow around an ellipsoid oriented arbitrarily with respect to the flow, one can find the velocity field potential $\varphi(\mathbf{r})$ in the vicinity of the spheroid. Unlike the case of axisymmetric flow around a spheroid considered above, this case is nonaxisymmetric and an azimuth dependence appears, which greatly complicates the problem. Now, the aerodynamic (and also the bubble) pressure must be expanded in spherical functions rather than in Legendre polynomials, with regard for the azimuth dependence. Moreover, it is necessary to make an asymptotic expansion in two independent parameters characterizing the deformation of the drop in two mutually orthogonal directions. From general physical considerations, it is clear that the equilibrium shape of the drop will be ellipsoidal and have three different axes under these conditions and that the effect of the laminar transverse flow will show up largely in extending the drop along the *OZ* axis transversely to the flow and to the external electric field. In this situation, the effect of the gas flow on the stability against self-charge and field-induced charge is not obvious. We can only argue that both the very fact of the flow (with the tangential discontinuity of the velocity field on the surface of the drop) and the transformation of the drop into a triaxial spheroid reduce its stability [7, 8, 23].

(8) To conclude, we note that the aerodynamic pressure on the surface of an extended spheroidal nonviscous incompressible charged liquid drop moving in a laminar nonviscous incompressible fluid and subjected to an external electrostatic field parallel to the velocity vector of the drop decreases its eccentricity and reduces its stability because of the tangential discontinuity of the velocity field on the surface. Thus, we have every reason to resume the construction of physical models of lightning initiation.

The above considerations are valid for laminar flows, that is, for thunderclouds with drops of radius $R \le 100$ μm. The majority of drops in the cloud meet this condition, since, as is known [13], cloud water concentrates mostly in drops with a radius from 3–4 to 20–30 μm.

APPENDIX

CALCULATION OF AERODYNAMIC PRESSURE ON THE SURFACE OF A SPHEROID IN A LAMINAR GAS FLOW

Let a spheroid of volume equal to that of a sphere of radius R have its axis of symmetry along the *OX* axis. The *OX* axis is directed along the laminar flow of a nonviscous incompressible fluid of density ρ around the spheroid. The flow velocity is U. The solution to the problem of the laminar potential flow of a nonviscous incompressible fluid around a rigid ellipsoid can be found elsewhere [10, 22]. Using the results of [10, 22] for the ellipsoid, we write the expression for the velocity potential $\varphi(\mathbf{r})$ in the vicinity of the ellipsoid:

$$\varphi(\mathbf{r}) = -Ux\left(1 + \frac{C(\xi)}{2 - C_0}\right),$$

$$C(\mathbf{r}) = \frac{2b^2 a}{(a^2 - b^2)^{3/2}}\int_{\nu}^{\infty}\frac{d\nu}{(\nu^2 - 1)\nu^2} \equiv \frac{2(1 - e^2)}{e^3}\left(\operatorname{arctanh}\frac{1}{\nu} - \frac{1}{\nu}\right), \tag{15}$$

$$\nu \equiv \left(\frac{a^2+\xi}{a^2-b^2}\right), \quad C_0 \equiv \frac{2(1-e^2)}{e^3}(\operatorname{arctanh} e - e),$$

where ξ is a spheroidal coordinate that is a positive root $\lambda \equiv \xi$ of the equation

$$\frac{x^2}{a^2+\lambda} + \frac{\rho^2}{b^2+\lambda} = 1, \quad \rho^2 = y^2 + z^2. \tag{16}$$

This spheroidal coordinate specifies a family of confocal spheroids. The surface of the flowed-around spheroid is defined by the relationship $\xi = 0$. Equation (16) is quadratic in λ, and its second root $\lambda \equiv \zeta$ ($-b^2 \geq \zeta \geq -a^2$) defines the second spheroidal coordinate, which specifies a family of two-sheet hyperboloids orthogonal to a family of spheroids specified by the coordinate ζ. The unit vector $\mathbf{n}_\xi$ is normal to the surface of the spheroid, and the unit vector $\mathbf{n}_\zeta$ is tangential to it.

The calculation of the flow velocity component V_ξ tangential to the spheroid will suffice in finding the pressure the nonviscous fluid flow exerts on the surface of the spheroid, since the normal component V_ζ vanishes on its surface.

Let us find the tangential component V_ξ by differentiating potential (15) along the vector $\mathbf{n}_\xi$ and taking into account the independence of the spheroidal coordinates ζ and ξ:

$$V_\zeta = -U\left(1 + \frac{C(\xi)}{2-C_0}\right)\frac{1}{h_\zeta}\frac{\partial x(\xi,\zeta)}{\partial z}\bigg|_{\xi=0}$$

$$\equiv -\frac{U}{e}\left(1 + \frac{C_0}{2-C_0}\right)\left(\frac{\zeta+a^2}{\zeta}\right)^{1/2} \tag{17}$$

$$\equiv -U\left(1 + \frac{C_0}{2-C_0}\right)\frac{\sin\Theta}{\sqrt{1-(2-e^2)e^2\cos^2\Theta}},$$

where e is the eccentricity and h_ζ is the Lamé coefficient of the coordinate ζ [24]:

$$h_\zeta \equiv \frac{1}{2}\left(\frac{(\zeta+b^2)(\zeta-\xi)}{(\zeta+a^2)}\right)^{1/2}.$$

When deriving relationship (17), we took into consideration that [24]

$$x = \mp\left(\frac{(\zeta+a^2)(\xi+a^2)}{(a^2-b^2)}\right)^{1/2},$$

$$\rho = \left(\frac{(\zeta+b^2)(\xi+b^2)}{(b^2-a^2)}\right)^{1/2},$$

and also that the condition

$$\xi = r(\Theta) - a^2(2-e^2)$$

is fulfilled on the spheroidal surface at $\xi = 0$ [$r(\Theta)$ is given by (8)].

The aerodynamic pressure on the surface of the spheroid is given by (up to terms $\sim e^4$)

$$p_U^s = \frac{1}{2}\rho V_\zeta^2 = \frac{9}{8}\rho U^2\left\{\frac{2}{3}(P_0(\mu) - P_2(\mu))\right.$$

$$+ e^2\frac{16}{35}(P_2(\mu) - P_4(\mu)) + e^4\left(-\frac{8}{175}P_0(\mu) + \frac{128}{525}P_2(\mu)\right. \tag{18}$$

$$\left.\left. + \frac{212}{1155}P_4(\mu) - \frac{64}{231}P_2(\mu) + O(e^6)\right)\right\}.$$

Expression (18) is written accurate to terms $\sim e^4$, because this form is convenient for calculating the addition δp_U^s by the relationship $\delta p_U^s \approx (\partial p_U^s/\partial e^2)\delta e^2$ when the stability of the drop is studied (see Section 4).

REFERENCES

1. A. I. Grigor'ev and S. O. Shiryaeva, Izv. Akad. Nauk, Mekh. Zhidk. Gaza, No. 3, 3 (1994).
2. D. F. Belonozhko and A. I. Grigor'ev, Elektrokhim. Obrab. Met., No. 4, 17 (2000).
3. V. A. Dyachuk and V. M. Muchnik, Dokl. Akad. Nauk SSSR **248**, 60 (1979).
4. A. I. Grigor'ev and S. O. Shiryaeva, Phys. Scr. **54**, 660 (1996).
5. I. P. Mazin, A. Kh. Khrgian, and I. M. Imyanitov, *Clouds and Cloud Atmosphere: Handbook* (Gidrometeoizdat, Leningrad, 1989).
6. A. I. Grigor'ev, S. O. Shiryaeva, and E. I. Belavina, Zh. Tekh. Fiz. **59** (6), 27 (1989) [Sov. Phys. Tech. Phys. **34**, 602 (1989)].
7. A. I. Grigor'ev, V. A. Koromyslov, and S. O. Shiryaeva, Zh. Tekh. Fiz. **69** (5), 7 (1999) [Tech. Phys. **44**, 486 (1999)].
8. A. I. Grigor'ev, V. A. Koromyslov, and S. O. Shiryaeva, Zh. Tekh. Fiz. **70** (7), 26 (2000) [Tech. Phys. **45**, 840 (2000)].
9. A. I. Grigor'ev and S. O. Shiryaeva, Zh. Tekh. Fiz. **57**, 1863 (1987) [Sov. Phys. Tech. Phys. **32**, 1119 (1987)].
10. N. E. Kochin, I. A. Kibel', and N. V. Roze, *Theoretical Hydrodynamics* (Fizmatgiz, Moscow, 1963), Part 1.
11. Rayleigh Lord (J. V. Strett), Philos. Mag. **14**, 184 (1882).
12. A. I. Grigor'ev, Zh. Tekh. Fiz. **55**, 1272 (1985) [Sov. Phys. Tech. Phys. **30**, 736 (1985)].

13. A. I. Grigor'ev and O. A. Sinkevich, Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza, No. 6, 10 (1985).
14. A. I. Grigor'ev and S. O. Shiryaeva, Zh. Tekh. Fiz. **61** (3), 19 (1991) [Sov. Phys. Tech. Phys. **36**, 258 (1991)].
15. A. N. Zharov, S. O. Shiryaeva, and A. I. Grigor'ev, Zh. Tekh. Fiz. **69** (12), 26 (1999) [Tech. Phys. **44**, 1420 (1999)].
16. I. P. Mazin and S. M. Shmeter, *Clouds: Structure and Physics of Formation* (Gidrometeoizdat, Leningrad, 1983).
17. A. L. Gonor and V. Ya. Rivkind, Itogi Nauki Tekh., Ser. Mekh. Zhidk. Gaza **17**, 86 (1982).
18. A. I. Grigor'ev, Zh. Tekh. Fiz. **70** (5), 22 (2000) [Tech. Phys. **45**, 543 (2000)].
19. M. Roulleau and M. Desbois, J. Atmos. Sci. **29**, 565 (1972).
20. T. B. Matthews, J. Geophys. Res. **72**, 3007 (1967).
21. E. L. Ausman and M. Brook, J. Geophys. Res. **72**, 6131 (1967).
22. H. Lamb, *Hydrodynamics* (Cambridge Univ. Press, Cambridge, 1932; Gostekhizdat, Moscow, 1947).
23. A. I. Grigor'ev and S. I. Shchukin, Zh. Tekh. Fiz. **68** (11), 48 (1998) [Tech. Phys. **43**, 1314 (1998)].
24. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1984).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 7, 2002, pp. 841–844. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 48–51.

GAS DISCHARGES, PLASMA

Surface Erosion of Tungsten and the Morphology of Erosion Products in Experiments Simulating Plasma Disruption

M. I. Guseva, V. M. Gureev, A. G. Domantovskiĭ, Yu. V. Martynenko, P. G. Moskovkin, V. G. Stolyarova, V. M. Strunnikov, L. N. Plyashkevich, and V. I. Vasil'ev

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

e-mail: martyn@nfi.kiae.ru

Received February 23, 2001; in final form, February 12, 2002

Abstract—The surface erosion of different sorts of tungsten subjected to high-power pulsed plasma streams simulating plasma disruption is studied. With W-13I polycrystalline and (111) single-crystal tungsten samples used as examples, the size distributions for the erosion products collected at different angles to the target are compared. The typical drop erosion of the surface is observed. Fine drops either return to the surface or fly away in a direction parallel to the surface. Coarse drops leave the surface nearly at right angles to the surface. The single-crystal surface displays the absence of fine (<0.125 μm) drops typical of a polycrystalline tungsten surface. The erosion of the single-crystal samples is least among the tungsten sorts considered. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

In the design of the International Thermonuclear Experimental Reactor (ITER), carbon–fiber composites are intended for in the most stressed parts of the divertor; tungsten, in divertor elements operating under less severe conditions. In view of the high threshold energies of tungsten sputtering by light ions, the disruption of plasma streams, which results in surface melting, can be considered the most dangerous source of tungsten erosion. In [1], we studied the surface erosion and the morphology of erosion products for domestic tungsten irradiated by a high-power pulsed deuterium plasma simulating plasma disruption.

In this study, the surface erosion for a number of tungsten materials exposed to high-power pulsed plasma streams is studied. With W-13I polycrystalline and (111) single-crystal tungsten samples used as examples, we compare the size distributions of the erosion products deposited on collectors made of single-crystal silicon and SiO_2 cloth, as well as on the surface of the targets irradiated.

EXPERIMENTAL

Tungsten specimens (W-13I, PLANSEE doped by La_2O_3 to ~1%, W(111) single crystal, and W–Re alloy) were electrochemically polished and then irradiated in the MKT plasma electrodynamical accelerator by ten shots of a deuterium plasma with an energy density of 300 kJ/m^2 and a duration of 60 μs. The experimental setup is presented in Fig. 1. Single-crystal silicon samples were used as collectors of the erosion products. The microstructure of the tungsten and silicon surfaces irradiated was studied with a JEOL scanning electron microscope. Chemical analysis of the erosion products was performed by X-ray micrography with the use of a Link attachment to an electron microscope.

RESULTS

Figure 2 shows the surface topography of the electrochemically polished tungsten samples. As is seen from Fig. 2a, the electrochemical polishing of W-13I uncovers grains from ~5 to 16 μm in size. La_2O_3 precipitates, most of which are spherical with a diameter varying between 2 and 6 μm, were revealed on the surface of the W + 1% La_2O_3 target (Fig. 2b). Dislocation networks are visible on the W–Re surface (Fig. 2c).

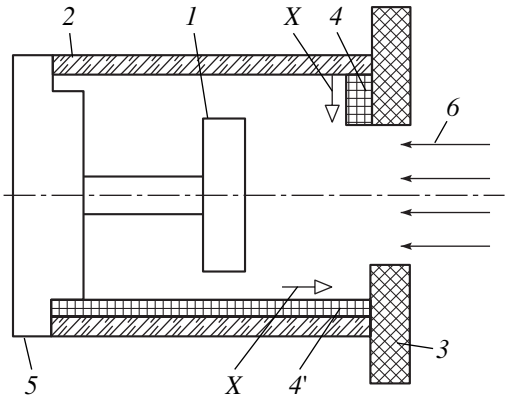


Fig. 1. Positions of the silicon collector relative to the target: (*1*) target, (*2*) quartz detector, (*3*) tungsten diaphragm, (*4*) basalt filter, (*4'*) silicon collector, (*5*) base, and (*6*) plasma stream.

1063-7842/02/4707-0841$22.00 © 2002 MAIK "Nauka/Interperiodica"

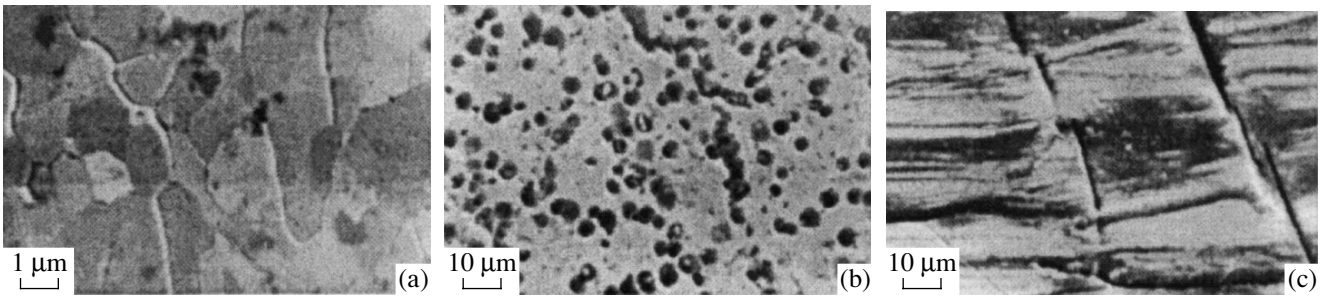


Fig. 2. Surface topography of the electrochemically polished tungsten targets: (a) W-13I, (b) W + 1% La_2O_3, and (c) W–Re.

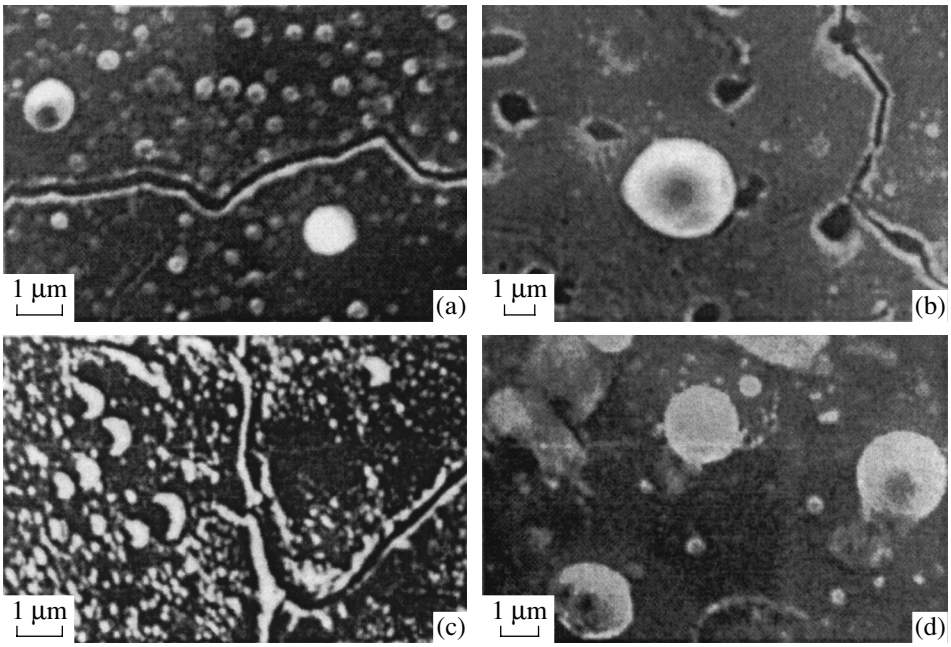


Fig. 3. Surface microstructure of the tungsten targets exposed to ten shots of the deuterium plasma with an energy density of 300 kJ/m^2: (a) W-13I, (b) W + 1% La_2O_3, (c) W–Re, and (d) W(111) face.

The microstructure of the same targets exposed to the deuterium plasma shots is shown in Fig. 3. The W-13I and W–Re targets crack along grain boundaries and dislocations, respectively (Figs. 3a, 3b). Cracks are also seen on the W + 1% La_2O_3 surface, which, in addition, exhibits voids formed at the sites of La_2O_3 precipitates. As for the (111) single-crystal tungsten surface, it shows no visible damage, the surface layer is molten, and some of the tungsten drops return from the screening plasma to the target. Tungsten drops are present on the surface of all the samples. On the micrographs of the irradiated targets, some drops lie on cracks, which indicates that the drops appear after the melting and cracking of the surface.

Figure 4a illustrates the surface of the silicon collector in experiments with polycrystalline W-13I. The collector is placed parallel to the target. Its surface is covered by fine drops with a maximal diameter of no more than 1 μm. The drops on the target surface and those on the collector are close in size. This is clearly seen in Fig. 5, which demonstrates the drop size distribution on the Si and W-13I surfaces (curves *1* and *2*, respectively). Both distributions *1* and *2* peak between 0.1 and 0.125 μm. Figure 5 also shows the drop size distribution for polycrystalline tungsten (curve *3*) from our previous study [1], where the arrangement of the erosion product collectors (basalt filter and silicon) differed from that used in these experiments.

Then, the collector was positioned in front of the target at a small angle to the normal to the surface. A micrograph of the silicon collector surface with tungsten drops captured is shown in Fig. 4b. From Figs. 4a and 4b, it follows that both the size of the drops captured and the position of the distribution peak depend essentially on the position of the collector. In contrast to distribution *1* in Fig. 5, drops with a diameter of above 1 μm are detected in the case illustrated by Fig. 4b. Curve *4* in Fig. 5 displays the size distribution

of the erosion products that returned from the screening plasma to the W(111) target.

The diffraction pattern of the drop (Fig. 6) suggests that the erosion product is tungsten (the molybdenum impurity is likely to peel off from the chamber walls, since molybdenum screws were used to fix the samples).

Thus, from Fig. 5, it follows that

(1) fine drops either return (curve *2* peaks in the size range from 0.1 to 0.125 μm) or fly away in a direction nearly parallel to the surface (curve *1* peaks in the same range of sizes);

(2) coarse drops leave the surface in the near-normal direction (curve *3*). The size distribution of the drops that are captured by the filter placed in front of the target at a nearly right angle to the surface has a maximum in the range from 0.7 to 1.0 μm;

(3) drops of a diameter <0.125 μm are absent on the single-crystal tungsten surface (curve *4*).

DISCUSSION

Waves appearing at some sites of the molten metal surface exposed to the plasma streams simulating plasma disruption are capillary waves due to the tangential instability on the liquid surface over which the vapor (plasma) flows. This flow arises because of the pressure gradient in the screening plasma and is directed outward from the center. According to the theory developed in [2], the frequency of such waves is given by

$$\omega = \kappa U(\rho'/(\rho + \rho')), \quad (1)$$

where U and ρ' are the velocity and density of the vapor above the surface, ρ is the density of the liquid, and κ is the wave vector.

The waves with $\kappa_{max} = 2\rho' U^2/3\alpha$ and $\lambda_{max} = 2\pi/\kappa_{max} = 3\pi\alpha/\rho' U^2$ have a maximal increment which equals

$$\gamma = 2\rho' U^3/3\alpha(\rho'/3\rho)^{1/2}.$$

For the parameters of the screening plasma $n_i = 10^{17}$ cm^{-3} and $U = 3 \times 10^5$ cm/s, $\omega = 10^3$ s^{-1}, $\gamma = 10^6$ s^{-1}, and $\lambda = 20$–30 μm, which is in close agreement with the wavelengths observed in the experiment.

Typical sizes of the escaping drops are less than the wavelength by 1–1.5 orders of magnitude, which indicates that only the tops, or crests, of the waves are sputtered. The wind pressure $P = \rho' U^2$ acts only on the leeward of the wave and blows off the crest. The drop sizes thus estimated are actually one order of magnitude less than the wavelength.

In this case, the escaped drop moves parallel to the surface and has a velocity

$$V_{\parallel} = U(\rho'/2\pi\rho)^{1/2} \cong 3 \times 10^2 \text{ cm/s}.$$

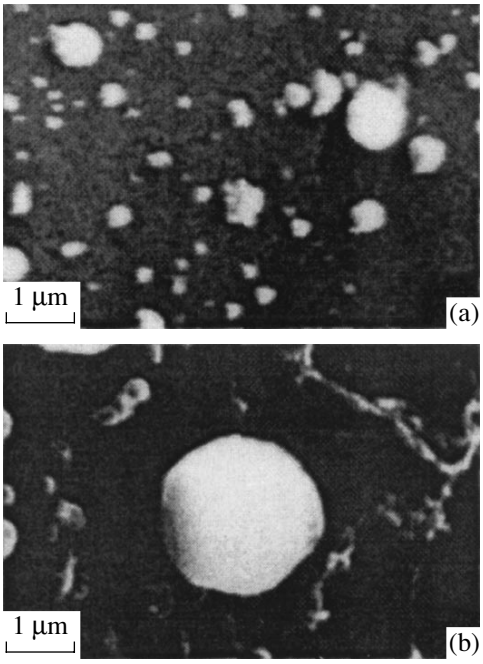


Fig. 4. Surfaces of the silicon collectors positioned near the target (a) parallel and (b) normal to it for the experiment with polycrystalline tungsten.

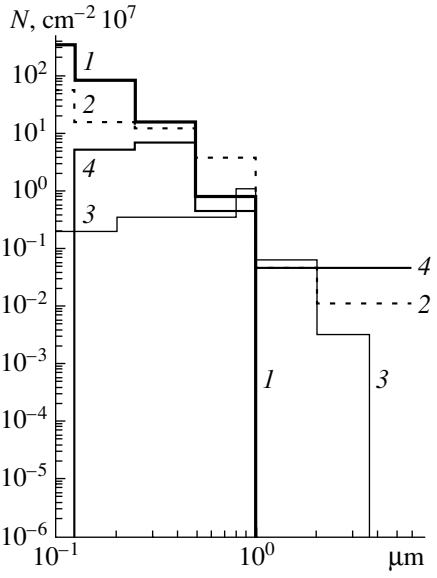


Fig. 5. Size distribution of the drops (*1*, *4*) on the Si collector surface and on the (*2*) W-13I and (*3*) polycrystalline W [1] targets irradiated.

The normal component of the velocity (the rise of the wave) is also nonzero due to the development of instability:

$$V_{\perp} = a\gamma = 2a\rho' U^3/3\alpha(\rho'/3\rho)^{1/2}$$
$$= U\pi^2(\rho'/3\rho)^{1/2} \cong 10^3 \text{ cm/s},$$

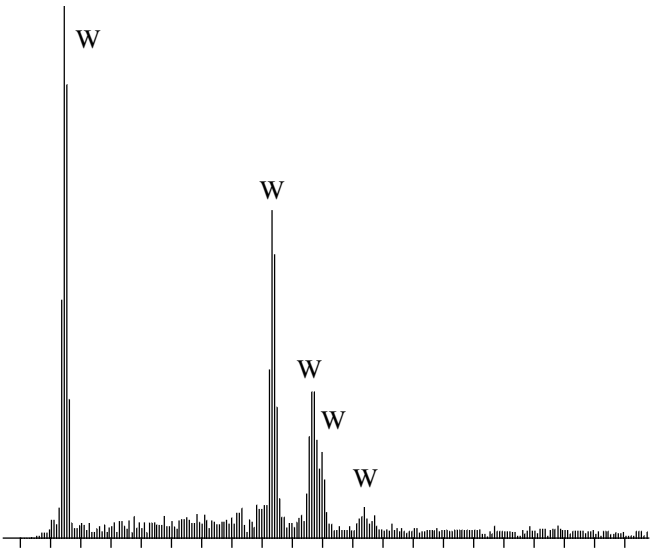


Fig. 6. Diffraction pattern of the drop shown in Fig. 3b.

where a is the wave amplitude taken to be $a = \lambda/4$ at the instant of separation.

Thus, the total velocity of the drop is the vector sum $V = V_\perp + V_\parallel$, $V \cong 10^3$ cm/s.

It has been shown [3] that the normal velocity of the drop decreases so that, after the drop has passed the Knudsen layer (which is on the order of the ion or atom range), its velocity in the screening plasma becomes

$$V_1 = (V_\perp^2 + 1.5pl/(\rho\Delta))^{1/2}.$$

Hence, the finer the drop (smaller Δ), the more effective the deceleration. The minimal size of a drop that can overcome the Knudsen layer is

$$\Delta_{\min} = 1.5pl/(\rho V_0^2).$$

Drops of size $\Delta < \Delta_{\min}$ return to the surface. That is why fine drops prevail among those that return to the sample surface or settle on the silicon collector placed near the sample.

Note that the velocity of the tungsten drops, $\sim 10^3$ cm/s, is about one order of magnitude lower than that of graphite fragments caused by brittle fracture. As a result, the drops travel a small distance ($\approx 5 \times 10^{-2}$ cm) from the surface within the pulse duration. At such distances, the temperature of the screening plasma is not very high and the drops do not burn out during the pulse. Because of this, the tungsten drops are easier to observe than the graphite dust, for example, in an MK-200 setup.

Sometimes, the surface of the samples subjected to the plasma shots was covered by the drops but the waves were absent. This may be attributed to the fact that the time it takes for the molten layer to cool,

$$t = (h^2/\chi)[(T_0^2/T_m^2 - 1)]$$

(h is the melt thickness, χ is the thermal diffusivity, T_0 is the surface temperature during the pulse, and T_m is the melting point) is sufficient for the waves to decay. The decay time is estimated as

$$\gamma^{-1} = 2\nu\kappa^2 \cong 10^{-4} \text{ s}.$$

Usually, the drops are seen if the molten layer is thin (h is small).

CONCLUSIONS

(1) Under conditions simulating plasma disruption, the tungsten surface erodes in the dropwise manner.

(2) Fine drops either return to the surface or fly away in a direction nearly parallel to the surface. Coarse drops leave the surface at nearly right angles to the surface.

(3) The dropwise erosion depends on the type of the tungsten material. In particular, drops with sizes <0.125 μm, which are typical of polycrystalline tungsten, are absent on the single-crystal tungsten surface.

(4) The experimentally found wavelengths and sizes of the drops are in close agreement with the predictions of the theory of capillary waves generated by screening plasma wind and with the model of detachment of the wave tops (crests).

(5) The screening plasma affects essentially the angular distribution of escaping drops and the probability of their return to the surface.

ACKNOWLEDGMENTS

Yu.V. Martynenko and P.G. Moskovkin acknowledge support from the Grant Council of the President of the Russian Federation (grant no. 00-15-96526) and from the leading scientific schools of the Russian Federation.

REFERENCES

1. M. I. Guseva, V. M. Gureev, Yu. V. Martynenko, *et al.*, *Morphological Features of Erosion Products for Graphite and Tungsten Targets Exposed by Powerful Plasma Flows. JCT/HT's Task Related Meeting on Materials Effects of Disruptions for Task T438* (Institute for Innovations and Fusion Research, Troitsk, 1999).
2. L. D. Landau and E. M. Lifshitz, *Mechanics of Continuous Media* (Gostekhizdat, Moscow, 1954).
3. Yu. V. Martynenko and P. G. Moskovkin, Vopr. At. Nauki Tekh., Ser. Termoyad. Sintez, No. 2, 31 (1999).

Translated by A. Sidorova

Technical Physics, Vol. 47, No. 7, 2002, pp. 845–850. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 52–57.

GAS DISCHARGES, PLASMA

Azimuthal Surface Waves at the Plasma–Metal Boundary in a Nonuniform Toroidal Magnetic Field

I. A. Girka

Karazin National University, pl. Svobody 4, Kharkov, 61077 Ukraine

Received May 25, 2001

Abstract—The propagation of electromagnetic surface waves with extraordinary polarization across the circular axis of a metal ring in a magnetized plasma is studied using perturbation theory. The amplitude of the fundamental mode of these waves is determined to within second-order terms, and the amplitudes of the nearest satellite modes are found correct to first-order terms. The correction to the eigenfrequency that comes from the nonuniformity of a constant toroidal magnetic field is shown be a second-order quantity. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Interest in studying the properties of electromagnetic surface waves (SWs) propagating near plasma–metal boundaries is related to the promising outlook for the use of these waves in plasma electronics [1–3]. The small toroidicity of the column of a gaseous plasma [4, 5] manifests itself in the following factors: first, spatial variations of a constant external toroidal magnetic field; second, radial displacement of magnetic surfaces in the direction of the major radius [6]; and, third, deviation of the shape of the poloidal cross sections of magnetic surfaces from circular. The effect of the second and third factors on the dispersion properties of azimuthal SWs (ASWs) propagating across a uniform axial magnetic field (and, accordingly, along the minor azimuthal angle) in a cylindrical plasma waveguide was thoroughly investigated in [7], where it was mentioned, in particular, that changing the sign of the curvature of the plasma–metal boundary surface leads to insignificant changes in the dispersion properties of ASWs. The distinctive features of the propagation of ASWs near the surface of a metal cylinder immersed in a magnetized plasma were also pointed out in [8].

The objective of the present paper is to investigate the effect of the spatial variations of an external magnetic field on the dispersion properties of ASWs propagating in the poloidal direction near the surface of a metal ring in a homogeneous magnetized plasma.

The paper is arranged as follows. In Section 1, the model is constructed and the basic equations are written out. In Section 2, the main dispersion properties of ASWs propagating around a cylindrical metal rod [8] are outlined in order to provide a starting point for further analysis. In Section 3, Maxwell's equations are solved to first order in the small toroidicity parameter. In Section 4, Maxwell's equations are solved and the dispersion relation is derived to second order in this parameter. In the conclusion, the main results are summarized.

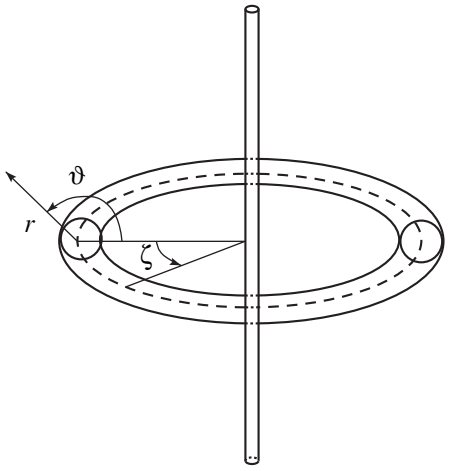
1. FORMULATION OF THE PROBLEM

We consider a perfectly conducting metal ring such that the ratio of its minor radius a to its major radius R is small, $\varepsilon_t = a/R \ll 1$. A constant toroidal magnetic field $\mathbf{B}_0 = B_{0\zeta}\mathbf{e}_\zeta$ is generated by a constant current in a straight conductor placed at the center-line of the ring, perpendicular to its symmetry plane (see figure):

$$B_{0\zeta} = B_0/[1-(r/R)\cos\vartheta], \qquad (1)$$

where (r, ϑ, ζ) are quasitoroidal coordinates.

The radial coordinate r in the poloidal cross section is measured from the circular axis of the ring, and ζ is the major azimuthal angle of the torus. The minor azimuthal angle ϑ is measured from the direction toward



Schematic of the waveguide structure and the coordinate system.

1063-7842/02/4707-0845$22.00

the center of symmetry of the ring. Assuming that this waveguide structure is symmetric with respect to the circular axis, $\partial/\partial\zeta \equiv 0$, we investigate the propagation of an E-wave, i.e., a wave with the electromagnetic field components E_r, E_ϑ, and B_ζ (recall that the surface H-wave, having the components E_ζ, B_r, and B_ϑ, cannot be excited in such a waveguide [8]).

Let us also assume that the density of the plasma surrounding the ring is uniform. The reason for making this assumption is twofold. First, it concerns an *n*-semiconductor plasma. In semiconductor physics, the SWs under consideration are called magnetoplasma polaritons and the above orientation of the external magnetic field with respect to the plasma–metal boundary corresponds to the Voigt geometry. Second, a metal ring may serve as an antenna for the excitation of ASWs in order to maintain gas discharges. In this case, over distances approximate the penetration depth of ASWs into the discharge plasma, the plasma density can also be assumed to be uniform.

The vectors of the electric induction and electric field strength are assumed to be related by the dielectric tensor of a cold, weakly collisional, magnetized plasma. In what follows, we will use the following two elements of the dielectric tensor:

$$\varepsilon_{11} = \varepsilon_0 - \sum_\alpha \frac{\omega_{p\alpha}^2}{\omega^2 - \omega_{c\alpha}^2} \equiv \varepsilon_1, \tag{2}$$

$$\varepsilon_{12} = i\sum_\alpha \frac{\omega_{c\alpha}\omega_{p\alpha}^2}{\omega(\omega^2 - \omega_{c\alpha}^2)} \equiv i\varepsilon_2, \tag{3}$$

where $\omega_{p\alpha}$ and $\omega_{c\alpha}$ are, respectively, the Langmuir and cyclotron frequencies of the particles of species α.

The plasma around the ring is assumed to be dense enough to satisfy the inequality $\omega_{pe}^2 \gg \varepsilon_0\omega_{ce}^2$. (The characteristic features of the propagation of ASWs in cylindrical magnetized plasma waveguides satisfying the opposite inequality $\omega_{pe}^2 \ll \varepsilon_0\omega_{ce}^2$ were investigated in [9].) For a gaseous plasma, we have $\varepsilon_0 = 1$. For an *n*-semiconductor plasma, the dielectric constant of the lattice is $\varepsilon_0 > 1$; as a result, the only nonzero terms in the sums over the charged-particle species in definitions (2) and (3) are those describing the conduction electrons.

We expand each of the dielectric tensor elements ε_{ij} in a series in the small toroidicity parameter ε_t:

$$\varepsilon_{1,2} = \varepsilon_{1,2}^{(0)} + \varepsilon_{1,2}^{(1)}\cos\vartheta + \varepsilon_{1,2}^{(2)}. \tag{4}$$

The main terms in expansions (4),

$$\varepsilon_1^{(0)} = \varepsilon_0 - \sum_\alpha \frac{\omega_{p\alpha}^2}{\omega^2 - \omega_{c\alpha}^{(0)2}}, \quad \varepsilon_2^{(0)} = \sum_\alpha \frac{\omega_{c\alpha}^{(0)2}\omega_{p\alpha}^2}{\omega(\omega^2 - \omega_{c\alpha}^{(0)2})}, \tag{5}$$

are coordinate-independent, $\varepsilon_{1,2}^{(0)} = \varepsilon_{1,2|B_{0\zeta}=B_0}$.

The first-order corrections are linear in the minor radius:

$$\varepsilon_1^{(1)} = -2\varepsilon_t\frac{r}{a}\sum_\alpha \frac{\omega_{p\alpha}^2\omega_{c\alpha}^{(0)2}}{(\omega^2 - \omega_{c\alpha}^{(0)2})^2},$$
$$\varepsilon_2^{(1)} = -\varepsilon_t\frac{r}{a}\sum_\alpha \frac{\omega_{c\alpha}^{(0)}\omega_{p\alpha}^2(\omega^2 - \omega_{c\alpha}^{(0)2})}{\omega(\omega^2 - \omega_{c\alpha}^{(0)2})^2}. \tag{6}$$

The second-order corrections are quadratic in the minor radius:

$$\varepsilon_1^{(2)} = \varepsilon_t^2\frac{r^2}{2a^2}\sum_\alpha \frac{\omega_{p\alpha}^2\omega_{c\alpha}^{(0)2}(\omega_{c\alpha}^{(0)2} + 3\omega^2)}{(\omega_{c\alpha}^{(0)2} - \omega^2)^3},$$
$$\varepsilon_2^{(2)} = \varepsilon_t^2\frac{r^2}{2a^2}\sum_\alpha \frac{\omega_{c\alpha}^{(0)}\omega_{p\alpha}^2\omega(\omega^2 + 3\omega_{c\alpha}^{(0)2})}{(\omega_{c\alpha}^{(0)2} - \omega^2)^3}. \tag{7}$$

In expansions (4), we can neglect the second-order terms that are proportional to $\exp(2i\vartheta)$, because they make third-order (and higher order) contributions to the eigenfrequency of the ASWs. In expressions (5)–(7), the cyclotron frequency $\omega_{c\alpha}$ is defined in terms of the uniform (nonvarying) toroidal magnetic field B_0: $\omega_{c\alpha}^{(0)} = \omega_{c\alpha|_{\varepsilon_t=0}}$, which corresponds to the zeroth order approximation.

The toroidal component of the RF magnetic field is determined from the equation

$$\frac{1}{r}\frac{\partial}{\partial r}\left(\frac{r\varepsilon_1}{\varepsilon_\perp}\frac{\partial B_\zeta}{\partial r}\right) + \frac{i}{r}\frac{\partial}{\partial\vartheta}\left(\frac{\varepsilon_2}{\varepsilon_\perp}\right)\frac{\partial B_\zeta}{\partial r} + \frac{\omega^2}{c^2}B_\zeta$$
$$+ \frac{1}{r^2}\frac{\partial}{\partial\vartheta}\left(\frac{\varepsilon_1}{\varepsilon_2}\frac{\partial B_\zeta}{\partial\vartheta}\right) - \frac{i}{r}\frac{\partial}{\partial r}\left(\frac{\varepsilon_2}{\varepsilon_\perp}\right)\frac{\partial B_\zeta}{\partial\vartheta} = 0, \tag{8}$$

where $\varepsilon_\perp = \varepsilon_1^2 - \varepsilon_2^2$.

In the zeroth approximation, Eq. (8) describes the independent propagation of ASWs with different azimuthal mode numbers. This circumstance and the symmetry of the problem [see expansions (4)] allow us to represent the sought-for solution to Eq. (8) in the form

$$B_\zeta = [B_\zeta^{(0)}(r) + B_\zeta^{(2)}(r) + B_\zeta^{(+1)}(r)e^{i\vartheta}$$
$$+ B_\zeta^{(-1)}(r)e^{-i\vartheta}]\exp(im\vartheta - i\omega t). \tag{9}$$

The electric field components of ASWs are described by expressions analogous to formula (9). In addition to the fundamental mode (represented by the term $\propto \exp im\vartheta$), this formula describes the two nearest satellite modes (represented by the terms $\propto \exp i(m \pm 1)\vartheta$). As will be shown below, the toroidal variations of the external magnetic field give rise to only a second-order correction $B_\zeta^{(2)}$ to the amplitude of the fundamental mode of ASWs.

For further analysis, we need the relationship of the poloidal component of the wave electric field E_ϑ to B_ζ:

$$E_\vartheta^{(0)}(r) = \frac{c}{i\omega\varepsilon_\perp^{(0)}}\left[\varepsilon_1^{(0)}\frac{dB_\zeta^{(0)}}{dr} + \frac{m\varepsilon_2^{(0)}}{r}B_\zeta^{(0)}\right], \quad (10)$$

$$E_\vartheta^{(\pm1)}(r) = \frac{c}{i\omega\varepsilon_\perp^{(0)}}\left[\varepsilon_1^{(0)}\frac{\partial B_\zeta^{(\pm1)}}{\partial r} + \frac{m\pm1}{r}\varepsilon_2^{(0)}B_\zeta^{(\pm1)} + \frac{\varepsilon_\perp^{(0)}\varepsilon_1^{(1)} - \varepsilon_\perp^{(1)}\varepsilon_1^{(0)}}{2\varepsilon_\perp^{(0)}}\frac{\partial B_\zeta^{(0)}}{\partial r} + \frac{m(\varepsilon_\perp^{(0)}\varepsilon_2^{(1)} - \varepsilon_\perp^{(1)}\varepsilon_2^{(0)})}{2r\varepsilon_\perp^{(0)}}B_\zeta^{(0)}\right], \quad (11)$$

$$\begin{aligned}E_\vartheta^{(2)}(r) = \frac{1}{\varepsilon_\perp^{(0)}}\Bigg\{&-\frac{\varepsilon_\perp^2}{2}E_\vartheta^{(0)} - \frac{\varepsilon_\perp^{(1)}}{2}(E_\vartheta^{(+1)} + E_\vartheta^{(-1)}) \\ &+ \frac{c\varepsilon_1^{(1)}}{2i\omega}\frac{\partial}{\partial r}(B_\zeta^{(+1)} + B_\zeta^{(-1)}) + \frac{c\varepsilon_1^{(0)}}{i\omega}\frac{\partial B_\zeta^{(2)}}{\partial r} \\ &+ \frac{c\varepsilon_1^{(2)}}{2i\omega}\frac{\partial B_\zeta^{(0)}}{\partial r} + \frac{cm\varepsilon_2^{(0)}}{i\omega r}B_\zeta^{(2)} + \frac{cm\varepsilon_2^{(2)}}{2i\omega r}B_\zeta^{(0)} \\ &+ \frac{cm\varepsilon_2^{(1)}}{2i\omega r}[(m+1)B_\zeta^{(+1)} + (m-1)B_\zeta^{(-1)}]\Bigg\},\end{aligned} \quad (12)$$

where we have introduced the notation

$$\begin{aligned}\varepsilon_\perp^{(0)} &= \varepsilon_1^{(0)2} - \varepsilon_2^{(0)2}, \quad \varepsilon_\perp^{(1)} = 2\varepsilon_1^{(0)}\varepsilon_1^{(1)} - 2\varepsilon_2^{(0)}\varepsilon_2^{(1)}, \\ \varepsilon_\perp^{(2)} &= \varepsilon_1^{(1)2} + 2\varepsilon_1^{(0)}\varepsilon_1^{(2)} - \varepsilon_2^{(1)2} - 2\varepsilon_2^{(0)}\varepsilon_2^{(2)}.\end{aligned} \quad (13)$$

The solution to Eq. (8) should satisfy the following boundary conditions. First, the wave field is finite everywhere outside the ring. Second, in order not to formulate the boundary conditions at the major (straight) axis of the torus, we assume that the wave fields decrease away from the ring and become negligible at distances comparable to its major radius or longer; in other words, we assume that the major radius of the ring is much larger than the penetration depth of the wave fields into the plasma. Third, the tangential (in the case at hand, poloidal) component of the wave electric field vanishes at the surface of the ring, $E_\vartheta|_{r=a} = 0$.

2. AZIMUTHAL SURFACE WAVES PROPAGATING AROUND A CIRCULAR CYLINDRICAL ROD (THE ZEROTH APPROXIMATION)

As a starting point for our analysis, we use the theory of ASWs propagating in the azimuthal direction around a cylindrical metal rod with a circular cross section of radius a [8]. The rod is assumed to be immersed in a homogeneous plasma, and a constant external magnetic field B_0 is assumed parallel to the axis of the rod. Let us briefly outline the main dispersion properties of such ASWs.

In the zeroth approximation, the amplitude $B_\zeta^{(0)}(r)$ of the fundamental harmonic is expressed through the modified Bessel function of the second kind [10]:

$$B_\zeta^{(0)}(r) = K_m(k_\perp r). \quad (14)$$

The penetration depth $k_\perp^{-1}$ of the wave field into the plasma is defined by the relationship

$$k_\perp^2 = -(\omega/c)^2\varepsilon_\perp^{(0)}/\varepsilon_1^{(0)} > 0. \quad (15)$$

The solution $\omega_m^{(0)}$ to the zero-order dispersion relation

$$D^{(0)}(\omega_m^{(0)}) = 0 \quad (16)$$

is assumed to be known. The dispersion relation (16) follows from the condition that the amplitude $E_\vartheta^{(0)}(r)$ of the fundamental harmonic of the poloidal electric field of an ASW vanishes at the plasma–metal boundary, $D^{(0)} \equiv E_\vartheta^{(0)}(a)$. The dependence of the eigenfrequency $\omega_m^{(0)}$ of the ASW on the plasma parameters in the zeroth approximation, as well as on the magnitude of the external magnetic field and the dimensions of the rod, was thoroughly investigated in [8].

The frequencies of ASWs with negative azimuthal mode numbers are low,

$$\sqrt{\omega_{pi}^2/\varepsilon_0 + \omega_{ci}^2} < \omega_m^{(0)} < |\omega_{ce}|, \quad (17)$$

while the frequencies of ASWs with positive mode numbers m are high,

$$\omega_U < \omega_m^{(0)} < 0.5|\omega_{ce}| + \sqrt{0.25\omega_{ce}^2 + \omega_{pe}^2/\varepsilon_0}, \quad (18)$$

where $\omega_U = \sqrt{\omega_{ce}^2 + \omega_{pe}^2/\varepsilon_0}$ is the upper hybrid frequency.

3. AMPLITUDES OF THE SATELLITE MODES (THE FIRST APPROXIMATION)

We assume that the amplitudes $B_\zeta^{(\pm1)}$ of the satellite modes are first-order quantities. We substitute expression (9) for the field amplitude B_ζ and expansions (4) for the dielectric tensor elements ε_{ij} into Eq. (8). In the resulting equation, we single out first-order terms, which turn out to be proportional either to $\exp[i(m+1)\varphi]$ or to $\exp[i(m-1)\varphi]$. In other words, taking into account first-order terms does not change the fundamental mode amplitude. Consequently, in the first approximation, the toroidal variations of the external magnetic field do not contribute to the eigenfrequency of the ASW. These first-order terms enter the inhomo-

geneous Bessel equations for $B_\zeta^{(\pm 1)}(r)$ with the known right-hand side. The inhomogeneous Bessel equations have the following solutions, which satisfy the boundary condition for the ASW field to be finite at infinite distances from the ring:

$$B_\zeta^{(\pm 1)}(r) = K_{m\pm 1}(k_\perp r)$$

$$\times \int_a^r r I_{m\pm 1}(k_\perp r)\hat{L}_\mp^{(1)} K_m(k_\perp r)dr + C_\pm K_{m\pm 1}(k_\perp r) \quad (19)$$

$$- I_{m\pm 1}(k_\perp r)\int_\infty^r r K_{m\pm 1}(k_\perp r)\hat{L}_\mp^{(1)} K_m(k_\perp r)dr,$$

where the operator $\hat{L}_\pm^{(1)}$ has the form

$$\hat{L}_\pm^{(1)} = (2\varepsilon_1^{(0)}\varepsilon_\perp^{(0)})^{-1}\left[\frac{\varepsilon_\perp^{(1)}}{r}\frac{\partial}{\partial r}\left(r\varepsilon_1^{(0)}\frac{\partial}{\partial r}\right)\right.$$

$$+ \frac{\varepsilon_\perp^{(0)}}{r}\frac{\partial}{\partial r}\left(r\varepsilon_1^{(1)}\frac{\partial}{\partial r}\right) - \varepsilon_1^{(0)}\frac{\partial\varepsilon_\perp^{(1)}}{\partial r}\frac{\partial}{\partial r}$$

$$- \frac{\varepsilon_1^{(0)}}{r}\frac{\partial\varepsilon_2^{(1)}}{\partial r}\hat{\Theta} + \frac{\varepsilon_2^{(0)}}{r}\frac{\partial\varepsilon_\perp^{(1)}}{\partial r}\hat{\Theta} - \frac{\varepsilon_\perp^{(1)}}{r^2}\varepsilon_1^{(0)}\hat{\Theta}^2 \quad (20)$$

$$- \frac{\varepsilon_\perp^{(0)}\varepsilon_1^{(1)}}{r^2}\hat{\Theta}^2 + 2\frac{\omega^2}{c^2}\varepsilon_\perp^{(0)}\varepsilon_\perp^{(1)} \mp \frac{\varepsilon_\perp^{(0)}\varepsilon_\perp^{(1)}}{r^2}\hat{\Theta} \pm \frac{\varepsilon_1^{(0)}\varepsilon_\perp^{(1)}}{r^2}\hat{\Theta}$$

$$\left. \pm \frac{\varepsilon_2^{(1)}\varepsilon_\perp^{(0)}}{r}\frac{\partial}{\partial r} \mp \frac{\varepsilon_2^{(0)}\varepsilon_\perp^{(1)}}{r}\frac{\partial}{\partial r}\right], \quad \hat{\Theta} I_m = -mI_m.$$

The integration constants $C_\pm$ can be found from the condition that the amplitudes of the first satellite harmonics of the poloidal electric field vanish at the metal surface of the ring, $E_\vartheta^{(\pm 1)}(a) = 0$:

$$C_\pm = \left[\frac{m+1}{a}\varepsilon_2^{(0)} K_{m\pm 1}(k_\perp a) - \varepsilon_1^{(0)} k_\perp K'_{m\pm 1}(k_\perp a)\right]^{-1}$$

$$\times \left\{\varepsilon_1^{(0)} k_\perp I'_{m\pm 1}(k_\perp a)\int_\infty^a r K_{m\pm 1} k_\perp r \hat{L}_\mp^{(1)} K_m(k_\perp r)dr\right.$$

$$+ \frac{m\pm 1}{a}\varepsilon_2^{(0)} I_{m\pm 1}(k_\perp a)\int_\infty^a r K_{m\pm 1}(k_\perp r)\hat{L}_\mp^{(1)} K_m(k_\perp r)dr \quad (21)$$

$$+ k_\perp K'_m(k_\perp a)\frac{\varepsilon_\perp^{(0)}\varepsilon_1^{(1)} - \varepsilon_\perp^{(1)}\varepsilon_1^{(0)}}{2\varepsilon_\perp^{(0)}}$$

$$\left. - \frac{m(\varepsilon_\perp^{(0)}\varepsilon_2^{(1)} - \varepsilon_\perp^{(1)}\varepsilon_2^{(0)})}{2a\varepsilon_\perp^{(0)}} K_m(k_\perp a)\right\}\Big|_{\omega = \omega_m^{(0)}}.$$

4. CORRECTION TO THE EIGENFREQUENCY OF THE ASW IN THE CASE OF TOROIDAL VARIATIONS OF THE EXTERNAL MAGNETIC FIELD (THE SECOND APPROXIMATION)

The second-order terms in Eq. (8) are proportional to one of the three phase factors: $\exp im\vartheta$ or $\exp i(m \pm 2)\vartheta$. Retaining only the terms with $\exp i(m \pm 2)\vartheta$ yields equations for the amplitudes of the second satellite modes. In expansion (9), these modes are neglected because they make third-order (and higher-order) contributions to the eigenfrequency of the ASW. Keeping the second-order terms that are proportional to $\exp im\vartheta$ gives an inhomogeneous differential Bessel equation for the correction $B_\zeta^{(2)}(r)$ to the amplitude of the fundamental harmonic of the ASW field. A solution to this equation satisfying the condition for the ASW fields to be finite at infinite distances from the ring can be found by the method of variation of constants:

$$B_\zeta^{(2)}(r) = K_m(k_\perp r)$$

$$\times \int_a^r r I_m(k_\perp r)(\hat{L}_-^{(1)} B_\zeta^{(-1)} + \hat{L}_+^{(1)} B_\zeta^{(+1)} + \hat{L}^{(2)} B_\zeta^{(0)})dr \quad (22)$$

$$+ C_2 K_m(k_\perp r) - I_m(k_\perp r)$$

$$\times \int_\infty^r r K_m(k_\perp r)(\hat{L}_-^{(1)} B_\zeta^{(-1)} + \hat{L}_+^{(1)} B_\zeta^{(+1)} + \hat{L}^{(2)} B_\zeta^{(0)})dr,$$

where the operator $\hat{L}^{(2)}$ is defined as

$$\hat{L}^{(2)} = [2\varepsilon_1^{(0)}\varepsilon_\perp^{(0)}]^{-1}\left\{\frac{\varepsilon_\perp^{(2)}}{r}\frac{\partial}{\partial r}\left[r\varepsilon_1^{(0)}\frac{\partial}{\partial r}\right] + \frac{\varepsilon_\perp^{(1)}}{r}\frac{\partial}{\partial r}\left[r\varepsilon_1^{(1)}\frac{\partial}{\partial r}\right]\right.$$

$$+ \frac{\varepsilon_\perp^{(0)}}{r}\frac{\partial}{\partial r}\left[r\varepsilon_1^{(2)}\frac{\partial}{\partial r}\right] - \frac{\varepsilon_1^{(1)}\varepsilon_\perp^{(1)}}{r}\frac{\partial}{\partial r} - \frac{2\varepsilon_1^{(0)}\varepsilon_\perp^{(2)}}{r}\frac{\partial}{\partial r} - \frac{\varepsilon_\perp^{(1)}\varepsilon_2^{(1)}}{r^2}\hat{\theta}$$

$$- \frac{2\varepsilon_\perp^{(0)}\varepsilon_2^{(2)}}{r^2}\hat{\theta} + \frac{\varepsilon_2^{(1)}\varepsilon_\perp^{(1)}}{r^2}\hat{\theta} + \frac{2\varepsilon_2^{(0)}\varepsilon_\perp^{(2)}}{r^2}\hat{\theta} - \frac{\varepsilon_\perp^{(2)}\varepsilon_1^{(0)}}{r^2}\hat{\theta}^2 \quad (23)$$

$$\left. - \frac{\varepsilon_\perp^{(1)}\varepsilon_1^{(1)}}{r^2}\hat{\theta}^2 - \frac{\varepsilon_\perp^{(0)}\varepsilon_1^{(2)}}{r^2}\hat{\theta}^2 + 2\frac{\omega^2}{c^2}\varepsilon_\perp^{(2)}\varepsilon_\perp^{(0)} + \frac{\omega^2}{c^2}\varepsilon_\perp^{(1)}\varepsilon_\perp^{(1)}\right\}.$$

The boundary conditions formulated in Section 2 are insufficient to determine the integration constant C_2 in solution (22). This constant, which is a second-order quantity, can be found from the condition

$$\int_a^\infty [2B_\zeta^{(2)} K_m(k_\perp r) + (B_\zeta^{(+1)})^2 + (B_\zeta^{(-1)})^2] r dr = 0, \quad (24)$$

implying that the energy of the wave magnetic field calculated in the second approximation should coincide with the energy calculated in the zeroth approximation.

In quantum mechanics, this requirement is analogous to the normalization condition for the wave function [11]. At the same time, the form of the second-order dispersion relation, which will be derived below, is independent of the value of the constant C_2.

In the second approximation, the boundary condition for the fundamental harmonic of the poloidal component of the wave electric field at the metal surface of the ring has the form

$$D^{(0)} + D^{(2)} = 0, \tag{25}$$

where $D^{(2)} = E_{\vartheta}^{(2)}(a)$.

The toroidal geometry of the ring gives rise to only a second-order correction $\omega_m^{(2)}$ to the eigenfrequency of the ASW:

$$\omega_m^{(2)} = -D^{(2)}(\partial D^{(0)}/\partial\omega)^{-1}\Big|_{\omega=\omega_m^{(0)}}. \tag{26}$$

The applicability condition for the approach developed here to determine the dispersion properties of ASWs propagating around a metal ring immersed in a magnetized plasma implies the smallness of the second-order correction with respect to the zero-order eigenfrequency: $|\omega_m^{(2)}| \ll \omega_m^{(0)}$.

The main mathematical advantage of the resulting formula (26) is that it reduces the problem with two-dimensional variations of the toroidal magnetic field (in the poloidal cross section) to the problem with one-dimensional variations (in the radial direction). The obvious drawback of formula (26) is its complexity. In the limiting case of thick rings ($R \gg a \gg |m|/k_\perp$), the eigenfrequency of the ASW can be calculated in the low-frequency (LF) range (17) using the asymptotic expressions

$$\omega_m^{(0)} = |m\omega_{ce}|\delta/a, \tag{27}$$

$$\omega_m^{(2)} = 0.25\varepsilon_t^2\omega_m^{(0)}. \tag{28}$$

In the high-frequency (HF) range (18), the eigenfrequency can be calculated from the asymptotic expressions

$$\omega_m^{(0)} = \sqrt{\omega_U^2 + m^2c^2/a^2}, \tag{29}$$

$$\omega_m^{(2)} = -\frac{\varepsilon_t^2}{4}\frac{\delta^2 m^2}{a^2}\left(1 + \frac{3|\omega_{ce}|k_\perp a}{\omega_{pe}m}\right)\omega_m^{(0)}. \tag{30}$$

In these expressions, $\delta = c/\omega_{pe}$ is the skin depth. Note that, in the case of thick rings, the amplitudes of the satellite modes of an ASW have essentially the same radial dependence. Of course, analytic solution (29) to dispersion relation (16) refers to the HF range (18). This necessary condition for the existence of HF ASWs can be formulated as the inequality

$$(\delta m/a)^2 < |\omega_{ce}|/\omega_{pe}. \tag{31}$$

This inequality implies that HF ASWs can propagate only around sufficiently thick rings.

For the eigenfrequencies of LF ASWs propagating around thin ($a \ll |m|/k_\perp \ll R$) rings, we obtain the following asymptotic expressions:

$$\omega_m^{(0)} = |\omega_{ce}|[1 - 0.5a^2/(m\delta)^2], \tag{32}$$

$$\omega^{(2)} = -\frac{\varepsilon_t^2}{8m^2}(2m^2 + m + 3)\omega_m^{(0)}. \tag{33}$$

In the limiting case of thin rings, the radial dependences of the amplitudes of the satellite modes of an ASW are essentially different, $B_\zeta^{(\pm1)}(r) \propto r^{-|m\pm1|}$; however, at the metal surface of the ring, these amplitudes are of the same order of magnitude because of the corresponding difference in the numerical factors $C_\pm$.

CONCLUSION

It has been established that ASWs can exist in a magnetized plasma surrounding a metal ring. Spatial distributions of the ASW fields have been derived [see expressions (10)–(12), (19), and (22)]. It has been shown that the toroidal variations of the external magnetic field in this waveguide structure cause ASWs to propagate as wave packets. In such a packet, the amplitudes A_{m+N} of the satellite modes are proportional to $\propto\exp[i(m+N)\vartheta]$, so that they are small ($A_{m+N} \sim \varepsilon_t^N A_m$) in comparison with the fundamental mode amplitude A_m, which is proportional to $\propto\exp[im\vartheta]$. The corrections to the fundamental mode amplitude that come from the toroidal variations of the external magnetic field are second-order in the small toroidicity parameter ε_t.

The correction to the eigenfrequency of ASWs caused by toroidal variations of the external magnetic field has been determined [see expression (26)]. This correction is found to be a second-order quantity. In the limiting cases of thick and thin rings, simple analytic expressions (27)–(30), (32), and (33) have been obtained for the eigenfrequencies of ASWs in the two frequency ranges in which the waves can exist.

It has been shown that HF ASWs cannot propagate around relatively thin metal rings whose minor radius does not satisfy inequality (31).

Note that ASWs can also exist in toroidal metal waveguides entirely filled with plasma. In this case, the correction to the eigenfrequency of ASWs that is introduced by the toroidal variations of the external magnetic field is equal in magnitude but opposite in sign to the correction obtained in the present paper.

REFERENCES

1. E. P. Kurushin and E. I. Nefedov, *Electrodynamics of Anisotropic Waveguide Structures* (Nauka, Moscow, 1983).
2. N. N. Beletskiĭ, A. A. Bulgakov, S. I. Khankina, and V. M. Yakovenko, *Plasma Instabilities and Nonlinear Phenomena in Semiconductors* (Naukova Dumka, Kiev, 1984).
3. N. A. Azarenkov, A. N. Kondratenko, and K. N. Ostrikov, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **63**, 335 (1993).
4. I. A. Girka and K. N. Stepanov, Ukr. Fiz. Zh. **36**, 1051 (1991).
5. I. A. Girka and V. A. Girka, in *Proceedings and Contributed Papers of the International Conference on Plasma Physics Combined with the 6th Latin American Workshop on Plasma Physics, Foz do Iguassu, Brazil, 1994*, Vol. 1, p. 21.
6. B. B. Kadomtsev and O. P. Pogutse, in *Reviews of Plasma Physics*, Ed. by M. A. Leontovich (Atomizdat, Moscow, 1967; Consultants Bureau, New York, 1970), Vol. 5.
7. V. A. Girka, I. A. Girka, and I. V. Pavlenko, Fiz. Plazmy **23**, 1037 (1997) [Plasma Phys. Rep. **23**, 959 (1997)].
8. V. A. Girka, I. A. Girka, A. N. Kondratenko, and V. I. Tkachenko, Radiotekh. Élektron. (Moscow) **34**, 296 (1989).
9. I. A. Girka and P. K. Kovtun, Zh. Tekh. Fiz. **68** (12), 25 (1998) [Tech. Phys. **43**, 1424 (1998)].
10. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (National Bureau of Standards, Washington, 1964; Nauka, Moscow, 1979).
11. L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Nonrelativistic Theory* (Nauka, Moscow, 1989; Pergamon, New York, 1977).

Translated by O. Khadin

Technical Physics, Vol. 47, No. 7, 2002, pp. 851–855. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 58–62.

GAS DISCHARGES, PLASMA

Surface Streamer Microwave Discharge

K. V. Aleksandrov, L. P. Grachev, I. I. Esakov, and K. V. Khodataev

State Unitary Enterprise Moscow Radiotechnical Institute, Russian Academy of Sciences, Moscow, 113519 Russia

e-mail: esakov@dataforce.net

Received September 24, 2001

Abstract—An atmospheric-pressure pulsed microwave discharge on the surface of an insulating plate and between two insulating films is considered. The discharge is initiated in a quasi-optical microwave beam away from the beam-forming elements. The TEM field of the beam is less than the critical breakdown value. The initiator is a metallic dipole placed either directly on the plate surface facing away from the radiation source or between the films. The discharge has the form of branching streamer channels adjacent to the surface and filling up the cross section of the microwave beam with a field level higher than the value separating the subcritical and highly subcritical forms of the microwave streamer discharge. © *2002 MAIK "Nauka/Interperiodica".*

It is known that a microwave discharge can be initiated in a quasi-optical electromagnetic (EM) cm-wave beam even at electric field amplitudes E_0 lower than the critical breakdown value E_c [1]. The discharge can be initiated, for example, by a thin metallic dipole placed into an EM beam [2]. If E_0 is slightly below E_c and the air pressure p exceeds several tens of torr, a subcritical discharge arising at the ends of the initiating dipole propagates counter to the radiation in the form of branching streamer channels growing with a velocity V_{str} of about 10^6 cm/s and gradually fills the beam volume [3]. The front of this discharge travels towards the radiation source with a velocity V_{fr} of about 10^5 cm/s, shielding the initiating dipole and the interior of the discharge. For such a self-developing subcritical volume discharge separated from the initiating dipole, intense ionization processes take place at its approximately $\lambda/2$ thick front, where λ is the radiation wavelength.

When the initial field E_0 becomes lower than E_c, the microwave discharge cannot separate from the initiating dipole. At air pressures p higher than several tens of torr and a pulsed field duration τ_{pul} of several tens of microseconds, such a highly subcritical discharge splits into a set of streamer channels. However, unlike a subcritical discharge with $E_0 < E_c$, the streamer channels of a highly subcritical discharge ($E_0 \ll E_c$) are in contact with the ends of the dipole all through τ_{pul}. It has been found [1] that, at atmospheric pressure (E_c = 42 kV/cm), the threshold field E_{th} separating the subcritical and highly subcritical forms of a microwave discharge is 2.5 kV/cm at λ = 8.9 cm.

In our experiments, a dielectric transparent shutter was placed across the beam propagating toward the discharge source in a subcritical initial field $E_{th} < E_0 < E_c$. The discharge was initiated by a dipole placed directly on the shutter surface facing away from the radiation source. The shape of the discharge (from which its structure can be judged), as well as its ability to separate from the initiator, penetrate the barrier, and propagate towards the source, were studied. In the experiments, the material and thickness of the shutter, as well as the number of initiators, were changed. Sometimes, the initiator was placed between two closely spaced dielectric surfaces.

The experimental setup was described in detail in [3]. A linearly polarized radiation with the TEM field, λ = 8.9 cm and τ_{pul} = 40 μs, is focused. The focus is several tens of centimeters distant from the nearest parts of the setup. In the focal plane, which is perpendicular to the EM radiation direction, the field is roughly Gaussian with the maximum amplitude at the focus, E_0 = 6.5 kV/cm. The sizes of the focus spot measured at the site where the field amplitude diminishes e times are x_0 = 5.2 cm along the vector $\mathbf{E}_0$ and y_0 = 2.5 cm perpendicular to the vector $\mathbf{E}_0$. The vector $\mathbf{E}_0$ is vertical. We used microwave shots generated at an interval of 1 min or more. The experiments were performed in atmospheric air. A dipole initiating a microwave discharge was placed in the focus spot of the EM beam parallel to $\mathbf{E}_0$. The dipole, made of a tungsten wire with a diameter $2a = 3 \times 10^{-2}$ cm and a length $2l$ = 1 cm, was fixed to the surface of a thin (in comparison with $\lambda/2$) radioparent insulator on the side facing away from the radiation source. The insulator was either KV quartz glass with a permittivity $\varepsilon \approx 4$ and a loss tangent $\tan\delta \approx 10^{-4}$ or polyethylene with $\varepsilon \approx 2.2$ and $\tan\delta \approx 10^{-4}$ [4]. The insulating film was placed normal to the radiation direction.

The discharge area was photographed with an exposure time more than τ_{pul}. In such "integral" photos, the radiation propagates from left to right, while the vector $\mathbf{E}_0$ is vertical as in the experiments. The photos were

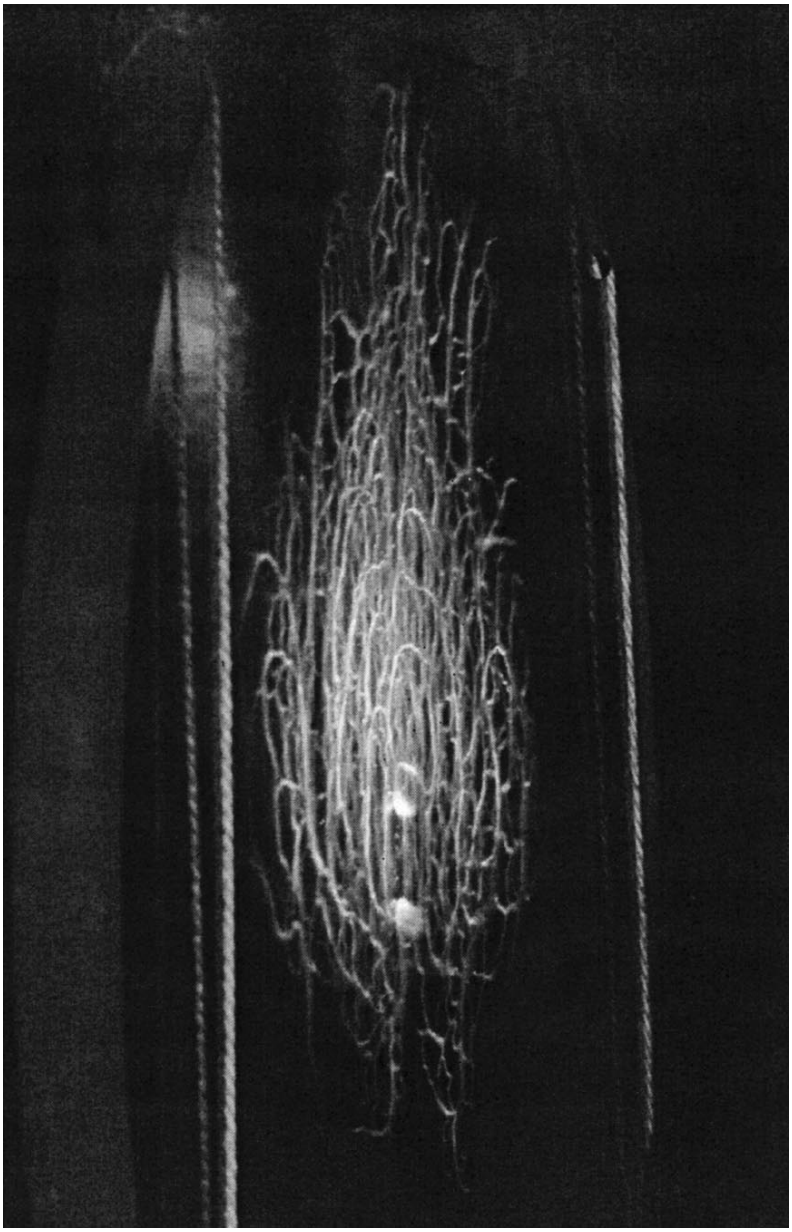


Fig. 1. Microwave streamer discharge initiated by one dipole on the surface of a quartz plate 1 cm thick (p = 760 torr).

taken on the dipole side along a line lying in the plane almost perpendicular to $\mathbf{E}_0$ and crossing the EM beam axis at a small angle to the plane of the insulator. The size $2l$ of the initiator can serve as the vertical scale of the image. To determine the horizontal scale of the image, the aspect angle should be taken into account. Therefore, sizes allowing one to estimate this scale are given for each of the images.

In all the experiments, the discharge was found to develop only on the surface where the initiator was fixed. It had the form of streamer channels with a maximum diameter of several hundredths of a centimeter filling up the cross sections of the beam. In other words, the discharge can leave the initiator and extend as a self-sustained surface streamer subcritical microwave discharge. The streamer channels were blue violet (which is also typical of streamer channels of volume microwave discharge in air) with bright white portions extended largely along $\mathbf{E}_0$.

A typical discharge on the surface of the quartz disk with a diameter of 20 cm and a thickness d = 1 cm is shown in Fig. 1. In the image, the disk side surface is at the left. On both sides of the discharge, one can see two vertical filaments confining the discharge in the EM beam. The spacing between the filaments, 15.5 cm, can be used as the horizontal scale. The discharge area is symmetric in the direction transverse to the vector $\mathbf{E}_0$. The total size of the discharge area is about 11 cm:

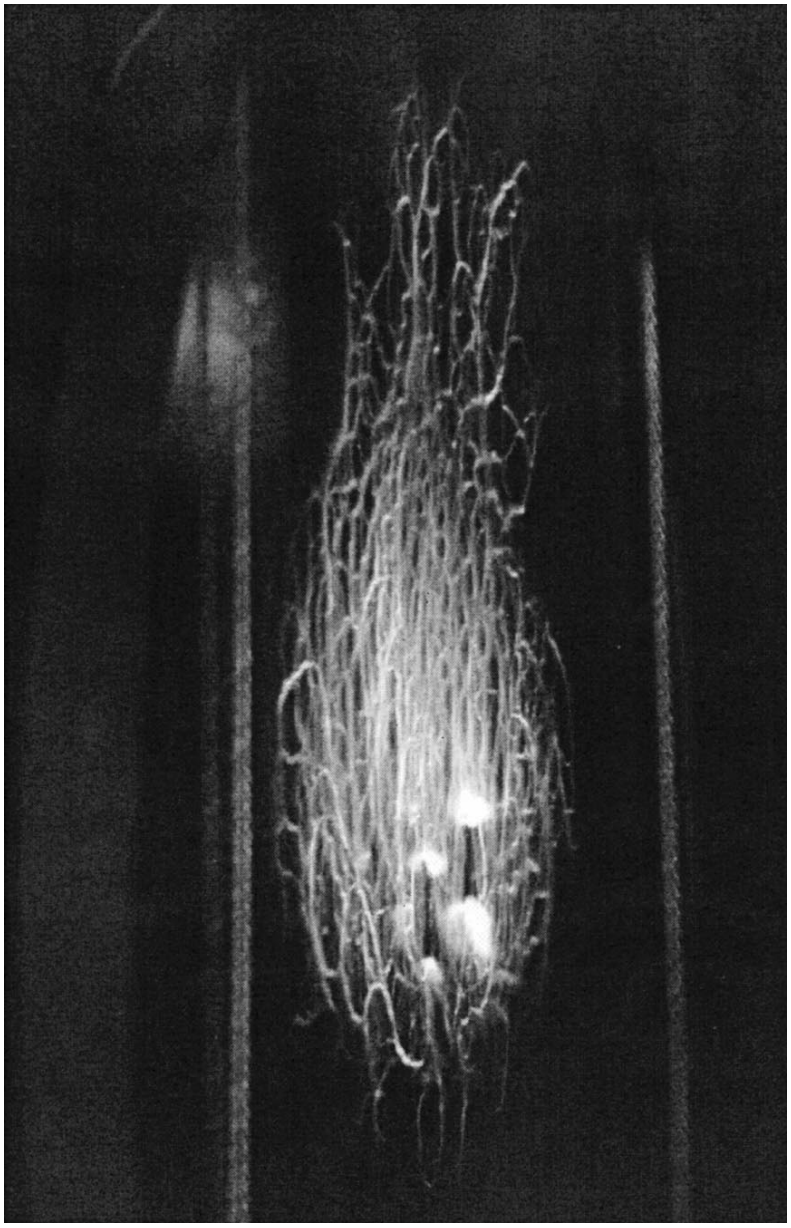


Fig. 2. Microwave streamer discharge initiated by two dipoles on the surface of a quartz plate 1 cm thick (p = 760 torr).

about 3 cm downward and 7 cm upward from the dipole center. The ends of the dipole are brighter, indicating the evaporation of its surface. The melting point of tungsten is $T \approx 3400$°C, and its boiling temperature is 5700°C [4]. Hence, the temperature of the streamer channels in contact with tungsten lies in this range.

A discharge initiated by two dipoles fixed on the same quartz disk within the focus spot of the beam is shown in Fig. 2. The dipoles are 1.5 cm apart. The shape of the discharge and its size remain almost unchanged. As before, one sees streamer channels adjacent to the quartz surface and filling up the beam cross section.

A discharge ignited between two polyethylene sheets with a thickness $d \approx 10^{-2}$ cm is shown in Fig. 3. The sheets were punched and freely hung by their upper ends. On the left of the image, one can see elliptical holes (1×0.6 cm) used for hanging. The discharge in the *folder* is seen to be qualitatively similar to the surface microwave discharge shown in Figs. 1 and 2. However, the former is more symmetric about the initiator and has brighter and thicker streamer channels in its upper left part. Obviously, this is associated with the melting of the film surface due to high-temperature discharge channels because of the gap between the films being narrower in this area. At the center of the folder, the gap Δ is defined by the diameter $2a$ of the dipole,

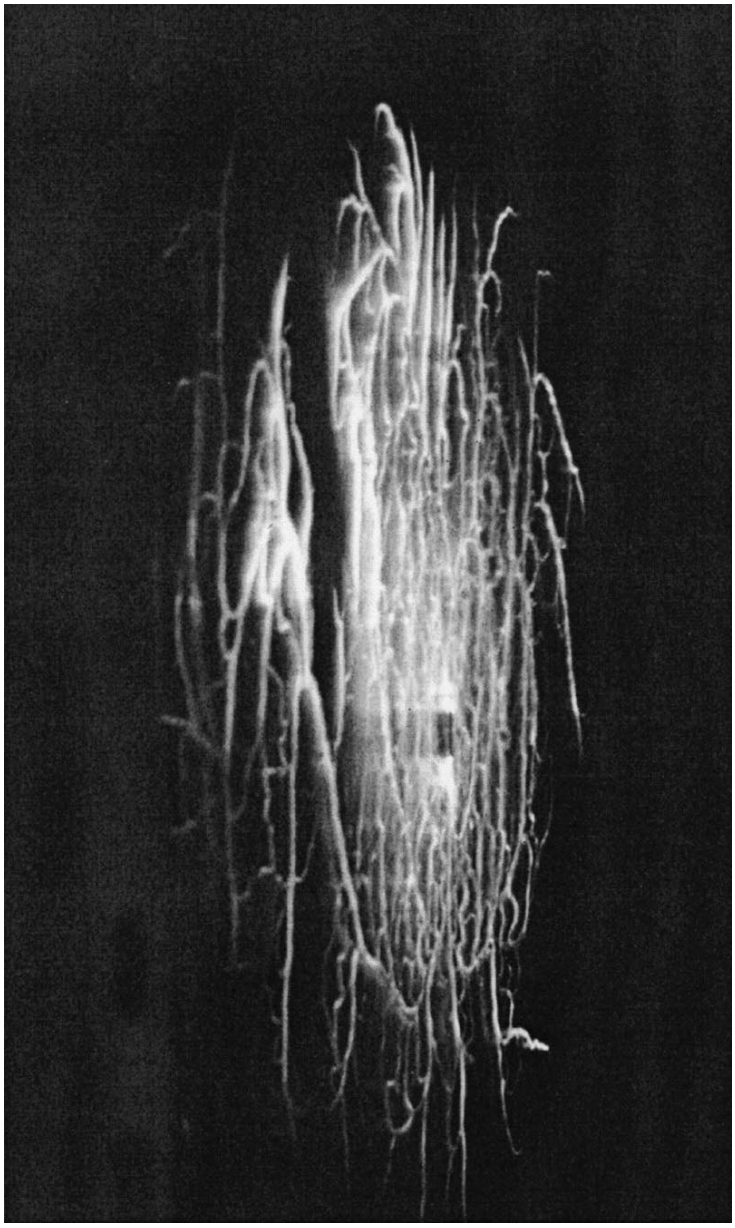


Fig. 3. Microwave streamer discharge initiated by one dipole in the gap between two polyethylene sheetsness 10^{-2}–9×10^{-3} cm thick (p = 760 torr).

$\Delta = 2a$. From this image, one can estimate the maximum vertical and horizontal sizes of the discharge area: $2x_{dis}$ = 10.5 cm and $2y_{dis}$ = 6 cm.

As was mentioned, for the given λ of the EM field in air, the threshold field separating the subcritical and highly subcritical forms of the discharge is E_{th} = 2.5 kV/cm. Under our experimental conditions, the field amplitude in the focal plane of the beam falls to this value when the focus is shifted by about x_0 along the vector $\mathbf{E}_0$ and by y_0 perpendicular to it. Obviously, the coincidence of the values x_{dis} and y_{dis} for the discharge in the *folder* shown in Fig. 3 with x_0 and y_0 is not accidental. This means that the same value of the field E_{th} specified the boundary of the surface streamer subcritical discharge with the developed structure.

For the discharges on the quartz shown in Figs. 1 and 2, the vertical size also approximately equals $2x_0$, while their vertical asymmetry about the dipole is probably due to its inaccurate position in the EM beam focus. The reason for the large horizontal size of these discharges is unclear and calls for further investigations. It can be assumed that the thick quartz disk influences the field distribution near the focus.

The propagation of streamer discharges in various ranges of the EM field wavelength, including microwave and constant fields, is caused by the enhancement

mechanism field at the ends of a streamer. The area of enhanced field with $E > E_c$ has a certain length. Our experiments with polyethylene films having a thickness d of several hundredths of a centimeter did not reveal the transition of the discharge to the side opposite to the initiator. Hence, the size of the enhanced field area at the ends of the streamer channels is less than d. This means that the diameter of the head of the streamer channels is also no more than 10^{-3} cm at atmospheric pressure.

Thus, we observed a surface subcritical pulsed microwave discharge with a developed structure in air at atmospheric pressure. The discharge was initiated in a linearly polarized quasi-optical beam of the EM field with a wavelength $\lambda = 8.9$ cm and a pulse duration $t_{pul} = 40$ μs. The initiator is a microwave dipole placed on the side of a radioparent shutter (quartz disk or polyethylene film) that faces away from the microwave radiation source. The discharge develops in the form of growing branching streamer channels with a maximum diameter of several hundredths of a centimeter. The channels can separate from the initiator and fill up the beam cross section if the field amplitude is higher than or equal to the threshold level $E_{th} = 2.5$ kV/cm. At atmospheric pressure, this threshold coincides with the known value separating the subcritical and highly subcritical forms of a volume streamer microwave discharge. This corresponds to the subcriticality ratio $E_c/E_{th} \approx 17$. The subcritical microwave discharge can also be initiated by several dipoles placed on the dielectric surface. Also, it is initiated in the gap between two dielectric surfaces with a gap size Δ no greater than several hundredths of a centimeter. If the experimental subcriticality $E_c/E_0 > 6.5$, the discharge propagates neither to the side of the polyethylene film facing away from the initiator nor counter to the radiation for a film thickness d exceeding several thousandths, of a centimeter.

REFERENCES

1. L. P. Grachev, I. I. Esakov, and K. V. Khodataev, Zh. Tekh. Fiz. **69** (11), 19 (1999) [Tech. Phys. **44**, 1276 (1999)].
2. L. P. Grachev, I. I. Esakov, G. I. Mishin, and K. V. Khodataev, Zh. Tekh. Fiz. **65** (7), 60 (1995) [Tech. Phys. **40**, 666 (1995)].
3. L. P. Grachev, I. I. Esakov, G. I. Mishin, and K. V. Khodataev, Zh. Tekh. Fiz. **65** (5), 21 (1995) [Tech. Phys. **40**, 416 (1995)].
4. *Handbook of Physical Quantities*, Ed. by I. S. Grigoriev and E. Z. Meĭlikhov (Énergoatomizdat, Moscow, 1991; CRC Press, Boca Raton, 1997).

Translated by M. Astrov

Technical Physics, Vol. 47, No. 7, 2002, pp. 856–861. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 63–67.

GAS DISCHARGES, PLASMA

Concerning the Electromagnetic Instability of a Homogeneous Anisotropic Plasma

V. A. Antonov, A. S. Baranov, and Yu. N. Gnedin

Pulkovo Observatory, Russian Academy of Sciences, Pulkovskoe shosse 65, St. Petersburg, 196140 Russia

Received October 31, 2001

Abstract—A study is made of the electromagnetic instability of a homogeneous plasma with a highly anisotropic one- or two-dimensional velocity distribution in the absence of a magnetic field. It is shown that, if the velocity distribution has a center of symmetry, then the plasma is always unstable against sufficiently long electromagnetic waves. This instability plays an important role in both laboratory plasma devices and cosmic rays.

INTRODUCTION

As is well known, electromagnetic perturbations in plasmas are those that generate plasma magnetic fields strong enough to bend the particle trajectories. It follows from simple energy considerations that a plasma with a spherically symmetric velocity distribution is stable against electromagnetic perturbations. On the contrary, an anisotropic plasma is subject to peculiar electromagnetic instabilities, which have been studied both theoretically (see, e.g., [1], Section 15) and numerically (see, e.g., [2]).

However, in our opinion, Mikhailovskii [1] investigated the simplest example of two counterpropagating plasma flows in the absence of an external magnetic field somewhat incompletely, because he assumed that the electric field perturbations are directed exactly along the flows. For this reason, it remains unclear whether the relevant problem has been solved in full measure and whether it can be generalized. In the present paper, we analyze this issue in more detail and suggest certain generalizations.

BASIC EQUATIONS

We consider a model in which a homogeneous plasma consists of plasma flows designated as $i = 1, 2, \ldots, N$ and assume that the velocity vectors $\mathbf{v}_i(u_i, 0, w_i)$ of the flows lie in the same plane. We denote the total density of the electrons by n (as usual, we take into account only their motion) and the partial densities of the flows by $\gamma_i n$. Then, we have

$$\gamma_i > 0, \quad \gamma_1 + \gamma_2 + \ldots + \gamma_N = 1. \tag{1}$$

We consider an electromagnetic wave propagating in the z direction or, equivalently, in the plane of the plasma flows. We treat an electromagnetic perturbation in the linear approximation, in which case all local perturbation parameters contain the factor $\exp(\lambda t + ikz)$, where t is time.

In the absence of perturbation, each particle moves by inertia; therefore, its initial position vector $\mathbf{r}_{0i}$ changes according to the equation

$$\mathbf{r} = \mathbf{r}_{0i} + \mathbf{v}_i t. \tag{2}$$

In fact, we must take into account small perturbations of both the electric (**E**) and magnetic (**H**) fields. In order to obtain an equation describing the evolution of the velocity vector of an individual particle, it is necessary, first, to write the familiar equation for the particle momentum,

$$\frac{d\mathbf{p}}{dt} = e\left(\mathbf{E} + \frac{\mathbf{r}\times\mathbf{H}}{c}\right) \tag{3}$$

(where e is the charge of an electron and c is the speed of light), and, then, to take into account the algebraic relationships between the momentum and the velocity,

$$\mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1-\frac{v^2}{c^2}}}, \quad \mathbf{v} = \frac{\mathbf{p}}{\sqrt{m^2+\frac{p^2}{c^2}}},$$
$$\sqrt{m^2+\frac{p^2}{c^2}} = \frac{m}{\sqrt{1-\frac{v^2}{c^2}}}. \tag{4}$$

Following what was said above, we can single out the time and coordinate dependences in the expressions for **E** and **H**:

$$\mathbf{E} = \boldsymbol{\varepsilon} e^{\lambda t + ikz}, \quad \mathbf{H} = \mathbf{h} e^{\lambda t + ikz}. \tag{5}$$

1063-7842/02/4707-0856$22.00

Linearizing Eq. (3) yields

$$\frac{d\delta\mathbf{p}}{dt} = e\left(\boldsymbol{\varepsilon} + \frac{\mathbf{v}\times\mathbf{h}}{c}\right)e^{\lambda t + ikz}. \quad (6)$$

Here and below, the symbol δ denotes the perturbation of a given quantity.

However, note that the z coordinate on the right-hand side of Eq. (6) refers to the instantaneous position of a particle and thus should be determined from Eq. (2), in which case the perturbation gives rise to only a small (second-order) correction to Eq. (2). With these remarks in mind, we integrate Eq. (6) from the infinite past to find

$$\delta\mathbf{p} = e\left(\boldsymbol{\varepsilon} + \frac{\mathbf{v}\times\mathbf{h}}{c}\right)\int_{\infty}^{t} e^{ik(z_0 + wt) + \lambda t}dt = e\left(\boldsymbol{\varepsilon} + \frac{\mathbf{v}\times\mathbf{h}}{c}\right)\frac{e^{\lambda t + ikz}}{\lambda + ikw}. \quad (7)$$

Here, the sought-for quantity is expressed in terms of the z coordinate of the particle in the absence of perturbation. Since we are interested exclusively in unstable perturbations, we set $\lambda > 0$. According to relationships (4), we obtain from Eq. (7) the following equation for the velocity perturbation:

$$\delta\mathbf{v} = \frac{e\sqrt{1 - \frac{v^2}{c^2}}}{m(\lambda + ikw)}e^{\lambda t + ikz}\left[\boldsymbol{\varepsilon} + \frac{\mathbf{v}\times\mathbf{h}}{c} - \frac{(\mathbf{v}\boldsymbol{\varepsilon})\mathbf{v}}{c^2}\right]. \quad (8)$$

We insert the instantaneous z coordinate from Eq. (2) into Eq. (8) and integrate the resulting equation over t to obtain

$$\delta\mathbf{k} = \frac{e\sqrt{1 - \frac{v^2}{c^2}}}{m(\lambda + ikw)}e^{\lambda t + ikz}\left[\boldsymbol{\varepsilon} + \frac{\mathbf{v}\times\mathbf{h}}{c} - \frac{(\mathbf{v}\boldsymbol{\varepsilon})\mathbf{v}}{c^2}\right]. \quad (9)$$

Equations (8) and (9) refer to any particle in any plasma flow. In what follows, in order to distinguish between the flows, we will mark them with the subscript i. The current at each point should be found from the relationship

$$\mathbf{J} = \mathbf{j}e^{\lambda t + ikz}, \quad (10)$$

in which case the corresponding inverse relationship has the form

$$\mathbf{j} = \frac{ke^{-\lambda t}}{2\pi}\int_0^{2\pi/2}\mathbf{J}e^{-ikz}dz. \quad (11)$$

The current $\mathbf{J}$ is the sum of all the contributions $em\mathbf{v}$ from each particle in each plasma flow; of course, the contributions should also be summed over i. In the expansion of $em\mathbf{v}\exp(-ikz)$, the zero-order terms entirely cancel each other, reflecting the fact that, in an unperturbed state, there are no currents. Accordingly, we have

$$\delta\mathbf{j} = \mathbf{j} = \frac{keme^{-\lambda t}}{2\pi}\sum_{i=1}^{N}\int_0^{2\pi/2}(\delta\mathbf{v} - ik\mathbf{v}\delta z)e^{-ikz}dz$$

or

$$\delta\mathbf{j} = \frac{e^2}{m}\sum_{i=1}^{N} n_i\left\{\left[\boldsymbol{\varepsilon} + \frac{\mathbf{v}_i\times\mathbf{h}}{c} - \frac{(\mathbf{v}_i\cdot\boldsymbol{\varepsilon})\mathbf{v}_i}{c^2}\right]L - ik\mathbf{v}_i\left[\varepsilon_z + \frac{(\mathbf{v}_i\times\mathbf{h})_z}{c} - \frac{\mathbf{w}_i(\mathbf{v}_i\cdot\boldsymbol{\varepsilon})}{c^2}\right]L^2\right\}\sqrt{1 - \frac{v_i^2}{c^2}} \quad (\mathrm{Re}\lambda > 0). \quad (12)$$

This way of deriving the perturbed current is advantageous in that it clearly illustrates how the factor $L = 1/(\lambda + ikw)$ arises in expression (12). The sought-for current can also be obtained by linearizing the hydrodynamic equations for each of the flows; in this way, we again arrive at the above expression (12) for δj. Note that, if k and λ are assumed to be of the same order, then the generated current is inversely proportional to k.

Now, we must combine expression (12) into Maxwell's equations. However, in the case at hand, we have $H_z = 0$; as a result, Maxwell's equations become

$$\frac{\lambda}{c}h_x = ik\varepsilon_y, \quad \frac{\lambda}{c}h_y = -ik\varepsilon_x, \quad ikh_y + \frac{\lambda}{c}\varepsilon_x + \frac{4\pi}{c}\delta j_x = 0, \quad -ikh_x + \frac{\lambda}{c}\varepsilon_y + \frac{4\pi}{c}\delta j_y = 0, \quad (13)$$

$$\frac{\lambda}{c}\varepsilon_z + \frac{4\pi}{c}\delta j_z = 0. \quad (14)$$

Solving the first of Eqs. (13) and the second of Eqs. (14) separately from the remaining equations and using expression (12), we obtain

$$\delta j_y = \frac{e^2 n}{m\lambda}\varepsilon_y, \quad \left(\frac{4\pi e^2 n}{cm\lambda} + \frac{\lambda}{c} + \frac{k^2 c}{\lambda}\right)\varepsilon_y = 0.$$

We can see that, regardless of the structure of the velocity distribution, there are two possible cases. The first case is that with

$$\lambda = \pm i\sqrt{k^2c^2 + \frac{4\pi e^2 n}{m}}.$$

This indicates that the system of particles undergoes oscillations with the frequency $\omega = \sqrt{k^2c^2 + w_p^2}$, which is equal to the frequency of electromagnetic oscillations in a cold plasma (here, ω_p is the electron plasma frequency). The second case is that with $\varepsilon_y = h_x = 0$. This latter case will be the subject of the further analysis.

DERIVATION AND ANALYSIS OF THE DISPERSION RELATION

Now, we are left with the first and third of Eqs. (14). Substituting h_y from the second of Eqs. (13) and the current from expression (12) into these two equations, we arrive at the set of linear homogeneous equations for ε_x and ε_z, which can be simplified to the form

$$P\varepsilon_x + Q\varepsilon_z = 0, \quad R\varepsilon_x + S\varepsilon_z = 0, \tag{15}$$

where

$$P = \frac{k^2 c}{\lambda} + \frac{\lambda}{c} + \frac{4\pi e^2}{mc}\sum_{i=1}^{N} n_i \left[\left(1 + \frac{ikw_i}{\lambda} - \frac{u_i^2}{c^2}\right)L + iku_i^2\left(\frac{ik}{\lambda} + \frac{w_i}{c^2}\right)L^2\right]\sqrt{1 - \frac{v_i^2}{c^2}},$$

$$Q = \frac{4\pi e^2}{mc}\sum_{i=1}^{N} n_i u_i \left[\frac{w_i}{c^2}L + ik\left(1 - \frac{w_i^2}{c^2}\right)L^2\right]\sqrt{1 - \frac{v_i^2}{c^2}},$$

$$R = \frac{4\pi e^2}{mc}\sum_{i=1}^{N} n_i u_i \left[\left(-\frac{ik}{\lambda} - \frac{w_i}{c^2}\right)L + ikw_i\left(\frac{ik}{\lambda} + \frac{w_i}{c^2}\right)L^2\right]\sqrt{1 - \frac{v_i^2}{c^2}},$$

$$S = \frac{\lambda}{c} + \frac{4\pi e^2}{mc}\sum_{i=1}^{N} n_i \left(1 - \frac{w_i^2}{c^2}\right)(1 - ikw_i L) \times L\sqrt{1 - \frac{v_i^2}{c^2}} \quad (v_i^2 = u_i^2 + w_i^2).$$

The sought-for equation for λ can be obtained by equating the determinant of Eqs. (15) to zero. Of greatest interest to the present discussion is the asymptotic limit $k \longrightarrow 0$. If, in this limit, λ remains proportional to k, then the first two terms in the expression for P are small in comparison with the third term and the first term in the expression for S is small in comparison with the second term. As a result, after some simplifications, the dispersion relation takes the form

$$\sum_{i=1}^{N} n_i \sqrt{1 - \frac{v_i^2}{c^2}}\left[1 - u_i^2\left(k^2 + \frac{\lambda^2}{c^2}\right)L^2\right] \times \sum_{i=1}^{N} n_i \sqrt{1 - \frac{v_i^2}{c^2}}\left(1 - \frac{w_i^2}{c^2}\right)L^2 - \left[\sum_{i=1}^{N} n_i \sqrt{1 - \frac{v_i^2}{c^2}}u_i\left(\frac{\lambda w_i}{c^2} + ik\right)L^2\right]^2 = 0. \tag{16}$$

For the particular case of two counterpropagating plasma flows of equal density, we can set $N = 2$, $n_{1,2} = n/2$, $u_{1,2} = \pm u$, and $w_{1,2} = \pm w$. In this case, the dispersion relation (16), being simplified to the ultimate possible extent, has only one pair of roots:

$$\lambda = \pm k\sqrt{\frac{(u^2 + w^2)\left(1 - \frac{w^2}{c^2}\right)}{1 - \frac{u^2 + w^2}{c^2}}}.$$

This formula clearly shows that the instability does exist and that the perturbation grows exponentially with time. The particular case considered by Mikhailovskii [1] is that with $w = 0$ (in our notation) and without the relativistic correction u^2/c^2. Below, we will verify directly that the relevant dispersion relation presented in [1], in a sense, follows from the formulas obtained in the present work.

Now, we turn to the general case of a two-dimensional distribution of the velocities of the electron beams. We again assume that k and λ are both small and are of the same order of magnitude. In this situation, in contrast to the case of oblique perturbations in a plasma with a one-dimensional velocity distribution, we cannot speak of the separation of the unknowns ε_x and ε_z. Nevertheless, under the additional assumption that the two-dimensional velocity distribution has a center of symmetry, it is also an easy matter to show the exponential growth of the instability. In other words, we impose the requirement that each beam with the velocity vector $v(u, 0, w)$ be counterbalanced by another beam with the velocity vector $v(-u, 0, -w)$ and with the same density. Then, for a real λ, the complex-conjugate terms in each of the sums on the left-hand side of the dispersion relation (16) are added together; thereby, the left-hand side itself, which will be denoted by $F(\lambda)$, is real. In particular, we have

$$F(0) = \frac{1}{k^2}\left[-\sum_{i=1}^{N} n_i \sqrt{1 - \frac{v_i^2}{c^2}}\left(1 + \frac{u_i^2}{w_i^2}\right) \times \sum_{i=1}^{N} n_i \sqrt{1 - \frac{v_i^2}{c^2}}\left(1 - \frac{w_i^2}{c^2}\right)\frac{1}{w_i^2} + \left(\sum_{i=1}^{N} n_i \sqrt{1 - \frac{v_i^2}{c^2}}\frac{u}{w_i^2}\right)^2\right].$$

The physically obvious requirement $u^2 + w^2 < c^2$ leads to

$$\frac{1}{w^2}\left(1 + \frac{u^2}{w^2}\right)\left(1 - \frac{w^2}{c^2}\right) - \left(\frac{u}{w^2}\right)^2 = \frac{1}{w^2}\left(1 - \frac{u^2 + w^2}{c^2}\right) > 0,$$

which gives

$$\left|\frac{u_i}{w_i^2}\right| < \sqrt{\left(1+\frac{u_i^2}{w_i^2}\right)\frac{1}{w_i^2}\left(1-\frac{w_i^2}{c_i^2}\right)}\ (i = 1, 2, \ldots, N).$$

Employing the Bunyakowsky inequality yields $F(0) < 0$. As $\lambda \longrightarrow \infty$, the quantity $F(\lambda)$ asymptotically approaches zero according to the law

$$F(\lambda) \approx \frac{1}{\lambda^2}\left[\sum_{i=1}^{N} n_i\sqrt{1-\frac{v_i^2}{c^2}}\left(1+\frac{u_i^2}{w_i^2}\right)\right.$$
$$\left.\times\sum_{i=1}^{N} n_i\sqrt{1-\frac{v_i^2}{c^2}}\left(1-\frac{w_i^2}{c^2}\right)-\left(\sum_{i=1}^{N} n_i\sqrt{1-\frac{v_i^2}{c^2}}\frac{u_i w_i}{c^2}\right)^2\right].$$

Since

$$\left(1-\frac{v^2}{c^2}\right)\left(1-\frac{w^2}{c^2}\right)-\left(\frac{uw}{c^2}\right)^2 = 1-\frac{u^2+w^2}{c^2} > 0,$$

the inequality $F(\lambda) > 0$ for large λ can be proved in the same way as was done above. For all intermediate positive values of λ, the function $F(\lambda)$ is regular; consequently, the equation $F(\lambda) = 0$ has at least one positive root, which was to be proved.

THE CASE OF A CONTINUOUS VELOCITY DISTRIBUTION

The instability described in the previous section can partially originate from the discrete structure of the velocity distribution of the electron flows. For this reason, it is expedient to present at least one example of an electron flow whose velocity distribution is continuous in both n and w. In this case, all the sums in the dispersion relation (16) should be replaced by the corresponding integrals with a certain velocity distribution function $f(u, w)$. To be specific, we use a Gaussian distribution with the dispersions σ_u^2 and σ_w^2 and with the correlation coefficient ρ. We restrict ourselves to nonrelativistic motions and, accordingly, omit corrections proportional to w^2/c^2 and v^2/c^2. The origin of the coordinate system in velocity space may be chosen to be located at the center $u = w = 0$ of the velocity distribution, in which case the function $F(\lambda)$ is real. We discard the corresponding relativistic corrections and switch from the sums to the integrals in dispersion relation (16):

$$\iint(1-k^2u^2L^2)f(u, w)dudw\iint L^2 f(u, w)dudw + k^2\left[\iint uL^2 f(u, w)dudw\right]^2 = 0. \quad (17)$$

The integrals in Eq. (17) can be converted into a more convenient form by using the familiar representation

$$L^2 = \frac{1}{(\lambda+ikw)^2} = \int_0^\infty te^{-\lambda t-ikwt}dt\ (\lambda > 0). \quad (18)$$

In particular, with allowance for this representation, the last of the integrals in Eq. (17) becomes

$$T_1 = \iint uL^2 f(u, w)dudw = \iiint ute^{-\lambda t-ikwt}dudwdt. \quad (19)$$

Here and below, the integration is carried out over the entire (n, w) plane and over the positive values of time t. Integral (19) can be further transformed using the characteristic function φ. In the above particular case with a Gaussian distribution, this function is known to have the form [3]

$$\varphi(\xi, \eta) = \iint e^{i(\xi u+\eta w)}f(u, w)dudw = ne^{-(\sigma_u^2\xi^2+2\rho\sigma_u\sigma_w\xi\eta+\sigma_w^2\eta^2)/2}. \quad (20)$$

We differentiate both sides of expression (20) with respect to the parameter ξ and substitute $\xi = 0$ and $\eta = kt$ into the resulting expression to obtain

$$i\iint ue^{-ikwt}f(u, w)dudw = \rho\sigma_u\sigma_w ktne^{-(k^2\sigma_w^2t^2/2)}.$$

Therefore, we obtain

$$T_1 = -i\rho\sigma_u\sigma_w kn\int_0^\infty t^2\chi dt,$$

where we have introduced the notation $\chi = \exp[-\lambda t - (k^2\sigma_w^2 t^2)/2]$.

Analogously (but without differentiating with respect to the parameter), we find

$$T_2 = \iint L^2 f(u, w)dudw = n\int_0^\infty t\chi dt,$$

in which case the double differentiation gives

$$T_3 = \iint u^2L^2 f(u, w)dudw = n\sigma_u^2\int_0^\infty(1-k^2\rho^2\sigma_w^2t^2)t\chi dt.$$

Hence, Eq. (17) takes the form

$$F(\lambda) = n^2\left[1-k^2\sigma_u^2\int_0^\infty(1-k^2\sigma_w^2\rho^2t^2)t\chi dt\right]$$
$$\times\int_0^\infty t\chi dt-\rho^2\sigma_u^2\sigma_w^2n^2k^4\left(\int_0^\infty t^2\chi dt\right)^2 = 0.$$

Simple manipulations yield

$$F(0) = \frac{n^2\left\{\sigma_w^2 + \left[\left(2-\frac{\pi}{2}\right)\rho^2 - 1\right]\sigma_u^2\right\}}{k^2\sigma_w^4},$$

the corresponding asymptotic behavior at large values of λ being $F(\lambda) \approx n^2/\lambda^2$. A sufficient condition for the instability has the form

$$\sigma_u^2\left[1-\left(2-\frac{\pi}{2}\right)\rho^2\right] > \sigma_w^2.$$

If, in the case of an anisotropic distribution (i.e., of different dispersions σ_1^2 and σ_2^2), the wave vector is oriented in the direction of the principal axis with the smaller dispersion σ_2^2, then we have $\sigma_u = \sigma_1$, $\sigma_w = \sigma_2$, and $\rho = 0$; therefore, the sufficient condition for the instability is satisfied. The only exception to this is the case of an isotropic distribution, for which we have $\sigma_u = \sigma_w$ and $\rho = 0$ regardless of the orientation of the wave vector.

Recall that we are dealing with a two-dimensional velocity distribution. The instability under consideration also takes place in intermediate situations, in which the dispersion in one direction (in the case at hand, this is σ_v^2) is nonzero and is smaller than the other dispersions. The corresponding calculations will be presented in a separate paper.

DISCUSSION OF THE RESULTS

We have proved that any two-dimensional velocity distribution is electromagnetically unstable, at least when it has a center of symmetry. However, this instability refers only to sufficiently long electromagnetic waves, thereby raising a question as to the wavelengths of unstable waves. In order to answer this question, we again turn to the continuous velocity distribution, retaining now all the terms in the general dispersion relation

$$\Psi(\lambda) \equiv PS - QR = 0. \tag{21}$$

Then, we have

$$\Psi(0) = k^2 + \frac{4\pi e^2}{m}\left[-\sum_{i=1}^{N} n_i\sqrt{1-\frac{v_i^2}{c^2}}\left(1-\frac{w_i^2}{c^2}\right)\right.$$
$$\left. + \frac{1}{c^2}\sum_{i=1}^{N} n_i\sqrt{1-\frac{v_i^2}{c^2}}\left(1+\frac{u_i^2}{w_i^2}\right)\right] + \left(\frac{4\pi e^2}{mc}\right)^2 F(0),$$

so that the condition $\Psi(0) < 0$, under which the above analysis is valid, holds for roughly $k < k^*$, where

$$k^{*2} \sim \frac{ne^2}{mvc}\sqrt{1-\frac{v^2}{c^2}}. \tag{22}$$

Here, v should be understood to be the characteristic velocity of the beams. For large λ values, the first two terms in the expression for P and the first term in the expression for S produce positive terms in the expression for $\Psi(\lambda)$; however, this positive contribution plays an insignificant role in the problem under analysis. Hence, we can suggest that unstable electromagnetic waves are those with the wavelengths $k < k^*$, therefore, for the electromagnetic instability to occur, the system must have sufficiently large dimensions. As for the instability growth rate, the dispersion relation (16) yields the estimate

$$\lambda \sim kv, \tag{23}$$

which has already been presented in the literature (see, e.g., [1]).

Now, let us again turn to the above particular case of two plasma flows. We assume that the flows are both perpendicular to the wave vector (i.e., $w = 0$) but use the general dispersion relation (21). One can easily see that this dispersion relation splits into two equations, $P = 0$ and $S = 0$, because, in the case at hand, we have $Q = R = 0$. Equating S to zero yields $\lambda = \pm i\omega_p$. The equation $P = 0$ is far more complicated:

$$\lambda^2(k^2c^2 + \lambda^2)$$
$$+ \omega_p^2\left[\lambda^2\left(1-\frac{v^2}{c^2}\right)^{3/2} - k^2u^2\left(1-\frac{v^2}{c^2}\right)^{1/2}\right] = 0. \tag{24}$$

Omitting the relativistic correction and taking into account an obvious difference in the notation (in particular, in [1], the electron plasma frequency is defined for each individual flow, whereas, in our study, it refers to the entire system), we can reduce Eq. (24) to the corresponding equation presented in [1], which was derived earlier by Weibel [4]. Recall that Mikhailovskii [1] considered an electric field having the only component along the flows. From a geometric standpoint, he limited himself to perturbations that are antisymmetric with regard to mirror reflections in a plane perpendicular to the original motion of the flows. However, as was shown above, symmetric perturbations manifest themselves as Langmuir oscillations rather than as electromagnetic waves. These symmetry considerations apply also to the relativistic limit. If a one-dimensional velocity distribution is asymmetric (e.g., $N = 2$, but $n_i = n_2$), then, neglecting the relativistic correction v^2/c^2 all the same yields $Q = R = 0$. However, this is not the case in the relativistic limit, in which the components ε_x and ε_z are coupled to one another.

Let us estimate the critical wavelength at which electromagnetic waves become unstable. For the conditions prevailing in fusion experiments with magnetic

confinement systems [3], we can set the plasma temperature to be about 10^7 K and take $n \sim 10^{18}$ cm^{-3}. The critical wavelength estimated for these values from formula (19) is about several hundredths of a millimeter, indicating that it is necessary to keep in mind the possible onset of electromagnetic instability in laboratory devices.

The question as to the stability of a system of galactic cosmic rays is somewhat more complicated. Based on the above calculations, we can conclude with a high degree of confidence that the anisotropy of the velocity distribution of a system of cosmic rays cannot be large, at least for stability reasons. In order to provide a stronger justification for this conclusion, it is necessary to carry out calculations for a significantly non-Gaussian velocity distribution. We hope to do this in future work.

Another interesting example is a system of ion beams with an energy $\varepsilon = 1$ MeV per nucleon emitted by the Sun during solar flares and interacting with the Sun's atmosphere. Setting the density of an ion beam to be 10% of the typical thermal plasma density in a coronal loop, we find that the characteristic critical wavelength is about 1 m. However, the corona plasma, in which the ion beams propagate, may have a stabilizing effect on them. This stabilization is sometimes provided by the efficient absorption of excited long electromagnetic waves in the thermal plasma as well as by the transfer of the wave energy in k space.

REFERENCES

1. A. B. Mikhailovskii, *Theory of Plasma Instabilities*, Vol. 1: *Instabilities of a Homogeneous Plasma* (Atomizdat, Moscow, 1971; Consultants Bureau, New York, 1974).
2. Ch. K. Birdsall and A. B. Langdon, *Plasma Physics Via Computer Simulations* (McGraw-Hill, New-York, 1985; Énergoatomizdat, Moscow, 1989).
3. L. A. Artsimovich and R. Z. Sagdeev, *Plasma Physics for Physicists* (Nauka, Moscow, 1973).
4. E. S. Weibel, Phys. Rev. Lett. **2**, 83 (1959).

Translated by O. Khadin

Technical Physics, Vol. 47, No. 7, 2002, pp. 862–865. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 68–71.

SOLIDS

Anisotropy of Magnetoelastic Wave Attenuation in (111) Crystal Plates with Combined Anisotropy

R. M. Vakhitov, V. V. Grinevich†, and O. G. Ryakhova

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

e-mail: VakhitovRM@bsu.bashedu.ru

Received August 20, 2001

Abstract—The effect of dissipation on the propagation of magnetoacoustic waves in a cubic ferromagnet under the action of applied stress $\boldsymbol{\sigma} \parallel [111]$ is studied theoretically. When the ferromagnet is in one of two symmetric phases ($\mathbf{M}_0 \parallel [111]$ or $\mathbf{M}_0 \parallel [\bar{1}10]$), the magnetoelastic waves weakly decay and may transform into modes of another type as the material approaches the point of spin-reorientation transition. It is also found that the propagation velocity and attenuation rate of quasi-phonons exhibit anisotropy, which can be controlled by the applied stress. © *2002 MAIK "Nauka/Interperiodica".*

It is known that magnetoelastic interaction is rather weak in most magnetic materials and exerts a minor effect on their magnetic properties. However, near the spin-reorientation phase transition (SRPT), where the magnetic anisotropy is negligible, the magnetoelastic interaction dominates. This enhances magnetoacoustic effects in magnets and, in particular, softens the magnetoacoustic mode [1]. In the presence of elastic and magnetoelastic anisotropy in crystals, this softening depends on the crystallographic direction.

The allowance for dissipation may also affect significantly the softening of the magnetoacoustic mode and the propagation of magnetoelastic waves on the whole. This is because, spin waves in a dissipative magnetic subsystem decay and their propagation depends to a great extent on how close the subsystem is to the SRPT point [2].

In this paper, we study the effect of relaxation processes on the propagation of magnetoelastic waves in a cubic ferromagnet with induced uniaxial anisotropy (IUA), which is typical of many magnets used in practice, including garnet ferrites [3]. We assume that uniaxial anisotropy is induced by an applied stress $\boldsymbol{\sigma} \parallel [111]$ (as in a (111) garnet ferrite plate).

The energy density in this magnet can be represented as

$$\begin{aligned} E &= A\left(\frac{\partial m}{\partial x_i}\right)^2 + K_1(m_x^2m_y^2 + m_x^2m_z^2 + m_y^2m_z^2) \\ &+ K_2m_x^2m_y^2m_z^2 + \frac{1}{8\pi}H_{\text{dem}}^2 + B_1(m_x^2U_{xx} + m_y^2U_{yy} + m_z^2U_{zz}) \\ &+ 2B_1(m_xm_yU_{xy} + m_xm_zU_{xz} + m_ym_zU_{yz}) \\ &+ \frac{1}{2}C_{11}(U_{xx}^2 + U_{yy}^2 + U_{zz}^2) \\ &+ C_{12}(U_{xx}U_{yy} + U_{xx}U_{zz} + U_{yy}U_{zz}) \\ &+ 2C_{44}(U_{xy}^2 + U_{yz}^2 + U_{xz}^2) + \sigma_{ij}U_{ij}, \end{aligned} \tag{1}$$

where A, B, and C are the exchange, magnetoelastic, and elastic interaction factors; K_1 and K_2 are the first and second cubic anisotropy constants; $\mathbf{m} = \mathbf{M}/M_s$ is the unit vector of magnetization; M_s is the saturation magnetization; U_{ij} and σ_{ij} are the strain and stress tensors ($\boldsymbol{\sigma} \parallel [111]$); and H_{dem} is the demagnetizing field determined from the magnetostatics equation

$$\nabla \cdot (\mathbf{H}_{\text{dem}} + 4\pi\mathbf{M}_s) = 0, \quad \nabla \times \mathbf{H}_{\text{dem}} = 0. \tag{2}$$

To find the magnetoelastic oscillation spectrum, it is necessary to simultaneously solve the Landau–Lifshitz equations with the Gilbert relaxation term and the equations of motion for the elastic displacement vectors $\mathbf{U}$:

$$\begin{aligned} \frac{d\mathbf{M}}{dt} &= -\gamma\mathbf{M} \times \mathbf{H}^{\text{eff}} - \frac{r}{M_s}\mathbf{M} \times \frac{\partial \mathbf{M}}{\partial t}, \\ \rho U_i &= \frac{\partial}{\partial x_j}\left[\frac{\partial E}{\partial U_{ij}}\frac{1+\delta_{ij}}{2}\right], \end{aligned} \tag{3}$$

where γ is the gyromagnetic ratio, ρ is the density of the crystal, $\mathbf{H}^{\text{eff}}$ is some effective magnetic field, and r is the dimensionless attenuation parameter.

Equations (3) constitute a complete system of magnetoelastic equations. It can be solved in the linear approximation with respect to the dynamic variables $\mathbf{M}^{(1)} = \mathbf{M} - \mathbf{M}^{(0)}$ and $\mathbf{U}^{(1)} = \mathbf{U} - \mathbf{U}^{(0)}$ [1]. The spontaneous strains $U_{ij}^{(0)}$ and the equilibrium directions of the

† Deceased.

magnetization vector $\mathbf{M}^{(0)}$($\mathbf{M}^{(0)} = \mathbf{M}_0$) are found from the minimum condition for the energy density E in the homogeneous state. Phenomena associated with surface effects are neglected.

Calculations for the ground state show that six magnetic phases may exist in the (111) plate. Two of them are symmetric phases with $\mathbf{M}_0 \parallel [111]$ (phase $F_{[111]}$) and $\mathbf{M}_0 \parallel [\bar{1}10]$ (phase $F_{[10\bar{1}]}$), three are angular phases, and one is a phase of type $[uvw]$ [4, 5]. Hereafter, we assume that the magnet is in one of the symmetric phases.

Consider the case $\mathbf{M} \parallel [\bar{1}10]$. This phase is stable if

$$K_u < 0,$$
$$\frac{\sqrt{K_2^2 - 16K_0K_2/3} - K_2}{4} > K_1 > -\frac{\sqrt{K_2^2 - 16K_uK_2/3} + K_2}{4}, \quad (4)$$

where $K_u = 3\sigma\lambda_{111}/2$ is the IUA constant and λ_{111} is the magnetostriction constant in the [111] direction.

The dispersion relation for the coupled oscillations is written as

$$(r^2 + 1)\omega^4 + ir\omega^3((\omega_{2k} + 4\pi M_s\gamma\sin\alpha) + (\omega_{3k} + 4\pi M_s\gamma\cos\alpha)) - \omega^2(\omega_{skII}^2 + \omega_t^2(1 + r^2)) - ir\omega\omega_t^2((\omega_{2k} + 4\pi M_s\gamma\sin\alpha) + (\omega_{3k} + 4\pi M_s\gamma\cos\alpha)) - \omega_t^2\omega_{skII}^2\xi_7^2 + \omega_t^2\omega_{skII}^2(1 - \xi_6^2) = 0, \quad (5)$$

where $\xi_7 = \gamma(b_I^2 + b_{II}^2)/\omega_{skII}M_s[(C_{11} - C_{12})(\sqrt{2}\sin\alpha + \cos\alpha)^2 + 2C_{44}(\sin\alpha - \sqrt{2}\cos\alpha)^2]$ is the magnetoelastic coupling coefficient and $\omega_{skII} = \sqrt{(\omega_{2k} + 4\pi M_s\gamma\sin\alpha)(\omega_{3k} + 4\pi M_s\gamma\cos\alpha) - \omega_{1s}^2}$. The remaining parameters and the velocities for the uncoupled components are given in [6].

When magnetoelastic coupling is absent, a solution to Eq. (5) has the form

$$\omega_{1,2} = \pm\left[\omega_{skII}^2 - \frac{1}{4}r^2((\omega_{3sk} + 4\pi M_s\gamma\cos\alpha) - (\omega_{2sk} + 4\pi M_s\gamma\sin\alpha))^2 + \omega_{1s}^2)\right]^{1/2} - \frac{1}{2}ir[(\omega_{3sk} + 4\pi M_s\gamma\cos\alpha) + (\omega_{2sk} + 4\pi M_s\gamma\sin\alpha)], \quad (6)$$

where $\omega_{3,2sk} = \omega_{3,2k} - \omega_{3,2ml}$ and $\omega_{3,2ml}$ is the minimum width of the magnon gap.

As follows from formula (6), away from the SRPT points ($\omega_{2s0} \longrightarrow 0$ or $\omega_{3s0} \longrightarrow 0$), where $\omega_{skII}^2 \gg r^2[(\omega_{3sk} + 4\pi M_s\gamma\cos\alpha) - (\omega_{2sk} + 4\pi M_s\gamma\sin\alpha)]^2 + \omega_{1s}^2$, the spin waves weakly decay. Near the SRPT points (for example, at the point $\omega_{2s} \longrightarrow 0$, where $\omega_{2sk} \ll \omega_{3sk}$ and $\omega_{1s} \longrightarrow 0$, or when $\mathbf{k} \longrightarrow 0$), the situation may change radically. When $\omega_{skII}^2 \ll r^2[(\omega_{3sk} + 4\pi M_s\gamma\cos\alpha) - (\omega_{2sk} + 4\pi M_s\gamma\sin\alpha)]^2 + \omega_{1s}^2$, solutions (6) describe purely relaxation oscillations:

$$\omega_1 = -\frac{i(\omega_{3sk} + 4\pi M_s\gamma\cos\alpha)(\omega_{2sk} + 4\pi M_s\gamma\sin\alpha)}{r[(\omega_{3sk} + 4\pi M_s\gamma\cos\alpha) - (\omega_{2sk} + 4\pi M_s\gamma\sin\alpha)]}, \quad (7)$$

$$\omega_2 = -ir(\omega_{3sk} + 4\pi M_s\gamma\cos\alpha). \quad (8)$$

These expressions define the inverse relaxation times for the transverse magnetization components. The relaxation mode ω_1 is soft, because its frequency tends to zero at the stability limit as $\mathbf{k} \longrightarrow 0$. Near the SRPT point ($\omega_{3s0} \longrightarrow 0$), the solution is given by formulas (7) and (8) with the subscripts 3 and 2 interchanged.

If magnetoelastic interaction is taken into consideration, the spectrum of coupled oscillations near the SRPT point ($\omega_{2s0} \longrightarrow 0$) at $k = 0$ becomes

$$\omega_{1,2} = \pm\left[\left(\omega_{s0II}^2 - \frac{1}{4}r^2((\omega_{30} + 4\pi M_s\gamma\cos\alpha) - (\omega_{20} + 4\pi M_s\gamma\sin\alpha))^2 + \omega_{1s}^2\right)\right]^{1/2} - \frac{1}{2}ir[(\omega_{30} + 4\pi M_s\gamma\cos\alpha) + (\omega_{20} + 4\pi M_s\gamma\sin\alpha)], \quad (9)$$

$$\omega_{3,4} = 0.$$

Therefore, with magnetoelastic interaction taken into account, the solution $\omega_{1,2}$ describes a decaying magnetization precession both away from and near the SRPT points. The solutions for the other two modes (at $k \neq 0$) are

$$\omega_{1,2} = \pm\left[\left(\omega_{skII}^2 - \frac{1}{4}r^2((\omega_{3k} + 4\pi M_s\gamma\cos\alpha) - (\omega_{2k} + 4\pi M_s\gamma\sin\alpha))^2 + \omega_{1s}^2\right)\right]^{1/2} - \frac{1}{2}ir[\omega_{3k}\omega_{2k} + 4\pi M_s\gamma(\sin\alpha + \cos\alpha)], \quad (10)$$

$$\omega_{3,4} = \pm\omega_t[1 - \xi_6^2]^{1/2} - \frac{ir}{2}\frac{\omega_t^2}{\omega_{skII}^2}[(\omega_{3k} + 4\pi M_s\gamma\cos\alpha) + (\omega_{2k} + 4\pi M_s\gamma\sin\alpha)\xi_6^2 - \omega_{skII}\xi_7]. \quad (11)$$

As follows from formulas (10) and (11), near the SRPT points, the coupled oscillation spectrum consists of weakly decaying quasi-spin ($\omega_{1,2}$) and quasi-elastic ($\omega_{3,4}$) branches. When relaxation is taken into consid-

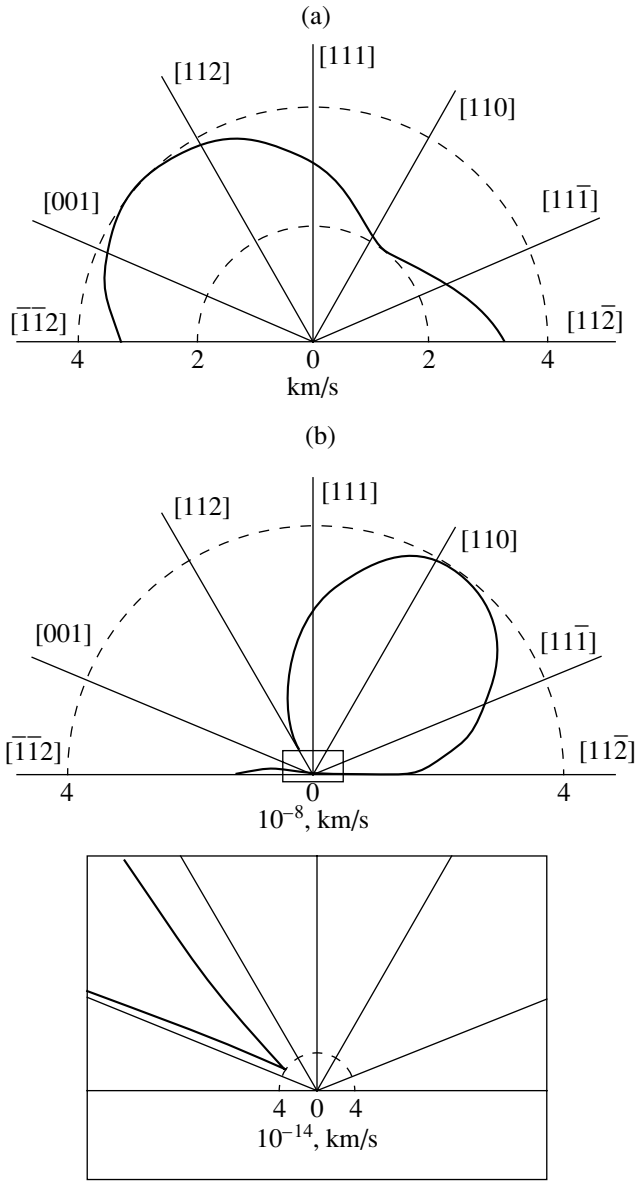


Fig. 1. Surface of the quasi-elastic wave phase velocities in $Sm_3Fe_5O_{12}$ (T = 77 K) cut by the plane ($\bar{1}10$) [8, 9]: (a) real and (b) imaginary parts of the spectrum.

eration, waves of both types decay. One can show that the propagation of such waves is pronouncedly anisotropic (Fig. 1). In the ($\bar{1}10$) plane, the magnetoacoustic waves propagate with extremal velocities in certain directions that are not crystallographic. These results contradict those of [6], where it was erroneously argued that these directions coincide with crystallographic directions. The difference is caused by the fact that the applied stress $\boldsymbol{\sigma} \parallel [111]$ breaks the cubic symmetry. This effect can be avoided by applying the stress in the directions along which the magnetoelastic waves propagate with extremal velocities.

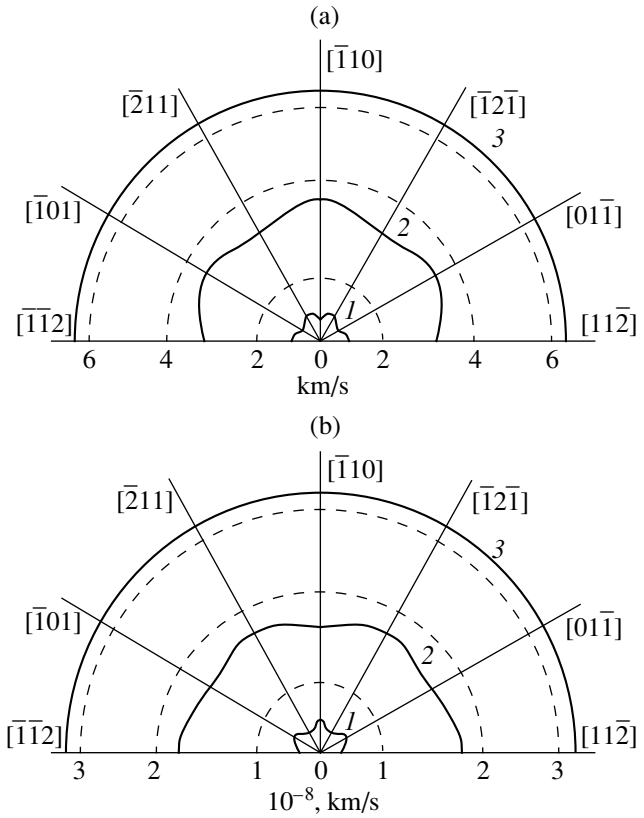


Fig. 2. Surface of the magnetoacoustic wave phase velocities in $Tb_3Fe_5O_{12}$ (T = 4.2 K) cut by the plane (111) [8, 9] for the symmetric phase $F_{[111]}$ ($\mathbf{k} \perp [111]$): (a) real and (b) imaginary parts of the spectrum.

For the symmetric phase $F_{[111]}$, whose stability is defined by the conditions

$$K_u > 0, \quad K_1 + K_2/3 - 3/2K_u > 0, \tag{12}$$

the situation is essentially different. In the absence of magnetoelastic interaction, magnetization oscillations away from the SRPT have the form of weakly decaying spin waves; however, near the SRPT point, at the stability limit of the phase $F_{[111]}$ ($\mathbf{M}_0 \parallel [111]$ and $\mathbf{k} \parallel [111]$), these modes become soft. These modes are responsible for the SRPT. With magnetoelastic interaction taken into consideration and $k \neq 0$, the spectrum of the magnet near the SRPT point consists of weakly decaying quasi-spin and transverse quasi-elastic branches.

Our calculations have shown that the propagation of such waves in the (111) plane ($\mathbf{M}_0 \parallel [111]$ and $\mathbf{k} \perp [111]$) is also anisotropic (Fig. 2). The anisotropy increases as the system approaches the boundary of the region where the phase exists, because the contributions of the elastic and magnetoelastic interactions depend on the crystallographic direction. In preferential directions, such as [11$\bar{2}$] and [01$\bar{1}$], about which the plots are symmetric, the behavior of the modes changes. In the [11$\bar{2}$] direction, mode *2* (Fig. 2a) is due

to only weakly decaying transverse elastic oscillations and has the minimum velocity of quasi-phonons. In this and other directions, the variation of the attenuation rate of this mode (Fig. 2b) is opposite in sign to that of the propagation velocity. Mode *1* and mode *3* are described, respectively, by quasi-transverse and quasi-longitudinal weakly decaying elastic oscillations in plane $(\bar{1}10)$. In $[01\bar{1}]$, these modes are longitudinal (mode *3*) and transverse (modes *1* and *2*) weakly decaying elastic oscillations.

Our results suggest that the anisotropy of the velocity is specified by the symmetry of the plane in which the magnetoelastic wave propagates. The symmetry of the plots in Fig. 2 is associated with the fact that the [111] direction is the hexagonal axis for the plane considered. At the same time, the anisotropy of the quasi-phonon velocity in Fig. 1 is of a different symmetry, because the $[\bar{1}10]$ direction is the binary axis.

It should be noted that the above results agree with experimental data obtained in [7], where the anisotropy of the propagation velocity and attenuation rate of magnetoacoustic waves in Mn–Zn spinel was found.

Thus, when the attenuation parameter is small, the spin waves decay weakly away from the phase transition points. Near these points, they may remain weakly decaying or turn into purely relaxation waves. At the stability limit of the phase $F_{[111]}$ in the absence of magnetoelastic coupling ($\mathbf{k} \parallel \mathbf{M}_0$), the weakly decaying spin mode softens completely, whereas for the $F_{[10\bar{1}]}$ phase ($\mathbf{M}_0 \parallel [\bar{1}10]$, $\mathbf{k} \perp \mathbf{M}_0$), the relaxation mode becomes soft. In both cases, these modes are responsible for the transition. However, the motion of the lattice does not become purely relaxation and remains weakly decaying near the SRPT point (when magnetoelastic coupling is taken into account). These results differ greatly from those obtained theoretically in [2], where a biaxial ferromagnet with isotropic elastic and magnetoelastic properties was studied. It was found that, near the SRPT points, the motions of the magnetization and lattice may be purely relaxation oscillations. This difference is due to the fuller consideration of the symmetry of elastic and magnetoelastic interactions in this work.

In conclusion, the propagation velocity of quasi-elastic oscillations in a (111) ferromagnet plate exhibits anisotropy, which depends on the symmetry of the plane in which the magnetoelastic waves propagate. The difference in the propagation velocities of quasi-phonons in various crystallographic directions stems from the anisotropy of magnetoelastic and elastic interactions in the crystal. This effect is enhanced when the system approaches the phase stability limit through variation of the temperature-dependent anisotropic constant or the magnitude and direction of the applied elastic stress.

REFERENCES

1. E. A. Turov and V. G. Shavrov, Usp. Fiz. Nauk **140**, 429 (1983) [Sov. Phys. Usp. **26**, 593 (1983)].
2. V. D. Buchel'nikov and V. G. Shavrov, Pis'ma Zh. Éksp. Teor. Fiz. **60**, 534 (1994) [JETP Lett. **60**, 548 (1994)].
3. S. Chikasumi, *The Physics of Ferromagnetism: Magnetic Characteristics and Engineering Applications* (Syokabo, Tokyo, 1984; Mir, Moscow, 1987) (translated from Japanese).
4. K. P. Belov, A. K. Zvezdin, A. M. Kadomtseva, and R. Z. Levitin, *Orientation Transitions in Rare-Earth Magnets* (Moscow, 1979).
5. R. M. Vakhitov, R. M. Sabitov, and Z. V. Gabbasova, Phys. Status Solidi B **165**, K87 (1991).
6. R. M. Vakhitov and V. V. Grinevich, Fiz. Met. Metalloved. **80** (4), 168 (1995).
7. S. G. Abarenkova, S. Sh. Gendelev, L. K. Zarembo, *et al.*, Fiz. Tverd. Tela (Leningrad) **27**, 2450 (1985) [Sov. Phys. Solid State **27**, 1467 (1985)].

Translated by A. Khzmalyan

Technical Physics, Vol. 47, No. 7, 2002, pp. 866–868. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 72–74.

SOLIDS

On Elastic Deformation of the Indentation Surface in Alkali-Haloid Crystals

V. A. Feodorov, L. G. Karyev, and A. N. Glushkov

Derzhavin State University, Tambov, 392622 Russia

e-mail: feodorov@tsu.tmb.ru

Received September 13, 2001; in final form, December 11, 2001

Abstract—It is demonstrated experimentally that in the microindentation process, elastic deformation takes place of the crystal surface region under load containing the forming imprint. A technique is proposed for determining this elastic deformation in alkali-haloid crystals. *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Deformation of a material under an indenter includes a plastic component, which is due to irreversible normal and tangential displacements of the material, and an elastic component, which is considered to be an elastic recovery of the imprint [1]. The latter is evident as the difference in the shape and size of the indenter imprint after it has been lifted. One of the promising methods of studying the properties of plastic deformation under indentation is continuous registration of the indenter penetration depth with increasing load [2–6].

However, in their analysis of plasticity mechanisms as they change with indenter penetration, the above studies did not take into account the contribution of elastic deformation [elastic bending of the surface (Fig. 1)] to the total indenter displacement.

In view of the aforesaid, we aim to develop a method for registration and qualitative estimation of the elastic bending of the surface under different loads on the indenter.

MATERIALS AND EXPERIMENTAL TECHNIQUE

Single crystals of LiF (10^{-3} wt %, Ca^{+2}, Mg^{+2}, Ba^{+2}), NaCl (10^{-2}–10^{-5} wt %, Fe^{+2}), KCl 10^{-2}–10^{-5} wt %, Mg^{+2}) were used in the experiments. The sample dimensions were $4 \times 8 \times 20$ mm. The experiments were carried out at a temperature of 293 K in a PMT-3 indentometer with a standard indenter. The indenter load was measured in the range of $0.1 \le P \le 1.0$ N. Samples in all experiments had freshly cleaved surfaces. Changes of the crystal surface relief were controlled with an MII-4 Linnik microinterferometer.

To solve the problem, a (001) crystal face was indented in the immediate vicinity of the banks of the cracks intentionally introduced in the crystal.

In the first series of experiments, cleaved 10–15 mm long macrocracks were produced by applying a knife blade to the sample end surface (Fig. 2a). In the second series, microscopic cracks in LiF single crystals were introduced in the {110} planes following a technique described in [7] (Fig. 2b).

In the indentation process, the (001) crystal surface near the crack bank subjected to the action of the indenter experiences an elastoplastic bending. The opposite bank of the crack does not bend and will, therefore, be left with a trace from the indenter edge, a notch (the imprint is oriented with its diagonal at right angles to the crack plane).

In all experiments, the (010) crack banks before indentation were on the same horizontal level.

EXPERIMENTAL RESULTS

Results of the first series of experiments are shown in Fig. 3a: between the imprint and the notch from the indenter on the opposite bank of the crack there is a crystal surface region missing in the imprint. Such a picture is observed not only near the crack tip, where the distance between the banks is $l \sim 1$ μm, but also in the region where the crack opening is large ($l \ge 10$ μm).

In the second series of experiments, as in the first, the banks opposite the indented ones also bear notches left by the indenter (Fig. 3b).

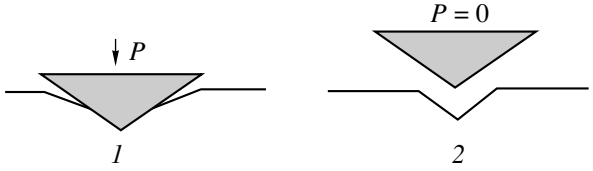


Fig. 1. Elastic deformation of the surface of the crystal region under the indenter. (*1*) Under load, (*2*) unloaded.

1063-7842/02/4707-0866$22.00

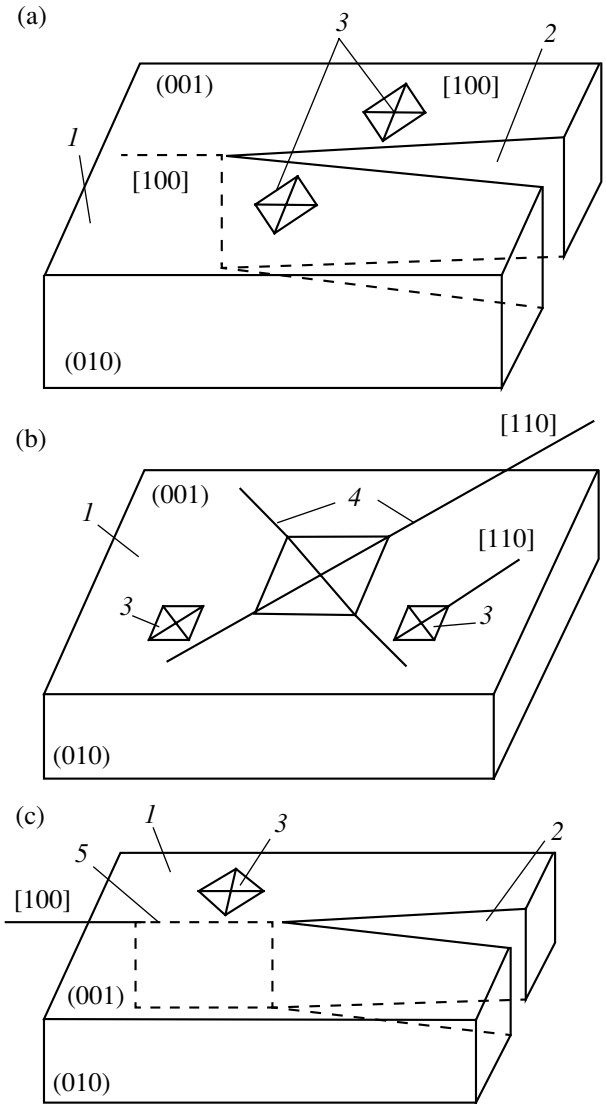


Fig. 2. Schematics of the (a) first, (b) second, and (c) third series of experiments. (*1*) (001) crystal plane to be indented; (*2*) cleaved crack in the (010) plane; (*3*) indenter imprints on the (001) surface regions under study; (*4*) microcracks along the {110} plane produced by the technique described in [7]; and (*5*) continuation trace of the cleaved crack.

DISCUSSION OF RESULTS

The experimental results can be explained by combined displacement of the indentured region of the crack bank containing plastic and elastic components.

In order to essentially exclude the effect of the plastic deformation that appears in the above experiments, a third series of experiments was carried out.

On the (001) surface of NaCl and KCl crystals (in these crystals, elastic recovery of the imprint does not take place [1]), an imprint was indented near the continuation of the trace of a cleaved crack in the (010) plane produced in the crystal (Fig. 2c). Then, the crack growth was initiated and its size made the same as in

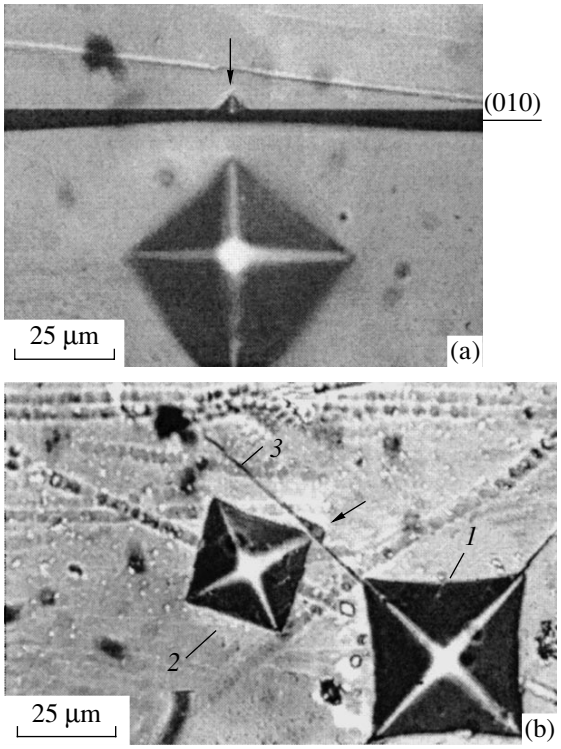


Fig. 3. Indented surfaces of the single crystals: close to a microcrack, NaCl (10^{-2} wt %), P = 0.3 N (a); close to microcrack *3* along the {110} plane resulting from imprint *1* (b); *2* is the indentation that left a marking on the opposite bank of the microcrack, LiF (10^{-3} wt %), P = 0.5 N (marks are indicated by arrows).

Fig. 2a. Further, the indenter was applied again to the initial imprint at the same orientation and load as in the first instance. On the opposite bank of the crack, a notch from the indenter similar to that seen in Fig. 3a was observed.

This experiment proves that plastic deformation cannot be the cause of the notch emerging at microindentation. Therefore, the experimental results can be explained by elastic bending of the region of the surface being indented under load.

The elastic deformation (h, Fig. 4) of the crystal under an indenter can be found from the expression

$$h = y\cot\frac{\alpha}{2}.$$

If β is the angle between the pyramid faces (β = 136°), then, expressing α in terms of β, we ultimately get

$$h \approx 0.29y.$$

Thus, by determining y experimentally, we get the amount of elastic bending of the crystal surface under the indenter. It is important to emphasize that the $h(P)$ dependences obtained by the methods illustrated in Figs. 2a and 2c practically coincide.

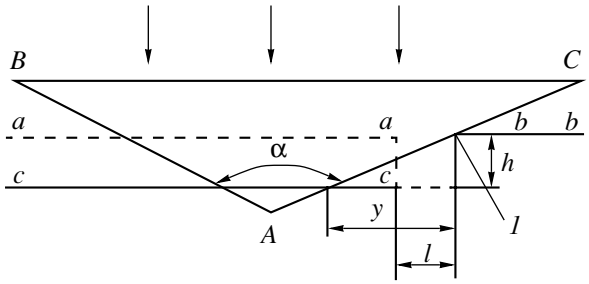


Fig. 4. Schematic used in the calculation of elastic deformation of the crystal surface by the indenter: *AB, AC*—indenter edges; *y*—distance from the imprint edge to the marking; α—the angle contained by the pyramid edges; *aa* and *bb*—initial levels of the crack banks; *cc*—surface level resulting from elastic displacement; *h*—magnitude of the elastic displacement; and *1*—marking.

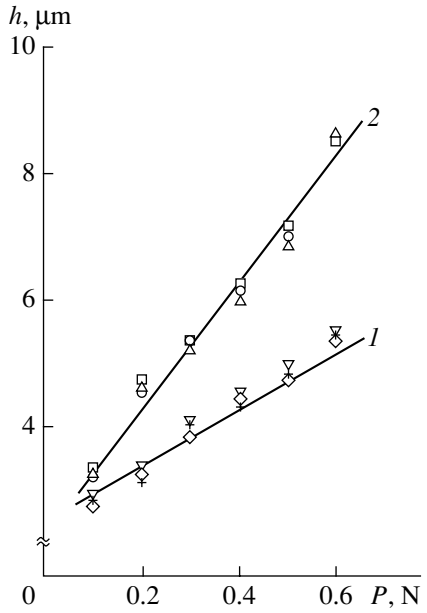


Fig. 5. Dependence of the elastic displacement of the indented surface in alkali-haloid crystals with different content of impurities on the indenter load: (*1*) NaCl ((▽) 10^{-5}, (◇) 10, (+) 10^{-2} wt %); (*2*) KCl ((□) 10^{-5}, (○) 10^{-3}, (△) 10^{-2} wt %).

Figure 5 shows $h(P)$ dependences for NaCl and KCl single crystals, in which elastic recovery of the imprint does not take place [1]. The measurement error for these quantities does not exceed 10%. It is seen that the difference in impurity concentrations in the crystals has almost no effect on the magnitude of elastic deformation, but in crystals with a high Young's modulus (NaCl), the magnitude of h is appreciably greater than in samples with a lower Young's modulus (KCl).

At the instant of indentation of the sample surface at points far from the crystal edge, the elastic deformation values (h) appear to be slightly less.

CONCLUSIONS

(1) In indentation of crystal surfaces, the total displacement of the indenter is a sum of (i) the plastic flow of the material (formation of an imprint) and (ii) the elastic bending of the loaded surface region together with the imprint being formed. The magnitudes of these displacements are comparable.

(2) In indentation of crystals, the imprint is formed in the elastically deformed region of the sample, which apparently affects the quantitative estimates of not only microhardness but also of the activation characteristics of the plastic deformation processes occurring under the indenter.

REFERENCES

1. Yu. S. Boyarskaya, D. Z. Grabko, and M. S. Kats, *Physics of Microindentation Processes* (Shtiintsa, Chisinau, 1986).
2. S. I. Bulychev, V. P. Alekhin, *et al.*, Fiz. Khim. Obrab. Mater., No. 2, 110 (1981).
3. V. P. Alekhin, G. S. Berman, *et al.*, Zavod. Lab. **38**, 488 (1972).
4. V. I. Degtyarev and V. Ya. Lanveshkin, Zavod. Lab. **43** (8), 999 (1977).
5. Yu. I. Golovin and A. I. Tyurin, Vestn. Tambov. Gos. Univ. **5**, 249 (2000).
6. V. I. Orlov *et al.*, Vestn. Tambov. Gos. Univ. **5**, 367 (2000).
7. L. G. Karyev, V. A. Fedorov, *et al.*, Vestn. Tambov. Gos. Univ. **5**, 378 (2000).

Translated by B. Kalinin

Technical Physics, Vol. 47, No. 7, 2002, pp. 869–872. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 75–78.

SOLIDS

Temperature Dependences of Thermal Expansion and Saturation Magnetization in Fe(67.0%)–Ni(32.5%)–Co(0.5%) Invar Alloy with Nanocrystalline Structure

Kh. Ya. Mulyukov, I. Z. Sharipov, I. Kh. Bitkulov, and R. R. Mulyukov

Institute of Problems of Metal Superplasticity, Russian Academy of Sciences, ul. Khalturina 34, Ufa, 450001 Bashkortostan, Russia

e-mail: radik@anrb.ru

Received December 4, 2001

Abstract—The effect of the nanocrystalline structure of Fe(67.0%)–Ni(32.5%)–Co(0.5%) Invar alloy on its thermal expansion is considered. It is found that the structure with grain mean sizes of about 100 nm increases its temperature coefficient of thermal expansion in the range of "invarness," i.e., in the temperature range where the alloy offers the Invar properties. Reasons for this behavior are analyzed by taking the temperature dependence of the saturation magnetization. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

In recent years, materials with nanocrystalline structure (grain mean sizes of about 10–100 nm) have attracted much attention. The physical properties of these materials differ considerably from those of coarse-grained ones [1–3]. Therefore, the investigation of nanocrystalline alloys used in practice (including Invar alloys) is of current interest. Because of the anomalously low coefficient of thermal expansion in the temperature range of invarness, these materials are widely applied in instrument making [4]. It should be noted that the nature of the Invar effect still remains unclear even though the properties of the alloy have been examined over the century. The study of the physical properties of this alloy in different structural states, including nanocrystalline one, might provide additional insights into this effect. In this work, we investigate the thermal expansion of Fe(67.0%)–Ni(32.5%)–Co(0.5%) Invar alloy with nanocrystalline structure as a function of temperature.

EXPERIMENTAL TECHNIQUE

The nanocrystalline structure of the alloy was formed by its severe plastic deformation on Bridgman anvils at room temperature [2]. Specimens were cut from a rolled sheet and annealed *in vacuo* at a temperature of 1070 K for 2 h.

The temperature dependence of the saturation magnetization was taken with a vacuum self-correcting magnetic balance [5] upon heating and cooling the specimens at a rate of 15 K/min in the range from 78 to 1070 K. The thermal expansion of the specimens was studied with a dilatometer at the same temperatures. The microstructures of the specimens were examined with a JEM 2000EX electron microscope.

RESULTS

As follows from the electron microscopy data, the severe plastic deformation of the Invar alloy produces a nanocrystalline structure with a mean grain size of about 100 nm (Fig. 1). The structure is heavily distorted and cold-hardened. The dislocation density in nanocrystalline structures is difficult to estimate. The ring

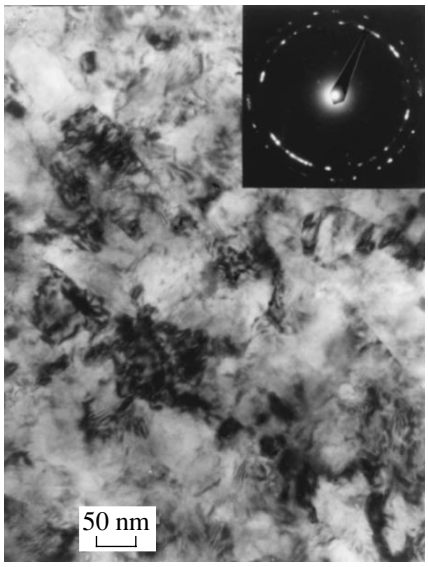


Fig. 1. Microstructure of the Fe(67.0%)–Ni(32.5%)–Co(0.5%) nanocrystalline alloy.

1063-7842/02/4707-0869$22.00

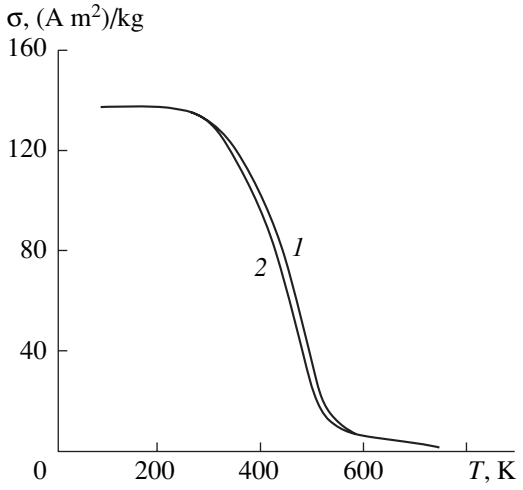


Fig. 2. Temperature dependence of the saturation magnetization for the coarse-grained specimen of the Fe(67.0%)–Ni(32.5%)–Co(0.5%) alloy: (*1*) heating and (*2*) subsequent cooling.

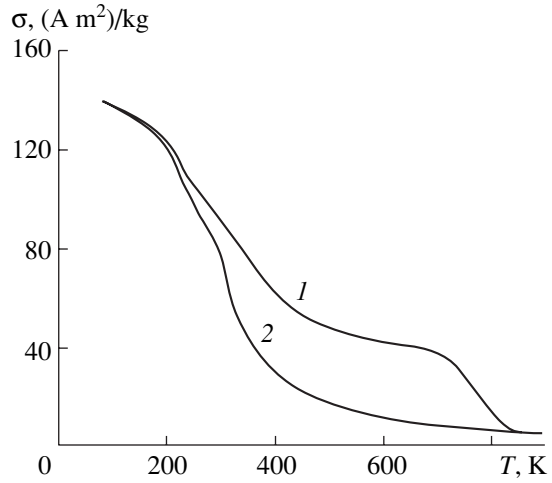


Fig. 3. Temperature dependence of the saturation magnetization for the nanocrystalline specimen of the Fe(67.0%)–Ni(32.5%)–Co(0.5%) alloy: (*1*) heating and (*2*) subsequent cooling.

electron diffraction pattern taken from an area of 0.5 μm² indicates large-angle misorientations between the grains. Two-hour annealing *in vacuo* at a temperature of 800°C recovers the structure. The grain size increases to several microns.

Figure 2 shows the temperature dependence of the saturation magnetization, $\sigma(T)$, for the coarse-grained structure. The shape of the curve is similar to that for ferromagnets. The magnetization in the range between 78 and 250 K remains virtually unchanged and then decreases. The decrease is the most significant in the range from 270 to 520 K. At higher temperatures, the magnetization smoothly tends to zero. The shapes of the curves $\sigma(T)$ upon (*1*) heating and (*2*) cooling of the specimen are similar. However, there is a temperature hysteresis: the curve obtained upon cooling shifts toward lower temperatures by 15 K.

The curves $\sigma(T)$ for the nanocrystalline specimen recorded under the same conditions are presented in Fig. 3. Their shape differs noticeably from that for the coarse-grained structures (Fig. 2). First, note the difference between the heating curves (curves *1* in Figs. 2 and 3). The dependence $\sigma(T)$ for the nanocrystalline structure, in contrast to the corresponding curve for the coarse-grained specimen, starts decreasing from the very beginning of the temperature range. One can distinguish four portions in it. Within the first portion (80–500 K), the magnetization decreases rapidly. Next, between 500 and 700 K, a "plateau" on which it changes insignificantly is observed. Within the third portion (700–850 K), the magnetization rapidly decreases again. At temperatures exceeding 850 K, the magnetization smoothly tends to zero. For the nanocrystalline specimen, the dependence $\sigma(T)$ recorded upon cooling (curve *2* in Fig. 3) differs significantly from that taken upon heating. In the former, only two distinct portions can be distinguished. When the specimen is cooled from 850 to 450 K, the magnetization slowly grows. Within the second portion (between 450 and 80 K), the magnetization rapidly reaches its initial value.

The curves $\sigma(T)$ demonstrate one more difference between the nanocrystalline and coarse-grained structures. As was noted above, the coarse-grained specimens (Fig. 2) exhibit a minor temperature hysteresis: curves *1* and *2* are shifted relative to each other by 15 K along the temperature axis. For the nanocrystalline specimens, the temperature hysteresis is as high as several hundreds of degrees. Finally, it should be noted that the magnetizations of both structures at 80 K are practically the same.

The measurements of the thermal expansion $[\Delta l/l](T)$ for specimens of both types are depicted in Fig. 4. These curves have two characteristic sections: with small and large slopes, i.e., with small and large thermal expansion coefficients (TECs), respectively. The basic requirement for Invar alloys is that they have a small TEC value in a certain temperature range in order to meet the Invar properties. In curve *1* (Fig. 4), corresponding to the coarse-grained structure, the section with a relatively small TEC extends to 400 K, whereas in curve *2*, corresponding to the nanocrystalline structure, it is limited by a temperature of 250 K.

Curves *1* and *2* in Fig. 4 differ not only by the temperature ranges with small TECs but also by the slopes of these sections. The slope of the small-TEC section in the curve $[\Delta l/l](T)$ for the nanocrystalline state is six times larger than that for the coarse-grained state of the specimen. Hence, the TEC in the former case is six times larger than in the latter. Thus, the formation of the nanocrystalline structure in the alloy adversely affects its Invar properties, namely, increases the thermal

expansion coefficient and narrows the temperature range of invarness.

DISCUSSION

Thus, the properties of the Invar alloy depend strongly on its structure. This dependence is of particular interest in the range of small TECs. Severe plastic deformation of the Fe(67%)–Ni(32.5%)–Co(0.5%) alloy forms a microstructure with grain mean sizes of about 100 nm and a high density of dislocations. As a result, the Invar properties degrade, the temperature of ferromagnetic disordering lowers, and the temperature hysteresis of the saturation magnetization becomes more pronounced.

Analysis of the results obtained is difficult because there is no reliable physical model describing the behavior of Invars. Invarness is often approximated by the relationship [4]

$$\alpha = -\frac{\sigma'_T}{3K} + \frac{\gamma c_V}{3VK}. \quad (1)$$

Here, the first term is the contribution from the potential energy of atomic interaction into the thermal expansion ($\sigma' = \partial\sigma/\partial T$ reflects the variation of the atomic interaction forces with temperature, K is the compression modulus) and the second one due to the anharmonicity of crystal lattice vibrations is the contribution of the kinetic energy (γ is the Grüneisen constant, c_V is the specific heat at constant volume). The competition between interatomic attractive forces growing with temperature below the Curie point and the Grüneisen thermal expansion specifies the Invar dilatation.

Earlier, it was shown [6–8] that the plastic deformation of Fe(64%)–Ni(36%) Invar alloy decreases the TEC. In [8], the alloy was made nanocrystalline by severe plastic deformation. For our Invar alloy, it was found that the severe plastic deformation increases the TEC. This difference is most likely to be associated with the fact that, in [6–8], the thermal expansion due to anharmonic atomic vibrations is partially compensated for by a decrease in the interatomic spacing with temperature due to magnetostriction. At the same time, it is known that the heating of Fe–Ni Invars with less than 34% Ni causes the $\alpha \longrightarrow \gamma$ martensitic transformation, which takes place at temperatures lower than those of the magnetic transformation [4]. Consequently, the reduction of interatomic attractive forces in our alloy with increasing temperature is explained by the martensitic transformation.

The temperature hysteresis of the saturation magnetization in our specimens also suggests the martensitic transformation. In other words, the shapes of the curves $\sigma(T)$ for the nanocrystalline specimen (Fig. 2) and for ferromagnets are similar only formally. A decrease in the magnetization with increasing temperature is caused by the transition from the ferromagnetic martensitic α phase to the paramagnetic austenitic γ phase.

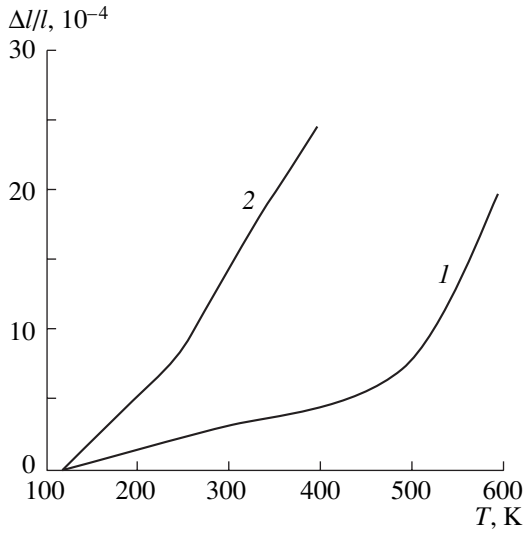


Fig. 4. Thermal expansion $[\Delta l/l](T)$ of the Fe(67.0%)–Ni(32.5%)–Co(0.5%) alloy with (*1*) coarse-grained and (*2*) nanocrystalline structures.

In the coarse-grained alloy, the ferromagnetic (martensitic) phase is almost stable up to 270 K and intense $\alpha \longrightarrow \gamma$ transformation with weak irreversibility is observed only in the temperature range of 270–520 K.

From the curve $\sigma(T)$ for the nanocrystalline specimen, it follows that its ferromagnetic α phase is of low temperature stability and that the $\alpha \longrightarrow \gamma$ transformation in it is a rather complicated two-stage process. The first stage begins at a temperature of about 200 K and ends at 450 K. Next, in the range of 450–700 K, the phase transformation nearly ceases and is completed at higher temperatures. The lower temperature of phase transformation in the nanocrystalline specimen as compared to the coarse-grained one agrees with the well-known fact that plastic deformation facilitates phase transformations [2]. The poor temperature stability of the α phase in the former specimen may be explained by both the small grain sizes and high density of lattice defects in small grains formed by severe plastic deformation.

The complex transformation in the nanocrystalline specimen is likely associated with the fact that severe plastic deformation of the alloy causes α-phase modifications with different transition temperatures to appear. It should be noted that some modification of this phase disappears after heating to 1000 K and does not appear again in the curve $\sigma(T)$ taken on cooling the specimen. It seems that this high-temperature α modification disappears because of the recovery of the system and the noticeable growth of the grains at high temperatures.

In addition, it is worth noting that the formation of the nanocrystalline structure breaks the translation symmetry, increases the volume fraction of grain boundaries and other lattice defects, and generates long-range internal stresses in the material [2]. All

these factors enhance the anharmonicity of the crystal lattice and, accordingly, deteriorate the Invar properties, i.e., raise the thermal expansion coefficient.

CONCLUSION

Thus, the formation of a nanocrystalline structure in Fe(67.0%)–Ni(32.5%)–Co(0.5%) Invar by severe plastic deformation significantly modifies its properties: the TEC increases and the temperature range of the Invar properties narrows. This is associated with the effect of martensitic transformation on the Invar properties, specifically, with a decrease in the temperature of this transformation when the alloy passes into the nanocrystalline state. In addition, the nanocrystallinity enhances the anharmonicity of the crystal lattice, raising the thermal expansion.

ACKNOWLEDGMENTS

This work was supported by the Complex Program "Nanocrystals and Supramolecular Systems" of the Russian Academy of Sciences (grant "Structure and Properties of Nanocrystals Obtained by Severe Plastic Deformation") and by the Russian Foundation for Basic Research (grant no. 00-02-17723).

REFERENCES

1. C. Suryanarayana, Int. Metall. Rev. **40**, 41 (1995).
2. R. R. Mulyukov and M. D. Starostenkov, Acta Metall. Sin. **13**, 301 (2000).
3. Kh. Ya. Mulyukov, Doctoral Dissertation (Ufa, 1998).
4. A. I. Zakharov, *Physics of Precision Alloys with Special Thermal Properties* (Metallurgiya, Moscow, 1986).
5. Kh. Ya. Mulyukov, I. Z. Sharipov, and S. S. Absalyamov, Prib. Tekh. Éksp., No. 3, 149 (1998).
6. V. P. Voroshilov, A. I. Zakharov, V. M. Kalinin, and A. S. Uralov, Fiz. Met. Metalloved. **35**, 953 (1973).
7. A. I. Zakharov, B. V. Molotilov, and L. V. Pastukhova, Fiz. Met. Metalloved. **37**, 654 (1974).
8. I. Kh. Bitkulov, V. A. Kazantsev, V. I. Kopylov, and R. R. Mulyukov, Izv. Vyssh. Uchebn. Zaved., Fiz., No. 2, 69 (2001).

Translated by Yu. Vishnyakov

Technical Physics, Vol. 47, No. 7, 2002, pp. 873–879. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 79–85.

SOLIDS

Motion of Domain Walls under the Action of Spin-Polarized Current in a Magnetic Junction

Yu. V. Gulyaev*, P. E. Zil'berman*, R. J. Elliott, and É. M. Épshteĭn***

** Institute of Radio Engineering and Electronics (Fryazino branch), Russian Academy of Sciences, pl. Vvedenskogo 1, Fryazino, Moscow oblast, 141190 Russia*

*** University of Oxford, Department of Physics, Theoretical Physics, Oxford OX1 3NP, United Kingdom*

Received December 11, 2001

Abstract—The effect of spin-polarized current on a domain structure in a magnetic junction consisting of two ferromagnetic metallic layers separated by an ultrathin nonmagnetic layer is studied within a phenomenological theory. The magnetization of one ferromagnetic layer (layer 1) is assumed to be fixed, while that of the other ferromagnetic layer (layer 2) can be freely oriented both parallel and antiparallel to the magnetization of layer 1. Layer 2 can be split into domains. Charge transfer from layer 1 to layer 2 is not attended with spin scattering by the interface but results in spin injection. Due to *s–d* exchange interaction, injected spins tend to orient the magnetization in the domains parallel to layer 1. This causes the domain walls to move and "favorable" domains to grow. The average magnetization current injected into layer 2 and its contribution to the *s–d* exchange energy are found by solving the continuity equation for carriers with spins pointing up and down. From the minimum condition for the total magnetic energy of the junction, the parameters of the periodic domain structure in layer 2 are determined as functions of current through the junction and magnetic field. It is shown that the spin-polarized current can magnetize layer 2 up to saturation even in the absence of an external magnetic field. The associated current densities are on the order of 10^5 A/cm^2. In the presence of the field, its effect can be compensated by such a high current. Current-induced magnetization reversal in the layer is also possible. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Magnetic junctions involving two ferromagnetic metallic electrodes and a nonmagnetic spacer have recently become the subject of extensive investigation from both the scientific and applied viewpoints [1–3]. In particular, of great interest is the effect of spin-polarized current on the magnetic states of the junctions. Two (surface and volume) mechanisms behind this effect have been discussed. In the former case, charge carriers experience spin-dependent scattering by the interfaces. As a result, the angular momentum of the spin is transferred from the charge carriers to a set of localized spins. This causes magnetization reversal in the electrodes. Such a mechanism was first considered in theoretical works [4–8] and was then used in the interpretation of experiments on magnetization switching [9–15].

The volume mechanism is based on spin injection, which was first described in [16, 17] and later in [18, 19]. On the one hand, the carrier spins add to the magnetic moment of the electrode volume, thereby changing the magnetic states. On the other hand, they influence the magnetization orientation in the electrode through *s–d* exchange interaction. Such an effect of exchange was first discussed in terms of the microscopic theory in [20, 21]. It was shown that this effect can be viewed as a specific exchange interaction between the electrodes via charge carriers, the so-called nonequilibrium exchange interaction (NEXI). According to [21], the NEXI energy may be several orders of magnitude higher than the energy of exchange interaction between electrodes in equilibrium. Within the NEXI concept, another explanation (other than that in [4–8]) of the magnetization switching was suggested [22]. Experimental data for current hysteresis [23–25] can also be considered in favor of the volume mechanism. In our opinion, however, the issue of the prevailing mechanism is still an open question. This work is aimed at gaining additional information about possible manifestations of the volume mechanism.

In pioneering works on the effect of current on the magnetic state of the junction, no consideration was given to the domain structure (DS) of the electrodes. Nevertheless, such a structure has been repeatedly discovered in experiments [2, 26–29]. It appears that, theoretically, the effect of current on the DS of a magnetic junction has not been studied so far. However, a theoretical approach is necessary for the interpretation of experiments. Here, we use a simple model of the junction and a phenomenological theory to show for the first time that spin-polarized current can rearrange the DS through the motion of the domain walls by means of spin injection and *s–d* exchange interaction.

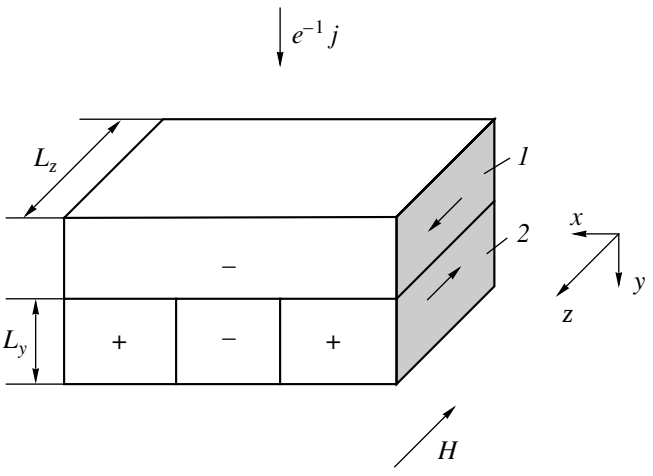
1063-7842/02/4707-0873$22.00

MODEL OF THE JUNCTION

A three-layer magnetic junction is depicted in the figure. All three layers are assumed to be extended in the plane *xy*; therefore, edge effects can be neglected. The side surface $z = L_z/2$ shows magnetic charges. Taking account of these charges with the total magnetic energy of the layer minimized causes the DS to appear. The structure is absent in layer 1, and it invariably remains uniformly magnetized because of the contact with an extra antiferromagnetic layer (omitted in the figure), which produces high unidirectional anisotropy in layer 1 (see, e.g., [30, 31]). Such "biased" layers are widely used in experiments [2]; therefore, our assumption is quite acceptable. It greatly simplifies the situation, since the current has an effect only on layer 2 in this case.

Let layer 2 be an ultrathin ferromagnetic film of a magnetically soft material with a high magnetization (Fe, Co, Ni, or Pt). Such films are routinely used in magnetic junctions. Also, let the easy magnetic axis lie in the plane of the film parallel to the *z* axis. Then, the magnetization vector is also parallel to the *z* axis.

For simplicity, we assume that layer 2 is extended along the *x* axis (its length $L_x \longrightarrow \infty$), its width L_z is limited, and its thickness L_y is small ($L_y \ll L_z$). Under these conditions, the DS represents a periodic (along the *x* axis) set of domains with alternating magnetization directions (parallel and antiparallel to the *z* axis; see the figure). In total equilibrium (without the current) and in the absence of an external magnetic field ($H_{ext} = 0$), the width w_a of the "antiparallel" domains (*a* domains), where the magnetization is directed oppositely to that in layer 1, is somewhat larger than the width w_p of the "parallel" domains (*p* domains). This is associated with the need for compensating for the magnetic field from layer 1 far away from the junction in order to minimize the total magnetic energy.



Schematic representation of the magnetic junction. *1*, ferromagnetic metallic electrode with the fixed orientation of the magnetic moment; *2*, ferromagnetic metallic electrode with the domain structure. Arrows show the directions of the magnetization vectors in layer 1 and in domains of layer 2, the direction of the carrier flux $e^{-1}j$ (*e* is the charge of a carrier), and the direction of the external field *H*. Shown also are the coordinate axes, junction width L_z, and the thickness L_y of layer 2.

Following [32], we will call DSs of this type Cotton structures. Their properties differ markedly from those of DSs where the magnetization is normal to the developed surface (Faraday structures in terms of [32]). The differences were discussed in our previous article [33], to which we will often refer below. Specifically, the width of Cotton domains varies with L_z and the domain wall energy linearly (not as the square root of them) and is inversely proportional to the magnetization in the second (not first) power.

Now, consider the effect of current on the magnetic state of layer 2. We will take advantage of the fact that the current passing normally to the interface from layer 1 to any of the domains in layer 2 depends on the magnetization orientation in this domain. This is an experimental observation (see, e.g., [34]) that reflects the fact that the direction of this vector specifies the position of the energy bands and the density of states of charge carriers. Thus, the currents passing into *a* and *p* domains will differ from each other. Moreover, such polarization currents will variously affect the associated domains, i.e. variously modify their width and the integral (over the cross section) conductivity of the junction. Below, we will evaluate only the width of the domains in the presence of current.

The partial current densities j_p and j_a passing, respectively, to *p* and *a* domains, as well as the external magnetic field *H*, will be considered as factors influencing the DS. This means that we will analyze and evaluate changes in the DS as functions of j_p, j_a, and *H*. We will not find the dependence of the currents j_p and j_a on the external voltage *V*, since this is an independent problem having little to do with the aim of this work. Note in addition that these currents have been theoretically evaluated within several specific models (the model of uniformly magnetized electrodes under the assumption of current tunnel passage through the spacer [35, 36] is an example).

It is of interest that our model totally denies the surface mechanism of spin scattering. Indeed, in our model, the magnetization vectors in layers 1 and 2 and the *z* axis are collinear. Therefore, the quantum state of a charge carrier remains unchanged after it has been transferred through the spacer from layer 1 to layer 2. In other words, a transferred carrier will be in one of the energy subbands of layer 2 and have a particular energy. For such a carrier, the probability amplitude of quantum transitions between various spin subbands equals zero. Accordingly, for this amplitude, phase averaging, which causes the effect of spin surface scattering [4–9], does not occur. Thus, our model leaves room only for the spin-injection mechanism of the effect of current on the magnetic state of the junction.

EFFECT OF SPIN INJECTION ON THE JUNCTION ENERGY

For subsequent calculations, it is necessary to determine the magnetization of layer 2 as a function of the current through the junction. This magnetization has two components: one from the localized spins and the other from the spins of charge carriers injected into layer 2. We assume that the magnetization of the localized spins $M_0 > 0$ in *p* domains and $M_0 < 0$ in *a* domains. From what follows, it will be clear that the widths of both domains, w_a and w_p, are much larger than the thicknesses of layers 1 and 2 and than the spin relaxation length l_s; that is,

$$w_p, w_a \gg l_s, L_y. \tag{1}$$

Under these conditions, the problem of spin injection into an individual domain and current passage through this domain can be considered to be one-dimensional. Then, the magnetization of the charge carriers in the domains is

$$m_{p,a}(y) = \mu_B n P_{p,a}(y), \tag{2}$$

where μ_B is the effective Bohr magneton (i.e., the magnetic moment of a single charge carrier). The quantity

$$P(y) = \frac{n_\uparrow(y) - n_\downarrow(y)}{n} \tag{3}$$

is the carrier polarization at a point *y* in layer 2. In (3), $n_\uparrow(y)$ and $n_\downarrow(y)$ are the respective concentrations of carriers with spins parallel and antiparallel to the magnetization in layer 1; $n = n_\uparrow(y) + n_\downarrow(y)$ is the carrier total concentration, which is assumed to be constant because of the metal quasi-neutrality; and $P_{p,a}$ is quantity (3) for *p* and *a* domains. The polarization of the carriers in layer 1 at the interface with layer 2 will be denoted by P_1; the equilibrium (without injection) polarization of the carriers in layer 2 (in domains of both types), by P_2. As was already noted, under the injection, the polarization in layer 2 varies.

One more assumption is that the current passing through the junction is stationary (time-invariable). The spin flux density is given by

$$J = \frac{\hbar}{2e}(j_\uparrow - j_\downarrow), \tag{4}$$

where $j_\uparrow$ and $j_\downarrow$ are the densities of the electrical currents for carriers with different spin polarizations.

To simplify the calculations, the mobility, diffusion coefficient *D*, and the spin relaxation time τ_s of the carriers are taken to be independent of the spin polarization. Then, for the currents $j_\uparrow$ and $j_\downarrow$, we have

$$\begin{aligned} j_\uparrow &= -eD\frac{\partial n_\uparrow(y)}{\partial y} + en_\uparrow(y)v, \\ j_\downarrow &= -eD\frac{\partial n_\downarrow(y)}{\partial y} + en_\downarrow(y)v, \end{aligned} \tag{5}$$

where *v* is the drift velocity in the external electrical biasing field.

The concentrations $n_\uparrow(y)$ and $n_\downarrow(y)$ must satisfy the steady-state continuity equations

$$\begin{aligned} \frac{1}{e}\frac{\partial j_\uparrow(y)}{\partial y} &= -\frac{(n_\uparrow(y) - n_\uparrow^e)}{\tau_s}, \\ \frac{1}{e}\frac{\partial j_\downarrow(y)}{\partial y} &= -\frac{(n_\uparrow(y) - n_\downarrow^e)}{\tau_s}, \end{aligned} \tag{6}$$

where $n_\uparrow^e$ and $n_\downarrow^e$ are the equilibrium concentrations and

$$\begin{aligned} [n_\uparrow(y) + n_\downarrow(y)] &= (n_\downarrow^e + n_\uparrow^e) = n, \\ [j_\uparrow(y) + j_\downarrow(y)] &= env \equiv j. \end{aligned} \tag{7}$$

Subtracting one equation in (6) from the other, substituting (5) in the resultant expression, and taking into account (7), we find

$$\frac{\partial^2 P_{p,a}(y)}{\partial^2 y} - \frac{v}{D}\frac{\partial P_{p,a}(y)}{\partial y} - \frac{[P_{p,a}(y) \mp P_2]}{D\tau_s} = 0, \tag{8}$$

where the upper and lower signs refer to *p* and *a* domains, respectively.

As boundary conditions to Eq. (8), we take condition (4) of spin flux continuity at the interface $y = 0$ and the condition of spin balance at $y = L_y$. Estimating the spin flux at $y = 0$ with the electron concentration gradient in layer 1 neglected, we obtain

$$J_{p,a}(y, 0) = \frac{\hbar}{2e} j_{p,a} P_1.$$

Eventually, the boundary conditions take the form

$$\begin{aligned} \left[-enD\frac{\partial P_{p,a}(y)}{\partial y} + j_{p,a}P_{p,a}(y)\right]_{y=0} &= j_{p,a}P_1, \\ P_{p,a}(L_y) &\approx \pm P_2. \end{aligned} \tag{9}$$

The solution to (8) that satisfies boundary conditions (9) has the form

$$\begin{aligned} P_{p,a}(y) = \pm P_2 &+ \frac{2r_{p,a}(P_1 \mp P_2)}{r_{p,a} + \sqrt{1 + r_{p,a}^2}} \\ &\times \exp\left[-\frac{-y\sqrt{1 + r_{p,a}^2} - r_{p,a}}{l_s}\right], \end{aligned} \tag{10}$$

where

$$r_{p,a} = \frac{j_{p,a}\sqrt{\tau_s}}{2en\sqrt{D}}, \quad l_s = \sqrt{D\tau_s} \tag{11}$$

and it is assumed that

$$L_y > \frac{l_s}{(\sqrt{1 + r_{p,a}^2} - r_{p,a})}. \quad (12)$$

Owing to inequality (12), the second condition in (9) is also satisfied.

In essence, the parameter $r_{p,a}$ characterizes the significance of drift in charge transfer. If $r_{p,a} \gg 1$, drift dominates over diffusion; then, within layer 2 (i.e., $0 < y \ll L_y$), we have $P(y) = P_1$ from (10). In other words, the carrier polarizations in the layers equal each other and are current-independent. This is the maximal possible effect of spin injection. Such a situation, however, requires very high current densities, namely,

$$j_{p,a} \gg j_0 \equiv 2en\sqrt{\frac{D}{\tau_s}}. \quad (13)$$

Since $D \sim v_F^2 \tau_p$, the Fermi velocity $v_F \sim 10^8$ cm/s, and the momentum relaxation time τ_p of carriers is $\approx 3 \times 10^{-15}$ s, we find that $j_0 \approx 5.4 \times (10^{10}–10^{11})$ A/cm^2 for $n \sim 10^{22}–10^{23}$ cm^{-3} and $\tau_s \sim 10^{-13}$ s. Thus, for not very high currents, we can consider that $r_{p,a} \ll 1$; this relationship will be used in subsequent calculations.

Now, let us consider the factors contributing to the magnetic energy of the junction. The energy of *s–d* exchange interaction in *p* and *a* domains is proportional, respectively, to the magnetization products $m_p M_0$ and $-m_a M_0$. We introduce the relative widths of the *p* and *a* domains:

$$\xi = \frac{w_p}{W} \quad \text{and} \quad 1 - \xi = \frac{w_a}{W}, \quad (14)$$

where $0 \le \xi \le 1$ and $W = (w_p + w_a)$ is the DS period.

Then, the mean density of the exchange energy (that is, the energy of the whole layer 2 divided by its volume) is given by

$$E_{s-d} = -\alpha M_0[\bar{m}_p \xi - \bar{m}_a(1 - \xi)], \quad (15)$$

where $\alpha \sim A/\mu_b^2 n$ is the dimensionless constant characterizing the *s–d* exchange. Typical values of *A* lie between 0.1 and 10 eV [37]. Then, $\alpha \approx 2 \times (10^4–10^7) \gg 1$. In addition, expression (15) includes the mean magnetization of charge carriers

$$\bar{m}_{p,a} = \mu_B n \bar{P}_{p,a} \equiv \frac{\mu_B n}{L_y} \int_0^{L_y} P_{p,a}(y)\,dy. \quad (16)$$

With expression (10) for $P_{p,a}(y)$ ($r_{p,a} \ll 1$) substituted into (16), we eventually arrive at

$$\bar{m}_{p,a} = \mu_B n \left[\pm P_2 + \frac{j_{p,a}\tau_s}{enL_y}(P_1 \mp P_2) \right]. \quad (17)$$

As follows from (17), the contribution of the injected carriers to the magnetization of layer 2 depends linearly on the current through the junction. This contribution is sufficiently small if

$$j_{p,a} \ll j_c \equiv \frac{enL_y}{\tau_s}. \quad (18)$$

Substituting the numerical values of the associated parameters (see the discussion following formula (13)) and $L_y \approx 50$ nm, we find that $j_c \approx 8 \times (10^{10}–10^{11})$ A/cm^2; that is, $j_c \ge j_0$ in this case, so that condition (18) is well met. Thus, the contribution of the injected carriers to the total mean magnetization $\bar{M}_{p,a} \equiv (\pm M_0 + \bar{m}_{p,a})$ is rather small. Yet, in expression (15), this contribution is multiplied by the large parameter $\alpha \gg 1$. This circumstance is of fundamental significance: it shows that the effect of the carriers injected on the magnetic energy of the junction is largely through strong exchange interaction. Hereafter, we will therefore leave the spin-injection energy term in exchange energy (15), rejecting it in the rest of the contributions to the energy. Specifically, the effect of spin injection on magnetodipole interaction in layer 2 is insignificant. When calculating the energy of this interaction, we will use a simplified expression for the total magnetization:

$$M_{p,a} \approx \bar{M}_{p,a} \approx \pm(M_0 + \mu_B n P_2) \equiv \pm\tilde{M}_0. \quad (19)$$

We must determine the mean density of the magnetodipole interaction energy E_M. It was calculated in [33]. In the case we are interested in most, where the relationships

$$L_z \gg W \gg L_y \quad (20)$$

are fulfilled, the formula derived in [33] has the form

$$E_M = \frac{2L_y}{L_z}\tilde{M}_0^2 \left\{ (2\xi - 1)^2 \left(\ln\frac{\pi L_z}{W} + C \right) - \ln\frac{\pi L_y}{W} - C + \frac{3}{8} - \frac{4}{\pi^2}\sum_{n=1}^{\infty} \frac{\ln n}{n^2}(1 - \cos 2\pi n\xi) \right\}, \quad (21)$$

where $C = 0.5772\ldots$ is the Euler constant and $\tilde{M}_0$ is substituted for M_0 according to (19).

The mean energy density for uniaxial magnetic anisotropy with the easy axis aligned with the *z* axis is given (in our designations) by

$$E_A = -K[M_p^2 \xi + M_a^2(1 - \xi)] \approx -K\tilde{M}_0^2, \quad (22)$$

where substitution (19) is also made.

Thus, E_A in this situation does not depend on the parameters *W* and ξ, which characterize the domains.

Magnetic field H, produced by layer 1 and the external source, adds to the energy density the value

$$E_H = -H[M_p\xi + M_a(1-\xi)] \approx -H\tilde{M}_0(2\xi - 1). \quad (23)$$

We should also take into account the energy contribution from the domain walls. This contribution per unit volume can be represented as

$$E_D = \frac{2\gamma}{W}, \quad (24)$$

where the parameter γ is the energy per unit area of the domain wall. The factor 2 reflects the fact that each DS period involves two domain walls.

Thus, the total magnetic energy of the junction is the sum of the components dependent on the DS parameters W and ξ:

$$E(W, \xi) = E_{s-d}(\xi) + E_M(W, \xi) + E_H(\xi) + E_D(W). \quad (25)$$

EFFECT OF CURRENT ON THE DOMAIN STRUCTURE

Since the current passing through the junction is spin-polarized, the distribution of the carriers over momenta and spins may be steady-state but nonequilibrium. This does not mean that some partial equilibrium in terms of other parameters is impossible when the variations of the spins and momenta from equilibrium are fixed. We suppose that there is partial equilibrium with respect to the partition into magnetic domains. Then, in the presence of the current, the parameters W and ξ are found from the conditions

$$\frac{\partial E(W, \xi)}{\partial W} = 0, \quad \frac{dE(W, \xi)}{\partial \xi} = 0. \quad (26)$$

The former allows us to determine the DS period as a function of ξ:

$$W(\xi) = \frac{\gamma L_z}{4L_y\tilde{M}_0^2\xi(1-\xi)}. \quad (27)$$

Period (27) is minimal, $W_0 = \gamma L_z/L_y\tilde{M}_0^2$, when $\xi = 1/2$, i.e., $w_p = w_a$. This value is reached in equilibrium when the current and the field are absent. The period grows indefinitely when $\xi \longrightarrow 0.1$. We however imposed conditions (20) and, in particular, the condition $L_z \gg W$. Therefore, the calculation is valid when $\xi(1-\xi) \gg W_0/4L_z$. The quantity W_0/L_z is much less than unity; therefore can be as close as possible but not equal to 0.1.

Consider now the latter condition in (26). According to (21), the contribution $E_M(W, \xi)$ to the magnetic energy $E(W, \xi)$ includes the series. In taking the derivative, this series will also be differentiated. The sum of the series resulting from differentiation is known (see, e.g., handbook [38]).

Substituting expression (27) for the DS period into the derivative thus found yields

$$\frac{\partial E_M(w, \xi)}{\partial \xi} = \frac{8L_y}{L_z}\tilde{M}_0^2 f(\xi, R), \quad (28)$$

where

$$f(\xi, R) = (2\xi - 1)\ln(R\xi(1-\xi)) - \ln\left(\frac{1}{\pi}\sin\pi\xi\right) - 2\ln\Gamma(\xi), \quad (29)$$

$R \equiv 2L_z/W_0$, and $0 \le \xi \le 1$. Then, from the second condition in (26), we come to the equation for ξ:

$$f(\xi, R) = \frac{L_z}{4L_y\tilde{M}_0}\left(H + \frac{M_0\mu_B n}{\tilde{M}_0}\frac{\alpha j_e}{2j_c}\right). \quad (30)$$

Equation (30) contains the "effective current density"

$$j_e = j_p(P_1 - P_2) + j_a(P_1 + P_2), \quad (31)$$

which also includes the degree of spin polarization.

The function $f(\xi, R)$ [see (29)] is similar to that introduced in [33]. However, Eq. (30), where this function is equated to an expression containing both the field and the current, is new. Thus, with (30), we can describe the combined effect of the field H and current j_e on the DS.

To do this, it is sufficient to consider the function reciprocal to (29), that is, $\xi(f, R)$. From (30), the function f is a linear combination of the field and current. In other words, the function reciprocal to (29) provides a direct relationship between the DS parameter ξ, H, and j_e. A family of the functions $\xi(f, R)$ for various R is universal and is given in [33]. It has been shown [33] that as $|H|$ or $|j_e|$ grows, a specific (saturation) state appears where the parameter ξ may take either of its two maximal values ($\xi = 0$ or $\xi = 1$). As the magnitude of the field (or current) increases further, this parameter remains unchanged. The saturation is reached if

$$\left|H + \frac{M_0\mu_B n}{\tilde{M}_0}\frac{\alpha j_e}{2j_c}\right|_{sat} = \frac{4L_y\tilde{M}_0}{L_z}\ln\frac{2\tilde{M}_0^2 L_y}{\gamma}. \quad (32)$$

For example, for $j_e = 0$, the saturation field is given by

$$H_{sat} = \frac{4L_y\tilde{M}_0}{L_z}\ln\frac{2\tilde{M}_0^2 L_y}{\gamma} \quad (33)$$

and for $H = 0$, the saturation current is

$$j_{e,sat} = \frac{8j_c}{\alpha}\frac{L_y}{L_z}\frac{\tilde{M}_0^2}{M_0\mu_B n}\ln\frac{2\tilde{M}_0^2 L_y}{\gamma}. \quad (34)$$

Let us evaluate these parameters. Equation (24) for γ can be represented as $\gamma \sim \delta\tilde{M}_0^2$, where δ is the domain

wall width. According to [39], $\delta = \sqrt{\alpha_{exc}/\beta}$, where α_{exc} is the nonuniform exchange constant (its typical value is $\sim 10^{-12}$ cm^2) and β is the anisotropy dimensionless parameter, which specifies the anisotropy field $H_{an} \approx \beta \tilde{M}_0$. Let us apply our numerical values of the parameters to a Co film for which [40] $M_0 \approx 0.14$ T, $n \approx 10^{22}$ cm^{-3}, $\alpha \approx 2 \times 10^5$, $\beta \approx 4.2$, $\delta \approx 5$ nm, $L_z \approx 50$ μm, and $L_y \approx 50$ nm. Then, we will find that $W_0 \approx 5$ μm, $H_{sat} \approx 16.8$ Oe, and $j_{e,\,sat} \approx 1.5 \times 10^5$ A/cm^2.

These values of W_0 and H_{sat} agree in order of magnitude with experimental data reported in [2]. From the estimate of the saturation current, a very significant fact follows: namely, such a value is quite achievable and is even low relative to those applied in experiments on magnetization switching [9–15] (in those experiments, the current densities were as high as 10^8–10^9 A/cm^2).

Note also that the right-hand side of (30) vanishes at certain values of H and j_e. This means that the effects of the field and current in layer 2 may cancel. Under the cancellation conditions, $\xi = 1/2$ and the DS does not differ (in our model) from that existing in total equilibrium in the absence of the field and current. The cancellation current is proportional to the field H:

$$j_{e,\,canc} = -\frac{2j_c}{\alpha}\left(\frac{1}{M_0 + \dfrac{1}{\mu_B n}}\right)H. \tag{35}$$

For $H = 20$ Oe in the Co film, $j_{e,\,canc}$ is estimated at $\sim 1.6 \times 10^5$ A/cm^2.

Obviously, we can not only compensate the effect of H in layer 2 but also reverse the magnetization of this layer by current. Let a field $H < -H_{sat}$ be applied to layer 2. This field sets the magnetization of this layer antiparallel to that of layer 1. Now, we apply a current, which injects the spins from layer 1 into layer 2. With this current sufficiently high, namely,

$$j_e > 2j_{e,\,sat}, \tag{36}$$

magnetization reversal will take place; that is, the magnetization in all of layer 2 will be aligned with that in layer 1.

DISCUSSION

Thus, we have described the effect of spin-polarized current on the DS in the magnetic junction in terms of the phenomenological theory. The proposal [25] that current-injected spins may rearrange the DS by displacing the domain walls has been substantiated. *s–d* exchange (not magnetodipole) interaction of injected and localized spins plays a dominant part in this process. As for magnetodipole interaction, we showed in the previous section that its role is negligible provided the currents through the junction are not too high ($j_{p,\,a} \ll j_c$ [24]).

The significance of the induced (self-magnetic) field H_{ind} deserves special attention. This field is defined by the Ampere law $\mathrm{curl}\,\mathbf{H}_{ind} = (4\pi/c)\mathbf{j}_{p,\,a}$ and can be estimated in our geometry as

$$|H_{ind}| \sim (4\pi/c)L_z j, \tag{37}$$

where j is the current density averaged over the cross section ($j \sim j_p \sim j_a$) and c is the velocity of light.

The induced field should be compared with the field H_{s-d} of *s–d* exchange interaction. Differentiating the energy E_{s-d} [see (15)] with respect to M_0, we derive an estimate for the nonequilibrium (that is, proportional to the current $j_{p,\,a}$) contribution to the effective *s–d* exchange field:

$$H_{s-d} \sim \left(\frac{A}{\mu_B}\right)\frac{j_{p,\,a}}{j_c}(P_1 \mp P_2), \tag{38}$$

where j_c is given by (18) if $L_y > l_s$.

From (37) and (38), the *s–d* exchange field prevails if

$$\frac{H_{s-d}}{H_{ind}} \sim \left(\frac{A}{\mu_B}(P_1 \mp P_2)/\left(\frac{2\pi}{c}\right)j_c L_z > 1\right). \tag{39}$$

Let $A \approx 0.1$ eV, $(P_1 \mp P_2) \approx 1$, $j_c = enL_y/\tau_s$, $n \sim 10^{22}$ cm^{-3}, $L_y = l_s \equiv \sqrt{D\tau_p} = v_F^2\tau_p\tau_s$, $v_F \sim 10^8$ cm/s, $\tau_p \approx 3 \times 10^{-15}$ s, and $\tau_s \sim 10^{-13}$ s. Then, the ratio of the fields in (39) is close to unity if $L_z = 10$ μm. With the transverse dimension of the junction $L_z \leq 5$–3 μm, condition (39) is satisfied.

To conclude, let us discuss in more detail the reversal of the magnetization in one of the electrodes [see formula (36)]. The junction is exposed to an external magnetic field that sets the magnetization over all of layer 2 antiparallel to that in layer 1 (in the absence of current). The application of the current, which injects the spins from layer 1 into layer 2, first causes the DS to appear. Then, favorable domains grow and eventually the single-domain state with reversed magnetization arises.

Physically, this effect is likely to differ from the switching effect predicted in [4, 6] and discovered experimentally in [10, 11]. In the works cited, the current resulted in unstable orientation of the magnetization vector, which rotated within a domain. In our case, the magnetization reversal is caused by the displacement of the domain walls and the growth of favorable domains. It appears that the threshold current of magnetization reversal in our work is three to four orders of magnitude lower than in [4, 6, 10, 11] because of this difference.

ACKNOWLEDGMENTS

This work was supported by the International Research Center (grant no. 1522) and the Russian Foundation for Basic Research (grant no. 16 384).

REFERENCES

1. R. Meservey and P. M. Tedrow, Phys. Rep. **238**, 173 (1994).
2. X. Portier and A. K. Petford-Long, J. Phys. D **32**, 1 (1999).
3. P. V. Levy and Sh. Zhang, Curr. Opin. Solid State Mater. Sci. **4**, 223 (1999).
4. J. C. Slonczewski, J. Magn. Magn. Mater. **159**, L1 (1996).
5. J. C. Slonczewski, J. Magn. Magn. Mater. **195**, L261 (1999).
6. L. Berger, Phys. Rev. B **54**, 9353 (1996).
7. L. Berger, J. Appl. Phys. **81**, 4880 (1997).
8. L. Berger, IEEE Trans. Magn. **34**, 3837 (1998).
9. M. Tsoi, A. J. M. Jausen, J. Bass, *et al.*, Phys. Rev. Lett. **80**, 4281 (1998); **81**, 493 (1998).
10. M. Tsoi, A. J. M. Jausen, J. Bass, *et al.*, Nature **406**, 46 (2000).
11. E. B. Myers, D. C. Ralph, J. A. Katine, *et al.*, Science **285**, 867 (1999).
12. J. Z. Sun, J. Magn. Magn. Mater. **202**, 157 (1999).
13. J. A. Katine, F. J. Albert, R. A. Burhman, *et al.*, Phys. Rev. Lett. **84**, 3149 (2000).
14. S. M. Resende, F. M. de Aguiar, M. A. Lucena, and A. Azevedo, Phys. Rev. Lett. **84**, 4212 (2000).
15. J. Z. Sun, Phys. Rev. B **62**, 570 (2000).
16. A. G. Aronov, Pis'ma Zh. Éksp. Teor. Fiz. **24**, 37 (1976) [JETP Lett. **24**, 32 (1976)].
17. A. G. Aronov and G. E. Pikus, Fiz. Tekh. Poluprovodn. (Leningrad) **10**, 1177 (1976) [Sov. Phys. Semicond. **10**, 698 (1976)].
18. M. E. Flatte and J. M. Byers, Phys. Rev. Lett. **84**, 4220 (2000).
19. G. Schmidt, D. Ferrand, L. W. Molenkamp, *et al.*, Phys. Rev. B **62**, R4790 (2000).
20. N. F. Schwabe, N. S. Wingreen, and R. J. Elliott, Phys. Rev. B **54**, 12953 (1996).
21. C. Heide and R. J. Elliott, Europhys. Lett. **50**, 271 (2000).
22. C. Heide, P. E. Zilberman, and R. Elliott, Phys. Rev. B **63**, 064424 (2001).
23. C. Heide, A. I. Krikunov, Yu. F. Ogrin, *et al.*, J. Appl. Phys. **87**, 5221 (2000).
24. A. M. Baranov, Yu. V. Gulyaev, P. E. Zil'berman, *et al.*, Radiotekh. Élektron. (Moscow) **46**, 102 (2001).
25. A. M. Baranov, A. I. Chmil, R. J. Elliott, *et al.*, Europhys. Lett. **53** (5), 625 (2001).
26. V. S. Gornakov, V. I. Nikitenko, L. H. Bennett, *et al.*, J. Appl. Phys. **81**, 5215 (1997).
27. S. J. C. H. Theeuwen, J. Caro, K. P. Wellock, *et al.*, Appl. Phys. Lett. **75**, 3677 (1999).
28. H. J. Hardner, M. J. Hurben, and N. Tabat, IEEE Trans. Magn. **35**, 2592 (1999).
29. P. K. Wong, J. E. Evetts, and M. G. Blamire, Phys. Rev. B **62**, 5821 (2000).
30. N. M. Salanskiĭ and M. Sh. Erukhimov, *Physical Properties and Application of Magnetic Films* (Nauka, Novosibirsk, 1975).
31. C. Tiusan, T. Dimopoulos, K. Ounadjela, *et al.*, Phys. Rev. B **61**, 580 (2000).
32. V. D. Buchel'nikov, V. A. Gurevich, and V. G. Shavrov, Fiz. Met. Metalloved. **52**, 298 (1981).
33. Yu. V. Gulyaev, P. E. Zil'berman, R. J. Elliott, and É. M. Épshteĭn, Fiz. Tverd. Tela (St. Petersburg) **44** (6), 1064 (2002) [Phys. Solid State **44**, 1111 (2002)].
34. M. Julliere, Phys. Rev. Lett. **54**, 225 (1975).
35. J. C. Slonczewski, Phys. Rev. B **39**, 6995 (1989).
36. Y. Utsumi, Y. Shimizu, and H. Miyazaki, J. Phys. Soc. Jpn. **68**, 3444 (1999).
37. É. L. Nagaev, *Physics of Magnetic Semiconductors* (Nauka, Moscow, 1979).
38. A. P. Prudnikov, Yu. A. Brychkov, and O. I. Marichev, *Integrals and Series* (Nauka, Moscow, 1981; Gordon and Breach, New York, 1986), Vol. 1.
39. L. D. Landau and E. M. Lifshitz, Phys. Z. Sowjetunion **8**, 153 (1935); L. D. Landau, *Collection of Works* (Nauka, Moscow, 1969), Vol. 1, p. 129.
40. I. A. Privorotskiĭ, Usp. Fiz. Nauk **108**, 43 (1972) [Sov. Phys. Usp. **15**, 555 (1972)].

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 7, 2002, pp. 880–885. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 86–91.

SOLIDS

On the Determination of the Energy of Activation of Relaxation Transitions in Polymers by Differential Scanning Calorimetry

A. I. Slutsker*, Yu. I. Polikarpov, and K. V. Vasil'eva****

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

e-mail: Alexander.Slutsker@pop.ioffe.rssi.ru

*** St. Petersburg State Technical University, ul. Politekhnicheskaya 29, St. Petersburg, 195251 Russia*

Received January 8, 2002

Abstract—The determination of the energy of activation (barrier height) of elementary events involved in relaxation (fluctuation) transitions in polymers from the temperature dependence of the specific heat is discussed. The dependence is derived by the method of differential scanning calorimetry. It is emphasized that the correct determination of the energy of activation must include the temperature variation of the barrier. The deviation of the preexponential in the Arrhenius temperature dependence from the value predicted theoretically demonstrates that the barrier does depend on temperature. Experimental data from which the realistic energy of activation of the α relaxation in polymers can be found are given. *© 2002 MAIK "Nauka/Interperiodica".*

As the temperature of polymers grows from low to high values, different forms of molecular mobility (relaxation transitions) successively come into play [1]. Among these transitions, so-called β and α transitions are the most typical [2, 3]. The former involve the motion of relatively short segments of chain molecules. β relaxation proceeds most effectively at an associated temperature T_β. α transitions imply the motion of longer segments (or even molecules) in cooperation with the motion of segments of adjacent molecules. The onset of these transition is characterized by a temperature T_α, usually called the vitrification temperature. In the context of this work, the term "devitrification temperature" would be more adequate, since a polymer passes from the glassy state to the high-elasticity state (as the temperature grows) near T_α [4]. It can be said that β relaxation opens the way to α relaxation [3]. Both transitions take place owing to energy fluctuations, which help to overcome the barriers U_β and U_α to elementary transition events. The energy of activation, the commonly accepted characteristic of relaxation processes, is the value of these barriers. Knowledge of the energy of activation is of great importance for constructing a model of process elementary events. Therefore, the correct experimental determination of the energies of activation for β and α transitions is an important physical problem.

Among the many methods for studying relaxation processes (mechanical, dielectric, and acoustic spectroscopies; X-ray diffraction analysis; etc.), differential scanning calorimetry (DSC) occupies a prominent place. The potentialities of DSC for the study of relaxation processes in polymers were comprehensively described in monograph [3]. DSC information on the temperature dependence of the specific heat at constant pressure $C_p(T)$ allows the identification and characterization of the relaxation processes. For example, the onset of various (previously "frozen") forms of molecular mobility in polymers shows up in a nonmonotonic run of the curve $C_p(T)$ (Fig. 1).

The relaxation nature of the α and β transitions appears most vividly in the temperature position depending on the rate of change of temperature $V_T = dT/dt$ (the scan rate). The temperatures T_β and T_α grow with V_T (Fig. 1) [3]. From the dependences $T_\beta(V_T)$ and

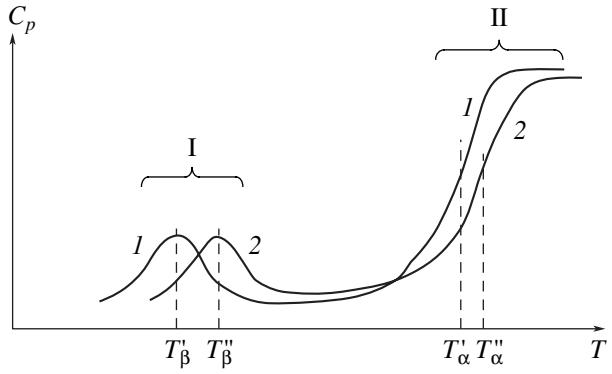


Fig. 1. Schematic depiction of the temperature dependence of the polymer specific heat at constant pressure. I, range of β relaxation; II, range of α relaxation; *1*, scan rate V'_T; and *2*, scan rate $V''_T > V'_T$.

1063-7842/02/4707-0880$22.00 © 2002 MAIK "Nauka/Interperiodica"

$T_\alpha(V_T)$, the energies of activation of the transitions are determined [3]. These dependences are usually plotted in the form $\ln V_T(1/T_\beta)$ and $\ln V_T(1/T_\alpha)$. The slopes of the resulting straight lines yield the value of $U_f = -k(\Delta \ln V_T/\Delta(1/T_i))$, which is frequently called the energy of activation (here, $k = 8.6 \times 10^{-5}$ eV/K is the Boltzmann constant).

Since the range of V_T is usually small (about one order of magnitude), the plots $\ln V_T(1/T_i)$ are nearly linear, which argues for the use of this method in finding the energy of activation.

However, the energy of activation of the α relaxation thus obtained turns out to be exceedingly too high (3–5 eV or even higher [3]) in comparison with the energy characteristics of molecular interaction. The reason for such high values of U_f has long been discussed, and various explanations for this fact have been provided [3, 5]. In particular, it is speculated that the value of U_f (energy of activation) extracted from the slope of the plot $\ln V_T(1/T_\alpha)$ does not have the sense of a barrier to an elementary event of α relaxation. It was suggested to call the parameter U_f the "apparent" energy of activation. Naturally, it is proposed to find the actual energy of activation of polymer devitrification from DSC data.

In this work, we discuss a possible reason for unrealistically high values of U_f and substantiate a technique that provides true values of energy of activation (transition barrier) from the dependence $\ln V_T(1/T_\alpha)$.

PHYSICAL GROUNDS

The theory [6] sets the correlation between the transition temperature T_i and the (constant) rate of rise of temperature V_T:

$$V_T\left(\frac{d\tau}{dT}\right)_{T_i} = -1, \tag{1}$$

where τ is the relaxation time or the mean expectation time for a transition elementary event.

The transitions occur owing to local energy fluctuations when the energy fluctuation $E_{fl} \geq U$, where U is the transition barrier. The mean expectation time for a local energy fluctuation E_{fl} (per atom or per small group of atoms) is given by the fundamental expression [7]

$$\tau_{fl} \cong \tau_0 \exp\frac{E_{fl}}{kT},$$

where $\tau_0 = 10^{-13}$–10^{-12} s is the period of atom oscillation in condensed media.

This expression for energy fluctuation has been supported by direct computer simulation of the atom dynamics in solids [8]. Sharp localization of the fluctuations in time ($\sim\tau_0$) and in space (atom-size range) has been shown [9]. Then, for τ, we get

$$\tau \cong \tau_{fl} \cong \tau_0 \exp\frac{U}{kT}. \tag{2}$$

Expressions (2) and (1) establish a correlation between the transition temperature T_i, transition barrier U_i, and rate of change of temperature V_T in the DSC method.

(i) Consider first the case where the barrier is temperature-independent; that is, U = const. Then,

$$\left(\frac{d\tau}{dT}\right)_{T_i} = -\frac{\tau_0 U_i}{kT_i^2}\exp\left(\frac{U_i}{kT_i}\right). \tag{3}$$

Substituting (3) into (1) and taking the logarithm yields

$$\ln V_T = \ln\left(\frac{kT_i^2}{\tau_0 U_i}\right) - \frac{U_i}{k}\frac{1}{T_i}. \tag{4}$$

The first term on the right of (4) is logarithmic and slowly varies with the parameters entering into it. This is especially true in view of the fact that T_i actually varies by as little as several degrees (typical values of T_i lie between 300 and 400 K) for the range of the scan rate V_T used [3]. The actual barrier values U_i are close to 1 eV [1, 3]. Then, putting $\tau_0 = 10^{-13}$–10^{-12} s, $U_i = 0.5$–1.0 eV, and $T_i = 300$–400 K and designating $kT_i^2/\tau_0 U_i = V_0$, we find

$$V_0 \cong 10^{13}\text{–}10^{14}\text{K/s} \tag{5}$$

and, from (1) and (3),

$$V_T \cong V_0 \exp\left(-\frac{U_i}{kT_i}\right). \tag{6}$$

Thus, in view of (4) and (5), the plot of the dependence $\ln V_T(1/T_i)$ must be a straight line, which is extrapolated to $\ln V_0 \cong 30$–32 (or $\log V_0 = 13$–14) at $1/T_i = 0$ (V_0 is expressed in K/s) (Fig. 2, curve *1*). The slope of this line,

$$\frac{\Delta \ln V_T}{\Delta\left(\frac{1}{T_i}\right)} = -\frac{U_i}{k},$$

defines the energy of activation (transition barrier height) U_i.

We stress that the mere linearity of the dependence $\ln V_T(1/T_i)$ does not yet mean that the slope of the dependence $\ln V_T(1/T)$ has the significance of the energy of activation (as a rule, the linearity or quasi-linearity takes place because of the small interval of $\Delta \ln V_T$ and $\Delta(1/T_i)$, as was noted above). It is also necessary that the value of V_0 extrapolated to $1/T_i = 0$ equal 10^{13}–10^{14} K/s (certainly with a logarithmic accuracy). The

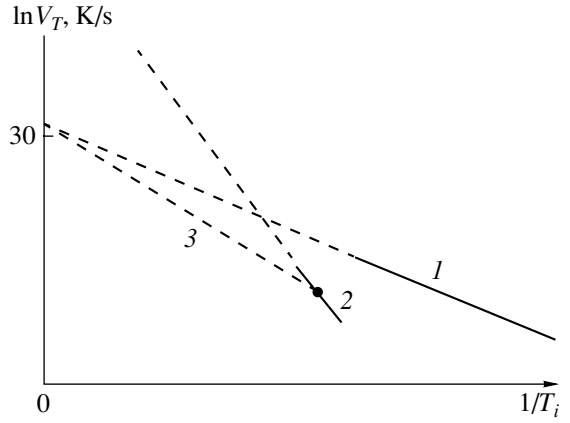


Fig. 2. Schematic depiction of dependences $\ln V_T(1/T_i)$. (*1*) Barrier U = const and (*2*) $U = U(T)$.

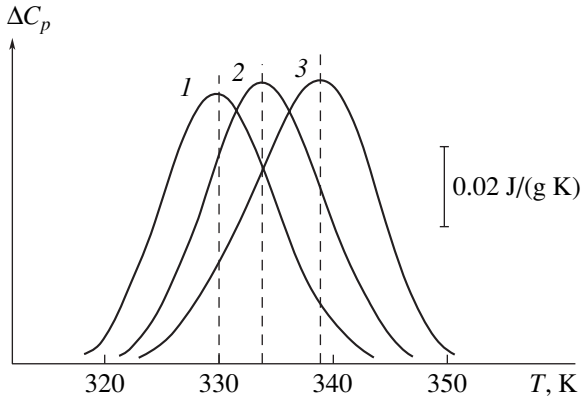


Fig. 3. β relaxation in polystyrene at V_T = 0.083 (*1*), 0.166 (*2*), and 0.333 K/s (*3*) [3].

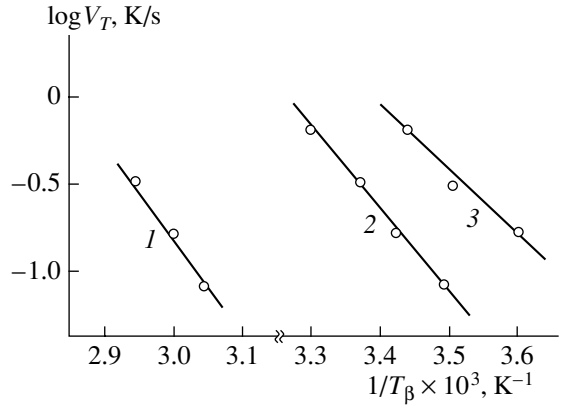


Fig. 4. Temperature of the β transition vs. scan rate: (*1*) polystyrene (from Fig. 3), (*2*) isopropylbenzene [3], and (*3*) polydimethylsiloxane [3].

fulfillment of this condition indicates that the energy of activation is determined correctly. Otherwise (the value of V_0 differs from that predicted theoretically), the problem of finding the energy of activation becomes more complicated (even if the experimental dependence $V_T(1/T_i)$ is linear).

(ii) Let the barrier be temperature-dependent; that is, $U = U(T) \neq$ const. In this case,

$$\tau = \tau_0 \exp\frac{U(T)}{kT},$$

where the value of τ_0 is the same as indicated above.

The validity of such an expression for $\tau(T)$ at $U(T) \neq$ const was justified in [5, 10].

Taking into account the weak dependence of the preexponential V_0 in (6) on the parameters entering into it, we can write the expression for V_T at $U(T) \neq$ const as

$$V_T \cong V_0 \exp\left(-\frac{U_i(T)}{kT_i}\right), \tag{7}$$

where V_0 remains constant and has its previous value, 10^{13}–10^{14} K/s. Then, the slope of the dependence $\ln V_T(1/T_i)$ will have the form

$$\frac{d\ln V_T}{d\left(\frac{1}{T_i}\right)} = -\frac{U_i(T_i)}{k} + \frac{T_i}{k}\left(\frac{dU_i(T)}{dT}\right)_{T_i}; \tag{8}$$

that is, the slope of $\ln V_T(1/T_i)$ is specified not only by the barrier U_i but also by the derivative of the barrier with respect to temperature. This important point was indicated in [5].

If the derivative

$$U_i'(T_i) = \left(\frac{dU_i(T)}{dT}\right)_{T_i}$$

is negative (i.e., the barrier decreases with increasing temperature), which is typical of polymers [11], the slope of $\ln V_T(1/T_i)$ exceeds U_i/k. Clearly, the value of "the energy of activation" derived from this slope will exceed the barrier height. The quasi-linear portion of $\ln V_T(1/T_i)$ for this case is schematically depicted in Fig. 2 (curve *2*).

The linear extrapolation of the dependence $\ln V_T(1/T_i)$ to $1/T_i = 0$ yields

$$\ln V_{0f} = \ln V_0 + \frac{U_i'(T_i)}{k}, \tag{9}$$

i.e., a value greater than V_0.

Thus, if the value of $\ln V_{0f}$ obtained by extrapolating the experimental dependence is much higher than $\ln V_0 = 30$–32 (V_0 in K/s), one can conclude that the energy of activation varies with temperature. From the difference $\ln V_{0f} - \ln V_0$, the derivative of the barrier height U_i' with respect to temperature at a temperature T_i can be estimated.

The very energy of activation (barrier height) should apparently be determined according to (7):

$$U_i(T_i) = kT_i \ln \frac{V_0}{V_T(T_i)}. \quad (10)$$

This expression corresponds to the slope of the straight line connecting the points $\ln V_T(1/T_i)$ and $\ln V_0$ (at $1/T = 0$) (Fig. 2, curve *3*).

EXPERIMENTAL RESULTS

Here, we use the DSC data for polymers reported in [3]. An example of the experimental dependences $C_p(T)$ at various V_T for the β transition in polystyrene is shown in Fig. 3. The peaks of $C_p(T)$ correspond to the temperatures T_β of the β transition. From Fig. 3, one can find the dependence $\log V_T(1/T_\beta)$ (Fig. 4, curve *1*). Figure 3 also demonstrates the dependences $\log V_T(1/T_\beta)$ for the β transition in two other polymers.

It should be emphasized that the nearly linear portions of $\log V_T(1/T_\beta)$ shown in Fig. 4 are extrapolated (at $1/T_\beta = 0$) to values of 13–16; i.e., they are in satisfactory agreement with the theoretical value $\log V_0 =$ 13–14 ($V_0 = 10^{13}$–10^{14} K/s).

Thus, the dependence of the temperature T_β of the β transition on the rate of change of the temperature V_T fits expressions (4) and (6). The slopes of $\log V_T(1/T_\beta)$ in this case give the energies of activation (barrier height) U_β of β-transition elementary events. These energies are listed in the table, from which (and also from [3]) it follows that U_β fall into the interval 0.5–1.0 eV. Such values appear to be quite reasonable for the motion of short segments of chain molecules.

Note that an elementary event of the β transition involves ten or more atoms. Therefore, if the β transition obeys theoretical dependence (6), corresponding to expression (2), this means that an elementary event is due to one energy fluctuation, which activates a group of atoms participating in the β transition. The entropy factor is of minor significance here, since the preexponential in (6) is close to the fundamental value $\tau_0 = 10^{-13}$ s.

Now, consider data for α transitions. The dependence $C_p(T)$ for a typical amorphous polymer, polymethyl methacrylate, near the α transition (devitrification) is shown in Fig. 5 for various scan rates. It is assumed that the half-height of the step in the $C_p(T)$ curve corresponds to the transition temperature T_α [3]. From Fig. 5, we find the dependence $\log V_T(1/T_\alpha)$, which is plotted in Fig. 6 (curve *1*). Curves *2* and *3* in Fig. 6 depict the same plot for two other polymers.

The dependences $\log V_T(1/T_\alpha)$ in Fig. 6 are similar to those for the β transitions in Fig. 4 in that they are

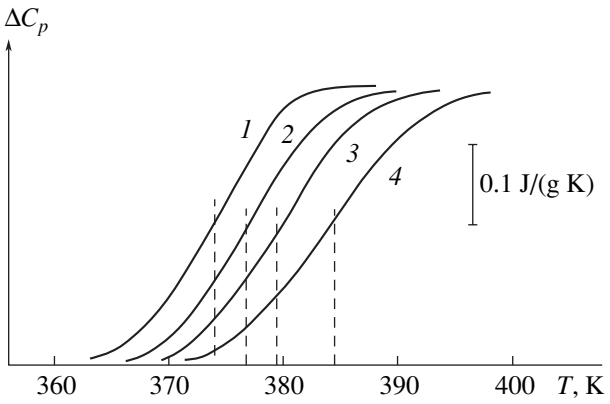


Fig. 5. α relaxation in polymethyl methacrylate for $V_T =$ 0.042 (*1*), 0.166 (*2*), 0.333 (*3*), and (*4*) 0.666 K/s [3].

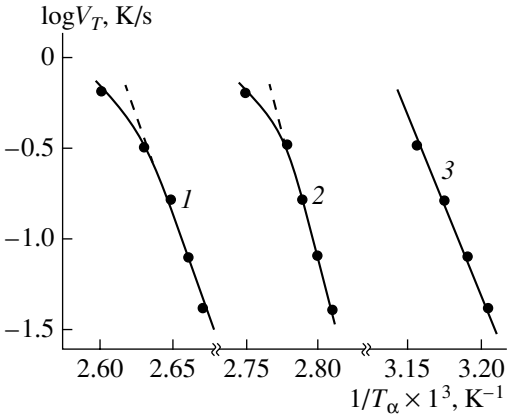


Fig. 6. Temperature of the [alpha] transition vs. scan rate: (*1*) polymethyl methacrylate (according to Fig. 5), (*2*) polyvinylchloride [3], and (*3*) polyvinylacetate [3].

largely quasi-linear. At the same time, the upper (high-temperature) parts of curves *1* and *2* in Fig. 6 deviate markedly from a straight line. Such behavior is typical of α relaxation, where the nonlinearity increases with temperature (see below).

The slopes of $\log V_T(1/T_\alpha)$ are several times larger than for the β transitions. However, the most important feature is that the linear extrapolation of $\log V_T(1/T_\alpha)$ to $1/T = 0$ yields $\log V_0 = 60$–70, i.e., values far exceeding 13–14. Such a large discrepancy between the extrapolated and theoretical values of V_0 suggests that the parameter

$$U_f = -k\frac{\Delta \ln V_T}{\Delta(1/T_\alpha)}$$

does not have the meaning of the energy of activation in this case and is much higher than the true energy of activation U_α.

Table

Polymer	U_β, eV	$U_\alpha(T_\alpha)$, eV	$U_f(T_\alpha)$, eV	$U'_\alpha(T_\alpha)$, eV/K
Polystyrene	1.1			
Isopropylbenzene	0.95			
Polydimethylsiloxane	0.7			
Polymethyl methacrylate	0.9	1.2	5.8	-1.3×10^{-2}
Polyvinylchloride	0.7	1.1	6.0	-1.4×10^{-2}
Polyvinylacetate	0.6	1.0	4.0	-0.9×10^{-2}

According to (7) and (10), U_α is found from the expression

$$U_\alpha(T_\alpha) = kT_\alpha[\ln V_0 - \ln V_T(T_\alpha)]. \quad (11)$$

The validity of (11), as applied to analysis of the α transitions, is strengthened by the following considerations. An elementary event of the α transition obviously involves a larger number of atoms than in the β transition. It has been indicated, however, that the number of atoms participating in the latter is also appreciable (ten or more). Nevertheless, the β transition is associated with a single local energy fluctuation $E_{fl} \cong U_\beta$, as follows from the fact that expression (6) is valid for this transition. When the number of participating atoms increases (several times), i.e., when we are dealing with the cooperative α transition, it appears that there are no reasons to discard the same single-fluctuation mechanism of an elementary α transition event. We believe, therefore, that one can determine the energy of activation (barrier height) of the α transition from expression (11), rather than from the slope of $\log V_T(1/T_\alpha)$ as many authors do without regard for the value of the preexponential.

To analyze DSC data with expression (11), it is necessary to take a value of $\log V_T$ from the $\log V_T$ interval measured (e.g., the central value of the interval) and the associated value of T_α. Note that the selection of any other value from this interval will affect the value of U_α determined insignificantly, because the $\log V_T$ and $1/T_\alpha$ intervals are narrow (Fig. 6).

The values of U_α thus obtained are listed in the table, where the "barriers U_f" found from the slope of $\log V_T(1/T_\alpha)$ are shown for comparison.

It is seen that U_α much exceed U_β. This seems to be natural because the α transition involves a larger number of atoms than the β transition. At the same time, U_α is considerable lower than U_f.

As was noted above, the reason why the true energy of activation exceeds U_f may be the temperature dependence of the barrier height of the α transition with the derivative

$$U'_\alpha(T_\alpha) = \left(\frac{dU_\alpha}{dT}\right)_{T_\alpha}.$$

Then, from (8),

$$U_f(T_\alpha) = U_\alpha(T_\alpha) - T_\alpha U'_\alpha(T_\alpha). \quad (12)$$

The values of $U'_\alpha(T_\alpha)$ found from the tabulated data with (12) are also summarized in the table. They are close to values that take into account the temperature dependence of barriers in the range of devitrification (α relaxation). These values were obtained by the methods of dielectric spectroscopy and mechanical spectroscopy, as well as by measuring the viscosity [5, 12]. Unlike DSC, these methods feature wide rate (V) and frequency (ν) ranges (up to 10–12 orders of magnitude). In this case, the curves $\log V(1/T)$ or $\log \nu(1/T)$ become essentially nonlinear, which directly indicates that the barriers to the α transition vary with temperature. It should be noted that for the β transitions, the dependences $\log V(1/T)$ and $\log \nu(1/T)$ remain linear in the same wide ranges of V and ν and are extrapolated to their theoretical values for $1/T = 0$ [12].

The temperature dependence of the barriers near the point of polymer devitrification (α transition) is explained by the extensive molecular rearrangement and the change in the molecular dynamics (mobility) at this point [4, 11].

With expression (10) used to determine the barriers to both the α and β transitions, U_α and U_β, we obtain, for the same scan rate V_T,

$$U_\alpha(T_\alpha) = \left(\frac{T_\alpha}{T_\beta}\right)_{V_T} U_\beta. \quad (13)$$

Obviously, the ratio T_α/T_β does not remain unchanged when V_T is varied over a wide range. For V_T varying between 5×10^{-2} and 1 K/s (which is typical of DSC), this ratio is roughly 1.5 for many polymers [3]. As V_T grows, $T_\alpha/T_\beta \longrightarrow 1$; as V_T decreases, $T_\alpha/T_\beta \longrightarrow \infty$. These limits for the temperature ratio (and, accordingly, for the barrier ratio) have been found by other methods covering a wide range of rates and frequencies [3].

Thus, for scan rates V_T typical of DSC, the ratio between the energies of activation of the α and β transitions (≈1.5) is not as large as the ratio U_f/U_β found directly from the slope $\log V_T(1/T_\alpha)$ (U_f/U_β = 4–5 for many polymers [3]). It seems that treating U_f as the true energy of activation is not quite correct, since the true energy of activation of the α transition is lower.

At the same time, the idea that the barrier to the α relaxation varies with temperature and the ability of DSC to detect this variation (by determining the derivative of the barrier with respect to temperature) favor a deeper insight into the molecular dynamics in polymers at the point of devitrification.

To conclude, we emphasize the need for taking into account the value of the preexponential in the analysis of temperature and rate dependences obtained with the differential scanning calorimetry method. Such an approach provides correct determination of the energy of activation of relaxation transitions in polymers.

REFERENCES

1. Yu. K. Godovskiĭ, *Thermal Physics of Polymers* (Khimiya, Moscow, 1982).
2. *Transitions and Relaxations in Polymers*, Ed. by R. Boyer (Interscience, Easton, 1966; Mir, Moscow, 1968).
3. V. A. Bershteĭn and V. M. Egorov, *Differential Scanning Calorimetry in Physics and Chemistry of Polymers* (Khimiya, Leningrad, 1990).
4. G. M. Bartenev and V. G. Nikol'skiĭ, *Encyclopedia of Polymers* (Sov. Éntsiklopediya, Moscow, 1977), Vol. 3, pp. 489–498.
5. P. P. Kobeko, *Amorphous Materials* (Akad. Nauk SSSR, Moscow, 1952).
6. M. V. Vol'kenshteĭn and O. B. Ptitsyn, Zh. Tekh. Fiz. **26**, 2204 (1956).
7. Ya. I. Frenkel', *Kinetic Theory of Liquids* (Nauka, Moscow, 1975).
8. A. I. Slutsker, A. I. Mikhaĭlin, and I. A. Slutsker, Usp. Fiz. Nauk **164**, 357 (1994) [Phys. Usp. **37**, 335 (1994)].
9. A. I. Slutsker, A. I. Mikhaĭlin, and I. A. Slutsker, in *Problems of Theoretical Physics: Collection of Articles* (Peterb. Inst. Yad. Fiz., St. Petersburg, 1994), pp. 42–69.
10. Yu. S. Lazurkin, *Encyclopedia of Polymers* (Sov. Éntsiklopediya, Moscow, 1972), Vol. 1, pp. 62–70.
11. J. D. Ferry, *Viscoelastic Properties of Polymers* (Wiley, New York, 1961; Mir, Moscow, 1963).
12. P. F. Veselovskiĭ and A. I. Slutsker, Zh. Tekh. Fiz. **25**, 1204 (1955).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 7, 2002, pp. 886–888. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 92–94.

**OPTICS,
QUANTUM ELECTRONICS**

On 2D Focusing of X Rays Bragg-Diffracted by a Biaxially Bent Crystal

T. Tchen

Lomonosov State Academy of Fine Chemical Technology, pr. Vernadskogo 86, Moscow, 117571 Russia
e-mail: docent65@mtu-net.ru, ttchen@e-mail.ru

Received November 30, 2001

Abstract—2D focusing of an X-ray wave Bragg-diffracted by a biaxially bent crystal is considered in terms of quasi-back diffraction scattering in the meridional plane. Analytical expressions for the focusing geometrical condition in the meridional plane and for the spatial distribution of the wave intensity in the same plane in the neighborhood of the point source image are derived. They differ from those currently available from the theory.

Bent crystals are widely used as focusing monochromators, spectrometers, and collimators incorporated into X-ray optics. The collection solid angle of X-ray beams is the highest if a biaxially bent crystal, providing point focusing, is used. The fundamental possibility of focusing X rays when a plane wave is Bragg-diffracted by a biaxially bent crystal was demonstrated in [1]. In [2], the rigorous theory of 2D focusing of a spherical wave Bragg-diffracted by a perfect biaxially bent crystal was elaborated. In [3], the theory of 2D diffraction focusing was extended for the case of Johann–Hamos X-ray spectrometers. In this work, we consider the focusing of a spherical wave by a biaxially bent crystal with a geometry other than that used in [2, 3].

By analogy with [2], the plane of diffraction scattering where the Bragg angle $\theta_B \neq \pi/2$ and the crystal is bent with a bending radius R_x will be called the sagittal plane (see the figure). The plane normal to it will be called meridional (in this plane, the bending radius of the crystal is R_y). We use the parabolic expansion of the phases of the incident (spherical) and diffracted waves. The infinite intensity arising in the Johann scheme [2] in this case is removed by including terms $\sim x^3$ in the expressions for the wave phase. Then, for the Johann scheme, the intensity distribution in the neighborhood of the focus is defined by the Airy function magnitude squared [4].

According to [2], the amplitude of the wave diffracted in a vacuum at a point $\mathbf{r}_p(\xi_p, y_p)$ is proportional to the product of two integrals: the former is the integral over plane harmonics in the sagittal plane, and the latter is the integral along the coordinate y in the meridional plane:

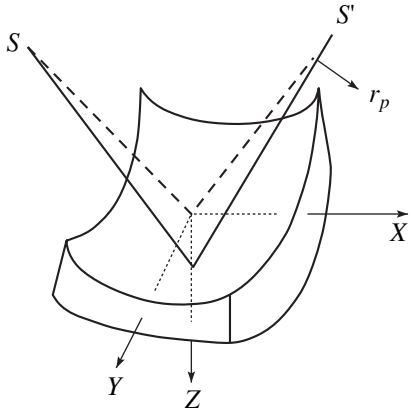
$$E_h(\mathbf{r}_p) = E_h(\xi_p)E_h(y_p), \quad (1)$$

where

$$E_h(\xi_p) = \int_{-\infty}^{+\infty} dk G_h(k + q_0) \times \exp[-ik^2(1/\alpha_0 + 1/\alpha_h)/2\kappa + ik\gamma_h\xi_p/\alpha_h L_h], \quad (2)$$

$$E_h(y_p) = \int_{-y_{\text{eff}}/2}^{+y_{\text{eff}}/2} dy \exp[-i\kappa y^2 \times (1/L_0 + 1/L_h - (\gamma_0 - \gamma_h)/R_y) - i\kappa y y_p/L_h]. \quad (3)$$

Here, $G_h(k + q_0)$ is the Fourier component of the Green function, q_0 is a quantity dependent on the Fourier component of the X-ray polarizability and taking into account refraction (for the expression for χ_0, see [2]), $\alpha_0 = \gamma_0^2/L_0 - \gamma_0/R_x$, $\alpha_h = \gamma_h^2/L_h + \gamma_h/R_x$, $\gamma_{0,h}$ are the direction cosines for the incident and diffracted waves, $L_{0,h}$



2D focusing of a spherical X-ray wave by a biaxially bent crystal. *S*, point source; *S'*, source image; and $\mathbf{r}_p$ is the radius vector of the point of observation.

1063-7842/02/4707-0886$22.00

are the distances from the spherical wave point source to the crystal and from the crystal to the source image, $\kappa = 2\pi/\lambda$, λ is the incident radiation wavelength, $y_{\text{eff}} \leq R_y|\chi_{hr}|^{1/2}$, and χ_{hr} is the Fourier component of the X-ray polarizability. In addition, in (2) and (3), we put $\xi_S = y_S = 0$ for definiteness.

The distance y_{eff}, introduced in [5], is the distance within which the diffracted beam still meets the condition of total Bragg reflection:

$$y_{\text{eff}} = \{L_0 2\Delta\theta\cos\theta/(|\sin\theta - L_0(1 + \sin^2\theta)/R_y|)\}^{1/2}, \quad (4)$$

where $\Delta\theta = |\chi_{hr}|/\sin 2\theta$ is the half-width of the Bragg reflection curve for $\theta < \pi/2$.

It is seen that y_{eff} defined by formula (4) differs from y_{eff} used in [2], but the typical square-root dependence of the X-ray polarizability on the Fourier component is retained.

From (2) and (3), one finds the geometrical condition for focusing in the sagittal and meridional plane with the method of stationary phase:

$$\begin{aligned} \gamma_0^2/L_0 + \gamma_h^2/L_h &= (\gamma_0 - \gamma h)/R_x, \\ 1/L_0 + 1/L_h &= (\gamma_0 - \gamma_h)/R_y. \end{aligned} \quad (5)$$

From (5), we see that point-to-point focusing takes place when R_x, R_y, and L_0 are rigidly related:

$$\begin{aligned} R_y &= R_x L_0 \\ &\times (\gamma_0 - \gamma_h)\gamma_h^2/\{(1 - \gamma_0^2/\gamma_h^2)R_x\gamma_h^2 + L_0(\gamma_0 - \gamma_h)\}. \end{aligned} \quad (6)$$

The condition for stigmatic focusing, which follows from (6) in the symmetrical case, has the form $R_y = R_x\sin^2\theta_B$. Conventional geometrical optics treats Eqs. (5) as conditions under which Fresnel diffraction becomes Fraunhofer diffraction.

Our approach to 2D focusing is as follows. In the two-wave approximation of the problem posed, the diffracted wave in the meridional plane can be considered to be a wave quasi-backscattered in this plane. The intensity distribution in the meridional plane is then described by an expression like (2), where $L_0\gamma_0$ and $L_h|\gamma_h|$ should be substituted for L_0 and L_h, respectively. The corrected condition for focusing in the meridional plane takes the form

$$1/L_0\gamma_0 + 1/L_h|\gamma_h| = 2/R_y. \quad (7)$$

From the first equation in (5) and Eq. (7), it follows that the astigmatism is absent if the bending radii of the crystal and the crystal–source distance are related as

$$R_y = 2R_xL_0\gamma_0|\gamma_h|^3/\{(\gamma_0 - \gamma_h)\gamma_0 L_0 + (|\gamma_h|^3 - \gamma_0^3)R_x\}. \quad (8)$$

It is easy to see that formulas (6) and (8) coincide only if the diffraction geometry is symmetric. In this case, the crystal surface is biaxially bent in the form of a torus and $R_y = R_x\sin^2\theta$.

In the asymmetric case, where the relationship between $R_{x,y}$, L_0, and θ_B is arbitrary, the beam being focused has an astigmatism and the diffraction image of the point source is a spot. From (8), it follows that the focusing of a plane wave to a point takes place if the bending radii meet the condition

$$R_y = 2R_x|\gamma_h|^3/(\gamma_0 - \gamma_h). \quad (9)$$

The intensity distribution of the diffracted wave in the neighborhood of the image point (symmetric diffraction) is given by

$$\begin{aligned} I_h(\mathbf{r}_p) &= I_h(\xi_p)I_h(y_p) \\ &\sim |J_1(t_\xi)/t_\xi|^2|J_1(t_y)/t_y|^2\Theta(t_\xi)\Theta(t_y), \end{aligned} \quad (10)$$

where $J_1(t)$ is the first-order Bessel function of the real argument,

$$t_\xi = 2\xi_p\gamma_h\sigma_h/\alpha_h L_h, \quad \sigma_h = \kappa\chi_h/4\cos\theta_B,$$

$$t_y = \pi y_p|\chi_{hr}|^{1/2}/(\lambda|1 - L_h/R_x|),$$

and $\Theta(t)$ is the Heaviside step function.

The diffraction broadenings of the focus in the sagittal and meridional planes are derived from (10) if $t_{\xi,y} \approx \pi$:

$$\begin{aligned} \Delta\xi_p &= \lambda\sin 2\theta_B|1 - L/R_x\sin\theta_B|/(2|\chi_{hr}|), \\ \Delta y_p &\cong \lambda|1 - L_h/R_x|/|\chi_{hr}|^{1/2} \\ &\cong \lambda\left|\{1/|\chi_{hr}|^{1/2} - L_h/y_{\text{eff}}\}\right|. \end{aligned} \quad (11)$$

Formulas (11) are valid for $L_h \neq R_x\sin\theta_B$.

The second term in our formula for Δy_p is the focus size along the Y axis in [2].

Turning to formula (10), we note that in the parabolic approximation for the phase of the incident spherical and diffracted waves, the position of the geometrical focus (point source image) coincides with the maximum of intensity (10).

With the phase terms $\sim x^3$ taken into account, the intensity spatial distribution (in either plane) will be specified by the Airy function. If the phase is expanded up to terms $\sim x^4$, the intensity distribution is easy to obtain in the Pearsey integral squared. In both cases, the intensity maximum and the geometrical focus, which is located at the diffracted beam axis, are displaced relative to each other.

For back scattering ($\theta_B \approx \pi/2$), it follows from (11) that the focus size in both planes are the same (if it is taken into account that $\cos\pi/2 \leq \Delta\theta = |\chi_{hr}|^{1/2}$: $\Delta\xi_p = \Delta y_p$). In this case, set (5) of two equations, which specifies the geometrical condition for 2D (point) focusing, is reduced to one equation, in which one should put $\gamma_0 = |\gamma_h| = 1$. For back scattering, point-to-point focusing takes place for a spherically bent crystal ($R_x = R_y$).

If back Bragg reflection is symmetric, the problem is symmetric under rotation about the Z axis.

If the beam divergence in the sagittal plane is small and in the meridional one is considerable (synchrotron radiation), the geometrical condition for 2D focusing is as follows:

$$\begin{aligned} \gamma_h^2/L_{h(x)} &= (\gamma_0 - \gamma_h)/R_x, \\ 1/(L_{0(y)}\gamma_0) + 1/(L_{h(y)}|\gamma_h|) &= 2/R_y. \end{aligned} \tag{12}$$

Here, $L_0(y)$ is the distance from the X-ray wave source to the crystal in the meridional plane and $L_{h(x)}$ and $L_{h(y)}$ are the distances between the crystal and the point source image in the sagittal and meridional planes, respectively. From (12), it follows that 2D focusing of synchrotron radiation with the formation of a point focus is possible if

$$R_y = 2R_x L_{0(y)}\gamma_0|\gamma_h|^3\{(\gamma_0 - \gamma_h)L_{0(y)}\gamma_0 + R_x|\gamma_h|^3\}. \tag{13}$$

When synchrotron radiation is incident on the crystal, the intensity distribution near the focus is given by (10) and (12).

To conclude, formulas (7)–(13) can be used for calculations in real physical experiments. In this respect, the focusing of synchrotron radiation is of special importance in view of the new-generation high-power synchrotron sources recently commissioned.

REFERENCES

1. K. T. Gabrielyan, D. I. Piskunov, F. N. Chukhovskiĭ, and G. O. Demirchyan, Pis'ma Zh. Éksp. Teor. Fiz. **46**, 411 (1987) [JETP Lett. **46**, 517 (1987)].
2. K. T. Gabrielyan, F. N. Chukhovskiĭ, and D. I. Piskunov, Zh. Éksp. Teor. Fiz. **96**, 834 (1989) [Sov. Phys. JETP **69**, 474 (1989)].
3. K. T. Gabrielyan, G. O. Demirchyan, and F. N. Chukhovskiĭ, Zh. Tekh. Fiz. **60** (1), 170 (1990) [Sov. Phys. Tech. Phys. **35**, 102 (1990)].
4. T. Tchen, Pis'ma Zh. Tekh. Fiz. **27** (21), 1 (2001) [Tech. Phys. Lett. **27**, 889 (2001)].
5. T. Tchen, V. A. Bushuev, and R. N. Kuz'min, Zh. Tekh. Fiz. **60** (10), 60 (1990) [Sov. Phys. Tech. Phys. **35**, 1148 (1990)].

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 7, 2002, pp. 889–893. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 95–99.

OPTICS, QUANTUM ELECTRONICS

Nonlinear Optical Parameters of Colloidal Silver at Various Stages of Aggregation

R. A. Ganeev*, A. I. Ryasnyanskiĭ, Sh. R. Kamalov*, M. K. Kodirov**, and T. Usmanov***

** Akadempribor Research and Production Association, Tashkent, 700143 Uzbekistan*

*** Navoi State University, Samarkand, 703004 Uzbekistan*

e-mail: ryasn2000@yahoo.com

Received September 12, 2001

Abstract—The *Z*-scan method is used to study the variation in the nonlinear susceptibility $\chi^{(3)}$ (–ω; ω, ω, –ω) and nonlinear refractive index of colloidal silver with the degree of aggregation under the action of picosecond and nanosecond radiation (λ = 1064 nm). The values of $\chi^{(3)}$ for aggregated silver that are defined by Kerr nonlinearity and thermal self-focusing are found to be -1.5×10^{-14} and -4.4×10^{-11} esu, respectively. The time evolution of the picosecond pulses passing through the colloid is investigated. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Colloidal silver is one of the most extensively studied colloidal solutions. Its optical and nonlinear optical parameters have been the subject of much investigation [1–7]. This colloid is of considerable promise for use in optical limiters, optoelectron devices, etc. The aggregation of microparticles in such structures increases the nonlinear optical response, in particular, the efficiency of degenerate four-photon scattering [8]. This is related to the growth of the local field amplitude in fractal clusters with a high polarizability. The optical characteristics of colloidal silver are governed by the dependence of its plasma oscillation frequency (the corresponding wavelength is about 415 nm) and the long-wave edge of its absorption spectrum on the degree of aggregation [9]. Such a conclusion was drawn by Karpov *et al.* [10] from theoretical [11] and experimental [6, 7] studies on the absorption spectrum of colloidal silver in the vicinity of plasma resonance. The interaction of multipoles induced by an electromagnetic field in metallic clusters causes broadening of the absorption spectra of aggregated colloids [12]. The degree of aggregation specifies the spatial parameters of the aggregates, making it possible to analyze their dimensional characteristics without having to use electron microcopy methods.

The purpose of this work is to set a correlation between the nonlinear optical characteristics of colloidal silver and the degree of aggregation from the spectral characteristics of the colloidal solution.

EXPERIMENTAL

The nonlinear optical characteristics of silver colloids at the wavelength λ = 1064 nm were studied by the conventional *Z*-scan technique [13]. A Nd:YAG laser generated trains of picosecond pulses. A pulse (t = 35 ps and E = 1 mJ) selected from the train was focused onto 5-mm-thick quartz cuvette *2* with colloidal silver by lens *1* with a focal length of 25 cm (Fig. 1). Micropositioner *8* moved the cuvette along the *Z* (optical) axis so that it crossed the focal spot. The laser maximum intensity (4×10^{11} W/cm^2) was below the optical breakdown threshold for the colloidal solutions studied. The energy of the pulses was measured by FD-24K calibrated photodiode *4* and then recorded by V4-17 digital voltmeter *5*. Calibrated neutral density filters were used to vary the energy of the laser radiation.

A Q-switched Nd:YAG laser generated nanosecond pulses of wavelength λ = 1064 nm, energy E = 18 mJ, and duration t = 28 ns. The maximum intensity of the nanosecond radiation at the focal waist was 8×10^9 W/cm^2.

An iris with a limiting diameter of 1 mm transmitting about 3% of the laser radiation was placed at a distance of 150 cm from the focal spot. The energy of the

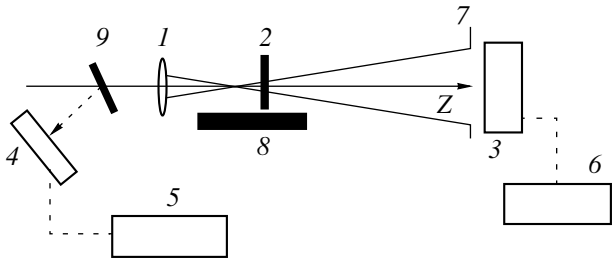


Fig. 1. Experimental scheme: *1*, focusing lens; *2*, cuvette with colloidal solution; *3* and *4*, photodiodes; *5* and *6*, digital voltmeters; *7*, iris; *8*, micropositioner; and *9*, deflector.

1063-7842/02/4707-0889$22.00

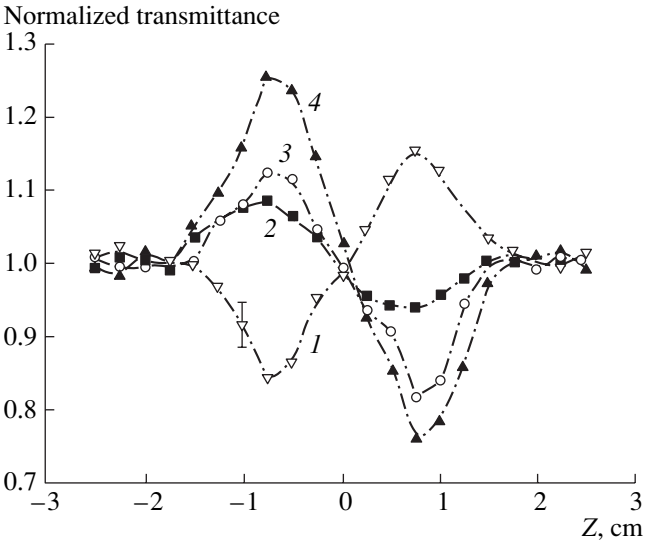


Fig. 2. Normalized transmittance versus the position of the cuvette with colloidal silver in the limiting-iris scheme. The degree of aggregation A is (*1*) 0, (*2*) 0.16, (*3*) 0.54, and (*4*) 0.8.

laser radiation transmitted by the iris was measured by FD-24K photodiode *3*. The signal from the photodiode was applied to V4-17 digital voltmeter *6*. The signal detected by photodiode *3* was normalized to the signal detected by photodiode *4* to eliminate laser output fluctuations. The use of a limiting iris enables one to determine both the sign and the absolute value of the nonlinear refractive index n_2, as well as the nonlinear susceptibilities $\chi^{(3)}$, of silver colloidal solutions. To determine the nonlinear absorption coefficient β, we opened the iris to maximum and measured the transmission of the colloidal solution vs. cuvette position relative to the focal spot. The relatively large size of the detector measuring the transmitted energy away from the cuvette made it possible to detect the entire radiation transmitted. Therefore, the decrease in the transmission in the experiments with the open iris was due to the nonlinear absorption in the cuvette.

The standard solution of silver was prepared by dissolving 20 mg of collargol in 200 mg of bidistilled water. The solution was heated until it turned yellow. The aggregation of the colloidal silver took place under natural conditions (in the absence of UV light, stimulating aggregation).

We recorded the time waveforms of the nanosecond pulses that passed through the cuvette with the colloidal silver by means of an FK-39 coaxial phototube and an S8-19 oscilloscope with a time resolution of no worse than 2 ns. The sizes of the aggregates were determined by electron microscopy. The absorption spectra of the samples at various stages of aggregation were taken by an SF-26 spectrophotometer.

RESULTS AND DISCUSSION

Below, we present the results of the *Z*-scan experiments, which demonstrate, in particular, the self-defocusing of nanosecond pulses due to thermal energy transfer from metal particles to the solvent (water). This nonlinearity is too "slow" to be detected in experiments with picosecond pulses. Our measurements show that the nonlinear optical response of the colloidal structure exhibits both "slow" and "fast" components related to the thermal nonlinearity of the surrounding insulator and the Kerr nonlinearity, respectively. We will analyze mechanisms of self-defocusing in silver colloidal solutions exposed to laser pulses of various duration and determine the nonlinear susceptibilities of associated processes.

Earlier, thermal self-defocusing in colloidal gold (a medium similar to that under study) [14] was analyzed in terms of the time characteristics of a nanosecond laser pulse (λ = 527 nm) transmitted through the solution and limiting iris in *Z*-scan experiments and heat transfer from the metallic clusters to the surrounding insulator. In these experiments, the nonlinear susceptibility was calculated ($\chi^{(3)} = 4.5 \times 10^{-11}$ esu) and it was concluded that the Kerr mechanism of self-focusing is of minor importance in the field of picosecond pulses with a medium amplitude ($I \sim 10^9$ W/cm^2). Also, the upper limit of the Kerr nonlinearity ($\chi^{(3)} = 2.9 \times 10^{-14}$ esu) was estimated. The intensity of the picosecond radiation in the experiments cited was relatively low [14] (1.4×10^9 W/cm^2); as a result, the dependence $T(Z)$ typical of media with Kerr self-focusing was not found. Subsequent experiments with picosecond pulses of higher intensity allowed us to validate the conclusion that the Kerr nonlinearity of colloidal gold is low ($\chi^{(3)} = 1.8 \times 10^{-14}$ esu). Below, we report the results of similar experiments with colloidal silver.

Figure 2 shows a family of plots $T(Z)$ for picosecond pumping and various degrees of aggregation of colloidal silver. Note the change of the sign of n_2 (it is positive in curve *1* and negative in curves *2–4*). A variable parameter here is the degree of aggregation that is determined from the absorption spectra [9, 10] by calculating the difference between the integral absorptions in the long-wave part of the spectrum for aggregated and nonaggregated colloidal solutions. The results are normalized by the maximum possible degree of aggregation of the sol (that is, immediately prior to its precipitation).

The nonlinear optical characteristics of the colloidal silver were determined by the relationship [13]

$$\Delta T_{p-v} = 0.404(1-S)^{0.25} \times |(2\pi/\lambda)n_2 I[1-\exp(-\alpha L)]/\alpha, \quad (1)$$

where ΔT_{p-v} is the normalized difference between the peak and valley in the dependence $T(Z)$, I is the intensity of radiation, S is the transmittance of the iris (the fraction of radiation detected by the diode), λ is the

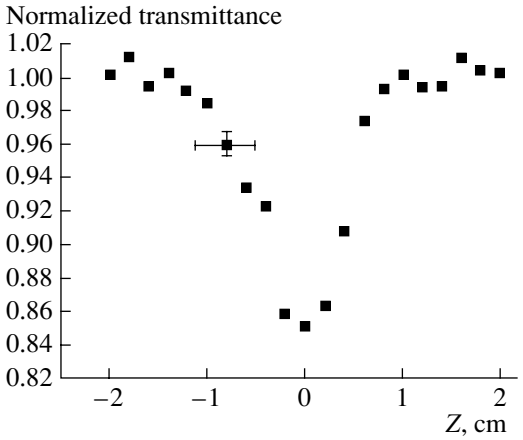


Fig. 3. The same as in Fig. 2 in the open-iris scheme for $A = 0.8$.

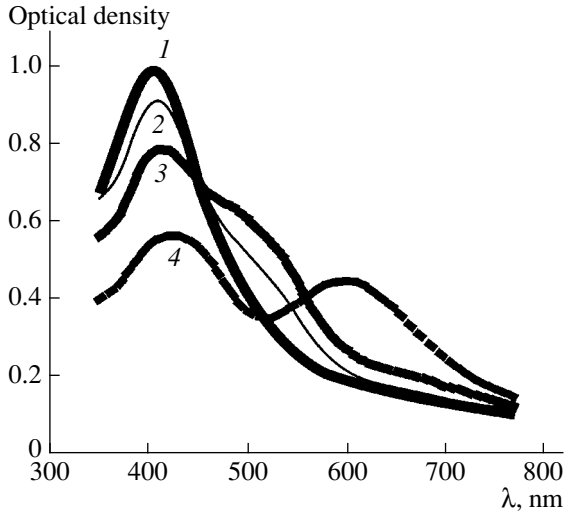


Fig. 4. Absorption spectra of colloidal silver for various degrees of aggregation. The values of A are the same as in Fig. 2.

wavelength of radiation, L is the length of the sample, and α is the linear absorption coefficient.

The nonlinear susceptibility of the silver colloid is represented as [15]

$$\chi^{(3)} = n_0 n_2 / 3\pi, \tag{2}$$

where n_0 is the linear refractive index.

The nonlinear absorption coefficient is given by [16]

$$\beta = q_0 / I L_{\text{eff}}, \tag{3}$$

$$T_0 = q_0^{-1} \ln(1 + q_0), \tag{4}$$

where T_0 is the minimum normalized transmission in the open-iris scheme and q_0 is a dimensionless parameter. The value of T_0 was determined from the plot $T(Z)$ in the open-iris scheme (Fig. 3).

Figure 4 shows the absorption spectra of samples with the same degrees of aggregation as in Fig. 2. It is seen that the increase in the degree of aggregation broadens the plasma resonance peak (415 nm) and gives rise to a spectral feature at longer wavelengths. As the degree of aggregation grows, the absorption coefficient increases (from 0.8 to 2.0 cm^{-1}). Therefore, we successively decreased the picosecond radiation intensity to avoid optical breakdown of the sample. This was taken into account in the calculations of the nonlinear susceptibility, nonlinear refractive index, and nonlinear absorption. The maximum nonlinear absorption ($\beta = 3.8 \times 10^{-11}$ W/cm) corresponded to the degree of aggregation $A = 0.8$. Note the absence of two-photon absorption at the initial stages of aggregation.

Figure 5a plots the nonlinear refractive index against the degree of aggregation A for the picosecond pumping pulses. The values of $\chi^{(3)}(-\omega; \omega\omega - \omega)$ and n_2 for the aggregated silver were found to be 1.5×10^{-14} and -1.1×10^{-13} esu, respectively. The same measurements were made for the nanosecond pulses. Generally, the variations in the nonlinear optical parameters with increasing degree of aggregation (n_2 changes sign and grows) remained the same (Fig. 5b). Note the absence of nonlinear absorption in the field of the nanosecond pumping radiation when the aggregation is maximum.

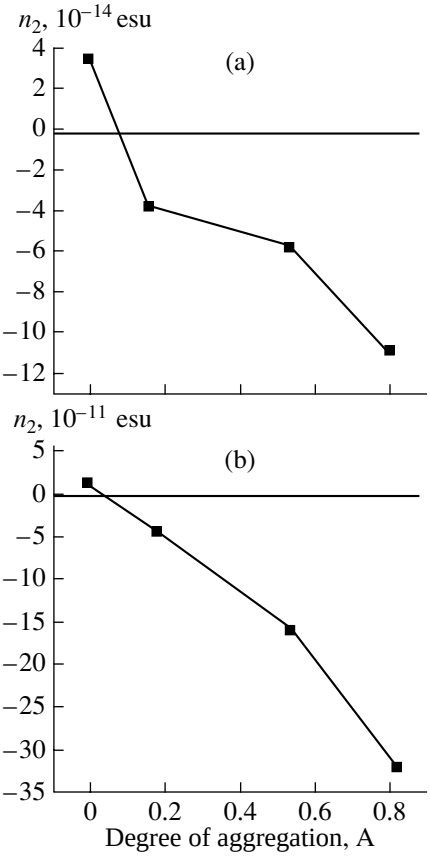


Fig. 5. Nonlinear refractive index versus the degree of aggregation for (a) picosecond and (b) nanosecond pumping.

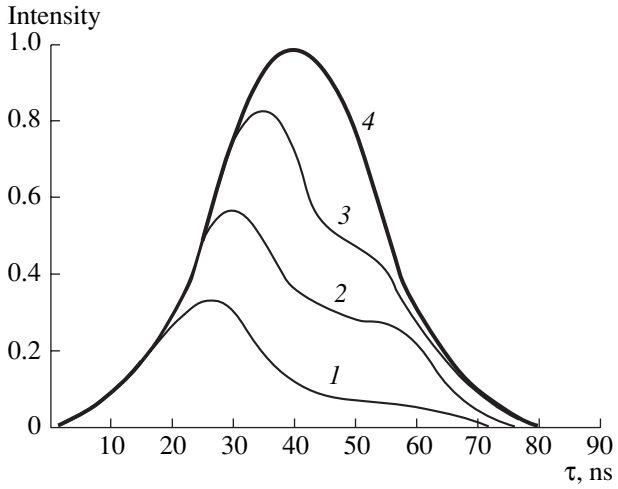


Fig. 6. Time waveforms of laser pulses having passed through the cuvette and the iris at various positions of the cuvette relative to the focus: (*1*) 0.8, (*2*) 1.5, (*3*) 1.9, and (*4*) 3.5 cm.

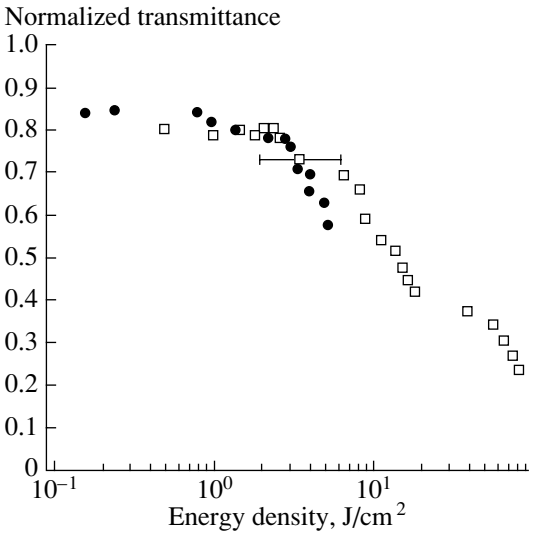


Fig. 7. Normalized transmittance of the cuvette with colloidal silver versus the intensity of picosecond (closed circles) and nanosecond (open squares) radiation with the wavelength $\lambda = 1064$ nm.

This would be natural for the nanosecond intensity $I_{max} = 8 \times 10^9$ W/cm^2, which is substantially less than the intensity of the picosecond radiation $I_{max} = 4 \times 10^{11}$ W/cm^2, for which two-photon absorption (Fig. 3) related to the appearance of the absorption band at 520–550 nm was observed (curve *4* in Fig. 4). In Fig. 5b, the values of $\chi^{(3)}(-\omega; \omega\omega - \omega)$ and n_2 were found to be -4.4×10^{-11} and -3.2×10^{-10} esu, respectively. The positiveness of n_2 at low degrees of aggregation is likely associated with the strong effect of the solvent (water), having a positive nonlinear refractive index, and with the low concentration of the colloidal silver. The concentration characteristics of the nonaggregated silver showed that n_2 changes sign from positive to negative when the initial concentration rises five- and tenfold.

Optical and electron micrographs of colloidal silver particles settled on a substrate at various stages of aggregation suggest that their size varies from 10 nm (initial stages of aggregation) to 40–100 nm.

To discover the physical mechanisms of self-defocusing of the nanosecond radiation in the aggregated silver, we recorded the time waveforms of the laser pulses that passed through the sample and the limiting iris for different cuvette positions relative to the focal spot ($Z = 0$). At a certain energy density, defocusing took place, showing up as a decrease in the signal intensity at the FK-39 photodetector. Figure 6 shows the waveforms at various cuvette–focus distances. The closer the cuvette to the position $Z(T_{min})$ corresponding to the minimum transmission, the earlier the maximum of the pulse detected. Such a scenario is typical of thermal defocusing. The thermal lens forms with increasing the absorption. Therefore, such a lens mostly affects the trailing edge of the pulse. For colloidal gold, the picture is similar [14, 17].

Note the decisive role of the energy (not power) density of the nanosecond radiation in the development of thermal self-defocusing. Our results can be interpreted as thermal self-defocusing when part of the energy absorbed by the clusters is transferred to water molecules. The degree of thermal self-defocusing depends on the integral value of the absorbed energy. That is why the energy density, rather than the power density, is a prime consideration.

As was mentioned above, media such as that considered in this article are of great promise for optical limiters of laser radiation that provide an inertialess decrease in the transmission when the energy density rises. We examined the applicability of colloidal silver for this purpose. In the *Z* scheme with the limiting aperture, the cuvette was placed in the position of minimum transmission and the normalized transmission as a function of the pumping energy density was measured (Fig. 7). When the energy density was increased to a certain threshold value (about 3 J/cm^2), the fraction of the energy transmitted through the iris sharply decreased. The faster decline of the picosecond pulses is associated with the joint effect of Kerr self-defocusing and nonlinear absorption. In the case of the nanosecond pulses, the latter effect is insignificant. Note that the efficiency of the limiters based on colloidal silver is less than that of the existing devices, employing fullerenes and semiconductor structures.

CONCLUSION

Based on the spectral characteristics of the solutions, we derived relationships between the nonlinear optical parameters of colloidal silver exposed to laser pulses of various duration and the degree of aggrega-

tion. The time waveforms of the nanosecond pulses that passed through the colloidal solution at various positions of the sample in the *Z* scheme suggest the thermal nature of defocusing in this medium. The behavior of colloidal silver exposed to picosecond radiation is analyzed in terms of Kerr self-defocusing. The results are compared with experimental data on colloidal gold obtained by a similar technique. The promise of aggregated silver for optical limitation of picosecond and nanosecond laser pulses was investigated. The characteristic feature of the results obtained is the change in sign of n_2 with increasing degree of aggregation for both the picosecond and nanosecond pulses. The values of the nonlinear refractive indices corresponding to Kerr and thermal self-defocusing were found to be -1.1×10^{-13} and -3.2×10^{-10} esu, respectively.

The nonlinear absorption becomes significant at the final stages of aggregation, which are characterized by the appearance of a long-wave branch in the absorption spectrum and on increase in the probability of two-photon absorption for the case of picosecond radiation.

REFERENCES

1. D. Ricard, P. Roussignol, and C. Flytzanis, Opt. Lett. **10**, 511 (1985).
2. F. Hache, D. Ricard, C. Flytzanis, and U. Kreibig, Appl. Phys. A **A47**, 347 (1988).
3. Yu. É. Danilova, V. P. Drachev, S. V. Perminov, and V. P. Safonov, Izv. Akad. Nauk, Ser. Fiz. **60** (3), 18 (1996).
4. S. G. Rautian, V. P. Safonov, P. A. Chubakov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **47**, 200 (1988) [JETP Lett. **47**, 243 (1988)].
5. A. V. Butenko, Yu. É. Danilova, S. M. Ishikaev, and V. P. Safonov, Izv. Akad. Nauk SSSR, Ser. Fiz. **53**, 1195 (1989).
6. S. V. Karpov, A. K. Popov, and V. V. Slabko, Pis'ma Zh. Éksp. Teor. Fiz. **66**, 97 (1997) [JETP Lett. **66**, 106 (1997)].
7. S. V. Karpov, A. K. Popov, V. V. Slabko, and G. B. Shevnina, Kolloidn. Zh. **57**, 199 (1995).
8. A. V. Butenko, V. M. Shalaev, and M. I. Shtokman, Zh. Éksp. Teor. Fiz. **94** (1), 107 (1988) [Sov. Phys. JETP **67**, 60 (1988)].
9. S. V. Karpov, A. A. Bas'ko, S. V. Koshelev, *et al.*, Kolloidn. Zh. **59**, 765 (1997).
10. S. V. Karpov, A. K. Popov, and V. V. Slabko, Izv. Akad. Nauk, Ser. Fiz. **60**, 43 (1996).
11. V. M. Shalaev and M. I. Shtokman, Zh. Éksp. Teor. Fiz. **92**, 509 (1987) [Sov. Phys. JETP **65**, 287 (1987)].
12. V. A. Markel', L. S. Muradov, and M. I. Shtokman, Zh. Éksp. Teor. Fiz. **98**, 819 (1990) [Sov. Phys. JETP **71**, 455 (1990)].
13. M. Sheik-Bahae, A. A. Said, T. Wei, *et al.*, IEEE J. Quantum Electron. **26**, 760 (1990).
14. S. C. Mehendale, S. R. Mishra, K. S. Bindra, *et al.*, Opt. Commun. **133**, 273 (1997).
15. J. F. Reintjes, *Nonlinear Optical Parametric Processes in Liquids and Gases* (Academic, New York, 1984; Mir, Moscow, 1987).
16. C. H. Kwak, Y. L. Lee, and S. G. Kim, J. Opt. Soc. Am. B **16**, 600 (1999).
17. G. A. Swartzlanger, B. L. Justus, A. L. Huston, *et al.*, Int. J. Nonlinear Opt. Phys. **2**, 577 (1993).

Translated by A. Chikishev

Technical Physics, Vol. 47, No. 7, 2002, pp. 894–899. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 100–105.

OPTICS, QUANTUM ELECTRONICS

Pulsed–Periodic Discharge in Cesium as an Effective Source of Light

F. G. Baksht and V. F. Lapshin

Ioffe Physicotechnical Institute, ul. Politekhnicheskaya 26, St. Petersburg, 194021 Russia

e-mail: baksht@pop.ioffe.rssi.ru

Received December 19, 2001

Abstract—The spectrum of visible radiation from an inhomogeneous cesium-plasma column is evaluated in the approximation of locally thermodynamically equilibrium plasma. The plasma parameters correspond to a pulsed–periodic low-power discharge at pressures of 0.1–1.0 atm and the axial temperature $T = 5500$ K. Under these conditions, the visible spectrum of the cesium plasma changes from discrete (line) to continuous as the pressure varies from 0.1 to 1.0 atm. This is associated largely with an increase in the intensity of the 6*P* and 5*D* recombination continua and an appreciable shift of the thresholds of the continua toward larger wavelengths (by ≈100 nm for 6*P* and by ≈150 nm for 5*D*) when the plasma density rises to $\approx 4 \times 10^{17}$ cm^{-3}. In this case, the optical thickness of the plasma column approaches unity and the average luminous flux per unit length of the arc column is close to 6500 lm/cm at the column radius $R = 2$ mm and a duty ratio of 0.1. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The fast development and cheapening of start-control electron equipment open up possibilities of fabricating environmentally safe effective radiation sources based on a pulsed–periodic discharge (PPD) in alkali metal vapors. Such a discharge has been shown to offer a number of advantages over a static arc [1–5]. For example, one can produce a high-density plasma with an electron concentration $n_e > 10^{17}$ cm^{-3} by delivering a relatively low power $W = 10$–60 W/cm at temperatures of 5000–6000 K. This greatly increases luminous fluxes from the plasma, as well as improves the energy efficiency of the arc as a source of light and its color characteristics. Moreover, the requirements for the buffer gas considerably loosen: its basic function is only to provide the initial breakdown of the discharge gap. Therefore, one can do away with mercury sources of light. Finally, it should be noted that conventional sources of light have largely line spectra. It has been reported [6] that the spectrum of the cesium plasma demonstrates a progressively increasing fraction of continuous radiation as the plasma density rises. This is because bright 6*P* and 5*D* recombination continua are produced and the thresholds of these continua shift toward larger wavelengths. As a result, with the PPD used, the optical thickness of the cesium plasma may approach unity at $T = 5000$–6000 K, at which the peak of the Planck radiation intensity lies in the range of the maximum sensitivity of the eye. Today, the continuous spectrum is usually provided with a xenon arc plasma, which becomes optically dense at temperatures as high as $T > 10^4$ K [7]. Calculations of the radiation spectrum and luminous fluxes carried out in this work show convincingly that a PPD in cesium is promising as an efficient source with the recombination mechanism of radiation. Preliminary results were reported in [8].

LOCAL THERMODYNAMIC EQUILIBRIUM IN A RADIATING CESIUM PLASMA

The visible spectrum from an inhomogeneous column of the Cs plasma was calculated at pressures $p = 0.1$–1.0 atm. The plasma parameters, specifically, the radial temperature distribution $T(r)$, correspond to a low-power (≈30 W/cm) PPD [4, 5]. In this discharge, current pulses of certain shape and an amplitude of ≈10 A pass through the plasma of a low-current (≈0.1 A) pilot discharge with a repetition rate $\nu \sim 10^3$ Hz. Figure 1 shows the analytical temperature profile $T(r) = T_1[1 - a(r/R)^2] + T_2 \exp\{-(r/bR)^4\}$ ($T_1 = 3500$ K, $T_2 = 2000$ K, $a = 9/14$, and $b = 0.6$), which approximates well the results of numerical simulations [4, 5]. The spectrum was calculated for a locally thermodynamically equilibrium (LTE) plasma.

Let us see how the conditions of local thermodynamic equilibrium are disturbed when the plasma emits in the bright 6*P* and 5*D* recombination continua, as well as produces a line spectrum corresponding to the 6*S*–6*P* and 5*D*–4*F* transitions in a Cs atom. We will also take into consideration all disturbances of the LTE conditions due to radial plasma inhomogeneities. To do this, we introduce the parameters δ_k ($k = 1$–5):

$$\delta_k = n_e n_i \langle v_e \sigma^{ph}_{rec,\gamma} \rangle / \alpha_\gamma n_e^2 n_i; \quad k = 1, 2,$$

$$\delta_k = A_{\gamma\gamma'} \Theta_{\gamma\gamma'} / n_e \langle v_e \sigma_{\gamma\gamma'} \rangle; \quad k = 3, 4,$$

1063-7842/02/4707-0894$22.00

$$\delta_5 = (D_{ia}\tau_{rec})^{1/2}/\left(\frac{n_i}{\partial n_i/\partial r}\right).$$

Here, $n_e = n_i$ is the charge carrier concentration in the plasma, $\sigma^{ph}_{rec,\gamma}(v_e)$ is the cross section of radiation recombination, α_γ is the coefficient of three-body collisional recombination into the final state γ in the Cs spectrum ($\gamma = 5D$ for $k = 1$ and $\gamma = 6P$ for $k = 2$), $A_{\gamma\gamma'}$ is the associated Einstein coefficient [9], $\Theta_{\gamma\gamma'}(r)$ is the probability of a photon escaping from the plasma (see [10, p. 81]), and $\sigma_{\gamma\gamma'}$ is the cross section of de-excitation by electron impact ($\gamma = 4F$, $\gamma' = 5D$ for $k = 3$ and $\gamma = 6P$, $\gamma' = 6S$ for $k = 4$). The symbol <…> means averaging over the Maxwell distribution of the electrons, D_{ia} is the coefficient of Cs^+ ambipolar diffusion, and τ_{rec} is the time of three-body electron–ion recombination. The values of $\sigma_{\gamma\gamma'}$, α_γ, and $\sigma^{ph}_{rec,\gamma}$ were calculated via the principle of detailed balance with the theoretically found values of the cross sections of excitation and ionization by electron impact [11] and those of photoionization [12]. The values of D_{ia} were found [13] from the cross sections of resonant charge exchange [14]. The results for δ_k are shown in Fig. 1. To check the results for δ_4, we also used experimental values of the cross sections for electron impact excitation [15] (dotted lines). It is seen that the values of δ_4 obtained by using the theoretical and experimental cross sections virtually coincide.

The LTE conditions occur in the range $\delta_k \ll 1$.[1] As follows from Fig. 1, the processes that disturb the plasma equilibrium are of minor importance. The plasma is in the nearly LTE state in the most part of the gas-discharge tube volume. The LTE conditions are disturbed only near the walls, where the plasma is cold ($T \le 2000$ K). The cold plasma effectively absorbs the radiation only near the resonance lines of cesium (852.1 and 894.6 nm) and is almost transparent to visible radiation. Therefore, when calculating the visible spectrum, we ignored the disturbances of the LTE conditions and assumed that these conditions are fulfilled throughout the discharge tube. The decrease in the atom ionization energy was evaluated in the Debye approximation, and the partition function for a Cs atom was calculated by the Planck–Larkin method [16].

[1] The value of δ_5 is the ratio of the ion recombination diffusion length to the characteristic scale of plasma inhomogeneity. As was shown in [4], the condition $(D_{ia}\tau_{rec})^{1/2} \sim V_i\tau_{rec}$, where V_i is the hydrodynamic velocity of the ions, holds almost throughout the pulse duration. Therefore, the condition $\delta_5 \ll 1$ also means that the convective length of ion recombination is much smaller than the scale of inhomogeneity.

SPECTRAL RADIATION FLUX FROM PLASMA IN THE VISIBLE RANGE

The spectral radiation flux per unit surface area of a plasma column under the LTE conditions is given by

$$F_\lambda = \int_{(2\pi)} I_\lambda(s, \mathbf{\Omega})(\mathbf{\Omega}\mathbf{e}_r)d\mathbf{\Omega} = F_{\lambda P}(T_w)\varepsilon_\lambda. \quad (1)$$

Here, $I_\lambda(s, \mathbf{\Omega})$ is the spectral radiation intensity at a point s on the plasma column surface, $\mathbf{\Omega}$ and $\mathbf{e}_r$ are the unit vectors specifying the direction of the radiation beam and the normal to the surface (Fig. 2), T_w is the tube wall temperature, $F_{\lambda P}(T) = \pi I_{\lambda P}(T) = 2\pi hc^2\lambda^{-5}/[\exp(hc/\lambda k_B T) - 1]$ is the radiation flux from the black body surface, λ is the wavelength, $I_{\lambda P}(T)$ is the Planck radiation intensity,

$$\varepsilon_\lambda = \frac{4}{\pi}\int_0^{\pi/2}\cos\Theta d\Theta\int_0^{\pi/2}\cos\psi d\psi\int_0^{I_0}k'_\lambda(l)\frac{I_{\lambda P}(T(l))}{I_{\lambda P}(T_w)} \times \exp\left\{-\frac{1}{\cos\psi}\int_0^l k'_\lambda(l')dl'\right\}dl, \quad (2)$$

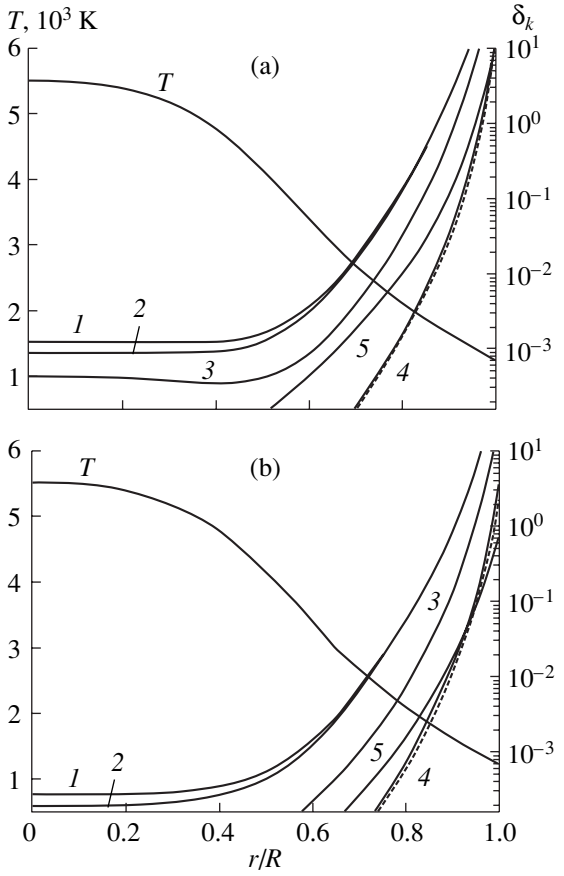


Fig. 1. Radial distribution of the plasma temperature T and the values of the (*1*) δ_1, (*2*) δ_2, (*3*) δ_3, (*4*) δ_4, and (*5*) δ_5 for plasma pressures of (a) 0.1 and (b) 1.0 atm.

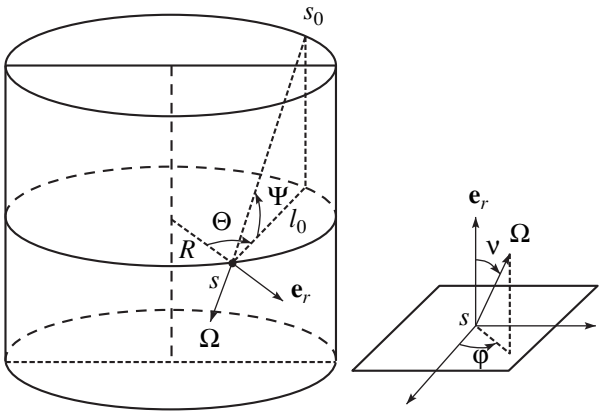


Fig. 2. Geometry of the problem. **Ω** and $\mathbf{e}_r$ are the unit vectors of the radiation direction and the normal to the surface, respectively.

k'_λ is the reduced absorption coefficient (taking into account induced emission), $l_0 = 2R\cos\Theta$, and l is the distance from the point s measured along the projection of the beam onto the plane normal to the discharge axis ($0 \le l \le l_0$). In the inhomogeneous plasma, $0 < \varepsilon_\lambda < \infty$. For the homogeneous plasma, ε_λ is the conventional spectral emissivity of the column:

$$\varepsilon_\lambda^{(h)}(\tau_\lambda) = \frac{4}{\pi}\int_0^{\pi/2} d\Theta\cos\Theta\int_0^{\pi/2} d\psi\cos^2\psi \times [1 - \exp(-\tau_\lambda\cos\Theta/\cos\psi)], \quad (3)$$

where $\tau_\lambda = 2Rk'_\lambda$.

In this case, $\varepsilon_\lambda^{(h)}$ meets the usual inequality $0 \le \varepsilon_\lambda^{(h)} \le 1$. If the plasma is optically dense ($Rk'_\lambda(R) \gg 1$), one can expand ε_λ in the parameter $\xi = [(hc/\lambda)/k_BT]/Rk'_\lambda(R)$ by integrating (2) by parts:

$$\varepsilon_\lambda = 1 - \frac{2}{3}\frac{(dT/dr)|_{r=R}}{T_w/R}\xi + O(\xi^2), \quad \xi \ll 1. \quad (4)$$

Expressions (3) and (4) were used to check the calculation of ε_λ by formula (2). The absorption coefficient k'_λ was calculated (as in [6]) as the sum of the absorption coefficients corresponding to free–free (f–f), bound–free (b–f), and bound–bound (b–b) transitions:

$$k'_\lambda = (k_\lambda^{(ff)} + k_\lambda^{(bf)} + k_\lambda^{(bb)})[1 - \exp(-hc/\lambda k_BT)]. \quad (5)$$

The absorption coefficient of bremsstrahlung radiation was calculated in the Kramers approximation [17]. Note that the contribution of the f–f transitions to k'_λ is small under our conditions. When calculating $k_\lambda^{(bb)}$, we took into account the following transitions in the discrete spectrum of a Cs atom: $6S_{1/2} - nP_{1/2}$, $6S_{1/2} - nP_{3/2}$ (n = 7, 8), $6P_{1/2} - nS_{1/2}$, $6P_{3/2} - nS_{1/2}$ (n = 8–13), $6P_{1/2} - nD_{3/2}$, $6P_{3/2} - nD_{5/2}$ (n = 7–13), $5D_{3/2} - nF_{5/2}$, and $5D_{5/2} - nF_{7/2}$ (n = 5–12). Accordingly, it was assumed that

$$k_\lambda^{(bb)} = \sum_{\gamma,\gamma'} n_\gamma\sigma_{\gamma\gamma'}^{(ph)}(\lambda),$$

where

$$\sigma_{\gamma\gamma'}^{(ph)}(\lambda) = \frac{\pi e^2}{mc^2}\lambda_{\gamma\gamma'}^2 f_{\gamma\gamma'}\varphi_{\gamma\gamma'}(\lambda)$$

is the cross section of photoabsorption for the γ–γ' transition and $f_{\gamma\gamma'}$ is the oscillator strength [9].

The shape of the line $\varphi_{\gamma\gamma'}(\lambda)$ was assumed to be Lorentzian and only the Stark broadening due to electrons was taken into account [18]. When calculating $k_\lambda^{(bf)}$, we considered the photoionization (PI) of the 7*S*, 6*P*, 7*P*, 8*P*, 5*D*, 6*D*, 7*D*, 4*F*, 5*F*, and 5*G* levels in the visible range. The calculation procedure for $k_\lambda^{(bf)}$ was the same as in [6]. The concentration n_γ of excited Cs atoms and the plasma concentration were determined from the Saha–Boltzmann equation.

As was noted, an important feature of the Cs spectrum is the existence of bright 6*P* and 5*D* recombination continua in the visible range (the associated thresholds are $\lambda_{th}(6P)$ = 504 nm and $\lambda_{th}(5D)$ = 594 nm). Let us discuss some points regarding the calculation of k_λ for $\lambda > \lambda_{th}$ slightly exceeding these thresholds. In the dense plasma, the broadening of the spectral lines causes higher order terms approaching the PI threshold to merge. We assumed in this work that adjacent lines overlap if the sum of their half-height widths w_n and w_{n+1} is larger than the their center distance: $w_n + w_{n+1} > \lambda_n - \lambda_{n+1}$. In either spectral series (6*P*–*nD* or 5*D*–*nF*), one can find lines for which $w_{n-1} + w_n < \lambda_{n-1} - \lambda_n$ but $w_n + w_{n+1} > \lambda_n - \lambda_{n+1}$. In the range $\lambda_{th} \le \lambda \le \lambda_{n+1}$, the spectrum was considered to be continuous and the absorption coefficient for the given spectral series γ (γ = 6*P* or 5*D*) was found from the relationship $k_{\lambda,\gamma}^{(bb)} \equiv k_{cont}^{(\gamma)} = n_\gamma\sigma_\gamma^{(ph)}(E)$, where the cross section of photoionization $\sigma_\gamma^{(ph)}$ from the state γ was determined through the spectral density of the oscillator strengths, as in [6, 12]. In the range $\lambda \ge \lambda_n$, the γ spectrum was considered to be discrete; thereby, $k_{\lambda,\gamma}^{(bb)} \equiv k_{disc}^{(\gamma)} = \sum_{\gamma'} n_\gamma\sigma_{\gamma\gamma'}(\lambda)$. In the intermediate range $\lambda_{n+1} < \lambda < \lambda_n$, we used the linear interpolation

$$k_{\lambda,\gamma}^{(bb)} = k_{disc}^{(\gamma)}(\lambda - \lambda_{n+1})/(\lambda_n - \lambda_{n+1}) + k_{cont}^{(\gamma)}(\lambda_n - \lambda)/(\lambda_n - \lambda_{n+1}).$$

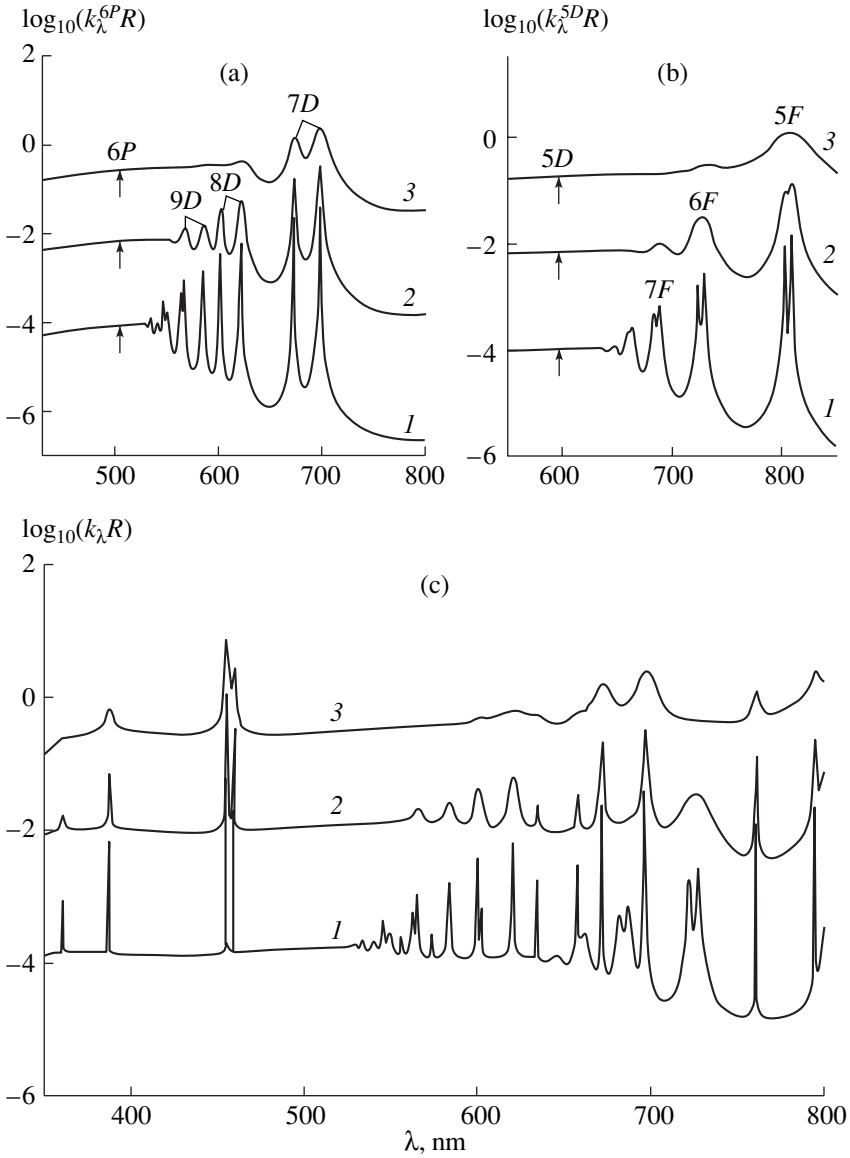


Fig. 3. Absorption spectrum of the cesium plasma under thermodynamic equilibrium at $T = 5500$ K for pressures of (*1*) 0.01, (*2*) 0.1, and (*3*) 1.0 atm. (a) Spectral series 5*D*–*nF*, (b) spectral series 6*P*–*nD*, and (c) absorption spectrum. $R = 2$ mm.

RESULTS OF COMPUTATION

Figure 3 shows the calculated values of the absorption coefficient for the cesium plasma at $T = 5500$ K and pressures of 0.01, 0.1, and 1.0 atm. The associated electron concentrations are $n_e = 6.58 \times 10^{15}$, 5.89×10^{16}, and 3.67×10^{17} cm^{-3}. The individual absorption coefficient for the spectral series 6*P*–*nD* and 5*D*–*nF* near the continuum threshold are shown in Figs. 3a and 3b, respectively. The threshold wavelengths of the 6*P* and 5*D* continua for an isolated Cs atom are 504 and 594 nm (indicated by arrows in Figs. 3a and 3b). It is seen that the continua shift toward larger wavelengths with increasing pressure. This is because two effects set in simultaneously: the atom ionization energy diminishes and higher terms of the spectral series converging to the PI threshold merge (due to the broadening). As a result, the continua overlap nearly in the entire visible range of the spectrum at $p \approx 1.0$ atm (Fig. 3c). In this case, the optical thickness of the plasma approaches unity in most of the spectrum. Figure 4 demonstrates the calculated spectral flux E_λ (continuous curve) per unit surface area of the inhomogeneous plasma column for the temperature profile $T(r)$ mentioned above and also the Planck flux $F_{\lambda P}(T_*)/3$ (dashed curve) from the black body surface at $T_* = 5000$ K. At $p = 0.1$ atm (Fig. 4a), the optical thickness of the plasma in the continuum is small. The plasma radiation is volume, and its spectrum is discrete. The radiation flux F_λ is smaller than the Planck flux. At $p = 1.0$ atm (Fig. 4b), the optical thickness approaches unity and the spectrum

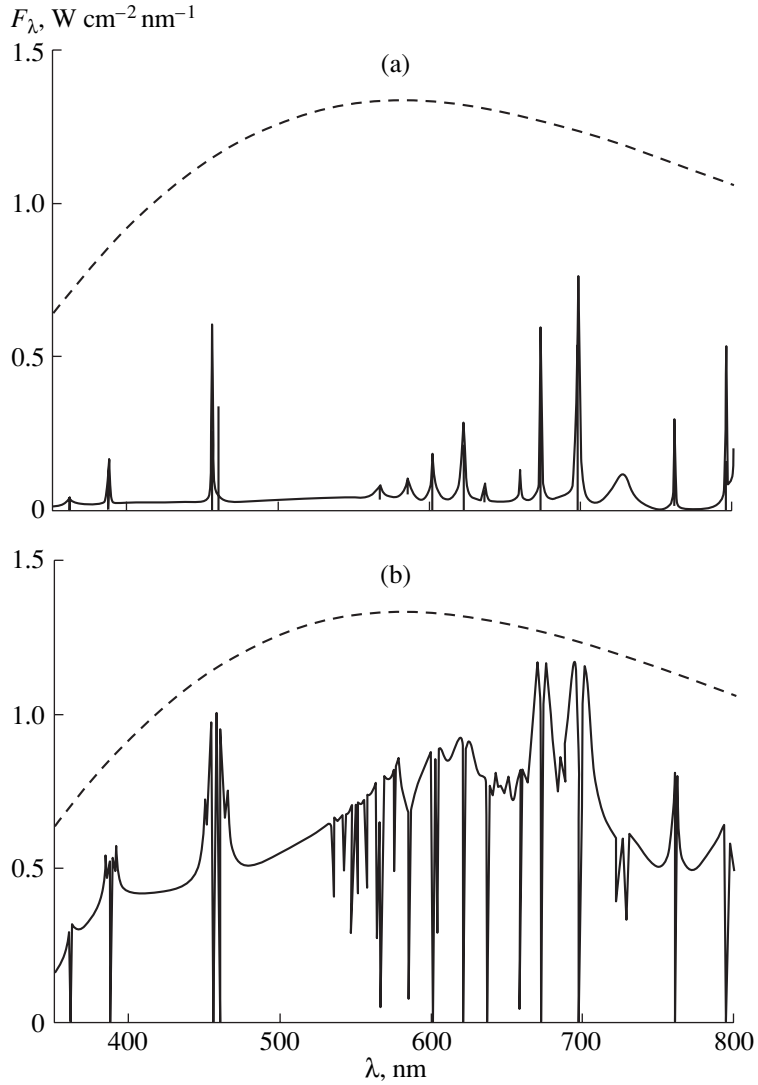


Fig. 4. Spectral density of the radiation radial flux in the visible range for pressures of (a) 0.1 and (b) 1.0 atm.

becomes almost continuous. In this case, the flux F_λ amounts to a considerable fraction of the Planck flux $F_{\lambda P}(T_*)$. Note that the emission spectrum of the inhomogeneous arc plasma is much smoother than the spectrum from a plasma homogeneous column [6]. This is because the outer (colder) plasma region intensely absorbs the line radiation.

An important parameter of radiation sources is the luminous flux Φ per unit length of an arc column: $\Phi = 2\pi RK_{max}\int F_\lambda V(\lambda)d\lambda$, where $V(\lambda)$ is the spectral luminous efficacy of the radiation and K_{max} = 680 lm/W is the maximal value of the luminous efficacy (which is achieved at λ = 555 nm). The mean values of the luminous flux, $\Phi^* = \nu t_p \Phi$, calculated for the characteristic duty ratio of current pulses in the PPD $\nu t_p = 0.1$ [4, 5] (t_p is the current pulse duration) and R = 2 mm were found to be $\Phi^* = 0.48 \times 10^3$ lm/cm at p = 0.1 atm and $\Phi^* = 0.65 \times 10^4$ lm/cm at p = 1.0 atm.

CONCLUSION

Thus, under the PPD conditions, a rise in the cesium plasma pressure to 1 atm causes qualitative changes in the arc radiation spectrum. The intensities of the bright 6*P* and 5*D* continua rise and their thresholds shift markedly (by more than 100 mm) toward larger wavelengths, resulting in a virtually continuous radiation spectrum in the visible range. Under these conditions, the optical thickness of the plasma approaches unity, which provides high mean values of the luminous fluxes Φ^* from an arc column. The calculated values of Φ^* far exceed luminous fluxes provided by continuous-spectrum radiation sources at a comparable energy consumption, W = 30–50 W/cm (for example, by high-

pressure xenon lamps [19]). Therefore, a PPD in cesium vapor seems to be a promising effective source of recombination-induced radiation.

REFERENCES

1. C. L. Chalek and R. E. Kinsinger, J. Appl. Phys. **52**, 716 (1981).
2. R. Schafer and H. P. Stormberg, J. Appl. Phys. **57**, 2512 (1985).
3. K. Gunther, H.-G. Kloss, T. Lehmann, *et al.*, Contrib. Plasma Phys. **30**, 715 (1990).
4. F. G. Baksht and V. F. Lapshin, Zh. Tekh. Fiz. **66** (11), 170 (1996) [Tech. Phys. **41**, 1177 (1996)].
5. F. G. Baksht and V. F. Lapshin, Zh. Tekh. Fiz. **67** (9), 22 (1997) [Tech. Phys. **42**, 1004 (1997)].
6. F. G. Baksht and V. F. Lapshin, Pis'ma Zh. Tekh. Fiz. **23** (24), 40 (1997) [Tech. Phys. Lett. **23**, 961 (1997)].
7. G. N. Rokhlin, *Discharge Light Sources* (Énergoatomizdat, Moscow, 1991).
8. F. G. Baksht and V. F. Lapshin, in *Proceedings of the Conference on Low Temperature Plasma, Petrozavodsk, 2001*, Vol. 1, p. 122.
9. P. M. Stone, Phys. Rev. **127**, 1151 (1962).
10. L. M. Biberman, V. S. Vorob'ev, and I. T. Yakubov, *Kinetics of Nonequilibrium Low-Temperature Plasmas* (Nauka, Moscow, 1982; Consultants Bureau, New York, 1987).
11. I. I. Sobelman, L. A. Vainshtein, and E. A. Yukov, *Excitation of Atoms and Broadening of Spectral Lines* (Nauka, Moscow, 1979; Springer-Verlag, Berlin, 1981).
12. J. Lahiri and S. T. Manson, Phys. Rev. A **33**, 3151 (1986).
13. A. A. Radtsig and B. M. Smirnov, in *The Mobility and Diffusion of Ions in Gases*, Ed. by E. W. McDaniel and E. A. Mason (Wiley, New York, 1973; Mir, Moscow, 1976).
14. S. Sakabe and Y. Izawa, Phys. Rev. A **45**, 2086 (1992).
15. N. S. Scott, K. Bartschat, P. G. Burke, *et al.*, J. Phys. B **17** (6), 191 (1984).
16. V. E. Fortov and I. T. Yakubov, *Nonideal Plasma* (Énergoatomizdat, Moscow, 1994),
17. I. I. Sobelman, *Atomic Spectra and Radiative Transitions* (Fizmatgiz, Moscow, 1963; Springer-Verlag, Berlin, 1979).
18. H. R. Griem, *Spectral Line Broadening by Plasmas* (Academic, New York, 1974; Mir, Moscow, 1978).
19. *Handbook of Lighting Technology*, Ed. by Yu. B. Aĭzenberg (Énergoatomizdat, Moscow, 1995).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 7, 2002, pp. 900–905. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 106–112.

RADIOPHYSICS

Collective Processes in a Microsecond Relativistic Electron Beam: Basic Rules and Mechanisms

E. V. Ganichev, N. V. Dvoretskaya, and G. G. Sominskiĭ

St. Petersburg State Technical University,
ul. Politekhnicheskaya 29, St. Petersburg, 195251 Russia
e-mail: sominski@prhf.spbstu.ru

Received January 8, 2002

Abstract—Collective processes taking place in the space charge of a microsecond relativistic electron beam with magnetic insulation are considered. The space–time characteristics of the low-frequency and high-frequency oscillations are found, and the effect of the magnetic compression of the beam near the cathode on these oscillations is studied. It is shown that the basic source of the low-frequency oscillations is the collective motion of the space charge, which takes place in crossed electric and magnetic fields near the cathode, while the primary reason for the high-frequency oscillations is two-stream instability in the beam. The possibility of suppressing both types of oscillations by compressing relativistic electron beams near the cathode is demonstrated. The effect of nonuniform magnetic fields, including their effect on the cathode plasma motion, is elucidated. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The successful development and application of high-power microwave devices imply the generation of high-quality relativistic electron beams (REBs) (see, e.g., [1]). Space-charge oscillations occurring in the beam worsen its quality, since they disturb the magnetic insulation and cause spurious signals. Early investigations [2–4] discovered a number of basic properties of space-charge oscillations in long-pulsed (microsecond) REBs. In a typical REB-forming system, a type-1 coaxial diode with magnetic insulation (CDMI-1) and an explosive emission cathode, two branches of space-charge oscillations (SCOs) typical of such beams have been revealed: low-frequency (LF) oscillations in the frequency interval 100–300 MHz and high-frequency (HF) oscillations with maximal amplitudes in the frequency interval 400–800 MHz. It has been found that the HF oscillations are related to the motion of the cathode plasma. Another reason for the HF oscillations is an electron stream propagating with a large transverse velocity from the side surface of the plasma emitter in the REB halo. Two mechanisms behind the HF oscillations where the role of the electrons with a high transverse velocity is significant have been considered [4]. One is the development of two-stream instability due to the interaction of the low-density stream in the REB halo with the main electron stream from the end face of the plasma emitter. The other mechanism is diocotron instability in the electron stream from the side surface of the plasma emitter. Both mechanisms must lead to the development of SCOs in an REB at frequencies where the intensity of the HF oscillations observed experimentally is the highest.

Reasons for the LF oscillations have also been considered [4]. It was suggested that they are excited in a stream of electrons reflecting from a collector and move between the collector and the cathode for a long time. However, the nature of the LF oscillation is also entirely understood. Experimental data obtained in experiments with the CDMI-1 are scarce because of the device's particular geometry and fixed ranges of electron energies (200–240 keV) and magnetic fields (≈1 T). This does not allow one to judge the generality of results reported in [2–4]. Another difficulty associated with the interpretation of the results found is that we do not know the properties of the collective processes in the near-cathode zone. Moreover, little is known of how magnetic field nonuniformities near the cathode, which are used to compress REBs, influence these processes. In this work, we report new data that might clarify the nature of the collective processes in REBs. Our experiments were performed with a type-2 CDMI (CDMI-2), which has an essentially different geometry of REB formation and transport. SCOs were detected both in the beam transport channel and near the emitter. Specially designed magnetic-field-generating devices allow for control of the field strength and the beam compression at the cathode.

EXPERIMENTAL SETUP AND MEASURING TECHNIQUES

Our experiments were carried out on an SÉR-1 setup [2–4]. Table 1 compares the parameters of CDMI-1 and CDMI-2.

The properties of the oscillations were studied in devices with edge explosive-emission cathodes made

1063-7842/02/4707-0900$22.00

of stainless steel over a wide range of REB-confining magnetic fields (0.6–1.4 T). Accelerating voltage pulses had a peak amplitude $U \approx 220$ kV. Their duration was limited by the magnetic insulation of the REBs and varied roughly between 2 and 5 μs.

The cross section of the CDMI-2 is shown in Fig. 1. Electrons emitted by the cathode are accelerated in the anode–cathode space; move in the grounded transport channel electrically connected to the anode; and were collected on the inner surface of the collector, which is insulated from the channel. The distribution of the magnetic field B is specified by solenoids $S1$–$S5$. With solenoid $S5$ switched off, the magnetic field is roughly uniform throughout the accelerating and transport regions, except in the collector area. The field of correction coil $S4$ prevents electrons of the cathode from penetrating to the grounded electrodes placed on the side opposite of the entrance to the transport channel. With solenoid $S2$ switched off, the beam area shrinks nearly twofold. With solenoids $S2$ and $S3$ switched off simultaneously, the compression increases fourfold. Solenoid $S5$ is used to produce a locally nonuniform magnetic field in front of the collector.

Under these conditions, the beam current I_c toward the collector varied from 0.7 to 2 kA. The leading edge duration of the pulse U was no more than ≈0.4 μs. On the top of the pulse, the voltage slowly declined because of the discharge of the voltage pulse generator capacitors. A plasma short, formed between the cathode and the grounded electrodes as a result of cathode plasma expansion [1], caused a voltage sharp drop at the trailing edge of the pulse U.

To characterize SCOs in the REB, probes $P1$, $P2$, and $P3$ were used. The first two were placed at distances of 40 and 80 cm from the cathode. Probe $P3$ was placed over the cathode holder. Detecting the probe signals at various time instants during a current pulse, one can keep track of the SCOs. The Fourier analysis of short (30–120 ns) temporal realizations [2–4] made it possible to find the shape of "instantaneous" oscillation spectra. Probes $P1$ and $P2$ were connected to the REB transport channel through small (diameter 5 mm) holes and detected signals coming from the nearest zone [2–4] at distances of ~40 and ~80 cm, respectively. Probe $P3$ detected HF signals coming from the near-cathode zone along the coaxial line formed by the cathode holder and the external electrode (sheath) of the device. Signals of frequencies less than 1 GHz coming from distant (relative to the cathode) areas of the transport channel could not reach this coaxial line, since the cutoff wavelength of the circular drift tube of the beam did not exceed roughly 20 cm.

RESULTS AND DISCUSSION

The spectral characteristics of SCOs in the REBs turned out to be qualitatively similar throughout the range of magnetic fields and current beams. To describe them, we will make use of experimental data obtained for the magnetic field $B = 1$ T. These data are convenient to correlate with those obtained with the CDMI-1 [2–4]. First, we will consider SCOs in an REB confined by an axially uniform magnetic field.

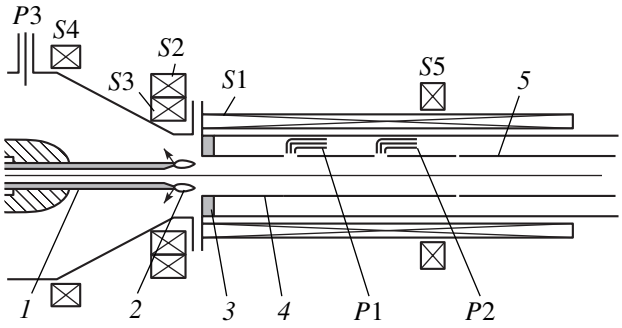


Fig. 1. CDMI-2 used in the experiments: *1*, cathode; *2*, cathode plasma; *3*, anode; *4*, drift channel; *5*, collector; $S1$–$S5$, solenoids; $P1$–$P3$, probes. Arrows indicate the direction of the lines of the electric field at the "inner" surface of the cathode plasma.

In remote areas (at distances $L \geq 40$ cm from the cathode), the oscillation spectra were determined with probes $P1$ and $P2$. The spectra recorded at various time instants (with different delays τ from the onset of the REB) are shown in Figs. 2 and 3. In this and subsequent figures, $U = 220$ kV. For $0.3 \leq \tau \leq 1.0$–1.2 μs, one can distinguish two frequency domains (100–300 and 400–800 MHz) wherein the oscillation amplitudes are increased.[1] The maximal amplitudes A_m in both domains vary with time nonmonotonically and peak (without compression at the cathode) at $\tau = 0.5$–0.8 μs, that is, with delay relative to the leading edge of the voltage pulse U, under all conditions studied. Figures 4a and 4b show the typical maximal amplitude A_m vs. record instant τ dependences that follow from the signals coming to probe $P1$. The ratios between the maximal amplitudes of the LF and HF branches (A_{LFm} and A_{HFm}, respectively) at probes $P1$ and $P2$ are different. At probe $P1$ (that nearest to the cathode), the amplitudes are roughly equal to each other, while at remote probe $P2$, $A_{HFm} \gg A_{LFm}$.

To summarize the measurements performed with probes $P1$ and $P2$, one can conclude that the basic characteristics of the SCOs in the REB passing in the transport channel are qualitatively similar in the CDMI-1 and CDMI-2, which differ in size, if there are no strong variations of the magnetic field in the beam formation and transport regions. The remarkable general features are almost identical structure of the spectra, including the HF and LF oscillation branches, and an increase in

[1] Along with these oscillations, those at frequencies below 50 MHz are sometimes observed. Such poorly reproducible oscillations can be associated with fluctuations of the electron current from the explosive emission cathode [1]. Here, we consider only the range $f > 50$ MHz.

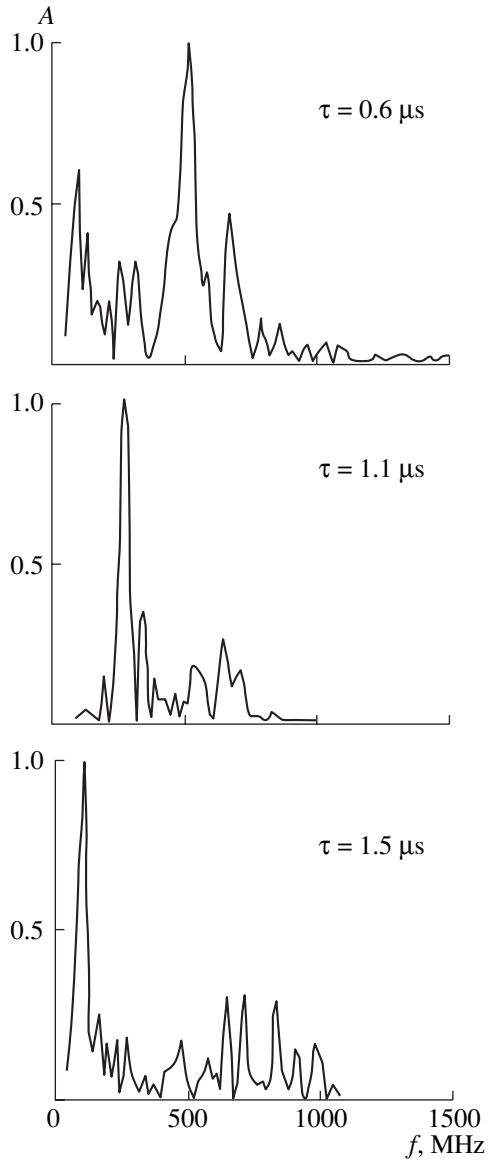


Fig. 2. Typical SCO spectra from probe *P*1 at the magnetic field $B = 1$ T at various time instants after the onset of the voltage pulse.

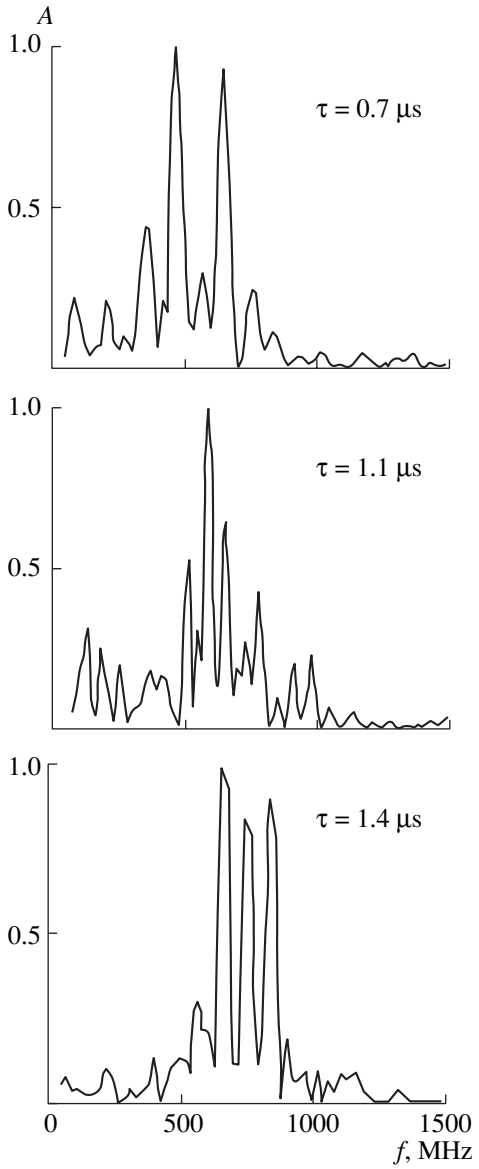


Fig. 3. The same as in Fig. 2 for probe *P*2.

the oscillation amplitude after the leading edge of the accelerating voltage pulse. As was already noted, such an increase, first discovered in [4] with the use of the CDMI-1, has been explained by the motion of the cathode plasma in the accelerating space and, accordingly, with the appearance of electrons having high transverse velocities. The difference between the ratios A_{LFm}/A_{HFm} at probes *P*1 and *P*2 may be related to an HF oscillation build-up from *P*1 to *P*2 (see also [4]). Basically, such a variation of the HF oscillation amplitude may be observed in the case of both two-stream instability in the REB and diocotron instability at the side surface of the cathode plasma [4]. The data found in [4] provide no grounds to decide between the two models. Our measurements indicate that the HF oscillation amplitudes at probes *P*1 and *P*2 are always maximal at frequencies in the interval 400–800 MHz irrespective of the magnetic field. This is illustrated by typical spectra for probe *P*2 taken at magnetic fields of 1, 0.8, and 0.6 T (Figs. 3, 5). The fact that the typical frequencies

Table 1

Basic parameters	CDMI-1	CDMI-2
Cathode diameter, mm	20	37
Transport channel diameter, mm	32	52
Cathode–anode spacing, mm	12.27	30
Transport channel length, mm	≈1200	≈1450

of the HF oscillations do not depend on the magnetic field seems to be in favor of the "two-stream" mechanism of their origination. If the HF oscillations were related to diocotron instability and the associated azimuth motion of space-charge bunches in the crossed fields at the side surface of the plasma emitter, one could expect a rise in their frequency with decreasing magnetic field because of the enhanced drift velocity V_d ($V_d = E_\perp/B$) of the azimuth motion of the bunches.

Thus, the data obtained with the CDMI-2 suggest the generality of the collective processes in greatly differing REB formation systems and, hence, under essentially different experimental conditions. In addition, our data shed light on a model of the HF oscillation branch.

Probe *P*3 provided information on the SCOs in the near-cathode zone of REB formation. Associated spectra are shown in Fig. 6. As for probes *P*1 and *P*2, the oscillation amplitude is maximal for τ = 0.5–0.8 μs. However, the LF oscillations in the signal coming to probe *P*1 far exceed the HF ones in amplitude if $\tau \le$ 1.0–1.1 μs. The difference is especially noticeable at τ = 0.5–0.8 μs, where the ratio A_{LFm}/A_{HFm} exceeds 100. At longer delays ($1.1 \le \tau \le 1.5$ μs), the amplitude of the HF components in the signal from *P*1 grows and the ratio A_{LFm}/A_{HFm} decreases, yet remaining greater than unity. In *P*3 spectra taken at $\tau \ge$ 1.1–1.2 μs, one can distinguish the components at frequencies that are roughly multiples of the fundamental low frequency. Taking into account that signals at the frequencies of the basic spectrum peaks cannot reach probe *P*3 from the beam transport channel, one can relate these signals to the development of nonlinear oscillations and the subsequent formation of higher harmonics of the LF SCOs in the immediate vicinity of this probe.

In our opinion, the spectra obtained with probes *P*1–*P*3 (Figs. 2–6) suggest that the near-cathode zone is the basic source of the LF oscillations in the interval $\tau \le$ 1.0–1.2 μs, while the electron beam in the transport channel is responsible for the HF oscillations in this delay range. The oscillations at 100–300 MHz may be associated with the collective motion of electrons in a specific magnetron diode formed by the cathode holder and outer metallic enclosure of the experimental setup (Fig. 1). However, the electron cloud may initially be formed by an electron stream generated by the plasma emitter and aimed away from the cathode. In fact, because of the radial expansion of the cathode plasma, electric fields directed toward the cathode holder must arise at the plasma surface facing away from the transport channel. Arrows in Fig. 1 indicate the direction of the lines of the electric field at the "inner" surface of the cathode plasma. If the velocity of the plasma radial expansion is assumed to be $\approx 5 \times 10^5$ cm/s [1], then a specific plasma emitter–igniter is produced by the time instant the amplitude of the LF oscillations becomes maximal (0.5–0.8 μs). The igniter rises above the cathode surface by 2.5–4.0 mm. Even a low-density electron stream from the igniter is capable of causing secondary emission from the cathode holder. This emission will favor the space charge accumulation at the side surface of the holder. The LF signals detected by probe *P*3 may be due to the bunching of the electron space charge in both the azimuth and the axial directions. The mechanism of space charge accumulation and development of oscillations in cold-cathode magnetron systems was described elsewhere (see, e.g., [5, 6]). Oscillation frequencies in the magnetron diode considered above are difficult to determine exactly. Estimates made with the relationships derived in [6] show that the typical frequencies 100–300 MHz may be observed at the lower order azimuth (n_φ = 1–3) modes and/or at higher order axial modes.

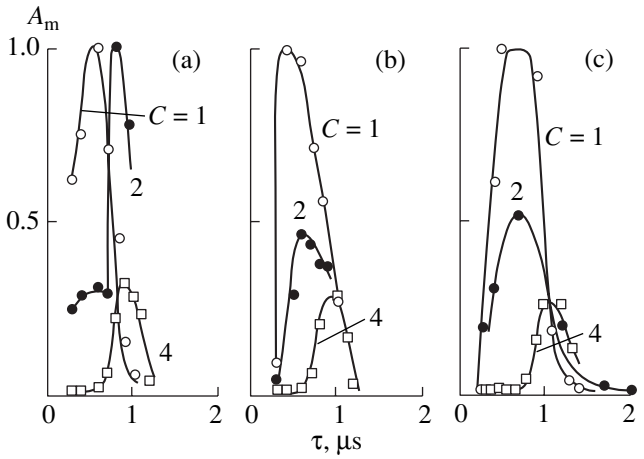


Fig. 4. Maximal amplitude A_m of the components of the SCO spectrum vs. record time τ at various levels of compression *C* in the magnetic field *B* = 1 T. (a, b) Probe *P*1 and (c) probe *P*3. (a, c) LF component of the spectrum and (b) HF component of the spectrum.

Data that can provide a better insight into the SCO nature were obtained when the oscillations were subjected to magnetic field local nonuniformities in the CDMI-2 collector region that were created with solenoid *S*5 (Fig. 1). By decreasing the magnetic field near this solenoid almost to zero (the solenoid produced a field of reverse polarity relative to the field of the main solenoid), we did not affect the collective processes in the REB in the delay interval $\tau >$ 1.0–1.2 μs. However, at the beginning of the accelerating voltage pulse ($\tau <$ 1.0 μs), this nonuniformity markedly diminished the rate of LF and HF oscillation build-up.

Table 2

Probe	*P*1			*P*2			*P*3		
C	1	2	4	1	2	4	1	2	4
A_{LFm}/A_{HFm}	0.6	1.5	1.0	0.2	2	–	115	10	1.3

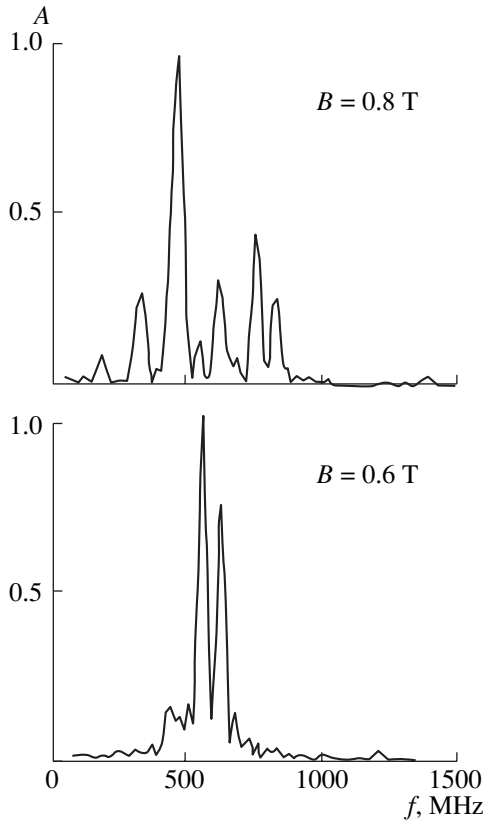


Fig. 5. SCO spectra from probe *P*2 at the time instant τ = 1.1 μs after the onset of the voltage pulse *U* for two values of the magnetic fields.

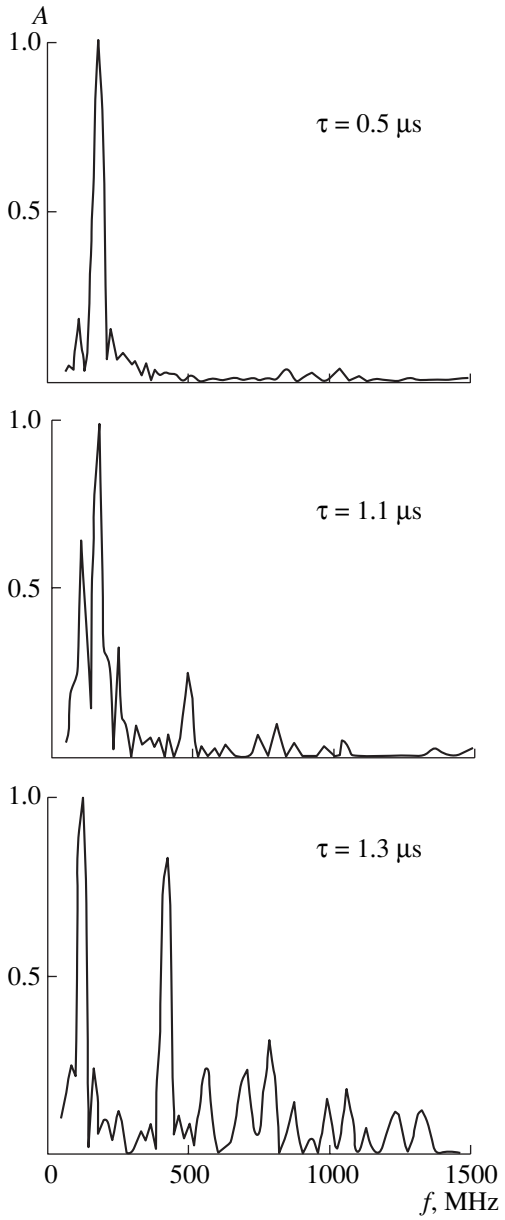


Fig. 6. Typical SCO spectra from probe *P*1 in the magnetic field *B* = 1 T at various time instants after the onset of the voltage pulse.

Such an effect on the HF oscillations seemingly argues against the assumption that they are localized largely in the magnetron diode. In fact, the local magnetic disturbance due to the short (≈15 cm) coil at a large distance (≈100 cm) from the cathode could not directly affect the motion of electrons in the magnetron diode. However, the indirect effect of this disturbance on the collective processes at the cathode, such as a change in the effective REB length, may still take place. Indeed, by introducing the "demagnetizing" magnetic field, we were able to deposit the beam on the transport channel surface nearer to the cathode, thereby cutting the beam effective length roughly by 30%. The effect of this factor on the LF oscillations at the cathode can be explained if the stream of fast secondary electrons from the collector that arises due to collector bombardment is taken into consideration. If the oscillations produce near-cathode space-charge bunches in the stream of electrons that initially move toward the collector, the backward stream of electrons reflected from the collector (or from the beam transport channel) will also be density-modulated. In other words, the system contains feedback. Its effect on the LF oscillations in the magnetron diode will depend on the phase with which the electron bunches fall into the near-cathode zone and, hence, on their travel. If the modulation of the backward electron stream is low, the effect of the feedback is the strongest at the early stage of the LF oscillation amplitude growth, as was observed in our experiments.

The experimentally observed decrease in the rate of HF SCO growth at the beginning of the pulse *U* when the magnetic disturbance at the collector is introduced can be explained in a similar way (that is, through the production of feedback).

Finally, consider the effect of beam compression at the cathode on collective processes in the REB. Figure 4 plots the SCO maximal amplitude A_m against record time τ without compression (C = 1) and with compres-

sion ($C = 2$ and 4). The effect of compression on the signals from probes $P1$ and $P2$ was similar; therefore, the dependence $A_m(\tau)$ for the LF and HF oscillations is shown only for probe $P1$ (Figs. 4a and 4b, respectively). For probe $P3$, we give only the variations of the fundamental LF component of the spectrum (Fig. 4c). The effect of compression is found to be extremely strong. As the compression increases, the oscillation amplitude decreases sharply. Simultaneously, the position of the peaks in the $A_m(\tau)$ curves and the ratio A_{LFm}/A_{HFm} of the peak amplitudes change (Table 2). Typical values of A_{LFm}/A_{HFm} for all three probes at different C are listed in Table 2. The oscillation amplitudes reached the maximal values at high levels of compression. The delay in the oscillation development seems to be associated with the effect of the magnetic disturbance on the cathode plasma motion. In fact, the increase in the compression diminishes the axial velocity of the cathode plasma. This must inevitably lead to a decrease in the HF oscillation growth rate if the oscillation enhancement mechanism put forward in [4] (the advance of the cathode plasma toward the anode) is valid. The rate of LF oscillation build-up declines possibly because the plasma emitter shrinks in a magnetic field that increases toward the collector. Because of this, it is harder for the electrons to penetrate into the magnetron diode (the source of the LF oscillations) from the back side of the collector.

CONCLUSION

Our results can be summarized as follows.

Collective processes taking place in long-pulsed REBs of different geometry and in a wide range of beam-confining magnetic fields are of a general nature.

Collective processes play an essential role in the crossed fields of a magnetron diode formed by a cathode system and the enclosure of a device, as well as in the generation of LF SCOs in magnetically insulated coaxial diodes.

Data obtained indicate that the HF SCOs are related to two-stream instability in the beam transport channel.

SCOs can be suppressed by means of REM compression in the near-cathode zone. The effect of magnetic disturbances produced in this zone, including their effect on the cathode plasma motion, is clarified.

ACKNOWLEDGMENTS

The authors thank A.V. Arkhipov for his assistance in performing the experiments and for valuable discussions.

This work was supported by the Russian Foundation for Basic Research (grant no. 01-02-17081).

REFERENCES

1. S. P. Bugaev, V. I. Kanavets, V. I. Koshelev, and V. A. Cherepenin, *Relativistic Multiwave Microwave Oscillators* (Nauka, Novosibirsk, 1991).
2. L. Yu. Bogdanov and G. G. Sominskiĭ, Zh. Tekh. Fiz. **65** (12), 77 (1995) [Tech. Phys. **40**, 1245 (1995)].
3. L. Yu. Bogdanov, N. V. Dvoretskaya, G. G. Sominskiĭ, and A. Ya. Fabirovskiĭ, Zh. Tekh. Fiz. **67** (8), 83 (1997) [Tech. Phys. **42**, 930 (1997)].
4. L. Yu. Bogdanov, G. G. Sominskiĭ, and A. Ya. Fabirovskiĭ, Zh. Tekh. Fiz. **68** (1), 102 (1998) [Tech. Phys. **43**, 439 (1998)].
5. É. A. Gel'vich and G. G. Sominskiĭ, Elektron. Prom-st, Nos. 7–8, 20 (1981).
6. G. G. Sominskiĭ, in *Lectures on Microwave Electronics: 4th Winter School-Workshop of Engineers* (Saratovskiĭ Gos. Univ., Saratov, 1978), p. 119.

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 7, 2002, pp. 906–909. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 113–116.

ELECTRON AND ION BEAMS, ACCELERATORS

On the Stationary States of Electron Beams in a Drift Space

N. F. Kovalev

Institute of Applied Physics, Russian Academy of Sciences, ul. Ul'yanova 46, Nizhni Novgorod, 603950 Russia

e-mail: kovalev@appl.sci-nnov.ru

Received September 6, 2001; in final form, November 20, 2001

Abstract—The results of a number of studies on virtual cathodes in electron beams subjected to a high magnetic field are shown to be erroneous. The possibility of a virtual cathode occurring in coaxial hollow beams at currents below the limiting value is proved rigorously. © *2002 MAIK "Nauka/Interperiodica".*

(1) Recently, a series of studies devoted to virtual cathodes was published, including the work in this journal [1]. This interesting issue, having a great variety of applications, was also discussed in earlier publications (see, e. g., [2]). Nevertheless, some results of the recent studies and conclusions reached therein seem to be questionable, if not faulty. We believe that the erroneous results can be traced back to inadequate initial statements. In this paper, we consider only one of the doubtful points in [1], concerning the stationary states of highly magnetized electron beams in regular drift channels. It is concluded in [1] that a virtual cathode can exist only if the injection current I_0 exceeds the limiting value $I_{\lim}$ and the current I that passes the drift space after the virtual cathode has been formed is equal to the limiting current:

$$I = I_{\lim}. \tag{1}$$

According to [1], both assertions are valid for plane (Fig. 1a) and thin-walled tubular beams (Fig. 1b). It is well known [2], however, that the dependence

$$I = I(I_0) \tag{2}$$

is essentially different; for example, it may be similar to that shown in Fig. 2 for the planar model of the drift space and nonrelativistic electron energies. As is seen from the diagram of currents in Fig. 2, function (2) is ambiguous; as a result a virtual cathode may also exist when the injection current is below the limiting value ($0.5I_{\lim} < I_0 < I_{\lim}$) and the current that crossed the virtual cathode is necessarily smaller than the limiting current:

$$I \in \left(\frac{1}{8}I_{\lim}, \frac{1}{2}I_{\lim}\right) < I_{\lim}.$$

On the assumption of an adiabatically smooth increase in the injection current $I_0(t)$, the transition from the single-stream state to that with a virtual cathode (dashed arrow in Fig. 2) occurs at $I_0 > I_{\lim}$ due to the Bursian instability. In the opposite case of a slow decline of $I_0(t)$, the virtual cathode disappears also because of the Bursian instability but at another value of the injection current ($I_0 = I_F$):

$$I_F = 0.5I_{\lim} < I_{\lim}. \tag{3}$$

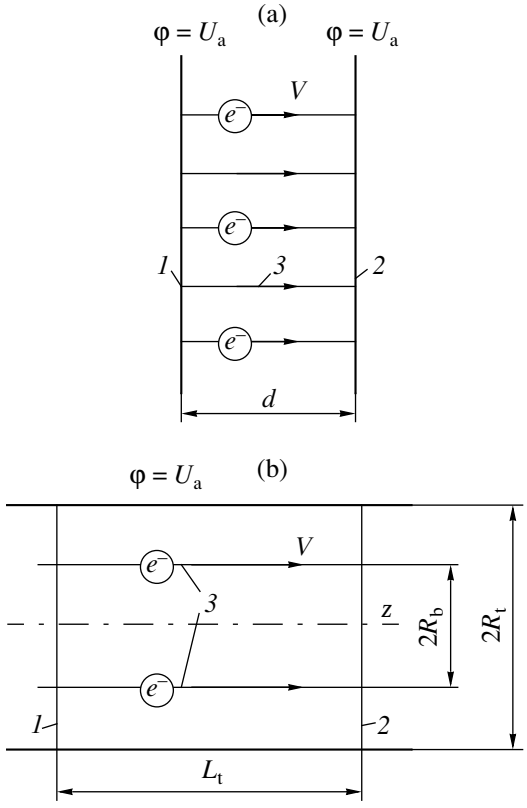


Fig. 1. (a) Planar and (b) coaxial models of the drift space. U_a is the accelerating voltage. (*1*) Electron (e^-) injection end face, (*2*) collector end face, and (*3*) electron beam.

1063-7842/02/4707-0906$22.00

The virtual cathode position is also I_0-dependent; in particular, the cathode appears at a distance $z \approx 0.3d$ from the injection plane ($z = 0$) and disappears at $z = 0.5d$.

(2) The inclusion of relativistic effects does not change the current diagram in Fig. 2 qualitatively. A similar diagram applies also to a highly magnetized cylindrical thin-walled beam in a coaxial tube (Fig. 3) [3]. As the length of the drift tube increases ($L_t/R_t \longrightarrow \infty$), the limiting current asymptotically approaches the value

$$I_{\lim} = \frac{mc^3}{2e\ln\dfrac{R_t}{R_b}}(\gamma_a^{2/3} - 1)^{3/2}, \qquad (4)$$

where

$$\gamma_a = 1 + \frac{eU_a}{mc^2};$$

U_a is the accelerating voltage applied to the electron gun, m and e are the mass and the charge of an electron, c is the velocity of light, R_t is the drift tube radius, and R_b is the thin-walled beam radius.

At $L_t \longrightarrow \infty$, the electron-beam current I behind the virtual cathode also decreases to its asymptotic value (Fig. 3),

$$I = \bar{I}_F < I_{\lim}. \qquad (5)$$

The need for introducing the current $\bar{I}_F$ for coaxial systems was first substantiated in [4] (see also [3]), where it was derived analytically:

$$\bar{I}_F = \frac{mc^3}{2e\ln\dfrac{R_t}{R_b}}(\gamma_a - \gamma_F)\frac{\sqrt{\gamma_F^2 - 1}}{\gamma_F}, \qquad (6)$$

$$\gamma_F = \sqrt{2\gamma_a + 0.25} - 0.5. \qquad (7)$$

(3) For nonrelativistic electron beams, current diagram (2) for the planar model can be constructed analytically by the formulas from [2]. For the coaxial model with a relativistic electron beam, on which much attention was focused in [1], the approximate current diagram was constructed in [3]. However, the existence of virtual cathodes at currents below the limiting value (a fact that was disregarded in [1]) is easy to prove for the cylindrical model as well.

For this purpose, consider the stationary injection of a thin-walled circular electron beam into the tube through its left end face, which must be nontransparent to electric fields (let the end face be covered by a thin metallic foil may) (Fig. 4). A uniform longitudinal magnetic field $\mathbf{H}_0$ is applied to the system. At the end of the drift space, the electrons meet a thin-walled edge collector that is under the cathode (or slightly higher-than-cathode) potential:

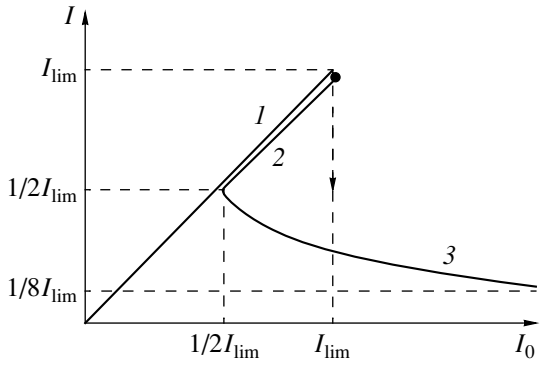


Fig. 2. Current I having passed through the equipotential planar region versus injection current I_0: (*1*) stable and (*2*) unstable single-stream states and (*3*) states with virtual cathode.

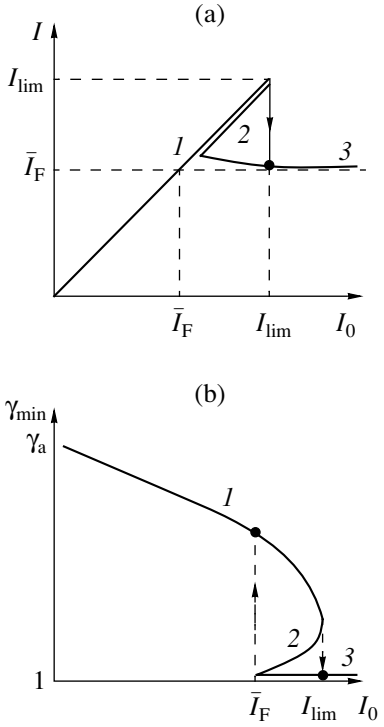


Fig. 3. (a) Current I having passed through the drift space and (b) minimal energy of the electrons versus injection current I_0 for the coaxial model. (*1*)–(*3*) The same as in Fig. 2.

$$\varphi_{col} = +0. \qquad (8)$$

The drift channel length L_t is taken to be much greater than the stopping range of the electrons at both the end face and the collector; as a result the highly uniform state of the beam in the central part of the drift tube is set ($\lim\limits_{L_t \to \infty} d\gamma/dz = 0$).

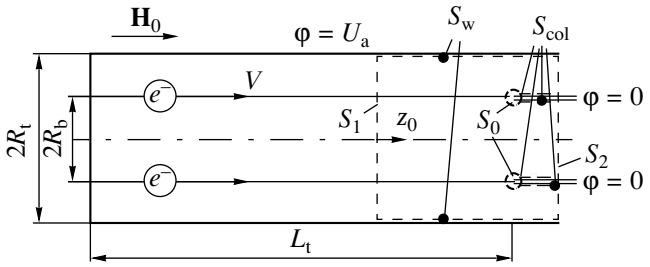


Fig. 4. Coaxial model with a slowing-down collector. $S = S_0 + S_1 + S_2 + S_w + S_{col}$ is the surface of integration.

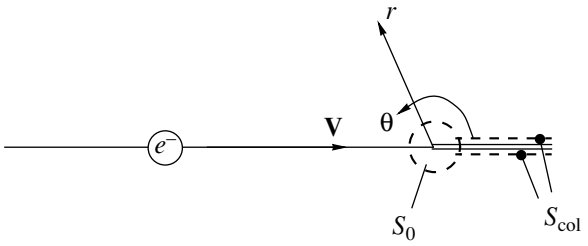


Fig. 5. Local coordinate system (r, θ) at the edge of the collector; S_0 and S_{col} are the surfaces of integration.

As in [4], we apply the law of conservation of the z component of the momentum,

$$\oint\left\{\frac{E^2+H^2}{2}\mathbf{z}_0 - E_z\mathbf{E} - H_z\mathbf{H} + 4\pi\frac{\rho}{l}\gamma m V_z\mathbf{V}\right\}d\mathbf{S} = 0, \quad (9)$$

to the volume bounded by the closed surface ($S = S_0 + S_1 + S_2 + S_w + S_{col}$) shown in Fig. 4 by the dashed line. It is assumed that the electron motion is unidirectional and that there is a spread in their velocities. The transverse component of the electron motion may be disregarded because of the high value of the magnetic field ($|\mathbf{H}_0| \longrightarrow \infty$). In (9), $\mathbf{E}$ and $\mathbf{H}$ are the electric and the magnetic field vectors, ρ is the volume charge density in the beam, $\mathbf{V}$ is the velocity of the electrons, and

$$\gamma = \left(1 - \frac{V^2}{c^2}\right)^{-1/2}$$

is their relativistic factor.

The electric field has a singularity at the edge of the collector; therefore the contribution to integral (9) over the surface S_0 (with its radius to tending to zero) is given by only a part of the field. The field potential, in view of (8), can be conveniently represented in the form

$$\varphi = Ar^{1/2}\sin\frac{\theta}{2}; \quad (10)$$

accordingly,

$$\lim_{r_0\to 0}\oint_{S_0}\{\ldots\}d\mathbf{S} = -\frac{\pi^2}{2}R_b A^2. \quad (11)$$

The local coordinate system (r, θ) used in (10) and (11) is shown in Fig. 5. The cross sections S_1 and S_2 are considered to be within uniform regions, which allows us to calculate all the integrals in (9). With the energy conservation law taken into account, we arrive at the following system of equations:

$$\frac{\pi}{2}R_b\ln\frac{R_t}{R_b}\left(\frac{eA}{mc^2}\right)^2 = (\gamma_a - 1)^2 - (\gamma_a - \gamma)^2$$

$$-4\ln\left(\frac{R_t}{R_b}\right)\frac{eI_0}{mc^3}\sqrt{\gamma^2-1}, \quad (12)$$

$$2\ln\left(\frac{R_t}{R_b}\right)\frac{eI_0}{mc^3} = (\gamma_a - \gamma)\frac{\sqrt{\gamma^2-1}}{\gamma},$$

from which the value of A^2 can be calculated with I_0 known. In (12), γ is the normalized energy of beam electrons at the cross section S_1, i. e., in the uniform part of the drift channel (Fig. 4). With an increase in the injection current I_0, the amplitude squared (A^2) decreases monotonically and vanishes at $I_0 = \bar{I}_F < I_{lim}$ [see (6)]. With further growth of the current ($I_0 > \bar{I}_F$), the amplitude A becomes imaginary, which is physically impossible. Therefore, at high currents ($I_0 > \bar{I}_F$), a virtual cathode due to the Bursian instability appears and reflects a part of the current back to the injection plane.

The foregoing also suggests that the current of the electron beam having passed through the virtual cathode cannot be equal to the limiting value ($I \neq I_{lim}$) as was argued in [1]. In accordance with (12), this current does not exceed $\bar{I}_F$ at $L_t \longrightarrow \infty$.

(4) In [1] and in earlier works (see Refs. in [1]), no mention was made of the current $\bar{I}_F$, which, as has been shown above, is of no less importance than I_{lim} in considering the virtual cathode properties. Nor has attention been given to the uncertain current behavior in the drift space and related hysteresis phenomena typical of the systems in question. In our opinion, all these omissions cast doubt on the results and conclusions in [1] and call for their refinement or even revision.

The above discussion touches upon only the stationary states of highly magnetized electron beams in drift channels. This point is certainly of importance for understanding and systematizing the properties of high-current electron beams and associated devices. It is evident, however, that further advances in the theory of the beams are related to investigation of their dynamic

states. Problems, such as stability, the excitation of various linear and nonlinear, including stochastic, oscillations, the behavior under the action of external forces, etc. should be of primary concern. There is still much we need to learn in this intriguing field.

REFERENCES

1. A. E. Dubinov and I. A. Efimova, Zh. Tekh. Fiz. **71** (6), 80 (2001) [Tech. Phys. **46**, 723 (2001)].
2. L. N. Dobretsov, *Electron and Ion Emission* (GPTI, Moscow, 1950).
3. S. P. Bugaev, N. I. Zaĭtsev, A. A. Kim, *et al.*, in *Proceedings of the II All-Union Workshop "High-Frequency Relativistic Electronics"* (Inst. Prikl. Fiz. Akad. Nauk SSSR, Gorki, 1981), p. 36.
4. A. I. Fedosov, E. A. Litvinov, S. Ya. Belomyttsev, *et al.*, Izv. Vyssh. Uchebn. Zaved., Fiz., No. 10, 134 (1977).

Translated by A. Sidorova

Technical Physics, Vol. 47, No. 7, 2002, pp. 910–914. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 117–121.

SURFACES, ELECTRON AND ION EMISSION

Erosion of the Field Emitter Surface Exposed to Low-Energy Ions

T. I. Mazilova and I. M. Mikhaĭlovskiĭ

Kharkov Institute of Physics and Technology, Ukrainian Scientific Center, Akademicheskaya ul. 1, Kharkov, 61108 Ukraine

Received March 19, 2001; in final form, October 5, 2001

Abstract—The radiation-induced erosion of the tungsten field emitter surface exposed to low-energy ions is studied by field ion microscopy and electron microscopy. During the bombardment, surface atoms are displaced to sites with lower coordination numbers and nanoasperities, generating a locally enhanced electric field, arise on the surface in a jump-like manner, which modifies the characteristics of the emitters. The field evaporation of the asperities produces cavities; hence, the erosion can be described in terms of blistering. Quasi-static surface erosion mechanisms are considered. It is shown that nanoblistering can be related to helium absorption in metal surface layers. *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Radiation-induced surface erosion plays an essential part in discharge-initiating processes and is a basic mechanism of damage to field emitters [1–4]. Interest in surface erosion has quickened in recent years because modern nanotechnology extensively uses the process of tunnel electron emission and, accordingly, highly localized coherent field-ion and electron emitters [5–7]. The evolution of the atomic topography of probes with a submicron radius of curvature of the emitting surfaces (which is inevitable during ion bombardment) is governed largely by radiation-induced diffusion processes [1, 4].

During ion bombardment, the emitters become atomically rough, which is associated with atom collisions on the surface [1–4, 8] and the surface exposure of radiation-induced interstitials [9]. The atomic roughness enhances the mobility of surface atoms. However, the use of this effect for the formation of appropriate nanoemitters that are characterized by low operating voltages and, accordingly, low mean energies of bombarding ions [7, 8] is a challenge, because there is no reliable collisional mechanism of surface erosion. In this work, we study the radiation-induced erosion of tungsten field emitters exposed to low-energy ions by the methods of field ion microscopy and electron microscopy.

EXPERIMENTAL

The study was carried out in a two-chamber field emission microscope where samples and cryogenic sorption ultra-high-vacuum pumps are cooled by liquid hydrogen. In the ionic regime, the operating voltage was varied from 2 to 15 kV; in the field-emission regime, from 0.2 to 1.0 kV. The residual gas pressure in the inner chamber was kept at 10^{-7}–10^{-6} Pa; the pressure of the image gas, at $(2–4) \times 10^{-3}$ Pa. Needle-like samples were made from 99.98%-pure VA-3 tungsten by electrochemical etching. The radius of the tip curvature was 5–40 nm. The samples were heated to temperatures in the interval 21–300 K by passing the current through a supporting ear. The temperature was determined from the threshold field of evaporation [10]. The emitting surface of the field emitters was cleaned and then finished by evaporation at 21 K at a rate of 10^{-3}–10^{-2} nm/s in a positive-polarity field of 5.8×10^{10} V/m.

The parameters of the helium-ion irradiation of the field emitters at a field strength $E = (2–4) \times 10^9$ V/m were determined by the technique described in [1, 11]. The energy of bombarding ions was varied over a wide range, from the ionization energy of helium atoms to an energy corresponding to the operating voltage (24.5 eV). The mean helium ion energy is roughly eEr_0, where e is the charge of an electron and r_0 is the curvature radius of the hemispherical part of the emitter surface. In our experiments, the mean energy was varied between 25 and 200 eV. To prevent vacuum breakdown, resistors limiting the short-circuit current at a level of 10–15 μA were connected in series with the emitters. Once the emitters had been irradiated, the voltage polarity was reversed and the surface erosion was studied in the field ion regime.

EXPERIMENTAL RESULTS

The ion bombardment of needle-like emitters at temperatures of 21–300 K results in an increase in the electron current and, as a rule, in the breakdown of the emitters at moderate mean current densities (10^9–10^{10} A/m^2). At fluences above 5×10^{18} ions/m^2, amplitude jumps by up to 50% are observed against the back-

1063-7842/02/4707-0910$22.00

ground of a monotonically increasing current (Fig. 1). At the instants of the jumps, the emission is highly localized on the surface. Figure 2 shows the field ion micrographs of a needle-like emitter of radius 30 nm (a) before and (b, c) after irradiation by helium ions of mean energy 150 eV at 21 K. The fluences were (b) 3.5×10^{18} and (c) 8.8×10^{18} ions/m^2. After the irradiation, the regular atom arrangement in the surface layer and the ring-shaped configuration of atomic steps typical of atomically smooth surfaces are disturbed. At fluences below 5×10^{18} ions/m^2, emission centers are randomly arranged on the surface (Fig. 2b). Most of them are images of individual atoms. The slight local increase in the contrast is due to the fact that, during the bombardment, these atoms are displaced to sites with lower surface coordination numbers, where the field strength is locally enhanced. As the fluence rises to 5×10^{18}–10^{19} ions/m^2, microasperities 1–5 nm across, in addition to the already existing individual adatoms, appear on the surface (Fig. 2c). Unlike the gradual growth of these asperities due to irradiation-stimulated diffusion, which was observed in [1, 4], some of them grow jumpwise. Their occurrence causes associated jumps of the field-emission current (Fig. 1). Detailed study of the surface atomic structure near the microasperities by field ion microscopy is difficult, since attempts to image them in the ion regime usually fail because of their destruction by the action of electric-field-induced mechanical stresses. However, field evaporation of asperities (domes) under controllable conditions discovers microcavities (Fig. 3). The depth of microcavities, which are discovered by field evaporation or mechanical destruction, is nearly proportional to the diameter of the microasperities, 3–6 nm.

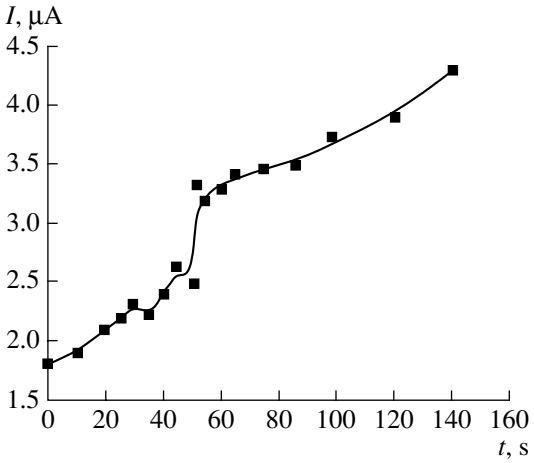


Fig. 1. Time variation of the field-emission current from the tip of the tungsten emitter of initial radius 30 nm at a helium pressure of 4.8×10^{-4} Pa.

DISCUSSION

The atomic topography of the nanoemitter surface exposed to slow ions did not display any traces of cumulative cathode sputtering. This is consistent with the investigations into the radiation resistance of field-emission cathodes with micron and submicron sizes of the emitting surface [1–4]. However, the description of field emitter surface erosion within the model of radiation-induced surface self-diffusion needs considerable refinement. The atomic roughness due to the ion bombardment (that is, the displacement of atoms to low-coordination sites on the surface) observed in this work (Fig. 2b) increases the atom mobility and can be considered to be a mechanism of radiation-induced surface self-diffusion. When the surface of low-voltage emitters is modified, the energy imparted to tungsten surface atoms by irradiation does not exceed 10% of the helium ion energy. In most experiments performed in this work, the transferred energy was much less than the radiation energy threshold for the formation of stable Frenkel pairs (the displacement threshold), heat of sublimation, or the energy of surface self-diffusion all controlling the radiation-induced reconstruction of the surface. Thus, the radiation resistance of the emitters bombarded by ions with a low (subthreshold) mean energy cannot be described in terms of collisional mechanisms of radiation-induced surface erosion.

When the nanoemitters are irradiated by slow ions (i.e., when dynamic channels are virtually absent), relaxation processes in surface layers containing a large amount of embedded helium atoms may play an essential role. Specifically, it is known that blistering (the formation of gas-filled domes on a metal surface) takes place as a result of internal stress relaxation when helium fluences are relatively high. Blistering in metals has been studied at length only for high-energy (10^4–10^6 eV) ions of inert gases [12–14]. However, the jump-like formation of asperities discovered in this work (Fig. 2c) seems to be of a similar nature. Metal damage due to the pressure of helium embedded deep into the material is well studied in relation to reactor building problems [12–14]. At the same time, the behavior of surface helium has been poorly studied. It is known that the interaction between helium and metal atoms is repulsive and the incorporation of helium atoms into the crystal lattice causes local displacements of surrounding matrix atoms. The energy of formation of a helium interstitial depends on its location in the lattice, varying from 5.5 to 7.4 eV in tungsten [12].

The quasi-static behavior of helium atoms in metal surface layers is convenient for description starting from the dependence of the energy of helium atom incorporation into the electron gas. It has been shown [15] that the energy of incorporation into the electron gas of local concentration n at a point r is related to the kinetic energy of conduction electrons as

$$E_{he} = 2\pi L n(r), \quad (1)$$

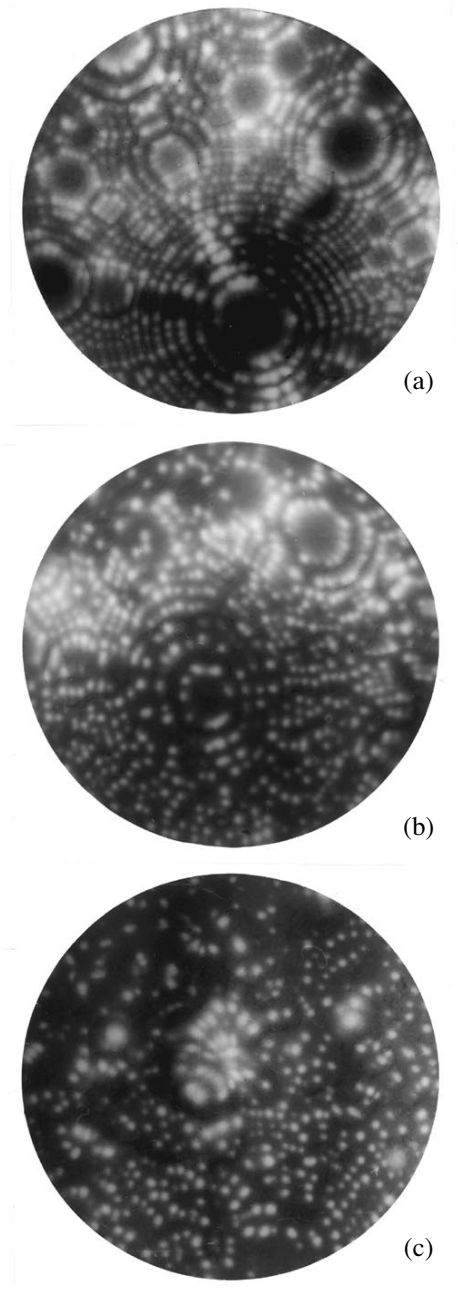


Fig. 2. Field ion images of the field emitter (a) before irradiation and after irradiation with helium ion fluences of (b) 3.5×10^{18} and (c) 8.8×10^{18} ions/m^2.

where L is the length of scattering of an electron by a helium atom (in this relationship, the atomic system of units is used: $\hbar = e = m = 1$).

At the metal surface, the electron concentration drops from the bulk value to zero virtually jumpwise. The width of the transition region is on the order of the Debye screening radius. The local electron concentration at some point k where the atom being analyzed is placed can be determined by summing the electron concentrations n_a produced at this point by individual

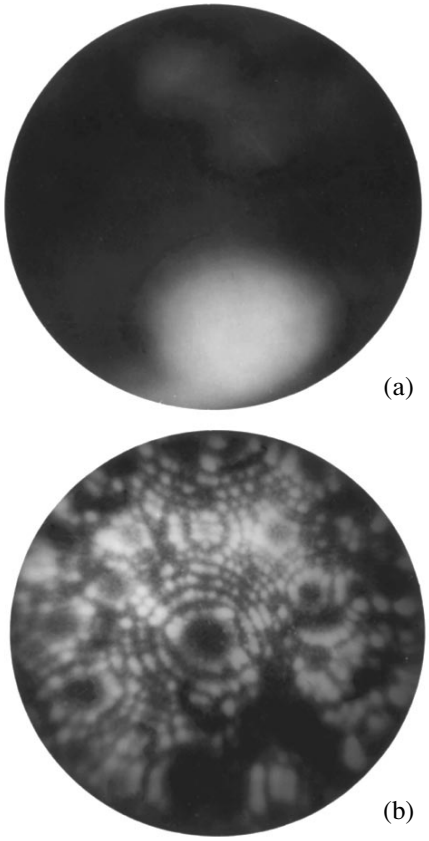


Fig. 3. (a) Electron image of a surface microasperity and (b) field ion image of a microcavity discovered after removal of the microasperity.

atoms [16]:

$$n_k = \sum_j n_a(r_{jk}). \quad (2)$$

When quantitatively analyzing the distribution of the electron concentration near the surface, one must take into account the oscillatory character of the relaxation displacements of atomic planes in the surface layer. Most quantitative data for the surface structure and its interaction with various lattice defects have been obtained in the pair potential approximation. The recent method of many-particle potentials has made it possible to refine considerably the character of atom displacements in surface layers. The relaxation of the surface layers and the distribution of the electron concentration in them can be described in terms of the embedded atom method [16, 17].

To find the electron concentration n_k, we replace the atomwize summation in (2) by the calculation of the overall contribution from layers parallel to the surface to the electron density. An analytical expression for electron concentration can be derived by the method elaborated in [18, 19]. Let $\{\mathbf{X}^j\}$ be a set of sites in the jth atomic layer and $\{\mathbf{G}^j\}$ be a set of vectors in the

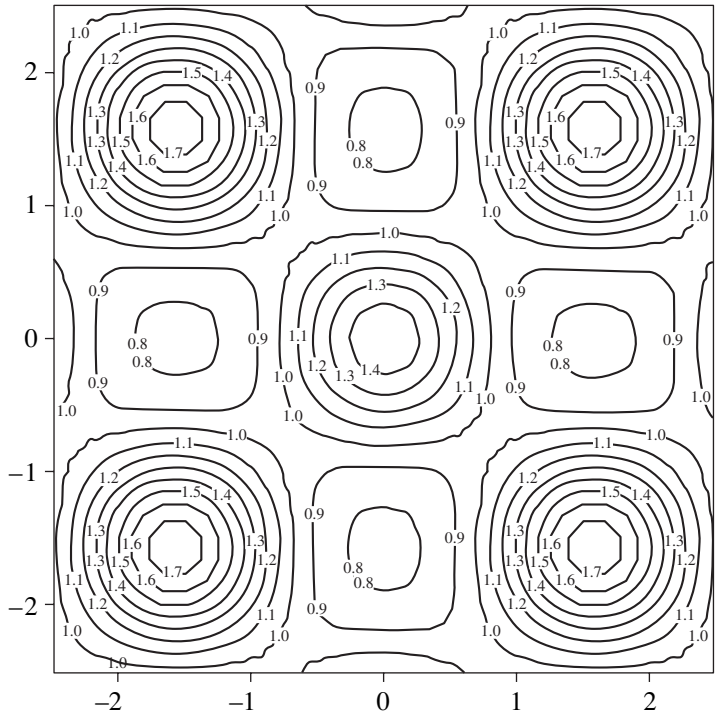


Fig. 4. Electron concentration distribution in the (100) plane 0.233 nm distant from the surface. $n = 1$ corresponds to the average concentration in this section.

reciprocal space. Atoms in the jth plane contribute to the electron density at the point k as

$$n_{jk} = \frac{1}{A_j}\sum_{\mathbf{G}^c}\tilde{n}(\mathbf{G}, z_{jk})\exp(i\mathbf{G}\cdot\mathbf{T}_{jk}). \qquad (3)$$

Here, A_j is the area of a primitive cell of the planar lattice, z_{jk} is the distance from the point k to the jth lattice, $\mathbf{G}$ is a 2D vector of the reciprocal lattice, $\mathbf{T}_{jk}$ is the vector of displacement of the point k from the origin related to the planar lattice, and $\tilde{n}$ is the 2D Fourier transform of the electron concentration.

For an arbitrary 2D wave vector, the Fourier transform of the electron concentration is given by

$$\tilde{n}(\mathbf{G}, z) = 2\pi\int_0^\infty yJ_0(Gy)n[(y^2+z^2)^{1/2}]dy, \qquad (4)$$

where J_0 is the zero-order Bessel function.

Let us introduce the orthogonal coordinate system (x, y) related to the ⟨100⟩ crystallographic directions. For {100} planes, the lattice spacings in the directions x and y are designated as a_x and a_y, respectively ($a_x = a_y = a$), and $|\mathbf{G}| = \sqrt{g_{xl}^2 + g_{ym}^2}$, where $g_{xl} = 2\pi l/a_x$ and $g_{ym} = 2\pi m/a_y$ (here, l and m are the indices of summation in the reciprocal space and a is the 3D lattice spacing). The displacement vector magnitude for the point k is $|\mathbf{T}_{jk}| = \sqrt{T_{xjk}^2 + T_{yjk}^2}$.

To derive the dependence of n_a on r_{jk} in tungsten, we make use of the empirical relationship for the function $n_a(r_{jk})$ obtained in [17]. For distances far exceeding the spacing between the nearest neighbors, we can apply the exponential approximation

$$n_a(r_{jk}) = n_0\exp(-cr_{jk}), \qquad (5)$$

where $n_0 = 417.04$ and $c = 2.2046$.

Substituting (5) and (4) into (3) and subsequent integration yield

$$n_k = 2\pi c n_0\sigma\sum_j\sum_{l,m}\left(\frac{1+t_{lm}|z_{jk}|}{t_{lm}^3}\exp(-t_{lm}|z_{jk}|)\right) \times\cos(g_{xl}T_{xjk})\cos(g_{ym}T_{yjk}), \qquad (6)$$

where $t_{lm} = \sqrt{c^2 + g_{xl}^2 + g_{ym}^2}$.

Near the {100} surface of tungsten single crystals, multilayer atomic relaxation of the oscillatory type is observed: the spacing between the first and second layers diminishes by 5.5%; between the second and third layers, increases by 2.3%; and between the third and fourth layers, diminishes again by 0.6% [17]. Expression (6) shows that the minimal electron concentrations lie on a straight line passing parallel to one of the ⟨100⟩ directions through the centers of octahedral and tetra-

hedral voids in the crystal lattice. The deepest minimum of the electron concentration is observed at a distance of 0.233 nm from the surface. Figure 4 demonstrates the distribution of the electron concentration in the (100) plane 0.233 nm distant from the surface. The concentration is normalized to the average concentration in the given cross section. Its minimum is at the center of an octahedral void; hence, as follows from relationship (1), a helium atom that occupies this position must have the highest binding energy. Thus, {100} layers of the tungsten surface act as traps for helium atoms. To find the absolute binding energy of helium interstitials in surface layers by formula (1), one must perform a self-consistent calculation of the relaxation in the neighborhood of the helium atom. At sufficiently low temperatures, one can expect the appearance of helium clusters near the surface, which cause groups of tungsten atoms to be displaced, i.e., nanoblistering. In this work, nanoblistering on the field emitter surface was observed during low-energy bombardment with fluences on the order of 10^{19} ions/m^2. This value is much less than the critical fluences at which blistering was observed in macrocrystals exposed to fast helium ions [13]. It should be noted that low-energy (subthreshold) ion bombardment does not produce point trapping centers (vacancies and intrinsic interstitials) for helium atoms in the bulk of the microtarget. Therefore, a major fraction of the helium atoms may be captured in the surface layer of the crystal even at low fluences. In this case, the local helium concentration in the surface layer may reach a level close to the saturation limit in bcc metals (0.6 helium atom per metal atom [13]). This is a possible explanation for the low fluences causing nanoblistering

ACKNOWLEDGMENTS

The authors thank V.I. Gerasimenko for valuable discussions.

REFERENCES

1. Zh. I. Dranova and I. M. Mikhaĭlovskiĭ, Fiz. Tverd. Tela (Leningrad) **12**, 132 (1970) [Sov. Phys. Solid State **12**, 104 (1970)].
2. Zh. I. Dranova and I. M. Mikhaĭlovskiĭ, Fiz. Met. Metalloved. **31**, 1108 (1971).
3. J. Y. Cavaille and M. Drechsler, Rev. Phys. Appl. **12**, 1631 (1977).
4. J. Y. Cavaille and M. Drechsler, Surf. Sci. **75**, 342 (1978).
5. G. N. Furseĭ, D. V. Glazanov, L. M. Baskin, *et al.*, Vak. Mikroélektron. **26** (2), 89 (1997).
6. L. M. Baskin, G. G. Vladimirov, and V. N. Shrednik, Poverkhnost, No. 7, 67 (1999).
7. T. I. Mazilova, Zh. Tekh. Fiz. **70** (2), 102 (2000) [Tech. Phys. **45**, 243 (2000)].
8. A. Knoblauch, Ch. Miller, and S. Kalbizer, Nucl. Instrum. Methods Phys. Res. B **139**, 20 (1998).
9. V. I. Gerasimenko, Zh. I. Dranova, and I. M. Mikhaĭlovskiĭ, Fiz. Tverd. Tela (Leningrad) **25**, 2456 (1983) [Sov. Phys. Solid State **25**, 1410 (1983)].
10. M. K. Miller, A. Cerezo, M. G. Hetherington, and G. D. W. Smith, *Atom Probe Field Ion Microscopy* (Clarendon, Oxford, 1996).
11. P. A. Bereznyak and V. V. Slezov, Radiotekh. Élektron. (Moscow) **17**, 354 (1972).
12. A. G. Zaluzhnyĭ, Yu. N. Sokurskiĭ, and V. N. Tebus, *Helium in Reactor Materials* (Énergoatomizdat, Moscow, 1988).
13. *Sputtering by Particle Bombardment*, Ed. by R. Behrisch (Springer-Verlag, New York, 1983; Mir, Moscow, 1986), Vol. II.
14. V. V. Kirsanov, A. L. Suvorov, and Yu. V. Trushin, *Formation Processes of Radiation-Induced Defects in Metals* (Énergoatomizdat, Moscow, 1985).
15. V. I. Gerasimenko, Fiz. Tverd. Tela (Leningrad) **19**, 2862 (1977) [Sov. Phys. Solid State **19**, 1677 (1977)].
16. M. Foiles, M. I. Baskes, and M. S. Daw, Phys. Rev. B **33**, 7983 (1986).
17. S. M. Foiles, Phys. Rev. B **48**, 4287 (1993).
18. A. P. Sutton, Philos. Mag. A **63**, 793 (1991).
19. V. I. Gerasimenko, T. I. Mazilova, and I. M. Mikhaĭlovskiĭ, Fiz. Met. Metalloved. **91** (4), 15 (2001).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 7, 2002, pp. 915–920. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 122–128.

SURFACES, ELECTRON AND ION EMISSION

Emission of Cadmium Excited Ions upon α Particle Bombardment

A. I. Mis'kevich and Zhao Xiaolin

Moscow State Engineering Physics Institute (Technical University), Kashirskoe sh. 31, Moscow, 115409 Russia

e-mail: miskev@htsc.mepi.ru

Received November 27, 2001

Abstract—Ion emission due to the sputtering of metallic cadmium by α particles from ^{238}Pu is studied. Experiments are carried out in helium at different cadmium foil temperatures (from 20 to 280°C) and gas pressures (from 75 to 2200 torr). The sputtering of the metallic cadmium causes the emission of Cd(II) excited ions in the $4d^9 5s^{22}\ D_{3/2,\ 5/2}$ and $4d^{10}6s^2S_{1/2}$ states. Above 160°C, the population of these stares grows exponentially. At a temperature of 240°C and a helium pressure in the chamber of 600 torr, the sputtering ratio of metallic cadmium is found to be 6.26×10^{-14} g per α particle; i.e., one α particle knocks out about 10^8 cadmium atoms from the foil. From spectroscopy and microphotography examinations of the metallic sample surface, a two-step model of ion emission is suggested. The model involves (1) the formation of a high-temperature wedge, which ejects a cadmium droplet, and (2) self-diffusion of displaced cadmium atoms in the droplet toward the surface. © *2002 MAIK "Nauka/Interperiodica".*

The bombardment of a solid surface by energetic particles sputters the target material and causes the emission of secondary particles, such as electrons, ions, and neutral atoms [1, 2]. This effect is widely used in applications. Beams of heavy Ar or Kr ions with energies of 10–14 keV are applied in etching semiconductors [3]. Cathode sputtering by electrons produces metal vapors of density ~10^{14} at/cm^3 in hollow-cathode laser tubes [14]. Such a vapor density is sufficient for the operation of copper-vapor [5] and other metal-vapor lasers [6].

Materials can also be sputtered by nuclear particles. It is known that an uranium fission fragment escaping the fisser entrains different particles, in particular, particles of the unfissioned substrate, in amounts up to ten atoms per fission [7].

Material sputtering is accompanied by photon emission, because some ions and neutrals turn out to be excited [1]. Intense photon emission was observed when Cd, Zn, Pb, Ag, Bi, and Cs films applied on quartz glass were sputtered by the energetic products of the ^{3}He(n, p)^3T nuclear reaction [8, 9]. The intensity of the emission grew upon heating the quartz substrates to 70°C (Cs), 140–240°C (Zn, Pb, Bi, and Cd), and 560°C (Ag). It was also shown [9] that a nuclear-pumped laser that has a cadmium film and is excited by the products of the ^{3}He(n, p)^3T nuclear reaction may have a low lasing threshold at substrate temperatures of 162–206°C when utilizing the $5s^2\ ^2D_{5/2}$–$5p^2P_{3/2}$ transition of a Cd(II) ion (λ = 442 nm).

In this work, we study luminescence arising when a 1-mm-thick cadmium metal foil is sputtered by α particles from 238 plutonium. For many applications, the use of a thick metallic foil may turn out to be more efficient than a thin foil applied on a quartz substrate.

The process was carried out in the setup described in [10]. A cadmium target (a disk of diameter 40 mm) and a planar source (diameter 25 mm) of α particles were placed in a stainless steel chamber (diameter 80 mm, length 80 mm). The target and the source mounted vertically and parallel to each other were 40 mm apart. Their planes were parallel to the chamber axis and to the entrance slit of the monochromator. The target was placed near the chamber axis, but its surface could not be scanned by the photodetector. The source of α particles was near the side wall of the chamber and also

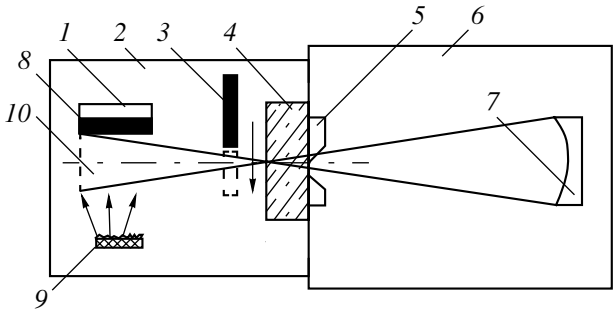


Fig. 1. Experimental setup for the study of luminescence arising when cadmium metal is irradiated by alphas of energy 5.3 MeV. *1*, heater; *2*, chamber; *3*, movable shutter; *4*, quartz window; *5*, entrance slit; *6*, MDR-23 monochromator; *7*, objective lens of monochromator (100 × 100-mm mirror); *8*, cadmium metal (foil); *9*, α source (^{238}Pu); and *10*, light collection region.

1063-7842/02/4707-0915$22.00 © 2002 MAIK "Nauka/Interperiodica"

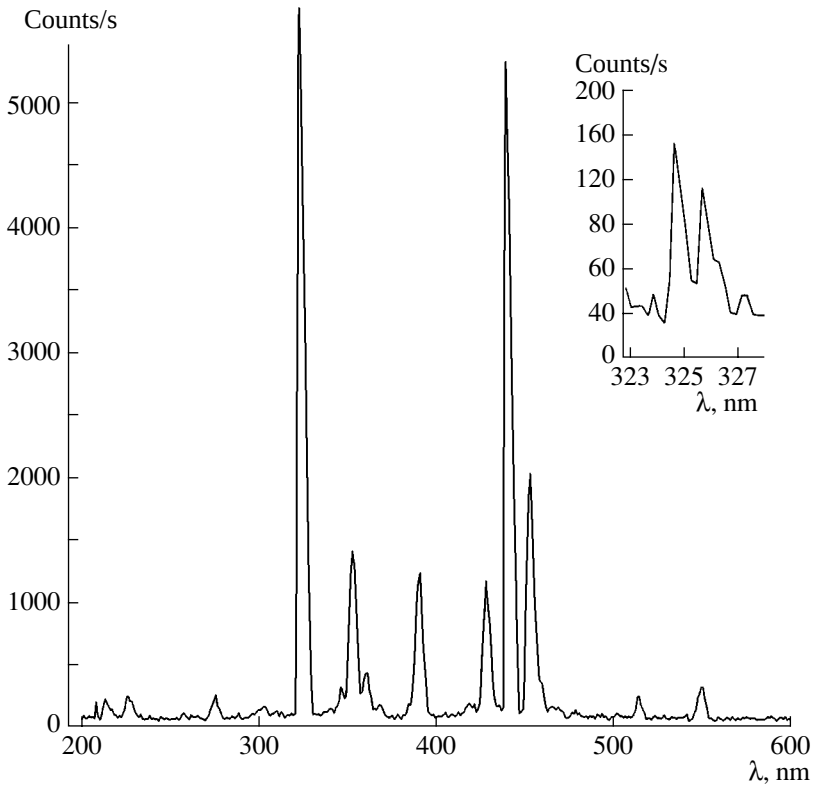


Fig. 2. Luminescence spectrum observed when cadmium metal is sputtered by alphas in helium. The He pressure is 600 torr, the Cd temperature is 240°C, and the monochromator slit width is 2.2 mm. 1200-II diffraction grating is used (1200 lines/mm). The insert shows a part of the spectrum near 325 nm taken with a high resolution.

could not be scanned by the photodetector. The optical emission was incident on the photodetector from a region between the target and the source. The boundaries of the region were offset by ±10 mm from the chamber axis.

A movable shutter making it possible to vary the location of the boundary on the side of the cadmium target was placed between the entrance slit of the monochromator and the target (Fig. 1). The total α particle flux toward the target surface was 3.75×10^7 particles/s. The energy of the particles was defined by the gas pressure in the chamber.

The chamber was equipped with a heater, which raised the cadmium foil temperature from 20 to 280°C. The foil temperature was monitored by a Chromel–Copel thermocouple.

The photon emission was detected by FÉU-100 (wavelengths of 200–800 nm) and FÉU-112 (wavelengths of 600–1200 nm) photomultipliers. Both photomultipliers were cooled to –60°C to suppress the noise level and operated in the photon counting regime. After being evacuated to a high vacuum, the chamber was filled with a gas or a gas mixture to a pressure of 75–2200 torr. During the time of measurement, the gas was continuously purified with sponge titanium heated to 600–700°C to prevent the quenching effect of the impurity gases (H_2O, N_2, CO_2, etc.) desorbed from the chamber walls. Usually, pure He was used, since other inert gases (Ne, Ar, or their mixtures) have a strong quenching effect on the luminescence associated with the Cd(II) $5s^2\ ^2D$ levels. Note that the range of α particles in helium exceeded 4 cm even at a pressure of 2200 torr and they reached the cadmium target in all cases because of their high initial energy (5 MeV).

A typical luminescence spectrum arising when a cadmium foil is irradiated by α particles with He as a buffer gas is shown in Fig. 2. The spectrum contains largely Cd(II) ion lines, which are related to the Cd(II) $5s^2\ ^2D_{5/2,\,3/2}$ Beitler levels at 325, 441.6, 214.4, 226.5, and 353.6 nm directly or through cascade transitions.

Of atom lines, the spectrum contains only Cd(I) resonance lines at wavelengths 326.1 and 228.8 nm; other ion lines are transitions from the Cd(II) $6s^2S_{1/2}$ level. The energy diagram for the cadmium atom and the cadmium ion is displayed in Fig. 3, wherein these transitions are depicted.

The ion line intensity depends on the Cd foil temperature. As the temperature grows, the intensity of all the lines increases sharply but the spectrum as a whole and the intensity ratio between individual lines remain unchanged. The temperature growth of the intensity of the line at 441.6 nm and of the two unresolved lines at

325 + 326.1 nm is shown in Fig. 4 in the logarithm of intensity–reciprocal temperature coordinates. The exponential growth of the luminescence starts at temperatures exceeding 150–160°C and is well approximated by the expression

$$I \sim \exp(-A/kT), \quad (1)$$

where A is the energy of activation of thee process, $k = 8.6153 \times 10^{-5}$ eV/K is the Boltzmann constant, and T is the absolute temperature of the cadmium foil. For the unresolved lines at (325 + 326.1 nm), A = 0.85 eV or 20.41 kJ/mol. For the 441.6-nm line, A = 0.749 eV or 17.98 kJ/mol. All these values agree well (within the accuracy of measurement) with the energy of activation Q of self-diffusion in metallic cadmium, 18.2–19.1 kcal/mol [2]. Thus, one can suppose that the growth of the Cd(II) ion line intensity after the α irradiation is related to the enhanced self-diffusion rate of cadmium atoms in the foil.

The range of alphas in the cadmium foil depends on their energy and the angle of incidence. Therefore, the growth of the Cd line intensity with increasing buffer gas pressure (Table 1) can be associated with a decrease in the alpha range in the substrate, since the particles lose their initial energy when passing through the buffer gas.

We performed tests experiments to find the contribution from the luminescence due to the Cd vapor excitation, since the Cd vapor partial pressure in helium was other than zero at the experimental temperatures (200–260°C) and equaled 10^{-3} torr or less. Under such conditions, the Cd(II) Beitler levels can be occupied via the Penning reactions or via charge exchange with He^+, He_2^+, He*, or He_2^* produced by alphas passing through the gas. Our experiments showed that in the case when alphas could not strike directly with the metallic foil but freely passed through the gas space, the luminescence due to the cadmium vapor excitation at temperatures of 200–260°C was five to six times lower than that due to direct sputtering by alphas.

The spectrum of the luminescence associated with the α particle bombardment depends on the choice of the buffer gas. Figures 5 and 6 demonstrate the spectra obtained when the chamber was filled with pure Ne and pure Ar, respectively. The difference in these spectra is explained by the quenching effect the buffer gas has on the Cd ion emission from the foil. The presence of the excited states of the buffer gas may also be a factor. To estimate the quenching effect of Ne and Ar atoms on the ion level $5s^2\ ^2D$, we measured the intensity of the line at 325 nm (the Cd(II) $5s^2\ ^2D_{3/2}$–$6p$P transition) when a small amount of Ne and Ar was added to the He buffer gas. As the concentration of the impurity gas rises, the luminescence is quenched smoothly in both cases. The rates of quenching of the Cd(II) $5s^2\ ^2D_{3/2}$ level were estimated at 1.8×10^{-10} cm^3/s (Ar) and 10^{-10} cm^3/s (Ne). For the $5s^2\ ^2D_{5/2}$ level, the rates were found to be

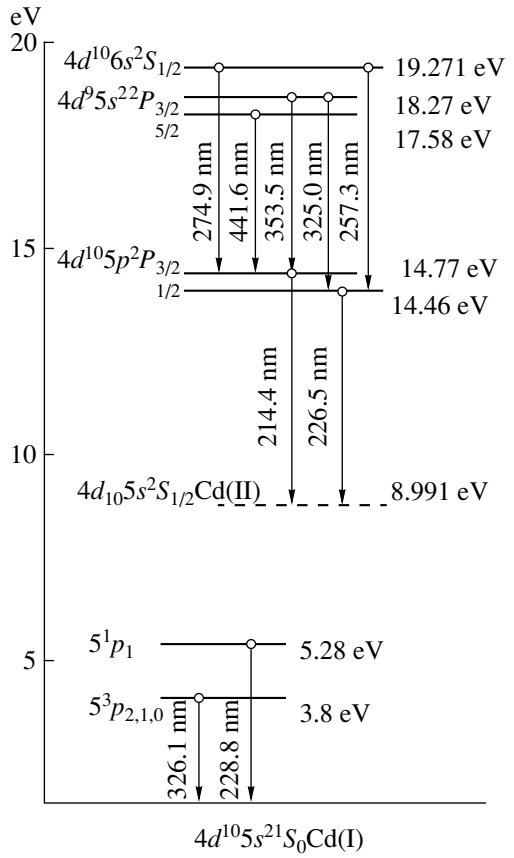


Fig. 3. Energy level diagram for a Cd(I) atom and a Cd(II) ion. The transitions observed are indicated by arrows.

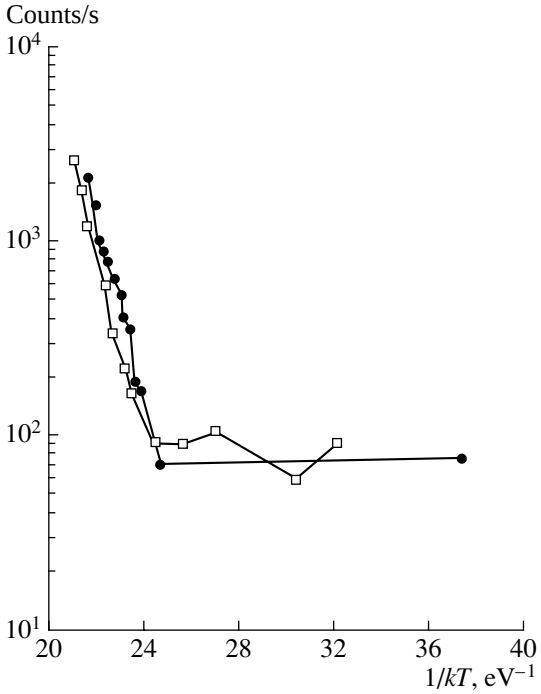


Fig. 4. Temperature dependence of the Cd(II) Beitler line intensities. (□) 441.6 nm, He buffer gas, pressure 300 torr and (●) 325.0 nm (together with the unresolved resonance line at 326.1 nm), He buffer gas, pressure 50 torr.

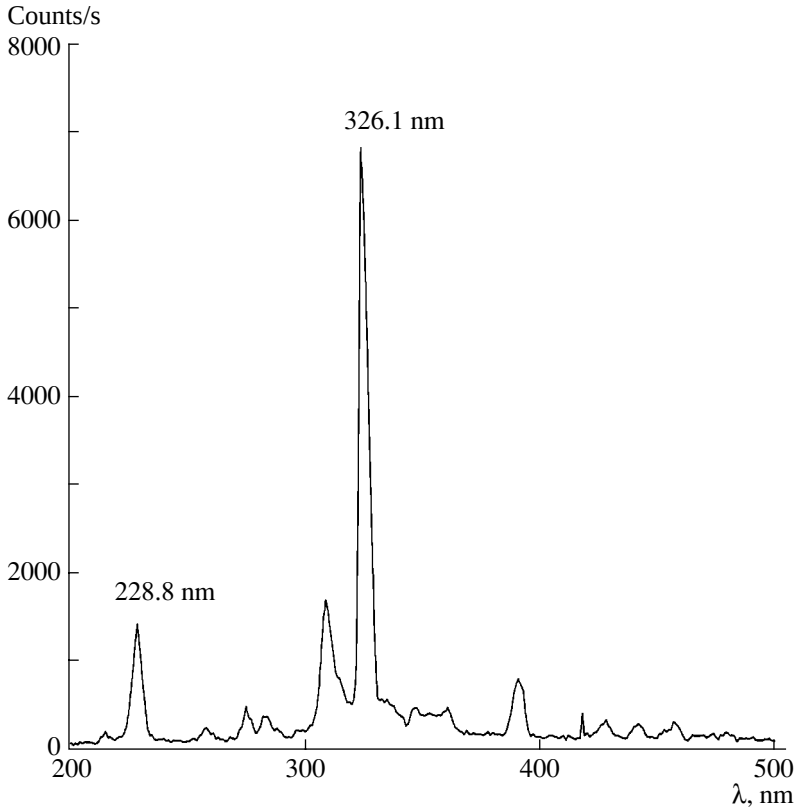


Fig. 5. Luminescence spectrum observed when cadmium metal is sputtered by alphas in neon. The Ne pressure is 360 torr, the Cd temperature is 240°C, and the monochromator slit width is 2.2 mm. A 1200-II diffraction grating is used (1200 lines/mm).

1.3×10^{-10} cm^3/s (Ar) and 10^{-10} cm^3/s (Ne). These findings correlate well with those in [13], where an increase in the threshold and a decrease in the lasing power for the Cd(II) ion line at 441.6 nm were observed with small amounts of Ar or Ne added to He.

The lifetime of the Cd(II) $5s^2\ ^2D$ Beitler levels is long. According to [11, 12], the spontaneous emissivity for the line at 442 nm is 1.4×10^6 s^{-1} (the Cd(II) $5s^2\ ^2D_{5/2}$–$5p\ ^2P_{3/2}$ transition) and 2.85×10^6 s^{-1} for the line at 325 nm (the Cd(II) $5s^2\ ^2D_{3/2}$–$5p\ ^2P_{1/2}$ transition). These levels are used for obtaining the lasing effect, including in lasers pumped by nuclear reaction products [13, 14]. The population efficiency for the Cd(II) ion levels upon sputtering of a metallic foil can be estimated by comparing the Cd spectra with luminescence spectra taken in He–Ne–Ar (He, 760 torr; Ne, 15 torr; and Ar, 7.5 torr) and Ar–Xe–CCl_4 (Ar, 760 torr, Xe, 76 torr; and CCl_4, 3.8 torr) mixtures under identical conditions. With the technique described in [10], we found that the lasing threshold of a nuclear-pumped cadmium laser utilizing the emission of excited ions due to bombardment by nuclear particles is comparable to the lasing threshold of a nuclear-pumped He–Ne–Ar laser, about 10^{15} Td/(cm^2 s) [15]. It was shown that the populations of the Ne(I) $2p_1$ level, Cd(II) $5s^2\ ^2D$ level, and *B* state of a XeCl* molecule are related as 0.011 (Ne) : 1 (Cd) : 18.2 (XeCl*). Thus, the sputtering mechanism of populating Cd(II) ion levels is close to excimer media in nuclear-to-optical energy conversion efficiency.

Table 1. Intensity of Cd(II) ion lines vs. buffer gas pressure. The Cd temperature is 220°C

He pressure, torr	Line intensity, arb. units		
	λ = 325 nm	λ = 442 nm	λ = 275 nm
150	2584	1494	775
300	3170	1780	643
600	3930	1839	842
1900	4443	1658	724

By varying the acceptance solid angle (Fig. 1) with the shutter, we measured the spatial distributions of the density of excited ions emitted at wavelengths of 325, 441.6, and 274.9 nm (Fig. 7). In spite of the great difference in the de-excitation times of these levels (roughly by 300 times), all three curves are nearly coincident with each other. Since the line at 274.9 nm (the Cd(II) $6s\ ^2S_{1/2}$–$5p\ ^2P_{3/2}$ transition) can by no means be related to the Beitler long-lived levels through cascade transitions, this fact can be explained only if a two-step

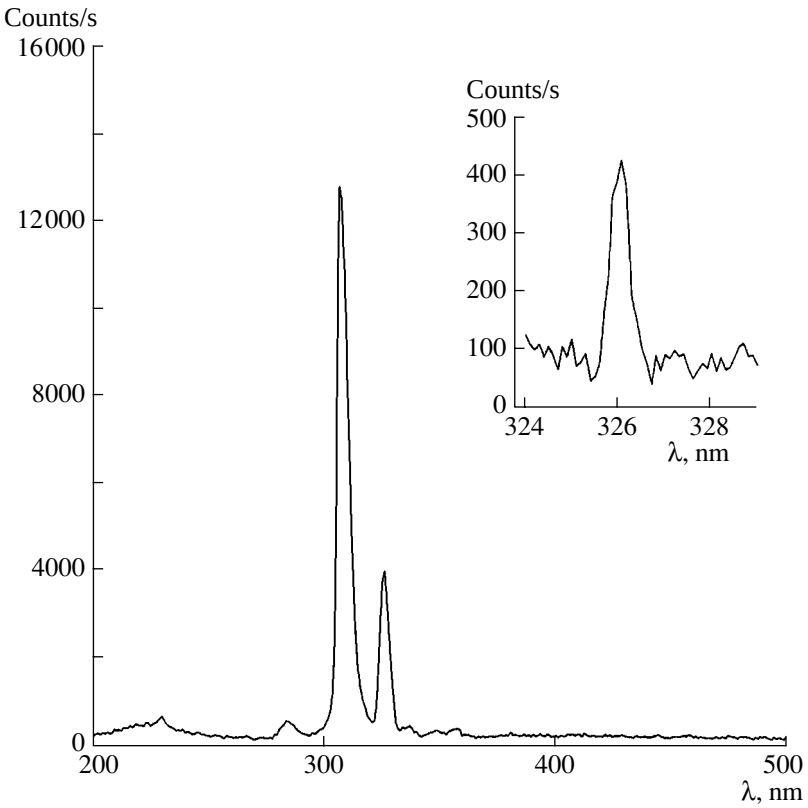


Fig. 6. Luminescence spectrum observed when cadmium metal (foil) is sputtered by alphas in argon. The Ar pressure is 300 torr, the Cd temperature is 240°C, and the monochromator slit width is 2.2 mm. A 1200-II diffraction grating is used (1200 lines/mm). The insert shows a part of the spectrum near 325 nm taken with a high resolution.

ion emission mechanism is assumed. At the first step, an α particle, when being decelerated in the cadmium, interacts with the lattice atoms, displacing some of them into interstices. Some the atoms displaced may be in the excited state. The kinetic energy of the particle rapidly heats a limited region of the lattice, causing the ejection of a heated metal droplet containing cadmium atoms displaced from the lattice sites.

The size and initial velocity of the droplet depend on the particle energy, foil temperature, and the angle of incidence of the particle on the surface.

At the second step, due to the self-diffusion of the atoms displaced, the flying droplet emits excited ions in the states Cd(II) $5s^2\,{}^2D$ (λ = 325 and 442 nm) and Cd(II) $6s\,{}^2S_{1/2}$ (λ = 274.9 and 257.3 nm). The atoms diffuse into the droplet to combine with vacancies. At cadmium temperatures of 160–260°C, the self-diffusion coefficient increases by many orders of magnitude and the displaced atoms readily emerge on the droplet surface.

The fact that the Cd(II) $4d^9 5s^2\,{}^2D$ states of the excited ions prevail can be associated with the decay of the autoionized states of the excited cadmium atoms [16]. The autoionized states of cadmium, which have the configuration Cd(I) $4d^9 5s^2 5p$, appear when an electron from the $4d^{10}$ closed shell is excited. Then, the

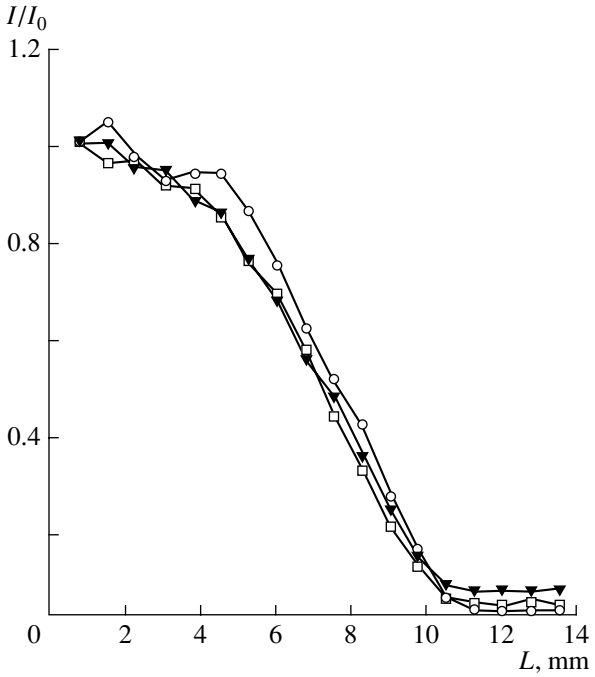


Fig. 7. Spatial distributions of the density of Cd(II) excited ions when cadmium metal is sputtered by alphas in helium. L is the distance from the Cd foil surface. The He pressure is 1900 torr, and the Cd foil temperature is 220°C. λ = (□) 325, (o) 441.6, and (▲) 274.9 nm.

Table 2. Sputtering ratio of cadmium metal exposed to alphas with an energy of 5.3 MeV in a helium atmosphere for various gas pressures

He pressure, torr	Cd temperature, °C	Sputtering ratio, g/particle
0	20	4.43×10^{-18}
150	200	2.55×10^{-14}
300	200	3.57×10^{-14}
600	200	6.26×10^{-14}
1900	200	3.21×10^{-14}

transitions $Cd^*(4d^9 5s^2 5p) \longrightarrow Cd^+(4d^9 5s^2) + e + \Delta E$ will produce the desired ions.

We estimated the size of the droplets by weighing the alpha-melted metal at various energies of alphas. The cadmium film was weighed on an analytical balance before and after it had been irradiated in a chamber filled with helium to various pressures. The irradiation was carried out at a constant temperature of 240°C within a fixed time interval. Since the flux and energy of alphas striking the target were known, one could determine the weight of the metal knocked out by one particle. The results are summarized in Table 2, which lists the sputtering ratio for the cadmium metal at room temperature. The sputtering ratio was measured after two-month exposure *in vacuo.*

The microphotographic examination of the cadmium foil irradiated surface showed that the metal is ejected from the interior of the film. Each of the alphas produces a near-cylindrical ejection region about 1.0 μm across. The longitudinal size of the region is on the order of the particle range. Thus, the droplets have various sizes and temperatures depending on the initial energy of the particle. These parameters govern the self-diffusion rate of the atoms and their escape from the metal.

REFERENCES

1. G. F. Thomas, Surf. Sci. **90**, 381 (1979).
2. *Handbook of Physical Quantities*, Ed. by I. S. Grigoriev and E. Z. Meilikhov (Énergoatomizdat, Moscow, 1991; CRC Press, Boca Raton, 1997), Chap. 25, p. 590.
3. Z. Yu. Gotra, *Reference Book on the Technology of Microelectronic Devices* (Radio i Svyaz', Moscow, 1991), p. 528.
4. B. E. Warner, K. B. Persson, and G. J. Collins, J. Appl. Phys. **50**, 5694 (1979).
5. A. M. Shukhtin, G. A. Fedotov, and V. G. Mishakov, Opt. Spektrosk. **40**, 411 (1976) [Opt. Spectrosc. **40**, 237 (1976)].
6. E. K. Karabut, V. S. Mikhalevskiĭ, V. F. Papakin, and M. F. Sém, Zh. Tekh. Fiz. **39**, 1923 (1969) [Sov. Phys. Tech. Phys. **14**, 1447 (1969)].
7. B. M. Aleksandrov, I. A. Baranov, A. S. Krivokhvatskiĭ, and G. A. Tutin, At. Énerg. **33**, 821 (1972).
8. A. P. Kopaĭ-Gora, A. A. Mavlyutov, A. I. Mis'kevich, and B. S. Salamakha, Opt. Spektrosk. **67**, 526 (1989) [Opt. Spectrosc. **67**, 307 (1989)].
9. A. A. Mavlyutov, A. I. Mis'kevich, and B. S. Salamakha, Laser Phys. **3**, 94 (1993).
10. A. A. Mavlyutov, A. I. Mis'kevich, and Zhao Xiaolin, Prib. Tekh. Éksp., No. 3, 114 (2001).
11. J. I. Collins, J. Appl. Phys. **44**, 4633 (1973).
12. P. P. D'yachenko, Yu. B. Dorofeev, E. D. Poletaev, and M. Z. Tarasko, in *Proceedings of the Conference "Physics of Nuclear-Excited Plasma and Problems of Nuclear-Pumped Lasers," Obninsk, 1992*, Vol. 1, p. 166.
13. É. P. Magda, in *Proceedings of the Conference "Physics of Nuclear-Excited Plasma and Problems of Nuclear Pumped Lasers," Obninsk, 1992*, Vol. 1, p. 65.
14. A. I. Mis'kevich, V. S. Il'yashenko, B. S. Salamakha, *et al.*, Zh. Tekh. Fiz. **52**, 402 (1982) [Sov. Phys. Tech. Phys. **27**, 260 (1982)].
15. A. P. Kopaĭ-Gora, A. A. Mavlyutov, A. I. Mis'kevich, and B. S. Salamakha, Pis'ma Zh. Tekh. Fiz. **16** (11), 23 (1990) [Sov. Tech. Phys. Lett. **16**, 411 (1990)].
16. V. T. Cherepin and M. A. Vasil'ev, *Secondary Ion–Ion Emission in Metals and Alloys* (Kiev, 1975).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 7, 2002, pp. 921–925. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 129–133.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Control of 3D Structure Growth by the Example of a Biocell: 2. From Theory to Experiment

S. L. Grigor'ev

Received November 13, 2001

Abstract—Based on the data obtained by developmental biology, molecular biology, and biochemistry, growth processes in living cells are described. These processes are discussed in terms of radiation chemistry, integrated optics, and nanooptics. Approaches to experimental fiber growth for the purposes of informatics and nanomechanics are proposed. *© 2002 MAIK "Nauka/Interperiodica".*

Membranes and fibers are the basic structures of living organisms, since they are involved in energy supply, locomotion, information processing, and other basic functions. Biochemists and microbiologists have developed technologies for producing artificial nanomembranes and membrane particles, such as micelles, liposomes, and vesicles, ranging in size from 250 Å to 300 μm [1]. Data obtained in studies on these particles explain many mechanisms of the growth and functioning of cellular and noncellular organisms (viruses) [3]. Lipids are the basic components of membranes.

Amphiphilic molecules (including lipids) in aqueous solutions form spontaneously ordered monolayers on a water surface and bilayered liposomes (vesicles) in a water column. The kind of structures for a given substance (lipid) depends mainly on its concentration in water and on the water temperature. Figure 1 schematically shows the formation of an artificial lipid bilayer structurally similar to the membrane of a living cell. The properties and structure of lipid vesicle membranes in aqueous solutions can be modified by adding other types of lipids, cholesterol, proteins, calcium ions, etc. to these solutions.

In living organisms, the growth of cell membranes depends both on external factors (such as the composition of the extracellular medium and the properties of neighboring cells) and on the functional properties of intracellular organelles and substances (nucleus, ribosomes, cytoplasmic proteins, etc.). Of the many phenomena governing growth and evolutionary processes in a cell, we will concentrate on the formation of axons, the long output fibers of nerve cells (neurons), bearing in mind the growth of artificial neural networks.

One of the basic ideas of neurobiology is that an axon is a long fibrous outgrowth of the nerve cell body: in neurons with a diameter of about 100 μm, the axon length may exceed 1 m [2].

Neurons differentiate from neuroblasts, which are rounded cells whose neural fate is specified at early stages of embryonic development (Fig. 2). These cells subsequently elongate due to the deployment of microtubules and actin filaments, which are cytoskeletal pro-

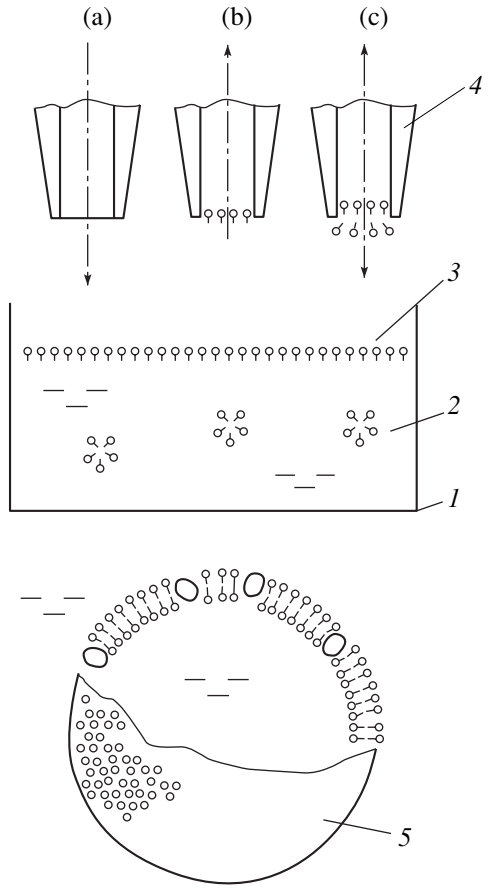


Fig. 1. Production of a lipid bilayer: *1*, vessel with *2*, solvent (water); *3*, surface film with a lipid monolayer; *4*, pipette (a) before immersion, (b) after first immersion, and (c) after the second immersion into the solvent; *5*, vesicle with protein molecules incorporated into it (if these molecules are present in the solution).

1063-7842/02/4707-0921$22.00

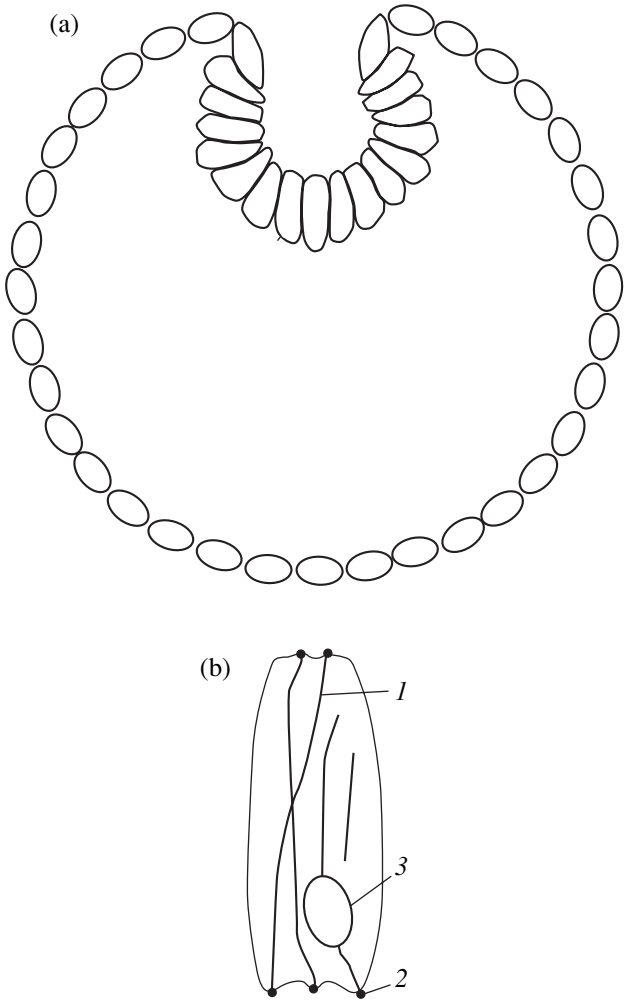


Fig. 2. Formation and structure of neuroblasts. (a) Formation of neuroblasts in the ectoderm; (b) neuroblast structure: *1*, microtubules; *2*, membrane-bound protein microfilaments; and *3*, cell nucleus.

tein structures connected to transmembrane linker proteins of the integrin family (Figs. 2, 6). At the point of connection, an extension (spike) is formed on the membrane, facilitating the growth of a fiber out of a cell in the corresponding direction. The growth processes interconnect neurons, and the development of this system continues virtually throughout a human's life. The most general (thermodynamic) theory explains the formation of the branching cell network through the resulting decrease in its free energy, which is due primarily to cell–cell adhesion forces.

To construct a more complete model of fiber growth, let us also consider the growth of flagella in some bacteria, which is similar to the growth of axons [2, vol. 1] (Fig. 3). A flagellum is the simplest biological machine providing bacterial motility and some other functions.

The growth of flagella proceeds in several stages: (1) the synthesis of tubulin, their basic construction material, on an RNA template by the protein-synthesizing cell machinery; (2) tubulin polymerization and

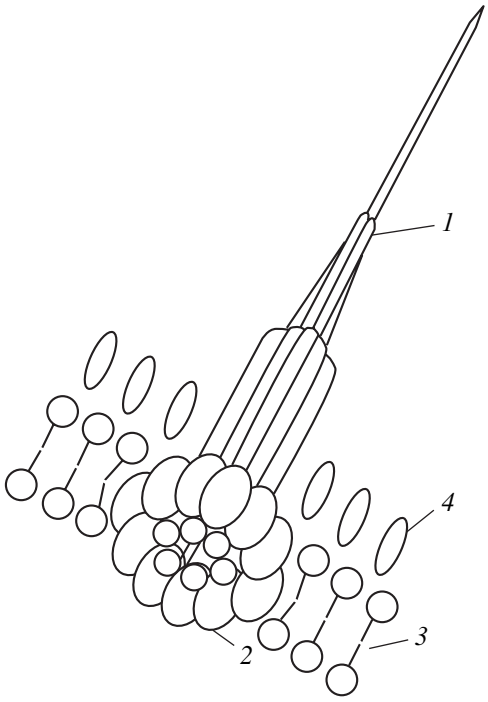


Fig. 3. Schematic representation of bacterial flagellum: *1*, tubulin microtubules; *2*, myosin globules; *3*, lipid bilayer (inner cell membrane); and *4*, glycogen layer (outer cell membrane).

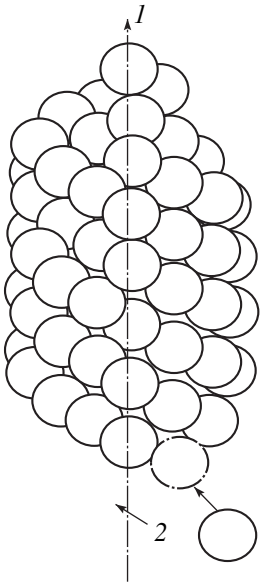


Fig. 4. Microtubule assembling by polymerization of protein globules (macromolecules): *1*, growth direction; *2*, local electric vector of the wave stimulating the process.

arrangement into microtubules 10–20 Å in diameter; (3) the stochastic formation of growth points (outward extensions, or spikes) in the membrane; and (4) flagellum formation from the spike in which microtubules grow most steadily. Figure 4 shows microtubule formation via protein polymerization in the flagellum.

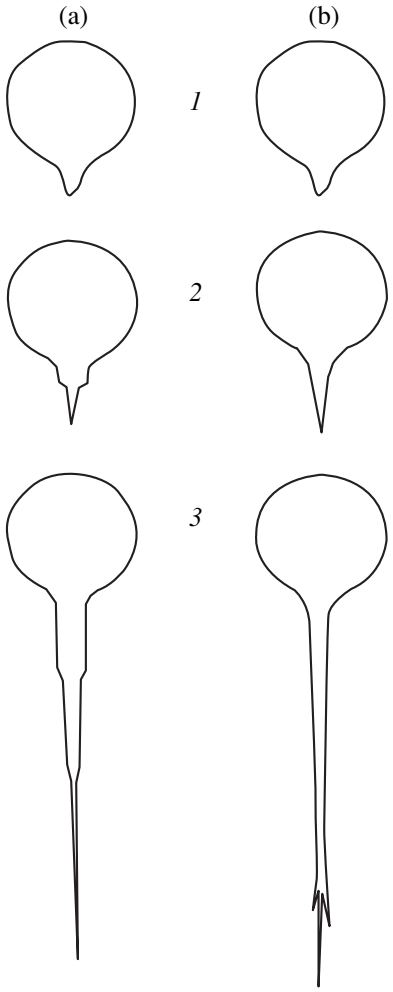


Fig. 5. (*1–3*) Growth stages of (a) flaggellum and (b) axon.

The mechanisms (stages) of axon and flagellum growth are compared in Fig. 5. The initial stages are similar, but subsequently some differences appear: the flagellum grows as a multilayer tube; the axon, as a single-layer branching tube. However, the growing tips in both processes remain similar in the following aspects: (1) the pointed growth cone at the tip of each process is formed due to membrane deformation by protein microtubules; (2) "building blocks" for the growth are supplied to the growth cone from the cell body; (3) the membrane of the growth cone contains anchor and receptor proteins that (which) provide its adhesion to the extracellular matrix and define the growth direction (via the gradient of adhesion forces or chemotaxis); and (4) building blocks are synthesized in the cell body under the genetic control of the cell nucleus (DNA–RNA–protein); this process is corrected by feedback signals coming from the growth cone proteins.

Figure 6 shows a general scheme of the "construction process." In the simplest terms, this process can be described as follows:

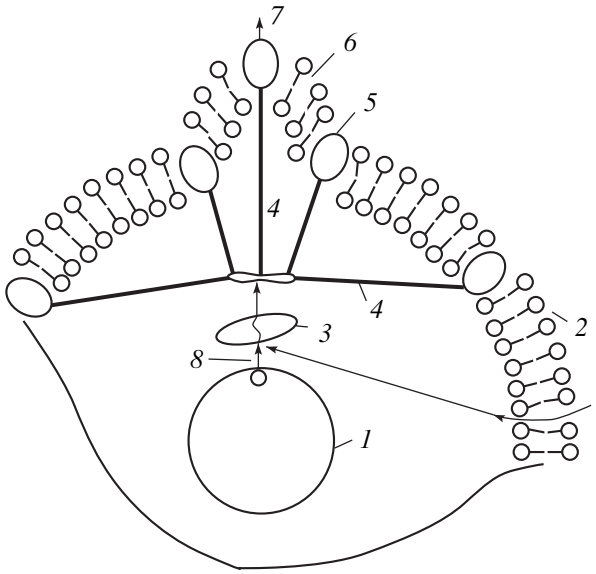


Fig. 6. Transport and synthesis of substances in the cell (the exocytosis path) and transport of electromagnetic energy in the cell (radiotaxis; feedback omitted): *1*, cell nucleus containing DNA molecules; *2*, cell membrane; *3*, ribosome; *4*, actin microtubules (cytoskeleton); *5*, membrane proteins (integrins); *6*, growing spike giving rise to a fiber; *7*, wave output in the direction of absorption; and *8*, control wave (photon).

(1) Nutrients (building blocks) enter the cell through the membrane by diffusion. Inside the cell, they are distributed over various organelles, including ribosomes.

(2) Ribosomes use these nutrients (amino acids) for synthesizing protein molecules (e.g., actin and tubulin). This synthesis occurs on RNA templates and is hence controlled by the cell nucleus.

(3) Actin and tubulin molecules are transported to the "assembling site," i.e., to the spike that will give rise to an axon or a flagellum.

(4) Actin and tubulin molecules are polymerized to form the extension of microtubules and filaments initially present on the as yet smooth membrane, using them as a substrate.

(5) New microtubules and filaments grow, extending the membrane, and so on.

Now, using the model of radiation chemical processes in a cell [3], we consider the role of sub-mm electromagnetic waves in axon growth and, more specifically, in microtubule polymerization. The possible role of electromagnetic waves in the feedback between membrane protein and nuclear DNA and RNA molecules was discussed in general terms elsewhere [3].

Figure 4 shows the scheme of actin polymerization resulting in microtubule formation [2, vol. 2]. Unpolymerized actin is a globular monomer with a molecular weight of about 40 000. The polypeptide chain of a monomer carries one molecule of bound guanosine triphosphate (GTP), which loses one phosphate residue

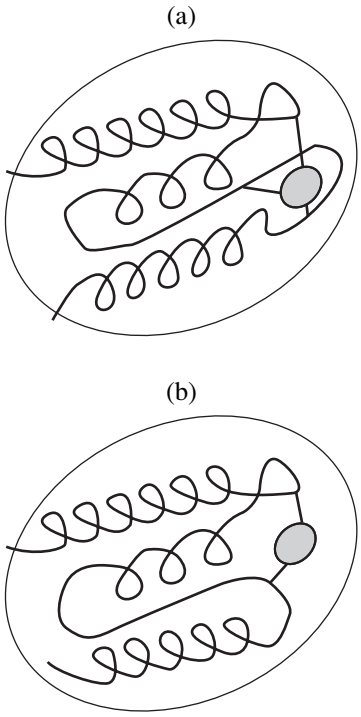


Fig. 7. Actin globules: (a) the dark globule contains a GTP molecule and may be involved in polymerization; (b) the dark globule contains a GDP molecule and is not involved in microtubule formation.

upon hydrolysis to form guanosine diphosphate (GDP). Two forms of actin globules are schematically shown in Fig. 7. A change in the number of bonds between this molecule and protein loops changes the shape and ability of the loops to polymerize and form a microtubule. This supposedly occurs under the effect of a certain external mitogenic signal of chemical nature (i.e., a special substance) [2, vol. 3]. However, a quantum of electromagnetic radiation can also cause such a change in the protein conformation [4]. This process has been studied in detail for the protein systems involved in photosynthesis [5]. The same applies to the system of light detection in the vertebrate eye, where G protein called transducin (which also carries a GTP molecule) relates the reception of a quantum of light by rhodopsin to an electric change in the photoreceptor cell. Rhodopsin, or visual pigment, is a complex protein consisting of opsin (a glycoprotein–lipid complex) and a small molecule of retinene, vitamin A aldehyde. According to [1], rhodopsin converts a photon with an energy of about 1 eV to a quantum of the rotational transition of atoms (with an energy of about 0.01 eV) in the polar groups of its retinene-bound glycoprotein component. It is also known that EHF radiation stimulates ATP synthesis in microorganisms and causes the conformational conversion of proteins to an active form in aqueous solutions [6, 7]. Apparently, quanta with the rotational transition energy strongly affect the activation energy of GTP–GDP transformation by hydrolysis

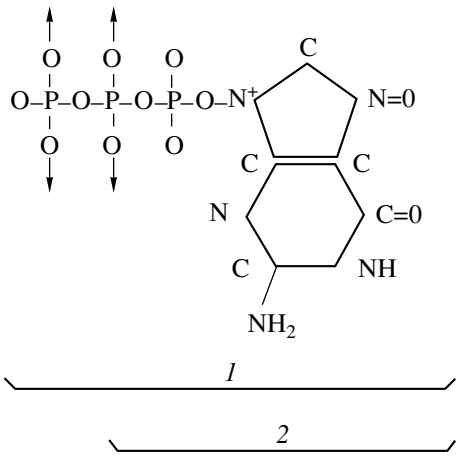


Fig. 8. (*1*) GTP and (*2*) GDP molecules. Arrows show the points of attachment to protein.

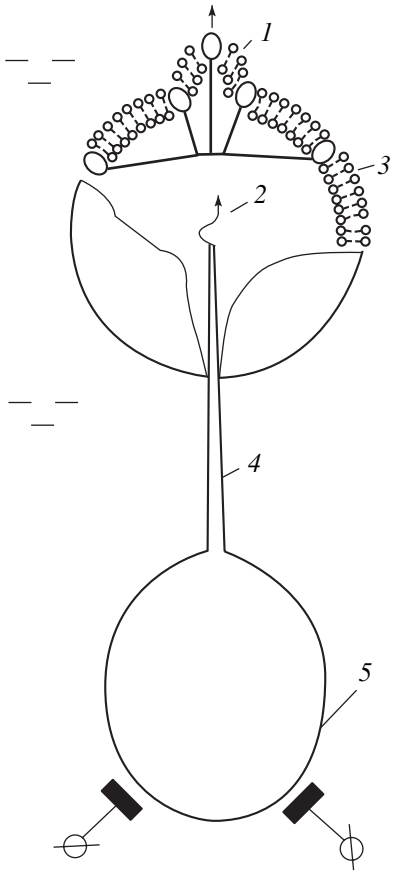


Fig. 9. Scheme of experiment: *1*, growing spike giving rise to a fiber; *2*, wave (phonon or soliton); *3*, vesicle in solution; *4*, tubular waveguide; and *5*, sub-mm wave pulsed maser.

(Fig. 8), which is similar to the catalytic effect of enzymes (hydrolases) [8].

Taking into account the aforesaid, let us discuss the possibility of simulating the electroradiochemical aspect of fiber growth from a rounded cell (vesicle). To

this end, it is necessary to maximally simplify the scheme shown in Fig. 6. For example, we can consider a simple bilayered lipid vesicle (Fig. 9) containing membrane protein integrin connected to actin microtubules which are in equilibrium with a spherical vesicle. Assume that some amount of unpolymerized actin and lipids is in the intracellular (intravesicular) space and that the vesicle is in a solution well permeable to sub-mm waves (i.e., in water with an appropriate pH). Such initial conditions can be provided by biochemical methods (see above). Let us now introduce a thin polymer waveguide made of a polar dielectric [1] into the cell. The waveguide is structurally similar to the actin filament and is connected to a pulsed maser (a source of sub-mm radiation) [4]. This device is intended to functionally substitute for the nucleus, ribosomes, and other cell organelles. Let the frequency and intensity of the electromagnetic radiation provide actin polymerization (in the range 50–200 cm^{-1}). This will cause microtubule growth in the direction of emission from the waveguide and change the shape of the smooth membrane; i.e., a spike will form. Membrane deformation will increase its permeability to lipids dissolved in the intracellular liquid, allowing new molecules to incorporate into the membrane and participate in cell formation.

Apparently, the problem has only been formulated in this paper, but formulating the problem is no less important than solving it.

CONCLUSIONS

(1) Quantized electromagnetic pulses with the rotational transition energy can stimulate the growth (polymerization) of microtubules and the formation of fiberlike structures in bioorganic systems.

(2) Combining the methods of biochemistry, microbiology, nanotechnology, and nanooptics, researchers can experimentally simulate the growth of fibers in living cells.

REFERENCES

1. R. B. Gennis, *Biomembranes: Molecular Structure and Function* (Springer-Verlag, New York, 1989; Mir, Moscow, 1997).
2. S. F. Gilbert, *Developmental Biology* (Sinauer Associates, Sunderland, 1988; Mir, Moscow, 1993), Vols. 1–3.
3. S. L. Grigor'ev, Zh. Tekh. Fiz. **71** (11), 110 (2001) [Tech. Phys. **46**, 1457 (2001)].
4. L. T. Bugaenko, M. G. Kuz'min, and L. S. Polak, *High Energy Chemistry* (Khimiya, Moscow, 1988).
5. V. M. Stepanov, *Molecular Biology. Structure and Functions of Albumens* (Vysshaya Shkola, Moscow, 1996).
6. S. L. Grigor'ev, Zh. Tekh. Fiz. **72** (2), 16 (2002) [Tech. Phys. **47**, 156 (2002)].
7. O. V. Betskiĭ and N. D. Devyatkov, Radiotekhnika, No. 9, 4 (1996).
8. B. D. Berezin and D. B. Berezin, *Course of Modern Organic Chemistry* (Vysshaya Shkola, Moscow, 1999).

Translated by N. Gorgolyuk

Technical Physics, Vol. 47, No. 7, 2002, pp. 926–928. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 134–136.

BRIEF COMMUNICATIONS

Electric Strength of the Accelerating Gap of a Plasma Electron Source at Rough Vacuum

V. A. Burdovitsin, M. N. Kuzemchenko, and E. M. Oks

Tomsk State University of Control Systems and Radioelectronics, Tomsk, 634050 Russia

e-mail: burdov@fet.tusur.ru

Received August 14, 2001; in final form, November 10, 2001

Abstract—Conditions for the electric breakdown of the accelerating gap of a plasma electron source are determined. It is shown that, depending on the gas pressure and the size of openings in the emission electrode, two types of breakdown can occur. One of them is due to the ignition of a low-voltage discharge between the electrodes of the accelerating system, and the other one is caused by switching of the main discharge from the anode to the accelerating electrode. © *2002 MAIK "Nauka/Interperiodica".*

The production of emitting medium in plasma electron sources (PESs) [1] requires either puffing of the working gas (in systems based on low-pressure discharges [2]) or the use of a cathode material vapor as a plasma-forming medium (in vacuum-arc sources [3]). In both cases, the gas pressure in the accelerating gap is elevated and, consequently, the electric strength of the gap decreases. This problem is of special importance for electron sources operating at rough vacuum [4, 5]. The aim of this study is to investigate the conditions and cause for the electric breakdown of the PES accelerating gap at elevated gas pressures.

In the experiments, we used an electron source based on a hollow cathode discharge and a plane-parallel accelerating system (Fig. 1), which are described in detail in [5]. To enable visualization, the high-voltage insulator was made of Plexiglass. Electric breakdown of the gap was initiated by an increase in either the accelerating voltage or the discharge current. Experiments were carried at different gas pressures and electrode configurations. The instant of breakdown was detected by the sharp increase in the current I_e through a high-voltage rectifier connected to the accelerating gap and the onset of a bright glow between the electrodes of the accelerating system. The dependences obtained are the result of statistical treatment of multiple measurements.

Two types of breakdown of the accelerating gap were observed. The first type of breakdown, conventionally called "interelectrode breakdown," occurs between the accelerating electrode and the outer surface of the anode. In this case, a sharp increase in the current is observed only in the high-voltage rectifier circuit. The second type of breakdown occurs between the discharge plasma and the accelerating electrode. The breakdown current, which flows in the accelerating electrode–plasma–hollow cathode circuit, is recorded in both the high-voltage rectifier circuit and the power supply circuit. This type of breakdown, further called "plasma breakdown," is preceded by the onset of low-frequency (0.1–10 kHz) discharge-current fluctuations.

The type of breakdown is mainly determined by the size h of the PES emission openings and the gas pressure in the accelerating gap. Relatively small openings and low pressure result in interelectrode breakdown, whereas large openings and high pressure lead to plasma breakdown.

In our experiments, a 0.5-mm-thick Ti plate with openings 0.6 mm in diameter was used as an emission electrode. The first type of breakdown was forced by increasing the voltage across the accelerating gap at a constant emission current. Experiments with different

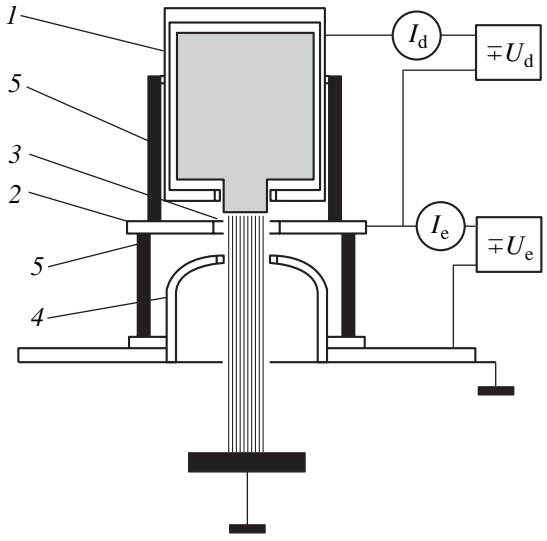


Fig. 1. Schematic of the electron source: (*1*) hollow cathode, (*2*) anode, (*3*) grid or perforated plate, (*4*) extractor, and (*5*) insulator.

1063-7842/02/4707-0926$22.00 © 2002 MAIK "Nauka/Interperiodica"

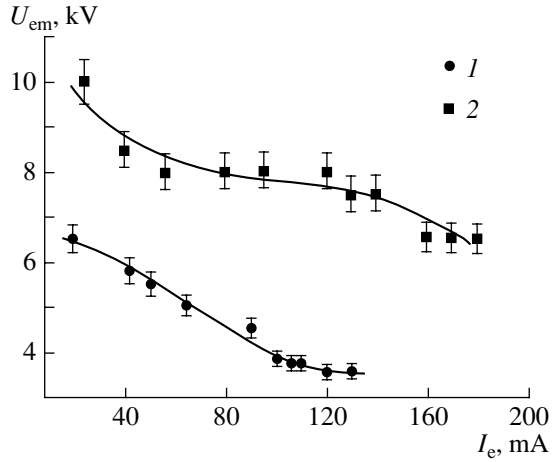


Fig. 2. Maximum accelerating voltage U_{em} vs. emission current I_e for p = (*1*) 10 and (*2*) 9 Pa.

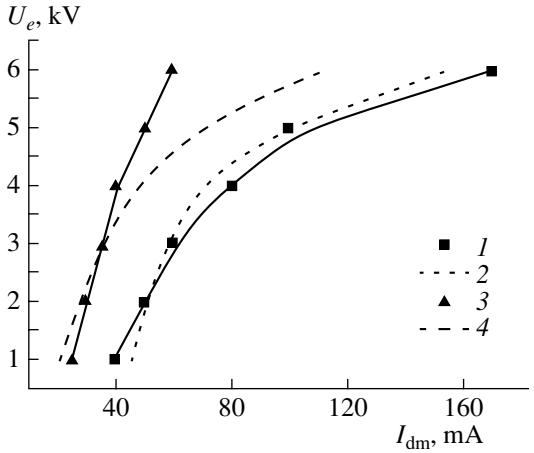


Fig. 3. The greatest possible discharge current I_{dm} vs. accelerating voltage U_e for p = (*1*, *2*) 14.6 and (*3*, *4*) 15.3 Pa. Curves *1* and *3* show the experimental data, and curves *2* and *4* show the results of calculations.

anode electrodes (grids or plates with a different number of openings) show that it is the emission current that affects breakdown. Figure 2 presents the dependences of the accelerating-gap breakdown voltage U_{em} on the electron emission current I_e (recorded just before breakdown) for different gas pressures. The experimental results allow us to conclude that the electron beam provokes the breakdown of the accelerating gap. When investigating the interelectrode breakdown, it was also established that the increase in the emission current over a certain threshold value increased the breakdown voltage U_{em}, which was first mentioned in [4].

The second type of breakdown occurred when the discharge current was increased at a constant voltage across the accelerating gap. The maximum discharge current I_{dm} versus the accelerating voltage U_e is plotted in Fig. 3. It is seen that the behavior of the current–voltage characteristic is opposite to that for the first type of breakdown. The increase in the anode transparency by increasing the number of openings results in an increase in the current I_{dm}. The experimental data indicate that, in this case, it is the discharge current (rather than the emission current) that is responsible for breakdown.

The different operating conditions and different behaviors of these types of breakdown point to their different physical mechanisms. It should be taken into account that, if there is no emission (i.e., the discharge current is $I_d = 0$) in the case of interelectrode breakdown, then applying the accelerating voltage to the accelerating gap results in ignition of the well-known high-voltage glow discharge with a current of several milliamperes. Actually, electric breakdown corresponds to a transition from the high-voltage to low-voltage discharge. This transition is caused by an additional ionizer (in our case, an electron beam).

The data on the second type of breakdown can be explained based on the model proposed in [4]. This model assumes that the breakdown of the accelerating gap occurs when the plasma from the discharge region penetrates into the gap. Two conditions must be satisfied for this penetration to occur. First, the thickness of the anode sheath, which separates the discharge plasma from the anode, must be much less than the emission opening. Second, the distance between the plasma and the accelerating electrode (which can be estimated from the Child–Langmuir law) must be shorter than the accelerating gap length. When these conditions are satisfied, the discharge current switches from the anode to the accelerating electrode and the voltage across the accelerating gap sharply falls [6], which is treated as breakdown. The dependence of the maximum pressure on the emission opening size noted in [4], as well as the fact that the discharge current (rather than the emission current) governs the breakdown, shows that the overlap of the beam and discharge plasmas cannot be treated as the main mechanism of breakdown (as was proposed in [7]).

The main reason for the decrease in the anode sheath thickness is the increase in the plasma density due to both the increase in the discharge current and the influx of gas ions from the accelerating gap. However, the increase in the plasma potential caused by the increase in the potential of the accelerating electrode increases the sheath thickness. The plasma potential is assumed to be positive with respect to the anode, which is confirmed by direct measurements with the help of an emission probe. Taking into account the current balance and the possibility of the accelerating-electrode field penetrating into the emission openings, we obtain the following expression for the plasma potential φ_p:

$$\varphi_p = \frac{kT_e}{e}\ln\left[\frac{S_a j_{ex}\left(1+\xi\left(\exp\left(\frac{eD\varphi_e}{kT_e}\right)\right)\right)}{I_d}\right], \qquad (1)$$

where j_{ex} is the density of electron current fluctuations in the plasma, φ_e is the accelerating electrode potential, D is the anode "electric transparency," S_a is the anode area, and ξ is the ratio of the emission area to the anode area.

Using the expression for the anode sheath thickness l from [4] and suggesting that the condition for plasma penetration from the discharge region into the accelerating gap is $l = \beta h$, we obtain a formula relating the maximum discharge current to the parameters of the discharge–emission system, as well as to the gas species and pressure:

$$I_{dm} = \frac{4}{(\beta h)^2}\left[\frac{\varepsilon_0 \varphi_p^{3/2}}{\sqrt{ekT_e}}\left(1 - \frac{3}{4} n_n Q_e d \frac{Q_i}{Q_n}\sqrt{\frac{MT_e}{mT_i}}\right)\right] \times \left(0.4 e S_c \sqrt{\frac{2kT_e}{M}}\right), \quad (2)$$

where n_n is the density of neutrals in the accelerating gap, Q_e is the cross section for electron–impact ionization of gas molecules by fast electrons, Q_i is the total cross section for the interactions of a slow ion in plasma, Q_n is the cross section for charge exchange of fast ions, M is the ion mass, T_i is the temperature of plasma ions, d is the accelerating gap length, the parameter β is determined experimentally ($\beta < 1$), and the anode potential is assumed to be zero.

The results of calculations by formulas (1) and (2) presented in Fig. 3 agree satisfactorily with the experimental data, which confirms the proposed model of the second type of breakdown. The model also accounts for the increase in I_{dm} with increasing number of openings in the anode diaphragm because, according to Eq. (1), the increase in the transparency D leads to an increase in φ_p, which, according to Eq. (2), increases I_{dm}.

Thus, the experimental data obtained allow us to conclude that the main reason for the breakdown of the PES accelerating gap at elevated pressures is the penetration of plasma from the discharge region into the accelerating gap.

REFERENCES

1. Yu. E. Kreĭndel', *Plasma Electron Sources* (Énergoatomizdat, Moscow, 1977).
2. S. I. Belyuk, Yu. E. Kreĭndel', and N. G. Rempe, Zh. Tekh. Fiz. **50**, 203 (1980) [Sov. Phys. Tech. Phys. **25**, 124 (1980)].
3. E. M. Oks and P. M. Schanin, Phys. Plasmas **7**, 1649 (1999).
4. Yu. A. Burachevskiĭ, V. A. Burdovitsin, A. V. Mytnikov, *et al.*, Zh. Tekh. Fiz. **71** (2), 48 (2001) [Tech. Phys. **46**, 179 (2001)].
5. V. A. Burdovitsin and E. M. Oks, Rev. Sci. Instrum. **70**, 2975 (1999).
6. A. V. Zharinov and Yu. A. Kovalenko, Zh. Tekh. Fiz. **56**, 681 (1986) [Sov. Phys. Tech. Phys. **31**, 410 (1986)].
7. S. Yu. Udovichenko, Zh. Tekh. Fiz. **70** (3), 19 (2000) [Tech. Phys. **45**, 304 (2000)].

Translated by N. Ustinovskiĭ

Technical Physics, Vol. 47, No. 7, 2002, pp. 929–931. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 137–140.

BRIEF COMMUNICATIONS

Comparison of the Shannon Information, Tsallis Information, and Number of System States in Diagnosing Regular and Chaotic Motions

A. L. Tukmakov

Institute of Mechanics and Mechanical Engineering, Kazan Scientific Center, Russian Academy of Sciences, Kazan, 420011 Tatarstan, Russia

e-mail: tukmakov@mail.knc.ru

Received October 30, 2001

Abstract—The properties of the Shannon and Tsallis informations in diagnosing laminar and turbulent motions are compared by solving the set of Lorentz equations. A method for analyzing the mode of dynamic behavior by constructing the function of number of states is proposed. © *2002 MAIK "Nauka/Interperiodica".*

The Shannon information is employed for the identification of nonlinear-system dynamic behavior and makes it possible to distinguish laminar and turbulent motions [1]:

$$I_S = -\sum_{i=1}^{N} p_i \log p_i. \quad (1)$$

Here, N is the number of system states and p_i is the probability of the system being in the ith state. In order to apply (1) in the analysis of the time realization of a signal, one has to define the state space of a dynamic system and find the probabilities of these states. The time slot principle [2] is commonly used to find time-dependent information. First, the slot ΔT is chosen on the time axis and information (1) found for this slot is assigned to the time instant corresponding to either the center or the right boundary of the slot. Then, the slot moves along the time axis, allowing one to construct the time dependence of the information. The space of states may be defined, for instance, by superimposing a mesh with a given partition on the attractor in the phase space (Fig. 1). The state when a point of the attractor falls into a cell (j, k) is taken to be the system state with the number $(k-1)J + j$. Within the current slot, the total number N of different system states constituting the complete set of events is determined, the probabilities p_i $[i = (1, N)]$ of the realization of these states are found, and the information at $t = \Delta T/2$ is calculated. Then, the slot is shifted by $\Delta\tau$, the information for $t = \Delta T/2 + \Delta\tau$ is calculated, etc. The Shannon information is a specific case of the Tsallis information [2],

$$I_T = \frac{1}{1-q}\left(\sum_{i=1}^{N} (p_i)^q - 1\right), \quad (2)$$

having a wider variety of properties than the Shannon information. As was shown in [3], $I_T \longrightarrow I_S$ as the numerical parameter q tends to unity. Let us compare the properties of the Shannon and Tsallis informations, as applied to the diagnostics of laminar and turbulent modes, by the example of solving the set of Lorentz equations

$$\dot{x} = \sigma(y-x), \quad \dot{y} = rx - xz - y, \quad \dot{z} = xy - bz, \quad (3)$$

with the control parameters $\sigma = 10$, $b = 8/3$, and $r = 170.5$. In this case, the supercriticality parameter $r - r^{\nu} = 170.5 - 166.07 = 4.43$.

Set (3) was solved numerically. The values of the functions x, y, and z at the time instant t^{n+1} were represented as a truncated Taylor series about the time point t^n. Higher derivatives were expressed in terms of first-order derivatives. This made it possible to turn to the $(n+1)$th time layer. The first four terms of the Taylor series were used, so that the method was accurate to the 4th order. In the calculations, the time step was $\Delta t = 0.0001$.

Figure 2a shows the time realization of the x component of the solution involving alternating laminar and

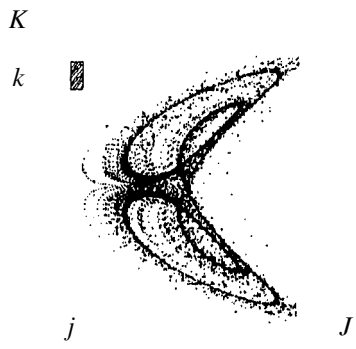


Fig. 1. Introduction of the state space of a dynamic system by superimposing a mesh on the phase space area occupied by the attractor. J and K are the numbers of cells in the x and y directions, respectively.

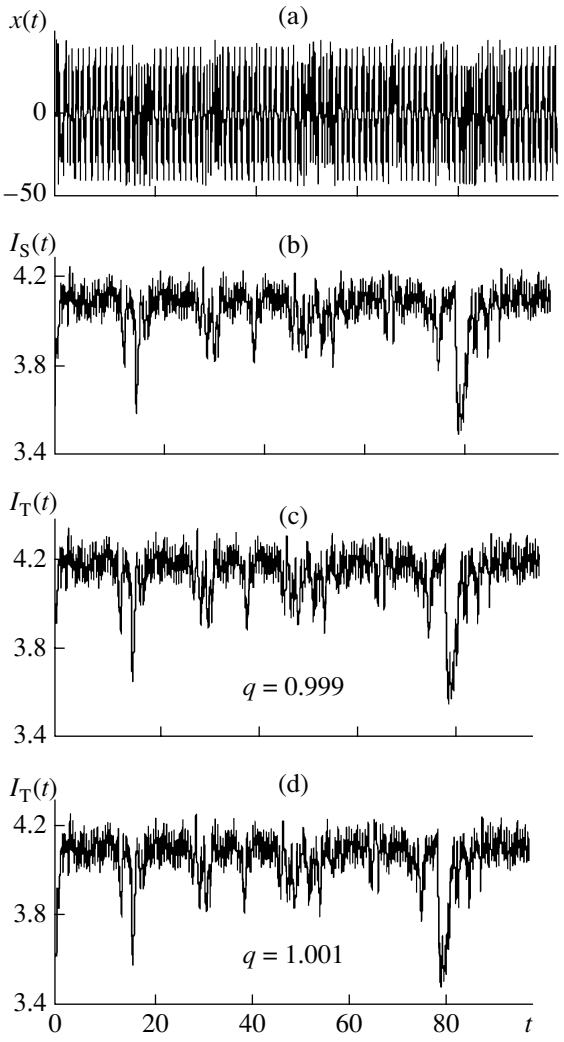


Fig. 2. (a) Time realization $x(t)$ plotted by 10^4 points; (b) the Shannon information; and (c) and (d) the Tsallis information at $q = 0.999$ and 1.001, respectively. The state space is introduced in the plane $x(t)$–$y(t)$ by superimposing a mesh with $J = 30$ and $K = 30$ on the area occupied by the attractor. The time slot consists of 100 points.

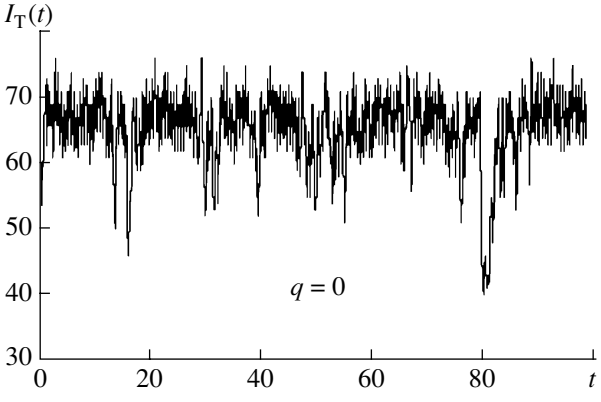


Fig. 4.

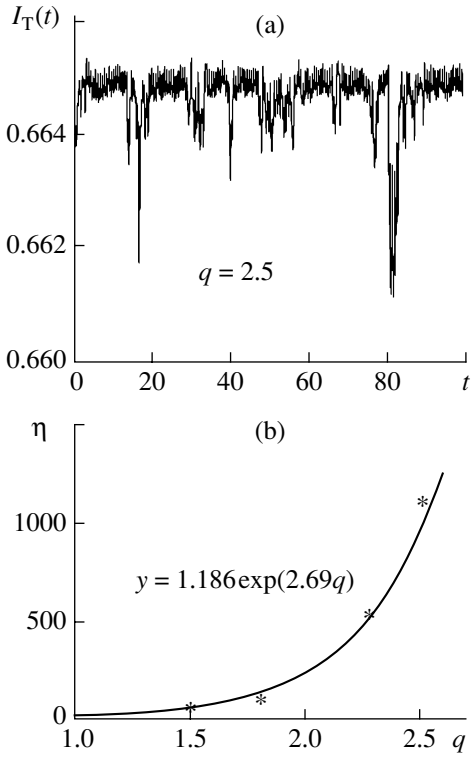


Fig. 3.

turbulent modes for $x(0) = 0$, $y(0) = 0.01$, and $z(0) = 0$. In Figs. 2b–2d, the Shannon and Tsallis informations for q close to unity are presented. As is seen, to the laminar modes in the time realization, there correspond information oscillations about the slowly varying mean value. The turbulent modes are described by information oscillations with amplitudes substantially exceeding the noise. At $q \approx 1$, the Shannon and Tsallis informations are close to each other (Figs. 2b–2d). As q increases, the signal-to-noise ratio for the Tsallis information changes (Figs. 3a, 3b).

The effect of noise suppression stems from the fact that the noise in the current slot has the form of a series of states with a low relative frequency of occurrence. When raised to a power $q > 1$, the weight of low-frequency states decreases compared to that of higher frequency states (corresponding to the signal). At $q = 0$, the Tsallis information $I_T(t) = N(t) - 1$, where $N(t)$ is the number of system states at a time instant t. Thus, when the turbulent motion arises, the number of system states changes abruptly compared to that in the laminar mode (Fig. 4). The further decrease in the parameter q (in the range of negative values) does not affect qualitatively the behavior of $I_T(t)$ when the laminar and turbulent modes alternate. With decreasing q, the signal level (Fig. 5) and the noise level increase. The increase in the noise intensity for $q < 1$ is because the weight of low-frequency states grows.

Turbulent and laminar modes differ from each other in number of possible system states. Therefore, the cur-

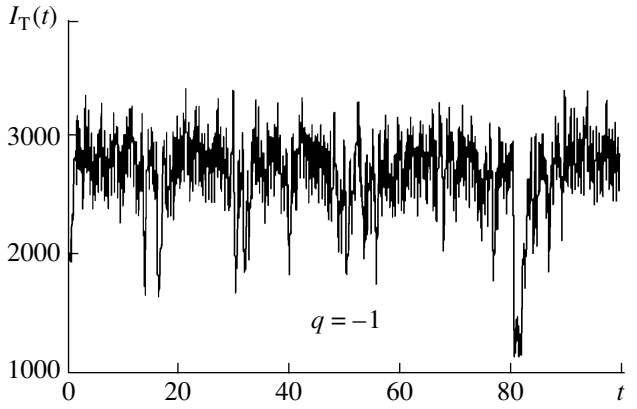


Fig. 5.

rent number of system states is an appropriate parameter for identifying the dynamic mode. Let us use the following rule to count the number of system states: the number of system states S increases by unity if some state occurs for the first time. Depending on the depth of analysis, by a "state" one means either a state at the current time instant, $S(t^n)$, or a set of states, $(S(t^n), S(t^{n-1}), \ldots, S(t^{n-k}))$.

The initial growth of the function $S(t)$ (Fig. 6a) is associated with the accumulation of nonrepeating states. In the course of time, the function $S(t^n)$ saturates: the space of its states becomes exhausted rather rapidly. For the more complicated notion of state, the saturation proceeds more slowly (curves *2* and *3* in Fig. 6a). To laminar slots of the function $y(t)$ (Fig. 6c), there correspond horizontal segments of the function S. In ranges of turbulence, S increases because new states appear in the dynamic process. It is worth noting that in the framework of this approach, a turbulent burst observed in previous time instants will be perceived as laminar in subsequent time instants. This causes the saturation of $S(T)$ and leads one to analyze the next (in terms of the depth of embedding) notion of system state.

The saturation of $S(T)$ can be avoided by limiting the set of states being analyzed. Let us introduce a time slot of width ΔT. In order to relate the number of states and time, $S(t^n)$, we consider the states realized within the slot in the range $t^n - \Delta T < t < t^n$. Information about the states having occurred by a time $t < t^n - \Delta T$ is lacking. The function $S(t^n)$ increases by unity if the state corresponding to the time instant t^n has not been occurred in the current slot (within the interval $t^n - \Delta T < t < t^n$). Figure 6b shows the results of analysis of the laminar and turbulent modes for the y component of the Lorentz system. They were obtained by calculating the number of system states that occur in the time slot of width $\Delta T = 3$ moving along the time axis. A comparison of curves *1* and *2* in Figs. 6b with the curve $y(t)$ in Fig. 6c shows that the slot function $S(t^n)$ does not saturate. This makes it possible to diagnose turbulent modes more precisely. We suppose that the slot width ΔT is defined

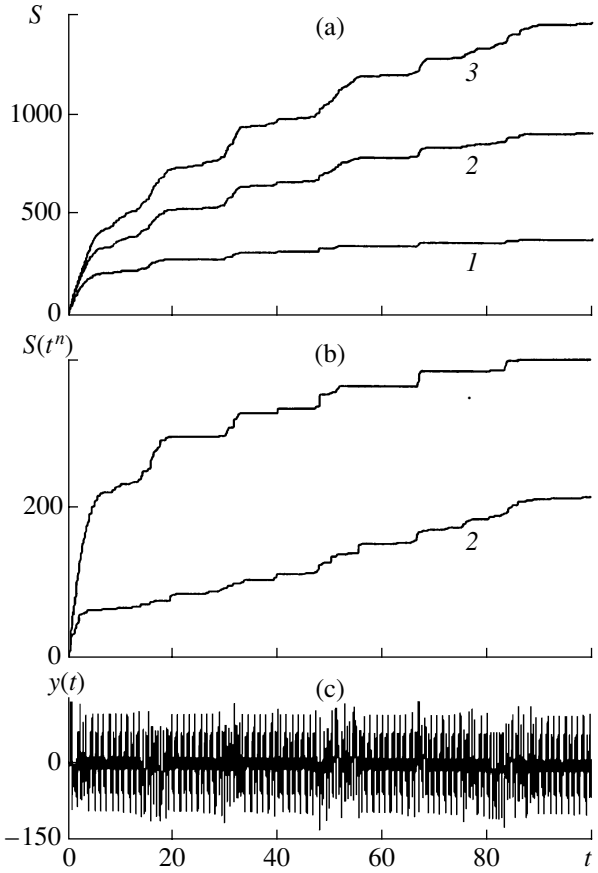


Fig. 6. (a) Number of states of a dynamic system for various depths of analysis: (*1*) $S = S(t^n)$, (*2*) $S = S(t^n, t^{n-1})$, and (*3*) $S = S(t^n, t^{n-1}, t^{n-2})$. (b) Comparison of the numbers of states $S(t^n)$ found (*1*) without and (*2*) with the use of the time slot. (c) Time dependence of the y component of the Lorentz system.

by the setting time of the laminar mode: the slot has to cover at least one period of the laminar motion.

Thus, it is possible to specify the signal-to-noise ratio for the time-dependent Tsallis information by varying the parameter q. The analysis of the number of system states with the use of the Tsallis information at $q = 0$ has shown that laminar and turbulent motions differ in the number of states. Our results offer possibilities of developing an effective method for diagnosing laminar and turbulent modes. This method is based on the calculation of the number of system states within a current time slot and is easy to implement.

REFERENCES

1. H. G. Schuster, *Deterministic Chaos* (Physik-Verlag, Weinheim, 1984; Mir, Moscow, 1988).
2. A. Capurro, L. Diambra, D. Lorenzo, *et al.*, Physica A (Amsterdam) **265**, 235 (1999).
3. S. Abe, Phys. Lett. A **271**, 74 (2000).

Translated by M. Fofanov

Technical Physics, Vol. 47, No. 7, 2002, pp. 932–933. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 7, 2002, pp. 141–142.

BRIEF COMMUNICATIONS

Nonequilibrium Semiconductor–Metal Phase Transition Due to Self-Heating

A. V. Melkikh and A. A. Povzner

Ural State Technical University, ul. Mira 19, Yekaterinburg, 620002 Russia

e-mail: mav@dpt.ustu.ru

Received November 9, 2001

Abstract—A model of the nonequilibrium semiconductor–metal phase transition in narrow-gap transition metals due to self-heating is considered. It is shown that self-heating may diminish the energy gap and, as a consequence, increase the current. Parameters resposible for the bistable behavior of the system are found. © *2002 MAIK "Nauka/Interperiodica".*

There is a wide group of narrow-gap transition metals that undergo the reversible insulator–metal transition without changing their aggregative state [1, 2].

At the phase transition temperature T_{tr}, the temperature dependence and the value of the conductivity change drastically. Below T_{tr}, it varies by the exponential law (which is typical of insulators and semiconductors); above T_{tr}, the conductivity weakly declines, which is characteristic of metals.

Mechanisms responsible for electron transitions still remain unclear (see, e.g., [3, 4]). However, since temperature variations may happen as a result of self-heating of a material (in this case, the conductivity changes stepwise by several orders of magnitude), the behavior of these substances can be described in terms of heat exchange with the environment and resulting current instabilities.

In insulators and semiconductors, the temperature dependence of the charge carrier concentration obeys the law

$$n = n_0 \exp\left(-\frac{E_g}{2kT}\right). \quad (1)$$

According to the electron theory of metal–semiconductor phase transition [1], the energy gap decreases with increasing temperature due to the growth of the charge carrier concentration n. This decrease can be represented approximately in the form

$$E_g = E_{g0} - Kn, \quad (2)$$

where E_{g0} is the energy gap at $n = 0$ and K is the electron–phonon interaction constant.

Then, the expression for the conductivity takes the form

$$\sigma = \sigma_0 \exp\left(-\frac{E_{g0} - M\sigma}{2kT}\right) = \sigma_0 \exp\left(-\frac{E_{g0}}{2kT}\right)\exp\left(\frac{M\sigma}{2kT}\right), \quad (3)$$

where $M = n_0/\sigma_0$ is a constant.

Let us consider a conductor with a circular cross section of radius R where the insulator–metal phase transition may take place and find its I–V characteristics. We assume that the heat being released in the sample is spent on its warming. Under steady-state conditions, the direct current power is equal to the thermal flux dissipating from the surface into the environment:

$$j\pi R^2 \Delta\varphi = \alpha 2\pi R(T - T_0), \quad (4)$$

where j is the current density in the sample and $\Delta\varphi$ is the potential difference across the sample.

We assume that the sample surface makes a major contribution to the thermal resistance and that the sample temperature is uniform. This situation corresponds to the case when the radius of a current filament [5] equals the radius of the conductor. Then, in view of Eqs. (3) and (4), the expression for the I–V characteristic of the sample can be written in the form

$$j = \sigma_0 \Delta\varphi \exp\left(-\frac{E_{g0} - M\frac{j}{\Delta\varphi}}{2k\left[T_0 + \frac{jR\Delta\varphi}{2\alpha}\right]}\right). \quad (5)$$

Due to the presence of feedback, the I–V characteristic is nonlinear.

1063-7842/02/4707-0932$22.00

To analyze the *S*-shaped regions of the *I–V* characteristic, we will pass to the dimensionless variables

$$\frac{Mj}{2kT_0\Delta\varphi} = Y, \quad \frac{R}{2\alpha T_0}j\Delta\varphi = X, \quad \frac{E_{g0}}{2kT_0} = E_0$$

and find its singularities:

$$\frac{\partial\Delta\varphi}{\partial j} = 0 = \frac{1}{\sigma_0}\exp\left(\frac{E_0 - Y}{1+X}\right)\left\{1 + \frac{-Y - E_0X}{[1+X]^2}\right\}. \quad (6)$$

In general, the number of singularities may equal two, one, or zero. The transition from the *S*-shaped curve to the monotonic (critical) *I–V* characteristic implies that the condition

$$\frac{\partial^2\Delta\varphi}{\partial j^2} = 0$$

must be satisfied at the critical point.

Then, from Eq. (6), it follows that

$$X^2 + (2 - E_0)X - Y + 1 = 0. \quad (7)$$

Equating the second derivative to zero, we find that

$$-E_0X + 4[1 + X] = Y. \quad (8)$$

Combining Eqs. (7) and (8), we come to the quadratic equation

$$X^2 - 2X - 3 = 0;$$

hence, $X_c = 3$.

Then, from the condition that the first and second derivatives equal zero, we have

$$16 - 3E_{0c} - Y_c = 0.$$

Using (5), we arrive at

$$\frac{kT_0}{M\sigma_0}(16 - 3E_0) = \exp(4 - E_0).$$

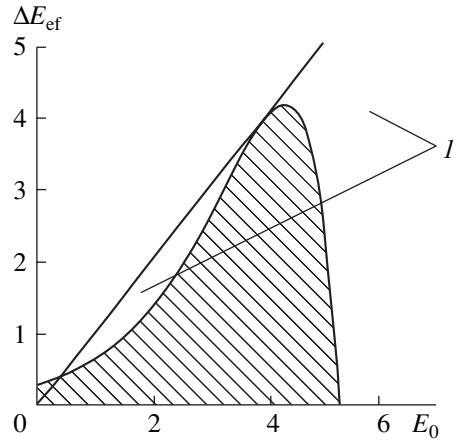
Earlier, for pure semiconductors, we estimated the critical energy gap at [6] $E_0 = 4$ or approximately 0.2 eV. Denoting

$$\frac{M\sigma_0}{kT_0} \equiv \Delta E_{ef},$$

we obtain the phase diagram of the system (see figure):

$$\Delta E_{ef} = \exp(E_0 - 4)(16 - 3E_0).$$

In the figure, the hatched region corresponds to the absence of bistability (nonequilibrium phase transitions); the straight line, to the zeroth gap width. Only the states below this line seem to have physical meaning.



Phase diagram of the system. *1*, regions of bistability

For example, for VO_2, the energy gap is about 0.5 eV [1]. Therefore, this material falls into the region of bistability. The formulas found above are valid in the case when the hot (or cold) phase occupies the entire conductor. The minimum temperature is expected on the sample surface. If this temperature is so low that the hot phase cannot exist, the transition to the cold phase occurs and a two-phase system forms. In this case, the *I–V* characteristic changes, since the interface shifts when the potential difference varies. We intend to consider this situation in subsequent works.

Thus, the analysis of the *I–V* characteristic for narrow-gap transition metals demonstrates that self-heating may cause a step in the conductivity. The parameters responsible for the bistable behavior of the system are found.

REFERENCES

1. A. A. Bugaev, B. P. Zakharchenya, and F. A. Chudnovskiĭ, *Metal–Semiconductor Phase Transition and Its Application* (Nauka, Leningrad, 1979).
2. N. F. Mott, *Metal–Insulator Transitions* (Taylor & Francis, London, 1974; Nauka, Moscow, 1979).
3. M. Imada, A. Fujimori, and Y. Tokura, Rev. Mod. Phys. **70**, 1059 (1998).
4. A. Krotkus and Z. Dobrovol'skis, *Electrical Conduction in Narrow-Gap Semiconductors* (Mokslas, Vilnius, 1988).
5. V. N. Andreev and F. A. Chudnovskiĭ, Fiz. Tverd. Tela (Leningrad) **17**, 2957 (1975) [Sov. Phys. Solid State **17**, 1966 (1975)].
6. A. V. Melkikh, A. A. Povzner, A. G. Andreeva, and I. N. Sachkov, Pis'ma Zh. Tekh. Fiz. **27** (6), 19 (2001) [Tech. Phys. Lett. **27**, 226 (2001)].

Translated by Yu. Vishnyakov