Technical Physics, Vol. 45, No. 10, 2000, pp. 1231–1237. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 1–7.

Critical Behavior and 1/f Noise in Lumped Systems Undergoing Two Phase Transitions

V. P. Koverda and V. N. Skokov

Institute of Thermal Physics, Ural Division, Russian Academy of Sciences, ul. Pervomaĭskaya 91, Yekaterinburg, 620219 Russia

E-mail: iva@itp.e-burg.su

Received February 12, 1999

Abstract—It is shown that the interaction of order parameters when subcritical and supercritical phase transitions take place simultaneously may result in a self-organized critical state and cause a $1/f^{\alpha}$ fluctuation spectrum, where $1 \le \alpha \le 2$. Such behavior is inherent in potential and nonpotential systems of nonlinear Langevin equations. A numerical analysis of the solutions to the proposed systems of stochastic differential equations showed that the solutions correlate with fractional integration and differentiation of white noise. The general behavior of such a system has features in common with self-organized criticality. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Fluctuations with a spectrum that is inversely proportional to frequency ($1/f^{\alpha}$ noise, $0.5 \le \alpha \le 1.5$) are observed in various systems. $1/f^2$ spectra are typical of current fluctuations in radiophysical devices, biological systems, and geophysical processes [1–4]. Recently, intense $1/f$ noise has been detected in a system of two interacting nonequilibrium phase transitions [5, 6]. This situation was realized when the superconductor–conductor phase transition due to Joule self-heating was imposed on the critical transition from nucleate boiling to film boiling of liquid nitrogen, which cools down the superconductor.

Despite the fact that $1/f$ noise occurs widely and interest in random processes with low-frequency spectral divergence has not been lost, a commonly accepted mathematical model is lacking and mechanisms of $1/f^{\alpha}$ noise generation are poorly understood. Different models are used to explain $1/f$ noise in various processes.

It is generally accepted that $1/f$ noise in radiophysics and solid-state physics arises from an exponentially wide distribution of relaxation times. In this model, the noise is treated as a superposition of random relaxation processes [1–4]. Electronic conductivity fluctuations with a $1/f^{\alpha}$ spectrum in solids are often related to heat conduction mechanisms as well [7]. However, this approach frequently fails to explain experimental results [1, 2]. Therefore, nonlinear heat sources in the heat conduction equation were taken into account in [8]. Thermal generation of $1/f^{\alpha}$ noise in metals was considered in [9], where the appearance of $1/f$ noise was associated with the nonlinear interaction of diffusive and heat-conducting modes. In [10], $1/f$ noise was explained by strong fluctuations of the relaxation and dissipation rates, whereas in [11], it was interpreted as the result of Brownian motion in a space of limited dimensionality.

It is well known that a random process with a $1/f$ spectrum can be obtained by the fractional (of the order 1/2) integration of white noise [3, 4]. This technique implies that a system remembers an infinite number of past events. Therefore, such an approach to the $1/f$ noise problem has not been adequately reflected in physical terms.

In [12], the idea of self-organized criticality was proposed to give insight into the appearance of $1/f$ noise. This concept was illustrated with the model of "sand-pile" cellular automata. The theory of self-organized criticality implies that systems with $1/f^{\alpha}$ noise should be in a critical state. However, the critical state is a steady state of the system, and there is no need for fine adjustment of control parameters. The concept of self-organized criticality [12] attracted considerable interest and became widely used in statistical physics.

Earlier, we proposed a mathematical model of two superimposed phase transitions to explain experimental data for the generation of intense $1/f$ noise in a system consisting of a current-carrying superconductor and a boiling coolant. The model is represented by a set of two coupled nonlinear Langevin equations that converts white noise into two modes of stochastic oscillations with spectra proportional to $1/f$ and $1/f^2$ [6, 13].

In this paper, we report a numerical study of the system of two coupled nonlinear Langevin equations. A correlation between the behavior of this system, fractional integration of white noise, and the concept of self-organized criticality was found.

1063-7842/00/4510-1231$20.00

1/*f* NOISE IN A SYSTEM UNDERGOING TWO PHASE TRANSITIONS

Intense 1/*f* noise was detected when thin-film high-temperature superconductors cooled by boiling nitrogen lost their superconducting properties when heated [5, 6, 14]. This noise was observed in the region where the subcritical nonequilibrium phase transition from nucleate boiling to film boiling of the nitrogen was superimposed on the critical or supercritical superconductor–conductor phase transition due to Joule self-heating. In these experiments, both transitions are localized in space and the system can be considered as lumped.

In the case of superimposed noninteracting supercritical and subcritical phase transitions with order parameters x and y, respectively, a lumped system is described by the potential

$$\Phi(x, y) = \frac{1}{4}a_1x^4 + \frac{1}{2}a_2x^2 + \frac{1}{2}b_1y^4 - b_2y^2, \qquad (1)$$

where the coefficients a_1, a_2, b_1, and b_2 are positive.

The interaction of the phase transitions can formally be taken into account by nonlinear transformation of variables. Let us introduce interacting order parameters X and Y that are related to x and y by the simple relations $x \equiv X$ and $y^2 \equiv XY$. Then, potential (1) takes the form

$$\Phi(X, Y) = \frac{1}{4}a_1X^4 + \frac{1}{2}a_2X^2 + \frac{1}{2}b_1X^2Y^2 - b_2XY. \quad (2)$$

Figure 1 shows the potential surface $\Phi(X, Y)$; and Fig. 2a, the potential "bottom" levels.

Under the additive action of white noise on the system, random walks in potential (2) are described by the set of Langevin equations

$$\begin{aligned} \frac{dX}{dt} &= -b_1XY^2 + b_2Y - a_1X^3 - a_2X + \varepsilon_1(t), \\ \frac{dY}{dt} &= -b_1X^2Y + b_2X + \varepsilon_2(t), \end{aligned} \qquad (3)$$

where $\varepsilon_1(t)$ and $\varepsilon_2(t)$ are Gaussian δ-correlated noises with a zero mean and a standard deviation D.

System (3) was solved numerically by the Euler method, where Gaussian random-number sequences were used as $\varepsilon_1(t)$ and $\varepsilon_2(t)$.

If system (3) is integrated in the absence of external noise and close-to-zero initial conditions, the variables $X(t)$ and $Y(t)$ first linearly increase and then, at $t \longrightarrow \infty$, behave asymptotically: $X(t) \longrightarrow t^{-1/2}$ and $Y(t) \longrightarrow t^{-1/2}$. As numerical integration shows, the introduction of weak noises $\varepsilon_1(t)$ and $\varepsilon_2(t)$ into system (3) just makes the above trajectories noisy but does not affect the system behavior qualitatively. The phase paths change qualitatively when the intensity of external noise exceeds a certain value. The transition from the relaxation behavior of the dynamic variables $X(t)$ and $Y(t)$ to steady random processes is observed. There exists a range of the dispersion of external white noise where the fluctuation spectral densities of the variables $X(t)$ and $Y(t)$ have the form $S_X(f) \sim 1/f$ and $S_Y \sim 1/f^2$, respectively. On further increasing the noise intensity, the divergent $1/f^\alpha$ spectra turn into Lorentz spectra with a characteristic "plateau" at low frequencies. The frequency range where the $1/f^\alpha$ spectra are observed rapidly extends with decreasing integration step. Simultaneously, so does the range of white noise intensities where the $1/f^\alpha$ spectra of the variables $X(t)$ and $Y(t)$ are observed. Figure 3 shows the $1/f$ spectrum of the variable $X(t)$ that results from the Fourier transform for the numerical solution of system (3) at $a_1 = a_2 = b_1 = b_2 = 1$ and the external noise dispersion $D2 = 25$ (the integration step $\Delta t = 0.01$, and the number of steps $N = 262\,143$). The corresponding spectrum of the variable $Y(t)$ has the form $S_Y \sim 1/f2$. Figure 4 shows the phase portrait of the system.

Thus, the interaction of subcritical and supercritical phase transitions that is defined by potential (2) (Fig. 1) may lead to the transformation of additive white noise into steady stochastic processes with $1/f^\alpha$-type spectral densities. In our model, fluctuations with a $1/f$ spectrum result from the Brownian behavior of a system subjected to two-valley potential (2).

Sets of Langevin equations yielding $1/f^\alpha$ noise need not be potential. A simple nonpotential system generating $1/f$ noise has the form [6]

$$\begin{aligned} \frac{dX}{dt} &= -XY^2 + Y + \varepsilon_1(t), \\ \frac{dY}{dt} &= -X^2Y + \lambda X + \varepsilon_2(t), \end{aligned} \qquad (4)$$

where the parameter $\lambda > 1$.

For $\lambda = 1$, system (4) describes random walks in the potential

$$\Phi(X, Y) = \frac{1}{2}X^2Y^2 - XY. \qquad (5)$$

The level lines of this potential are shown in Fig. 2b. However, in this case, the system does not exhibit $(\lambda = 1)1/f^\alpha$ behavior, and the spectra of the variables $X(t)$ and $Y(t)$ have the Lorentzian form. In order for $1/f^\alpha$ spectra with divergent low-frequency asymptotic to appear, one should break the symmetry of the determinate force components in the Langevin equations. In system (4), this is achieved by introducing an additional force not related to the potential (λ is other than 1). Physically, it may mean the presence of uncompensated fluxes in an open system. Numerical simulations showed that there are no qualitative differences in the behavior of systems (4) and (3). Moreover, solutions close to those of (4) are obtained if the second-degree factors in the first terms of the determinate forces in Eq. (4) are changed to fourth- or sixth-degree ones and

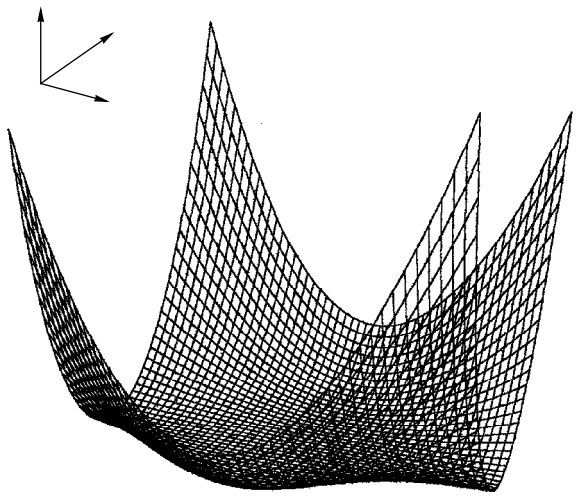


Fig. 1. Potential surface $\Phi(X, Y)$ defined by (2).

the linear terms are changed to third- or fifth-degree terms. In all the cases, the mean value $\langle X(t)Y(t)\rangle \simeq 1$, whereas the relaxation asymptotic for $t \longrightarrow \infty$ in noiseless systems of elevated degree is different but also fractional-power ($t^{-1/4}$ and $t^{3/4}$ or $t^{-1/6}$ and $t^{5/6}$, respectively).

Note also that the rotation of the phase plane in Eq. (4) through the angle $\pi/4$ by the linear transformation $u = (x + y)/2$ and $v = (x - y)/2$ produces the new system

$$\begin{aligned}\frac{du}{dt} &= -\frac{1}{2}u^3 + \frac{3}{2}u + \frac{1}{2}uv^2 + \frac{1}{2}v + \varepsilon_1(t),\\ \frac{dv}{dt} &= -\frac{1}{2}v^3 - \frac{3}{2}v + \frac{1}{2}u^2v - \frac{1}{2}u + \varepsilon_2(t),\end{aligned} \tag{6}$$

which yields similar (but not the same) stochastic processes with spectra $S_u(f) \sim f^{-1.5}$ and $S_v(f) \sim f^{-1.5}$ in a wide range of white noise intensities. Rotation through another angle may yield other values of the low-frequency divergence α in the range $1 \le \alpha \le 2$. The form of system (6) indicates that it, like potential system (3), corresponds to the superposition of two interacting phase transitions: subcritical (parameter u) and supercritical (parameter v). However, in contrast to system (3), the determinate force components in (6) contain nonpotential terms due to the presence of an uncompensated flux. It is clear that the same physical meaning may also be assigned to system (4). The interaction of the subcritical and supercritical phase transitions in these systems leads to the critical behavior in a rather wide range of white noise intensities, as evidenced by the low-frequency divergence of the spectral densities of the order parameters.

Thus, there is a wide class of sets of nonlinear Langevin equations that provide $1/f^{\alpha}$ spectra of variables. Their common features are nonlinear interaction and asymmetry of the equations. Further, we will analyze the simplest case of systems (4) at $\lambda = 2$.

CRITICALITY OF A SET OF LANGEVIN EQUATIONS

Numerically integrating system (4) by the Euler method, we found that computational instability appears with increasing integration step even in the absence of external noise [6]. The Euler system can be made stable for an arbitrary integration step and arbitrary initial conditions if one takes the value of X_i in the first term of the first equation at the finite point of a step Δt instead of at the initial one. In this case, we come to the system

$$\begin{aligned}X_{i+1} &= X_i + (-X_iY_i^2 + Y_i)\Delta t(1 + Y_i^2\Delta t)^{-1} + \xi_i\Delta t,\\ Y_{i+1} &= Y_i + (-X_i^2Y_i + 2X_i)\Delta t + \eta_i\Delta t,\end{aligned} \tag{7}$$

where ξ_i and η_i are sequences of Gaussian random numbers with a zero mean and a standard deviation D.

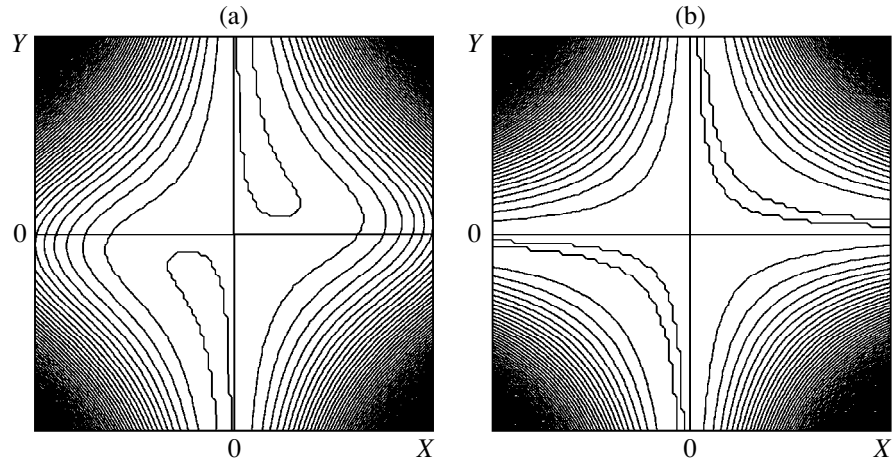


Fig. 2. Level lines of the potential "bottom" according to (a) formula (2) and (b) formula (5).

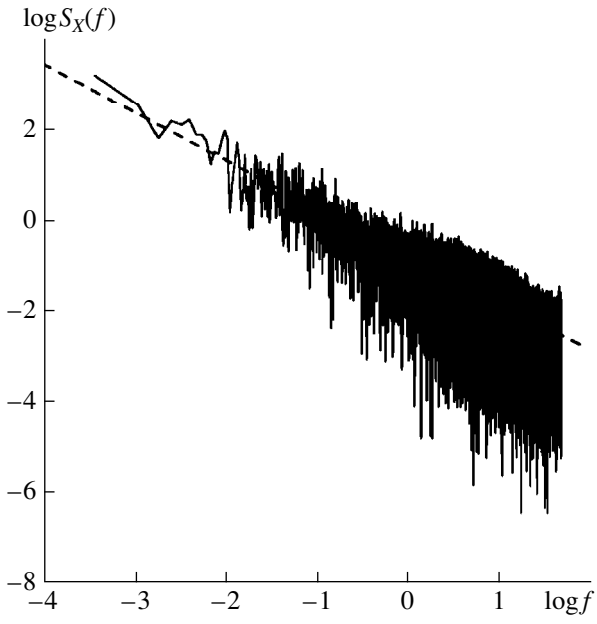


Fig. 3. 1/*f* spectral density of the order parameter $X(t)$ from system (3). The dashed line stands for $S_X(f) \sim 1/f$.

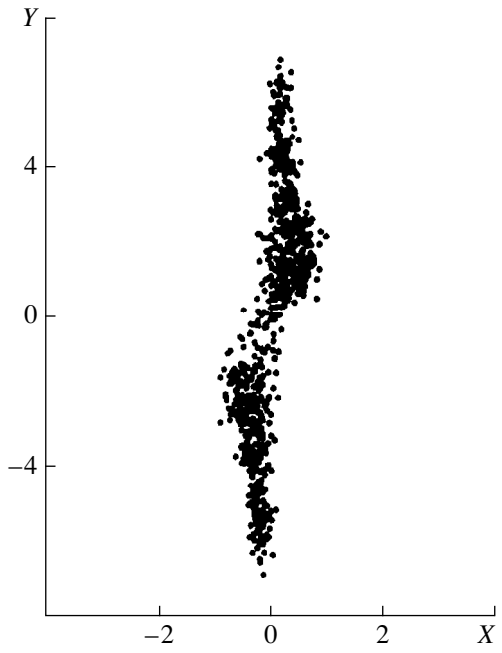


Fig. 4. Phase portrait of the system in the (X, Y) plane obtained by numerical integration of (3).

The same form of the first equation in (4) can be obtained more rigorously. One should take into account jumps not only of the first, but also of all subsequent derivatives while integrating with a finite step Δt:

$$X(t + \Delta t) = X(t) + \frac{dX}{dt}\Delta t + \frac{1}{2}\frac{d^2X}{dt^2}\Delta t^2 + \dots .$$

Having expressed the derivatives in terms of the functions X and Y and summed the convergent geometric series, we obtain (7).

Note that, if an integration step is changed n-fold, one should change the standard deviation of white noise D by the factor $n^{-1/2}$. If an integration step is decreased n-fold, the time sample size (number of steps) for which the $1/f^{\alpha}$ behavior is observed can be increased by a factor of more than n; i.e., one can trace the divergence of the spectra at lower frequencies.

Consider the characteristics of the solution of system (4) in greater detail. First, it is worth noting that the mean value $\langle X(t)Y(t)\rangle \simeq 1$ in the wide range of white noise parameters where the $1/f^{\alpha}$ spectra are observed. This means that the time-averaged determinate force in the first equation of system (4) is equal to 0. The zero value of the determinate force in the Langevin equations corresponds to the critical state. In other words, system (4) adjusts the first equation to the critical state.

As was mentioned in the Introduction, 1/*f* noise may result from the fractional integration of white noise; i.e., a stochastic process with a 1/*f* spectrum can be considered as the result of anomalous Brownian motion.

In [15], we noticed the correlation between time series $X(t)$ and fractional integration (of the order 1/2) of white noise that affects the first equation of system (4).

The 1/2-order fractional integral of white noise $\varepsilon(t)$ is given by the expression

$$I_{1/2}(t) = \frac{1}{\Gamma(1/2)}\int_0^t \frac{\varepsilon(t')}{\sqrt{t - t'}}dt', \qquad (8)$$

where Γ is the Euler gamma function.

A random process defined by (8) is steady and exhibits a 1/*f* spectrum, strictly speaking, only for $t \longrightarrow \infty$. Figure 5 shows the numerical solution $X(t)$ of system (7) and the result of calculation by formula (8) for $\Delta t = 0.1$, $D = 4.3$, and $N = 1000$. From Fig. 5, it follows that the fractional integral of white noise is a fairly good approximation of the numerical solution to the set (7) of Langevin equations for short realizations. The Pearson coefficient for the vectors X_i and I_i is $K_p \simeq 0.75$. A stronger correlation can hardly be achieved in this case, since formula (8) disregards the determinate force components appearing in sets (4) and (7).

Much stronger correlations were found between the rates of change of the dynamical variables appearing in the set (7) of Langevin equations and the functionals of white noise that contain fractional derivatives. This property may be used to find solutions of system (7). We will dwell on this feature.

The 1/2-order fractional derivative of white noise is defined as the first-order derivative of the fractional integral

$$D^{1/2}(\varepsilon) \equiv \frac{d^{1/2}\varepsilon(t)}{dt^{1/2}} = \frac{dI_{1/2}(t)}{dt}. \quad (9)$$

The evaluation of the fractional integral [and, hence, the fractional derivative by formula (9)] requires much computer time for a large number of integration steps. This time is much greater than that required for integrating the initial set of Langevin equations which is Markovian and has no memory. Therefore, for long realizations, we used the following approximation for the fractional derivative:

$$D^{1/2}(\xi_i) = \sum_{j=\begin{cases}0, & i \le r\\ i-r, & i>r\end{cases}}^{i-1} (\xi_{j+1} - \xi_j)(i-j)^{-3/2}, \quad (10)$$

where r limits the number of terms in the sum.

The results of calculation by formula (10) turned out to be almost identical to those obtained by directly computing the fractional derivative from formula (9), whereas the required time was much less. For steps $\Delta t > 0.1$, r can be within several tens.

A remarkable feature of system (7) is that the rate of change of the stochastic variable $\dot{X}_i = (X_{i+1} - X_i)/\Delta t$ obtained by the numerical solution of (7) strongly correlates with the 1/2-order fractional derivative of the sequence of random numbers ξ_i appearing in the first equation of (7). The Pearson coefficients for these quantities were $K_P = 0.92$–0.99 for arbitrary realizations and sequences ξ_i. This result was virtually independent of which formula [(9) or (10)] is used for evaluating the fractional derivative.

As for the second equation of system (7), the rate of change of the variable $\dot{Y}_i = (Y_{i+1} - Y_i)/\Delta t$ for small integration steps ($\Delta t \le 0.1$) was found to correlate strongly with the values of the random sequence numbers η_i appearing in this equation. Here, the Pearson coefficients were $K_P \simeq 0.95$ and, unlike the first equation, depended on Δt. Namely, they were close to 1 for small steps and decreased with increasing step (or with increasing sample size for a fixed step).

Thus, at least for small integration steps Δt, the determinate force component in the second equation of system (7) is automatically adjusted to zero under the action of external noises. Since such behavior is stable to variations of the system coefficients and is observed in a certain (rather wide) range of external noise intensities, we can speak about the self-organized critical state.

In order to find other properties of system (7), we studied the correlation of various combinations of the stochastic variables X_i and Y_i with combinations of the

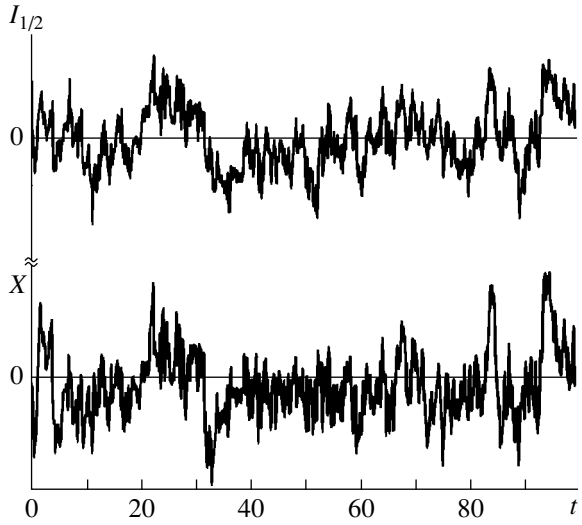


Fig. 5. Numerical solution $X(t)$ of system (7) and 1/2-order integral of white noise.

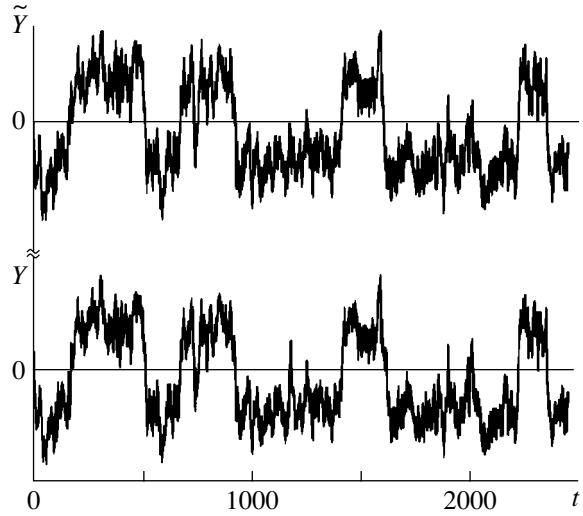


Fig. 6. Numerical solutions $Y(t)$ of system (7) and $\tilde{Y}(t)$ of stochastic Eq. (13).

noises ξ_i and η_i and their fractional derivatives. The vector X_iY_i was found to correlate strongly with the difference $\Delta\mu_i = D^{1/2}(\xi_i) - \xi_i$ between the 1/2-order derivative of white noise in the first equation of (7) and the noise itself. Hence, the approximate equality

$$X_iY_i \simeq 1 + \Delta\mu_i Y_i \quad (11)$$

is true.

Equality (11) holds with the Pearson coefficient $K_P = 0.92$–0.98. Expression (11) makes it possible to approximate one of the system variables with such an accuracy if the other variable and noise that acts on the first Eq. of (7) are known. However, approximated

equality (11) turned out to be insufficiently accurate to separate the variables in (7). Therefore, we considered the correlations of the difference $\dot{Y}_i - \eta_i$ with noise combinations and found that the approximate equality

$$1 - Y_i^2(\Delta\mu_i)^2 + mY_i\Delta\mu_i \simeq Y_i(\dot{Y}_i - \eta_i) \qquad (12)$$

was true with the correlation $K_P = 0.95$. The third term on the left-hand side of (12) involves the constant factor m. It varied from $m = 2$ for the step $\Delta t = 0.1$ to $m = 0.5$ for $\Delta t = 0.5$. This term does not affect the Pearson coefficient to the second decimal place. However, it takes into account the asymmetry of the probability distribution of the product X_iY_i about zero in the first order. Correlation (12) is equivalent to the equation

$$\tilde{Y}_{i+1} = \tilde{Y}_i + \tilde{Y}_i\frac{\Delta t}{1 + \tilde{Y}_i^2} - \tilde{Y}_i(\Delta\mu_i)^2\Delta t + m\Delta\mu_i\Delta t + \eta_i\Delta t. \qquad (13)$$

Equation (13) adequately approximates the second equation of (7) but, unlike (7), involves only one unknown stochastic function $\tilde{Y}(t)$. In the second term on the right-hand side of (13), $1/\tilde{Y}_i$ is replaced by $\tilde{Y}_i/(1 + \tilde{Y}_i^2)$ to avoid singularity. This substitution provides a better correlation between the noiseless behavior of (7) and (13). For short realizations (to 4000 points), the correlation between the solutions Y_i of (7) and $\tilde{Y}_i$ of (13) is $K_P \simeq 0.95$. However, if system (7) experiences the action of single noise ($\xi_i = \eta_i$) instead of two, the solutions Y_i and $\tilde{Y}_i$ virtually coincide for as many as several tens of thousands of steps (K_P = 0.98–0.99) (Fig. 6).

In fact, the analysis of various correlations between the stochastic variables of system (7) and combinations of noises made it possible to separate the variables in this system in the first order of the correlations. The corresponding stochastic equation for $\tilde{Y}(t)$ has the form

$$\frac{dY}{dt} = \frac{Y}{1 + Y^2} - Yz^2(t) + mz(t) + \varepsilon_2(t), \qquad (14)$$

where $z(t) = \varepsilon_1 - d^{1/2}\varepsilon_1/dt^{1/2}$ is the difference between the noise that acts on the first equation of (4) and the 1/2-order derivative of this noise. In the case of white noise, the fractional derivative should be treated in terms of generalized functions [16]. Stochastic differential Eq. (14) is multiplicative, and its physical interpretation is more complicated than that of system (4). However, it demonstrates the self-adjustment of the determinate force component to the zero value, so that the approximate equality $dY/dt \simeq \varepsilon_2(t)$ holds, providing self-organized criticality.

CONCLUSION

The superposition of subcritical and supercritical phase transitions with interacting order parameters may result in the self-organization of the critical state of the system. Random walks in corresponding two-valley potential (2) split additive white noise into two modes of stochastic oscillations with $1/f$ and $1/f^2$ spectral densities. To these random walks, there corresponds a class of sets of nonlinear Langevin equations with both potential and nonpotential determinate force. $1/f^\alpha$ stochastic processes with any α in the range from 1 to 2 can be obtained by the linear transformation of variables (rotation about the axis of symmetry). Although the derived set of stochastic differential equations is Markovian, their solutions correlate strongly with the fractional integral of white noise. The essential feature of this integral is the presence of system memory. The numerical analysis of the solutions to the set of nonlinear Langevin equations and the found correlations show that the system behavior has much in common with self-organized criticality. In contrast to conventional models of self-organized criticality, such as the sand pile model, which are spatially distributed, the proposed model of interacting phase transitions does not involve spatial coordinates, and the corresponding set of stochastic differential equations is a point set. The self-organization of the critical state and the generation of $1/f^\alpha$ noise (under conditions that two phase transitions in lumped systems interact) allow the interpretation of the experimental data [5, 6, 13–15] and the experimental search for systems with critical behavior and $1/f^\alpha$ noise.

ACKNOWLEDGMENTS

We thank V.P. Skripov for valuable discussions.

This work was supported by the Russian Foundation for Basic Research, project no. 99-02-16227.

REFERENCES

1. Sh. M. Kogan, Usp. Fiz. Nauk **4** (2), 285 (1985) [Sov. Phys. Usp. **28**, 170 (1985)].
2. M. B. Weissman, Rev. Mod. Phys. **60**, 537 (1988).
3. M. J. Buckingham, *Noise in Electronic Devices and Systems* (Ellis Horwood, New York, 1983; Mir, Moscow, 1986).
4. G. P. Zhigal'skiĭ, Usp. Fiz. Nauk **167**, 624 (1997) [Phys. Usp. **40**, 599 (1997)].
5. V. P. Koverda, V. N. Skokov, and V. P. Skripov, Pis'ma Zh. Éksp. Teor. Fiz. **63**, 739 (1996) [JETP Lett. **63**, 775 (1996)].
6. V. P. Koverda, V. N. Skokov, and V. P. Skripov, Zh. Éksp. Teor. Fiz. **113**, 1748 (1998) [JETP **86**, 953 (1998)].
7. R. F. Voss and J. Clarke, Phys. Rev. B **13**, 556 (1976).

8. Yu. S. Levitan, N. N. Panchenko, and O. A. Sinkevich, Dokl. Akad. Nauk SSSR **302**, 1359 (1988) [Sov. Phys. Dokl. **33**, 750 (1988)].
9. R. O. Zaĭtsev, Pis'ma Zh. Éksp. Teor. Fiz. **58**, 978 (1993) [JETP Lett. **58**, 915 (1993)].
10. Yu. E. Kuzovlev, Zh. Éksp. Teor. Fiz. **111**, 2086 (1997) [JETP **84**, 1138 (1997)].
11. Yu. L. Klimontovich, *Statistical Theory of Open Systems* (Yanus, Moscow, 1995; Kluwer, Dordrecht, 1995).
12. P. Bak, C. Tang, and K. Wiesenfeld, Phys. Rev. A **38**, 364 (1988).
13. V. P. Koverda, V. N. Skokov, and V. P. Skripov, Dokl. Akad. Nauk **356**, 614 (1997) [Phys. Dokl. **42**, 532 (1997)].
14. V. N. Skokov, V. P. Koverda, and V. P. Skripov, Cryogenics **37** (5), 263 (1997).
15. V. P. Koverda and V. N. Skokov, Physica A (Amsterdam) (1998).
16. W. Horsthemke and R. Lefever, *Noise Induced Transitions. Theory and Applications in Physics, Chemistry and Biology* (Springer-Verlag, Heidelberg, 1984; Mir, Moscow, 1987).

Translated by M. Fofanov

Technical Physics, Vol. 45, No. 10, 2000, pp. 1238–1240. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 8–10.

Polarization of Two Close Metal Spheres in an External Homogeneous Electric Field

I. E. Mazets

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

Received December 17, 1999

Abstract—The dipole moment of each of two uncharged conducting spheres with radius R in an external electric field was calculated. The distance between the centers of the spheres is $2l$. It was shown that, if $R/l \lesssim 0.8$, the influence of higher multipole moments is negligible. © *2000 MAIK "Nauka/Interperiodica".*

We consider two identical isolated uncharged, perfectly conducting spheres with radius R placed in an external homogeneous electric field $\mathbf{E}_0$ in vacuum. The centers of the spheres lie on the z-axis and stay at a distance $2l$ from each other. Each sphere acquires a dipole moment $\mathbf{p}$ related to the external field strength as follows:

$$\mathbf{p} = a\mathbf{E}_0. \tag{1}$$

Let us calculate the factor a called the effective polarizability. For an isolated sphere, this value is known to equal the radius cubed. When other conductors, e.g., another sphere, are present at finite distances, the homogeneous component of the electric field in the vicinity of the first sphere is different from $\mathbf{E}_0$. When the spheres are far from each other ($R \ll l$), only their direct dipole–dipole interaction is significant. When only this interaction is taken into account, the effective polarizability is easily calculated, resulting in

$$a_{zz} = \frac{R^3}{1-(R/l)^3/4}, \tag{2}$$

$$a_{xx} = a_{yy} = \frac{R^3}{1+(R/l)^3/8}. \tag{3}$$

Formulas (2) and (3) are valid only when the external field is directed along, or perpendicular to, the z-axis, respectively. In the general case, the influence of multipole moments of higher orders has to be taken into account. The field created by the dipole distribution of charge on the first sphere is inhomogeneous around the second sphere with a spatial inhomogeneity scale of the order of l. The second sphere acquires higher multipole moments, which then change the homogeneous component of the field around the first sphere. It is interesting to calculate their contribution to the effective polarizability numerically, because such effects turn out to be appreciable in the electrostatic interaction of close charged metal spheres [1].

A solution method based on a series expansion in spherical harmonics [2] results in an infinite system of linear algebraic equations for multipole moments which cannot be solved analytically. Therefore, this method is used only for approximate calculations of the polarization of clusters consisting of more than two metal particles. The problem with two spheres has an analytical solution in the bipolar coordinates [3] α, β, and ϕ related to the Cartesian coordinates as follows:

$$x = \frac{c\sin\alpha\cos\phi}{\cosh\beta-\cos\alpha}, \quad y = \frac{c\sin\alpha\sin\phi}{\cosh\beta-\cos\alpha}, \quad z = \frac{c\cos\alpha}{\cosh\beta-\cos\alpha}. \tag{4}$$

The surfaces of the spheres are defined by the equations $\beta = \pm\beta_0$, where $\cosh\beta_0 = l/R$. In addition, $c = (l^2 - R^2)^{1/2}$. The potential outside the spheres is written as a sum of the external field potential $-\mathbf{E}_0\mathbf{r}$ and the potential φ' created by the charge distribution on the surfaces of the spheres. When the external field is directed along the z-axis, the solution has the form [3]

$$\varphi' = (2\cosh\beta - 2\cos\alpha)^{1/2} E_0 c \sum_{n=0}^{\infty} (2n+1-\gamma) \times \exp[-(n+1/2)\beta_0]\frac{\sinh[(n+1/2)\beta]}{\sinh[(n+1/2)\beta_0]} P_n(\cos\alpha), \tag{5}$$

where $P_n(\cos\alpha)$ is the Legendre polynomial and

$$\gamma = F^{(-)}(1, \beta_0)/F^{(-)}(0, \beta_0). \tag{6}$$

Let us denote

$$F^{(-)}(q, \beta_0) = \sum_{n=0}^{\infty} \frac{(2n+1)^q \exp[-(n+1/2)\beta_0]}{\sinh[(n+1/2)\beta_0]}, \quad F^{(+)}(q, \beta_0) = \sum_{n=0}^{\infty} \frac{(2n+1)^q \exp[-(n+1/2)\beta_0]}{\cosh[(n+1/2)\beta_0]}. \tag{7}$$

1063-7842/00/4510-1238$20.00

As we have calculated, the potential created by the spheres in an external field directed along the x-axis has the form

$$\varphi' = 2(2\cosh\beta - 2\cos\alpha)^{1/2} E_0 c \sum_{n=0}^{\infty} \exp[-(n+1/2)\beta_0] \times \frac{\cosh[(n+1/2)\beta]}{\cosh[(n+1/2)\beta_0]} P_n^1(\cos\alpha)\cos\phi, \quad (8)$$

where $P_n^1(\cos\alpha)$ is the associate first-order Legendre polynomial.

Let us now analyze (5) and find the dipole moment **p** of each of the two spheres that is directed along the z-axis. Direct calculation using the charge distribution on the surface of a sphere is possible but rather cumbersome. It is easier to obtain it by analyzing the behavior of φ' at large distances. Indeed, at $x = y = 0$ and $z \longrightarrow +\infty$, i.e., at $\alpha = 0$ and $\beta \sim 2c/z \longrightarrow 0$, the potential produced by both spheres has the asymptotic form $2p/z^2 \sim \beta^2 p/(2c^2)$. Hence, using $P_n(0) = 1$, we find p and the effective polarizability

$$a_{zz} = c^3[F^{(-)}(2, \beta_0) - \gamma F^{(-)}(1, \beta_0)]. \quad (9)$$

Similarly, from the asymptotic expression of (8) at $\beta = 0$ and $\alpha \sim 2c(x^2 + y^2)^{-1/2} \longrightarrow 0$, the effective polarizability in an external field directed perpendicular to the line connecting the centers of the spheres is found as follows:

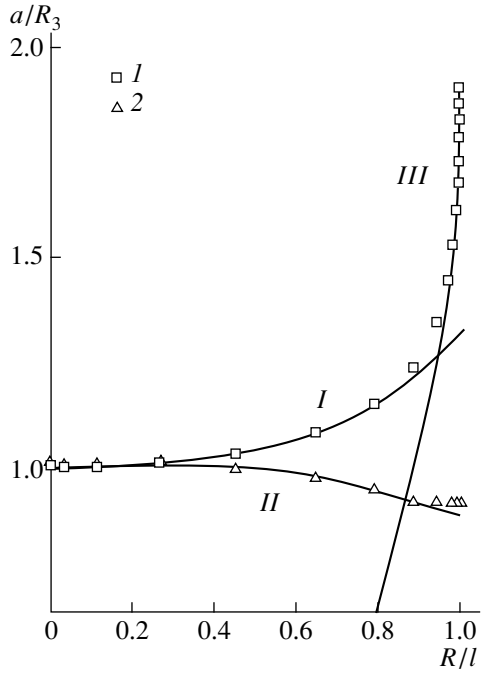
$$a_{xx} = a_{yy} = \frac{1}{2}c^3[F^{(+)}(2, \beta_0) - F^{(+)}(0, \beta_0)]. \quad (10)$$

For small values of R/l, formulas (9) and (10) lead to the approximate expressions

$$a_{zz} \approx R^3\left[1 + \frac{1}{4}(R/l)^3 + \frac{1}{16}(R/l)^6\right],$$
$$a_{xx} \approx R^3\left[1 - \frac{1}{8}(R/l)^3 + \frac{1}{64}(R/l)^6\right], \quad (11)$$

which are in excellent agreement with (2) and (3).

The figure contains the graph of the effective polarizability as a function of the distance between the centers of the spheres calculated by using (9) and (10). It is evident that the approximate expression (2) works well up to $R/l \approx 0.8$, while (3) is close to the exact expression for all possible values of R/l. It is in agreement with the fact that a description of the effective dielectric constant of thin metal films using the formulas where only the direct dipole–dipole interaction between separate metal islands is taken into account [4] (the influence of the substrate is also considered in the model of interaction between real dipoles and dipole images [5]) is adequate to interpret the experimental data in a wide domain of values of the ratio between the size of the metal islands and the spatial period of the periodic surface structure composed of them [6]. Of course, this



The dots *1* and *2* represent the values of the effective polarizabilities a_{zz} and a_{xx}, respectively, calculated by the exact formulas. The curves *I* and *II* are associated with approximate relations (2) and (3), respectively, while *III* is the asymptotic expression (12).

spatial period must be much smaller than the incident radiation wavelength in order to calculate the effective dielectric constant in the quasi-static approximation.

Finally, we will study the solutions in the limit of very close spheres. The parameter $\beta_0 = \ln[l/R + (l^2/R^2 - 1)^{1/2}]$ now tends to zero. The method of calculating the corresponding asymptotic behavior of the functions $F^{(\pm)}(q, \beta_0)$ by means of the Mellin integral transform is described in [3]; note that the result for $F^{(-)}(1, \beta_0)$ given there contains an arithmetic error. Using the asymptotic expressions from the Appendix, we find the following approximate expressions for the effective polarizability of each of two spheres in an external field directed along the z-axis:

$$a_{zz} \approx R^3\left[2\zeta(3) - \frac{\zeta(2)^2}{C + \ln[2/(l^2/R^2 - 1)^{-1/2}]}\right] \quad (12)$$

or perpendicular to it

$$a_{xx} \approx R^2\left[\frac{3\zeta(3)}{4} - \frac{\ln 2}{2}(l^2/R^2 - 1)\right], \quad (13)$$

where $\zeta(s)$ is the Riemann zeta function and $C = 0.577\ldots$ is the Euler constant. When the external field is directed perpendicular to the z-axis, the effective polarizability differs by less than 2% from the value given by approximate formula (3), however close the spheres are. When the field is parallel to the z-axis, the

effective polarizability remains finite but its derivative with respect to R/l has a singularity. However, the domain of sharp variation of a_{zz} is small. Formulas (12) and (2) yield the same a_{zz} values at $R/l \approx 0.95$. As the distance between the spheres increases further, the approximation taking into account only the direct dipole–dipole interaction gets better and its deviation from the exact result becomes less than 1%, starting at $R/l \approx 0.80$. We believe that the main qualitative result following from the exact solution of the electrostatic problem obtained in this paper is that the contribution to the effective polarizability of the indirect interaction, including the induction of higher multipoles at the neighboring sphere, is small.

APPENDIX

The asymptotic form of functions (7) for argument β_0 tending to zero are as follows:

$$F^{(-)}(0, \beta_0) \sim \frac{1}{\beta_0}[C + \ln(2/\beta_0)] + \frac{\beta_0}{72}, \qquad F^{(+)}(0, \beta_0) \sim \frac{\ln 2}{\beta_0},$$

$$F^{(-)}(1, \beta_0) \sim \frac{\zeta(2)}{\beta_0^2} - \frac{1}{12}, \qquad F^{(+)}(1, \beta_0) \sim \frac{\zeta(2)}{2\beta_0^2} + \frac{1}{12},$$

$$F^{(-)}(2, \beta_0) \sim \frac{2\zeta(3)}{\beta_0^3} + \frac{1}{6\beta_0}, \qquad F^{(+)}(2, \beta_0) \sim \frac{3\zeta(3)}{2\beta_0^3}.$$

REFERENCES

1. V. A. Saranin, Usp. Fiz. Nauk **169**, 453 (1999).
2. J. M. Gérardy and M. Ausloos, Phys. Rev. B **22**, 4950 (1980).
3. H. Buchholz, *Elektrische und magnetische Potentialfelder* (Springer-Verlag, Berlin, 1957; Inostrannaya Literatura, Moscow, 1961).
4. T. Yamaguchi, S. Yoshida, and A. Kinbara, Thin Solid Films **18**, 63 (1973).
5. T. Yamaguchi, S. Yoshida, and A. Kinbara, Thin Solid Films **21**, 173 (1974).
6. G. A. Niklasson and H. G. Graighead, Thin Solid Films **125**, 165 (1985).

Translated by V. Gursky

Technical Physics, Vol. 45, No. 10, 2000, pp. 1241–1250. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 11–20.

Quasi-adiabatic Compression of a Plasma Column by a Longitudinal Magnetic Field

V. F. Ermolovich, A. V. Ivanovskiĭ, A. P. Orlov, and V. D. Selemir

All-Russia Research Institute of Experimental Physics, Russian Federal Nuclear Center, Sarov, Nizhni Novgorod oblast, 607190 Russia

Received June 8, 1999

Abstract—Approximate 1.5-dimensional MHD equations are derived that describe the quasi-adiabatic compression of a thin plasma column by a longitudinal magnetic field. The parameters of the compressed plasma are obtained analytically as functions of the initial conditions and longitudinal field. The stability of plasma compression against the Rayleigh–Taylor instability is investigated. It is shown that, in the *Z*–Θ-pinch geometry, increasing the longitudinal magnetic field makes it possible to achieve radial compression ratios of 20–30 without violating the cylindrical symmetry of the column. The possibility of thermonuclear ignition in a thin plasma column in a *Z*–Θ-pinch configuration is studied. The ranges of the initial plasma densities and temperatures and the initial lengths of the plasma column that are needed to achieve ignition in a plasma compressed by a factor of 20–30 are determined. The parameters of the electromagnetic energy source required to achieve such a high plasma compression are estimated. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

In the 1960s, the problem of adiabatic compression of a plasma column by a longitudinal magnetic field was investigated both theoretically and experimentally (see, e.g., [1] and other monographs). The main difficulty in achieving thermonuclear ignition in this way is associated with the onset of instabilities in a compressed plasma, so that the linear compression ratio is restricted to $\delta \sim 3$–5. In addition, high temperatures are difficult to achieve because of the outflow of heat along open magnetic field lines toward the ends of the chamber via electron heat conduction.

Rapid technological progress in the 1980s and 1990s increased interest in the dynamic stability of plasma compression in various systems [2–4]. In particular, in the *Z*–Θ-pinch geometry, the plasma shell was compressed to $\delta_m \sim 22$ [5]; the theoretically predicted compression ratio was $\delta_m \sim 30$ [6].

The advances achieved in deep stable compression of plasma shells make it relevant to analyze the prospects for adiabatic compression of a plasma column by its own longitudinal magnetic field in a *Z*–Θ-pinch configuration. Making the longitudinal field pulse more than ten times peakier than the current pulse that compresses the plasma shell [7] reduces thermal plasma losses due to electron heat conduction. Investigations in this area are very important because of the possibility of generating current pulses with an amplitude of up to 200 MA (at present, 50-MA current pulses are achievable) and a front rise time of about 3.5 μs [8].

Compression with the velocity $v_0 \cong r_0/t_m$ (where r_0 is the initial plasma radius and t_m is the compression time), which is lower than the magnetosonic velocity $v_m \cong \sqrt{v_S^2 + v_A^2}$ (where v_S is the speed of sound and v_A is the Alfvén speed), is adiabatic in nature. The momentum and pressure are carried from the plasma boundary toward the plasma center by small-amplitude magnetosonic waves. The subject of our paper is the process of slow quasi-adiabatic compression with running magnetosonic waves. The small parameter of the problem, $r_0/(v_m t_m)$, makes it possible to reduce the two-dimensional MHD equations to a set of quasi-one-dimensional equations.

Among the instabilities that arise in a plasma compressed by a longitudinal magnetic field, the fastest growing one is the Rayleigh–Taylor instability, which does not perturb the magnetic field lines [4]. For an infinitely long plasma column, the quasi-one-dimensional MHD equations reduce to a set of ordinary differential equations, which permit the problem of the stability of compression to be solved analytically.

Since the acceleration rates during compression are high, small-amplitude magnetosonic waves can strongly affect the stability of compression without a significant impact on the adiabatically varying plasma density and pressure. When analyzing the prospects for thermonuclear ignition, this circumstance allows us to study plasma compression as an adiabatic process (i.e., to neglect inertial terms in describing the motion perpendicular to the magnetic field), in which case the maximum possible compression ratios should be estimated by solving the stability problem. An ignition factor of $W \cong 0.3$–1 serves as a criterion for thermonuclear ignition of a plasma [9].

1063-7842/00/4510-1241$20.00

The plasma may be preheated in different ways: through the absorption of laser light in a D–T gas, by a shock wave (in the case of a stepped Θ-pinch), via the absorption of laser light in a plasma ablated from a laser-irradiated solid target, etc.

BASIC EQUATIONS

We consider a plasma column of length L that is confined in vacuum by the axial magnetic field $B = B_s(t)$. If the magnetic field diffusion is neglected, then the plasma density ρ, plasma temperature T, and the radial (v_r, B_r) and axial (v_z, B_z) components of the plasma velocity and magnetic field should be determined by solving the ideal MHD equations with $\gamma = 5/3$:

$$\begin{aligned}
&\frac{\partial \rho}{\partial t} + \frac{1}{r}\frac{\partial}{\partial r}(r\rho v_r) + \frac{\partial}{\partial z}(\rho v_z) = 0,\\
&\rho\left(\frac{\partial v_r}{\partial t} + v_r\frac{\partial v_r}{\partial r} + v_z\frac{\partial v_r}{\partial z}\right)\\
&= -\frac{\partial P}{\partial r} - \frac{1}{\mu_0}B_z\frac{\partial B_z}{\partial r} + \frac{1}{\mu_0}B_z\frac{\partial B_r}{\partial z},\\
&\rho\left(\frac{\partial v_z}{\partial t} + v_r\frac{\partial v_z}{\partial r} + v_z\frac{\partial v_z}{\partial z}\right)\\
&= -\frac{\partial P}{\partial z} - \frac{1}{\mu_0}B_r\frac{\partial B_r}{\partial z} + \frac{1}{\mu_0}B_r\frac{\partial B_z}{\partial r},\\
&\frac{\partial P}{\partial t} + \frac{1}{r}\frac{\partial}{\partial r}(rv_rP) + \frac{\partial}{\partial z}(v_zP) = (\gamma-1)\\
&\times\left[\frac{\partial}{\partial z}\left(\lambda_\parallel^e\frac{\partial T}{\partial z}\right) - P\left(\frac{1}{r}\frac{\partial}{\partial r}(rv_r) + \frac{\partial v_z}{\partial z}\right) - Q_r\right],\\
&\frac{\partial B_r}{\partial t} = \frac{\partial}{\partial z}(v_rB_z - v_zB_r),\\
&\frac{\partial B_z}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}[r(v_zB_r - v_rB_z)],
\end{aligned} \tag{1}$$

where $\lambda_\parallel^e$ is the electron thermal conductivity along the magnetic field lines and Q_r is the power of the volume radiative energy losses. We neglect the ion and electron cross-field thermal conductivities in Eqs. (1) and determine the plasma pressure from the equation of state $P = 2A\rho T$, where A is the gas constant per unit mass. From the condition $\mathrm{div}(\mathbf{B}) = 0$, we estimate B_r as

$$\begin{aligned}
B_r &= -\frac{1}{r}\int_0^r r\frac{\partial B_z}{\partial z}dr \approx -\frac{r}{2}\frac{\partial B_z}{\partial z}(r=0)\\
&\approx \frac{r_s}{2L}B_z(r=0),
\end{aligned} \tag{2}$$

where r_s is the radius of the plasma column. For a thin column ($r_s/L \ll 1$), we can neglect the effect of B_r on the plasma motion and ignore the longitudinal gradient of the transverse velocity $v_z\partial v_r/\partial z$ in comparison with the transverse gradient $v_r\partial v_r/\partial r$. Then, Eqs. (1) can be simplified to

$$\begin{aligned}
&\frac{\partial \rho}{\partial t} + \frac{1}{r}\frac{\partial}{\partial r}(r\rho v_r) + \frac{\partial}{\partial z}(\rho v_z) = 0,\\
&\rho\left(\frac{\partial v_r}{\partial t} + v_r\frac{\partial v_r}{\partial r}\right) = -\frac{\partial P}{\partial r} - \frac{1}{\mu_0}B_z\frac{\partial B_z}{\partial r},\\
&\rho\left(\frac{\partial v_z}{\partial t} + v_r\frac{\partial v_z}{\partial r} + v_z\frac{\partial v_z}{\partial z}\right) = -\frac{\partial P}{\partial z},\\
&\frac{\partial P}{\partial t} + \frac{1}{r}\frac{\partial}{\partial r}(rv_rP) + \frac{\partial}{\partial z}(v_zP) = (\gamma-1)\\
&\times\left[\frac{\partial}{\partial z}\left(\lambda_\parallel^e\frac{\partial T}{\partial z}\right) - P\left(\frac{1}{r}\frac{\partial}{\partial r}(rv_r) + \frac{\partial v_z}{\partial z}\right) - Q_r\right],\\
&\frac{\partial B_r}{\partial t} = -\frac{1}{r}\frac{\partial}{\partial r}[rv_rB_z].
\end{aligned} \tag{3}$$

We assume that, at the initial instant, the plasma is homogeneous and immobile; i.e., the plasma density is $\rho|_{t=0} = \rho_0$, the plasma pressure is $P|_{t=0} = P_0$, the velocity components are $v_r|_{t=0} = v_z|_{t=0} = 0$, and the axial magnetic field component, which satisfies the condition

$$(P + B_z^2/2\mu_0)\big|_{r=r_s(t)} = B_s^2/2\mu_0, \tag{4}$$

is determined from the equality $B_z^2\big|_{t=0} = B_0^2 - 2\mu_0P_0$, where $B_0 = B_s|_{t=0}$.

A representative solution to Eqs. (3) for an infinitely long plasma column ($\partial/\partial z \equiv 0$, $v_z \equiv 0$) is illustrated in Fig. 1, which shows the waveforms of the dimensionless coordinate $r_s' = r_s/r_0$, velocity $\vartheta = v_rt_m/r_0$, and acceleration $g = \partial\vartheta/\partial\tau + \vartheta\partial\vartheta/\partial r'$ of the plasma boundary (here, r_0 is the initial plasma radius, t_m is the compression time, and $\tau = t/t_m$) at the dimensionless Alfvén speed $\vartheta_A = [B_0t_m/(\mu_0\rho_0r_0)]^{1/2} = 30$ for the plasma parameter $\beta = (2\mu_0P_0)/B_0^2 = 0.5$. Against the background of slow adiabatic compression, we can see high-frequency oscillations, which are especially pronounced in the waveform of the acceleration. The oscillations are associated with the excitation of magnetosonic waves, which carry the momentum and pressure from the plasma boundary toward the plasma center.

Analysis of the spatial profiles shows that, for $\vartheta_A \geq 10$, the amplitude of the magnetosonic waves is small. This indicates that the jumps in the plasma density and pressure associated with magnetosonic waves distort the adiabatic profiles only slightly. Consequently, as a first approximation, we can assume that the plasma is com-

pressed uniformly, so that $\rho(z, r, t) = \rho(z, t)$, $P(z, r, t) = P(z, t)$, and the continuity equation in (3) gives

$$v_r(r, z, t) = v_s(z, t)\frac{r}{r_s(z, t)},$$
$$v_s(z, t) = \frac{\partial r_s(z, t)}{\partial t}, \quad v_z(r, z, t) = v_z(z, t). \tag{5}$$

From expressions (5) and the last equation in (3), we obtain the following relationship of v_r to the axial component of the magnetic field in a plasma B_z or, equivalently, to the radius r_s of the plasma column:

$$v_r = -0.5S(z, t)r, \tag{6}$$

where

$$S(z, t) = \frac{1}{B_z(z, t)}\frac{\partial B_z(z, t)}{\partial t},$$
$$B_z(z, t) = \sqrt{B_0^2 - 2\mu_0 P_0}\frac{r_0^2}{r_s^2(z, t)}.$$

Substituting v_r from (6) into the first, third, and fourth equations in (3) yields

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial z}(\rho v_z) = \rho S(z, t),$$
$$\rho\left(\frac{\partial v_z}{\partial t} + v_z\frac{\partial v_z}{\partial z}\right) = -\frac{\partial P}{\partial z},$$
$$\frac{\partial P}{\partial t} + \frac{\partial}{\partial z}(v_z P) = (\gamma - 1)\left[\frac{\partial}{\partial z}\left(\lambda_\parallel^e \frac{\partial T}{\partial z}\right) - P\frac{\partial v_z}{\partial z} - Q_r\right] + \gamma P S(z, t), \tag{7}$$
$$S(z, t) = \frac{1}{B_z(z, t)}\frac{\partial B_z(z, t)}{\partial t},$$
$$B_z(z, t) = \sqrt{B_0^2 - 2\mu_0 P_0}\frac{r_0^2}{r_s^2(z, t)}.$$

In order to close the set of Eqs. (7), which describe quasi-adiabatic compression of a thin plasma column, we need to relate the radius r_s of the plasma column to the plasma parameters. This relationship can be found from the equation for radial plasma motion in (3) and expressions (5):

$$\rho r\frac{\partial v_s}{\partial t} = -r_s\frac{\partial}{\partial r}\left(P + \frac{B_z^2}{2\mu_0}\right). \tag{8}$$

We integrate (8) over dr from 0 to r_s and take into account boundary condition (4) to obtain

$$m\frac{\partial v_s}{\partial t} = 2\pi r_s\left(P(r = 0) + \frac{B_z^2(r = 0)}{2\mu_0} - \frac{B_s^2}{2\mu_0}\right). \tag{9}$$

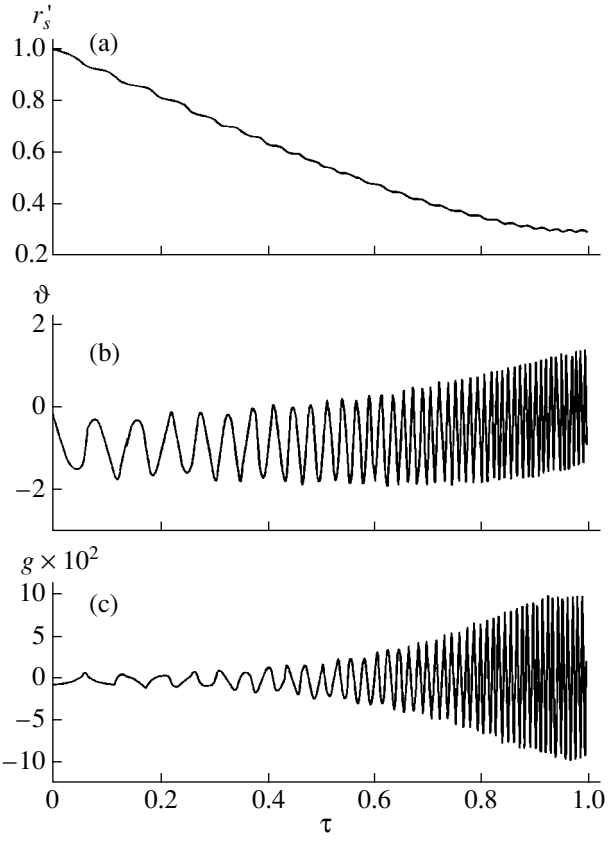


Fig. 1. Waveforms of (a) the coordinate, (b) velocity, and (c) acceleration of the boundary of the plasma column.

We estimate the plasma mass per unit length as $m = 2\pi\int_0^{r_s}\rho r dr \cong \pi r_s^2\rho$ ($r = 0$). In the uniform compression approximation, which implies that $\rho = \rho(r = 0)$, $P = P(r = 0)$, and $B_z = B_z(r = 0)$, we use relationship (6) to obtain

$$\rho r_s\frac{\partial^2 r_s}{\partial t^2} = 2\left(P + \frac{B_z^2}{2\mu_0} - \frac{B_s^2}{2\mu_0}\right). \tag{10}$$

To solve Eqs. (7) and (10), we need to specify the boundary conditions at the ends of the plasma column. We assume that the column terminates at the rigid end plates, where $-v_z(z = 0, L) = 0$ and the heat fluxes $-(\lambda_\parallel^e \partial T/\partial z)|_{z=0} = q_1$ and $-(\lambda_\parallel^e \partial T/\partial z)|_{z=L} = q_2$ or the temperatures $T|_{z=0} = T_1$ and $T|_{z=L} = T_2$ are assumed to be given.

A comparative analysis of numerical solutions to Eqs. (3), (4), (7), and (10) for an infinitely long plasma column shows that magnetosonic waves in the exact analytic solution correspond to uniform macroscopic oscillations of the system parameters in the quasi-adiabatic model. In other words, the quasi-adiabatic model yields a solution that is accurate to the time required for a magnetosonic wave to pass from the plasma center to

the plasma boundary. In the limit $\vartheta_A \longrightarrow \infty$, the quasi-adiabatic amplitudes of the oscillations of the velocity and acceleration asymptotically approach the exact amplitudes. For $\vartheta_A \geq 10$, the exact and quasi-adiabatic waveforms of the acceleration (which experiences the most pronounced oscillations) differ by no more than 40% (Fig. 1). This circumstance enables us to examine the stability of compression of the plasma column by applying the quasi-adiabatic model.

ANALYSIS OF THE SOLUTION

For an infinitely long plasma column and $Q_r = 0$, Eqs. (7) and (10) for the dimensionless parameters $r_s' = r_s/r_0$, $\tau = t/t_m$, $\rho' = \rho/\rho_0$, $p = P/P_0$, $H = B_z/B_0$, and $f(\tau) = B_s(\tau)/B_0$ reduce to

$$\ddot{r}_s' = \vartheta_A^2[(1-\beta)r_s'^{-3} + \beta r_s'^{-2\gamma+1} - r_s' f^2(\tau)],$$
$$\rho' = r_s'^{-2}, \quad p = r_s'^{-2\gamma}, \quad H = \sqrt{1-\beta}\, r_s'^{-2}, \qquad (11)$$
$$r_s'\big|_{\tau=0} = 1, \quad r_s'\big|_{\tau=0} = 0.$$

where we introduce the notation $\ddot{r}_s' = \partial^2 r_s'/\partial\tau^2$ and $\dot{r}_s'/\partial\tau$, $\dot{r}_s' = \partial r_s'/\partial\tau$.

To linearize the first equation in (11), we represent r_s' as (Fig. 1)

$$r_s'(\tau) = r_0(\tau) + r_1(\tau), \quad |r_1| \ll r_0, \qquad (12)$$

where $r_0(\tau)$ is the solution to Eqs. (11) in the limit $\vartheta_A \longrightarrow \infty$ (below, this solution will be referred to as the adiabatic solution).

We substitute (12) into (11) and expand the right-hand side of the resulting equation in a Taylor series in the vicinity of the point $r_s' = r_0$ to arrive at the following equations for $r_1(\tau)$:

$$\beta r_0^{-2\gamma} + (1-\beta)r_0^{-4} = f^2(\tau),$$
$$\ddot{r}_1 = \vartheta_A^2[-3(1-\beta)r_0^{-4} + \beta(1-2\gamma)r_0^{-2\gamma} - f^2(\tau)]r_1 - \ddot{r}_0. \qquad (13)$$

Here, the mean (adiabatic) radius r_0 of the plasma column is determined from the first equation and is then inserted into the second equation to find a rapidly oscillating correction from magnetosonic waves. For diamagnetic ($\beta = 1$) and magnetized ($\beta = 0$) plasmas, Eqs. (13) give

$$\ddot{r}_s' + 2\alpha_1\vartheta_A^2 f^2(\tau)r_s' = 2\alpha_1\vartheta_A^2 f^2(\tau)r_0, \qquad (14)$$

where

$$\alpha_1 = \begin{cases}\gamma, & \beta = 1,\\ 2, & \beta = 0,\end{cases} \quad r_0 = f^{-1/\alpha_1}.$$

Let us consider three ways of amplifying the axial magnetic field.

(i) The field component B_s is produced by a linearly increasing solenoidal current, as is the case with a classical Θ-pinch:

$$f(\tau) = 1 + (\alpha - 1)\tau. \qquad (15)$$

(ii) The magnetic flux is compressed by an accelerated cylindrical liner:

$$f(\tau) = \alpha[1 + (\alpha-1)(1-\tau)^2]^{-1}. \qquad (16)$$

This method is implemented in magnetocumulative generators (MCG) [10]. In (15) and (16), the factor $\alpha = B_m/B_0$ reflects the amplification of the magnetic field.

(iii) The magnetic flux is compressed by a cylindrical liner accelerated by the current, as in a *Z*–Θ-pinch [6]. In this case, for the current $J(t) = J_m h(\tau)$, the waveform of the magnetic field pulse can be represented as $f(\tau) = \xi^2$, where the dimensionless radius $1/\xi(\tau)$ of the liner is to be found by solving the equation [7]

$$\frac{d^2\xi}{d\tau^2} = \frac{a}{\xi}\left[-h^2(\tau) + \frac{(1-\xi^4)}{b^2\xi^2}\right],$$
$$a = \frac{\mu_0}{4\pi}\frac{J_m^2 t_m^2}{m a_a^2}, \quad b = \frac{\mu_0}{2\pi}\frac{J_m}{B_0 a_0}, \qquad (17)$$
$$\xi|_{\tau=0} = 1, \quad \left.\frac{d\xi}{d\tau}\right|_{\tau=0} = 0.$$

Here, m is the liner mass per unit length and a_0 is the initial radius of the liner. For a magnetized plasma ($\beta \ll 1$), the quantity $b \cong \alpha^{1/2}$ is close to the maximum compression ratio δ of a plasma. If, in the range $10 \leq \alpha \leq 100$, the current depends on time as $h(\tau) = \sin(\pi/2\tau)$, then the compression will be maximum for $a \approx 4$ at the time $\tau \approx 1.1$.

In the quasi-adiabatic approximation, Eq. (14) in the limit $\vartheta_A \longrightarrow \infty$ is solved by expanding in powers of $1/\vartheta_A$ [11]. For a Θ-pinch and MCG, we retain only the first terms of the series expansion to obtain

$$g(\tau) = \ddot{r}_0'(\tau) \cong \ddot{r}_0(\tau) + \sqrt{2\alpha_1}\vartheta_A \dot{r}_0(0) f^{3/2}(\tau) \times \sin\left(\sqrt{2\alpha_1}\vartheta_A \int_0^\tau f(\tau')d\tau'\right), \qquad (18)$$

where $\ddot{r}_0(\tau)$ is the acceleration during adiabatic motion and the initial velocity $\dot{r}_0(0)$ of the plasma boundary is equal to $\dot{r}_0(0) = -(\alpha-1)/\alpha_1$ for a Θ-pinch and $\dot{r}_0(0) = -2/\alpha_1(\alpha-1)/\alpha$ for an MCG.

For a *Z*–Θ-pinch, the first three derivatives of the adiabatic radius r_0 vanish at $\tau = 0$ and we have $\dddot{r}_0'(0) =$

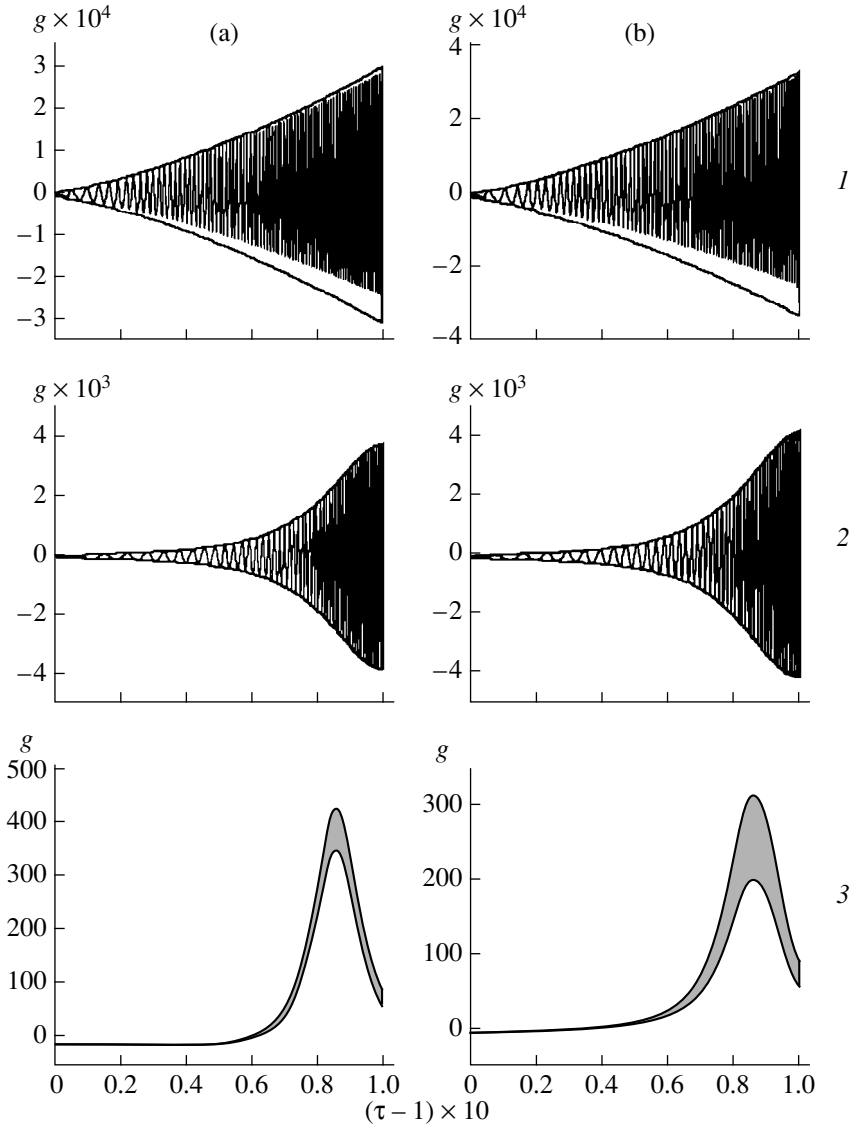


Fig. 2. Waveforms of the acceleration at $\vartheta_A^2 = 10^3$ and β = (a) 0 and (b) 1 for (*1*) classical Θ-pinch with $\alpha = 16$, (*2*) MCG with $\alpha = 16$, and (*3*) *Z*–Θ-pinch with $\alpha = 400$.

$-\pi^2 a/\alpha_1$ [see (7)]. To within terms of order ϑ_A^{-4}, we obtain

$$g(\tau) \cong \ddot{r}_0(\tau) + \frac{\dddot{r}_0(0)}{2\alpha_1\vartheta_A^2} f^{3/2}(\tau) \times \cos\left(\sqrt{2\alpha_1}\vartheta_A \int_0^\tau f(\tau')d\tau'\right). \quad (19)$$

Figure 2 shows the waveforms of the acceleration that were obtained by solving Eq. (11) and the envelopes of the waveforms that were obtained from (18) and (19) for sin(…) and cos(…) = ±1. We can see that the waveforms $g(\tau)$ for an MCG and *Z*–Θ-pinch are described well by expansions (18) and (19). For a Θ-pinch (Fig. 2a), the deviation of the asymptotic solution from the exact one stems from two causes: the increase in the expansion coefficients and the larger value of α. Refining the asymptotic solution requires taking higher order terms in the series expansion into account.

Hence, the solution to our problem can be represented as the sum of a slowly varying (adiabatic) part and a rapidly oscillating part, whose oscillation frequency increases with time. The coordinate of the plasma boundary is primarily determined by the first term of the series expansion. The rapidly oscillating part describes acceleration of the plasma boundary; i.e., it allows us to investigate whether or not the plasma compression is stable, depending on the compression mechanism. For a classical Θ-pinch and MCG, the acceleration of the plasma boundary is described by the rapidly oscillating part of the solution, whereas, for a *Z*–Θ-pinch, it is described by the adiabatic part. In other

words, for a classical Θ-pinch and MCG, the amplitude g_m of the rapidly oscillating component of the acceleration increases linearly with increasing Alfvén speed ϑ_A, whereas, for a *Z*–Θ-pinch, the amplitude g_m is inversely proportional to the square root of ϑ_A, so that the solution asymptotically approaches the adiabatic solution.

STABILITY OF COMPRESSION

In the linear stage of compression, the relative growth rate of the perturbations is estimated from the formula

$$\Gamma = \exp\left(\int_0^1 \sigma(\tau)d\tau\right). \tag{20}$$

In a plasma compressed by a longitudinal magnetic field, the growth rate of the Rayleigh–Taylor instability is

$$\sigma(\tau) = \begin{cases} 0, & g(\tau) \geq 0, \\ \sqrt{\dfrac{m}{r_s'(\tau)}|g(\tau)|}, & g(\tau) < 0. \end{cases} \tag{21}$$

We assume that the cylindrical symmetry of the compressed column is violated at $\Gamma = \Gamma_m = 100$ [2]. Assuming that $g(\tau)$ is governed by its rapidly oscillating component [see (18)] and that the radius of the plasma column changes adiabatically, i.e., $r_s'(\tau) \cong r_0(\tau) = f^{-1/\alpha_1}$, we obtain the following expression for a Θ-pinch and MCG:

$$\int_0^1 \sigma(\tau)d\tau \cong \frac{(2\alpha_1)^{1/4}}{2}\sqrt{m\vartheta_A|\dot r_0(0)|}\int_0^1 f^{3/4+1(2\alpha_1)} \times \sqrt{\left|\sin\left[\sqrt{2\alpha_1}\vartheta_A\int_0^1 f(\tau)d\tau\right]\right|}\,d\tau. \tag{22}$$

In the limit $\vartheta_A \gg 1$, the right-hand side of (22) can be estimated by the sum over the sine half-periods. Then, switching from summation to integration over $d\tau$, we obtain

$$\int_0^1 \sigma(\tau)d\tau = \frac{(2\alpha_1)^{1/4}}{2\sqrt{\pi}}\Gamma(3/4)\sqrt{m\vartheta_A|\dot r_0(0)|} \times \int_0^1 f^{3/4+1(2\alpha_1)}d\tau, \tag{23}$$

$$\Gamma(3/4) = 1.2254.$$

For a strongly magnetized plasma ($\beta = 0$; i.e., $\alpha_1 = 2$), this formula assumes an especially simple form:

$$\int_0^1 \sigma(\tau)d\tau = \frac{\Gamma(3/4)}{\sqrt{2\pi}}\sqrt{m\vartheta_A|\dot r_0(0)|}\int_0^1 f(\tau)d\tau. \tag{24}$$

From (20), (24), (15), and (16), we obtain the following stable compression conditions for a magnetized plasma:

$$\vartheta_A = \frac{2\pi}{m}\frac{\ln^2(\Gamma_m)}{\Gamma^2(3/4)}\begin{cases} \dfrac{8(\alpha-1)}{(\alpha^2-1)^2} & \text{for a Θ-pinch,} \\ \dfrac{1}{\alpha\arctan(\sqrt{\alpha-1})} & \text{for an MCG.} \end{cases} \tag{25}$$

As ϑ_A and m increase, the plasma compression becomes less stable. At $\vartheta_A = 10$ and $m = 1$, the maximum extent to which the magnetic field can be amplified without violating the cylindrical symmetry of the compressed column is $\alpha = 3.6$ ($\delta \cong 2$) for a Θ-pinch and $\alpha = 5.5$ ($\delta \cong 2.4$) for an MCG. This indicates that, for a Θ-pinch and MCG, the plasma column cannot be compressed quasi-adiabatically ($\vartheta_A \geq 10$) to $\delta \geq 3$ without violating the cylindrical symmetry. The compression of a diamagnetic plasma column may be inherently more stable, because the rapidly oscillating part of $g(\tau)$ is partially suppressed by electron heat conduction.

For a *Z*–Θ-pinch, $g(\tau)$ is primarily governed by the adiabatic part of the solution. The solution to Eqs. (14), (17), (20), and (21) is illustrated in Fig. 3, which shows the maximum possible amplification factor α of the magnetic field as a function of ϑ_A for different values of m. We can see that the values of α are much higher than those attainable in a Θ-pinch and MCG.

The maximum achievable compression ratio is governed by the highest growing mode of the Rayleigh–Taylor instability. The growth of Rayleigh–Taylor modes with large numbers m is partially suppressed by dissipative processes: transverse diffusion, plasma resistance, plasma viscosity, etc. [1, 12]. The compression conditions are improved by imposing a slightly sheared magnetic field, such that it has essentially no effect on the compression dynamics [4]. Let us estimate the maximum possible compression ratios δ when the Rayleigh–Taylor modes are partially suppressed by a sheared magnetic field.

In the presence of a constant current I_0 flowing along the plasma surface, the azimuthal magnetic field near the plasma boundary evolves according to the law

$$B_\varphi(t) = \frac{I_0}{2\pi r_s(t)}. \tag{26}$$

For uniform compression of a plasma with an azimuthal magnetic field (26), the growth rate $\sigma(t)$ of the

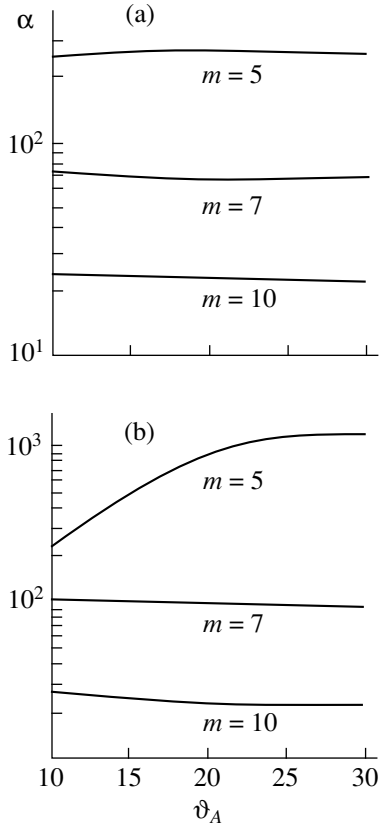


Fig. 3. Maximum possible amplification factor α of the magnetic field vs. the Alfvén speed ϑ_A for different values of *m*. Cases (a) and (b) are the same as in Fig. 2.

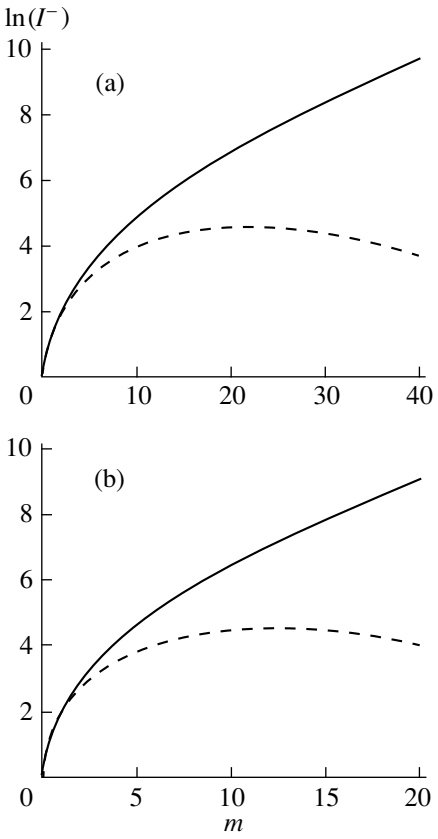


Fig. 4. Total growth rate of the Rayleigh–Taylor instability vs. the mode number at $\vartheta_A = 15$ and $\beta = 0$ for (a) $B_\perp = 0.02$ and $\alpha = 32.3$ and (b) $B_\perp = 0.03$ and $\alpha = 247$. The solid and dashed curves are for $B_\perp = 0$ and $B_\perp \neq 0$, respectively.

Rayleigh–Taylor instability has the form [4, 12]

$$\sigma^2(t) = \sqrt{k^2 + \frac{m^2}{r_s^2(t)}}|g(t)| - \frac{[kB(t)]^2 + \left[kB_s(t) + m\frac{mB_\varphi(t)}{r_s}\right]^2}{\mu_0\rho(t)}, \quad (27)$$

where $|k| = 2\pi/\lambda$ and λ is the perturbation wavelength along the plasma axis.

Taking into account (11), we can write the dimensionless growth rate $\sigma(\tau) = \sigma(t)t_m$ as

$$\sigma^2(\tau) = \sqrt{k'^2 + \frac{m^2}{r_s'^2}}|g(t)| - \frac{\vartheta_A^2}{r_s'^2} \times \{(1-\beta)k'^2 + [k' f(\tau) r_s'^2 + mB_\perp]^2\}, \quad (28)$$

where $k' = kr_0$, $g(\tau) = g(t)t_m^2/r_0$, $B_\perp = I_0/(2\pi r_0 B_0)$, and $f(\tau)$ satisfies Eq. (17).

The second and third terms on the right-hand side of (28) describe the stabilization of Rayleigh–Taylor modes with large numbers *m* due to additional tension of the lines of the sheared magnetic field in a plasma and vacuum, respectively.

Figure 4 shows a comparison of the growth rates of the Rayleigh–Taylor modes with different numbers *m* for a magnetized plasma ($\beta = 0$) at $\vartheta_A = 15$ for $B_\perp = 0$, 0.02, and 0.03. In the calculations performed for $B_\perp \neq 0$, we adjusted the parameter k' ($k' \ll 1$) so that the mode with a fixed number *m* grows at the highest possible rate. The larger the mode number *m*, the lower the level at which the mode is saturated by a sheared magnetic field. For $B_\perp = 0.02$, the profile of the total growth rate of the Rayleigh–Taylor instability peaks at the mode number $m = 21$, whereas, for $B_\perp = 0.03$, the fastest growing mode is the one with the number $m = 12$. As a result, the maximum possible compression ratio increases from $\delta = 5.69$ ($\alpha = 32.3$) to $\delta = 15.7$ ($\alpha = 243$). As ϑ_A increases during compression, the compression stability improves. For $B_\perp = 0.02$, the compression ratio $\delta = 30$ is achieved at $\vartheta_A \cong 27$; and for $B_\perp = 0.03$, this ratio is achieved at $\vartheta_A \cong 17.5$. In the *Z*–Θ-pinch geom-

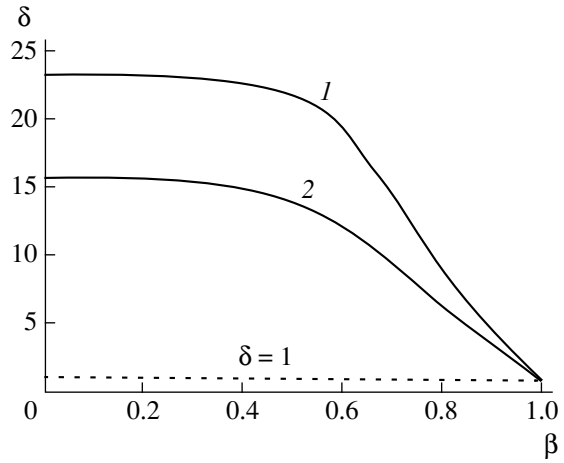


Fig. 5. Maximum possible compression ratio vs. degree of plasma magnetization for (*1*) $\vartheta_A = 25$ and $B_\perp = 0.02$ and (*2*) $\vartheta_A = 15$ and $B_\perp = 0.03$.

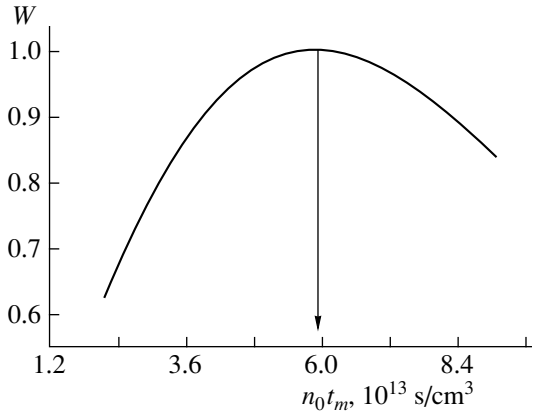


Fig. 6. Profiles of the ignition factor as a function of the product of the initial plasma density and the compression time for $\beta = 0.1$, $\delta = 20$, and $T_0 = 232$ eV.

etry for $\vartheta \geq 20$, the stability of compression of the plasma column is actually governed by the stability of compression of the plasma shell, where the compression ratio is $\delta_m \cong 22$–30 [5, 6].

For a diamagnetic plasma ($\beta = 1$), each mode with the number m is characterized by its own parameter

$$k' = -m\frac{B_\perp}{r_s'^2(\tau)f(\tau)} \approx -mB_\perp, \qquad (29)$$

which drives the stabilizing term in (28) to zero. This indicates that, in a diamagnetic plasma, the sheared magnetic field is ineffective in saturating the Rayleigh–Taylor modes with larger numbers m at lower levels. Figure 5 presents the maximum possible compression ratio as a function of β. We can see that the stable compression conditions are nearly optimum in plasmas with $\beta \leq 0.5$.

PROSPECTS FOR THERMONUCLEAR IGNITION

Here, we analyze whether it is possible to ignite a heated 50–50% D–T fuel plasma in the Z–Θ-pinch geometry. The criterion for thermonuclear ignition of a plasma with mass M can be formulated as the following condition on the ignition factor W [9]:

$$W = \frac{1}{M}\int\int_{V(t)}\frac{Q_{nf}}{E_T}dVdt \cong 0.3 - 1, \qquad (30)$$

where Q_{nf} is the released fusion power deposited in a unit plasma volume and E_T is the internal energy per unit mass of a compressed plasma (this energy is calculated for a cold gas, without consideration of fusion power release).

We assume that neutrons freely leave the plasma and do not contribute to Q_{nf}. In the longitudinal magnetic field $B_z = B_0(r_0/r_s)^2 = B_0\delta^2$ of the plasma column, the Larmor radius $r_\alpha = (2E_\alpha m_\alpha)^{1/2}/(eB_z)$ of an α-particle with the energy $E_\alpha \sim 3.5$ MeV is determined by

$$\frac{r_\alpha}{r_s} = \frac{\sqrt{2E_\alpha m_\alpha}}{eB_0 r_0}\frac{1}{\delta} \cong \frac{0.54}{H_0 r_0}\frac{1}{\delta}, \qquad (31)$$

where the initial magnetic field strength H_0 and the initial radius r_0 of the plasma column are expressed in MG and cm, respectively. For $H_0 \geq 0.1$ MG and $\delta \geq 10$, we have $r_s \geq r_\alpha$ ($r_0 \sim 1$ cm), so that α-particles always remain in the plasma column.

The power Q_r of the radiative energy losses is governed by internal bremsstrahlung, which is most intense in the soft X-ray spectral range. The energy losses associated with magneto-bremsstrahlung can be neglected, because the emitted infrared, visible, and ultraviolet radiations can propagate in a plasma only over short distances and are reflected by the liner system [13].

In studying thermonuclear ignition processes, we neglect the rapidly oscillating part of the solution (this is equivalent to omitting inertial terms when describing the motion perpendicular to the magnetic field). As a result, Eq. (10) reduces to the following equation for the axial magnetic field B_z:

$$P + \frac{B_z^2}{2\mu_0} = \frac{B_s^2}{2\mu_0}. \qquad (32)$$

For an infinitely long plasma column, investigating the prospects for thermonuclear ignition in terms of

Eqs. (7), (30), and (32) reduces to analyzing the solution to the equations

$$\frac{1}{T}\frac{\partial T}{\partial \tau} - (\gamma - 1)\frac{1}{\eta}\frac{\partial \eta}{\partial \tau} = -(\gamma - 1)\frac{q_r(T)}{2kT_\eta}(t_m n_0),$$
$$T_0 + \beta T\eta = T_0 \eta^2 f^2(\tau, b),$$
$$W = (\gamma - 1)(t_m n_0)\int_0^{\tau_k} E_\alpha \frac{q_{nf}(T)}{2kT\eta} d\tau, \quad (33)$$
$$T|_{\tau = 0} = T_0, \quad \eta|_{\tau = 0} = 1,$$

where n_0 is the initial plasma density; t_m is the time at which the field component $B_s(t)$ becomes maximum; $\eta = \rho/\rho_0$ is the degree of compression of the plasma column at the time τ; q_{nf} (cm^3/s) = $Q_{nf}/(E_\alpha n_D n_T)$ is the rate constant of the D–T reaction; $E_\alpha \cong 3.5$ MeV is the energy of an α-particle; q_r (erg cm^3/s) = Q_r/n^2 is the rate coefficient for energy losses due to bremsstrahlung; and n, n_D, and n_T are the densities of plasma particles and deuterium and tritium ions, respectively ($n = n_T + n_D$).

Solution (33) is determined by the parameters β, T_0, $n_0 t_m$, and $b \cong \delta$ ($\beta \ll 1$). Figure 6 displays a representative profile of the ignition factor as a function of the parameter $n_0 t_m$. During long-term plasma compression, bremsstrahlung-induced energy losses increase and the plasma density becomes higher. As a result, the function $W(n_0 t_m)$ increases, reaches its maximum W_m, and then decreases. For the above values of the parameter β and compression ratio δ, the maximum W_m and the value of $n_0 t_m$ at which the maximum is reached are both determined by the initial plasma temperature T_0. When $\delta = 20$ and $\beta = 0.1$, the maximum $W_m = 1$ is achieved for $T_0 \cong 230$ eV at $(n_0 t_m) = (n_0 t_m)_{opt} = 6 \times 10^{13}$ s/cm^3. This is the optimum parameter value for satisfying the thermonuclear ignition criterion for the compression ratio $\delta = 20$. For $\delta = 30$ ($\beta = 0.1$), we have $T_0 \cong 130$ eV, so that $(n_0 t_m)_{opt} = 4.2 \times 10^{13}$ s/cm^3. We can conclude that, for a fixed compression ratio, the parameter $n_0 t_m$ determines the optimum condition for igniting a pulsed system, in analogy with the Lawson criterion ($n\tau$) for controlled nuclear fusion. The parameters required to achieve ignition ($W_m \cong 1$) can be estimated from the formulas

$$B_0 = \sqrt{\frac{4\mu_0 k T_0 n_0}{\beta}}, \quad n_0 = \frac{(t_m n_0)_{opt}}{t_m},$$
$$J_m \cong 4\pi R_0 \delta \sqrt{\frac{n_0 k T_0}{\mu_0 \beta}}. \quad (34)$$

Thus, for the source described in [8] ($t_m \cong 3.5$ μs) and for $\delta = 20$, we have $n_0 = 1.7 \times 10^{19}$ cm^{-3}, $B_0 = 1.8$ MG, and $J_m \cong 360$ MA, where the initial radius of the outer liner is $R_0 = 2$ cm. For $\delta = 30$, the corresponding parameter values are $n_0 = 1.2 \times 10^{19}$ cm^{-3}, $B_0 = 1.1$ MG, and $J_m \cong 340$ MA. We can see that the current required to achieve $W_m \cong 1$ is slightly sensitive to the maximum possible compression ratio δ, because the increase in δ is canceled by a decrease in the initial temperature T_0 and density n_0 of the plasma. The above estimates refer to the value $\beta = 0.1$. For β equal to 0.4 (Fig. 5), the initial magnetic field B_0 and current J_m can be lower by a factor of approximately two [see formulas (34)].

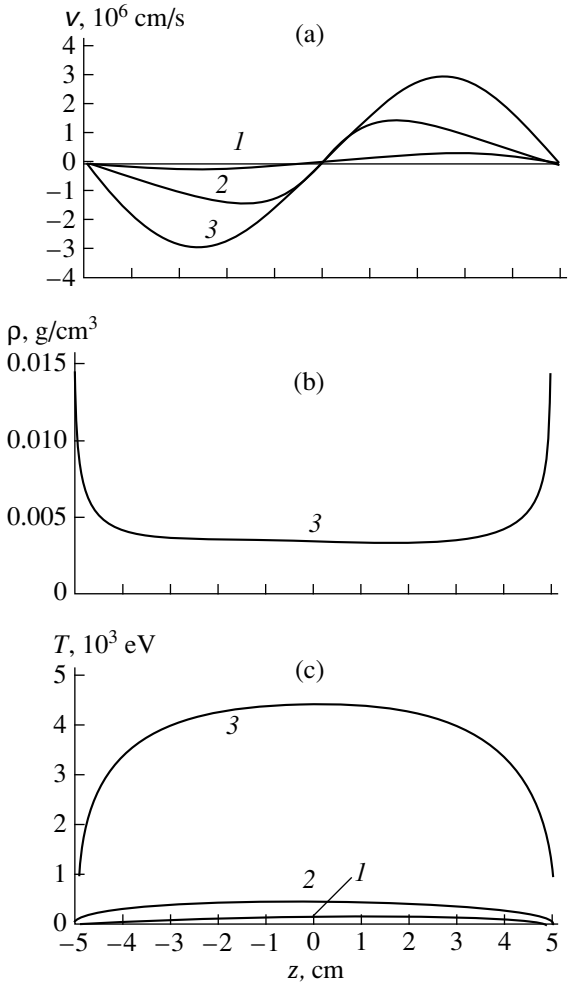


Fig. 7. Spatial profiles of (a) the velocity, (b) density, and (c) temperature of the plasma at the instants (*1*) $0.5t_m$, (*2*) t_m, and (*3*) $1.086t_m$ (the time of maximum compression).

In order to investigate how the energy losses associated with heat outflow toward the chamber walls via electron heat conduction affect plasma compression, we solved Eqs. (7) and (32) with the boundary conditions $T|_{z = \pm L/2} = 0$ (corresponding to the maximum losses) at the ends of the plasma column. Figure 7 shows representative spatial profiles of the plasma parameters at different times. The profiles were calculated for $r_0 = 1$ cm, $\beta = 0.1$, $\delta = 20$, and $t_m = 3.5$ μs, where the initial plasma density and temperature were

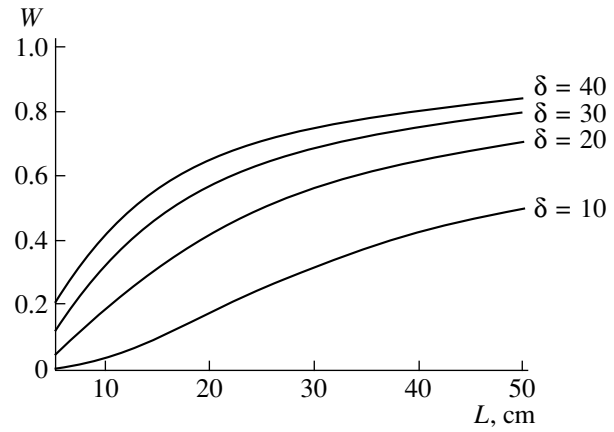


Fig. 8. Profiles of the ignition factor as a function of the length of the plasma column for $r_0 = 1$ cm.

$n_0 = 2.5 \times 10^{18}$ cm^{-3} and $T_0 = 100$ eV. As the end plates cool and the plasma pressure is equalized by hydrodynamic plasma motion, low-temperature dense plasma layers form near the plates.

The effect of the finite dimensions of the plasma column on the value of the ignition factor defined in (30) is illustrated in Fig. 8. For each value of δ, the initial plasma temperature and density were chosen so as to achieve the value $W_m = 1$ in the limit $L \longrightarrow \infty$. For compression ratios in the range $\delta = 20$–30, the values $W_m \geq 0.3$ are reached in a plasma column of length $L \geq$ 10–15 cm.

CONCLUSION

We have derived equations describing the compression of a thin plasma column by a longitudinal magnetic field $B_s(t)$ in the quasi-adiabatic approximation. We have investigated the stability of compression against the Rayleigh–Taylor instability. During the time t_m in which the field $B_s(t)$ is increased by a linearly increasing solenoidal current (in the case of a Θ pinch) or by compressing the magnetic flux by an accelerated cylindrical liner (e.g., in the case of an MCG), the compression stability worsens, whereas, in the Z–Θ pinch geometry, the compression stability improves as $B_s(t)$ increases.

For a Θ-pinch and MCG, the plasma column cannot be compressed quasi-adiabatically ($t_m \geq 10 r_0/v_m$) to $\delta \geq 3$ without violating the cylindrical symmetry. For a Z–Θ-pinch and $t_m \geq 15 r_0/v_m$, it is possible to achieve compression ratios of several tens without violating the cylindrical symmetry. In fact, in a Z–Θ-pinch configuration, the maximum possible plasma compression ratios are governed by the maximum achievable compression ratios $\delta_m \cong 20$–30 of the liner system.

To achieve thermonuclear ignition conditions for compression ratios of 20–30 in the Z–Θ-pinch geometry requires preheating of a D–T fuel plasma with the parameter $n_0 t_m \cong 4$–6×10^{13} s/cm^3 to a temperature of about $T_0 \cong 130$–230 eV. To ensure these conditions for $t_m = 3.5$ μs requires a plasma column of length $L \geq 10$–15 cm, initial axial magnetic field $B_0 \cong 0.5$–1 MG, and initial current $J_m \cong 170$–180 MA ($\beta = 0.4$). At present, it is possible in principle to generate current pulses with an amplitude of up to $J_m \cong 200$ MA (50-MA current pulses have already been produced) [8].

ACKNOWLEDGMENTS

We are grateful to V.V. Gorokhov, V.A. Demidov, V.A. Zhmaĭlo, V.I. Karelin, P.B. Repin, and V.T. Selyavskiĭ for fruitful discussions of some aspects of this work.

REFERENCES

1. N. A. Krall and A. W. Trivelpiece, *Principles of Plasma Physics* (Academic, New York, 1973; Mir, Moscow, 1975).
2. A. B. Bud'ko, A. L. Velikovich, A. I. Kleev, *et al.*, Zh. Éksp. Teor. Fiz. **95**, 496 (1989) [Sov. Phys. JETP **68**, 279 (1989)].
3. A. B. Bud'ko, F. S. Felber, A. I. Kleev, *et al.*, Phys. Fluids B **1**, 598 (1989).
4. A. B. Bud'ko, M. A. Liberman, A. L. Velikovich, and F. S. Felber, Phys. Fluids B **2**, 1159 (1990).
5. F. S. Felber, M. M. Malley, F. J. Wessel, *et al.*, Phys. Fluids **31**, 2053 (1988).
6. S. M. Golberg, M. A. Liberman, and A. L. Velikovich, Plasma Phys. Controlled Fusion **32**, 319 (1990).
7. F. S. Felber, F. J. Wessell, N. C. Wild, *et al.*, J. Appl. Phys. **64**, 3831 (1988).
8. V. A. Demidov, S. A. Kazakov, A. S. Kravchenko, *et al.*, in *Proceedings of 11th IEEE International Pulsed Power Conference, Baltimore, MD, 1997*, Vol. 2, p. 1459.
9. Yu. S. Vakhrameev, V. N. Mokhov, and N. A. Popov, At. Énerg. **49** (2), 121 (1980).
10. H. Knoepfel, *Pulsed High Magnetic Fields* (North-Holland, Amsterdam, 1970; Mir, Moscow, 1972).
11. M. V. Fedoryuk, *Ordinary Differential Equations* (Nauka, Moscow, 1985).
12. V. E. Golant, A. P. Zhilinskii, and I. E. Sakharov, *Fundamentals of Plasma Physics* (Atomizdat, Moscow, 1977; Wiley, New York, 1980).
13. S. Yu. Luk'yanov, *Hot Plasma and Controlled Nuclear Fusion* (Nauka, Moscow, 1975).

Translated by O. Khadin

Technical Physics, Vol. 45, No. 10, 2000, pp. 1251–1257. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 21–26.

Composition of a Non-self-sustained Discharge Plasma in an $N_2 : O_2 : H_2O$ Mixture at Atmospheric Pressure

M. F. Danilov

Research Institute of Integrated Testing of Electron-Optical Apparatus and Systems, Vavilov State Optical Institute, All-Russia Scientific Center, Sosnovyĭ Bor, Leningrad oblast, 188537 Russia

Received July 20, 1999

Abstract—The plasma composition of a discharge sustained by a pulsed ionization source of μs duration is computed. It is shown that, within a time interval of ~10^{-6} s after the ionization pulse, the dependences of the ion densities on the electric field and ionization source power show features that should be taken into account when developing laser systems for controlling electric discharges in long air gaps. The effect of the plasma composition on the efficiency of electron photodetachment from negative O_2^- ions is investigated by the example of a discharge initiation system consisting of two lasers with different pulse durations and wavelengths. Plasmochemical processes under conditions of enhanced electron photodetachment from negative O_2^- ions are simulated. It is shown that photodetachment can increase the electron density for a time of <10^{-5} s. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

In recent years, electric discharges in air have been actively studied with the aim of plasmochemical cleaning of atmospheric air to remove harmful impurities [1, 2]. Both experimental and theoretical investigations of long-gap discharges, including those controlled by lasers, are now being studied [3, 4]. The first field experiments on the initiation of lightning with powerful CO_2 and Nd lasers have been carried out (see, e.g., [5]). A high-power laser pulse causes an optical breakdown in air which is accompanied by the formation of a beaded plasma structure (long laser spark). Experiments have confirmed the space and time correlation between the creation of the spark plasma channel and the initiation of lightning [5]. One of the theoretical problems that needs to be solved for successful continuation of experiments on controlling electric discharges is an understanding of the physical processes occurring in air plasma. Most of the published data on this issue concern modeling of the plasma composition and an analysis of the processes occurring in either low- and intermediate-pressure discharges [1] or quasi-steady discharges [2]. At the same time, solving the problem of controlling a breakdown in a long air gap requires the investigation of a non-self-sustained discharge at atmospheric pressure.

When modeling a laser spark in the presence of an atmospheric electric field, laser spark beads are considered to be the sources of external photoionization of the surrounding air [6]. A photoionization pulse with a characteristic duration of ~10^{-6} s creates electrical conductivity, which determines the rate of polarization of the laser spark plasma in the electric field. After the external ionization source is switched off, the plasma conductivity drops by one order of magnitude in a time of $\tau_a \sim 10^{-7}$ s and is then controlled by ion–ion recombination. The rate of both ion–ion and electron–ion recombination is known to depend on the ion composition [1, 7], which influences the plasma channel conductivity and the effect of the laser spark on breakdown development.

Since the photodetachment cross sections depend substantially on the radiation wavelength and ion species [8], data on the plasma composition are important for evaluating the efficiency of electric-discharge laser initiation systems (under development) in which a second laser is used to cause photodetachment of electrons from negative ions [9].

The aim of this paper is to numerically simulate plasmochemical processes in a non-self-sustained atmospheric-pressure discharge in an $N_2 : O_2 : H_2O$ mixture that is close to air in composition. Attention is mainly focused on the analysis of the plasma composition dynamics in a discharge sustained by an ionization source of μs duration. The computed data on the plasma composition under various external conditions are important for the development of laser systems for controlling electric discharges.

DESCRIPTION OF THE MODEL AND AN ANALYSIS OF THE PROCESSES

The mathematical model of plasmochemical and charge kinetics in electric discharges in mixtures of molecular gases was described in detail in [10–12]. Based on this model, the CHARM (Chemical Reac-

1063-7842/00/4510-1251$20.00

tions of Molecules) numerical code was developed. The code analyzes the list of possible reactions and prepares a system of differential balance equations for the densities of the plasma components, taking into account the incorporated reactions and the data on their rate constants. The system is solved by Gear's procedure. The code allows modifications to account for changes in the chemical composition of the medium and the external conditions and to include external ionization sources and particle fluxes, in both continuous and pulsed regimes.

The model incorporates about 300 reactions for 50 plasma components. The rate constants for plasmochemical reactions were taken from [1, 2, 7, 8, 13–26]. To calculate the rates of reactions involving electrons, the Maxwellian electron energy distribution function and the corresponding cross sections were used. The cross sections for electron-impact ionization of N_2, O_2, H_2, and NO molecules; dissociative attachment to O_2, H_2O, NO, and N_2O molecules; and excitation of $N_2(A^3\Sigma_u^+)$ and $N_2(a'^1\Sigma_u^-)$ metastable nitrogen levels were taken from [17–20]. The cross sections for direct ionization of H_2O molecules and H, N, and O atoms and step ionization from metastable levels of N_2 molecules were calculated according to the Thomson formula [21].

In the computations, we took into account the dependences of the coefficients of three-body attachment [22] and electron–ion recombination [7, 14, 21] on the electron temperature T_c and the dependence of the detachment coefficients on the electric field E [2]. Altogether, the model incorporates 30 processes whose rates depend on T_c and E. The dependence of the electron temperature on the electric field in air [27] was used to coordinate these parameters. In accordance with [7, 26], it was assumed that recombination rates of $H_3O^+ \cdot H_2O$, $O_2^+ \cdot H_2O$, and $H_3O^+ \cdot OH$ ion clusters do not depend on the electron temperature and are equal to $\sim 10^{-6}$ cm^3 s^{-1}.

The initial mixture composition is $N_2 : O_2 : H_2O = 4 : 1 : 0.05$. The electric field $E(t)$, gas temperature $T_g(t)$, and ionization source power $q(t)$ are the model parameters that determine the external conditions. In the computations, the gas temperature $T_g = 300$ K and the electric field were time-independent; the latter was varied within the range 0–15 kV/cm.

Let us consider the main channels for the generation and loss of charged particles in air plasma, among which ionization, electron attachment, and electron detachment are of primary importance. The ionization rate in a non-self-sustained discharge is determined by an external source. In this study, the ionization source pulse $q(t)$ was assumed to have a rectangular shape, a duration $\tau = 10^{-6}$ s, and an amplitude $q_0 = 10^{17}$–10^{21} cm^{-3} s^{-1}. The q_0 value was chosen so as to fit the calculated electron density to the measured one [28]. The estimations based on photocurrent measurements show that the electron density in the shell of a plasma bead of the laser spark decreases with distance from the bead. The measured photocurrent densities of $\sim 10^{-1}$–10^2 mA/cm^2 are consistent with electron densities of 10^9–10^{12} cm^{-3} [28].

The processes of ionization by an external source are represented in the model by monomolecular reactions of the type $O_2 \longrightarrow O_2^+ + e$, for which the ionization rate $\nu_i(t)$ is regarded as the reaction rate constant. The source power $q(t)$ is related to the ionization rates as follows:

$$q(t) = \sum_j \nu_{i,j}(t) n_j,$$

where summation is performed over the neutral plasma components N_2, O_2, H_2O, N, O, and NO, whose photoionization is incorporated into the model; $\nu_{i,j}$ is the ionization rate (to calculate the photoionization rates, we used data from [10]); and n_j is the density of the jth component.

The main channels for electron losses in air at atmospheric pressure are the reactions of three-body electron attachment to O_2 molecules, with O_2, H_2O, or N_2 molecules as the third body,

$$O_2 + e + O_2 \longrightarrow O_2^- + O_2, \quad k(E)\ [22], \tag{1}$$

$$O_2 + e + H_2O \longrightarrow O_2^- + H_2O, \quad k(E)\ [22], \tag{2}$$

$$O_2 + e + N_2 \longrightarrow O_2^- + N_2, \quad k(E)\ [22], \tag{3}$$

and the reactions of dissociative attachment

$$O_2 + e \longrightarrow + O^- + O, \quad k(E)\ [19], \tag{4}$$

$$H_2O + e \longrightarrow + H^- + OH, \quad k(E)\ [19]. \tag{5}$$

In the above reactions, it is indicated that the rate constants $k(E)$ depend on the electric field discharge and the corresponding source references are given. Note that the rate constants of three-body attachment decrease and those of dissociative attachment increase with the electric field. At $E \approx 4$ kV/cm, the rates of three-body attachment and dissociative attachment become equal to each other. At lower values of E, the processes of O_2^- ion formation dominate over those of O^- and H^- ion formation, whereas, at higher values of E, the situation is the opposite.

The main electron detachment channels are collisions of negative ions with N_2 and O_2 molecules of the initial gas mixture:

$$O_2^- + O_2 \longrightarrow O_2 + O_2 + e, \quad k(E)\ [2], \tag{6}$$

$$O_2^- + N_2 \longrightarrow O_2 + N_2 + e, \quad k(E)\ [2], \tag{7}$$

$$O^- + O_2 \longrightarrow O_3 + e, \quad k = 1.0 \times 10^{-12}\ \text{cm}^3\ \text{s}^{-1}\ [14], \tag{8}$$

$$O^- + N_2 \longrightarrow N_2O + e, \quad k = 1.0 \times 10^{-12}\ \text{cm}^3\ \text{s}^{-1}\ [14], \tag{9}$$

$$H_- + O_2 \longrightarrow HO_2 + e, \quad k = 1.0 \times 10^{-9}\ \text{cm}^3\ \text{s}^{-1}\ [14]. \tag{10}$$

A large group of reactions consist of detachment reactions involving chemically active particles of O, N, H, and $O_2(a^1\Delta)$, which are created in the discharge [13, 14]. As a rule, their rate constants are two to three orders of magnitude higher than those of reactions (6)–(9); however, their role in charge kinetics depends on their production rate, which is determined by the specific energy input into the discharge.

Ion conversion reactions change the plasma composition. Conversion of positive ions proceeds via the following reaction chain:

$$N_2^+ + 2N_2 \longrightarrow N_4^+ + N_2, \quad k = 7.2 \times 10^{-29}\ \text{cm}^6\ \text{s}^{-1}\ [13], \tag{11}$$

$$H_2O^+ + H_2O \longrightarrow H_3O^+ + OH, \quad k = 1.0 \times 10^{-9}\ \text{cm}^3\ \text{s}^{-1}\ [13], \tag{12}$$

$$O_2^+ + 2O_2 \longrightarrow O_4^+ + O_2, \quad k = 2.8 \times 10^{-30}\ \text{cm}^6\ \text{s}^{-1}\ [13], \tag{13}$$

$$O_4^+ + H_2O \longrightarrow O_2^+ \cdot H_2O + O_2, \quad k = 1.2 \times 10^{-9}\ \text{cm}^3\ \text{s}^{-1}\ [13], \tag{14}$$

$$O_2^+ \cdot H_2O + H_2O \longrightarrow H_3O^+ \cdot OH + O_2, \quad k = 1.0 \times 10^{-9}\ \text{cm}^3\ \text{s}^{-1}\ [13]. \tag{15}$$

The conversion can proceed in several steps, whose characteristic times are about 10^{-9}–10^{-7} s [12], where the ion densities are determined by the gas mixture composition and the duration of the pulse exciting the discharge.

Conversion of negative ions via the reactions

$$O^- + O_3 \longrightarrow O_3^- + O, \quad k = 5.3 \times 10^{-10}\ \text{cm}^3\ \text{s}^{-1}\ [13], \tag{16}$$

$$O_2^- + O_3 \longrightarrow O_3^- + O_2, \quad k = 3.5 \times 10^{-10}\ \text{cm}^3\ \text{s}^{-1}\ [13], \tag{17}$$

$$O_2^- + NO_2 \longrightarrow NO_2^- + O_2, \quad k = 1.9 \times 10^{-9}\ \text{cm}^3\ \text{s}^{-1}\ [13], \tag{18}$$

$$NO_2^- + H_2O + M \longrightarrow NO_2^- \cdot H_2O + M, \quad k = 1.0 \times 10^{-28}\ \text{cm}^6\ \text{s}^{-1}\ [13] \tag{19}$$

proceeds more slowly than that of positive ions, because it involves the products of plasmochemical reactions O_3 and NO_2, which are absent in the initial gas mixture.

According to the energy approach to the consideration of ion reactions [26], in the course of exothermic reactions, positive ions convert into positive ions with a lower recombination energy and negative ions convert into other negative ions with a higher attachment energy (electron affinity). Ion conversion leads to substitution of the primary ions with large ion complexes (clusters), which increases the rate of electron–ion recombination [7]. The model under consideration does not incorporate the processes of the formation of $H_3O^+ \cdot nH_2O$ and $NO^+ \cdot nH_2O$ ion clusters for $n > 1$ or $O_2^- \cdot nH_2O$ and $OH^- \cdot nH_2O$ ion clusters [26]. These processes make the picture of ion conversion more complicated but do not change its main tendency.

An analysis of plasmochemical processes reveals a significant role of water vapor, even though its fraction is relatively low. For example, at $E = 5$ kV/cm and a water molecule concentration of 3 vol %, the contribution of water vapor to the total electron-impact ionization rate amounts to ≈10%. Moreover, the rate of three-body and dissociative attachment reactions involving H_2O molecules is about 40% of the total attachment rate. Hence, an increase in the water content of air increases electron losses. Electron detachment cannot compensate for electron losses due to attachment. This is confirmed by the experimental results of [2].

Note that a key reaction determining the important role of water in the discharge is the ion conversion reaction

$$H^- + H_2O \longrightarrow OH^- + H_2, \quad k = 3.8 \times 10^{-9}\ \text{cm}^3\ \text{s}^{-1}\ [2]. \tag{20}$$

If this reaction is not incorporated into the model, detachment reaction (10) almost completely compensates for the process of dissociative attachment (5). As a result, the balance between the generation and loss of electrons shifts toward generation. In this case, an increase in the water content of the mixture would lead to a higher electron density, which contradicts the experimental results [2].

RESULTS OF COMPUTATIONS

Figure 1 presents the calculated time evolution of the positive ion density obtained at an ionization source power of $q_0 = 7 \times 10^{18}\ \text{cm}^{-3}\ \text{s}^{-1}$ and electric field of $E = 5$ kV/cm. It is seen that the characteristic time of ion conversion is 10^{-9}–10^{-8} s. Within this time interval, the

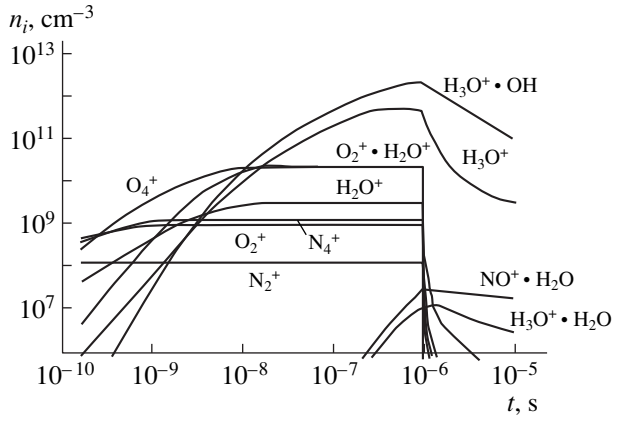


Fig. 1. Time evolution of the positive ion densities.

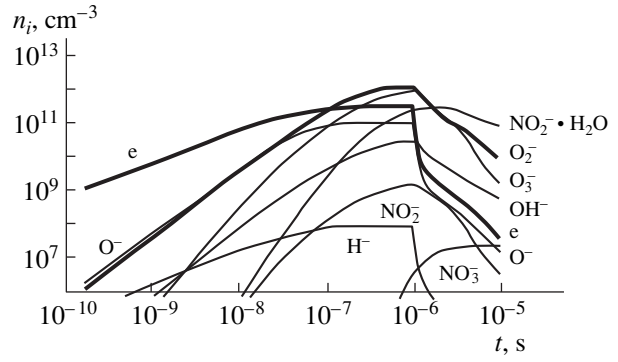


Fig. 2. Time evolution of the negative ion densities.

densities of most ions reach their steady value, except for H_3O^+, $H_3O^+ \cdot OH$, $H_3O^+ \cdot H_2O$, and $NO^+ \cdot H_2O$. The densities of these ions keep rising during the entire ionization pulse. Quasi-steady densities of N_2^+, O_2^+, N_4^+, H_2O^+, $O_2^+ \cdot H_2O$, and O_4^+ ions are determined by the ionization source power and the rates of ion conversion reactions (11)–(15), whereas those of H_3O^+ and $H_3O^+ \cdot OH$ ions are determined by the source power and the effective recombination coefficient $n_{0,i} = \sqrt{q/k_{r,i}}$. After the ionization pulse, the plasma composition changes rapidly due to processes (11)–(15). In a time $\tau \sim 10^{-8}$ s, the densities of O_4^+, $O_2^+ \cdot H_2O$, H_2O^+, N_4^+, O_2^+, and N_2^+ ions fall by several orders of magnitude. Further change in the ion composition is determined mainly by the effective recombination coefficients for the ion species. The densities of $H_3O^+ \cdot H_2O$ and $NO^+ \cdot H_2O$ ions decrease more slowly via the reactions

$$H_3O^+ + H_2O + H_2 \longrightarrow H_3O^+ \cdot H_2O + H_2,$$
$$k = 5.0 \times 10^{-27} \text{ cm}^6 \text{ s}^{-1} \text{ [26]},$$

$$NO^+ + H_2O + M \longrightarrow NO^+ \cdot H_2O + M,$$
$$k = 2.0 \times 10^{-28} \text{ cm}^6 \text{ s}^{-1} \text{ [13]}.$$

The calculated densities of negative ions O_2^-, O_3^-, O^-, $NO_2^- \cdot H_2O$, OH^-, H^-, NO_2^-, and NO_3^- are shown in Fig. 2. Several characteristic time intervals are seen in the dynamics of the negative ion composition. The time in which the steady-state densities of electrons and O^- and H^- ions are established is $\tau_1 \approx 10^{-7}$ s. The moment $\tau_2 \approx 10^{-6}$ s, at which the ionization source is switched off, coincides with the time in which the steady-state densities of O_2^- and OH^- ions are established. After the ionization pulse, in a time $\tau_3 \approx 5 \times 10^{-7}$ s, the electron density drops by two orders of magnitude due to attachment to oxygen molecules. As a result, electron detachment is balanced out by attachment. Further change in the electron density is determined by the balance between the processes of impact ionization and recombination. The increase in the $NO^- \cdot H_2O$ ion density due to reactions (18) and (19) continues for 2 μs after the ionization pulse.

The electric field is one of the basic parameters that determine the conditions of experiments on laser control of a breakdown in a long air gap. In nature, a thunder cloud can generate an electric field as high as ~10^2 V/cm near the earth's surface [5]. Since investigations of lightning control require high electric fields, special measures to increase the field are undertaken. For example, laser beams are focused on a tower tip [5], where the electric field can attain ~10^3–10^4 V/cm. In this connection, the investigation of the dependences of both electron and ion densities on the electric field (Fig. 3) is of interest, because these densities determine the plasma channel conductivity and the time constant of laser spark polarization [6]. The dependences computed for the instant $t = 5$ μs and $q_0 = 7 \times 10^{18}$ cm^{-3} s^{-1} show a slight variation of the electron density and the densities of O_3^-, $NO_2^- \cdot H_2O$, OH^-, and NO_2^- ions in the range $E = 3$–10 kV/cm. Here, the main process of electron generation is electron detachment with the participation of an oxygen atom:

$$O_2^- + O \longrightarrow O_3 + e,$$
$$k = 3.3 \times 10^{-10} \text{ cm}^3 \text{ s}^{-1} \text{ [14]}.$$

At $E = 5$ kV/cm, its contribution to the total rate of electron generation is higher than 60%. The electron losses are determined mainly by dissociative attachment reaction (4), whose contribution is higher than 50%. The contributions of electron-impact ionization and recombination to the general charge-kinetics balance are less than 2 and 3%, respectively. The drop in the density of O_2^- ions in the range $E = 1$–6 kV/cm is caused by the decrease in the rate constants of reactions

(1)–(3). The increase in the electron density with electric field at $E > 10$ kV/cm stems from the increasing role of detachment processes (6) and (7).

The ionization source power q_0 is related to the plasma density in the beads of a long laser spark [6, 28]. Experiments on laser initiation of lightning [5] require the formation of plasma channels with a high plasma density in the beads. Thus, investigation of the parameters of a non-self-sustained discharge as functions of q_0 is necessary to formulate the required characteristics of a laser spark and laser system for discharge control. Calculations of the plasma composition for $t = 5$ μs and $E = 5$ kV/cm show that, as the source power increases from 10^{19} to 10^{21} cm^{-3} s^{-1}, the growth of the ion densities saturates and the ion composition changes radically (Fig. 4). The increase in the ionization power stimulates conversion of negative ions in the discharge, which results in substitution of O_2^- ions by $NO_2^- \cdot H_2O$ and O_3^- ions. The saturation is related to the nonlinear dependence of the ion recombination rate on the ion density. Using the well-known solution to the recombination equation $dn_i(t)/dt = -k_r n_i^2(t)$, we find that the time required for the ion density to decrease from $n_i^{(0)}$ to $n_i^{(1)}$ is $t_1 = (n_i^{(0)} - n_i^{(1)})/n_i^{(0)}/n_i^{(1)}/k_r$. If $n_i^{(0)} \gg n_i^{(1)}$, then $t_1 \approx 1/n_i^{(1)}/k_r$; i.e., this time does not depend on $n_i^{(0)}$. At $n_i^{(1)} = 10^{12}$ cm^{-3} and $k_r = 10^{-6}$ cm^3 s^{-1}, we have $t_1 = 10^{-6}$ s; i.e., in a time of about 1 μs, the ion density drops to ~10^{12} cm^{-3}, whatever the initial density. In a time of 100 μs, the ion density drops to ~10^{10} cm^{-3}. Hence, the increase in the plasma density in the beads of a laser spark can increase the charged particle density to more than 10^{12} cm^{-3}, but only during the ionization pulse and for 1 μs after the pulse.

The data on the plasma composition of a non-self-sustained discharge allow us to consider in detail the other methods of forming conducting channels to control an electric discharge in air. From this point of view, the combination of UV and visible lasers is of most interest for initiating a high-voltage discharge at atmospheric pressure [9]. A UV laser pulse with wavelength of 248 nm and duration of less than 1 μs produces three-photon ionization of oxygen molecules and four-photon ionization of nitrogen molecules, thus forming a needle-shaped zone of ionized air in the region of a high electric field. A laser pulse with a wavelength of 750 nm and duration of more than 5 μs produces photodetachment of electrons from negative ions and thus maintains the free-electron current. According to the calculations of [9], the time required for plasma polarization, which is responsible for the formation of a fast ionization wave due to field amplification at the plasma channel end, amounts to several microseconds.

Let us consider the electron photodetachment processes, taking into account the plasma composition.

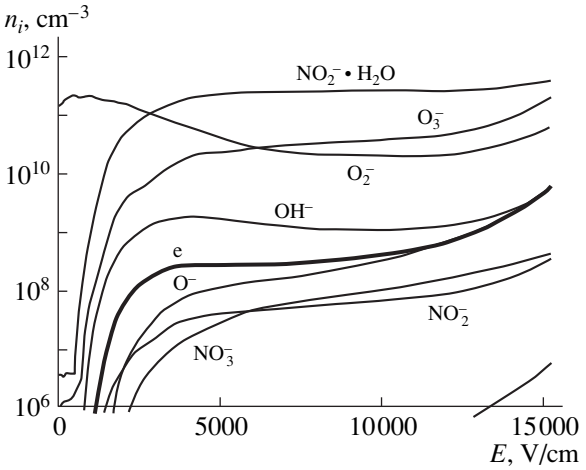


Fig. 3. Negative ion densities vs. electric field.

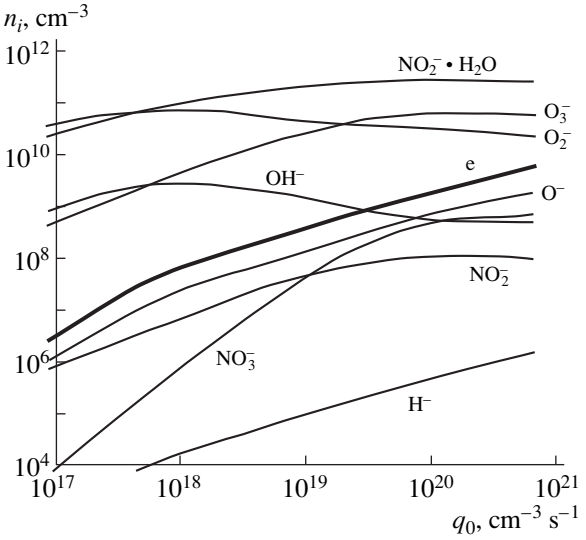


Fig. 4. Negative ion densities vs. ionization source power.

Radiation with a wavelength of 750 nm (photon energy of 1.64 eV) can produce photodetachment of electrons from O_2^-, H^-, and O^- ions, whose electron affinity is 0.44, 0.75, and 1.46 eV, respectively. Efficient electron photodetachment from OH^-, O_3^-, NO_2^-, and $NO_2^- \cdot H_2O$ ions, whose electron affinity ranges from 1.83 to 3 eV [26], is expected to occur with a radiation wavelength no longer than 700 nm [8]. Simulations of the plasma composition (Figs. 2–4) show that the effect of radiation with a wavelength of 750 nm is most pronounced in the detachment of electrons from O_2^- ions. To verify this result, we modeled plasmochemical processes in a discharge, taking into account the photodetachment reaction $O_2^- \longrightarrow O_2 + e$. For comparison, Fig. 5 presents the results obtained with and without incorporat-

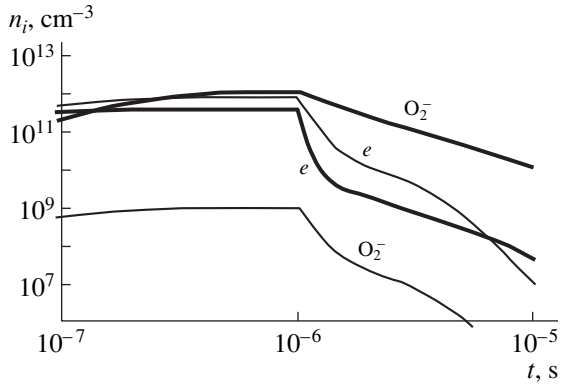


Fig. 5. Time evolution of electron density and O_2^- ion density with (light curves) and without (heavy curves) taking into account photodetachment for $E = 5$ kV/cm and $q_0 = 7 \times 10^{18}$ cm^{-3} s^{-1}.

ing photodetachment. For a photodetachment rate of 10^{10} s^{-1} and electric field of $E = 5$ kV/cm, the density of O_2^- ions decreases by three to four orders of magnitude. In this case, the electron density first increases several times, then rapidly drops, and, after $t \approx 6$ μs, becomes even less than when photodetachment is neglected. This can be explained by the fact that, at the high negative ion density in the discharge, the electron density is maintained via detachment. The disappearance of negative ions shifts the balance between detachment and attachment; this leads, first, to speeding up of electron generation and then to an increase in their loss rate. Thus, the calculations show that the use of the other laser for electron photodetachment can increase the electron density only for some limited time. A final conclusion about the efficiency of the method in question requires numerical simulations that take into account the real values of the laser intensity and cross sections for electron photodetachment from O_2^-, H^-, and O^- ions at a wavelength of 750 nm.

CONCLUSION

An analysis of the main channels for the generation and loss of charged particles in a non-self-sustained discharge plasma has shown that, regardless of the relatively low concentration of water molecules in the mixture, they play a significant role in the processes of attachment and ion conversion. The densities of primary ions are determined mainly by the balance between photoionization and ion conversion. The densities of complex ions originating via the chain of ion–molecular reactions are bound above by recombination. Calculation revealed a relatively high contribution from reactions involving oxygen atoms and molecular O_2^- ions to electron detachment.

The dependences of the ion densities on the electric field and ionization source power for ~10^{-6} s after the ionization pulse show features that should be taken into account when developing laser systems for controlling electric discharges in long air gaps. Slight variations in the electron and ion densities in the range $E = 3$–10 kV/cm lead to slight variations in both the plasma conductivity and the time constant of polarization of a long laser spark in the electric field. An increase in the source power from 10^{19} to 10^{21} cm^{-3} s^{-1} results in saturation of the growth of the ion density at a level of 10^{12} cm^{-3}. The saturation is related to the nonlinear dependence of the ion recombination rate on the ion density.

Modeling plasmochemical processes under conditions of enhanced electron photodetachment from negative O_2^- ions has shown that photodetachment can increase the electron density for a time of $<10^{-5}$ s.

ACKNOWLEDGMENTS

I am grateful to V. M. Gromovenko for fruitful discussions.

REFERENCES

1. N. L. Aleksandrov, S. V. Dobkin, A. N. Konchakov, and D. A. Novitskiĭ, Fiz. Plazmy **20**, 492 (1994) [Plasma Phys. Rep. **20**, 442 (1994)].
2. Yu. S. Akishev, A. A. Deryugin, and V. B. Karal'nik, Fiz. Plazmy **20**, 571 (1994) [Plasma Phys. Rep. **20**, 511 (1994)].
3. G. N. Aleksandrov, V. L. Ivanov, G. D. Kadzov, *et al.*, Zh. Tekh. Fiz. **47**, 2122 (1977) [Sov. Phys. Tech. Phys. **22**, 1233 (1977)].
4. M. Miki, A. Wada, and T. Shindo, Opt. Zh. **66** (3), 25 (1999) [J. Opt. Technol. **66**, 190 (1999)].
5. S. Uchida, Y. Shimada, H. Yasuda, *et al.*, Opt. Zh. **66** (3), 36 (1999) [J. Opt. Technol. **66**, 199 (1999)].
6. M. F. Borisov, V. M. Gromovenko, M. F. Danilov, and V. A. Lapshin, Opt. Zh. **66** (3), 48 (1999) [J. Opt. Technol. **66**, 210 (1999)].
7. M. A. Biondi, in *Applied Atomic Collision Physics*, Vol. 3: *Gas Lasers*, Ed. by E. W. McDaniel and W. Nigan (Academic, New York, 1982; Mir, Moscow, 1986).
8. H. S. W. Massey, *Negative Ions* (Cambridge Univ. Press, Cambridge, 1976; Mir, Moscow, 1979).
9. P. Rambo, J. Biegert, V. Kubecek, *et al.*, Opt. Zh. **66** (3), 30 (1999) [J. Opt. Technol. **66**, 194 (1999)].
10. M. F. Danilov and E. V. Kalashnikov, Teplofiz. Vys. Temp. **33**, 663 (1995).
11. M. F. Danilov and A. N. Malinin, Zh. Tekh. Fiz. **66** (2), 74 (1996) [Tech. Phys. **41**, 153 (1996)].
12. M. F. Danilov and A. N. Malinin, Zh. Tekh. Fiz. **68** (5), 27 (1998) [Tech. Phys. **43**, 502 (1998)].
13. K. Smith and R. M. Thompson, *Computer Modeling of Gas Lasers* (Plenum, New York, 1978; Mir, Moscow, 1981).

14. H. Hokazono and H. Fujimoto, J. Appl. Phys. **62**, 1585 (1987).
15. V. N. Kondrat'ev, *Gas-Phase Reaction Rate Constants* (Nauka, Moscow, 1970).
16. V. P. Agafonov, V. K. Vertushkin, A. A. Gladkov, and O. Yu. Polyanskiĭ, *Nonequilibrium Physicochemical Processes in Aerodynamics* (Mashinostroenie, Moscow, 1972).
17. J. J. Lovke, A. V. Phelps, and B. W. Irwin, J. Appl. Phys. **44**, 4664 (1973).
18. D. Rapp and P. Englander-Golden, J. Chem. Phys. **43**, 1464 (1965).
19. D. Rapp and D. D. Briglia, J. Chem. Phys. **43**, 1480 (1965).
20. A. G. Engelhardt, A. V. Phelps, and C. G. Risk, Phys. Rev. A **135**, 1566 (1964).
21. Yu. P. Raizer, *Gas Discharge Physics* (Nauka, Moscow, 1987; Springer-Verlag, Berlin, 1991).
22. N. L. Aleksandrov, Usp. Fiz. Nauk **154**, 177 (1988) [Sov. Phys. Usp. **31**, 101 (1988)].
23. A. V. Berdyshev, I. V. Kochetov, and A. P. Napartovich, Fiz. Plazmy **14**, 741 (1988) [Sov. J. Plasma Phys. **14**, 438 (1988)].
24. A. V. Berdyshev, A. S. Golovin, A. V. Gurashvili, *et al.*, Fiz. Plazmy **15**, 335 (1989) [Sov. J. Plasma Phys. **15**, 195 (1989)].
25. A. A. Deryugin, I. V. Kochetov, A. F. Pal', *et al.*, Fiz. Plazmy **17**, 1138 (1991) [Sov. J. Plasma Phys. **17**, 661 (1991)].
26. W. J. Wiegand, in *Applied Atomic Collision Physics*, Vol. 3: *Gas Lasers*, Ed. by E. McDaniel and W. Nighan (Academic, New York, 1982; Mir, Moscow, 1986).
27. S. C. Brown, *Basic Data of Plasma Physics* (MIT Press, Cambridge, Mass., 1959; Atomizdat, Moscow, 1961).
28. M. F. Borisov, V. M. Gromovenko, V. A. Lapshin, *et al.*, Opt. Zh. **66** (3), 40 (1999) [J. Opt. Technol. **66**, 203 (1999)].

Translated by N. Ustinovskiĭ

Technical Physics, Vol. 45, No. 10, 2000, pp. 1258–1264. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 27–33.

Development of a Self-sustained Gas Volume Discharge

V. V. Osipov and V. V. Lisenkov

Institute of Electrophysics, Ural Division, Russian Academy of Sciences, Komsomol'skaya ul. 34, Yekaterinburg, 620049 Russia

E-mail: osipov@ief.intec.ru

Received August 26, 1999

Abstract—The development of a self-sustained volume gas discharge is studied numerically. Special attention is paid to the formation and stability of the cathode sheath. It is shown that the high-voltage threshold for the ignition of a volume discharge is determined by the onset of cathode sheath instability. The methods for suppressing cathode sheath instability are discussed. © *2000 MAIK "Nauka/Interperiodica"*.

INTRODUCTION

The creation of pulsed high-pressure gas lasers has stimulated investigations of volume discharges (VDs). Within a short time, methods for VD initiation were developed [1–3], the energy and time characteristics of a VD were studied [3–5], mechanisms for the contraction of a VD were proposed [6–9], and VD-based lasers with an active volume of hundreds of liters were created [10, 11]. In spite of the large amount of data obtained, the physical picture of the development of a VD from the instant the voltage is applied to the formation of the plasma column and cathode potential drop is still poorly understood.

There are many theoretical and experimental studies devoted to this problem. In [12], it was found that, after applying a dc voltage to an air gap of $d = 0.3$ cm at atmospheric pressure, an intermediate stage with a higher discharge voltage and lower discharge current existed for about 1 μs before the formation of the discharge channel [12]. Electron-optical measurements revealed the existence of the cathode potential drop, negative glow, and Faraday dark space at this stage and identified this stage as an ordinary glow discharge [13–16].

The use of photomultipliers allowed a substantial increase in the sensitivity of the diagnostic equipment and revealed the following stages of discharge development: a series of avalanches leading to the creation of a positive space charge near the anode, giving rise to the first ionization wave; passage of several ionization waves, which increased the gap conductivity up to the value determined by the discharge circuit; and cathode sheath formation by the last but one ionization wave [17]. It is assumed that, after the last stage, a glow discharge is formed. Thus, in [12–17], the general physical picture of the development of both steady and pulsed glow discharges ignited within the Townsend overvoltage range was revealed. In contrast to ordinary low-pressure glow discharges, the pulsed high-pressure discharge was called a transient glow discharge [18]. Note that this discharge is more frequently called a pulsed glow discharge.

Unexpected results were obtained in [19, 20], where a discharge with a steplike voltage drop at a streamer overvoltage of ~100% was observed in a 0.2-cm air gap in which ~10^4 electrons were created 60 ns before the initiation of the main discharge and where the discharge would apparently take the form of a channel. Electron-optical measurements carried out at a high overvoltage and high initial electron density confirmed the existence of a discharge stage with a fairly homogeneous glow occupying nearly the entire working volume between the electrodes [21]. Further, this stage was observed in the nonspark mode and was referred to as a "volume" discharge. Here, the term "volume" stresses the fact that the discharge parameters are determined only by the discharge itself, and not by its interaction with the chamber wall (as in an ordinary glow discharge).

In [22], it was shown that, in fact, a VD is developed from a single ionization wave and that its realization requires preionization of the entire gas volume [22]. In [23], a model of the fast development of a VD was proposed. The model assumes that the discharge is formed due to the overlapping of electron avalanches whose sizes have reached a certain critical value. Later, based on this model, criteria were obtained for the development of a VD according to which an initial electron density of $n_0 > 10^6$ cm^{-3} must be produced in a gas before applying the high-voltage pulse [24–26].

The model of the formation of a plasma column proposed in [27, 28] agrees better with the experimental results. The model takes into account that electron drift is considerably faster than electron diffusion across the electric field; therefore, the initial electrons, which ionize the gas, give rise to a train of avalanches following one another and forming thin current filaments. As soon as the filament radius becomes equal to the Debye length, electron diffusion is replaced with ambipolar diffusion and the filament almost stops widening within

1063-7842/00/4510-1258$20.00

the time interval of interest (~10^{-7} s). The minimum electron density that enables the overlapping of the current filaments can be found by setting the filament radius equal to the Debye length. The model allows one to explain both the dependence of the maximum energy input in the discharge on the initial electron density and the presence of the current filaments in a discharge [29]. However, this model fails to explain the higher ignition voltage of a VD compared to both ordinary and transient glow discharges [24, 25]. It also does not provide a complete explanation of the possible contraction of a VD in the early stage of its development. For this reason, in this paper, we study the development of a VD numerically, under the assumption that the discharge is homogeneous (the current filaments overlap) and discuss the possible reasons for the above features of a VD.

FORMATION OF THE CATHODE SHEATH

To analyze the dynamics of discharge development, we solve the following set of equations consisting of Poisson's equation and the balance equations for the densities of electrons n_e, ions n_i, and excited molecules n_*:

$$\frac{\partial n_e}{\partial t} + \mu_e E\frac{\partial n_e}{\partial x} = (\nu_i(E) - \nu_a(E))n_e, \quad (1)$$

$$\frac{\partial n_i}{\partial t} - \mu_i E\frac{\partial n_i}{\partial x} = (\nu_i(E) - \nu_a(E))n_e, \quad (2)$$

$$\frac{\partial n_*}{\partial t} = \nu_e(E)n_e - \frac{N_*}{\tau_r}, \quad (3)$$

$$\frac{\partial^2 \varphi}{\partial x^2} = \frac{e}{\varepsilon_0}(n_e - n_i), \quad (4)$$

with the boundary conditions at the cathode

$$\varphi = 0, \quad j_e = \gamma_i \mu_i E n_i + \gamma_{ph}\int_0^d \frac{n_*}{\tau_r}dx \quad (5)$$

and at the anode,

$$\varphi = U(t), \quad (6)$$

where φ is the potential; E is the electric field; μ_e and μ_i are the electron and ion mobilities, respectively; ν_i, ν_a, and ν_e are the rates of ionization, dissociative attachment, and excitation, respectively; γ_i is the secondary ion–electron emission coefficient; and γ_{ph} is the photoemission coefficient.

A similar approach was used earlier, e.g., in [30] to analyze the processes in the cathode region of a VD in a XeCl laser, where the plasma column was considered within the 0D approximation. The model allows a self-consistent description of the processes occurring in both the plasma column and the cathode sheath. It

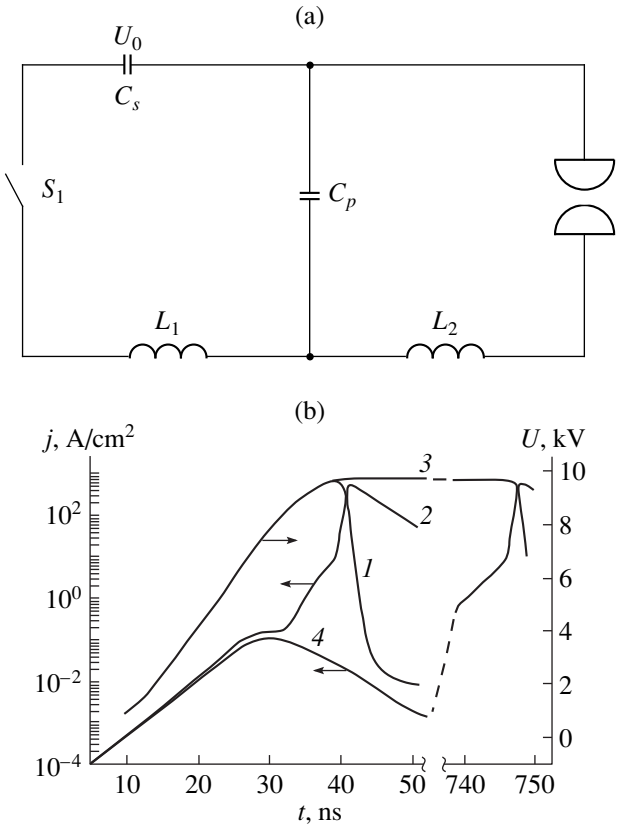


Fig. 1. (a) Schematic of the discharge circuit corresponding to experiment [31] (C_s is the storage capacitor, C_p is the peaking capacitor, and L_1 and L_2 are the inductances) and (b) the calculated time evolutions of (*1*, *3*) the discharge voltage and (*2*, *4*) current density with (curves *1*, *2*) and without (curves *3*, *4*) taking photoemission into account.

incorporates the equations for the discharge circuit (Fig. 1a) with parameters corresponding to the experimental conditions [31]: the discharge volume of $0.5 \times 0.5 \times 8$ cm^3 was filled with a $CO_2 : N_2 : He = 1 : 1 : 3$ mixture at atmospheric pressure. The initial electron density n_0 was 10^8 cm^{-3}.

The rate constants of the processes involving electrons were determined from the electron energy distribution function obtained by numerically solving Boltzmann's equation. The reaction cross sections were taken from [32, 33]. The parameter ν_i in Eqs. (1) and (2) is the total ionization rate of all the components. When determining this rate, the whole positive charge was assigned to CO_2^+ ions formed in the reactions

$$N_2^+ + CO_2 \longrightarrow CO_2^+ + N_2, \quad (7)$$

$$He^+ + CO_2 \longrightarrow CO^+ + O + He, \quad (8)$$

$$CO^+ + CO_2 \longrightarrow CO_2^+ + CO. \quad (9)$$

Since the rate constants of these reactions are fairly high (~10^{-9} cm^3/s [33]) and, therefore, the correspond-

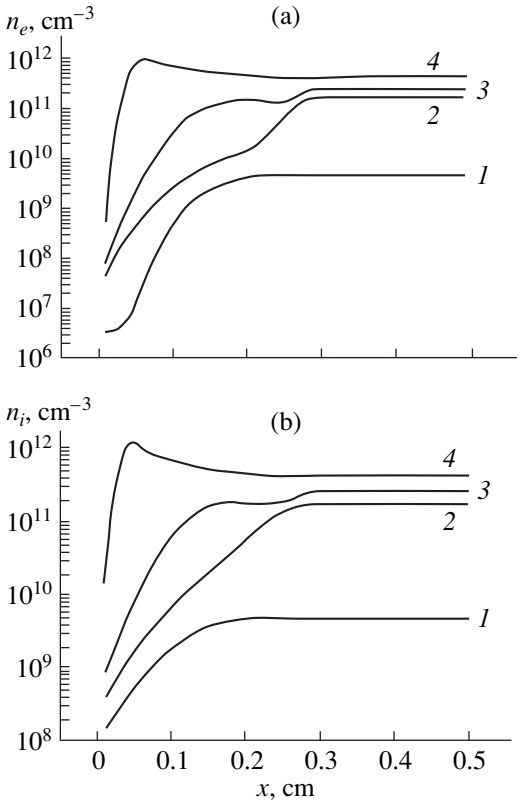


Fig. 2. The profiles of (a) the electron density and (b) the positive ion density in the discharge gap for t = (*1*) 22, (*2*) 34, (*3*) 36, and (*4*) 38 ns; x is the distance from the cathode.

ing characteristic times (~10^{-10} s) are one order of magnitude less than the ionization time, we exclude the N_2^+ and He^+ ions from consideration. Dissociative attachment was also not included in an explicit form. The point is that the created oxygen ions are quickly neutralized in the reactions

$$O^- + CO_2 + He \longrightarrow CO_3^- + He, \tag{10}$$

$$CO_3^- + CO_2^+ \longrightarrow 2CO_2 + O. \tag{11}$$

The rate constant of reaction (11) is 5×10^{-7} cm^3/s, which is one order of magnitude larger than the rate constant of electron–ion recombination

$$\begin{aligned} &CO_2^+ + e \longrightarrow CO + O, \\ &k = 4.0 \times 10^{-8}\ \text{cm}^3/\text{s}. \end{aligned} \tag{12}$$

Consequently the term in Eq. (2) accounting for the generation of positive ions can be rewritten in the same form as in Eq. (1). Note that the product $\nu_a n_e$ accounts for the loss of positive ions via reactions (10) and (11), which enables us to exclude the equations for negative ions.

Figure 1b presents the calculated time evolutions of the discharge voltage and current density. The results of the calculations agree well with the experimental data with respect to the amplitudes and durations of the pulses [31]. When studying the processes occurring at the stage of the voltage rise (the formation of the cathode sheath and the onset of instability), we eliminated the discharge circuit equation to reduce the computation time and approximated the discharge voltage by the function

$$U(t) = \begin{cases} \dfrac{U_0}{\tau_f} t, & t < \tau_f \\ U_0, & t \geq \tau_f, \end{cases} \tag{13}$$

where τ_f is the voltage rise time and U_0 is the voltage pulse amplitude.

The profiles of the electric field and electron and ion densities during the development of the discharge are shown in Figs. 2 and 3. It is seen that, at the initial stage, when the electron density in the gap is relatively low, the plasma column is 0.25 cm away from the cathode. The rapid increase in the electron density is accompanied by the formation of a charge-depleted region between the cathode and the plasma column, in which the ion density is higher than the electron density. The ions partially shield the external field, thus reducing the field in the plasma column and increasing it in the cathode region. As a result, the electron density and, consequently, the current density in the plasma column, grow at a lower rate (Fig. 1b). The increase in the ionization rate in the cathode region leads to a fast (in a time of 34–38 ns) propagation of the ionization wave (plasma column) toward the cathode, thus resulting in the formation of the cathode potential drop. At the peak of the discharge current, the electric field in the cathode sheath reaches a maximum value of $\approx 2 \times 10^6$ V/cm, whereas the depth of the cathode sheath drops to a minimum value of 2.4×10^{-3} cm (Fig. 3).

The fact that the ion density in the cathode region remains low (~10^8 cm^{-3}) for a relatively long time interval (Fig. 2b) determines the type of emission processes governing the conduction current in the discharge gap. The results of calculating the ratio between the photoemission and ion–electron emission currents are shown in Fig. 4. It is seen that, during the first 36 ns, the photoemission current is considerably higher than the ion–electron emission current. Note that, later on, in contrast to an ordinary glow discharge, photoemission also plays a significant role in extracting electrons from the cathode.

Neglecting the photoemission current (Fig. 1b, curves *3*, *4*) leads to a sharp decrease in the velocity of the ionization wave propagating toward the cathode. This is accompanied by a decrease in the current density to $<10^{-3}$ A/cm^2 and an increase in the discharge delay time. A delay of ≈740 ns between application of the voltage and the stage of rapid current rise is in distinct contradiction to the available experimental data

[31]. Note that, in this case, for better stability of the computation scheme, we eliminated the equations for the discharge circuit and assumed that the discharge voltage was constant during the propagation of the ionization wave toward the cathode.

The existence of a stage in which the conduction current is determined by the photoemission current distinguishes a VD from both ordinary and transient glow discharges, in which, due to the passage of several ionization waves until the cathode potential drop forms, the number of ions in the cathode region is fairly high, so as to ensure that electron emission due to bombardment of the cathode with ions is dominant [17]. During this stage, a VD becomes very sensitive to the presence of local sites with enhanced electron emission on the cathode. This viewpoint was confirmed in experiment [34], in which 1.2×10^{-3}-cm-thick nickel strips deposited onto a copper cathode caused discharge stratification as early as the stage of discharge development. Hence, filling the charge-depleted region with electrons should significantly increase VD stability. Indeed, this was the case in the experiments with electron injection into the cathode region [35] and experiments in which a plasma cathode [36] or high-emissive cathode [37] was used.

HIGH-VOLTAGE THRESHOLD FOR VD IGNITION

The existence of a high-voltage threshold for ignition of a VD between two metal electrodes, which was first revealed in experiments with preionized high-pressure gases, was erroneously attributed to a static breakdown [38]. Later, the existence of such a threshold was explained by the fact that a VD discharge necessarily develops in the form of a quasi-streamer [25]. In our opinion, the existence of a threshold voltage for VD ignition stems from the existence of cathode instability, which leads to discharge stratification at the discharge development stage. Let us consider this issue in more detail.

Naturally, the stage in which two clearly defined regions exist in the discharge, (namely, the plasma column and a charge-depleted region near the cathode) is of primary interest. In this stage, the current density is such that the discharge exists in subnormal mode; consequently, it cannot occupy the entire electrode surface and is stratified. The possibility of discharge stratification at the discharge development stage was pointed out in [39]. This process can lead to a sharp change in the current–voltage characteristic, provided that the conditions necessary for autoelectronic emission to exist are satisfied [40].

Figure 5 presents the calculated current–voltage characteristics for two cases: (i) the ionization wave (plasma column) propagating toward the cathode remains homogeneous when passing through the charge-depleted region, and the enhancement of the field at microinhomogeneities is taken into account by

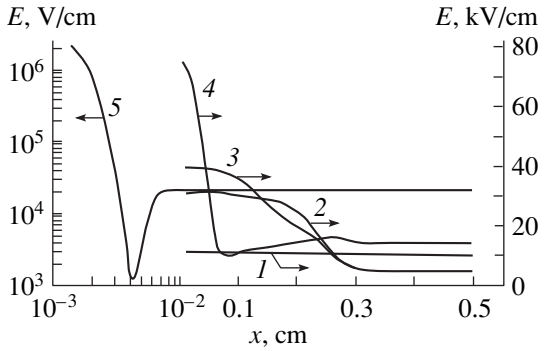


Fig. 3. The profiles of the electric field in the discharge gap for t = (*1*) 22, (*2*) 34, (*3*) 36, (*4*) 38, and (*5*) 43 ns.

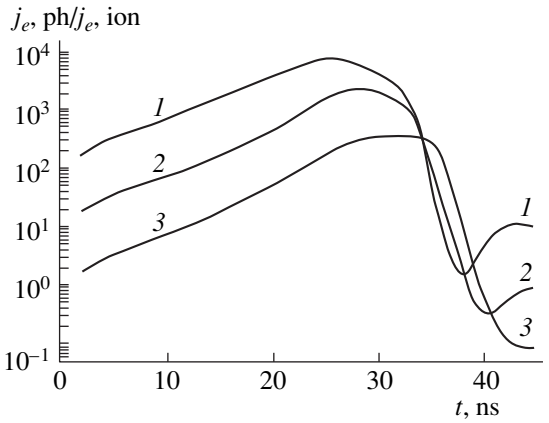


Fig. 4. Time evolution of the ratio between the photocurrent density $j_{e,\,ph}$ and the density of the secondary ion-electron emission current $j_{e,\,ion}$ for γ_{ph} = (*1*) 2×10^{-4}, (*2*) 2×10^{-5}, and (*3*) 2×10^{-6}.

the integral amplification coefficient [40] (curve *1*); (ii) the ionization wave undergoes filamentation in the charge-depleted region due to discharge stratification (curves *3–5*). In the latter case, it was supposed that one of the filaments of diameter λ (its definition will be given below) contracts to an area of $\sim 10^{-6}$ cm^2, provided that there is a $\sim 10^{-3}$-cm microinhomogeneity on the cathode that disturbs the external electric field. For example, this may be a micropoint of corresponding height (the fact that the field at the tip of a micropoint is enhanced by a factor of β is taken into account). The parameters of the cone-shaped cathode sheath were calculated in spherical coordinates; the angular distribution of the parameters was assumed to be uniform.

The density of the autoelectronic emission current in plane geometry is determined by the formula [40]

$$j_a = \sqrt{\pi} a b^2 \left(\frac{\beta}{\vartheta}\right)^{3/2} \exp\left(-\frac{2\vartheta}{\beta}\right) \times \left\{1 + 2\sqrt{\frac{\beta}{\pi\vartheta}} + \frac{\beta}{2\pi\vartheta}\right\}, \quad (14)$$

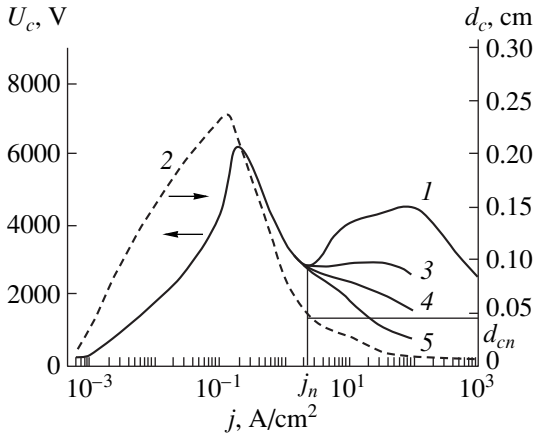


Fig. 5. Cathode potential drop in (*1*) plane geometry and at a micropoint with β = (*3*) 10, (*4*) 30, and (*5*) 100, and the cathode sheath depth (*2*) vs. the discharge current density.

where a and b are the numerical coefficients in the Fowler–Nordheim formula for the density of the autoelectronic emission current (they are equal to $1.55 \times 10^{-6}\varphi^{-1}$ and $6.85 \times 10^7\varphi^{3/2}$, respectively), φ is the work function in eV, β is the field gain averaged over the cathode surface, $\vartheta = \sqrt{b/(\beta E_k)}$, and E_k is the average field at the cathode surface in V/cm. Autoelectronic emission from the tip of a micropoint was taken into account using the Fowler–Nordheim formula.

It is seen that the filamentary structure of the plasma column and autoelectronic emission lead to the generation of local sites near the micropoints where the amplification factor is higher than 10 (Fig. 5, curves *4*, *5*). In these sites, the current–voltage characteristic is descending; hence, a cathode spot and then a channel are formed as a result of cathode instability.

The formation of a plasma column with a filamentary structure in the cathode region can occur during propagation of the ionization wave toward the cathode, because the subnormal mode can occur at this stage (Fig. 5, curve *1*). Generally, the dynamics of this process can be traced only by solving the corresponding two- or three-dimensional problem. However, using the criterion

$$\Omega = \int_{(t_{in})} \nu_{in}(t)dt = 1, \qquad (15)$$

where ν_{in} is a fluctuation growth rate, it is possible to determine whether the filamentary structure can be formed during propagation of the ionization wave toward the cathode.

The fluctuation growth rate is determined by the formula [41]

$$\nu_{in} = -\frac{\mu_i U_c}{d_c^2}\left(\frac{\partial U_c}{\partial j}\sigma\frac{2\pi}{\lambda} + 1\right), \qquad (16)$$

where U_c is the cathode potential drop, d_c is the cathode sheath depth, σ is the column conductivity, and λ is the distance between the maxima of the fluctuations.

To determine whether criterion (15) is satisfied, it is necessary to calculate the parameters U_c, d_c, σ, λ, and $\partial U_c/\partial j$ in formula (16). As was mentioned above, the cathode region is characterized by an enhanced electric field. Therefore, the depth of this region can be defined as the distance d_c from the cathode at which the field becomes equal to U/d (average field in the gap). U_c is the potential at this point. The parameter $\partial U_c/\partial j$ is equal to

$$\frac{\partial U_c}{\partial j} = \frac{\partial U_c}{\partial t}\left(\frac{\partial j}{\partial t}\right)^{-1}. \qquad (17)$$

The normal current density j_n is defined as the current density corresponding to the local minimum of the current–voltage characteristic. The value of d_c at $j = j_n$ is designated as d_{cn} (Fig. 5).

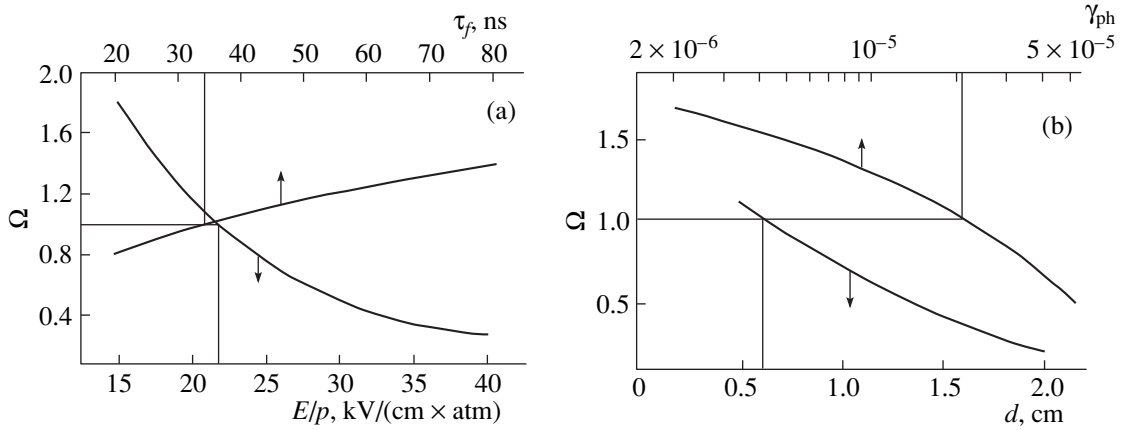


Fig. 6. Dependences of the parameter Ω on (a) the electric field and voltage rise time and (b) the interelectrode spacing and photoemission coefficient.

The parameter λ in formula (16) can be determined from the shielding criterion proposed in [42]: if the spacing between metal micropoints is less than their half-height, the perturbations introduced by micropoints into the external electric field are insignificant. The same is true for any perturbation with elevated conductivity and with a shape similar to that of a micropoint. It is assumed that, if $\lambda < \lambda_n = d_{cn}/2$ and $d_c \geq d_{cn}$, then the perturbations do not disturb the field; consequently, they do not cause the formation of a channel. At $d_c < d_{cn}$, the perturbations damp due to the positive slope of the current–voltage characteristic. When $\lambda \geq \lambda_n$, perturbations with $\lambda = \lambda_n$ grow with the highest rate. For this reason, we used the value $\lambda = \lambda_n$ to calculate Ω.

Figure 6a presents the calculated dependences of Ω on the applied voltage U (expressed in E/p) and the voltage rise time τ_f; the other parameters are kept constant. It is seen that, as E/p increases, Ω decreases. At $E/p > 20$ kV/(cm atm), the condition $\Omega < 1$, necessary to ignite a VD, is satisfied. In contrast, Ω increases with increasing τ_f. For $E/p = 20$ kV/(cm atm) and $d = 0.5$ cm, VD ignition becomes impossible at $\tau_f > 40$ ns because of the onset of cathode instability. Both of these tendencies were observed in experiments. For example, in [43], a VD was realized at $\tau_f = 10$ ns, $U = 200$ kV, and interelectrode spacing of 2.5 cm in the gas mixture $CO_2 : N_2 : He = 8 : 7 : 85$ at pressures up to 6 atm.

Figure 6b presents the dependences of Ω on the internal parameters of the discharge gap, namely, cathode emissivity and interelectrode spacing d. Both curves are seen to descend with γ_{ph} and d, which is not surprising, because an increase in both parameters leads to an increased electron photoemission current due to the increase in the cathode emissivity in the former case and to the increase in the photon flux onto the cathode caused by the increased discharge volume in the latter case. The effects in question were confirmed experimentally. It is known that a large-volume discharge can be excited by a voltage pulse with a fairly long rise time [10] and that the increase in the electrode emissivity results in a drastic increase in the energy input into the gas [37].

The fact that Ω decreases with increasing electric field (Fig. 6a, curve *1*) explains the existence of a high-voltage threshold for VD ignition. The dependence of the threshold on the pressure was calculated under conditions of experiment [31]. The value of γ_{ph} was chosen such that the threshold voltage obtained from the condition $\Omega = 1$ was equal to the measured one at $P = 1$ atm. Then, using this value of γ_{ph} and the condition $\Omega = 1$, we calculated the threshold voltage for different pressures. The results shown in Fig. 7 are in good agreement with the experimental data of [31]. They also show that the high-voltage threshold for VD ignition (curve *1*) is related to neither the transition from a Townsend to streamer discharge (curve *2*) nor the static breakdown voltage (curve *3*), but is determined by the onset of cathode instability.

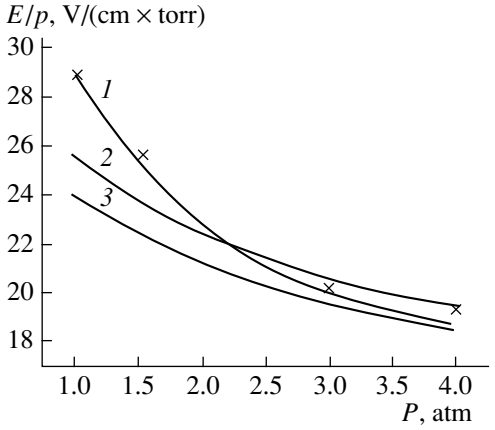


Fig. 7. (*1*) High-voltage threshold for VD ignition (crosses show the experimental data from [31]) and the threshold electric field for (*2*) an avalanche–streamer transition and (*3*) a static breakdown voltage vs. pressure.

CONCLUSIONS

The main results obtained within the proposed model of a VD are as follows. It is shown that, in the early phase of VD development, there is a stage in which the discharge current is determined by the cathode photoemission current, which distinguishes a VD from both ordinary and transient glow discharges. During this stage, a VD is least stable and, under certain conditions, becomes filamentary due to the onset of cathode instability. The threshold conditions for this instability determine the high-voltage threshold for VD ignition. Cathode instability can be suppressed by injecting electrons into the cathode region or by using a plasma cathode or high-emissive cathode, as well as by a decrease in the rise time or an increase in the amplitude of the high-voltage pulse applied to the discharge gap.

REFERENCES

1. G. A. Mesyats, Yu. I. Bychkov, and V. V. Kremnev, Usp. Fiz. Nauk **107**, 201 (1972) [Sov. Phys. Usp. **15**, 282 (1972)].
2. Yu. I. Bychkov, Yu. D. Korolev, G. A. Mesyats, *et al.*, *Injection Electronics* (Nauka, Novosibirsk, 1982).
3. E. P. Velikhov, V. Yu. Baranov, V. S. Letokhov, *et al.*, *Pulsed CO_2 Lasers for Channel Separation* (Nauka, Moscow, 1983).
4. V. Yu. Baranov, V. M. Borisov, and Yu. Yu. Stepanov, *Inert-Gas Halogenide Electric-Discharge Excimer Lasers* (Énergoatomizdat, Moscow, 1988).
5. G. A. Mesyats, V. V. Osipov, and V. F. Tarasenko, *Pulsed Gas Lasers* (Nauka, Moscow, 1991).

6. Yu. D. Korolev and G. A. Mesyats, *Autoelectronic Emission and Explosion Processes in a Gas Discharge* (Nauka, Novosibirsk, 1982).
7. E. P. Velikhov, A. S. Kovalev, and A. T. Rakhimov, *Physical Phenomena in a Gas Discharge Plasma* (Nauka, Moscow, 1987).
8. Yu. D. Korolev and G. A. Mesyats, *The Physics of Pulsed Gas Breakdown* (Nauka, Moscow, 1991).
9. A. P. Napartovich and A. N. Starostin, Khim. Plazmy, No. 6, 153 (1979).
10. V. V. Apollonov, G. G. Baĭtsur, *et al.*, Kvantovaya Élektron. (Moscow) **14**, 220 (1987).
11. A. I. Pavlovskiĭ, V. F. Basmanov, *et al.*, in *Initiation and Development of a High-Pressure Gas Discharge: Proceedings of the II All-Union Conference on the Physics of Electric Breakdown in Gases, Tartu, 1984*, Part 2, p. 330.
12. W. Rogowski and R. Tamm, Arch. Electrotek. **20** (107), 625 (1928).
13. R. F. Saxe, Br. J. Appl. Phys. **7** (9), 336 (1956).
14. G. Schroder, in *Proceedings of the 7th Conference on Phenomena in Ionized Gases, Belgrade, 1965*, p. 152.
15. A. A. Doran and J. Mayer, Br. J. Appl. Phys. **18**, 793 (1967).
16. M. C. Gavenor and J. Mayer, Aust. J. Phys. **22** (2), 155 (1969).
17. A. A. Doran, Z. Phys. **208**, 427 (1968).
18. I. D. Chalmers, J. Phys. D **4**, 1147 (1971).
19. G. A. Mesyats, Doctoral Dissertation (Tomsk, 1966).
20. Yu. I. Bychkov, G. A. Mesyats, and A. M. Iskol'dskiĭ, Zh. Tekh. Fiz. **38**, 1281 (1968) [Sov. Phys. Tech. Phys. **13**, 1051 (1968)].
21. Yu. I. Bychkov *et al.*, in *Proceedings of All-Union Conference on the Physics and Generators of Nanosecond Plasma, Alma-Ata, 1970*, p. 470.
22. K. A. Klimenko, A. V. Kozyrev, *et al.*, Fiz. Plazmy **10**, 336 (1984) [Sov. J. Plasma Phys. **10**, 107 (1984)].
23. P. S. Palmer, Appl. Phys. Lett. **25** (3), 138 (1974).
24. J. I. Levatter and S. C. Lin, J. Appl. Phys. **51**, 210 (1980).
25. V. N. Karnyushin and R. I. Soloukhin, Dokl. Akad. Nauk SSSR **236**, 347 (1977) [Sov. Phys. Dokl. **22**, 521 (1977)].
26. M. S. Kushner, IEEE Trans. Plasma Sci. **19**, 387 (1991).
27. V. V. Osipov and V. V. Lisenkov, Pis'ma Zh. Tekh. Fiz. **22** (19), 74 (1996) [Tech. Phys. Lett. **22**, 810 (1996)].
28. V. V. Osipov and V. V. Lisenkov, Opt. Atmos. Okeana **10**, 1260 (1997).
29. H. Shields, A. S. Alcock, and R. S. Taylor, Appl. Phys. **31**, 27 (1983).
30. A. Belasri, J. P. Boeuf, and L. C. Pitchford, J. Appl. Phys. **74**, 1553 (1993).
31. Yu. I. Bychkov, V. V. Osipov, *et al.*, Izv. Vyssh. Uchebn. Zaved., Ser. Fiz., No. 4, 89 (1986).
32. J. J. Lowke, A. V. Phelps, and B. W. Irwin, J. Appl. Phys. **44**, 4664 (1973).
33. H. Hokazono and H. Fujimoto, J. Appl. Phys. **62**, 1585 (1987).
34. R. Dreiskemper, G. Schroder, *et al.*, IEEE Trans. Plasma Sci. **23**, 180 (1995).
35. V. V. Apollonov, G. G. Baĭtsur, *et al.*, Pis'ma Zh. Tekh. Fiz. **11**, 1260 (1985) [Sov. Tech. Phys. Lett. **11**, 521 (1985)].
36. Yu. I. Bychkov, D. Yu. Zaroslov, *et al.*, Zh. Tekh. Fiz. **53**, 2138 (1983) [Sov. Phys. Tech. Phys. **28**, 1309 (1983)].
37. G. A. Mesyats, V. V. Osipov, *et al.*, Zh. Tekh. Fiz. **60** (4), 143 (1990) [Sov. Phys. Tech. Phys. **35**, 486 (1990)].
38. P. R. Pearson and H. M. Lambertson, IEEE J. Quantum Electron. **8**, 145 (1972).
39. S. Ya. Bronin, V. M. Kolobov, *et al.*, Teplofiz. Vys. Temp. **18**, 46 (1980).
40. A. V. Kozyrev, Yu. D. Korolev, and G. A. Mesyats, Zh. Tekh. Fiz. **57**, 58 (1987) [Sov. Phys. Tech. Phys. **32**, 34 (1987)].
41. A. V. Kozyrev and Yu. D. Korolev, Fiz. Plazmy **9**, 864 (1983) [Sov. J. Plasma Phys. **9**, 503 (1983)].
42. S. Ya. Belomyttsev, S. P. Bugaev, *et al.*, Izv. Vyssh. Uchebn. Zaved., Ser. Fiz., No. 11, 142 (1975).
43. D. I. Hidson, V. Makios, and R. Morrison, Phys. Lett. A **40**, 413 (1972).

Translated by N. Ustinovskiĭ

Technical Physics, Vol. 45, No. 10, 2000, pp. 1265–1270. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 34–39.

Breakdown in Air in a Rising Microwave Field

V. S. Barashenkov*, L. P. Grachev, I. I. Esakov**, B. F. Kostenko*, K. V. Khodataev*, and M. Z. Yur'ev***

** Joint Institute of Nuclear Research, Dubna, Moscow oblast, 141980 Russia*

*** Moscow Radiotechnical Institute, Russian Academy of Sciences, Varshavskoe sh. 132, Moscow, 113519 Russia*

E-mail: esakov@dataforce.net

Received October 25, 1999

Abstract—Results are presented from studies of a high-pressure electrodeless breakdown in air at the focus of a standing wave in a high-Q quasi-optical two-mirror resonator pumped by single microwave pulses. In the experiment, the breakdown occurred at the front of the pulse of the resonator field. The breakdown field substantially exceeded the critical level and, under fixed conditions, showed a scatter from pulse to pulse. It is shown that the experimentally found excess in the threshold breakdown field over the critical level is due to the fact that the resonator field increases as a discharge plasmoid forms during breakdown and that the appearance of an electron initiating breakdown in a gas is a random event. *© 2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Electrodeless, freely localized microwave discharges in quasi-optical beams of electromagnetic (EM) waves in gases at high pressures appear in the form of thin streamer channels. On reaching the resonance length, they have an EM energy absorption cross section several orders of magnitude higher than their geometric cross section [1, 2]. The energy absorbed by this resonant microwave plasma vibrator can be accumulated in a very small volume around its center [3]. These features of high-pressure microwave discharges suggest ways of applying them. At present, the possibility is being considered of using such discharges to create point pulsed sources of hard UV and X radiation, to inject plasmas into various types of plasma confinement systems, to employ them in studies of the nature of ball lightning, etc.

However, studies of high-pressure electrodeless microwave discharges in gases encounter certain difficulties. Thus, for a microwave breakdown in atmospheric air to occur, it is necessary to produce power densities of $\Pi > 1$ MW/cm^2. To achieve such levels of Π even in short-focus EM beams, in which the focus spot size is restricted by diffraction to the radiation wavelength λ, an EM beam power, e.g., for $\lambda \cong 10$ cm, as high as 100 MW is necessary. Such fields can be produced only by employing various types of microwave resonators, in particular, a high-Q two-mirror quasi-optical open resonator [3].

Because of technical problems related to the excitation of high-pressure electrodeless microwave discharges, these discharges are still rarely investigated. This particularly concerns the field amplitude needed to excite discharges under various conditions. To date, gas breakdown in a uniform continuous microwave field in the simplest configurations bounded by walls that absorb charges has been studied rather extensively [4]. Breakdown occurring in a field that is continuous and uniform but localized far from the surfaces that neutralize charges has also been investigated [5]. An examination of how the amplitude of a pulsed field needed for breakdown is affected by a dc field applied to the gas during the microwave pulse [6] has been made. A specific situation arises when a pulsed microwave discharge is excited, e.g., at the focus of a high-Q two-mirror resonator. In this case, the breakdown occurs in the region located at a considerable distance from the resonator mirrors and construction elements; the microwave field in this region is highly nonuniform, and its amplitude can vary substantially during the pulse.

In [7], the results were reported from studies of a pulsed breakdown in air at pressures of $0.1 \le p \le 400$ torr in a linearly polarized TEM-wave with $\lambda = 0.81$ cm formed at the focus of a quasi-optical two-mirror resonator with $Q = 3.4 \times 10^4$ pumped by single microwave pulses of duration $\tau_{pul} = 0.8$ μs. In experiments carried out at various pressures p, the pulsed power of the oscillator exciting the resonator was gradually elevated to the minimum level at which breakdown occurred and this value was used to calculate the threshold breakdown field E_{br}. The dependence $E_{br}(p)$ obtained in this way coincided with the dependence predicted by the theory developed by that time.

This paper is also devoted to the study of a pulsed breakdown in air at the focus of a two-mirror microwave resonator with $Q = 5 \times 10^4$. In contrast to [7], the measurements were conducted at substantially higher pressures (to $p = 760$ torr); a significantly longer wavelength ($\lambda = 8.9$ cm); a pulse duration of $\tau_{pul} = 40$ μs; and, what is important, a constant level of pulsed power of the oscillator exciting the resonator. In the experiments, breakdown occurred at the front of the envelope

1063-7842/00/4510-1265$20.00 © 2000 MAIK "Nauka/Interperiodica"

of the pulsed resonator field and the experimental dependence $E_{br}(p)$ differed substantially (by several times) with respect to the field level from a similar dependence obtained in [7]. Below, we explain this difference by the fact that breakdown occurs in a growing resonator field.

EXPERIMENTAL LAYOUT

A schematic diagram of the experiment is shown in Fig. 1. An open resonator *4* was formed by two coaxial concave spherical mirrors with radii of curvature of 35 cm and diameters of 55 cm; the maximum distance between them along the *z*-axis was equal to 50.4 cm. A microwave-transparent hermetic cell *5* was placed symmetrically in the center of the resonator and was oriented perpendicularly to the *z*-axis. The cell was a quartz pipe with an inner diameter of $2H = 8$ cm and flat glass ends for observations. The air pressure in the cell ranged from 3 to 760 torr and was measured to an accuracy of ± 1.5 torr. The remainder of the resonator volume was filled with air at atmospheric pressure.

The resonator was pumped by a magnetron microwave oscillator *1* through a waveguide transmission line with a circulator *2* and a coupler *3* with a power coupling coefficient on the order of 10^{-3}. The oscillator produced flat-top pulses with a duration of $\tau_{pul} = 40$ μs; the intervals between pulses amounted to several seconds. The output microwave power at a circular frequency of $\omega \cong 2 \times 10^{10}$ s^{-1} ($\lambda = 8.9$ cm) attained several megawatts. In the experiments carried out at various pressures, the pulsed power of the oscillator was kept constant.

The field along the resonator axis had the form of a standing, linearly polarized, focused TEM mode with a maximum at the center of the resonator (and, consequently, of the cell) and with the electric **E** perpendicular to both the *z*-axis and the axis of the cell. Near the focus, the distribution of the electric field amplitude along the *z* and *r* directions is described by [8]

$$E = E_0 \cos\left(\frac{\pi}{2}\frac{z}{\Lambda}\right) e^{-(r/f)^2}, \qquad (1)$$

with $\Lambda = 2.35$ cm and $f = 6$ cm.

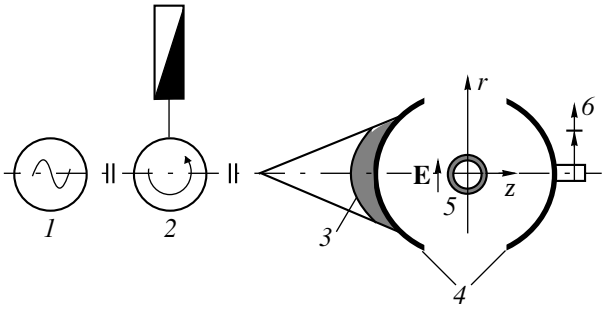


Fig. 1. Schematic diagram of the experiment.

A signal from the resonator was fed to the input of an oscillograph through a linear amplitude detector *6*. This allowed us to monitor the waveform of the field amplitude at the focus $E_0(t)$ and, after calibrating, to determine its absolute value.

After a microwave pulse was applied, the field in the resonator began to grow but its spatial structure remained unchanged. The observed time behavior of the resonator field is fairly well described by the dependence

$$E_0 = E_m(1 - e^{-t/\tau}) \qquad (2)$$

with a time constant $\tau = 5$ μs. If the air pressure in the cell was high and the field E_0 was insufficient for breakdown, then, after reaching the maximum value, the field in the resonator remained constant to the end of the microwave pulse, after which it decayed exponentially with the same time constant τ.

EXPERIMENTAL RESULTS

Measurements were carried out by the following procedure. A certain level of pressure (starting from low levels) was established in the cell; then, a microwave pulse was fed to the resonator input. The field strength at the resonator focus rose according to dependence (2), and, when it reached a certain level $E_0 = E_{br}$, breakdown of air occurred in the cell. No energy could then be supplied to the resonator, and the field in it rapidly fell with the characteristic decay time shorter than one microsecond. At low *p*, the breakdown occurred at the leading edge of the resonator field envelope observed on the oscillograph display. As *p* increased, the breakdown occurred at later times and the field E_{br} increased accordingly. In the experiments carried out at a fixed level of the pulsed power exciting the resonator, the breakdown of air in the cell occurred only up to $p_{max} = 760$ torr. At this pressure, rapid decay of the resonator field due to breakdown was observed at the plateau of the resonator field envelope. Experimental conditions at p_{max} were identical to those of [7]. At a pressure of 750 torr, the treatment of the measurement results (the treatment procedure is described in [7]) gave a maximum field amplitude $E_m = 32$ kV/cm at the focus. This point was used for absolute calibration of the oscillograph signal.

Because of the scatter in the value of E_{br} from pulse to pulse, this value was measured several times for each *p*. At $p = p_{max}$, the field value $E_{br} = E_m$ apparently did not change and a scatter in the breakdown time (within the plateau of the pulse envelope) was observed. The results of the $E_{br}(p)$ measurements are shown in Fig. 2. Visual observations through the cell end showed that the discharge was located near the cell axis. The azimuthally symmetric scatter in the positions of the streamer center in different shots was 1–2 cm.

DISCUSSION

The conventional method for measuring the breakdown microwave field is based on the assumption that the field is constant during the pulse and changes only due to the breakdown itself. The following inequality is accepted as a criterion for a breakdown [7]:

$$\nu_i \geq \nu_a + \nu_d + \gamma, \tag{3}$$

where ν_i, ν_a, and ν_d are the rates of ionization, electron attachment, and electron diffusion from the discharge region and γ is the electron-avalanche time constant. In criterion (3), the electron attachment rate in air ν_a can be estimated by the formula [9]

$$\nu_a = 2 \times 10^4 p, \text{ s}^{-1} \tag{4}$$

(here and in other similar formulas, p is expressed in torr).

If the microwave field in the discharge region bounded by the charge-neutralizing surfaces is uniform, then the rate ν_d in (3) can be estimated from the relation [4]

$$\nu_d = F^2 \frac{D}{L^2}, \text{ s}^{-1}, \tag{5}$$

where the electron diffusivity in air is equal to [10]

$$D = \frac{1.6 \times 10^6}{p}, \frac{\text{cm}^2}{\text{s}}. \tag{6}$$

For a plane layer with a gap d and a long cylinder or a sphere with radius a, the characteristic length L and the form factor F are $L = d$, a and $F = \pi$, 2.4, π, respectively. In criterion (3), one widely used approximation of the ionization rate for air has the form [9, 10]

$$\nu_i = \nu_a \left(\frac{E_0}{E_k}\right)^\beta, \text{ s}^{-1}, \tag{7}$$

with $\beta = 5.34$. Here, E_k is the critical breakdown field and the E_0/E_k ratio lies in the range $1.3 \leq E_0/E_k \leq 3$.

Under our experimental conditions ($p \leq 800$ torr), the electron energy does not follow cyclic variations of the field; i.e., the following inequality holds:

$$\nu_c \delta \leq \frac{\omega}{\pi}, \tag{8}$$

where $\delta = 2.1 \times 10^{-3}$ is the fraction of electron energy lost by electrons in collisions with molecules [11].

In this case, the amplitude of the critical breakdown field can be calculated by the formula [4]

$$E_k = 30 p \sqrt{2} \sqrt{1 + \left(\frac{\omega}{\nu_c}\right)^2}, \frac{\text{V}}{\text{cm}}, \tag{9}$$

where the frequency of electron–molecule collisions is [4]

$$\nu_c = 4 \times 10^9 p, \text{ s}^{-1}. \tag{10}$$

The commonly used models of low-, intermediate-, and high-pressure microwave discharges are based on the approximation of formula (9) [6]. Thus, at high pressures, when the inequality $\nu_c > \omega$ holds, we have $E_k \cong 42p$ [V/cm]. At low pressures, when the inequality $\nu_c < \omega$ is satisfied, we obtain $E_k \cong 10^{-8}\omega$ [V/cm]. In our experiments, we had $\nu_c = \omega$ at $p = 5$ torr; hence, formally, we investigated high- and intermediate-pressure microwave discharges.

Finally, the avalanche constant in criterion (3), which is equal to [6]

$$\gamma = \frac{\ln(N_m/N_0)}{\tau_{\text{pul}}}, \text{ s}^{-1} \tag{11}$$

should ensure the increase in the number (or density) of electrons from the initial level N_0 to the value N_m observed in the experiment during the time τ_{pul}. For estimates, the value $\ln(N_m/N_0)$ is usually assumed to be equal to 20–30 [12].

In view of (7), criterion (3) yields the following value of the threshold breakdown field:

$$E_{\text{br}} = E_k \left(1 + \frac{\nu_d}{\nu_a} + \frac{\gamma}{\nu_a}\right)^{1/\beta}. \tag{12}$$

It follows from this expression that the value of E_{br} is minimum if the field is uniform and constant ($\nu_d = \gamma = 0$); this value is called the critical field. At $E_0 = E_k$, ionization should only balance out electron losses due to attachment [see (7)].

Let us estimate the ν_d/ν_a and γ/ν_a ratios under our experimental conditions. As was noted above, a diffusion theory of microwave breakdown in a highly nonuniform field is still lacking; however, at relatively high p, we can use some approximations. In the experiments, a drop in the field amplitude, e.g., by 15%, occurs as the observation point shifts from the focus along the z-axis over a distance of $z_0 = 0.75$ cm and along r over $r_0 = 2.2$ cm. When estimating ν_d, we

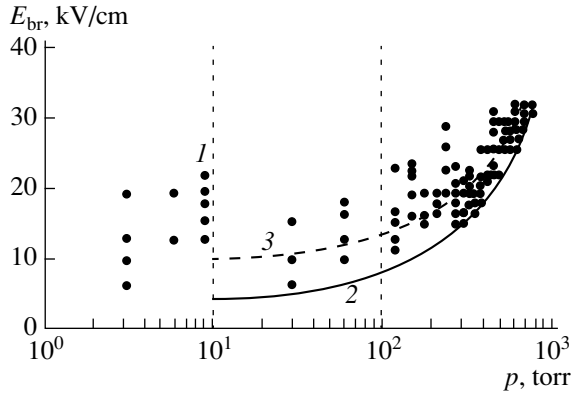


Fig. 2. Pressure dependence of the breakdown field: (*1*) experiment, (*2*) calculation, and (*3*) calculation by the approximate formula.

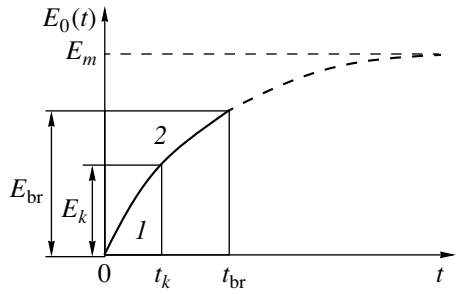


Fig. 3. Time dependence of the field at the resonator focus: (*1*) interval in which the multiplication of electrons is suppressed by attachment and (*2*) interval in which the discharge develops from a single electron to a formed plasma vibrator.

assume that breakdown occurs in this cylindrical region of length $2z_0$ and radius r_0 and that the field amplitude (and consequently the ionization rate) in this region is constant. Then, for this region, we can use formula (5) if the electron diffusion attachment length

$$l_a = \sqrt{\frac{D}{\nu_a}} = \frac{9}{p}, \text{ cm} \tag{13}$$

satisfies the inequality

$$l_a \ll 2z_0, r_0. \tag{14}$$

In this case, outside of the ionization region, at the boundaries of which the field is $E(r, z) = E_{\rm br}$, the electron density drops rapidly because of efficient electron attachment; i.e., air plays the role of an absorbing boundary surface. In the experiments, inequality (14) holds for pressures of $p \geq 60$ torr, for which formula (5) yields $\nu_d/\nu_a \leq 0.1$. Hence, under our experimental conditions (at least at $p \geq 60$ torr), we can ignore the influence of diffusion electron losses from the breakdown region on the value of $E_{\rm br}$. At the same time, it follows from Fig. 2 that, for $p = 60$ torr, the measured value of the breakdown field is $E_{\rm br} > 6$ kV/cm; i.e., it is higher than $E_{\rm k} = 42p = 2.5$ kV/cm by a factor of more than 2.5. From formula (12), taking into account expressions (4) and (11), in which, for definiteness, we set $\ln(N_m/N_0) = 20$ and use the experimental value for $\tau_{\rm pul}$, we find that the ratio γ/ν_a increases the value of $E_{\rm br}$ by 10% over E_k only at $p \leq 60$ torr. Therefore, taking into account a finite pulse duration cannot explain the experimental results.

It should be noted that the presence of the constant γ in criterion (4) makes the latter somewhat ambiguous. The term "breakdown" in the literature has the meaning of an effect which, in essence, is not a finite-duration process [4], but rather the condition that the field applied to the gas exceeds a certain threshold. At the same time, the process that starts when the breakdown voltage is applied to the gas is referred to as a gas discharge. The ambiguity is only eliminated in the course of an experiment by measuring certain discharge parameters. During the discharge, the gas begins to glow, it becomes conductive, the produced plasma disturbs the initial electric field, and so on. $E_{\rm br}$ measurements based on recording different discharge parameters can apparently give different quantitative results.

Let us interpret the processes occurring in our experiment using Fig. 3, which shows the time dependence of the field at the resonator focus, starting from the instant when the microwave pulse $E_0(t)$ is applied.

In the interval $0 \leq t < t_k$, the field is below threshold ($E_0(t) < E_k$) and the multiplication of electrons appearing in the focal region is suppressed by their attachment. The conditions for electron multiplication are only satisfied at $t \geq t_k$. At the same time, at $t > t_k$ (up to a certain instant), the discharge plasma that is produced does not influence the process of pumping EM energy into the resonator and the field in it continues to grow. As a result of discharge evolution, by the instant $t_{\rm br}$, a resonant microwave plasma vibrator forms and begins to intensively absorb the EM energy accumulated in the resonator [1]. This leads to unbalancing of the oscillator–resonator system; as a result, the field in the resonator drops abruptly. In the experiment, we in fact measure a value that is proportional to $E_{\rm br}(t_{\rm br})$ and is thus referred to as the threshold breakdown field.

We describe this process quantitatively, neglecting diffusion, whose influence, as was shown, is insignificant.

The continuity equation for electrons in a given EM field can be written in the form

$$\frac{d}{dt}\ln(N) = \nu_i(E_0) - \nu_a. \tag{15}$$

Differentiation of expression (2) with respect to time gives

$$\frac{dE_0}{dt} = \frac{E_m - E_0}{\tau}. \tag{16}$$

Dividing Eq. (15) by Eq. (16) and integrating over E_0, we obtain the integral equation which determines the breakdown field:

$$\ln\frac{N_m}{N_0} = \int_{E_k}^{E_{\rm br}} \frac{(\nu_i(E_0) - \nu_a)\tau}{E_m - E_0} dE_0. \tag{17}$$

The dependence $E_{\rm br}(p)$ obtained by numerically solving Eq. (17) for $E_m = 32$ kV/cm, $\ln(N_m/N_0) = 20$, and the function $\nu_i(E_0)$ approximating the actual dependence of the ionization rate on the field amplitude is shown in Fig. 2 by curve *2*. It is seen that the results of the calculations confirm the assumption that growth of the field is one of the reasons why the measurements disagree markedly with criterion (3).

We expand the dependence $\nu_i(E_0)$ (7) in series in the critical field, correct to first order. Then, taking into account criterion (2), we substitute the result obtained

in Eq. (15) and integrate it over t from t_k to t_{br}. As a result, we obtain a transcendental algebraic equation for calculating the relative breakdown-avalanche time $(t_{br} - t_k)/\tau$:

$$A + 1 - e^{-(t_{br}-t_k)/\tau} = \frac{t_{br}-t_k}{\tau}, \tag{18}$$

where

$$A = \frac{E_k \ln\dfrac{N_m}{N_0}}{\tau \nu_a \beta (E_m - E_k)}. \tag{19}$$

In turn, this quantity allows us to calculate E_{br} from criterion (2). Under our experimental conditions, we have

$$\frac{t_{br}-t_k}{\tau} < 1 \tag{20}$$

and formula (18) is simplified to

$$\frac{t_{br}-t_k}{\tau} = \sqrt{2A}. \tag{21}$$

Hence, E_{br} is determined by the approximate equality

$$E_{br} \cong E_k \left[1 + \left(\frac{E_m}{E_k} - 1\right)\sqrt{\frac{2\ln\dfrac{N_m}{N_0}}{\nu_a \tau \beta\left(\dfrac{E_m}{E_k} - 1\right)}}\right]. \tag{22}$$

Curve *3* in Fig. 2 shows the corresponding dependence $E_{br}(p)$. It is seen that this dependence coincides with the measured values within the limits of statistical scatter. All of the dependences $E_{br}(p)$ shown in Fig. 2 are calculated assuming that $\ln(N_m/N_0) = 20$. Note, however, that the results of the calculations depend only weakly on the logarithm value.

The suggested approach to determining the breakdown field also gives reasonable results in two limiting cases. If the field applied to the gas increases slowly ($\tau \longrightarrow \infty$), we obtain $E_{br} \longrightarrow E_k$. If the field is applied rapidly ($\tau \longrightarrow 0$), we obtain $E_{br} = E_m$, which, in turn, cannot be less than E_k.

In the above considerations, we assumed that, at the instant $t = t_k$, there are free electrons in the focal region. In fact, the question of their origin is rather complicated. In similar experiments [3], it was shown that a discharge is generally initiated by one electron ($N_0 = 1$). Obviously, the fact that the measured values of E_{br} lie above the theoretical curve *2* in Fig. 2 is a consequence of the fact that the appearance of electrons initiating the discharge is a random event.

In air, at a background radiation level near the earth's surface, the number of electrons generated in a cubic centimeter per second is equal to $J_q = (1.3\text{–}4) \times 10^{-2} p$. As an example, for $p = 760$ torr, this gives (10–30) $\text{cm}^{-3}\ \text{s}^{-1}$ [12]. At room temperature, the generated electrons rapidly attach to oxygen molecules, due to three-body attachment, at the rate

$$\nu_{a1} = (10p)^2\ \text{s}^{-1}. \tag{23}$$

For example, at 760 torr, the attachment time is equal to $1/\nu_{a1} = 1.7 \times 10^{-8}$ s. Assuming that, in our experiments, $E_0 \geq E_k$ in the volume $V = \pi r_0^2 2z_0 \cong 23\ \text{cm}^3$, we find that the average time between successive events of generating an electron is equal to $t_q = 1/(J_q V) \cong 4 \times 10^{-3}$ s. Evidently, the probability of generating a free electron during a 40-μs microwave pulse is negligibly small.

Most probably, the electron initiating the discharge is generated through its detachment from a negative oxygen ion. This occurs when the microwave field has already achieved its critical level. The initial density of the negative ions in air is

$$n_{O_2^-} = \sqrt{\frac{J_q}{\alpha_r}},\ \text{cm}^{-3}, \tag{24}$$

where

$$\alpha_r = 2 \times 10^{-6}\ \text{cm}^3\ \text{s}^{-1} \tag{25}$$

is the ion–ion recombination coefficient.

For example, at $p = 760$ torr, the ion density is $4 \times 10^3\ \text{cm}^{-3}$. The observed scatter in the E_{br} values at $p = \text{const}$ indicates that the process of generating the electron initiating the breakdown is random, with a microsecond time scale and centimeter space scale. An analysis of this process is beyond the scope of this paper.

CONCLUSION

Studies of electrodeless breakdown in air at intermediate and high pressures in the microwave field of a quasi-optical two-mirror resonator have shown that, under the given experimental conditions, the threshold breakdown field can substantially exceed the critical breakdown field.

The observed effect is explained by the fact that the breakdown occurs in the growing field. The higher the amplitude of the field applied to the gas and the higher the rate at which the field grows, the higher the excess value. When the applied field substantially exceeds the critical level, the effect in question is independent of the gas pressure and is dominant at high pressures when the influence of diffusion on the breakdown field is negligibly small.

The experiments have shown that, when the discharge is initiated at the front of the envelope of the microwave resonator field, the final stage of the discharge, in which the energy stored in the resonator is intensively absorbed by the discharge plasma, occurs in

a field whose value substantially exceeds the critical level. This fact is important for practical applications of this type of discharge, because, in this case, the rate of ionization processes forming the electrodynamic properties of the discharge is significantly higher than the rate at which the gas in the discharge channel is heated.

The described experimental procedure for measuring the value of the breakdown field and its statistical characteristics can be employed for further studies of some physical phenomena, e.g., the generation of free electrons in gases in microwave fields above the critical level and the formation of the resonant properties of a streamer microwave discharge.

ACKNOWLEDGMENTS

We thank K.V. Aleksandrov for assistance in the experiments.

REFERENCES

1. L. P. Grachev, I. I. Esakov, G. I. Mishin, and K. V. Khodataev, Zh. Tekh. Fiz. **66** (7), 32 (1996) [Tech. Phys. **41**, 652 (1996)].
2. A. L. Vikharev, V. B. Gil'denburg, S. V. Golubev, *et al.*, Zh. Éksp. Teor. Fiz. **94** (4), 136 (1988) [Sov. Phys. JETP **67**, 724 (1988)].
3. L. P. Grachev, I. I. Esakov, G. I. Mishin, and K. V. Khodataev, Zh. Tekh. Fiz. **64** (2), 26 (1994) [Tech. Phys. **39**, 130 (1994)].
4. A. D. MacDonald, *Microwave Breakdown in Gases* (Wiley, New York, 1966; Mir, Moscow, 1969).
5. V. B. Gil'denburg and V. E. Semenov, Fiz. Plazmy **14**, 494 (1988) [Sov. J. Plasma Phys. **14**, 292 (1988)].
6. G. M. Batanov, S. I. Gritsinin, I. A. Kossyĭ, *et al.*, Tr. FIAN **160**, 174 (1985).
7. A. L. Vikharev and B. G. Eremin, Zh. Éksp. Teor. Fiz. **68**, 452 (1975) [Sov. Phys. JETP **41**, 219 (1975)].
8. L. A. Vaĭnshteĭn, *Open Cavities and Open Waveguides* (Sov. Radio, Moscow, 1966).
9. Yu. A. Lupan, Zh. Tekh. Fiz. **46**, 2321 (1976) [Sov. Phys. Tech. Phys. **21**, 1367 (1976)].
10. A. V. Gurevich, Geomagn. Aéron. **19**, 633 (1979).
11. E. M. Bazelyan and Yu. P. Raizer, *Spark Discharge* (MFTI, Moscow, 1997; CRC, Boca Raton, 1997).
12. Yu. P. Raĭzer, *Modern Physical Foundations of Gas Discharge* (Nauka, Moscow, 1980).

Translated by N.F. Larionova

Technical Physics, Vol. 45, No. 10, 2000, pp. 1271–1275. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 40–44.

SOLID-STATE ELECTRONICS

Degradation of Gallium Arsenide under Irradiation with an Excimer Laser

V. Gradoboev and A. I. Fedorov

Institute of Atmospheric Optics Siberian Division, Russian Academy of Sciences, Akademicheskiĭ pr. 1, Tomsk, 634055 Russia

Received June 12, 1999; in final form, March 29, 2000

Abstract—The results of a study of degradation of the surface of gallium arsenide resulting from irradiation with a power excimer laser at power densities ranging from the threshold power to the power level causing local melting of the surface are presented. Two degradation mechanisms have been identified, one of which causes the formation of a thin near-surface layer of modified nonstoichimetric gallium arsenide at a power level higher than 1×10^7 W/cm^2 and the other of which causes the formation of a separate gallium phase. The formation of the separate gallium phase can be produced either by a single pulse of laser radiation with a power density exceeding 2.7×10^{11} W/cm^2 or by a few less powerful pulses. An empirical relationship has been established between the power density and the number of pulses causing the formation of the separate gallium phase. It has also been established that as a result of laser irradiation at the boundary of "cold" and "hot" gallium arsenide, periodically ordered defects in the form of blocks aligned along the [100] directions emerge. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

By now a considerable amount of work has been carried out on the degradation of a gallium arsenide surface under irradiation with a pulsed laser at different wavelengths [1–3]. These studies mainly deal with power densities that cause melting of the semiconductor surface. In [4, 5] we presented the results of a study of the degradation of a GaAs surface resulting from irradiation with a pulsed UV laser in a power density range near the threshold value. The threshold power is construed here as the lowest power density at which the degradation of a semiconductor surface can be detected. The aim of this work is to study the mechanisms that cause degradation of the surface of gallium arsenide under irradiation from a power excimer laser at power densities ranging from the threshold power to the power level causing local melting of the surface of gallium arsenide.

EXPERIMENTAL TECHNIQUE

A pulsed periodic XeCl laser with a wavelength of 308 nm, an aperture of 1×2.5 cm^2, pulse durations from 10 to 50 ns, pulse repetition rates from 1 to 10 Hz, and beam power from 0.01 to 1 MW [6, 7] was used as the source of pulsed radiation. The radiation power was measured with an IMO-2H calorimeter. Usually, various aperture diaphragms were used to improve the uniformity of the energy distribution over the beam cross-section. The shape of the radiation pulses was registered with a FEK-22CPUM photodetector and I2-7 or C8-14 oscilloscopes. The laser radiation pulse was characterized by the power density per pulse calculated from data on the pulse power and shape. The pulse power density was varied by changing the operating regimes of the laser or by focusing the beam with a quartz lens. The exposure of the surface of gallium arsenide to the laser radiation was characterized by the pulse power density and the number of incident pulses. The number of pulses was determined as the ratio of the measured time of exposure and the pulse repetition rate.

We studied light-emitting diodes whose active elements were AlGaAs layers produced by liquid-phase epitaxy on gallium arsenide substrates. The diodes were fabricated using standard procedures of photolithography, dicing into separate crystals, and thermocompression assembly into metal-plastic packages. The laser radiation was focused onto the gallium arsenide substrate of the diodes in such a way that the area of the focused spot was considerably larger than the substrate area. In this way, edge effects at the boundary between the "hot" gallium arsenide in the laser-illuminated area and the "cold" gallium arsenide outside this area could be excluded from consideration. The diodes were characterized by the integrated power of the optical emission coming out through the substrate and by its spectrum. These characteristics were measured for every diode at different phases of the investigations. It should be emphasized that in order to simplify the processing of the experimental data the diode emission power was normalized to its value prior to the tests.

In order to evaluate and control the temperature of the active region of the light-emitting diode during laser irradiation, the forward bias voltage at a specified operating current was monitored. This was done in

1063-7842/00/4510-1271$20.00

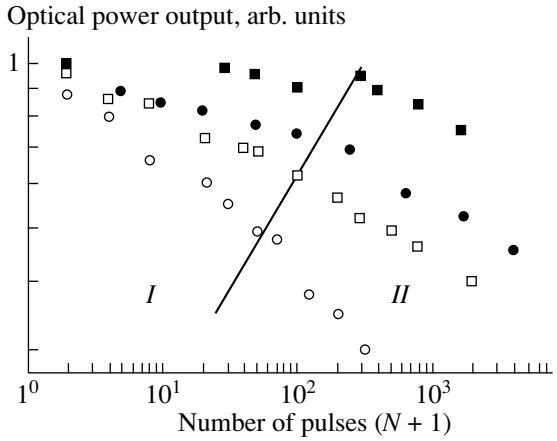


Fig. 1. Degradation of the diode optical power output as a result of pulsed laser irradiation. Laser radiation power density, W/cm^2: solid squares, 1.25×10^7; solid circles, 2.5×10^7; open squares, 6.7×10^7; open circles, 2.5×10^8. The straight line separates the regions where the two different degradation mechanisms operate.

order to avoid additional degradation of the diode optical power output that could result from a temperature rise in the active region of the diode during laser irradiation and thus isolate the degradation of the diode optical power output caused by the emergence of a nonstoichiometric layer near the gallium arsenide surface. In addition, the current–voltage characteristics of the diodes were measured before and after irradiation in order to eliminate cases of degradation of the active region of the diode due to laser radiation.

The experimental values given here represent an average of the data for batches of no less than five diodes. All experiments were conducted at normal atmospheric pressure and room temperature.

Epitaxial layers of gallium arsenide were also studied. The layers were grown by chloride gas epitaxy on heavily doped substrates and had an electron concentration ranging from 2×10^{17} to 2×10^{18} cm^{-3}. In these experiments, the area of the laser spot was considerably less than the area of the epitaxial structure, which gave rise to the edge effects mentioned above. These effects will be considered in detail in an analysis of the experimental results below.

In order to study the disturbed gallium arsenide layer produced near the exposed surface by UV laser radiation, the following methods were used. The optical spectra of the diode emission coming out through the substrate were recorded in order to study the nature of the defects in the disturbed nonstoichiometric layer of gallium arsenide. Two independent methods were employed to determine the thickness of the damaged layers. In the first method, the optical power output of a diode after exposure to laser radiation was measured, then a part of the damaged layer was removed using a polishing etchant based on $H_2SO_4 + H_2O_2$ [8], and the optical power output was measured again. This procedure was repeated until complete recovery of the optical power output, which was evidence that the disturbed nonstoichiometric layer of gallium arsenide modified by the laser radiation was completely removed. This method gives the thickness of the damaged layer to within 3–5 μm.

The second method of determining the thickness of the damaged layers uses measurements of the microcathodoluminescence profile across the cleaved surface of a gallium arsenide structure subjected to laser irradiation. These measurements were carried out with a REMMA-202 scanning electron microscope equipped with a special attachment for microcathodoluminescence measurements. The resolution over the cleavage plane in this instance was from 0.8 to 1 μm.

EXPERIMENTAL RESULTS AND DISCUSSION

UV laser radiation is outside the transparency window of gallium arsenide and is therefore absorbed in a thin near-surface layer. Constancy of the forward voltage across the light-emitting diode during irradiation is an indication that no additional generation of current carriers takes place in the active region and that the temperature gradient arising near the surface does not reach this region.

In Fig. 1, experimental results on degradation of the diode optical power output as a function of the number of laser radiation pulses of different power densities are presented. Analysis of these results suggests that two degradation mechanisms operate in the near-surface layer of gallium arsenide as a result of laser irradiation (regions *I* and *II* in Fig. 1).

In region *I*, laser irradiation of the surface of gallium arsenide causes its decomposition accompanied by the escape of arsenic to the environment [1–5]. As a result, a near-surface nonstoichiometric layer of modified gallium arsenide is formed, which contains a large number of arsenic vacancies V_{As} and gallium atoms in the arsenic sublattice sites Ga_{As}. These defects represent radiationless recombination centers in the active region of the light-emitting diode, which has been confirmed by examination of the emission spectra before and after laser irradiation. Introduction of these radiationless recombination centers into the near-surface layer of gallium arsenide (through which the radiation from the active region of the diode is brought out) results in degradation of the diode optical power output. Thus, the dependence of the diode optical power output on the laser radiation power can be used in investigations of the dynamics of degradation processes in the near-surface layer of gallium arsenide. The relationships established for region *I* make it possible to determine the threshold power density above which degradation processes in the near-surface layer of gallium arsenide develop. Figure 2 shows the dependence of the slope of the diode emission power degradation curves on the

laser radiation power density for region *I* (Fig. 1). From these data, the threshold laser radiation power density is found to be 9.6×10^6 W/cm^2, in agreement with the results published earlier [1–5].

Now consider the results of determining the thickness of the damaged layer produced near the surface of gallium arsenide by laser irradiation at a power density corresponding to region *I*. Typical results of the diode power output measurements before and after complete or partial chemical removal of the damaged layer are given in Table 1. It is seen that complete removal of the damaged layer leads to full recovery of the diode power output. Thus, the technique we employed to monitor the diode power output in the process of controlled removal of the damaged layer by chemical etching makes it possible to determine both the damaged layer thickness and the parameters of this layer as functions of the power characteristics of the laser radiation.

The results of measurements of the damaged layer thickness produced by laser irradiation near the surface of gallium arsenide, obtained both by chemical etching and by measuring microcathodoluminescence intensity profiles over a cleaved facet are given in Table 2. The results obtained in this work coincide with those known in the literature [9]. Analysis of the estimated damaged layer thickness values shows that there is only a weak dependence of the damaged layer thickness on the laser power density. This fact was noted earlier in a study using laser radiation in the visible range [10].

Measurements of the microcathodoluminescence intensity profiles in the near-surface layers following laser irradiation give thickness values 5–10 μm greater than those determined by layer-by-layer chemical etching. This discrepancy in the thickness values of the damaged layer indicates that the method which uses microcathodoluminescence from a cleaved facet is more sensitive to the presence of a damaged layer than the one based on controlled removal of the damaged layer by chemical etching until recovery of the diode power output. It should also be noted that removal of the damaged layer results in full recovery of the diode power output; this fact confirms that the degradation mechanism operates exclusively in the near-surface layer of gallium arsenide modified by the laser radiation.

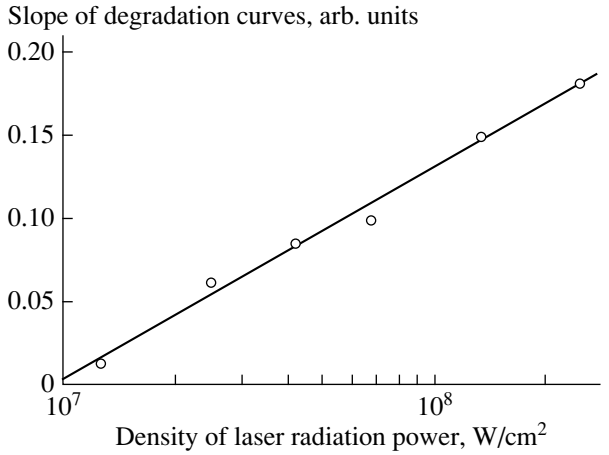


Fig. 2. Dependence of the slope of diode power output degradation curves on the laser radiation power density for region *I*.

Thus, the above investigations permit the following conclusions to be made about the degradation mechanism of the surface of gallium arsenide in region *I*. As a result of pulsed UV laser irradiation with a power density close to the threshold value, thermal decomposition of gallium arsenide takes place, with arsenic escaping to the atmosphere and the near-surface layer becoming rich in gallium. According to data available in the literature [11], the gallium arsenide surface temperature at the threshold power density of laser radiation is about 900 K, which is substantially lower than its melting point. This mechanism depends, not on the wavelength, but exclusively on the power density of the laser radiation, where the threshold power density is of the order of 1×10^7 W/cm^2. Comparing our results (Table 2) with data given in the literature [9, 10] we may conclude that the thickness of the damaged layer near the gallium arsenide surface is only weakly dependent on the laser radiation power density. It has also been established that the thickness of the damaged layer, as determined by layer-by-layer etching, is somewhat less than the value obtained from microcathodoluminescence measurements, which is explained by the higher sensitivity of the latter method.

If the laser radiation power density and/or the number of incident pulses are further increased, a different

Table 1. Recovery of the diode power output as a result of removal of the damaged layer by chemical etching

Laser radiation power density, W/cm^2	Diode power output, arb. units					
1.25×10^7	0.90	0.93	0.97	0.98	1.00	–
6.7×10^7	0.46	0.48	0.52	0.63	0.87	0.98
2.5×10^7	0.29	0.30	0.47	0.51	0.73	0.78
Thickness of the removed layer, μm	0	5	10	15	20	25

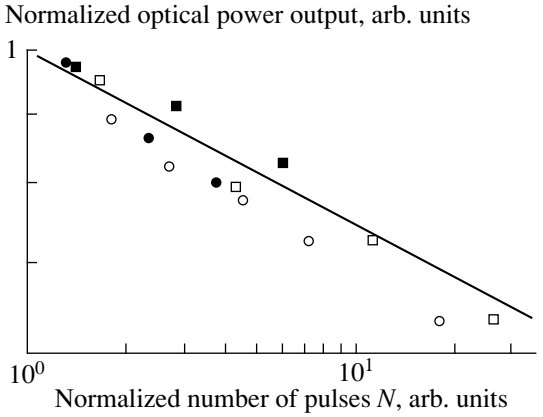


Fig. 3. Dependence of the normalized diode optical power output on the normalized number of incident pulses for region *II*. Laser radiation power density, W/cm^2: solid squares, 1.25×10^7; open squares, 2.5×10^7; open circles, 6.7×10^7; solid circles, 2.5×10^8.

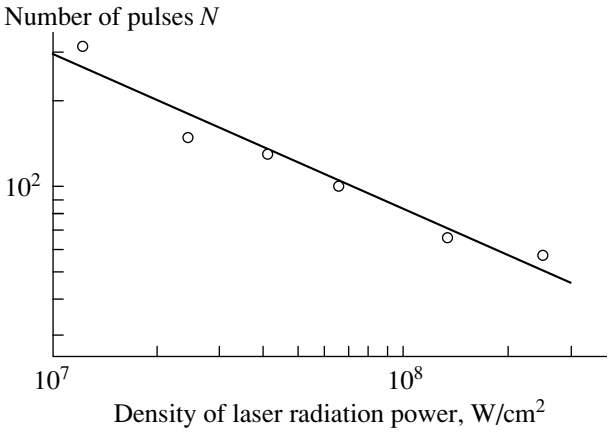


Fig. 4. Dependence of the number of incident laser pulses causing degradation by the second mechanism on the laser radiation power density.

degradation mechanism of the surface of gallium arsenide takes over (region *II* in Fig. 1). Let us consider the experimental results for this region in more detail. First, we need to be certain that a single degradation mechanism is in operation in this region. Shown in Fig. 3 is the dependence of the normalized diode power output on the normalized number of incident pulses in region *II* for various values of the laser radiation power density. The normalization was made to the diode power output and the number of incident laser radiation pulses corresponding to the onset of the second degradation mechanism. The region where the second mechanism is in operation begins at the intersection of the straight line in Fig. 1 with the corresponding degradation curves of region *I*. From these results, it can be seen that the degradation of the diode optical power output in region *II* can be described in terms of one relationship, giving evidence that a single degradation mechanism operates throughout region *II* whatever the laser radiation power density or the number of incident pulses.

Now, the energy conditions under which the second degradation mechanism of the modified layer are initiated should be defined. In Fig. 4, the dependence of the number of incident pulses needed to trigger the second degradation mechanism on the laser radiation power density is shown. From this dependence, the limiting value of the laser radiation power density has been found to equal 2.7×10^{11} W/cm^2. Starting with this value, gallium emerges at the surface as a separate metallic phase after a single pulse.

Thus, the impact of pulsed laser radiation on the surface of gallium arsenide can be described using two characteristic power density values: the threshold power density of the order of 1×10^7 W/cm^2, which causes decomposition at the surface at a temperature of about 900 K, and the limiting power density of the order of 2.7×10^{11} W/cm^2, which produces the metallic gallium phase at the semiconductor surface after a single pulse.

Next consider the physical nature of the identified degradation mechanism of the gallium arsenide surface in region *II*. Electron microscope investigations have shown that at the onset of this degradation mechanism gallium first appears as a separate phase in the form of drops. As the power density is increased, the separate drops coalesce into a continuous phase. The separate metallic phase of gallium causes reflection of the diode emission by the metallic inclusions, which is the major degradation mechanism of the diode power output (region *II* in Fig. 1).

Simultaneously, periodic defect structures in the form of rectangular blocks with their sides aligned along [110] directions emerge at the surface of gallium arsenide at the periphery of the laser spot; i.e., these structures arise at the boundary of the "cold" and "hot" gallium arsenide. Observations of such block structures have been reported in [12, 13]. Supposedly, these block structures form due to large temperature gradients occurring near the above boundary.

Formation of the separate metallic gallium phase as a result of laser-induced decomposition of the surface,

Table 2. Thickness of the damaged layer for various values of the laser radiation power density determined by different methods

Laser radiation power density, W/cm^2	Thickness of the damaged layer, μm	
	chemical etching	microcathodoluminescence
1.25×10^7	20	24
2.5×10^7	23	30
6.7×10^7	27	34
2.5×10^8	30	40

as well as the appearance of the block structures, can be caused by a single pulse of laser radiation having the limiting power density or by multiple radiation pulses of lower power density. The results presented in Figs. 1 and 4 allow one to establish the following relationship between the laser radiation power density and the threshold number of the incident pulses:

$$N_{\text{thr}} = (W_{p\text{lim}}/W_p)^{0.56},$$

where N_{thr} is the threshold number of the incident pulses for a given power density W_p capable of causing the formation of the separate gallium phase and the block structures, and $W_{p\text{lim}} = 2.7 \times 10^{11}$ W/cm^2 is the limiting value of the power density at which a separate gallium phase forms after a single laser pulse.

RESULTS AND CONCLUSIONS

(1) As the threshold power density of pulsed laser radiation, 1×10^7 W/cm^2, is reached, decomposition of the surface layer of gallium arsenide caused by the evolved heat begins and leads to the escape of arsenic to the environment and enrichment with gallium. The surface temperature of the gallium arsenide at this time is estimated at 900 K. These processes form a thin near-surface layer of modified gallium arsenide.

(2) The thickness of the modified layer is about 20–40 μm and is only weakly dependent on the laser radiation power density. Microcathodoluminescence measurements are more sensitive to disturbances in the modified layer than chemical etching and therefore give values of the damaged layer thickness that are 5–10 μm greater.

(3) A further increase in the laser radiation power density and/or the number of incident pulses causes formation of a separate phase of metallic gallium. A single laser pulse of power density 2.7×10^{11} W/cm^2 causes the formation of a separate phase of metallic gallium at the surface of gallium arsenide.

(4) An empirical relationship has been established between the laser radiation power density and the number of incident pulses required for the onset of the degradation mechanism causing formation of the separate gallium phase.

(5) As a result of laser irradiation at the boundary of the "cold" and "hot" gallium arsenide, periodic defect structures in the form of rectangular blocks aligned along [100] directions emerge at the semiconductor surface at the periphery of the laser spot. This phenomenon is observed if the laser spot size is less than the area of the structure being irradiated. The block structures can be the consequence of large temperature gradients occurring near the laser spot periphery under laser irradiation.

REFERENCES

1. G. M. Gusakov and T. N. Kondratova, Poverkhnost, No. 3, 152 (1990).
2. A. I. Efimova, P. K. Kashkarov, V. N. Petrov, and V. Yu. Timoshenko, Poverkhnost, No. 8, 94 (1990).
3. A. A. Ptashchenko and B. V. Tsap, Zh. Prikl. Spektrosk. **50**, 987 (1989).
4. A. V. Gradoboev and A. I. Fedorov, in *Proceedings IV International Conference "Pulsed Lasers on Atomic and Molecular Transition", Tomsk, 1999.*
5. A. V. Gradoboev and F. I. Fedorov, Opt. Atmos. Okeana **2**, 772 (1989).
6. A. I. Fedorov and S. A. Brichkov, Opt. Atmos. Okeana **2**, 772 (1989).
7. A. I. Fedorov, Opt. Atmos. Okeana **10**, 1274 (1997).
8. S. Iida and K. Ito, J. Electrochem. Soc. **118**, 768 (1971).
9. N. G. Dzhumamukhambetov and A. G. Dmitriev, Pis'ma Zh. Tekh. Fiz. **18** (4), 83 (1992) [Sov. Tech. Phys. Lett. **18**, 127 (1992)].
10. B. G. Gribov, G. M. Gusakov, T. N. Kondratova, *et al.*, Dokl. Akad. Nauk SSSR **314**, 618 (1990) [Sov. Phys. Dokl. **35**, 847 (1990)].
11. A. G. Dmitriev, Fiz. Tekh. Poluprovodn. (St. Petersburg) **27**, 583 (1993) [Semiconductors **27**, 323 (1993)].
12. P. K. Kashkarov, V. I. Petrov, D. V. Ptitsin, and V. Yu. Timoshenko, Fiz. Tekh. Poluprovodn. (Leningrad) **23**, 2080 (1989) [Sov. Phys. Semicond. **23**, 1287 (1989)].
13. A. A. Zakharova and V. I. Ryzhiĭ, Fiz. Tekh. Poluprovodn. (Leningrad) **23**, 1898 (1989) [Sov. Phys. Semicond. **23**, 1175 (1989)].

Translated by B. Kalinin

Technical Physics, Vol. 45, No. 10, 2000, pp. 1276–1280. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 45–49.

SOLID-STATE ELECTRONICS

Silicon Dual-Drain Strain-Sensitive FETs

G. G. Babichev, S. I. Kozlovskiĭ, and N. N. Sharan

Institute of Semiconductor Physics, National Academy of Sciences of Ukraine, Kiev, 252650 Ukraine

Received September 3, 1999

Abstract—Two types of silicon *p*-channel strain-sensitive FETs were investigated: an MIS strain-sensitive transistor and a junction-controlled strain-sensitive FET. The transistors are strain-sensitive lateral unipolar semiconductor devices with internal differential output. The optimum geometry of the devices was found, and their basic characteristics were calculated. © *2000 MAIK "Nauka/Interperiodica".*

One most promising approach to fabricating new semiconductor devices (electromechanical sensors and microactuators) is the integrated technology of microelectromechanical silicon systems (MEMS technology) [1]. Its extensive development makes researchers look for new types of strain-sensitive components.

Today's silicon integrated strain-sensitive resistive-output components can be subdivided into two groups: "unipolar" and "bipolar." An output signal in them is formed by the directional motion of majority and minority carriers, respectively. The first group covers strain-sensitive components based on the piezoresistive effect and the effect of piezoelectromotive force; the other, strain-sensitive components built on strain-sensitive transistors (tensotransistors) [2, 3]. Strain gauges with bipolar transistors offer high output signal and relative sensitivity but are much more power-hungry than their unipolar counterparts. Low power consumption and reasonably high output signal can be combined in strain gauges with an FET as a sensitive element.

In this work, we will discuss the operation of field-effect dual-drain tensotransistors and optimize their design. Two types of the transistors will be considered: an MIS transistor and a junction-controlled tensotransistor. Like conventional FETs [4], strain-sensitive FETs must offer a number of significant advantages over the bipolar tensotransistors. These are, in particular, weak temperature dependence of the device parameters and low noise level. In addition, the junction-controlled transistors must have higher radiation hardness.

By analogy with magnetic FETs, the sensitive element of an FET-based integrated strain gauge can be characterized by absolute sensitivity S_A and conversion efficiency, or relative sensitivity S_R [5]:

$$S_A = R_D\left|\frac{\partial(I_{D1} - I_{D2})}{\partial\sigma}\right|_{\sigma = 0}, \quad (1)$$

$$S_R = (I_{D1} + I_{D2})^{-1}\left|\frac{\partial(I_{D1} - I_{D2})}{\partial\sigma}\right|_{\sigma = 0}, \quad (2)$$

where σ is the mechanical strain in the vicinity of a tensotransistor and $I_{D1(D2)}$ are the drain ($D1$ or $D2$) currents. The load resistances in the drain circuits of the tensotransistor are assumed to be equal: $R_{D1} = R_{D2} = R_D$.

Figure 1 shows (a) the design and (b) connection diagram for a *p*-channel depletion-mode MIS tensotransistor (transistor 1). Figure 2 depicts the same for a *p*-channel junction-controlled field-effect tensotransistor (transistor 2).

We will assume that the field-effect tensotransistor is centrally located on a specially shaped silicon membrane that converts a load uniformly distributed over its surface to uniaxial elastic (compressive or tensile) strain in the [110] crystallographic direction [2].

The transistor operates as follows. When appropriate bias voltages (U_{DS}, U_G) are applied, holes, subjected to a sweeping field, drift in the channel. Note that the transistor design is mirror-symmetric about the plane passing through the channel center $y = L_y/2$ normally to the xy plane. Therefore, in the absence of strain, the drain currents are equal to each other, $I_{D1} = I_{D2} = I_D^0$, and the output V_0 (Figs. 1, 2) is zero if the load resistances are the same, $R_{D1} = R_{D2} = R_D$.

If the membrane is loaded uniformly, the mobility of majority carriers (holes) becomes anisotropic in the xy plane; as a result, extra, transverse and longitudinal, electric fields (aligned with the y and x directions, respectively, and proportional to the strain) arise in the channel.

Under steady-state conditions, the transverse potential difference due to mobility anisotropy prevents current density redistribution across the channel. Therefore, the practically significant transverse redistribution of the current will be observed only near the drains, where the transverse potential difference is short-circuited by the p^+ regions of the drains. Under the short-circuit conditions, the strain-induced anisotropy of hole mobility will cause current density redistribution

1063-7842/00/4510-1276$20.00

across the channel, drain current inequality, and the output signal to appear.

The operation of the tensotransistors is formally akin to that of many-drain magnetic FETs, where the Hall field redistributes the current between the drains [5, 6].

Let the FET channel be confined by $0 \le x \le L_x$; $0 \le y \le L_y$, and $0 \le z \le L_z$. Then, the drain current of transistor 1 (Fig. 1a) is given by

$$I_{D1(D2)} = C\mu_p \int_{0(L_y - L_x)}^{L_D(L_y)} \frac{\partial U}{\partial x}(x = L_x, y) \times [U_C - U_G - U(x = L_x, y)]dy, \quad (3)$$

where C is the gate–semiconductor capacitance per unit area, U_C is the cutoff voltage (the gate voltage at which the channel conductivity becomes zero), U_G is the gate voltage, and $U(x, y)$ is the channel potential at a point with coordinates x and y.

For transistor 2 with a uniformly doped channel (Fig. 2a), the drain current is expressed as

$$I_{D1(D2)} = \sigma_0 L_z \int_{0(L_y - L_D)}^{L_D(L_y)} \frac{\partial U}{\partial x}(x = L_x, y) \times \left[1 - \sqrt{\frac{U_G + U_i + U(x = L_x, y)}{U_C}}\right]dy. \quad (4)$$

Here, σ_0 is the metallurgical conductivity of the channel, U_i is the built-in potential of the junction, and U_C is the cutoff voltage. In (4) and (5), it is assumed that the drains $D1$ and $D2$ are of equal width L_D and are separated by $W = L_y - 2L_D$. In the presence of strain, the potential $U(x, y)$ is represented as

$$U(x, y) = U_0(x) + a\Phi(x, y), \quad (5)$$

where $U_0(x)$ is the channel potential at a point x without strain, $\Phi(x, y)$ is a potential change due to elastic strain, and a is the anisotropy parameter [2, 3].

In our case, a is linearly related to uniaxial elastic mechanical strain σ near the tensotransistor through the relationship [2]

$$a = \Pi_{44}\sigma/2, \quad (6)$$

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

For practical elastic strains, $|a| \ll 1$ and the inequality

$$U_0(x) \gg |a\Phi(x, y)| \quad (7)$$

holds.

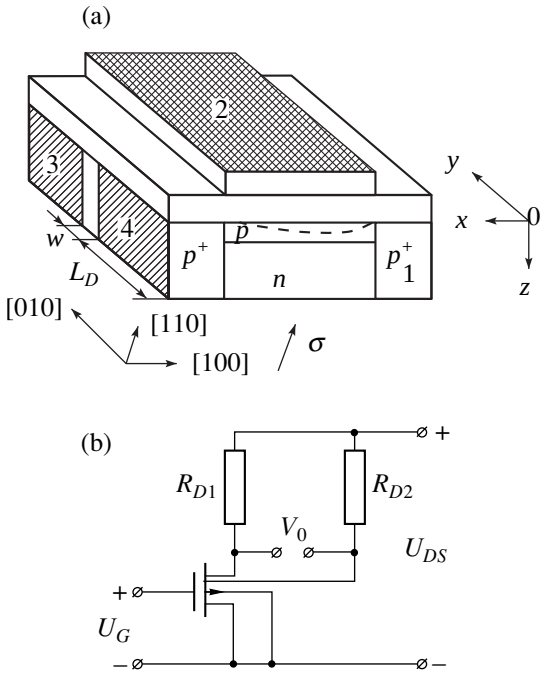


Fig. 1. Design, connection diagram, and crystallographic orientation of a dual-drain MIS tensotransistor. (*1*) source, (*2*) gate, and (*3*, *4*) drains.

Substituting expressions (3)–(6) into (2), we arrive at the expressions for absolute sensitivity S_A and conversion efficiency S_R for transistor 1:

$$S_A = \Pi_{44} C\mu_p R_D (U_C - U_G - U_{DS}) \times \int_0^{L_D} \frac{\partial \Phi}{\partial x}(x = L_x, y)dy, \quad (8)$$

$$S_R = \Pi_{44} C\mu_p \frac{(U_C - U_G - U_{DS})}{2I_D^0} \int_0^{L_D} \frac{\partial \Phi}{\partial x}(x = L_x, y)dy, \quad (9)$$

$$I_D^0 = \mu_p \frac{C}{2}\frac{L_y}{L_x}\left(U_C - U_G - \frac{U_{DS}}{2}\right)U_{DS}. \quad (10)$$

Here, $I_{D1}^0 = I_{D2}^0 = I_D^0$ are the drain currents without strain. It should be noted that formulas (8)–(10) are also valid for an MIS tensotransistor with an inversion p-channel if the signs of U_C and U_G are reversed. For transistor 2, the associated expressions have the form

$$S_A = \Pi_{44} R_D \sigma_0 L_z (1 - \sqrt{(U_G + U_i + U_{DS})/U_C}) \times \int_0^{L_D} \frac{\partial \Phi}{\partial x}(x = L_x, y)dy, \quad (11)$$

$$S_R = \frac{\Pi_{44}}{2}\sigma_0 L_z \frac{(1-\sqrt{(U_G+U_i+U_{DS})/U_C})}{I_D^0} \times \int_0^{L_D} \frac{\partial \Phi}{\partial x}(x=L_x, y)dy, \quad (12)$$

$$I_D^0 = \frac{L_z L_y}{2L_{xx}}\sigma_0 \times \left[U_{DS} + \frac{2}{3}\frac{(U_G+U_{U_i})^{3/2} - (U_G+U_i+U_{DS})^{3/2}}{\sqrt{U_C}}\right]. \quad (13)$$

Thus, the problem of finding the values of S_A and S_R is reduced to finding the change in the potential distribution $a\Phi(x, y)$ in the channel under load (strain).

Let us consider the limiting cases of uniform and nonuniform field distributions in the channel without strain.

1. Uniform field distribution ($\partial U_0/\partial x$ = const). In this case, ($|U_C - U_G| \gg |U_{DS}|, |U_i|$) and $\Phi(x, y)$ takes the form [3]

$$\Phi(x, y) = U_{DS}\frac{2}{\pi^2}\sum_{n=1}^{\infty}\frac{[(-1)^n - 1]}{n^2} \times \sin(\pi n x/L_x)\frac{\sinh[\pi n(2y - L_y/2L_x)]}{\cosh(\pi n L_y/2L_x)}. \quad (14)$$

With (14), one readily calculates a strain-induced change in the longitudinal electric field near the drains:

$$\delta E_x(L_x, y) = -a\frac{\partial \Phi}{\partial x}(x = L_x, y). \quad (15)$$

In Fig. 3, $\delta E_x(L_x, y)$ is plotted against the transverse coordinate. The field change is seen to have opposite signs at $0 \le y \le L_y/2$ and $L_y/2 \le y \le L_y$. The signs of δE_x and sweeping field coincide in the first region and differ in the second. Consequently, if the vicinity of the tensotransistor is elastically strained, the current of one drain will increase, while that of the other decreases.

Substituting $\Phi(x, y)$ into (9) and (12) yields the expressions for S_A and S_R of the tensotransistors.

For transistor 1,

$$S_A = \Pi_{44} C\mu_p R_D(U_C - U_G - U_{DS}) \times U_{DS}\frac{L_y}{L_x}F(L_x, L_y, L_D), \quad (16)$$

where

$$F(L_x, L_y, L_D) = \frac{2}{\pi^2}\frac{L_x}{L_y}\sum_{n=1}^{\infty}\frac{[1-(-1)^n]}{n^2} \times \frac{[\cosh(\pi n L_y/2L_x) - \cosh(\pi n L_y/2L_x)(2L_D/L_y - 1)]}{\cosh(\pi n L_y/2L_x)}. \quad (17)$$

For transistor 2,

$$S_A = \Pi_{44} R_D \sigma_0 L_z\left[1 - \sqrt{(U_G+U_i+U_{DS})/U_C}\right] \times U_{DS}\frac{L_y}{L_x}F(L_x, L_y, L_D). \quad (18)$$

The expression for S_R,

$$S_R = \Pi_{44}F(L_x, L_y, L_D) \quad (19)$$

is the same for both transistors.

Figures 4 and 5 plot the absolute and relative sensitivities of the tensotransistors vs. channel geometry. From these curves, one can obtain the optimum ratio L_y/L_x and the drain width at which S_A and S_R are maximum. The conversion efficiency (relative sensitivity) and the absolute sensitivity are seen to variously depend on L_y/L_x. If the former first grows with L_y/L_x, then peaks, and eventually monotonically drops, the latter first monotonically rises and then saturates at $L_y/L_x \ge 2$. As follows from Figs. 4 and 5, the value of

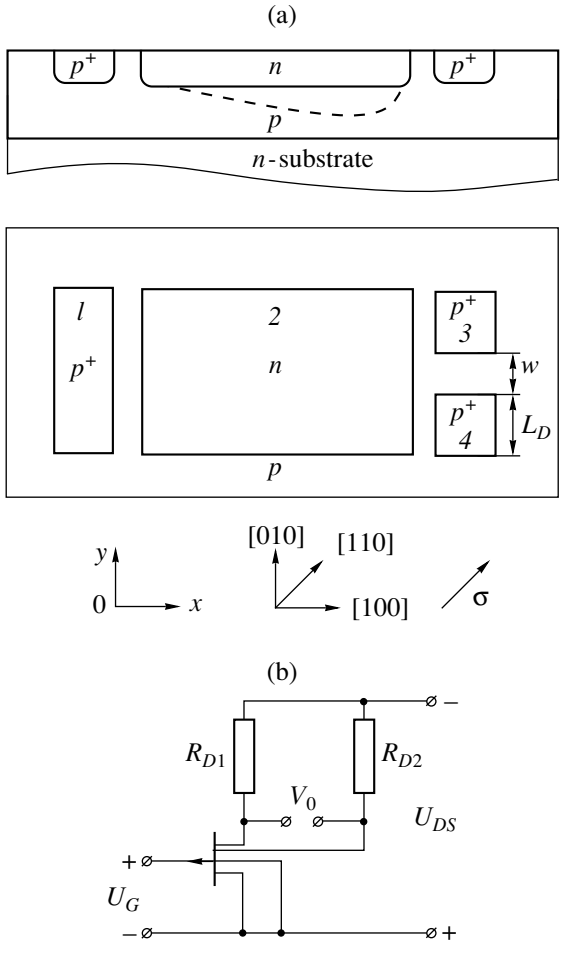


Fig. 2. The same as in Fig. 1 for a junction-controlled field-effect tensotransistor.

$L_y/L_x \approx 1$ is optimum; in this case, the drain spacing $W \leq 0.2L_y$.

Substituting the numerical values of the parameters into (19), we can estimate the conversion efficiency of the tensotransistors:

$$S_R \cong 5.0 \times 10^{-10}\ \text{Pa}^{-1}. \tag{20}$$

2. Nonuniform field distribution ($\partial U_0/\partial x \neq \text{const}$)**.** Without strain, the sweeping field distribution in the channel of transistor 1 is found from the constancy of the total current in any cross section of the channel (see, e.g., [4]):

$$E_x(x) = -\frac{\partial U_0(x)}{\partial x} = -\frac{(U_C - U_G - U_{DS}/2)U_{DS}}{L_x\sqrt{[(U_C - U_G)^2 - 2(U_C - U_G - U_{DS}/2)U_{DS}x/L_x]}}. \tag{21}$$

For arbitrary U_G's and U_{DS}'s, an analytic expression for $\Phi(x, y)$ is impossible to derive; therefore, numerical methods should be invoked [6–8]. According to [7], the elastic-strain-induced longitudinal component of the electric field can be represented in the form

$$\delta E_x(x, y) = -a\frac{\partial \Phi}{\partial x}(x, y) = -\frac{a}{2}\frac{\partial U_0(x)}{\partial x}\frac{y}{(L_x - x_M(L_x, L_y))}F_1(L_x, L_y, L_D), \tag{22}$$

where $x_M(L_x, L_y)$ is the distance from the source to a channel point where Φ is maximum and $F_1(L_x, L_y, L_D)$ is a function depending on the channel geometry and drain width.

Substituting (3), (21), and (22) into (2) in view of the calculated data for $x_M(L_x, L_y)$ and $F_1(L_x, L_y, L_D)$ [7], we obtain the estimation for the maximum absolute sensitivity:

$$S_R \cong 4.2 \times 10^{-10}\ \text{Pa}^{-1}. \tag{23}$$

The numerical analysis [7] shows that the drain and channel geometries that are optimum for the uniform case are likewise optimum for a nonuniform field distribution.

Below is given an estimate of the absolute sensitivity of transistor 1 under matched-load conditions (the output resistance of the device equals the load resistance R_D). Let $L_x = L_y$, $L_D = L_y/2$, $F_1(L_x, L_y, L_D)L_x/(L_x - x_M = 2.4)$, $U_{DS} = 10$ V, and $\Pi_{44} = 1.4 \times 10^{-9}$ Pa^{-1}. Substituting (21) and (22) into (8) yields

$$S_A = 8.4 \times 10^{-9}\ \text{V/Pa}. \tag{24}$$

It is easy to check that S_A and S_R for transistor 1 almost coincide with those for transistor 2 calculated by (11)–(13), all other things being equal.

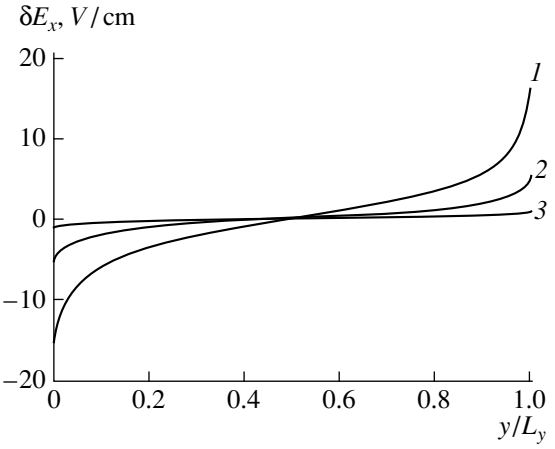


Fig. 3. Distribution of the strain-induced change in the longitudinal electric field across the channel. L_y = (*1*) 10, (*2*) 30, and (*3*) 150 μm. U_{DS} = 1 V, a = 0.01, and L_x = 100 μm.

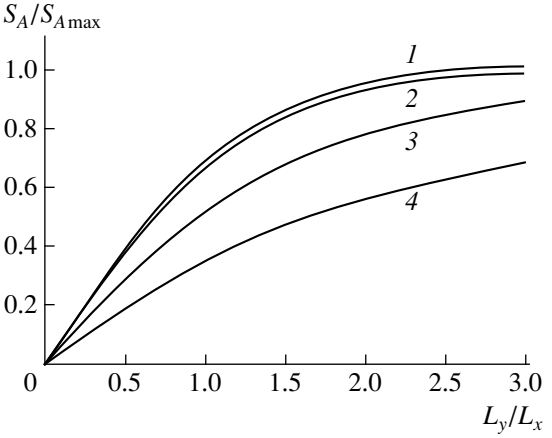


Fig. 4. Relative sensitivity of the tensotransistor vs. channel geometry. L_D/L_y = (*1*) 0.5, (*2*) 0.4, (*3*) 0.2, and (*4*) 0.1.

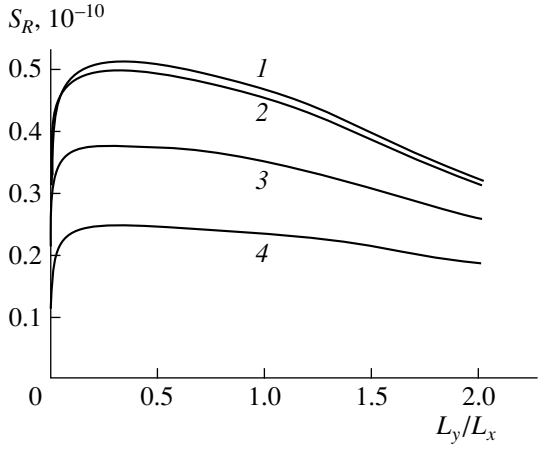


Fig. 5. Normalized-to-maximum absolute sensitivity of the tensotransistor vs. channel geometry. (*1–4*) The same as in Fig. 4.

It is of interest to compare the S_A and S_R values for the tensotransistors with those for a piezoresistive bridge. For the latter [9],

$$S_A = 3.5 \times 10^{-9}\ \text{V/Pa}, \quad S_R = 6.6 \times 10^{-10}\ \text{Pa}^{-1}. \quad (25)$$

As follows from (23)–(25), the values of S_R for the transistors and the bridge are nearly coincident; the absolute sensitivity of the transistors is, however, more than twice that of the bridge. Note, in conclusion, that the absolute sensitivity of field-effect tensotransistors can be further improved at cascade connection, as for magnetic FETs [10].

REFERENCES

1. J. Bryzek, Sens. Actuators A **56**, 1 (1996).
2. G. G. Babichev, V. N. Guz', I. P. Zhad'ko, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **26**, 1244 (1992) [Sov. Phys. Semicond. **26**, 694 (1992)].
3. I. I. Boĭko, I. P. Zhad'ko, S. I. Kozlovskiĭ, and V. A. Romanov, Optoélektron. Poluprovodn. Tekh. **27**, 94 (1994).
4. I. M. Vikulin and V. I. Stafeev, *The Physics of Semiconductors* (Radio i Svyaz', Moscow, 1990).
5. H. P. Baltes and R. S. Popovic, Proc. IEEE **74**, 1107 (1986).
6. P. W. Fry and S. F. Hoey, IEEE Trans. Electron Devices *ED-16*, 35 (1969).
7. J. Lau, P. K. Ko, and P. C. H. Chan, Sens. Actuators A **49**, 155 (1995).
8. N. Mathieu, P. Giordano, and A. Chovet, Sens. Actuators A **32**, 656 (1992).
9. V. N. Guz', I. P. Zhad'ko, *et al.*, Fiz. Tekh. Poluprovodn. (Leningrad) **24**, 409 (1990) [Sov. Phys. Semicond. **24**, 256 (1990)].
10. F. Ning and E. Bruun, Sens. Actuators A **58**, 109 (1997).

Translated by V. Isaakyan

Technical Physics, Vol. 45, No. 10, 2000, pp. 1281–1287. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 50–56.

SOLID-STATE ELECTRONICS

The Effect of GaAs(001) Substrate Roughness on the Magnetic Properties of Epitaxial Fe Films

S. L. Vysotskiĭ, A. S. Dzhumaliev, G. T. Kazakov, Yu. A. Filimonov, and A. Yu. Tsyplin

Institute of Radio Engineering and Electronics, Russian Academy of Sciences (Saratov Branch), Saratov, 410019 Russia
E-mail: fil@sfire.saratov.su

Received October 26, 1999

Abstract—The effect of GaAs(001) surface roughness on the magnetic properties of MBE-grown Fe films having a thickness t in the interval from 12 to 140 Å is investigated by the ferromagnetic resonance method. The films were deposited at room temperature with rates of 9 and 3 Å/min. For films grown on substrates with the rms deviation of the roughness $\sigma \approx 10$ and 30 Å, the spectrum is essentially dependent on the relationship between t and σ. At $t \le \sigma$ and $t \ge 3\sigma$, a single absorption line is observed, whereas at $\sigma \le t \le 3\sigma$, two absorption lines are present. These features of the spectra are related to the island growth of the films and the influence of roughness on island coalescence. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The magnetic properties of thin epitaxial metal films and multilayer structures are of interest both for applications and for basic research on magnetism. The effect of interface roughness on the magnetic properties of the films is among the topical problems [1–13]. Roughness has been shown to influence magnetization reversal processes in the thin films [1–3, 9, 10] and multilayer structures [4, 6–8], interlayer exchange parameter [6] and magnetoresistance [11] of exchange-coupled films, as well as the anisotropy and demagnetizing fields in the thin films [5, 9, 10]. Also, roughness characteristics of "magnetic boundaries" in thin-film structures may noticeably differ from those of the interface [12, 13]. It was noted [1–13] that the effect of roughness on the film properties depends not only on the degree of roughness but also on film growth parameters, such as lattice mismatch between a film and a substrate, deposition rate, substrate temperature, etc. Therefore, the influence of roughness on the magnetic properties needs detailed study in each particular case. The aim of this study is to investigate the properties of magnetic Fe films MBE-grown on GaAs(001) substrates of different roughness.

Of great interest are the magnetic properties of epitaxial Fe films on GaAs substrates [14–27]. This system attracts attention, since the lattice parameters of Fe and GaAs are multiples of each other with an accuracy of no worse than 5%. In addition, GaAs is widely used in planar integrated circuits, and Fe/GaAs structures are candidates for spin-dependent nanoelectronics. The Fe films are applied on GaAs(110) [14–17], (001) [18–26], and (111) [27] wafers. It was noticed in [18] that the quality of Fe films with t = 90–330 Å that are grown by MBE on (001) substrates is, in general, worse than that of films of the same thickness grown on (110) substrates. This is explained by a greater roughness of GaAs(001) substrates. In [26], roughness of the (001)GaAs substrate was assumed to be responsible for two absorption lines in the ferromagnetic resonance (FMR) spectrum of 70- to 100-Å-thick in-plane magnetized Fe films. It was also shown [24, 25] that films of thickness ≥700 Å grown on GaAs(001) substrates are as good as those on the (110) substrates. However, the influence of GaAs(001) roughness on the Fe film properties was not investigated.

In this work, the properties of Fe films on GaAs(001) substrates of different roughness were studied by the FMR method. The films with $t \approx$ 12–140 Å were grown by room-temperature MBE. The absolute, R_z, and rms, σ, values of roughness were used as parameters characterizing the condition of the substrates. They were determined with an atomic force microscope. The GaAs surface finish corresponded to class 14; i.e., σ < 120 Å[28]. The emphasis was on the difference in the magnetic properties of the films with close thicknesses grown on the substrates with $\sigma_1 \approx$ 30 Å and $\sigma_2 \approx$ 10 Å. In addition, we investigated the effect of deposition rate on the properties of magnetic films grown on the substrates with $\sigma_2 \approx$ 10 Å.

EXPERIMENTAL RESULTS

The Fe films were MBE-grown with a setup based on an ES 2301 electron spectrometer. The growth chamber and the analytical chamber, where the elemental composition of the surface was investigated by Auger electron spectroscopy, were evacuated. An arm carried samples from one chamber to the other. Before deposition, the residual pressure in the chamber did not

1063-7842/00/4510-1281$20.00

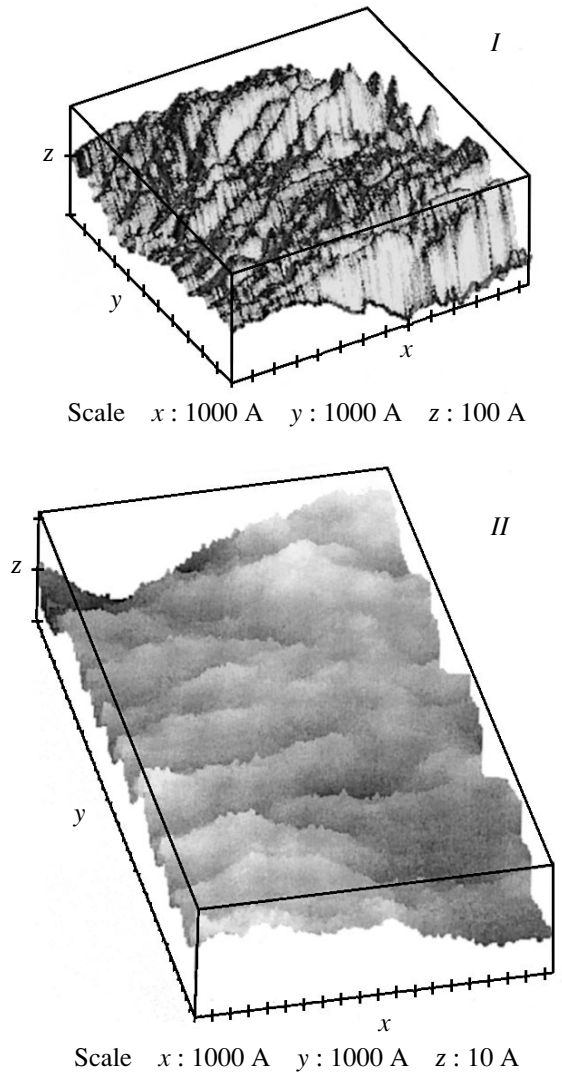


Fig. 1. Surface images of GaAs(001) substrates of type I and II.

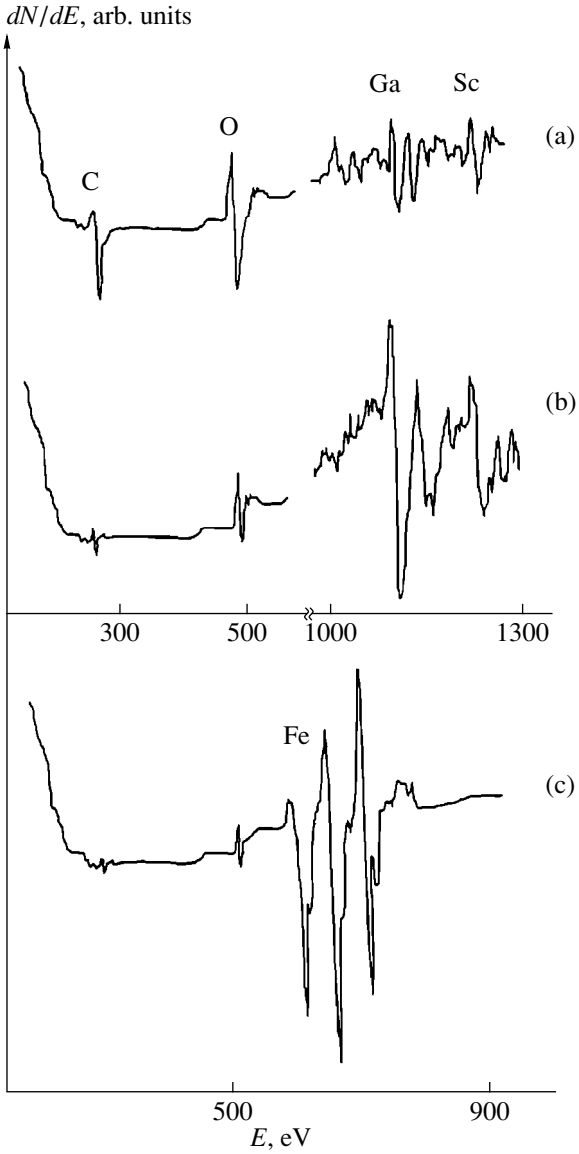


Fig. 2. Auger spectra from the GaAs(001) surface: (a) before cleaning, (b) after cleaning, and (c) after deposition of iron.

exceed 10^{-9} torr, and during the process, ~10^{-8} torr. The Fe films were grown at two rates: $v_1 \approx 9$ Å/min and $v_2 \approx 3$ Å/min. The deposition rate was determined by means of a quartz thickness meter. The substrate temperature was kept close to room temperature.

The GaAs(001) substrates had $\sigma_1 \approx 30$ Å (type-I substrates) and $\sigma_2 \approx 10$ Å (type-II substrates). The roughness was measured in two arbitrarily chosen substrate regions with a SOLVER P4-SPM-MDT atomic force microscope. The rms value of the roughness was calculated as in [29]. The surface relief of type-I and type-II substrates is shown in Fig. 1. One can see that the roughness amplitude R_z for the type-I and type-II substrates is ≤160 and ≤50 Å, respectively.

Prior to deposition, the substrates were chemically cleaned in a 40% KOH solution for 5 min, rinsed in distilled water, treated in a H_2SO_4 (conc.) : H_2O_2 (conc.) : H_2O = 16 : 1 : 1 solution, again rinsed in distilled water, and dried. Just before loading into the growth chamber, the GaAs surface was treated by UV radiation and then annealed for 30 min at ~800 K. As a result, the amount of carbon and oxygen impurities significantly decreased, as followed from the Auger spectra. The surface composition after the depositions indicated the formation of the Fe films (Fig. 2).

The films were investigated by the FMR method at the frequency $f_0 \approx 9.8$ GHz under room temperature. The samples in the resonator were placed in a such way that the magnetic field $H_0 = 0.1$–5 kOe lay in the surface

plane and formed an angle Θ with the [110]-axis (Fig. 3). We recorded the dependence of the resonance field H_r on angle Θ. The accuracy of Θ measurement was 2°–3°.

In our case, the FMR signal was observed only for the films whose thickness t exceeds a certain critical thickness t^*. For the films deposited at the rate v_1, the values of t^* were ≈30 Å and for those grown at the rate v_2, $t^* \approx 10$ Å. Notice that, for $\sigma \approx 10$–30 Å, we did not find an explicit relation between t^* and σ.

The experimental orientation dependence of the resonance field on rotation angle, $H_r(\Theta)$, was compared with the calculated dependence $H_r(\Theta)$. The calculations were performed for cubic ($H_c = K1/M_0$), perpendicular uniaxial ($H_\perp = 2K_\perp/M_0$), and planar uniaxial ($H_\parallel = 2K_\parallel/M_0$) anisotropy fields, where $K1$, $K_\perp$, and $K_\parallel$ are the cubic, perpendicular uniaxial, and planar uniaxial anisotropy constants, respectively. It was assumed that the direction of equilibrium magnetization M_0 forms an angle Ψ with the field H_0, and the axis **l** of planar uniaxial anisotropy forms an angle β with the [110] axis (Fig. 3). The angle Ψ was calculated from the relationship

$$-2H_0\sin\Psi + H_c\sin 4\alpha + H_\parallel\sin 2\eta = 0, \qquad (1)$$

where $\alpha = \Theta - \Psi$ and $\eta = \alpha - \beta$.

Then, the expression for $H_r(\Theta)$ takes the form

$$(H + 4\pi M_0 + H_\perp) \times (H - 3H_c\sin^2 2\alpha - H_\parallel\sin^2\eta) = f_0^2/\gamma^2, \qquad (2)$$

where $H = H_r(\Theta)\cos\Psi + H_c(2 - \sin^2\alpha) + H_\parallel\cos^2\eta$.

Figures 4 and 5 show typical experimental orientation dependences of the resonance field for the Fe films deposited at the rates v_1 and v_2, respectively. The films grown with v_1 were obtained on substrates of both types, while those grown with the smaller rate, only on the type-II substrates. The solid lines depict $H_r(\Theta)$ curves calculated from (1) and (2) for the film parameters listed in the table. In the inserts to Fig. 4, typical dependences of the derivative of the power reflected from the resonator with respect to constant magnetic field, $\frac{dP}{dH_0}(H_0)$, is shown. The distance ΔH between the extrema corresponds to the FMR linewidth, and the arrows point to the position of the resonance field H_r.

First, consider the resonance curves $\frac{dP}{dH_0}(H_0)$ and curves $H_r(\Theta)$ for the Fe films grown on the type-I substrates (Fig. 4) and for films 1–4 (see table). From Fig. 4a, one can see that, at $f_0 = 9.8$ GHz and $H_0 = 0.1$–5 kOe, films 1 and 4 show the single resonance absorption line, which should be assigned to the FMR fundamental mode. In film 1, having the minimal thickness $t_1 \approx 30$ Å, the dependence $H_r(\Theta)$ corresponds to an isotropic film (curve *1*), whereas for film 4 with the thickness $t_4 \approx 140$ Å, the orientation dependence suggests the presence of anisotropy (curve *2*). The dependence $H_r(\Theta)$ for film 4 meets the case when a cubically anisotropic film having the [100] direction of easy magnetization ($K1 > 0$) possesses also uniaxial planar anisotropy with the [110] axis of easy magnetization. The effective magnetizations $4\pi M_{\text{eff}} = 4\pi M_0 + 2K_\perp/M_0$ and anisotropy fields H_c and $H_\parallel$ obtained in view of (1) and (2) are given in the table.

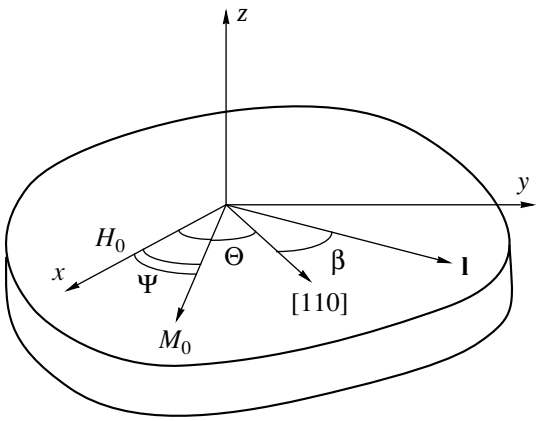


Fig. 3. Coordinate system.

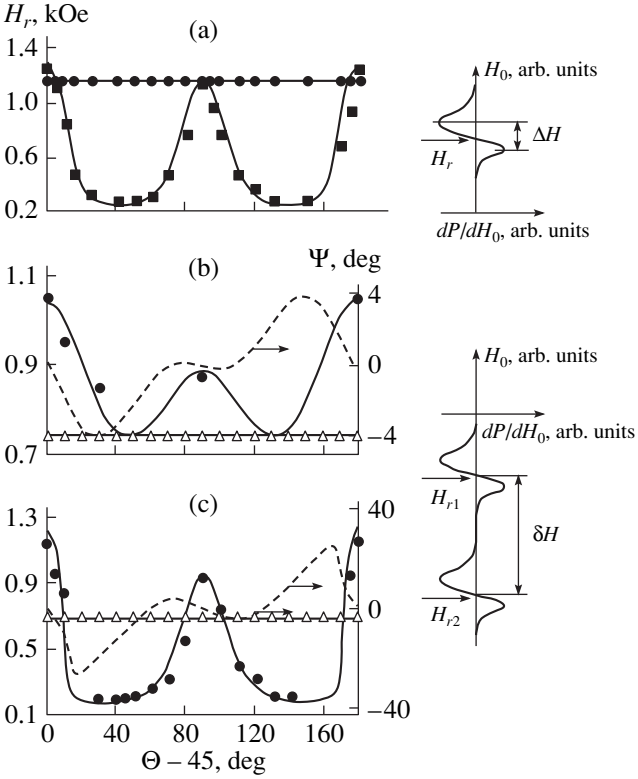


Fig. 4. $H_r(\Theta)$ curves for the films on type-I substrate. (a) (*1*) Film 1 and (*2*) film 4, (b) film 2, and (c) film 3; (b) and (c) $\Psi(\Theta)$ curves for the anisotropic phase of films 2 and 3, respectively (dashed curves).

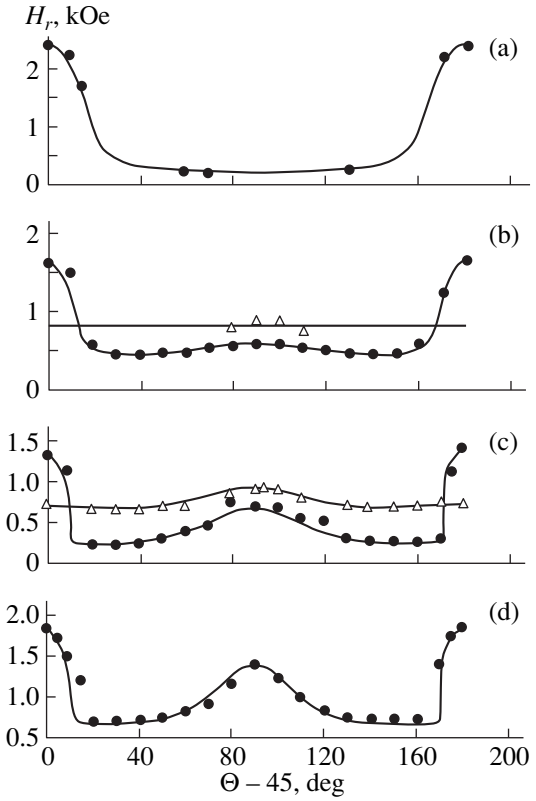


Fig. 5. $H_r(\Theta)$ curves for films (a) 8, (b) 9, (c) 10, and (d) 11 grown on type-II substrate.

The FMR spectra from the Fe films having the intermediate thicknesses $t_2 \approx 70$ Å (Fig. 4b) and $t_3 \approx 100$ Å (Fig. 4c) contain two absorption lines separated by $|\delta H_0| < 1$ kOe (the lower insert in Fig. 4). The dependences $H_r(\Theta)$ for these lines were different and corresponded either to the isotropic case (triangles) or to the case of a cubically anisotropic film with planar uniaxial anisotropy (circles).

Notice that the effective magnetizations for the anisotropic and isotropic responses [obtained with (1) and (2)] in the FMR spectrum are equal for film 2 and differ only slightly for film 3 (see table). Moreover, both the ratio of the isotropic and anisotropic response amplitudes and the interval of angles Θ in which the anisotropic response is observed (Figs. 4b, 4c) change as the film thickness increases.

For the films deposited on type-II substrates with the rate v_1, we failed to find an interval of thicknesses where two absorption lines exist in the FMR spectrum. In film 5 (the thickness $t_5 \approx 30$ Å), the dependence $H_r(\Theta)$ corresponded to the isotropic case, and its shape was similar to that of curve *1* in Fig. 4a. For films 6 and 7 ($t_6 \approx 40$ Å and $t_7 \approx 50$ Å), the dependence corresponds to the anisotropic case, and its shape was similar to curve *2* in Fig. 4a. The parameters of films 5–7, for which the best coincidence of the calculated and experimental dependences $H_r(\Theta)$ was obtained, are given in the table. The Fe films grown on the type-II substrate exhibit, in general, a narrower FMR line and stronger cubic and uniaxial planar anisotropy fields.

Consider now the resonance curves for the Fe films grown on the type-II substrates at the lower deposition rate v_2 (Fig. 5). The single absorption line is observed in the films with the minimal (film 8) (Fig. 5a) and maximal (film 11) (Fig. 5d) thicknesses. In film 8 with $t_8 \approx 12$ Å, the orientation dependence $H_r(\Theta)$ corresponds to the case of uniaxial anisotropy, while film 11 ($t_{11} \approx 50$ Å) exhibits the dependence demonstrating the presence both of uniaxial and of cubic anisotropy.

For films 9 and 10 of the intermediate thicknesses, there are two absorption lines in the FMR spectrum, which are typical of the anisotropic dependences $H_r(\Theta)$ (circles and triangles in Figs. 5b and 5c, respectively). In the film having the thickness $t_9 \approx 21$ Å, the second line (triangles) was observed only at Θ's between 70° and 110°. For this reason, the experimental dependence $H_r(\Theta)$ for the second absorption line was approximated with (1) and (2) for the effective magnetization corresponding to the first absorption line in the FMR spectrum (see table). The anisotropy fields were ignored in this case.

In the film with the thickness $t_{10} \approx 30$ Å, the second absorption line (triangles) was observed throughout the range of Θ. However, in this case, too, when using relationships (1) and (2), we failed to obtain satisfactory agreement between the calculated and experimental (Fig. 5c) dependences $H_r(\Theta)$. It is necessary to stress that the best coincidence between the experimental and calculated resonant fields for the second absorption line is obtained on assumption that the [110]-axis becomes the "hard" magnetic axis. This approximation corresponds to the reversal of sign of the uniaxial anisotropy field in the table. However, even in this case, several data points near the angles $\Theta \approx 0$, 90°, and 180° do not fit the calculated dependence $H_r(\Theta)$.

For absorption lines associated with an anisotropic dependence $H_r(\Theta)$, the maximum FMR linewidths as a function of Θ coincide with the maximal $|\partial H_r/\partial\Theta|$ values. This is demonstrated in Fig. 6 with the orientation dependences $\Delta H(\Theta)$ for the anisotropic absorption line in the FMR spectrum of films 3 and 11, having close-to-record minimal values of ΔH.

DISCUSSION

Note that the parameters $4\pi M_{\text{eff}}$, H_c, and $H_{\|}$ and their thickness dependences for our films are, in general, similar to those for the Fe/GaAs(100) and Fe/GaAs(110) films studied in [14–18]. We recall that the thickness dependence of the film parameters is related to the effect of Néel surface anisotropy. It arises when the symmetry of crystalline fields on the film surface and at the interface and that of magnetoelastic

fields (due to a lattice mismatch between a film and a substrate) break [14–18]. Detailed discussion of the thickness dependence of the film parameters is beyond the scope of this article. Note, however, that the above surface and interfacial effects in the film excite the nonuniform mode for which the resonant field can be represented as [17]

$$H_r = (H_r)_{\text{unif}} + \frac{H_s}{N}, \tag{3}$$

where $(H_r)_{\text{unif}}$ is the resonant field of uniform FMR [see (1) and (2)], H_s is the surface-induced field, and N is the number of monolayers in the film.

Now turn to the table and Figs. 4 and 5 in order to elucidate the effect of roughness on the FMR spectrum shape and film parameters. The anisotropic properties of the films having close thicknesses but grown on substrates with different roughness are seen to considerably diverge. Indeed, the films with t = 50–100 Å that are grown on type-I substrates exhibit both isotropic and anisotropic absorption lines in the FMR spectrum, whereas those grown on type-II substrates have only the anisotropic absorption line at $t \geq 40$ Å.

Notice that the additional absorption line observed in films 2, 3, 9, and 10 cannot be accounted for by the excitation of nonuniform magnetization oscillations, which are characterized by the whole number of half-waves across the film ($n \geq 1$) and in its plane ($k, m \geq 1$). Indeed, in Fe films with the magnetization $4\pi M_0$ = 12 kG, exchange hardness $A = 2 \times 10^{-6}$ erg/cm, and thickness $t \leq 100$ Å, the resonant fields of the principal

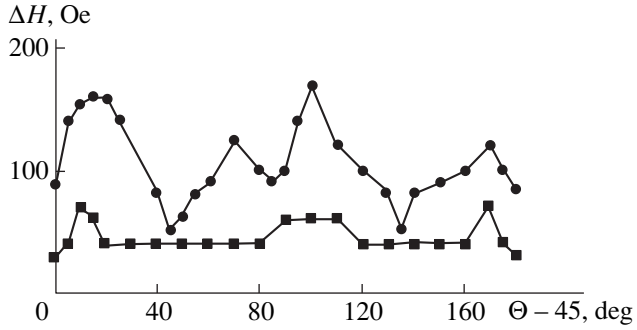


Fig. 6. $\Delta H(\Theta)$ curves for films (●) 3 and (■) 11.

($n = 0$) and first-order ($n = 1$) modes of spin-wave resonance differ by

$$\delta H_0^{(0.1)} = \frac{\pi^2 2A}{t^2 M_0} \geq 40 \text{ kOe}. \tag{4}$$

This far exceeds not only the observed values of δH_0 but also the H_0 range available in the experiment. The difference in the resonant fields of the principal mode and the mode having inhomogeneous in-plane magnetization can be evaluated with the dispersion relation for magnetostatic waves in an in-plane magnetized ferromagnetic film [30]

$$\delta H_0^{(0,1)} \approx \frac{\pi^2 k t M_0}{L(1 + H_0/4\pi M_0)} < 0.1 \text{ Oe}, \tag{5}$$

where L is the length of the film side that fits k half-waves.

Parameters of the films

Film no.	t, Å	σ, Å	$4\pi M_{\text{eff}}$, kG	H_c, Oe	$H_{\parallel}$, Oe	$\Delta H_{\min}$, Oe
1	30	30	8.48	–	–	210
2	70	30	14.2	–	–	150
			14.2	123	–80	100
3	100	30	15.5	–	–	190
			16.6	232	–156	50
4	140	30	14.5	252	–70	150
5	30	10	9.8	–	–	110
6	40	10	12	194	–202	70
7	50	10	16	220	–157	80
8	12	10	8.5	44	–583	90
9	21	10	13.0	152	–267	50
			13.0	–	–	100
10	30	10	13.5	32	45	60
			16.5	202	–172	30
11	45	10	17.5	240	–106	30

The appearance of the two absorption lines in the FMR spectrum of films 2, 3, 9, and 10 can be explained by the formation of nonuniform texture in these films at the chosen parameters of Fe deposition and substrate roughness. It is assumed that film growth starts with the formation of islands, which overlap after a certain thickness t_c has been achieved. In our case, the coalescence of islands must be accompanied by their recrystallization, since the orientation dependences $H_r(\Theta)$ of the absorption lines in the FMR spectrum for films 2, 3, 9, and 10 are different. It is evident that the roughness, along with the other growth parameters, must influence the value of t_c and the island size d. If it is supposed that a two-line FMR spectrum corresponds to the beginning of island coalescence and an FMR spectrum characteristic of films with pronounced cubic anisotropy ($K_c > K_\parallel$) correlates with the formation of the continuous film, then, for the chosen range of roughness and growth conditions, we obtain that $t_c \sim \sigma$ and that the continuous film is formed at $t \approx 3\sigma$.

It is clear that the FMR signal from discontinuous films is averaged over the islands. The resonance field H_r for each of the islands will be determined not only by the fields $(H_r)_{\text{unif}}(H_0, H_c, H_\parallel, H_\perp)$ and H_s, involved in (3), but also by the demagnetizing fields H_d. If $d = 100$–1000 Å [31], for films of critical thickness t^* and $4\pi M \approx 8$ kG, we obtain $H_d \sim 4\pi M t^*/d \approx 0.1$–$2$ kOe. Such values of H_d, being comparable to the fields H_0, H_c, and $H_\parallel$, can make, along with H_s, a noticeable contribution to the resonance field H_r [see (3)] of a single island. If the islands noticeably differ in sizes, one should expect a significant attenuation of the FMR signal from the film. That is why, in our opinion, the FMR signal is absent in the films of subcritical thickness $t < t^*$. It is known that the grain size d grows with decreasing deposition rate v [31]; this may explain the change in the critical thickness from $t_1^* \approx 30$ Å at $v_1 \approx 9$ Å/min to $t_2^* \approx 10$ Å at $v_2 \approx 3$ Å/min.

Inverse proportion between the island size and deposition rate may be responsible for the difference in the orientation dependences $H_r(\Theta)$ for films 1 and 5, on the one hand, and film 8, on the other. In the films grown at the higher deposition rate v_1, the size of the islands d seems to be so small that the contribution of the fields H_d and H_s to H_r becomes dominant and the averaged response from the films becomes Θ-independent due to a spread in d. For film 8, grown at the lower deposition rate v_2, the effect of the above factors decreases because of increased d's; hence, the contribution of the fields H_c and $H_\parallel$, characterizing epitaxial growth of the film, becomes noticeable.

Island growth can explain both a change in the amplitude ratio of the isotropic and anisotropic responses and the extension of the Θ interval where the anisotropic response is observed as the film thickens (Figs. 4b, 4c). The variation of the anisotropic FMR linewidth with Θ (Fig. 6) can also be attributed to island growth. The last two effects can be explained by the influence of exchange interaction at grain boundaries when the equilibrium magnetizations M_0 inside the grains are misaligned. If, for example, the isotropic and anisotropic responses in the FMR spectrum of films 2 and 3 are associated with the presence of the isotropic and anisotropic phases, the misalignment of the magnetizations in the phases will be determined by an angle Ψ. For the parameters corresponding to the anisotropic response in films 2 and 3, we obtain $|\Psi| \leq 5°$ and $|\Psi| \leq 25°$, respectively (dashed curve in Figs. 4b, 4c). It is seen that there is only the isotropic response from films 2 and 3 for angles Θ at which $|\Psi|$ is maximum. At $\Psi \longrightarrow 0$, to which corresponds the magnetization along the $\langle 100 \rangle$ and $\langle 110 \rangle$ directions ($\Theta = \pi N/4$, where N is an integer), both isotropic and anisotropic responses are observed. It is also clear that, in this case, the contribution to ΔH due to the nonuniformity of magnetization orientation is absent and the width of the resonance curve is minimal (Fig. 6).

Exchange interaction at island boundaries can be responsible for the difficulties in selecting the parameters H_c and $H_\parallel$ [(1) and (2)] to fit the orientation dependences $H_r(\Theta)$ of the additional absorption lines for films 9 and 10. In fact, the occurrence of exchange interaction between the islands with degenerate FMR spectra must remove degeneration. Then, the FMR spectra must move apart as it takes place in exchange-coupled films [32]. In this case, the resonant dependences $H_r(\Theta)$ do not intersect (Figs. 5b, 5c).

CONCLUSION

We investigated the influence of GaAs(001) substrate roughness on the FMR spectra taken at the frequency 9.8 GHz from Fe films having the thicknesses $t \approx 12$–140 Å. The films were grown by room-temperature MBE with the deposition rates $v_1 \approx 9$ Å/min and $v_2 \approx 3$ Å/min. For the substrates with the rms deviations of the roughness amplitude $\sigma_1 \approx 30$ Å and $\sigma_2 \approx 10$ Å and for the chosen growth parameters, the results are the following.

(1) The spectrum shape is essentially defined by the relationship between t and σ. At $t \leq \sigma$ and $t \geq 3\sigma$, the single absorption line is observed, whereas at $\sigma \leq t \leq 3\sigma$, the spectrum contains two lines.

(2) The FMR spectrum is observed in the films whose thickness t is greater than the critical thickness t^*. The latter is determined largely by the deposition rate and does not explicitly depend on the substrate roughness.

These features of the FMR spectra arise from the island mechanism of film growth and the effect of roughness on island coalescence.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant no. 97-02-18614), State Program "Topical Problems of Condensed-Matter Physics" (Subprogram "Surface Atomic Structures," grant nos. 97-3.16 and 2.9.99), and partially by the Program "Integration" (grant no. 696.3).

REFERENCES

1. R. J. Prosen, B. E. Gran, and J. Kivel, J. Appl. Phys. **34**, 1147 (1963).
2. M. Li, Y.-P. Zhao, G.-C. Wang, and H.-G. Min, J. Appl. Phys. **83**, 6287 (1998).
3. Y.-L. He and G. C. Wang, J. Appl. Phys. **76**, 6446 (1994).
4. X. Meng, X. Bian, W. B. Muir, *et al.*, J. Appl. Phys. **76**, 7084 (1994).
5. J. F. Cochran, W. B. Muir, J. M. Rudd, *et al.*, J. Appl. Phys. **69**, 5206 (1991).
6. D.-H. Han, J.-G. Zhu, J. H. Judy, and J. M. Sivertsen, J. Appl. Phys. **81**, 340 (1997).
7. C.-H. Chang and M. H. Kryder, J. Appl. Phys. **75**, 6864 (1994).
8. H. Takeshita, K. Hittori, Y. Fujiwara, *et al.*, J. Appl. Phys. **75**, 6415 (1994).
9. P. Bruno, G. Bayreuther, P. Beauvillain, *et al.*, J. Appl. Phys. **68**, 5759 (1990).
10. C. Chappert and P. Bruno, J. Appl. Phys. **64**, 5736 (1988).
11. W. Folkerts and F. Hakkens, J. Appl. Phys. **73**, 3922 (1993).
12. J. W. Freeland, V. Chakarian, K. Bussmann, *et al.*, J. Appl. Phys. **83**, 6290 (1998).
13. J. F. MacKay, C. Teichert, and M. G. Lagally, J. Appl. Phys. **81**, 4353 (1997).
14. G. A. Prinz and J. J. Krebs, Appl. Phys. Lett. **39** (5), 397 (1981).
15. F. J. Rachford, G. A. Prinz, J. J. Krebs, and K. B. Hathaway, J. Appl. Phys. **53**, 7966 (1982).
16. J. J. Krebs, F. J. Rachford, P. Lubitz, and G. A. Prinz, J. Appl. Phys. **53**, 8058 (1982).
17. G. A. Prinz, G. T. Rado, and J. J. Krebs, J. Appl. Phys. **53**, 2087 (1982).
18. J. J. Krebs, B. T. Jonker, and G. A. Prinz, J. Appl. Phys. **61**, 2596 (1987).
19. E. Gu, J. A. C. Bland, C. Daboo, *et al.*, J. Appl. Phys. **76**, 6440 (1994); Phys. Rev. B: Condens. Matter **51**, 3596 (1995).
20. C. Daboo, R. J. Hicken, E. Gu, *et al.*, Phys. Rev. B: Condens. Matter **51**, 15964 (1995).
21. A. Filipe and A. Schuhl, J. Appl. Phys. **81**, 4359 (1997).
22. R. W. Tustison, T. Varitimos, J. van Hook, and E. F. Schloemann, Appl. Phys. Lett. **51** (4), 285 (1987).
23. S. A. Oliver, C. Vittoria, E. Schloemann, *et al.*, J. Appl. Phys. **63**, 3802 (1988).
24. B. T. Jonker, E. M. Kneedler, P. Thibado, *et al.*, J. Appl. Phys. **81**, 4362 (1997).
25. J. M. Florczak and E. Dan Danlberg, Phys. Rev. B: Condens. Matter **44**, 9338 (1991).
26. S. L. Vysotskiĭ, S. S. Gul'bukh, A. S. Dzhumaliev, *et al.*, Pis'ma Zh. Tekh. Fiz. **25** (3), 36 (1999) [Tech. Phys. Lett. **25**, 79 (1999)].
27. Y. Endo, S. Okamoto, O. Kitakami, and Y. Shimada, J. Appl. Phys. **81**, 344 (1997).
28. Yu. M. Yakovlev and S. Sh. Gendelev, *Single Crystal Ferrites in Radio Electronics* (Sov. Radio, Moscow, 1975).
29. Y. Poon Chin and B. Bhushan, J. Appl. Phys. **79**, 5799 (1996).
30. A. G. Gurevich and G. A. Melkov, *Magnetic Oscillations and Waves* (Fizmatlit, Moscow, 1994).
31. *Thin Film Technology* (Sov. Radio, Moscow, 1977), Vol. 2, p. 768.
32. Yu. A. Filimonov, G. T. Kazakov, S. L. Vysotsky, *et al.*, J. Magn. Magn. Mater. **131**, 235 (1994).

Translated by Yu. Vishnyakov

Technical Physics, Vol. 45, No. 10, 2000, pp. 1288–1290. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 57–59.

OPTICS, QUANTUM ELECTRONICS

Transformation of an Optical Pulse in a Periodically Nonuniform Fiber with Amplification or Absorption

I. O. Zolotovskiĭ and D. I. Sementsov

Ul'yanovsk State University, Ul'yanovsk, 432700 Russia

E-mail: sements@quant.univ.simbirsk.ru

Received October 8, 1999; in final form, January 20, 2000

Abstract—Two-mode optical fibers in which the refractive index varies periodically along the axis are considered. The effect of the imaginary part of a complex refractive index on the transformation and dispersion of a propagating optical pulse is studied theoretically. It is demonstrated that the pulse can be compressed with no phase modulation at the fiber input. *© 2000 MAIK "Nauka/Interperiodica".*

There has recently been considerable interest in fibers in which the refractive index varies periodically along the axis [1–4]. This feature is viewed as an efficient means of controlling propagating laser radiation. For a real-valued refractive index, analysis of the linear and nonlinear regimes of optical-pulse transformation and compression has shown that these fibers possess unique dispersive properties [5, 6]. Consequently, they may serve as compact pulse compressors for integrated optics. In contrast, this study addresses the effect of the imaginary part of the complex refractive index on the dispersive properties of a periodic fiber and on the transformation of propagating optical pulses. In short, we investigate the effect of amplification or absorption in the fiber.

1. We will consider a two-mode optical fiber in which the permittivity is a complex quantity and is distributed as

$$\varepsilon(r, z) = \varepsilon_0\{1 - f(r)[1 + \gamma\cos(2\pi z/\Lambda)]\}. \quad (1)$$

Here, $\varepsilon_0 = \varepsilon_0' + i\varepsilon_0''$ is the permittivity on the fiber axis. For real fibers, $|\varepsilon''| \ll |\varepsilon'|$. The function $f(r)$ describes the optical nonuniformity in the transverse direction, whereas $\gamma \ll 1$ and Λ respectively evaluate the modulation depth and the period of the permittivity in the longitudinal direction. Since it is small, the term $\gamma\varepsilon_0''$ will be neglected: we assume that the imaginary part of the permittivity is uniform along the fiber axis. The propagation constant of the complex permittivity is $\beta_j = \beta_j' - i\beta_j''$, where $|\beta_j''| \ll |\beta_j'|$. Let us represent the field in the fiber as the superposition of the eigenmodes in a fiber that is uniform in the longitudinal direction:

$$\mathbf{E}(t, r, z) = \frac{1}{2}\sum_j [\mathbf{e}_j C_j(t, z) R_j(r) \exp i(\omega_0 t - \beta_j' z) + c.c.], \quad (2)$$

where $\mathbf{e}_j$ are the unit polarization vectors of the modes, $R_j(r)$ describe the distributions of corresponding modes over the fiber cross section, and ω_0 is the carrier frequency of a wave packet launched into the fiber.

Since the modal propagation constants are complex quantities, the time envelopes of the modal amplitudes have the form

$$C_j(t, z) = A_j(t, z)\exp(\beta_j'' z). \quad (3)$$

Intermodal coupling is pronounced if modal phase matching occurs at the carrier frequency. Since the permittivity distribution is periodic and the propagation constants are complex quantities, the phase-matching conditions are as follows:

$$\delta'(\omega_0) = 0, \quad \delta''(\omega_0) \cong 0,$$
$$\delta(\omega) = \beta_1(\omega) - \beta_2(\omega) - 2\pi/\Lambda. \quad (4)$$

For parameter values that approximately satisfy the phase-matching conditions ($\omega \cong \omega_0$), the coupled-wave equations for modal time envelopes of the pulse have the form

$$\frac{\partial A_1}{\partial z} - \frac{1}{v}\frac{\partial A_1}{\partial \tau} - i\frac{d_1}{2}\frac{\partial^2 A_1}{\partial \tau^2} = -i\sigma_{12}A_2\exp(i\delta z),$$
$$\frac{\partial A_2}{\partial z} + \frac{1}{v}\frac{\partial A_2}{\partial \tau} - i\frac{d_2}{2}\frac{\partial^2 A_2}{\partial \tau^2} = -\sigma_{21}^* A_1\exp(-i\delta z). \quad (5)$$

Here, $\tau = t = z/u$ with $2u = u_1 + u_2$, where $u_j \cong (\partial\beta_j'/\partial\omega)_0^{-1}$ are the modal group velocities. Furthermore, $1/v$ and d_j are complex parameters expressed as follows:

$$1/v \cong \left(\frac{u_1 - u_2}{u_1 + u_2}\right)(\partial\beta/\partial\omega)_0,$$

and $d_j = (\partial^2\beta_j/\partial\omega^2)_0$, where $2\beta = \beta_1 + \beta_2$. The intermodal-coupling coefficients σ_{ij} are expressed as

$$\sigma_{ij} \cong (k_0^2\varepsilon_0\gamma\int \mathbf{e}_i\mathbf{e}_j f(r)R_iR_j r dr)(2\beta_i\int R_i^2 r dr)^{-1}, \quad (6)$$

where $i, j = 1, 2$ with $i \neq j$; $k_0 = \omega/c$; and ω and c are the frequency and the speed of light *in vacuo*, respectively.

2. At the input of the fiber, the time envelopes of the pulse modal amplitudes are determined by the type of launching. Let us represent the amplitudes as $A_j(\tau, 0) = A_{j0}\phi(\tau)$. In most cases, single- or two-mode launching is used, whereby either $A_{10} \neq 0$ and $A_{20} = 0$ (or vice versa) or $A_{20} = \psi A_{10}$. If $\psi = \pm 1$, launching is symmetric or antisymmetric, respectively. With a Gaussian pulse,

$$\phi(\tau) = \exp(-\tau^2/2\tau_0^2), \quad (7)$$

where τ_0 is the pulse length at the input.

Let the intermodal coupling be strong so that, with a high degree of accuracy, $|A_1|^2 + |A_2|^2 = \text{const}$ over the intermodal-beat length $L_b = 1/|\sigma|$ and, therefore, $|\sigma_{12}| \cong |\sigma_{21}^*| \equiv |\sigma|$. System (5) then has the solution

$$\begin{aligned} A_1 &= a_1(\tau, z)\exp[(iq - \delta/2)z] \\ &+ a_2(\tau, z)\exp[(-iq - \delta/2)z], \\ A_2 &= \chi a_1(\tau, z)\exp[(iq + \delta/2)z] \\ &- \chi^{-1}a_2(\tau, z)\exp[(-iq + \delta/2)z], \end{aligned} \quad (8)$$

where the parameters a_1 and a_2 are slowly varying functions of z; they are determined by the launching conditions. Furthermore,

$$\chi = \frac{(2q + \delta)A_{20} - 2\sigma A_{10}}{(2q - \delta)A_{10} - 2\sigma A_{20}}, \quad q \equiv (\sigma^2 + \delta^2/4)^{1/2}. \quad (9)$$

Thus, since it is generated by the interacting modes, the pulse is the superposition of two component pulses. According to (5) and (8), the amplitudes of the component pulses obey the equations

$$\frac{\partial a_f}{\partial z} - \frac{(-1)^f\delta}{2qv}\frac{\partial a_f}{\partial\tau} - \frac{iD_f}{2}\frac{\partial^2 a_f}{\partial\tau^2} = 0; \quad f = 1, 2, \quad (10)$$

where

$$D_f = d + \frac{(-1)^f}{2vq}(1 - \delta p v^2) \quad (11)$$

is the complex effective dispersion of the fth component pulse and $p = (d_2 - d_1)/2$. In the input conditions $a_f(\tau, 0) = a_{f0}\phi(\tau)$, we have

$$a_{f0} = \frac{1}{2}\left[A_{10} + (-1)^f\left(\frac{\delta}{2q}A_{10} + \frac{\sigma}{q}A_{20}\right)\right] \quad (12)$$

due to (9).

With the above conditions at the input, the solution to system (10) has the form

$$\begin{aligned} a_f(\tau; z) &= a_{f0}(1 + izD_f/\tau_0^2)^{-1/2} \\ &\times \exp\left(-\frac{\tau_f^2}{2\tau_0^2(1 + izD_f/\tau_0^2)^2}\right), \end{aligned} \quad (13)$$

where $\tau_f = \tau + (-1)^f\delta z/2qv$.

If phase matching is perfect ($\delta = 0$ and $\chi = -1$), then

$$D_f = D_f' - iD_f'' = d' + \frac{(-1)^f}{v'^2|\sigma|} - i\left(d'' + \frac{2(-1)^f}{v'v''|\sigma|}\right). \quad (14)$$

We now recast solution (13) as follows:

$$\begin{aligned} a_f &= \frac{a_{f0}}{\sqrt{\xi_f}}\exp\left(\frac{1}{2}\left(i\theta_f - \frac{\tau^2}{\tau_0^2\xi_f^2}\right)\right), \\ \xi_f &= \left(\frac{(1 + b_f'')^2 + b_f'^2}{(1 + b_f'')}\right)^{1/2}, \\ \theta_f &= \frac{b_f'\tau^2/\tau_0^2}{(1 + b_f'')^2 + b_f''^2} - \arctan\left[\frac{b_f'}{(1 + b_f'')}\right], \end{aligned} \quad (15)$$

where $b_f \equiv zD_f/\tau_0^2$ and

$$a_{f0} = \frac{1}{2}[A_{10} + (-1)^f A_{20}].$$

Two-mode launching ($\psi = \pm 1$) is of practical value. In this case, the entire pulse can be represented by only one of the component pulses and the modal intensities $I_j = |C_j|^2$ are equal:

$$I_1 = I_2 = \frac{I_0}{2\xi_f}\exp\left(-\beta_f''z - \frac{\tau^2}{\tau_0^2\xi_f^2}\right). \quad (16)$$

The above expressions imply that pulse compression is possible if $D_f'' < 0$. Furthermore, once a pulse has traveled a distance $z = L_{sf}$, where

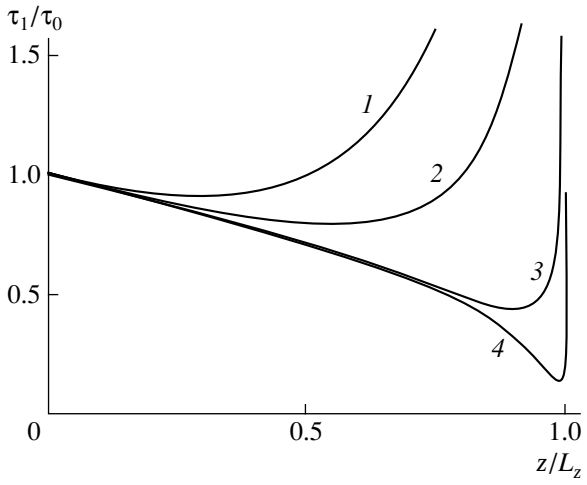
$$L_{sf} = L_z\left(1 - \frac{1}{\sqrt{1 + \eta_f^2}}\right), \quad L_z = \frac{\tau_0^2}{|D_f''|}, \quad (17)$$

the pulse length τ_1, generally expressed as $\tau_1 = \tau_0\xi_f$, takes on the minimum value

$$\tau_{\min} = \frac{\sqrt{2}\tau_0}{\eta_f}(\sqrt{1 + \eta_f^2} - 1)^{1/2}, \quad (18)$$

where $\eta_f = |D_f''/D_f'|$.

If $\eta_f \gg 1$, then $\tau_{\min} \cong \tau_0\sqrt{2/\eta_f}$. To be specific, consider absorbing quartz fibers. It has been established for



Pulse length vs. distance traveled for η_f = (*1*) 1, (*2*) 2, (*3*) 10, and (*4*) 100.

such fibers that the imaginary part of the refractive index has an insignificant effect on pulse transformation and dispersion. For example, perfect fibers have an absorption coefficient $\alpha = 2\beta''$ of about 10^{-4} m^{-1}, so that there is only a slight decrease in intensity after a pulse has traveled a relatively long distance [see (16)]. Typically, $|D''| \simeq \partial^2\beta''/\partial\omega^2)_0 \simeq \beta''/\omega_0^2 \simeq 10^{-35}$–$10^{-33}$ s^2/m, whereas $|D'| \simeq 10^{-26}$ s^2/m. Such a small $|D''|$ indicates that absorption is unimportant in terms of pulse dynamics. Therefore, we will focus on the dispersive properties of an *amplifying* fiber. The fiber is assumed to be fabricated from neodymium-doped glass [7, 8].

Let the pulse be made up of the LP_{01} and LP_{02} modes of the fiber core. Then the condition $\beta_1'' = \beta_2'' = \beta''$ is fulfilled with a high degree of accuracy. For such active fibers, the quantity $2\beta''(\omega)$ equals the gain per unit length and can be expressed as

$$2\beta''(\omega) = \rho N\left[1 + \frac{I_0}{I_n} + \left(\frac{\omega_0 - \omega_r}{\Delta\omega_l}\right)^2\right]^{-1}, \qquad (19)$$

where ω_r and ρ are, respectively, the frequency and the cross section of the stimulated transition; N is the concentration of active particles in the absence of lasing; $\Delta\omega_l$ is the spectral linewidth; and I_n is the saturation intensity [9].

In amplifying fibers, D'' is very close to the imaginary part d'' of the material dispersion, because $|D''| \geq |D'|$ in the cases of interest, where $|d''| \gg 2/|v'v''\sigma|$. Consequently, $D'' \cong (\partial^2\beta''/\partial\omega^2)_0$. We thus arrive at the expression

$$D'' = \frac{\rho N}{\Delta\omega_l^2}\frac{1 + I_0/I_n - 3(\Delta\omega/\Delta\omega_l)^2}{(1 + I_0/I_n + (\Delta\omega/\Delta\omega_l)^2)^3}, \qquad (20)$$

where $\Delta\omega = \omega - \omega_r$.

We will assume that $I_n \gg I_0$. According to (15), radiation compression in the fiber is possible only if $3\Delta\omega^2 > \Delta\omega_l^2$. This means that the carrier frequency must deviate somewhat from the stimulated-transition frequency. Otherwise, $D_f'' > 0$, indicating that uniform pulse stretching occurs. Let us consider the conditions of pulse compression, for which $D'' < 0$. For example, if $\Delta\omega_0 \cong \Delta\omega_l$, $\omega_0 \cong 1.8 \times 10^{15}$ s^{-1}, $\Delta\omega_l = 10^{12}$ s^{-1}, and $\rho N \cong 2$ m^{-1} [8], then one has $d'' \cong -0.8 \times 10^{-24}$ s^2/m. By choosing suitable fiber characteristics (the fiber diameter and the refractive-index profiles of the core and the cladding), one may obtain $D_f' \leq 10^{-26}$ s^2/m and $\eta_f \cong 100$ [5, 6].

The figure shows the normalized pulse length as a function of the normalized distance z/L_z for η_f = 1, 2, 10, or 100 (curves *1–4*, respectively). It is seen that the pulse length approaches infinity as z tends to L_z. Generally speaking, if $z \cong L_z$, one has to allow for dispersion terms of order higher than 2 so as to avoid unlimited pulse spreading at a finite z. Note that amplifying optical fibers serving as compressors must be operated at carrier frequencies differing from the resonance frequency of the active medium. This makes amplification less efficient, although the power of the pulse may be raised by means of pulse compression.

To sum up, in the context of amplifying optical fibers, the effects corresponding to complex values of the refractive index seem to be important for a detailed analysis of lasing dynamics when there is strong linear intermodal coupling in the pulse. In this connection, we highlight the possibilities for optical-pulse compression with no frequency modulation at the input. Consequently, two-mode amplifying fibers show potential for use as compact compressors of laser radiation in integrated optics.

REFERENCES

1. A. Galvanauskas, M. E. Fermann, and J. E. Roman, Appl. Phys. Lett. **66**, 1053 (1995).
2. S. A. Vasil'ev, E. M. Dianov, A. S. Kurkov, *et al.*, Kvantovaya Élektron. (Moscow) **24** (2), 151 (1997).
3. S. A. Vasil'ev, E. M. Dianov, D. S. Starodubov, *et al.*, Kvantovaya Élektron. (Moscow) **24** (2), 160 (1997).
4. E. M. Dianov, V. I. Karpov, A. S. Kurkov, *et al.*, Tech. Dig. Ser.-Opt. Soc. Am. **22** (14), 77 (1995).
5. I. O. Zolotovskiĭ and D. I. Sementsov, Opt. Spektrosk. **85**, 304 (1998) [Opt. Spectrosc. **85**, 281 (1998)].
6. I. O. Zolotovskiĭ and D. I. Sementsov, Opt. Spektrosk. **86**, 824 (1999) [Opt. Spectrosc. **86**, 737 (1999)].
7. J. H. Povlsen, A. Bjarklev, T. Rasmussen, *et al.*, Electron. Lett. **28**, 1807 (1992).
8. I. N. Duling, L. Goldberg, and J. F. Weller, Electron. Lett. **24**, 1334 (1988).
9. *Handbook on Lasers*, Ed. by A. M. Prokhorov (Sov. Radio, Moscow, 1978).

Translated by A. Sharshakov

Technical Physics, Vol. 45, No. 10, 2000, pp. 1291–1295. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 60–64.

A Photodissociation High-Power Iodine Laser with Improved Capacity to Break Down Solid Targets

Yu. M. Ivanenko, V. E. Sabinin, and G. A. Skorobogatov

St. Petersburg State University, Universitetskaya nab. 7/9, St. Petersburg, 198902 Russia

E-mail: gera@mail.moss.pu.ru

Received March 25, 1999; in final form, November 25, 1999

Abstract—A single-pulse high-power photodissociation iodine laser (PIL) was developed. It is pumped by the radiation of an open electric discharge (31.3 kJ, 20 kA). The total output energy at a wavelength of 1316 nm was found to be (30 ± 5) J for ~100-μs-wide pulses. To improve the breakdown capacity, each pulse was modulated with a train of five to six narrower pulses by applying a variable magnetic field from the discharge pinch to the laser medium. *© 2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

To date, single-pulse lasers with a pulse energy as high as many kJ have been implemented [1–5]. For applications such as laser-induced fusion reactions [6] or laser processing of solids [7, 8], it is essential that the output energy be absorbed to the greatest possible extent. Provided that the heat capacity C_s of a solid target is constant, the maximum temperature increment in the vicinity of a laser spot is given by

$$\Delta T_{\max} = \frac{\alpha}{C_s m} W_\Sigma, \quad (1)$$

where α is the absorption coefficient of the target surface and m is the mass of a heated part of the target. The total energy of a laser pulse is

$$W_\Sigma = \int_0^\tau P(t)dt, \quad (2)$$

where $P(t)$ is the beam power and τ is the pulse width.

From (1), the upper limit of molten mass within a laser spot is found from the formula

$$m_{\text{melt}} = \frac{\alpha}{C_s(T_m - 300\ \text{K}) + q_m} W_\Sigma, \quad (3)$$

where q_m is the heat of melting and T_m is the melting point of a target material.

In experiments on laser breakdown, the upper limit of evaporated material is given by

$$m_{\text{ev}} = \frac{\alpha W_\Sigma}{C_s(T_m - 300\ \text{K}) + q_m + C_l(T_b - T_m) + q_{\text{ev}}}, \quad (4)$$

where C_l is the heat capacity of the liquid (molten) target, q_{ev} is the evaporation heat, and T_b is the boiling point.

From (3) and (4), the breakdown depth is a linear function of α for any processing technique. Direct measurements [9], however, showed that a laser plasma due to target sputtering has a very high reflectivity, so that α does not exceed 0.01 in the visible spectrum range.

The value of α can be increased by decreasing τ, since only the leading edge of a radiation pulse is responsible for target damage. In addition, it was established that solid targets are eroded most rapidly when a train of short (to 5 ms) pulses (instead of one long) is applied. In this work, we used both approaches to enhance the erosive action of a single-pulse PIL on solid targets.

PHOTODISSOCIATION IODINE LASERS PUMPED BY AN OPEN ELECTRIC DISCHARGE

PILs employ the lasing transition

$$I(^2P_{1/2}) + h\nu(1316\ \text{nm}) \xrightarrow{\sigma_5} I(^2P_{3/2}) + 2h\nu. \quad (5)$$

In early PILs [10], the pumping of the $^2P_{1/2}$ level was accomplished by the photodissociation of perfluoroalkyloids due to pulse xenon lamps:

$$RI + h\nu(250\text{–}300\ \text{nm}) \xrightarrow{\sigma_6} \begin{cases} \xrightarrow{p_1} R + I(^2P_{1/2}), \\ \xrightarrow{p_2} R + I(^2P_{3/2}), \end{cases} \quad (6)$$

where $p_1 = 0.8$ (for $R = CF_3$), 0.95 (C_2F_5), 1.0 (C_3F_7), or 0.9 (iC_3F_7) [11].

In [12], a PIL was pumped with the radiation of an open electric discharge from an exploding conductor. The advantage of such pumping is the absence of reflecting (absorbing) surfaces between the glowing plasma of a pump source and the PIL lasing cavity. However, this technique makes it impossible to control the pumping duration and space–time characteristics of the pumping plasma. In this case, extra optical devices like an exciter–amplifier are needed to cut τ by one or two orders of magnitude. This inevitably results in

1063-7842/00/4510-1291$20.00

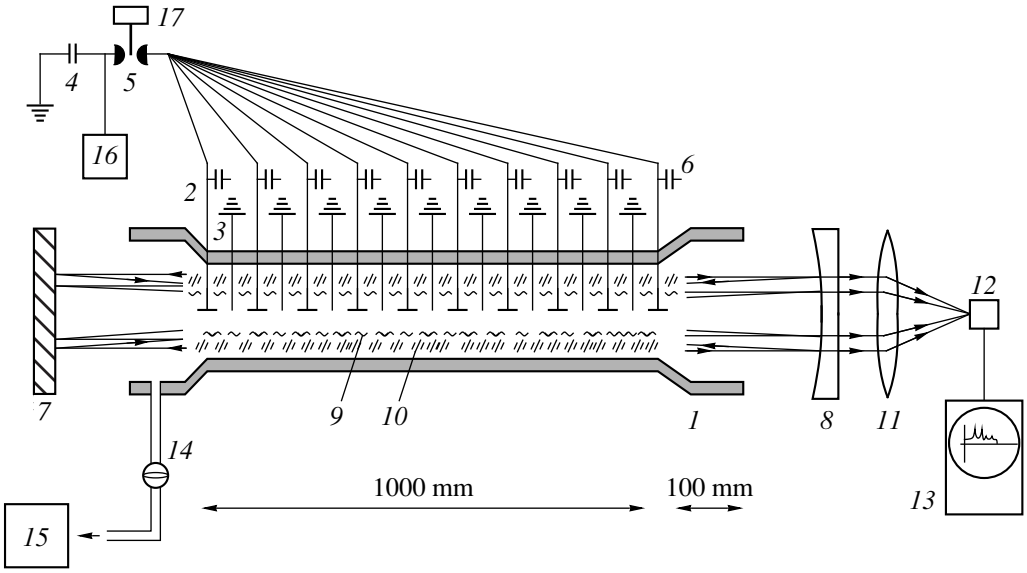


Fig. 1. Photodissociation iodine laser pumped by an open electric discharge. *1*, glass–textolite tube (length $l = 1200$ mm, inner diameter $d = 150$ mm) with plane-parallel end windows made of K-8 glass; *2* and *3*, tungsten discharge electrodes hermetically sealed into tube *1*; *3*, grounded electrodes; *4*, bank of high-voltage KVN capacitors (100 μF) connected to electrodes *2* through gap *5*; *5*, gap; *6*, high-voltage low-inductance capacitors controlling plasma cylinder shape; *7*, totally reflecting mirror (focal distance $f = 2$ m, reflectivity $r = 99.8\%$); *8*, output mirror ($f = 2$ m, $r = 98\%$); *9*, expanding plasma cylinder due to electric discharge; *10*, expanding cylinder of photodissociation and lasing zone; *11*, focusing glass lens; *12*, photodetector (also calorimeter or target); *13*, storage oscilloscope for photodetector; *14*, vacuum tract; *15*, vacuum system for preparation and storage of working gases; *16*, IK-50 power supply to apply 25-kV voltage to capacitors *4*; and *17*, unit to control gap, oscilloscope, etc.

great losses of excited atoms $I(^2P_{1/2})$ for the time during which the inversion persists. Moreover, the amplification cannot be modulated by applying external magnetic fields [13] because of the high magnetic field of the pump source.

We designed the electric-discharge single-pulse PIL shown in Fig. 1. The use of indestructible tungsten electrodes *2* and *3* and high-voltage low-inductance capacitors *6* makes it possible to form, with control system *17*, a specific space–time structure of the plasma surface. The plasma is initiated by a discharge of main capacitor bank *4*. In experiments, the voltage across bank *4* was 25 kV, which provided the stored energy $E_0 = 31.3$ kJ.

Prior to each pulse, lasing cavity *1* was evacuated and then filled by a fresh gas mixture prepared in vacuum system *15*. An SF_6 + perfluoroalkyloid mixture used in [14] allows for only one shot. We applied a train of short pulses, using an iC_3F_7I (25–37 wt %) + Xe (75–63 wt %) mixture. The total pressure was varied from 0.05 to 0.2 atm. With such an operating mixture, lasing zone *10* was away from expanding plasma cylinder *9*, which reduces refraction losses [15] of stimulated emission (5).

The time variation of the current pulse I in the discharge circuit (Fig. 2a) was recorded with a Rogovskiĭ loop. The total energy W_Σ of a laser beam was measured with an IKT calorimeter instead of target *12* (Fig. 1). The pulse shape was displayed on storage oscilloscope *13*, which receives the signal from the photocathode mounted in the path of a deflected beam. The time variation of the output power P is shown in Fig. 2b. The energy efficiency of the PIL in this case is

$$\eta_e = \frac{W_\Sigma}{E_0} = \frac{(30 \pm 5)\text{J}}{31.3\text{kJ}} = 0.1\%, \qquad (7)$$

which is comparable to $\eta_e = 0.4\%$, reported in [3]. Figure 2 implies that the beam power peaks when the magnetic field of the discharge vanishes. This is consistent with the results in [13], where the pulse appeared when the pump current dropped to near-zero values.

KINETIC MODEL OF PHOTODISSOCIATION IODINE LASER

If losses were absent and the iodide in the lasing cavity were fully dissociated, we would obtain $W_{\Sigma,\,\max} = 1492$ J. Thus, the photochemical efficiency in the experiment (Fig. 2b) was

$$\eta_{\text{ph}} = \frac{W_\Sigma}{W_{\Sigma,\,\max}} = (2.0 \pm 0.4\%). \qquad (8)$$

This means that, in principle, the efficiency can be improved. For example, η_{ph} and η_e can be doubled for the same PIL design if iC_3F_7I iodides are replaced by $(CF_3)_2PI$ [16] or $C_6F_{13}I$ [17, 18] compounds, which have the greater product $p_1\langle\sigma_6\rangle$. More accurate relationships between the PIL energy parameters and those of

a pumping source and active medium can be derived from a PIL kinetic model.

The construction of the kinetic model must begin with determining the magnitude of the pumping radiation $\varphi_p(t)$. It may be found fairly accurately by the method of autoactinometry [19] from laser emission characteristics.

In Fig. 2b, the lasing time is $\tau = 110$ μs. For the inner radius of tube *1* $d/2 = 7.5$ cm (Fig. 1), the expansion rate of plasma cylinder *9* is

$$v_{ex} = \frac{d}{2\tau} = 680 \text{ m/s}. \quad (9)$$

With such a velocity of the shock wave, the radius of the plasma cylinder $r(\Delta t)$ in a time $\Delta t = 25$ μs becomes equal to $\nu_{ex}\Delta t = 1.7$ cm. In the operating mixture used, the iodide photodissociates to a depth of 1.5–2.0 cm. As a result, at the time instant $\Delta t = 25$ μs, lasing occurs within a cylindrical layer of inner surface area

$$S = 2\pi r(\Delta t)l = (1000 \pm 100) \text{ cm}^2 \quad (10)$$

and volume

$$V = 2\pi(0.5 + r(\Delta t))l(1.75 \pm 0.25) = (2000 \pm 300) \text{ cm}^3. \quad (11)$$

In Fig. 2b, the power of the second laser peak is $P(25\ \mu s) = (6 \pm 1)$ MW, which is equivalent to the stimulated flux of quanta $h\nu_s = 0.9544$ eV that is averaged over $\bar{\varphi}_{st}(25\ \mu s) = 5.3 \times 10^{24}$ s^{-1}. This flux comes from volume (11); hence, the rate of lasing transition (6) is

$$\frac{\bar{\varphi}_{st}(25\ \mu s)}{V} = \frac{d[I(^2P_{1/2})]}{dt} = \frac{2}{3}p_1\left(-\frac{d[RI]}{dt}\right) = (2.7 \pm 0.5) \times 10^{21} \text{ cm}^{-3} \text{ s}^{-1}. \quad (12)$$

At the instant of the second peak, the flux φ_p in the absorption band of the iodide and parameters of the PIL active medium are fairly accurately [19] related as

$$\varphi_p(t_{max}) = \frac{3\bar{\varphi}_{st}}{2p_1V\langle\sigma_{RI}\rangle[RI]_{t_{max}}}, \quad (13)$$

where $\langle\sigma_{RI}\rangle$ is the absorption-band-averaged cross section of photodissociation (6).

For iC$_3$F$_7$I, $\langle\sigma_{RI}\rangle$ was measured to be 6.0×10^{-19} cm^2 in the band $\Delta_{iC_3F_7I} = 275 \pm 22$ nm. At

$$[RI]_{t_{max}} \cong 0.5[RI]_0 = 4.1 \times 10^{17}\frac{\text{mol}}{\text{cm}^3}, \quad (14)$$

substituting (12) into (13) yields the desired pumping radiation flux

$$\varphi_p(25\ \mu s) = (1.8 \pm 0.4) \times 10^{22}\frac{h\nu}{\text{cm}^2 \text{ s}}. \quad (15)$$

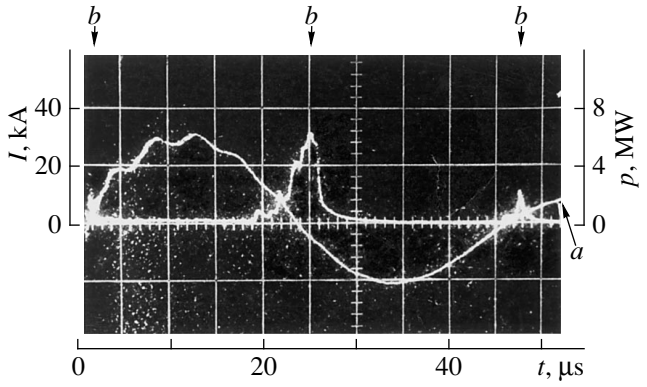


Fig. 2. (a) Current pulse I vs. time for an open discharge between electrodes *2* and *3* (see Fig. 1) and (b) time variation of the PIL output power P incident on target *12* (Fig. 1). The data are given for the iC$_3$F$_7$I + Xe working mixture. The partial pressures are 25 torr for iC$_3$F$_7$I and 125 torr for Xe; $W_\Sigma = 30 \pm 5$ J. The oscillogram shows the first three out of six emission peaks and the first out of three decaying current waves.

Let us check the reliability of the obtained $\varphi_p(25\ \mu s)$. For the average surface area given by (10), flux (15) provides the radiant energy release

$$E_\Delta = (h\nu = 4.51 \text{ eV})\varphi_p\tau S = (8 \pm 2) \times 10^{21} \text{ eV} = 1300\text{J} \quad (16)$$

for $\tau = 100$ μs in the band $\Delta_{iC_3F_7I}$. This comprises

$$\eta_\Delta = \frac{E_\Delta}{E_0} = (4 \pm 1)\% \quad (17)$$

of the stored energy $E_0 = 31.3$ kJ and is consistent with the Planck distribution of blackbody radiation with a gas-discharge plasma temperature of 14000 K [20]. The value $\eta_\Delta = 4 \pm 1\%$ is the limiting energy efficiency for PILs of the given design.

The shock wave having velocity (9) passes through the lasing layer of the active gas (~2 cm thick) for the time $\Delta t = 27$ μs. Within this time, flux (15) provides photodissociation to the level

$$\frac{[i\text{-}C_3F_7I]_{\Delta t}}{[i\text{-}C_3F_7I]_0} = e^{-\langle\sigma_{RI}\rangle\varphi_p\Delta t} = e^{-0.3} = 0.74. \quad (18)$$

This agrees with estimate (14), so that a refinement of $[RI]_{t_{max}}$ in formula (13) will leave flux (15) within the indicated error. According to [21], the velocity of a bleaching (i.e., iodide photodissociation) wave is

$$v_{ph} = \left(\frac{1}{c} + \frac{[RI]_0}{\Phi_0}\right)^{-1}, \quad (19)$$

where c is the speed of light and Φ_0 is the flux of photolyzing radiation in the absorption band of RI at the front of a moving pumping source.

Under the conditions of Fig. 2b, i.e., for $[RI]_0$ = 25 torr and photolyzing flux (15), it follows from (19) that

$$v_{ph} \cong \Phi_0/[RI]_0 = 220 \text{ m/s}. \quad (20)$$

Since shock wave velocity (9) far exceeds (20), we deal with a photodissociation wave of the second kind [21]. In this case, the iodide concentration and the stimulated emission output immediately at the front of a coming shock wave are expressed as

$$[iC_3F_7I]_{r = v_{ex}t} = [RI]_0 - \frac{\Phi_0}{v_{ex}} = 5.8 \times 10^{17} \frac{\text{mol}}{\text{cm}^3}, \quad (21)$$

$$G(r = v_{ph}t) = \frac{p_1/\langle\sigma_{RI}\rangle}{1.5}\Phi_0\left([RI]_0 - \frac{\Phi_0}{v_{ex}}\right) = 3.8 \times 10^{21} \frac{h\nu}{\text{cm}^3\ \text{s}}. \quad (22)$$

From (18), only a fraction of the initial iC_3F_7I molecules given by

$$\eta_{ph} = 0.26\frac{2}{3}p_1 = 0.156 \quad (23)$$

contributes to lasing. Then, the energy efficiency of our PIL would have to be

$$\eta_e = \eta_{ph}\eta_\Delta = 0.0064. \quad (24)$$

The experimentally found values [see (7) and (8)] turned out to be seven times less. In Fig. 2b, the fourfold decrease in W_Σ is due to the action of the magnetic field. The almost twofold decline in $P(t)$ may well be associated with chemical and refraction [15] losses. Thus, if the detrimental effect of the magnetic field were completely eliminated, W_Σ and, accordingly, η_e would increase fourfold (η_e would attain 0.4%). It is precisely this value that was reported for electric-discharge PILs designed in [3] and [20]. Note for comparison that η_e = 1.4% was attained in lamp-pumped PILs at the excitation energy E_{exc} = 1000 J [2]. The primary advantage of lamp pumping is that an electric discharge in the closed volume of a lamp provides a luminance temperature of pumping radiation as high as 22000 K.

INTERACTION OF LASER PULSES WITH SOLID TARGET

From the aforesaid, it becomes clear that a periodic magnetic field applied to the PIL lasing medium decreases the output W_Σ by a factor of 3 or 4. However, the erosive effect of a laser beam is enhanced in this case. A similar effect was also observed earlier: the plasma generation threshold in the laser focus decreases under repeat pulse irradiation of the sample [22].

Products of a laser jet absorb laser emission insignificantly [9], so that the breakdown energy is concentrated largely at the pulse front. If magnetic modulation were absent in the experiments represented by Fig. 2b, the total pulse energy would be W_Σ = 100–120 J, of which no more than 6–8 J would be absorbed by the steel target. Due to magnetic modulation, W_Σ was set at 30 ± 5 J; however, more than 75% of the energy penetrated into the target.

For steel, C_s = 0.11 cal/g/°C, T_m = 1536°C, q_m = 65.5 cal/g, $C_l \cong C_s$, T_b = 3000°C, and q_{ev} = 1511 cal/g. To melt steel, E_m should equal 984 J/g, and to evaporate steel, E_{ev} should be 8.02 J/g. In the experiments, laser pulses were focused into a 0.02-cm^2 spot, so that the energy density absorbed by the target was

$$W_a = \alpha W_\Sigma = \frac{(25 \pm 5)}{0.02\ \text{cm}^2} = (300 \pm 60)\frac{\text{cal}}{\text{cm}^2}. \quad (25)$$

Eventually, from (3) and (4), we obtain (for α = 1) the limiting molten steel thickness,

$$h_m - 1.28\ \text{g/cm}^2 = 0.16\ \text{cm} \quad (26)$$

and the limiting thickness of evaporated steel

$$h_{ev} = 0.157\ \text{g/cm}^2 = 0.02\ \text{cm}. \quad (27)$$

The depth of an erosion spot was found to be about 0.1 mm. This means that no less than 100 kJ are required to remove 1 cm^3 of the steel target. This value is comparable to 250 ± 50 kJ/cm^3, which was found for pulse CO_2 lasers [23]. It is known that the erosive effect of laser emission on solid targets is enhanced with decreasing wavelength [22].

REFERENCES

1. G. Brederlov, E. Fill, V. Fuss, *et al.*, Kvantovaya Élektron. (Moscow) **3**, 906 (1976).
2. A. S. Antonov, I. M. Belousova, V. A. Gerasimov, *et al.*, Pis'ma Zh. Tekh. Fiz. **4**, 1143 (1978) [Sov. Tech. Phys. Lett. **4**, 459 (1978)].
3. N. G. Basov, V. S. Zuev, V. A. Katulin, *et al.*, Kvantovaya Élektron. (Moscow) **6**, 311 (1979).
4. A. F. Gibson and M. H. Key, Rep. Prog. Phys. **43** (1), 1 (1980).
5. V. P. Arzhanov, B. L. Borovich, V. S. Zuev, *et al.*, Kvantovaya Élektron. (Moscow) **19**, 135 (1992).
6. N. G. Basov, N. I. Belousov, P. A. Grishunin, *et al.*, Kvantovaya Élektron. (Moscow) **24**, 2068 (1987).
7. V. P. Veĭko and M. N. Libenson, in *Laser Processing* (Lenizdat, Leningrad, 1973).
8. P. P. Pashanin, S. F. Rastopov, and A. T. Sukhodol'skiĭ, Kvantovaya Élektron. (Moscow) **14**, 869 (1997).
9. V. I. Mazhukin and G. A. Pestryakova, Dokl. Akad. Nauk SSSR **278**, 843 (1984) [Sov. Phys. Dokl. **29**, 791 (1984)].
10. J. V. V. Kasper, J. H. Parker, and G. C. Pimentel, J. Chem. Phys. **43**, 1827 (1965).
11. G. A. Skorobogatov, Doctoral Dissertation (Leningradsk. Gos. Univ., Leningrad, 1980), p. 225.

12. B. L. Borovich, V. S. Zuev, V. A. Katulin, *et al.*, Itogi Nauki Tekh., Ser. Radiotekh. **15**, 189 (1978).
13. I. M. Belousova, B. D. Bobrov, A. S. Grenishin, and V. M. Kiselev, Kvantovaya Élektron. (Moscow) **4**, 446 (1977).
14. B. V. Alekhin, B. V. Lazhintsev, V. A. Nor-Arevyan, *et al.*, Kvantovaya Élektron. (Moscow) **3**, 2369 (1976).
15. V. S. Zuev, K. S. Korol'kov, O. Yu. Nosach, and E. P. Orlov, Kvantovaya Élektron. (Moscow) **7**, 2604 (1981).
16. B. P. Dymov and G. A. Skorobogatov, Zh. Tekh. Fiz. **48**, 124 (1978) [Sov. Phys. Tech. Phys. **23**, 73 (1978)].
17. G. A. Skorobogatov, B. N. Maksimov, V. G. Seleznev, *et al.*, Kvantovaya Élektron. (Moscow) **9**, 253 (1982).
18. G. A. Skorobogatov, V. G. Seleznev, B. N. Maksimov, and O. N. Slesar', Zh. Tekh. Fiz. **45**, 2454 (1975) [Sov. Phys. Tech. Phys. **20**, 1533 (1975)].
19. V. S. Komarov, V. G. Seleznev, and G. A. Skorobogatov, Zh. Tekh. Fiz. **44**, 875 (1974) [Sov. Phys. Tech. Phys. **19**, 558 (1974)].
20. A. S. Kamrukov, G. N. Kashnikov, N. P. Kozlov, *et al.*, Kvantovaya Élektron. (Moscow) **3**, 2023 (1976).
21. G. A. Skorobogatov, Pis'ma Zh. Tekh. Fiz. **1**, 452 (1975) [Sov. Tech. Phys. Lett. **1**, 209 (1975)].
22. V. P. Ageev, A. A. Gorbunov, V. I. Konov, and A. M. Prokhorov, Izv. Akad. Nauk SSSR, Ser. Fiz. **49**, 732 (1985).
23. A. A. Vedenov, G. G. Gladush, S. V. Drobyazko, and Yu. M. Senatorov, Kvantovaya Élektron. (Moscow) **8**, 2154 (1981).

Translated by V. Isaakyan

Technical Physics, Vol. 45, No. 10, 2000, pp. 1296–1306. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 65–75.

OPTICS, QUANTUM ELECTRONICS

Interaction of Light with a Semiconductor Thin Film, Taking into Account Concentration Gain of the Dipole Moment of the Excitonic Transition

I. Khadzhi and L. V. Fedorov

Shevchenko State University, Pridnestrov'e, Tiraspol, 3300 Moldova

Received October 13, 1999; in final form, April 7, 2000

Abstract—The peculiarities of nonlinear transmission of ultrashort pulses of resonance laser radiation by a semiconductor thin film in the excitonic spectral region have been studied, taking into account the exciton–phonon interaction and concentration gain of the dipole moment of the excitonic transition. It has been shown that under exact resonance conditions the film transmits only the leading edge of the incident rectangular pulse and completely reflects the remaining part. For a nonzero off-resonance, a residual transmission occurs. A series of possibilities for transforming Gaussian pulses is predicted. The state equations for steady-state bistable transmission (reflection) have been derived, and the stability of the solutions has been studied. A theorem of areas, which predicts area restriction of the transmitted pulses, has been stated; namely, the transmitted pulse area cannot exceed $\pi/2$. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The nonlinear optical phenomena in a system of coherent excitons in semiconductors have usually been studied within the framework of an infinite medium model (see, for example, [1, 2] and references therein). However, in recent years the nonlinear properties of dimensionally confined semiconductor structures, in particular, thin films, have been attracting ever growing scientific and practical interest. A peculiar relationship between the field of an electromagnetic wave propagating through a semiconductor thin film (STF) and polarization of the medium gives rise to a number of interesting physical effects [3–9]. Theoretical studies have been carried out on the peculiarities of the non-steady-state interaction of ultrashort pulses (USP) of laser radiation with an STF consisting of two-level atoms [3–8] and in the excitonic spectral region, taking into account exciton–phonon interactions, two-photon excitation of biexcitons from the ground state of the crystal, optical exciton–biexciton conversion, and the effects of saturation of the dipole moment of the excitonic transition [9]. In the light of this, the study of nonlinear optical properties of STF would be of interest, since semiconductors possess high nonlinear susceptibilities, a variety of nonlinearity mechanisms, and small relaxation times of media excitations, which makes the prospects of their practical application highly attractive.

This paper presents the results of a theoretical study of nonlinear optical properties in the excitonic spectral region of an STF interacting with USP of resonant laser radiation that takes into consideration the effect of concentration gain of the dipole moment of the excitonic transition. This mechanism of nonlinearity was proposed in a study of the features of non-resonator laser effect in excitons [10]. Apparently, this mechanism is similar to the nonlinearity mechanism known in the theory of optical adsorption bistability, which is responsible for an increase in the absorption coefficient with the electron–hole pair density [11, 12]. If the frequency of the exciting laser radiation is tuned slightly below the excitonic resonance, only weak single-photon absorption occurs in the Urbach tail of the absorption band. Nevertheless, electron–hole pairs are generated in the crystal. With increasing pair density, many-body effects cause renormalization of the electrical spectrum, in particular, the red shift of the band edge. For a certain pair density, the band edge will coincide with the energy of excited photons and the absorption will be considerably increased. The increase in absorption can be interpreted as a concentration enhancement of the dipole moment of excitonic transition at high levels of laser excitation.

FORMULATION OF THE PROBLEM: BASIC EQUATIONS

Suppose that a USP of the resonant laser radiation is normally incident on an STF with a thickness L much less than the light wavelength λ in vacuum and the envelope of the electric field $E_i(t)$ of the electromagnetic wave varies slowly with a frequency ω and duration τ_p that is much less than the exciton relaxation time τ_{rel} but much longer than the oscillation period of the field of the wave ($\omega^{-1} \ll \tau_p \ll \tau_{rel}$). Under these conditions, the real (dissipative) light absorption is missing

1063-7842/00/4510-1296$20.00 © 2000 MAIK "Nauka/Interperiodica"

[13]. Photons of the pulse passing through the STF excite excitons from the ground state of the crystal and interact with them. Through radiative recombination of the excitons, energy is returned to the passing pulse, which causes distortion of its shape. At high excitation levels, the exciton–exciton interaction can produce various many-body effects in the high-density excitonic system [14], in particular, concentration gain of the dipole moment of the excitonic transition. The Hamiltonian for the interaction of excitons with the field of an electromagnetic wave in this case can be written in the form [10]

$$H = -\hbar g(a^+E^+ + aE^-) - \hbar\mu(a^+a^+aE^+ + a^+aaE^-), \quad (1)$$

where $E^+(E^-)$ is a positive- (negative-) frequency component of the wave field, a is the amplitude of the wave of excitonic polarization of the medium, the constant g describes interaction of the excitons with the wave field, and the constant μ describes the effect of the concentration gain of the transition dipole moment.

Using (1), it is easy to obtain the Heisenberg equation of motion for the amplitude a of the excitonic wave,

$$i\dot{a} = \omega_0 a - gE^+ - 2\mu a^+aE^+ - \mu aaE^-, \quad (2)$$

where ω_0 is the frequency of excitonic transition.

This equation must be supplemented with electrodynamics equations, which relate the amplitudes of slowly varying envelopes of incident $E_i^+(t)$, transmitted $E_t^+(t)$, and reflected $E_r^+(t)$ pulses, which are obtained from the conditions of conservation of tangential components of the fields at the crystal–vacuum interface [3–9],

$$E^+ = E_i + \frac{4\pi\omega L}{c}P^+, \quad (3)$$

$$E_r^+ + E_i = E^+, \quad (4)$$

where P^+ is the medium polarization

$$P^+ = \hbar\mu(g + \mu a^+a)a. \quad (5)$$

The system of related Eqs. (2)–(5) is closed. The problem consists in determining the dependence of the temporal evolution of the envelopes of the transmitted and reflected pulses, assuming that the envelope $E_i(t)$ of the incident pulse has been specified.

The solutions of the system of Eqs. (2)–(5) will be sought in the form of a product of slowly varying envelopes (containing real and imaginary components) and a rapidly oscillating high-frequency exponential factor:

$$E^+ = (E_t' + iE_t'')e^{-i\omega t},$$
$$E_r^+ = (E_r' + iE_r'')e^{-i\omega t}, \quad a = (u + iv)e^{-i\omega t}. \quad (6)$$

Let us introduce the following normalized (dimensionless) variables:

$$F_i = \sqrt{\mu g}\tau_0 E_i, \quad F_i^+ = F_t' + iF_t'' = \sqrt{\mu g}\tau_0 E^+,$$
$$F_r^+ = F_r' + iF_r'' = \sqrt{\mu g}\tau_0 E_r^+,$$
$$x = \sqrt{\kappa}u, \quad y = \sqrt{\kappa}v, \quad \kappa = \mu/g, \quad \tau = t/\tau_0, \quad (7)$$
$$\delta = \Delta\tau_0, \quad \Delta = \omega - \omega_0,$$

where Δ is the off-resonance and τ_0 is the main temporal parameter of the STF response to the external radiation, which is determined by the equation

$$\tau_0^{-1} = 2\pi\hbar\omega L g^2/c. \quad (8)$$

Then the system of Eqs. (2)–(4) can be reduced to the form

$$\frac{dx}{d\tau} = -\delta y - 2F_i xy - x(1 + x^2 + y^2)^2, \quad (9)$$

$$\frac{dy}{d\tau} = \delta x + F_i(1 + 3x^2 + y^2) - y(1 + x^2 + y^2)^2, \quad (10)$$

$$F_t' = F_i - y(1 + x^2 + y^2), \quad F_t'' = x(1 + x^2 + y^2), \quad (11)$$

$$F_r' = -y(1 + x^2 + y^2), \quad F_r'' = x(1 + x^2 + y^2). \quad (12)$$

Equations (8) and (9) contain only two dimensionless parameters, namely, the normalized amplitude $F_i(\tau)$ of the incident pulse field and the normalized off-resonance δ.

From (9) and (10) it follows that, in general, the excitonic wave amplitude, as well as the amplitudes of the fields of transmitted and reflected pulses, are complex quantities, that is, phase-modulated functions. If a normalized exciton density $z = x^2 + y^2$ is introduced and all quantities are expressed through their amplitudes and phases as $F_t^+ = F_i e^{i\varphi_t}$, $F_r^+ = F_r e^{i\varphi_r}$, and $a = \sqrt{\kappa z}\, e^{i\varphi}$, then Eqs. (9)–(12) can take the form

$$\frac{dz}{d\tau} = 2\sqrt{z}(1 + z)[F_i \sin\varphi - \sqrt{z}(1 + z)], \quad (13)$$

$$\frac{d\varphi}{d\tau} = \delta + \frac{1 + 3z}{\sqrt{z}}F_i \cos\varphi, \quad (14)$$

$$F_t = [F_i^2 - 2F_i\sqrt{z}(1 + z)\sin\varphi + z(1 + z)^2]^{1/2}, \quad (15)$$

$$F_r = \sqrt{z}(1 + z), \quad (16)$$

$$\tan\varphi_t = \sqrt{z}(1 + z)\cos\varphi[F_i - \sqrt{z}(1 + z)\sin\varphi]^{-1},$$
$$\varphi_r = \frac{\pi}{2} + \varphi. \quad (17)$$

We consider the features of the steady-state regime of transmission (reflection) in an STF when a pulse of laser radiation of constant amplitude F_i = const illumi-

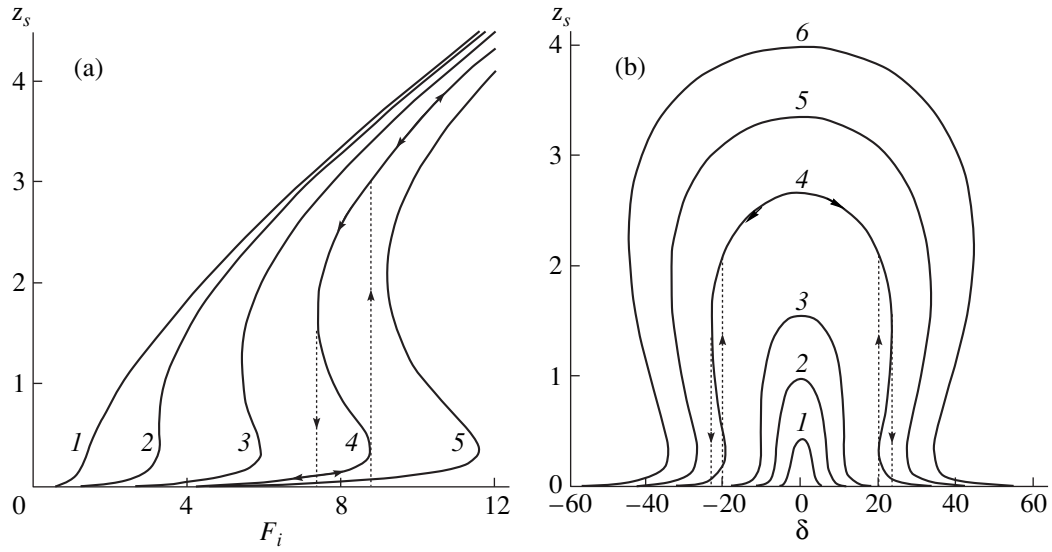


Fig. 1. Dependence of the exciton density z_s in a steady-state regime on (a) the field amplitude F_i of the incident pulse at $\delta = 4$ (*1*), 10.8 (*2*), 20 (*3*), 30 (*4*), 40 (amplitude hysteresis) (*5*) and on (b) the off-resonance δ at $F_i = 1$ (*1*), 2 (*2*), 3.23 (*3*), 6 (*4*), 8 (*5*), 10 (frequency hysteresis) (*6*).

nates the film for a time $1 \ll \tau \ll \gamma/\tau_0$. The steady-state regime, as will be shown below, sets in within a time t of the order of several τ_0. According to our estimates [9], $\tau_o \approx 10^{-13}$ s. Since $\gamma \approx 10^{-10}$–10^{-11} s, $\gamma/\tau_0 \approx 10^3$–$10^4$, so that the steady state has ample time to set in during the pulse time, which is less than the relaxation time. The steady-state values of the exciton density z_s and the phase of the excitonic wave φ_s are determined from the equations

$$F_i \sin\varphi_s = \sqrt{z_s}(1 + z_s), \tag{18}$$

$$\delta\sqrt{z_s} + F_i(1 + 3z_s)\cos\varphi_s = 0, \tag{19}$$

from which a nonlinear equation for z_s can be obtained in the form

$$z_s[(1 + z_s)^2 + \delta^2/(1 + 3z_s)^2] = F_i^2. \tag{20}$$

The steady-state exciton density z_s in an STF is a nonlinear function of the pumping amplitude F_i and δ. It follows from (20) that $z_s \sim F_i^2/(1 + \delta^2)$ at low excitation levels and $z_s \sim F_i^{2/3}$ at high excitation levels. Thus, with an increase in the excitation level, the kinetics of exciton generation change significantly. Considering F_i^2 as a function of z_s (at fixed δ) it is easy to show that at $|\delta| < \delta_c = \sqrt{(349 + 85\sqrt{17})/6} \approx 10.797$ it is a single-valued, monotonically increasing function with a point of inflection at $F_{ic} = \sqrt{(71 + 17\sqrt{17})/27} \approx 3.233$, where $z_s = (71 + 17\sqrt{17})/6 \approx 0.521$. At $|\delta| > \delta_c$, the function $F_i^2(z_s)$ has two extrema at positive values of z_s. As regards the inverse function $z_s(F_i)$, at $|\delta| < \delta_c$, it is triple-valued in a certain range of F_i values.

STEADY-STATE OPTICAL BISTABILITY

In Fig. 1, plots of $z_s(F_i)$ for fixed δ values and of $z_s(\delta)$ for fixed F_i values are presented, and in Fig. 2, a 3D plot of $z_s(F_i, \delta)$ is shown. From Fig. 1a it is seen that at $|\delta| < \delta_c$ a monotonic increase in the exciton density occurs with an increase in the field amplitude E_i; that is, the dependence $z_s(F_i)$ is nonlinear but single-valued (curves *1*, *2*). The situation changes at $|\delta| > \delta_c$. At low excitation levels, a unique linear relationship is found between the exciton density in the film and the intensity of incident radiation; that is, $z_s \sim F_i^2/(1 + \delta^2)$. However, at high levels of excitation, this relationship becomes ambiguous. From Fig. 1a it follows that with an increase in the field amplitude F_i the exciton density initially increases monotonically along the lower branch of the hysteresis curve. When a critical value of the field amplitude F_i is reached, a further increase in the amplitude results in an abrupt change in the exciton density. An additional increase in the field results in a monotonic increase in the excitonic density along the upper branch of the hysteresis curve. If the field amplitude now is decreased as it moves along the upper branch, then an abrupt change in the excitonic density occurs again, but at a smaller value of the field amplitude. Thus, during a cyclic change (increase and

decrease) in the amplitude of the incident field, abrupt changes in the exciton density occur, which produces a hysteresis in the $z_s(F_i)$ curve. These abrupt changes are indicated on curve *4* in Fig. 1a. The width of the hysteresis curve increases with δ. The physical reason for the hysteresis is the dependence on density of the dipole moment of the excitonic transition.

The hysteresis presented in Fig. 1a is an amplitude-related hysteresis, since it is caused by variation of the field amplitude F_i at a fixed off-resonance δ (frequency of the incident wave ω). However, a frequency-related hysteresis is also possible if the off-resonance δ (or frequency ω) is varied at a fixed amplitude F_i. Curves of the frequency hysteresis $z_s(\delta)$ are presented in Fig. 1b. It is seen that at low excitation levels no frequency hysteresis occurs and the curve $z_s(\delta)$ is a Lorentzian. The hysteresis arises at field amplitude values $F_i > F_{ic}$. The hysteresis curve width increases with F_i. The $z_s(\delta)$ curve in this case is a distorted Lorentzian. During a cyclic change (increase and decrease) in the off-resonance δ, two abrupt changes in the forward direction and two abrupt changes in the backward direction arise (curve *4* in Fig. 1b), which form two ranges of z_s hysteresis variation in the clockwise and counterclockwise directions. In Fig. 2, a three-dimensional profile of the function $z_s(F_i, \delta)$ is presented; its projections on the (z_s, F_i) and (z_s, δ) planes are the curves in Figs. 1a and 1b.

In Fig. 3, variations of the steady-state phase φ_s of the excitonic wave are shown as a function of the pumping amplitude F_i at fixed δ. It follows from (18) and (19) that $\varphi_s = \pi/2$ at $\delta = 0$. At low excitation levels ($F_i \longrightarrow 0$), we obtain $\sin\varphi_s = (1 + \delta^2)^{-1/2}$; that is, with increasing off-resonance, the phase φ_s rapidly decreases and goes to zero at $|\delta| \gg 1$. At $|\delta| < \delta_c$, the phase φ_s increases rapidly with the pumping amplitude F_i, while at $|\delta| > \delta_c$ it undergoes hysteresis variation for a cyclic change of F_i. At high values of the pumping amplitude ($F_i \gg 1$), the steady-state phase of the excitonic wave asymptotically approaches $\pi/2$. Note also that according to (17)–(19), in the steady-state regime, the phase φ_t of the wave transmitted through STF coincides with the phase φ_s of the excitonic wave.

The expressions obtained for $z_s(F_i, \delta)$ and $\varphi_s(F_i, \delta)$ directly determine the behavior of the steady-state functions of transmission $F_t(F_i, \delta)$ and reflection $F_r(F_i, \delta)$. In Fig. 4a, a plot of the amplitude of the transmitted wave F_t versus the amplitude F_i of the pumping field is shown for different values of δ derived from Eqs. (15) and (20). It can be seen that, at $|\delta| < \delta_c$, this dependence is nonlinear but single-valued, while at $|\delta| > \delta_c$, the function $F_t(F_i)$ shows hysteresis. The hysteresis curve width increases with δ. As the incident wave amplitude F_i increases, the amplitude of the transmitted wave F_t essentially increases in proportion to F_i; that is, the function $F_t(F_i)$ is linear. However, at certain

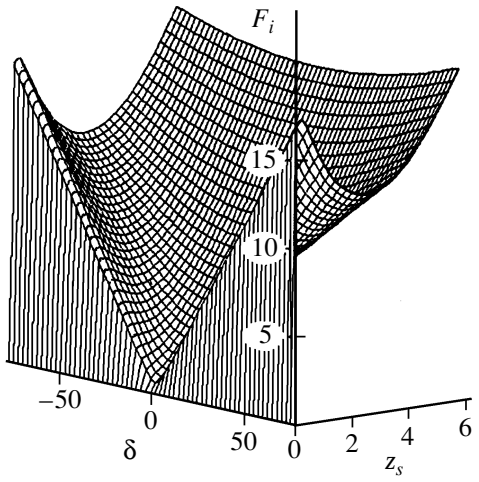


Fig. 2. Dependence of the exciton density z_s on F_i and δ.

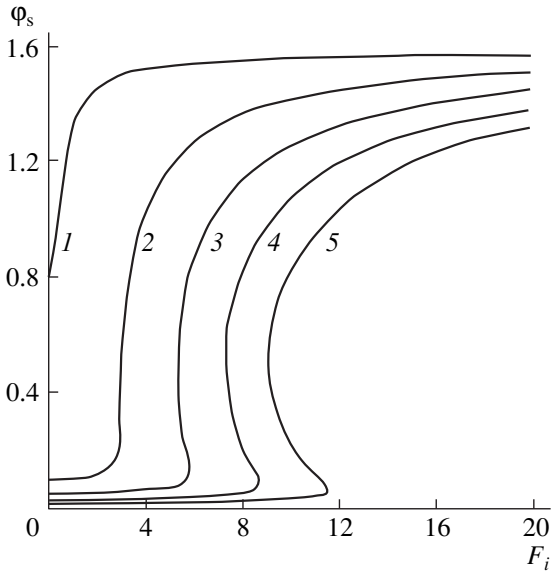


Fig. 3. Dependence of the phase φ_s of the wave of the excitonic medium polarization on F_i at $\delta = 1$ (*1*), 2 (*2*), 3.23 (*3*), 6 (*4*), and 8 (*5*).

critical values of F_i determined by δ, a jump from the upper to the lower branch of the hysteresis curve occurs and the film abruptly darkens; with a further increase in F_i, the transmission decreases monotonically. For a cyclic variation of F_i, the jump of the F_t function in the forward direction is greater than the jump in the backward direction. Similar plots for the intensity of the reflected wave are presented in Fig. 4b. In Fig. 5, the frequency hysteresis of the transmission function $F_t(\delta)$ is shown for fixed values of the incident wave amplitude F_i. It is seen that there is a critical value of the field F_i at which the transmission becomes bistable.

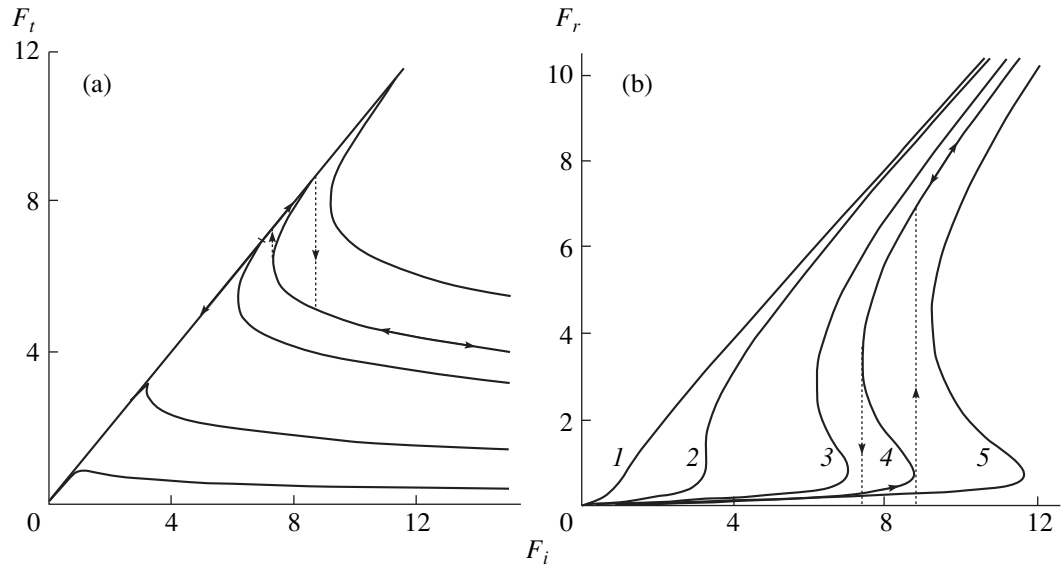


Fig. 4. Dependence of the field amplitude of the transmitted F_t (a) and the reflected F_r (b) pulses on F_i at $\delta = 4$ (*1*), 10.8 (*2*), 20 (*3*), 30 (*4*), and 40 (*5*).

STABILITY OF THE STEADY-STATE SOLUTIONS

Let us consider the stability of the steady-state solutions obtained for z_s and φ_s. To this end, we will use Eqs. (13), (14), (18), and (19). We assume that $z = z_s + x$ and $\varphi = \varphi_s + \psi$, where $|x| \ll z_s$ and $|\psi| \ll |\varphi_s|$, and linearize the system of Eqs. (13) and (14) with respect to small deviations of x and ψ from equilibrium:

$$\dot{x} = 2\sqrt{z_s}(1 + z_s)\left[F_i \cos\varphi_s \psi - \frac{3z_s + 1}{2\sqrt{z_s}}x\right], \qquad (21)$$

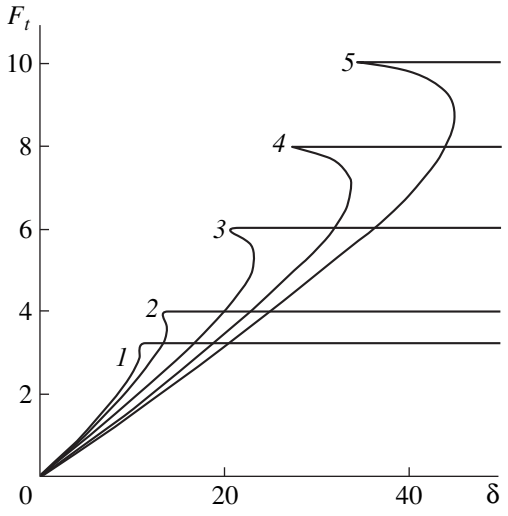


Fig. 5. Frequency hysteresis of the transmission function $F_i = 3.23$ (*1*), 4 (*2*), 6 (*3*), 8 (*4*), and 10 (*5*).

$$\dot{\psi} = \frac{F_i}{\sqrt{z_s}}\left[\frac{3z_s - 1}{2z_s}\cos\varphi_s x - (3z_s + 1)\sin\varphi_s \psi\right]. \qquad (22)$$

A search for solutions of the system in the form $x = Ae^{\lambda\tau}$, $\psi = Be^{\lambda\tau}$ yields a characteristic equation for the parameter λ, whose solutions have the form

$$\lambda = -(z_s + 1)(3z_s + 1) \pm \frac{|\delta|}{3z_s + 1}\sqrt{(3z_s - 1)(z_s + 1)}. \qquad (23)$$

Hence, the stability condition for the steady-state solutions is the inequality

$$(3z_s + 1)^4(z_s + 1) > \delta^2(3z_s - 1), \qquad (24)$$

which coincides with inequality $\partial z_s/\partial F_i > 0$. Thus, all steady-state solutions which have a positive slope of the tangent to the curve $z_s(F_i)$ are stable. Unstable solutions correspond to the middle part of the hysteresis curve. In addition, it follows from (23) that the curve $z_s = 1/3$, which in the space of the parameters (F_i, δ) is described by the equation

$$\delta^2 = 4\left(3F_i^2 - \frac{16}{9}\right), \qquad (25)$$

separates stationary points in the phase plane (z_s, φ_s), which on one side are stable focuses and on the other side are stable nodes. A bifurcational diagram in the (F_i, δ) plane is presented in Fig. 6. Region *III* of the values of the parameters F_i and δ between the two curves corresponds to unstable solutions located in the middle part of the bistable curve $z_s(F_i)$ with a negative slope of the tangent. Regions *I* and *II* of the values of the param-

eters F_i and δ correspond to stable solutions, which in phase plane (x, y) represent singular points of the stable node (region *I*) and stable focus (region *II*) types. In Fig. 7, an example of the behavior of the phase paths of the system of Eqs. (9)–(12) at $|\delta| > \delta_c$ is presented, from which it is seen how the phase paths in the non-steady-state regime converge on a stable node or focus and diverge from a saddle point.

PECULIARITIES OF NON-STEADY-STATE STF TRANSMISSION UNDER EXACT RESONANCE CONDITIONS

We now consider the peculiarities of the non-steady-state transmission of USP. Exact analytical solutions of the systems of Eqs. (9) and (10) or (13) and (14) in general cannot be found except in the particular case of the transmission of a rectangular pulse under exact resonance conditions $\delta = 0$. If it is assumed that at $\tau = 0$ the crystal is in the ground state, that is, there are no excitons, then from (9)–(13) it follows that $x(\tau) = 0$, $F_t''(\tau) = 0$, and $F_r''(\tau) = 0$; namely, under exact resonance conditions, the in-phase component of the medium polarization is exactly zero, as are the imaginary components of the transmitted and reflected pulses, which means that there is no phase modulation of these pulses. If the amplitude of the incident pulse $F_i = F_0 = \text{const}$, then according to (10), the rate of variation of the imaginary component of the medium polarization $y(\tau)$ asymptotically tends from below to the limiting value y_1, which depends only on F_0 and can be determined by solving the cubic equation $y_1(1 + y_1^2) = F_0$. This equation has one real root

$$y_1 = 2^{-1/3}\left\{\left[\left(F_0^2 + \frac{4}{27}\right)^{1/2} + F_0\right]^{1/3} - \left[\left(F_0^2 + \frac{4}{27}\right)^{1/2} - F_0\right]^{1/3}\right\} \quad (26)$$

and two complex conjugate roots. Then the solution of Eq. (10) takes the form

$$F_0\tau = \arctan y - \frac{1}{3y_1^2 + 1}\left\{y_1 \ln\left(1 - \frac{y}{y_1}\right) - \frac{y_1}{2} \times \ln\frac{y^2 + yy_1 + y_1^2 + 1}{y_1^2 + 1} + \frac{3y_1^2 + 2}{\sqrt{3y_1^2 + 4}} \times \left(\arctan\frac{2y + y_1}{\sqrt{3y_1^2 + 4}} - \arctan\frac{y_1}{\sqrt{3y_1^2 + 4}}\right)\right\}, \quad (27)$$

where y_1 is determined by expression (26).

From an examination of (27) and Fig. 8, it is seen that the amplitude of the excitonic wave $y(\tau)$ (normalized density of excitons $z = y^2(\tau)$) first increases rapidly with time then its increase slows down and y asymptotically approaches y_1. The amplitude of the transmitted pulse F_t first rapidly decreases and for long times $\tau \gg \tau_0$ goes to zero, while the amplitude of the reflected pulse increases and approaches F_0 at $\tau \gg \tau_0$. The steady-state regime ($\tau \gg \tau_0$) corresponds to total reflection of the incident pulse. The larger the F_0 value, the faster the steady state sets in. Therefore, under exact resonance conditions, the STF transmits only the leading edge of the incident pulse and almost totally reflects the trailing edge, forming an even shorter pulse in transmission. Qualitatively, this result agrees with the result obtained previously for an STF consisting of two-level atoms [3–7]. The reason is that the field of the pulse propagating

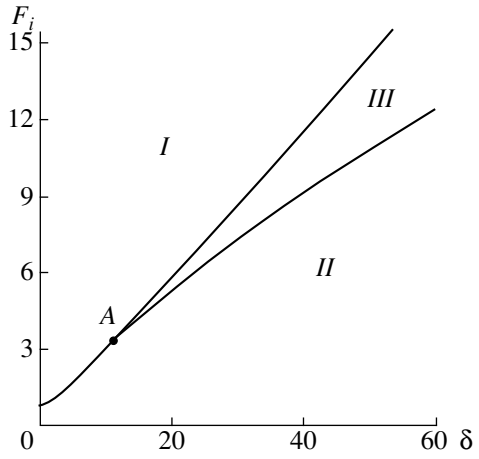


Fig. 6. Bifurcation curves of system (9)–(10) in the (F_i, δ) parameter space separating regions of stable and unstable solutions.

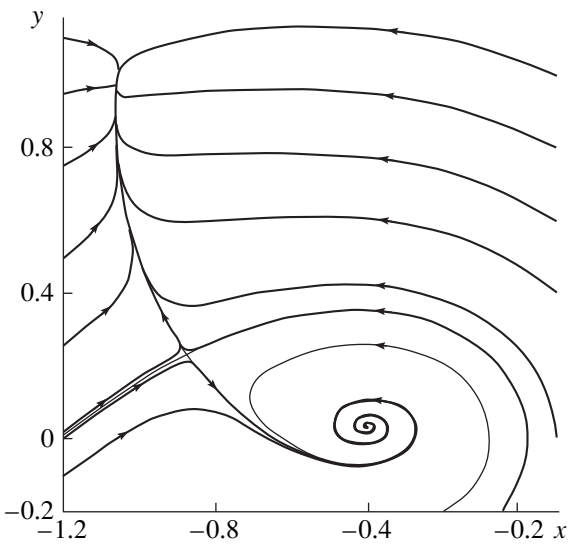


Fig. 7. Example of phase paths of system (9)–(10) in the (x, y) space at $\delta = 24$ and $F_i = 6.5$.

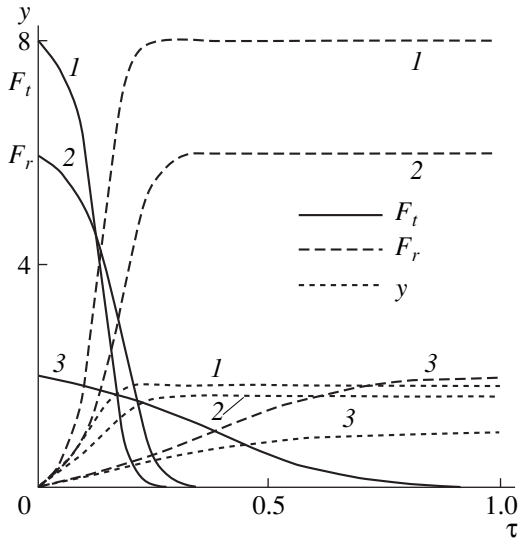


Fig. 8. Temporal evolution of the amplitude of the transmitted F_t and reflected F_r pulses and of the amplitude of the excitonic wave y at F_i = 8 (*1*), 6 (*2*), and 2 (*3*).

through the STF is the sum of the field of the incident pulse and the secondary field induced by the medium polarization, whereas the field of the reflected pulse is formed only by the secondary field. Since the polarization-induced field, whose phase is shifted by π relative to the incident pulse, increases with time and the amplitude of the incident pulse is constant, the field of the transmitted pulse can only decrease.

We now consider transmission of a Gaussian pulse of amplitude F_0 and halfwidth T:

$$F_i(\tau) = F_0 \exp(-\tau^2/T^2). \qquad (28)$$

The results of numerical integration of Eqs. (9)–(12) at δ = 0 are presented in Fig. 9. The transmitted pulse has a complex structure; in fact, it consists of two subpulses. If the length of the incident pulse $\tau_{pulse} \ll \tau_0$ ($T \ll 1$), the leading edge of the transmitted pulse initially follows the shape of the incident pulse; rapidly reaches a maximum; and then decays, goes to zero, changes sign, and forms a second pulse (Fig. 9a). The amplitude and halfwidth of the first subpulse at $T \ll 1$ are essentially equal to the amplitude and halfwidth of the incident pulse, while the second subpulse has a small amplitude and large halfwidth. The reason is that at $T \ll 1$ the transmission time of the pulse is too short for the medium to react, since its polarization lags far behind, affecting only the trailing edge of the incident pulse. Thus, at $T \ll 1$, the first subpulse is formed by the field of the incident pulse, while the second one is formed due to the secondary field being induced by medium polarization. As the halfwidth of the incident pulse increases (at fixed amplitude), the two subpulses will be formed as before; however, the amplitude of the first subpulse will decay much more rapidly than the amplitude of the second subpulse. At $T \geq 1$, with the increase in F_0, the amplitudes of both subpulses slowly increase; and at high F_0, this increase tends to saturate (Fig. 9b). Hence, we can conclude that an effect of power restriction of pulses transmitted through an STF exists.

PECULIARITIES OF NON-STEADY-STATE STF TRANSMISSION FOR NONZERO OFF-RESONANCE

We now consider nonlinear STF transmission for nonzero off-resonance. In this case, the transmitted (reflected) pulse is phase-modulated. In Fig. 10a, the results of a numerical simulation of the temporal evolution of the amplitude $F_t(\tau)$ of the transmitted wave following incidence on an STF of a rectangular pulse with amplitude F_i = 8 and different off-resonance amounts δ are presented. It is seen that at $\delta < \delta_c = 29.86$ (for the specified F_i value), that is, when δ is in range *I* in Fig. 6, the evolution of the transmitted field proceeds as a rapid decrease in its amplitude F_t with time and establishment of the steady state, which is described by Eqs. (15) and (20). The higher the δ value, the higher the steady-

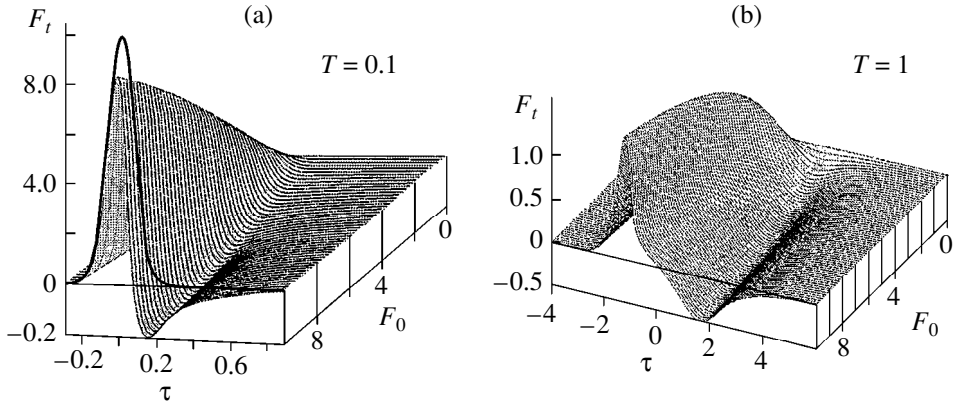


Fig. 9. Non-steady-state transmission of Gaussian pulses of duration T = 0.1 (a) and 1 (b).

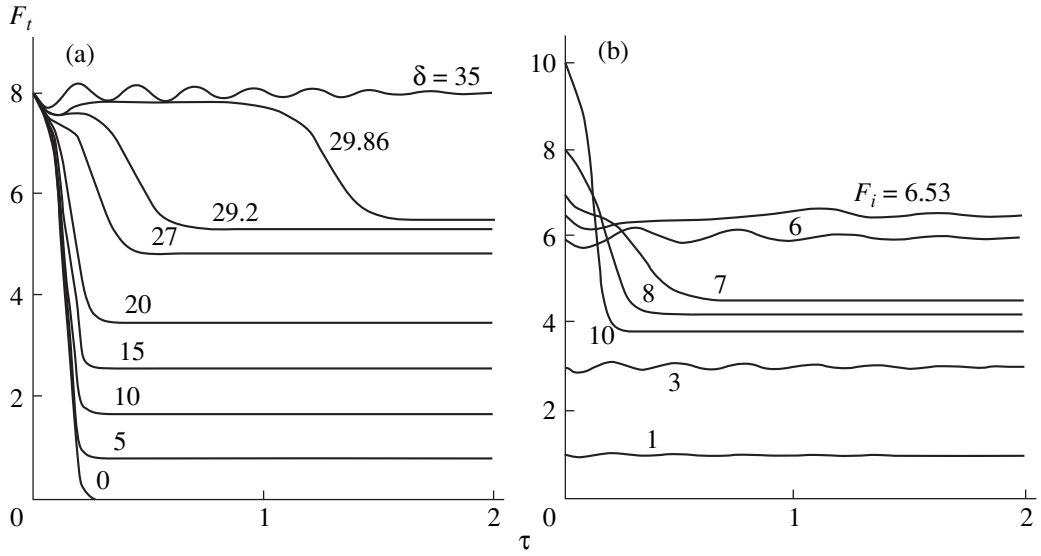


Fig. 10. Temporal evolution of the field amplitude F_t at $F_i = 8$ (a) and $\delta = 24$ (b).

state value of the amplitude F_t, since large off-resonance amounts correspond to the film transparency region. The establishment of the steady-state regime at $\delta < \delta_c$ proceeds in a time τ of the order of $\tau \simeq 0.2$–0.3 ($\tau \simeq (0.2$–$0.3)\tau_0$). For δ approaching δ_c, the time for the establishment of the steady state increases (compare the curves with $\delta = 29.2$ in Fig. 10a). Finally, for $\delta > \delta_c$, the steady state is approached in the oscillation regime, since these δ values correspond to region *II* in Fig. 6, where the stationary points are stable focuses. The system evolution for rectangular pulses of different amplitudes and with fixed off-resonance (that is, for the vertical section in Fig. 6) is shown in Fig. 10b. At low amplitudes F_i corresponding to region *II* in Fig. 6, the steady state is approached in the oscillation regime, whereas at large F_i corresponding to region *I*, a rapid decrease in the transmission takes place with the establishment of the steady-state regime in time $t < \tau_0$. Thus, it is seen that at $\delta \neq 0$ residual STF transmission occurs.

In Fig. 11, temporal evolution plots are presented for the amplitudes of transmitted and reflected pulses, as well as the exciton density in the STF for Gaussian incident pulses (28) with amplitude $F_0 = 10$, for off-resonance amount $\delta = 24$ and different halfwidths *T*. As follows from Fig. 4, this off-resonance amount corresponds to a developed bistability of the steady-state transmission (reflection). It is seen from Fig. 11 that at the first stage the shape of the transmitted pulse essentially coincides with the shape of the incident pulse and the exciton density is vanishingly small. At the time when the amplitude of the incident pulse F_i corresponds to a critical value (at which F_t jumps from the upper to the lower branch of the hysteresis curve in Fig. 4a), the amplitude of the transmitted pulse falls sharply, while the exciton density rises sharply (in accordance with Fig. 1a). Then, with the increase in the incident pulse amplitude, the amplitude of the transmitted pulse decreases slowly and reaches its minimum approximately at the time corresponding to the maximum amplitude of the incident pulse. At this time, the exciton density is also a maximum. With decreasing incident pulse amplitude, the amplitude of the transmitted pulse increases slowly and the exciton density decreases. When the pulse amplitude equals the critical value corresponding to the jump from the lower to the upper branch of the hysteresis curve in Fig. 4a, the transmitted pulse amplitude is rapidly rising and the density of excitons is decreasing. Then, the amplitude of the transmitted pulse decreases in the oscillation regime, which is the more pronounced the smaller the halfwidth *T* of the incident pulse. From the $F_t(F_i)$ and $F_r(F_i)$ curves in Fig. 11, it is seen that considerable deviations from the steady-state hysteresis (Figs. 4a, 4b) arise at small halfwidths of the incident pulse, whereas at large halfwidths (for example, $T = 20$), these curves nearly coincide with the corresponding steady-state curves (Figs. 4a, 4b).

Note that if the amplitude of the incident pulse F_0 is less than the critical one for a given δ, the shape of the transmitted pulse coincides with the shape of the incident pulse, the reflection is very weak, and the transmission function is linear.

THEOREM OF AREAS

It is generally believed that the theorem of areas is an important result in the theory of coherent nonlinear propagation of a USP of resonant laser radiation [15–17]. Following [17], let us derive such a theorem for the excitonic spectral region using the above nonlinearity. Let us write the areas of the transmitted Θ and incident

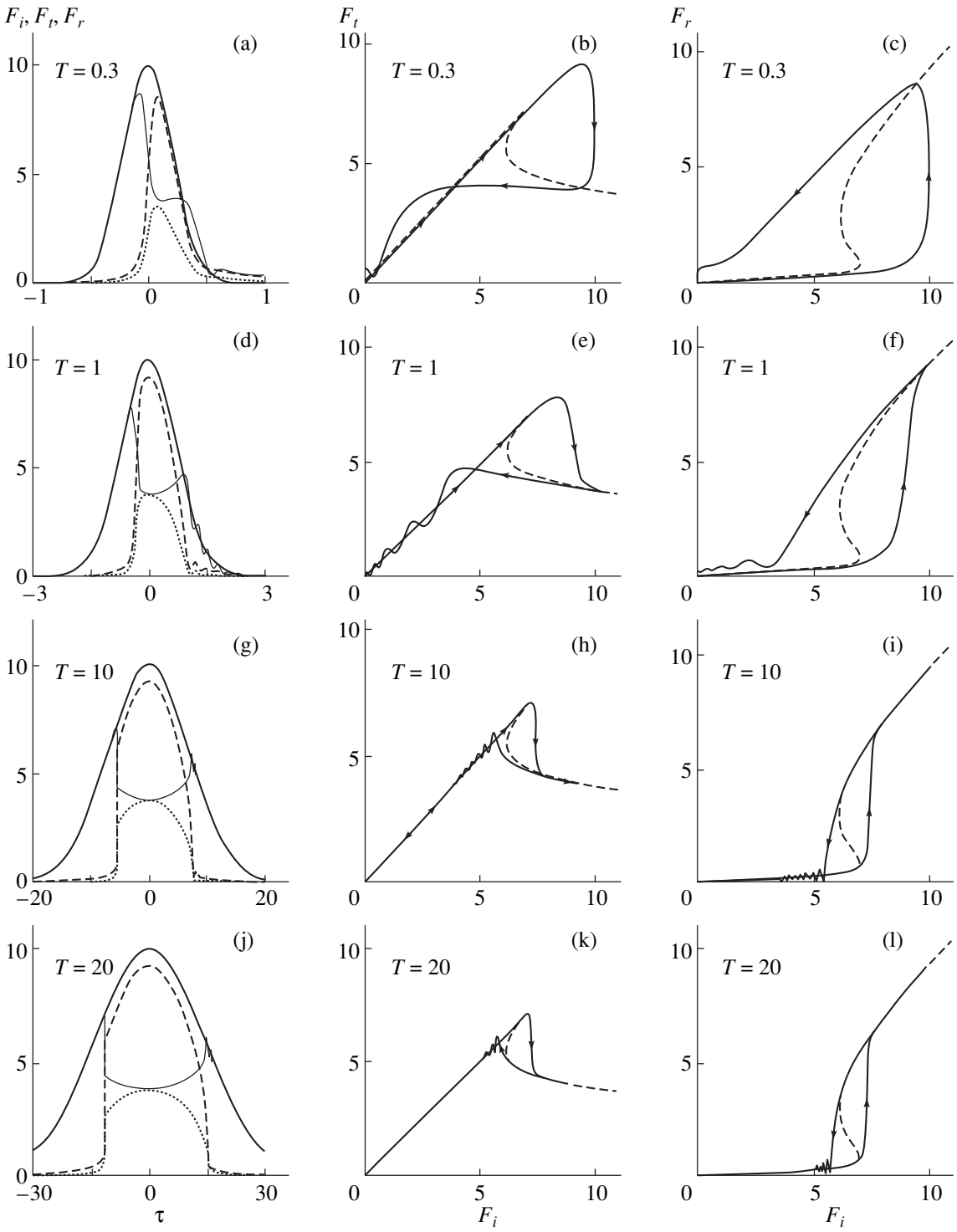


Fig. 11. Temporal evolution of the amplitudes of the transmitted F_t (thin solid curves) and reflected F_r (dashed curves) pulses and the exciton density z (dotted lines) in an STF for Gaussian incident pulses (thick solid curves) of amplitude $F_0 = 10$ and different halfwidths T (a, d, g, j) and the corresponding non-steady-state hystereses in transmission (b, e, h, k) and reflection (c, f, i, l). The dashed curves in figures (b, e, h, k) and (c, f, i, e) correspond to steady-state conditions.

Θ_i pulses in the form

$$\Theta = \sqrt{\mu g}\int_{-\infty}^{\infty} E_t(t)dt, \quad \Theta_i = \sqrt{\mu g}\int_{-\infty}^{\infty} E_i(t)dt. \qquad (29)$$

The duration of the propagating pulses τ_p is assumed to be much less than the exciton relaxation time τ_{rel} and the exciton system to be nonuniformly broadened with the distribution function $f(\Delta)$ with

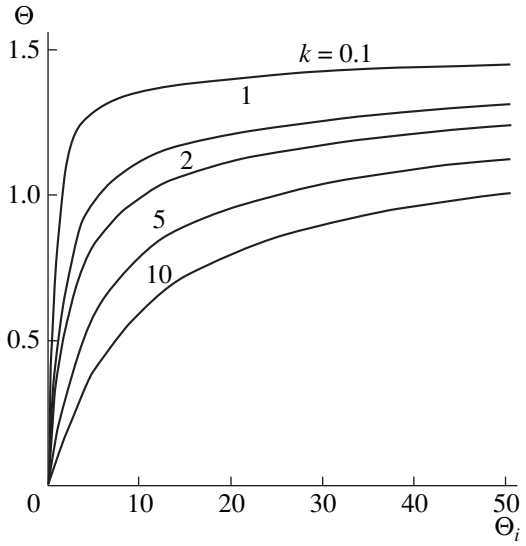


Fig. 12. Dependence of the area Θ of the transmitted pulse on the area Θ_i of the incident pulse at different values of the parameter k.

respect to the off-resonance Δ. Integrating system of equations (9)–(11) with respect to time, the following relationship between the areas of the incident Θ_i and the transmitted Θ pulses (theorem of areas) can be obtained:

$$\Theta_i = \Theta + k\tan\Theta\left(1 + \frac{5}{6}\tan^2\Theta\right), \qquad (30)$$

where $k = \pi f(0)/\tau_0$ and $f(0)$ is the distribution function at $\Delta = 0$.

If it is assumed that $f(\Delta)$ is a Gaussian with the half-width Δ_0, then $k = \sqrt{\pi}\,(\Delta_0\tau_0)^{-1}$. It can be seen from (30) that the area of the transmitted pulse is limited from above $\Theta \longrightarrow \pi/2$ at $\Theta_i \longrightarrow \infty$. Thus, the largest possible area of the transmitted pulse tends to $\pi/2$ from below. The dependence $\Theta(\Theta_i)$ for a series of values of the parameter k is presented in Fig. 12. It is seen that the area of the transmitted pulse initially increases monotonically with the area of the incident pulse Θ_i; however, this increase saturates rapidly. The saturation is slower at larger k. Thus, if the concentration gain of the dipole moment of the exciton transition is taken into account, the STF proves to be an efficient limiter of the area of incident pulses.

CONCLUSIONS

The transmission (reflection) of resonance laser radiation by an STF was investigated in a steady-state regime and in the regimes of ultrashort rectangular and Gaussian pulses, taking into account the exciton–phonon interaction and the exciton transition dipole moment gain at high levels of excitation. It has been shown that under exact resonance conditions an STF transmits only the leading edge of the incident rectangular pulse and almost completely reflects the trailing edge; in this case, the higher the amplitude of the incident pulse, the faster the film becomes "closed" to transmission. For a nonzero off-resonance, the film transmission is oscillatory and rapidly decays, giving way to a residual transmission which coincides in magnitude with the steady-state transmission. In the transmission of Gaussian pulses, there is a wide range of possible transformations depending on the off-resonance, amplitude, and halfwidth. It has been shown that as the halfwidth of the incident Gaussian pulse is increased, the difference between the non-steady-state and the bistable steady-state transmission functions diminishes. The derived theorem of areas predicts limitation of the transmitted pulse area, which can also be concluded from the properties of the non-steady-state transmission; namely, an arbitrarily large increase in the amplitude of ultrashort incident pulses does not result in an appreciable increase in the amplitude of the transmitted pulses.

Along with the studied effects of the transmission of pulses of simple shapes, investigations of the transmission (reflection) of short phase-modulated pulses, taking into account the chirping effect, and of weak pulses with harmonically varying amplitude under continuous pumping by powerful pulses near the switching threshold are of great interest. In this case, a broad spectrum of bifurcations of period doubling and the emergence of chaotic transmission regimes is possible. All these problems will be studied at a later time.

REFERENCES

1. P. I. Khadzhi, *Nonlinear Optical Processes in Exciton and Biexciton Systems in Semiconductors* (Shtiintsa, Kishinev, 1985).
2. P. I. Khadzhi, G. D. Shibarshina, and A. Kh. Rotaru, *Optical Bistability in a System of Coherent Excitons and Biexcitons in Semiconductors* (Shtiintsa, Kishinev, 1988).
3. V. I. Rupasov and V. I. Yudson, Kvantovaya Élektron. (Moscow) **9**, 2179 (1982); Zh. Éksp. Teor. Fiz. **93**, 494 (1987) [Sov. Phys. JETP **66**, 282 (1987)].
4. S. M. Zakharov and É. A. Manykin, Poverkhnost, No. 2, 137 (1988); Poverkhnost, No. 7, 68 (1989); Zh. Éksp. Teor. Fiz. **95**, 800 (1989) [Sov. Phys. JETP **68**, 457 (1989)]; Zh. Éksp. Teor. Fiz. **105**, 1053 (1994) [JETP **78**, 566 (1994)].
5. M. G. Benedict, A. I. Zaĭtsev, V. A. Malyshev, and E. D. Trifonov, Opt. Spektrosk. **68**, 812 (1990) [Opt. Spectrosc. **68**, 473 (1990)].
6. S. M. Zakharov, V. A. Maĭmistov, É. A. Manykin, *et al.*, Poverkhnost, No. 12, 60 (1989).
7. A. M. Samson, Yu. A. Logvin, and S. I. Turovets, Kvantovaya Élektron. (Moscow) **17**, 1223 (1990).
8. Yu. A. Logvin, A. M. Samson, and S. I. Turovets, Kvantovaya Élektron. (Moscow) **17**, 1521 (1990).
9. P. I. Khadzhi and S. L. Gaĭvan, Zh. Éksp. Teor. Fiz. **108**, 1831 (1995) [JETP **81**, 1003 (1995)]; Kvantovaya Élek-

tron. (Moscow) **22**, 929 (1995); **23**, 451 (1996); **23**, 837 (1996); **24**, 532 (1997).

10. A. S. Davydov, *The Physics of the Solid State* (Nauka, Moscow, 1976).
11. F. Ou, B. H. Wei, K. W. Yu, and C. H. Liu, J. Phys.: Condens. Matter **8**, 2957 (1996).
12. I. Abram, Phys. Rev. B **28**, 4433 (1983).
13. P. I. Khadzhi, *The Kinetics of Recombination Radiation of Excitons and Biexcitons in Semiconductors* (Shtiintsa, Kishinev, 1977).
14. S. A. Moskalenko, *An Introduction to the Theory of High-Density Excitons* (Shtiintsa, Kishinev, 1983).
15. A. Allen and J. H. Eberly, *Optical Resonance and Two-Level Atoms* (Wiley, New York, 1975; Mir, Moscow, 1979).
16. I. A. Poluéktov, Yu. M. Popov, and V. S. Roĭtberg, Usp. Fiz. Nauk **114** (1), 87 (1974).
17. V. A. Alekseev and B. Ya. Zel'dovich, Kvantovaya Élektron. (Moscow) **2**, 1078 (1975).

Translated by M. Lebedev

Technical Physics, Vol. 45, No. 10, 2000, pp. 1307–1311. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 76–80.

RADIOPHYSICS

Green's Function for a Pair Electromagnetic Interaction of Undulator Emitters

E. V. Bulyak, V. I. Kurilko†, A. P. Tolstoluzhskiĭ, and I. N. Shapoval

Kharkov Physicotechnical Institute, National Scientific Center, Kharkov, 310108 Ukraine

Received June 16, 1998

Abstract—An analytical representation of the longitudinal Lorentz force describing the electromagnetic interaction of two identical undulator emitters moving along the axis of a helical undulator is obtained under the dipole approximation. A functional dependence of the coherence area volume of the emitters on the external parameters of the system is determined. The applicability of the results and the main aspects of their physical meaning are discussed. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

It is known [1, 2] that the formation of the coherently emitting electron bunches is just what provides a positive feedback necessary for the development of collective instabilities in the electronic amplifiers and self-excited oscillators producing coherent microwave radiation. A longitudinal Lorentz force generated by the field of the natural undulator radiation (NUR) of the electron bunch provides such a grouping in the magnetic bremsstrahlung (undulator) free electron laser (FEL) [3, 4]. If the relative volume density of the bunch is sufficiently high and the frequencies are relatively low, one can use a fluid model of the bunch with neglect of the spatial structure of the Lorentz force, in which case the final theoretical results depend only on the mean volume density in the bunch. On the other hand, the above-mentioned structure eventually determines the number of the coherently emitting electrons in the case of the ultrashort wavelength range where the relative volume density of the ultrarelativistic electron beam (UREB) is relatively low. The number of the coherently emitting electrons determines the rate of the bunch deceleration by its microwave radiation field in the undulator and the parameters of the nonequilibrium UREB–undulator system at which it stops emitting incoherent magnetodeceleration (synchrotron) radiation and starts emitting coherent radiation of the FEL type [5, 6].

This work concentrates on the search for and analytical study of the Green function of the force characterizing a pair electromagnetic interaction of two charges via the field of the microwave NUR of one of them in a helical undulator. Below we demonstrate a possibility of performing such a study in the dipole approximation, provided that the undulator field strengths are relatively low.

† Deceased.

FORMULATION OF THE PHYSICAL PROBLEM AND THE METHODS TO SOLVE IT

The purpose of this work was to establish an analytical relationship between the characteristic volume of the coherence area for the grouping force of an individual emitter (V_{coh}) and the external parameters of the system, specifically, the undulator field amplitude (H_0) and period (D) and the relativistic factor of the emitting charge (γ). For this purpose, we calculate the Lorentz force for a microradiowave interaction of two identical charges moving along the axis of a helical undulator with equal velocities at a fixed distance from each other and interacting via NUR of one of them. Initial expressions for the NUR fields use their classical representation via the Lienard–Wiechert retarded potentials [6, 7]. For the physical interpretation, we use a receptor charge (affected by the microwave field of the emitting neighbor) as a test particle detecting field structure of the emitter radiation.

PROBLEM SOLUTION

1. NUR field of an individual charge in a helical undulator. Consider an ultrarelativistic charge ($\gamma \gg 1$) moving along the axis of a helical undulator. The undulator field is given by the formula

$$\mathbf{H}_w = H_0(\mathbf{e}_x \cos k_\omega z + \mathbf{e}_y \sin k_\omega z), \quad k_\omega \equiv \frac{2\pi}{D}. \qquad (1)$$

Here, $\mathbf{e}_\alpha$ are the unit vectors directed along the Cartesian axes OX ($\alpha = x$) and OY ($\alpha = y$), and k_ω is the undulator wavenumber. An emitter characterized by the total energy $W_{\text{tot}} \equiv m_0 c^2 \gamma_{\text{tot}}$ and the charge q moves along the helical trajectory circumventing the undulator axis:

$$\begin{aligned}\mathbf{R}_e[Z_e(t)] &\equiv \mathbf{e}_x[X_e(0) + a_\omega \sin k_\omega Z_e(t)] \\ &+ \mathbf{e}_y[Y_e(0) - a_\omega \cos k_\omega Z_e(t)] + \mathbf{e}_z Z_e(t),\end{aligned} \qquad (2)$$

1063-7842/00/4510-1307$20.00 © 2000 MAIK "Nauka/Interperiodica"

where $\mathbf{e}_z$ is the unit vector directed along the *OZ* axis, $Z_e(t) \equiv Z_e(0) + V_\parallel^{(0)} t$ is the trajectory of the emitter longitudinal motion, $\mathbf{R}_e(0)$ is the emitter position at the initial moment of the Euler time ($t = 0$), and $V_\parallel^{(0)} \equiv c\sqrt{1-\gamma_\parallel^{-2}}$ is the longitudinal velocity of charge. The corresponding longitudinal relativistic factor $\gamma_\parallel$ is uniquely determined by the total energy γ_{tot} and the dimensionless undulator parameter *K*:

$$\gamma_\parallel^2 \equiv \frac{\gamma_{\text{tot}}^2}{1+K^2}, \quad K \equiv \frac{qH_0D}{2\pi m_0 c^2}.$$

The charge rotation radius in the field (1) also depends on these parameters:

$$a_\omega = \frac{KD}{2\pi\gamma_{\text{tot}}}.$$

Accelerated emitter motion in the field (1) gives rise to a radiation field. Under the ultrashort wavelength approximation ($\gamma \gg 1$) this field can be represented in terms of the Lienard–Wiechert retarded potentials in vacuum (with neglect of the influence of boundaries of the interaction region [7]):

$$\mathbf{E}_{\text{rad}}^{(e)}(\mathbf{r}, t) = \frac{q}{c^2}\frac{[\mathbf{n}'_e, [\mathbf{n}'_e - \boldsymbol{\beta}'_e, \dot{\mathbf{V}}'_e]]}{\mathbf{R}'_e(1-\boldsymbol{\beta}'_e\mathbf{n}'_e)},$$
$$\mathbf{R}'_e \equiv \mathbf{r} - \mathbf{R}_e[Z_e(t')], \qquad (3)$$
$$\mathbf{H}_{\text{rad}}^{(e)}(\mathbf{r}, t) \equiv [\mathbf{n}'_e, \mathbf{E}_{\text{rad}}^{(e)}], \quad \mathbf{n}'_e \equiv \frac{\mathbf{R}'_e}{R'_e}, \quad \beta'_e \equiv \frac{\mathbf{V}'_e}{c}.$$

Here, primed values refer to the moment of the field emission *t'*. This moment is uniquely related to the moment of the field observation *t:*

$$t = t' + \frac{R'_e}{c},$$

A dot at the emitter transverse velocity denotes the derivative with respect to time *t.*

2. The force of the microwave interaction of two identical charges. In the general case (with arbitrary values of the undulator and emitter external parameters) the field (3) has an extremely complicated structure and cannot be studied by analytical methods. In order to explicitly represent this structure, we consider a limiting case of the dipole approximation when $K^2 \ll 1$. For the same purpose (explicit representation of the field structure) the detector is represented by another charge (receptor) possessing the same properties and moving along the trajectory differing from the emitter trajectory (2) only by the initial conditions:

$$\mathbf{R}_r(0) - \mathbf{R}_e(0) \equiv \mathbf{e}_x\rho_{\perp x} + \mathbf{e}_y\rho_{\perp y} + \mathbf{e}_z\rho_\parallel; \quad \rho_\parallel > 0. \quad (4)$$

This method of detection substantially simplifies the structural study of the bunching force generated by the emitter NUR field:

$$F_{\text{rad}}^{(r,e)} \equiv q\left\{\mathbf{E}_{\text{rad}}^{(e)}(\mathbf{R}_r, t) + \frac{1}{c}[\mathbf{V}_r, \mathbf{H}_{\text{rad}}^{(e)}(\mathbf{R}_r, t)]\right\}_z. \quad (5)$$

Indeed, in this case, the right-hand side in Eq. (5) depends only on the longitudinal ($\rho_\parallel$) and transverse ($\rho_\perp \equiv (\rho_x^2 + \rho_y^2)^{1/2}$) distances between the charges at the undulator entrance (but does not depend on time). Expanding the right-hand side in (5) in the powers of a small parameter $K^2 \ll 1$ and introducing new dimensionless variables $x \equiv k_\omega\rho_\parallel\gamma^2$ and $y \equiv k_\omega\rho_\perp\gamma^2$ we arrive at the following representation of the force (5):

$$F_{\text{rad}}^{(r,e)} = -2(qk_\omega K\gamma)^2\Phi(x, y),$$
$$\Phi(x, y) \equiv \frac{1}{2(x^2+y^2)}\{A(x, y)\sin(x + \sqrt{x^2+y^2})$$
$$- B(x, y)[1 - \cos(x + \sqrt{x^2+y^2})]\}, \qquad (6)$$
$$A(x, y) \equiv \sqrt{x^2+y^2}\left[x(x + \sqrt{x^2+y^2}) + \frac{1}{2}y^2\right],$$
$$B(x, y) \equiv \frac{x[2x(x + \sqrt{x^2+y^2}) - y^2]}{2(x + \sqrt{x^2+y^2})}.$$

Note that the magnetic and electric components of the total force (5) yield the first (decelerating) and the second (accelerating) terms in the right-hand side in Eq. (6), respectively.

The spherical symmetry of the problem under consideration in the *x, y* variables is manifested by the corresponding symmetry of the Green function $\Phi(x, y)$, which we assume to be normalized to unity at the $y = 0$ axis at $x \longrightarrow +0$. This symmetry makes it possible to introduce new spherical variables by substituting $x \equiv \rho\cos(2\psi)$ and $y \equiv \rho\sin(2\psi)$. The Green function can be represented in the new variables as:

$$G(\rho, \psi) \equiv \frac{\sin(\rho\cos^2\psi)}{\rho}\left\{2(\cos^4\psi)\cos(\rho\cos^2\psi)\right.$$
$$\left. - (\cos 2\psi)[\cos 2\psi - \sin^2\psi]\frac{\sin(\rho\cos^2\psi)}{\rho}\right\}. \qquad (7)$$

ANALYSIS OF THE SPATIAL STRUCTURE OF THE $G(\rho, \Psi)$ FUNCTION

As defined above, the parameter $\rho \equiv \sqrt{x^2+y^2}$ characterizes a dimensionless distance between the receptor and the emitter, whereas vector ρ makes an angle

$\Theta \equiv 2\Psi$ with the positive semiaxis of the undulator ($x > 0$, $y = 0$). At small values of the above distance ($\rho^2 \ll 1$), the corresponding asymptotics $G_0(\Psi) \equiv G(0, \Psi)$ is given by the formula:

$$G_0(\Psi) = [2\cos^2\Psi - \cos 2\Psi(\cos 2\Psi - \sin^2\Psi)]\cos^4\Psi. \quad (8)$$

It follows from this formula that, as the angle Θ modulus increases from $\Theta_{min} = 0$ to $\Theta_{max} = \pi/2$, the $G_0(\Psi)$ amplitude initially increases (up to $G_{max} \approx 1.024$ at $x \approx 0.29$) and then monotonically decreases from about unity (at $x \approx 0.3$) to 1/4 (at $\Theta = \pi/2$) (Fig. 1). The modulus of the right-hand side in (7) is inversely proportional to ρ and decreases at $\rho > 1$. This decrease is accompanied by the oscillations in G. The rate of these oscillations reaches a maximum at the emitter axis ($\Theta \longrightarrow 0$), where the period equals 2π, and exhibits a twofold decrease in the transverse direction (i.e., at $\Theta \longrightarrow \pi/2$, see Fig. 2). It is this decrease in the field of the magnetic retardation dipole radiation of the emitter in its far-field zone that appears to be the reason of a decrease in the coherency in the UREB electrons deceleration by the NUR field of these electrons at relatively small values of the bunch volume density ($Q \ll 1$; [5, 6]). In the latter limiting case ($Q \ll 1$) the nonequilibrium UREB–undulator system represents a source of incoherent NUR of UREB electrons. Equating the right-hand side in Eq. (7) to zero, we obtain a curve that separates the positive and the negative values of the $G(\rho, \Psi)$ function in the band ($0 \le \rho \le \pi$; $0 < \Psi < \pi/4$). In this curve, the expression in braces in the right-hand side of Eq. (7) becomes equal to zero. Physically, this curve implies equality of the accelerating and decelerating forces produced by the emitter electric and magnetic fields, respectively (see the right-hand side in Eq. (5)). Figure 3 shows equilevel lines of the $G(\rho, \Psi)$ surface over the aforementioned band. Once the shape of the $G(\rho, \Psi)$ surface is known, one can calculate the volume of the coherency domain of the bunching force produced by the emitter.

This volume is uniquely determined by the integral:

$$V_{coh} \equiv 2\pi\int \rho_\perp d\rho_\perp d\rho_\parallel G(\rho_\perp, \rho_\parallel) = \frac{D\lambda^2}{\pi^2}\int G(\rho, \Theta/2)\rho^2 d\rho \sin\Theta d\Theta = D\lambda^2 f_{coh}, \quad (9)$$

$$f_{coh}(G_{min}) \equiv \pi^{-2} \int\limits_{G \le G_{min}} G(\rho, \Theta/2)\rho^2 d\rho \sin\Theta d\Theta.$$

Note that the integration over the dimensionless length ρ and the angle Θ must be performed only over the area adjacent to the point $\rho = 0$ where $G(\rho, \Theta/2)$ is not only positive but is larger than a positive constant value G_{min}.

It is seen from the right-hand side in Eq. (9) that the product of the square of the NUR wavelength ($\lambda \equiv D/2\gamma^2$) and the undulator period (D) determines the functional dependence of the coherency volume, that is of the domain where the emitter bunching force depends on the external parameters of the system. The form factor f_{coh} is a functional of the admissible value of the decrease in coherency of emitters:

$$\delta G \equiv 1 - G_{min} > 0. \quad (10)$$

Figure 4 shows a plot of the form factor f_{coh} versus the quantity G_{min}.

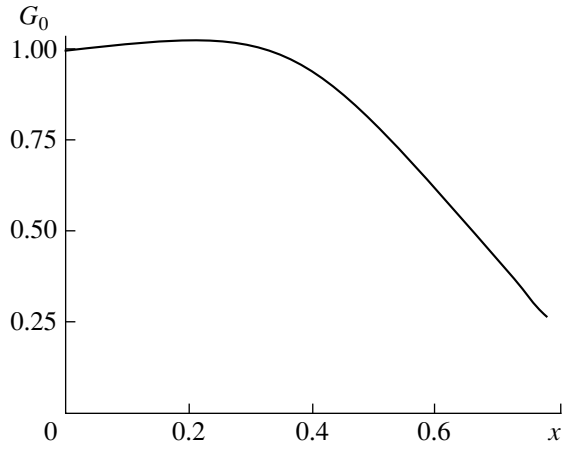


Fig. 1. The plot of G_0 versus $\Psi \equiv \Theta/2$.

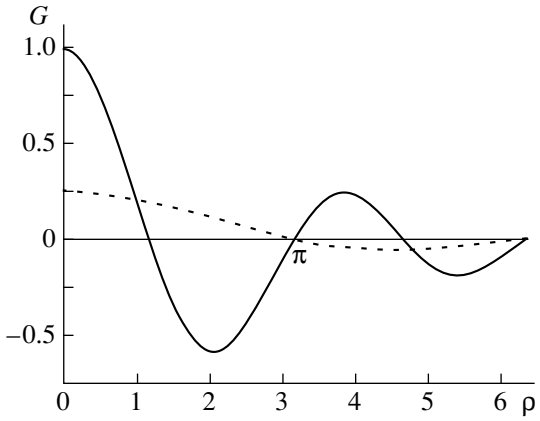


Fig. 2. Plots of the $G(\rho, \Psi)$ function versus ρ for $\Psi = 0$ (solid line) and $\pi/4$ (dashed line).

DISCUSSION

The main results are as follows:

(1) We proposed and realized an analytical method allowing one to describe the spatial structure of the bunching force produced by the NUR field of an ultrarelativistic electron in a helical undulator. The essence of the method lies in the explicit analytical pre-

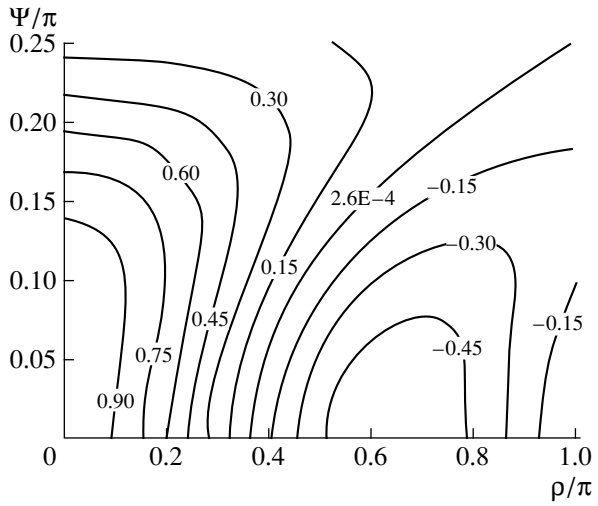


Fig. 3. Equilevel lines of the $G(\rho, \Psi)$ function over the band $(0 \le \rho < \pi;\ 0 \le \Psi \le \pi/4)$.

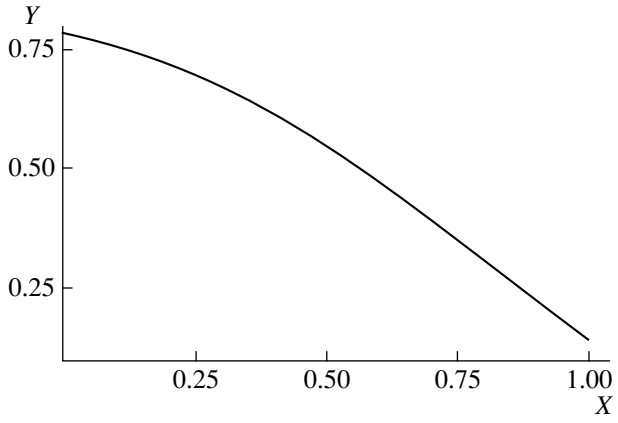


Fig. 4. The plot of form factor f_{coh} versus G_{min} ($X = G_{min}$ and $y = 2nf_{coh}$).

sentation of the emitter field (in the dipole approximation) using the Lienard–Wiechert retarded potentials [7] and the detection of this field by a receptor that is identical to the emitter.

(2) The proposed method allowed us not only to analytically determine an explicit functional dependence of the bunching force on the length and orientation of the **ρ** vector interconnecting receptor and emitter but also to calculate the dependence of the characteristic coherence volume of this force on the external parameters of the emitter–receptor NUR (see formula (9)).

Proceeding to the discussion of these results, note, first of all, the square decrease in the coherence volume with decreasing NUR wavelength. This decreasing hinders the advent of the X-ray FELs into the superhard X-ray range ($\lambda \ll 10^{-10}$ m). Specially note the independence of the rate of this decrease of the undulator parameter K. To additionally prove this, one can estimate the emitter coherence volume by a theoretical model for a traveling wave tube (TWT) using the same bunch ($\gamma_\parallel$) and radiation (λ) parameters. Indeed, equations for the spectra of the TWT and FEL natural wavelengths are similar to each other (see, e.g., [3, 6, 8]). This similarity allows one to qualitatively consider the NUR field coherence volume of an individual emitter in FEL as a cylinder of the height $H_{coh} = \lambda/2$ and the radius $R_{coh} = \gamma_\parallel\lambda/2\pi$ [9]:

$$V_{coh}^{(TWT)} \equiv \pi R_{coh}^2 H_{coh} = \frac{\lambda^2 D}{16\pi} = D\lambda^2 f_{coh}^{(TWT)},$$
$$f_{coh}^{(TWT)} \equiv \frac{1}{16\pi} \approx 2 \times 10^{-2}. \quad (11)$$

It follows from this formula that the coherence volume is really independent of the undulator parameter K. In this respect, formula (11) is analogous to the exact formula (9) obtained in the dipole approximation. The only difference between them lies in the value of the numerical factor (form factor f_{coh}). One can estimate this value in the following way. Strictly speaking, the coherence volume (9) must correspond to a sufficiently uniform amplitude of the microwave interaction force, which implies a small value of the right-hand side in Eq. (10). Assuming $\delta G = 0.30$ ($G_{min} = 0.70$), which corresponds to the increase in the Thomson scatterers bunch size up to a quarter wavelength (in the bunch-centered coordinates [6, 10]), we arrive at the form factor value that is close to the above estimate based on the analogy between FEL and TWT: $f_{coh} = 0.04$.

Further comments refer to the bunching force amplitude $F_{rad}^{(r,e)}$. It follows from the right-hand side in Eq. (6) that this force exhibits the same dependence on the external parameters of the emitting charge and undulator as the deceleration force produced by the emitter radiation (see, e.g., [7, 11]). The only difference lies in the value of the resulting numerical factor that is three times higher in the case of the Eqs. (6) and (7) than in the case of the emitter deceleration due to its intrinsic magnetic retardation [7, 11]. The main reason of this difference lies in the fact that, in deriving the analytical asymptotics (6) and (7), we have assumed the smallness of the dimensionless amplitude of the emitter and receptor rotational motion in comparison with the modulus of the distance between the emitters: $\rho \gg K$. At the same time, it is well known [7] that the radiation-induced decelerating force is determined by the radiation field structure just at a small distance from the emitter (i.e., at $\rho \le K$). Since the purpose of this work was to study the spatial structure of the force (6) at $\rho^2 \gg K^2$, the aforementioned difference is not essential.

To conclude this section, let us consider in brief the applicability of the results and the novelty of the proposed approach.

Note, first of all, that the inequality $\rho^2 \gg K^2$ provides for a small value of the Coulomb force

$$F_{z\text{coul}}^{(r,e)} = \frac{(q\gamma k_\omega)^2}{(x^2+y^2)^{3/2}}x$$

as compared to the force of the microwave interaction of the same charges (6). Thus, the Coulomb field can be neglected in the right-hand side in (6) as a small correction in the radiation field (5).

Note also that a classical description of the NUR field by the Lienard–Wiechert potentials is valid only in the limiting case when the number of the magnetic bremsstrahlung quanta emitted by the charge at the undulator half-period (where the emitter trajectory curvature is still finite)

$$N_{\hbar\omega} \equiv \frac{(qk_\omega K\gamma)^2 D}{3\hbar\omega} = \frac{\pi}{3}\frac{q^2}{\hbar c}K^2(1+K^2) \qquad (12)$$

is much greater than unity (see, e.g., [11] for the details). If the latter condition is not met one has to take into account a random character of the quanta emission and absorption by the moving charges.

Note that condition (12) can be valid even in the dipole approximation if the emitter represents a bunch containing a rather large number of electrons

$$Q \gg \min Q(N_{\hbar\omega}) \equiv \left|\frac{q}{e}\right| = \sqrt{\frac{3N_{\hbar\omega}}{\pi K^2(1+K^2)}\frac{\hbar c}{e^2}} \gg 1,$$

where e is the electron charge.

Finally, Kurilko and Ognivenko [12] simulated the electron bunch NUR field using the Lienard–Wiechert retarded potentials. The authors described a self-consistent bunch dynamics in the NUR field of its particles (electrons) restricting the consideration to the dipole approximation. Abandoning the dipole approximation allows wider application of the results but the consequent necessity in the numerical simulation implies the loss of their physical transparency (cf., e.g., [12–14]).

ACKNOWLEDGMENTS

We are grateful to V.A. Buts, N.I. Aĭzatskiĭ, A.N. Lebedev, and K.N. Stepanov for fruitful discussions. This work was supported by the Ukrainian Scientific-Technological Center, contract no. 279.

REFERENCES

1. Ya. B. Faĭnberg, At. Énerg. **11**, 313 (1961).
2. A. V. Gaponov, M. I. Petelin, and V. K. Yulpatov, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **10**, 1414 (1967).
3. Ph. Sprangle and R. A. Smith, Phys. Rev. A **21**, 293 (1980).
4. Th. C. Marshall, *Free-Electron Lasers* (McMillan, New York, 1985; Mir, Moscow, 1987).
5. V. I. Kurilko and Yu. V. Tkach, Nucl. Instrum. Methods Phys. Res. A **282**, 431 (1989).
6. V. I. Kurilko and Yu. V. Tkach, Usp. Fiz. Nauk **165**, 241 (1995) [Phys. Usp. **38**, 231 (1995)].
7. L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (Fizmatgiz, Moscow, 1960; Pergamon, Oxford, 1975).
8. L. A. Vaĭnshteĭn and V. A. Solntsev, *Lectures on Microwave Electronics* (Sov. Radio, Moscow, 1973).
9. E. V. Bulyak and V. I. Kurilko, Dokl. Nats. Akad. Nauk Ukrainy, No. 7, 82 (1998).
10. V. I. Kurilko and V. V. Ognivenko, Fiz. Plazmy **20**, 475 (1994) [Plasma Phys. Rep. **20**, 426 (1994)].
11. V. A. Bazylev and N. K. Zhivago, *Radiation of Fast Particles in Matter and External Fields* (Nauka, Moscow, 1987).
12. V. I. Kurilko and V. V. Ognivenko, Dokl. Akad. Nauk **335**, 437 (1994) [Phys. Dokl. **39**, 214 (1994)]; Fiz. Plazmy **20**, 633 (1994) [Plasma Phys. Rep. **20**, 568 (1994)]; Zh. Tekh. Fiz. **66** (10), 182 (1996) [Tech. Phys. **41**, 1068 (1996)].
13. M. Tecimer and L. Elias, in *Proceedings of the Conference on Particle Accelerators, Dallas, 1995*, Vol. 1, p. 254.
14. L. Elias, in *Proceedings of European Particle Accelerator Conference (EPAC-96), Barselona, 1996*, Vol. 1, p. 724.

Translated by A. Chikishev

Technical Physics, Vol. 45, No. 10, 2000, pp. 1312–1319. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 81–89.

ELECTRON AND ION BEAMS, ACCELERATORS

Construction of Optimum Profiles for Multipole Lenses

V. V. Vecheslavov and O. V. Loginova

Budker Institute of Nuclear Physics, Siberian Division, Russian Academy of Sciences, pr. Akademika Lavrent'eva 11, Novosibirsk, 630090 Russia

E-mail: vecheslavov@inp.nsk.su

Received September 29, 1999

Abstract—A flexible and easy to implement method is proposed for computing the multipole-lens potential and field. It integrates conformal-mapping technology and numerical approaches. A procedure for the sequential optimization of the geometric pole shape that is aimed at improving the lens characteristics is described and demonstrated with examples. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The discovery of the strong-focusing principle [1] has initiated a wide application of quadrupole (pole pairs $P = 2$) lenses in the acceleration technology for forming and transporting charged-particle fluxes. Along with quadrupoles, sextupoles ($P = 3$) and octupoles ($P = 4$) are currently used. Let a lens with the number of pole pairs P be called a P lens. The cross sections of quadrupoles and sextupoles and the coordinate system used are presented in Fig. 1.

The quality of the field produced by a lens depends on the pole geometry and the location of pole-face winding. Each P lens is known to have the theoretical profile, providing the "ideal" field: linear, quadratic, and cubic for a quadrupole, sextupole, and octupole, respectively (see Section 2 and Appendix 3). However, these profiles cannot be achieved in practice, since they leave no room for placing the winding and closing the magnetic flux. For this reason, fields produced by actual lenses always contain higher (as a rule, undesirable) harmonics together with the fundamental component.

Numerous works describe the search for "good" profiles (see References in [2, 3]). In most cases, the efforts are aimed at constructing an adequate approximation (e.g., a segment of a circle) of the "broken" ideal profile. The first successful attempt to radically change the approach seems to be made in [4]. The pole was bounded by a flat, instead of a hyperbolic, surface, and its size was selected such that the sixth harmonic was absent in the lens-potential spectrum.

In recent years, tolerances for the field deviation from ideal have become much more stringent, and the determination of optimum profiles providing a desired field while minimizing the lens size has become topical. The problem is conventionally solved with the algorithms of numerical field-pattern construction on sufficiently fine meshes, which requires high-perfomance computers and much computer time. These costs are fully justified at the final design stage; however, for initial rough estimations and comparative analysis of many designs, it is desirable to use simple and fast analytic methods.

We propose a flexible and simple method integrating conformal-mapping techniques and numerical approaches. By using it, it is possible to design lenses with any number of pole pairs and high field quality within the working aperture. The algorithms described below are applicable to both magnetic and electrostatic lenses. We will consider magnetic lenses because of their use in high-energy physics. The conditions conventional for analytical methods (long lens, full symmetry, and negligible iron saturation) are assumed to be satisfied.

1. MAPPING OF THE UPPER SEMIPLANE ONTO A LENS SECTOR

A sector of a P lens (the partial cross section that belongs to one pole; Fig. 1) is approximated by a polygon $\mathcal{A}$ with the total number of vertices $2M + 1$ (Fig. 2), onto which the upper semiplane of an auxiliary complex variable $w = u + iv$ is conformally mapped.

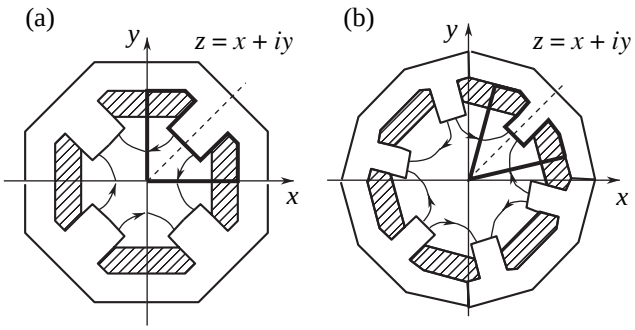


Fig. 1. Cross sections of (a) a quadrupole and (b) a sextupole. Heavy lines show the sector boundaries.

1063-7842/00/4510-1312$20.00 © 2000 MAIK "Nauka/Interperiodica"

The mapping is effected with the Christoffel–Schwartz integral [5]

$$z(w) = C_0 \int_{w_0}^{w} \prod_{m=1}^{2M+1} (w - a_m)^{\beta_m} dw + C_1, \qquad (1)$$

where $\beta_m = \alpha_m - 1$, α_m are the interior (relative to the domain under study) angles in terms of π, and a_m are points on the real number axis u on the w plane that are the images of the A_m-polygon vertices.

The images of three vertices may be assigned arbitrarily, and it is convenient to set up the following correspondence:

$$z = A_1 \longrightarrow w = a_1, \quad z = A_{-1} \longrightarrow w = -a_1,$$
$$z = A_{M+1} = 0 \longrightarrow w = \infty.$$

The constant C_1 is responsible for displacing the polygon as a unit in the z plane. For $C_1 = \exp(i\pi/4)$, the point $u = a_0 = 0$, $v = 0$ becomes the image of the pole center $x = y = 1/\sqrt{2}$ (this image can be considered as the zero vertex of an A_0 polygon with the power $\beta_0 = 0$).

Taking into account the symmetry of the domain about the bisectrix of the first quadrant and in view of the normalizations performed, Eq. (1) can be recast in a much simpler form

$$z(w) = C_0 \int_{w_0}^{w} \prod_{m=1}^{M} (w^2 - a_m^2)^{\beta_m} dw + C_1. \qquad (2)$$

The length of the side I_k between the vertices A_k and A_{k+1} is defined as

$$I_k = |C_0| \int_{a_k}^{a_{k+1}} \prod_{m=1}^{M} |u^2 - a_m^2|^{\beta_m} du; \qquad (3)$$

hence, up to similarity, the desired polygon can be described by specifying $M - 1$ relative lengths of its sides:

$$\lambda_k = I_k / I_0, \quad k = 1, 2, \ldots, M-1. \qquad (4)$$

Note that integrals (3) may have singularities at one or both extremes of the interval. They, if any, can be eliminated by making appropriate changes of variables (see Appendix 1).

To find the values of the still unknown parameters $a_2, a_3, \ldots, a_M$, in the domain $\mathcal{A}$ is the central and, in the general case, rather intricate problem governing the success in constructing mapping function (2). For this purpose, we use the procedure of successive approximations that was described in detail in [6].

We select arbitrarily the initial values of the free parameters $a_2^0, a_3^0, \ldots, a_M^0$ obeying the inequality $a_1 < a_2^0 < a_3^0 \ldots < a_M^0 < \infty$, which follows from the path-tracing rule, and integrate (2) along the real semiaxis $0 \le u < \infty$. Then (due to the symmetry), we come (on the plane z) to a $(2M+1)$-gon $\mathcal{A}^0$ with vertices at the points $A_1^0, A_{-1}^0, A_2^0, A_{-2}^0, \ldots, A_M^0, A_{-M}^0, A_{M+1}^0$. If the constant C_0 in (2) is selected such that the segments $A_0A_1^0$ and A_0A_1 coincide, like sides of the polygons $\mathcal{A}^0$ and A will be parallel to each other, but their relative lengths λ_k^0 and λ_k may diverge as much as desired.

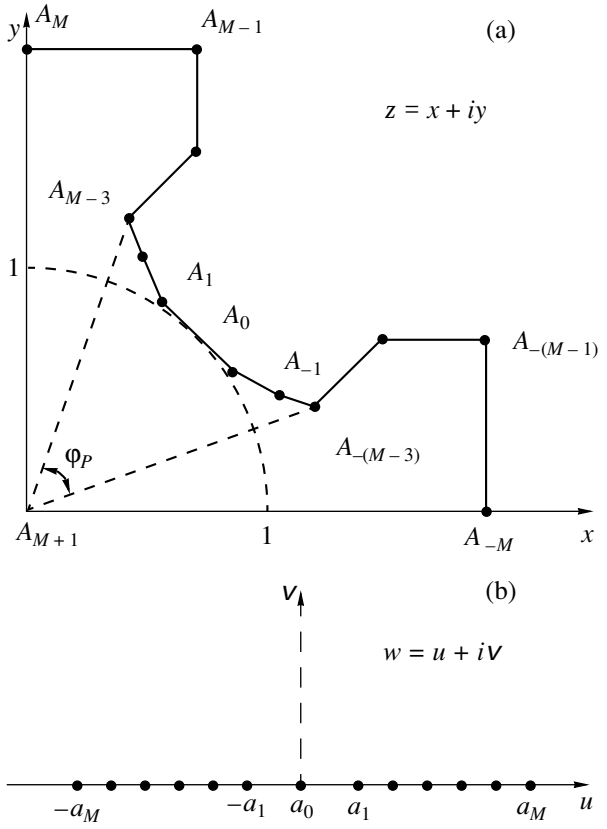


Fig. 2. Correspondence between the sector vertices and their images at conformal mapping.

Let us call the scalar quantity

$$G = \max\left|(\lambda_k^0 - \lambda_k)/\lambda_k\right|; \quad k = 1, 2 \ldots (M-1) \qquad (5)$$

the "distance" between the figures $\mathcal{A}^0$ and $\mathcal{A}$.

The construction of image (2) is considered to be completed when $G \simeq 0$.

If the distance between $\mathcal{A}^0$ and $\mathcal{A}$ is large, the approximation procedure should be partitioned into several stages. As shown in [6], for any value of the parameter K_λ from the interval [0; 1], there exists a polygon $\mathcal{A}^1$ between $\mathcal{A}^0$ and $\mathcal{A}$ with vertices at the points $A_1^1, A_{-1}^1, A_2^1, A_{-2}^1, \ldots A_M^1, A_{-M}^1, A_{M+1}^1$ and the relative side lengths $\lambda_k^1 = \lambda_k^0 - K_\lambda(\lambda_k^0 - \lambda_k)$ (at $K_\lambda = 0$, it coincides with the original polygon $\mathcal{A}^0$, and at $K_\lambda = 1$,

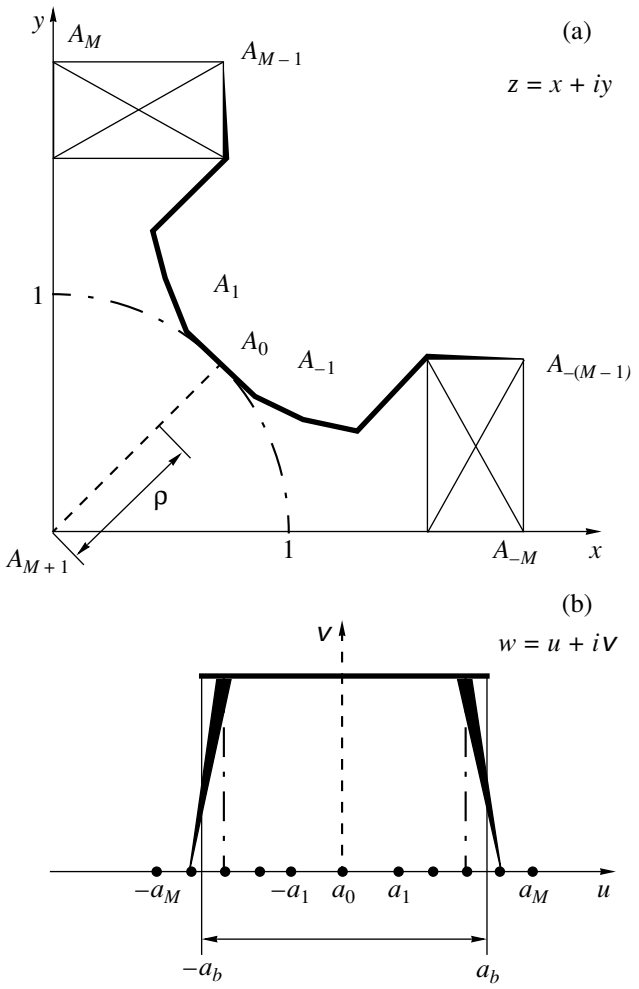


Fig. 3. Distribution of the potential along the domain boundaries.

with the desired polygon $\mathcal{A}$). At each of the stages, a transfer is made from $\mathcal{A}^0$ to this intermediate polygon $\mathcal{A}^1$. With this purpose in mind, a linear system of equations

$$\sum_{j=2}^{M} \frac{\partial \lambda_k^0}{\partial a_j^0}(a_j - a_j^0) = K_\lambda(\lambda_k - \lambda_k^0); \quad k = 1, 2\ldots, M-1 \tag{6}$$

is set up, and its solutions are sought for a series of K_λ values starting from unity (e.g., each next value of K_λ is taken to be half as large as the previous one). This process is terminated as soon as the condition $G(\mathcal{A}^1) < G(\mathcal{A}^0)$ is satisfied [see Eq. (5)]. The intermediate state thus found becomes the initial for the next stage. Note that the derivatives appearing in system (6) may be found either numerically or using the analytical relations in Appendix 2.

If the number of free polygon vertices is not small ($M > 5$), the convergence of this multiparametric optimization strongly depends on the proper selection of the starting values for the images of these vertices. The following technique proved to be useful. For any ideal *P* lens, its exact conformal map on the upper semiplane *w* can easily be constructed (the relevant formulas are given in Appendix 3). As starting values, we took the images of the ideal-profile points obtained with formula (3.6) in Appendix 3 (however, the final lens profile may radically differ from the ideal one; see Section 3).

2. DISTRIBUTIONS OF THE MAGNETIC POTENTIAL AND INDUCTION

The complex magnetic potential of a *P* lens can be expressed as

$$\mathcal{P}(z) = \mathcal{D}(z) + i\mathcal{F}(z) = p_0 z^P\left(1 + \sum_{n=1}^{\infty} \tilde{p}_n z^{2nP}\right), \tag{7}$$

where $\mathcal{D}(z)$ and $\mathcal{F}(z)$ are the vector and scalar magnetic potentials, respectively [2].

We prefer to deal with the scalar potential $\mathcal{F}(z)$. Its distribution over the quadrupole sector is shown in Fig. 3a. On the pole surface, $\mathcal{F}(z) = 1$ (heavy line). It tends to zero near the winding (line of varying thickness) and equals zero along the remainder of the boundary (fine lines).

The conformal map constructed in Section 1 allows us to transfer this distribution to the real axis *u* on the plane *w* (Fig. 3b) so that the relation between the value of the potential and the line thickness (Fig. 3a) holds. Even rough calculations demonstrated that the winding region is very small (<1% of the length of the region with $\mathcal{F}(u) = 1$). This fact allows us to replace the true distribution $\mathcal{F}(u)$ by its rectangular equivalent. Eventually, we have

$$\begin{aligned} \mathcal{F}(u) &= 1 \quad \text{for} \quad -a_b \le u \le a_b, \\ \mathcal{F}(u) &= 0 \quad \text{for any other } u. \end{aligned} \tag{8}$$

The potential at any point of the upper semiplane is found from the known Schwartz formula. In our case, it yields

$$\mathcal{F}(u, v) = \frac{v}{\pi}\int_{-\infty}^{\infty} \frac{\mathcal{F}(\bar{u})d\bar{u}}{(u-\bar{u})^2 + v^2} = \frac{v}{\pi}\int_{-a_b}^{a_b} \frac{d\bar{u}}{(u-\bar{u})^2 + v^2} = \frac{1}{\pi}\left[\tan^{-1}\left(\frac{a_b - u}{v}\right) + \tan^{-1}\left(\frac{a_b + u}{v}\right)\right]. \tag{9}$$

We remind that mapping (2) converts the length A_0A_{M+1} on the line of sector symmetry (Fig. 3) to the imaginary semiaxis $0 \le v < \infty$. Then, potential distribution (9) along this semiaxis takes the simple form

$$\mathcal{F}(v) = \frac{2}{\pi}\tan^{-1}\left(\frac{a_b}{v}\right). \tag{10}$$

Equation (9) becomes basic in what follows. We will begin with the induction distribution within the segment A_0A_{M+1} (Fig. 3). The distance to the lens center along this segment is

$$\rho = 1 - |C_0|\int_0^v \prod_{m=1}^{M} (v^2 + a_m^2)^{\beta_m} dv. \quad (11)$$

Differentiating with respect to ρ, we find the relation for the dependence $v(\rho)$:

$$\frac{dv}{d\rho} = -\frac{1}{|C_0|}\prod_{m=1}^{M} (v^2 + a_m^2)^{-\beta_m}. \quad (12)$$

The distribution of the induction within the segment A_0A_{M+1} is given by the formula

$$\mathcal{B}(\rho) = \frac{d\mathcal{F}}{d\rho} = \frac{d\mathcal{F}}{dv}\frac{dv}{d\rho} = \frac{2}{\pi}\frac{a_b}{v(\rho)^2 + a_b^2}\frac{1}{|C_0|}\prod_{m=1}^{M} (v(\rho)^2 + a_m^2)^{-\beta_m}. \quad (13)$$

It immediately follows from (13) that the exact value of the induction at the pole center ($v = 0$) is

$$\mathcal{B}_{\rho=1} = \frac{2}{\pi}\left(\frac{1}{a_b|C_0|}\right)\prod_{m=1}^{M} a_m^{-2\beta_m}. \quad (14)$$

Since $v \longrightarrow \infty$ at $\rho \longrightarrow 0$, it is possible to determine the induction near the lens center with Eq. (13):

$$\mathcal{B}_{\rho\to 0} = \frac{2a_b}{\pi|C_0|}v^{(1-P)/P} = \frac{2a_b}{\pi(|C_0|P)^P}P\rho^{P-1}. \quad (15)$$

This formula was obtained using the geometric equality

$$2\sum_{m=1}^{M}\beta_m = -2 - \beta_{M+1} = -(P+1)/P$$

and the relation

$$\frac{1}{v} = \left(\frac{\rho}{|C_0|P}\right)^P, \quad (16)$$

which follows from (11) when $v \longrightarrow \infty$.

Equations (15) and (7) help in computing the accurate coefficient of the fundamental harmonic for the potential

$$p_0 = \frac{2}{\pi}a_b\alpha_0, \quad (17)$$

where

$$\alpha_0 = (P|C_0|)^{-P}. \quad (18)$$

As shown in Appendix 4, the knowledge of p_0 and α_0 is the necessary and sufficient condition for the construction of a recurrent analytical procedure to find the rest of the coefficients $\tilde{p}_n$, $n = 1, 2, \ldots, N$ in Eq. (7) (the construction of the potential spectrum) up to any desired order N.

This remark could have concluded this section, because the known potential spectrum allows us to answer any question about the field formed by this lens. However, the computation of the spectrum is a rather laborious procedure (Appendix 4). Moreover, in our case, it is written in the REDUCE language, while the other package programs are in PASCAL, and their integration in the continuous computational process has failed.

For this reason, we searched for the field distributions $\mathcal{B}(\rho)$ and $\mathcal{B}(s)$ within the segments of unit length that passed from the lens center along the line of symmetry and the left boundary of the sector, respectively. The distribution providing the maximum deviation from the basic field was taken into account.

The computation of $\mathcal{B}(\rho)$ is performed by jointly using Eqs. (12) and (13). Similar dependences can be obtained for the construction of $\mathcal{B}(s)$ (coordinate s is reckoned from the lens center along the boundary of the sector; $s \equiv y$ for a quadrupole). To obtain them, the left boundary of the sector is mapped onto the positive semiaxis $a_M < u < \infty$, $v = 0$, where only the v component of the induction is other than zero. This component is computed through potential (9):

$$\mathcal{B}_v(u, 0) = -\lim_{v\to 0}\frac{\partial\mathcal{F}}{\partial v} = -\frac{2a_b}{\pi}\frac{1}{u^2 - a_b^2}.$$

This equality makes it possible to construct the distribution of the induction at the sector boundary:

$$\mathcal{B}(s) = \mathcal{B}_v(u, 0)\frac{du}{ds} = -\frac{2a_b}{\pi|C_0|}\frac{1}{u(s)^2 - a_b^2}\prod_{m=1}^{M}(u(s)^2 - a_m^2)^{-\beta_m}, \quad (19)$$

where the dependence $u(s)$ is found by integrating the equation

$$\frac{du}{ds} = \frac{1}{|C_0|}\prod_{m=1}^{M}(u^2 - a_m^2)^{-\beta_m}. \quad (20)$$

For both distributions (13) and (19), the relative deviation from the basic field is computed by the formula

$$\delta\mathcal{B}(\rho|s) = \frac{\mathcal{B}(\rho|s)}{Pp_0(\rho|s)^{P-1}} - 1. \quad (21)$$

3. LENS-PROFILE CONSTRUCTION: TECHNIQUE AND EXAMPLES

Let the width of the pole of a P lens be measured by the relative angle $\tilde{\varphi}_P$ defined as the ratio of the angle φ_P between the segments $A_{M+1}A_{M-3}$ and $A_{-M+3}A_{M+1}$ (Fig. 2) to the full angle of the lens sector:

$$\tilde{\varphi}_P = P\varphi_P/\pi. \quad (22)$$

This characteristic of the lens is assumed to be specified. Hereafter, we will consider only the left half-sector of the lens, because similar results will also be obtained in the other half-sector by virtue of symmetry.

First, we assume that the pole profile (the region between the points A_0 and A_{M-3} in Fig. 2) is ideal and contains $N = M - 4$ evenly spaced intermediate points $A_1, A_2, \ldots, A_N = A_{M-4}$. Each of these points is defined by angular and radial coordinates φ_n^0, r_n^0 ($n = 1, 2, \ldots, N$) found from. (3.6) in Appendix 3. Our experience suggests that the optimization process calls for a departure from the ideal profile, so the coordinates of the pole points must be changed. We emphasize that both coordinates of the zero (by convention) vertex ($\varphi_0 = \pi/4$, $r_0 = 1$) and the angular coordinate $\varphi_{M-3} = \varphi_{N+1}$ of the vertex A_{N+1} (the coordinate determining the angular size of the pole) are always fixed.

To change the coordinates, we make use of the relationships containing two angle- and radius-adjusting parameters T_φ and T_r ($n = 0, 2, \ldots, N + 1$):

$$\varphi_n = \varphi_0 + (\varphi_n^0 - \varphi_0)T_\varphi^{N+1-n}, \quad r_n = (r_n^0)^{T_r}. \quad (23)$$

One can easily see that the fixed coordinates remain fixed for any values of these parameters. One more constraint is due to the fact that the vertex A_0 is taken by convention and is not put in parentheses in integral (2); i.e., the angle at it should be equal to π. For this reason, the radial coordinates of the adjacent vertices were always corrected so that the segment A_1A_{-1} was perpendicular to the line of symmetry of the sector.

It is known (see, e.g., [7]) that sometimes the field quality is improved by using so-called shims, which are small flat sections located, as a rule, at the pole edge. In the case shown in Fig. 2, the presence of a shim would have to "split" the A_{M-3} vertex into two vertices A_{N+1} and A_{M-3} (which coincide in the absence of the shim). A shim (if any) will be characterized by the length l_{sh} and the slope φ_{sh} of the segment $A_{N+1}A_{M-3}$.

With the parameters T_φ and T_r and the shim characteristics l_{sh} and φ_{sh}, one can design lenses with high field quality. After each change in the lens geometry, conformal mapping onto the upper semiplane w is performed (Section 1). This makes it possible to calculate the induction distribution over the line of symmetry, $\mathcal{B}(\rho)$, and along the sector boundary, $\mathcal{B}(s)$, and find the maximum relative deviation $\delta\mathcal{B}_{max}$ of the magnetic induction from the basic value (Section 2) within the working aperture $|r_a \le 0.9|$.

Below are two examples of using this technique. For a quadrupole with the relative angular pole dimension $\tilde{\varphi}_P = 0.7$ and the number of free (unfixed) vertices $M = 10$, the maximum relative field deviation $\delta\mathcal{B}_{max} = 0.0021$ within the working aperture $|r_a \le 0.9|$ was obtained for $T_\varphi = 0.960$, $T_r = 0.967$, $l_{sh} = 0.02$, and $\varphi_{sh} = 0.39\pi$. Note that $\delta\mathcal{B}_{max} = 0.0042$ was obtained for the nonoptimized broken ideal profile of the same angular size.

For a sextupole with $\tilde{\varphi}_P = 0.6$ and $M = 10$, $\delta\mathcal{B}_{max} = 0.011$ within the working aperture $|r_a \le 0.9|$ for $T_\varphi = 0.99$ and $T_r = 0.86$ (the shim was absent). The broken ideal profile with the same angular size had $\delta\mathcal{B}_{max} = 0.071$.

CONCLUSION

In our opinion, this optimization procedure, aimed at improving the quality of the lens field, may be further developed in two directions. First, it seems promising to construct algorithms that significantly increase the number of free vertices of a polygon approximating the lens sector. As mentioned above, difficulties in the convergence of the conformal mapping process may arise in our approach when the number of these vertices is large (Section 1). This problem is known well from the theory and applications of multiparametric minimization when the deterministic algorithms like that of gradient descent are used: the system reaches a local minimum and cannot leave this state on its own [8]. This fact is a serious limitation and sends us in search of other methods for finding the global minimum. The first attempt was made in [6], where the current state of the system is "shaken up" in order to take it out of the local minimum. We are planning to combine regular- and random-search approaches [9].

Currently, the lens profile is sequentially improved (Section 3) "manually," i.e., through operator–machine dialog. The other way of evolving this technique is to develop software that would allow for a totally computerized optimization process.

APPENDIX 1

Computation of the Polygon Side Lengths

Let us write Eq. (3) for the polygon side length in the expanded form

$$I_k = |C_0| \int_{a_k}^{a_{k+1}} \prod_{m=1}^{M} |u - a_m|^{\beta_m} \prod_{m=1}^{M} |u + a_m|^{\beta_m} du. \quad (1.1)$$

Integration is along the positive semiaxis u only; therefore, singularities at negative β_k and β_{k+1} may arise only in the first product. To eliminate them, two

changes of variables are proposed. The first one [10],

$$\eta = (u - a_k)/(a_{k+1} - a_k),$$
$$u = \eta(a_{k+1} - a_k) + a_k, \quad du = (a_{k+1} - a_k)d\eta,$$

brings the integral to the standard form

$$I_k = C_k \int_0^1 Q_k(\eta)\eta^{\beta_k}(1-\eta)^{\beta_{k+1}}d\eta. \tag{1.2}$$

In this expression,

$$C_k = |C_0|(a_{k+1} - a_k)^{-\alpha_{M+1}},$$
$$Q_k(\eta) = \prod_{m=1}^{M}|\eta + \bar{\gamma}_{km}|^{\beta_m} \prod_{m=1,\, m\neq k,\, k+1}^{M}|\eta - \gamma_{km}|^{\beta_m},$$
$$\bar{\gamma}_{km} = (a_m + a_k)/(a_{k+1} - a_k),$$
$$\gamma_{km} = (a_m - a_k)/(a_{k+1} - a_k).$$

Let us present integral (1.2) as the sum of two terms: $I_k = I_{kl} + I_{kr}$, where

$$I_{kl} = C_k \int_0^{1/2} Q_k(\eta)\eta^{\beta_k}(1-\eta)^{\beta_{k+1}}d\eta, \tag{1.3a}$$

$$I_{kr} = C_k \int_{1/2}^1 Q_k(\eta)\eta^{\beta_k}(1-\eta)^{\beta_{k+1}}d\eta. \tag{1.3b}$$

The second change differs for different parts of the interval. For the left part, the change

$$\eta(\xi) = [(\beta_k + 1)\xi]^{1/(\beta_k+1)},$$
$$\xi = \eta^{\beta_k+1}/(\beta_k + 1), \quad d\xi = \eta^{\beta_k}d\eta$$

transforms the augend I_{kl} to the form

$$I_{kl} = C_k \int_0^{U_l} Q_k(\eta)(1-\eta)^{\beta_{k+1}}d\xi, \tag{1.4a}$$
$$U_l = (1/2)^{\beta_k+1}/(\beta_k + 1).$$

For the right part, the change

$$\eta(\xi) = 1 - [(\beta_{k+1} + 1)\xi]^{1/(\beta_{k+1}+1)},$$
$$\xi = (1-\eta)^{\beta_{k+1}+1}/(\beta_{k+1} + 1), \quad d\xi = (1-\eta)^{\beta_{k+1}}d\eta$$

allows I_{kr} to be expressed as

$$I_{kr} = C_k \int_0^{U_r} Q_k(\eta)\eta^{\beta_k}d\xi, \tag{1.4b}$$
$$U_r = (1/2)^{\beta_{k+1}+1}/(\beta_{k+1} + 1).$$

Both relations (1.4) are free of singularities throughout the closed interval of integration $[a_k, a_{k+1}]$ for any values of the exponents β_k and β_{k+1}.

APPENDIX 2

Computation of the $\partial\lambda_k/\partial a_j$ Factors for System (6)

The factors appearing in Eq. (6) are related to the derivatives of the polygon side lengths as follows [6]:

$$\frac{\partial\lambda_k}{\partial a_j} = \frac{1}{I_0}\left(\frac{\partial I_k}{\partial a_j} - \lambda_k\frac{\partial I_0}{\partial a_j}\right). \tag{2.1}$$

For $j \neq k, k+1$, the derivatives may be obtained by direct differentiation under the integration sign:

$$\frac{\partial I_k}{\partial a_j} = \left(\frac{\partial I_k}{\partial a_j}\right)_- + \left(\frac{\partial I_k}{\partial a_j}\right)_+,$$
$$\left(\frac{\partial I_k}{\partial a_j}\right)_- = |C_0|\int_{a_k}^{a_{k+1}}\prod_{m=1}^{M}\left|u^2 - a_m^2\right|^{\beta_m}\frac{\beta_j}{a_j - u}du, \tag{2.2a}$$

$$\left(\frac{\partial I_k}{\partial a_j}\right)_+ = |C_0|\int_{a_k}^{a_{k+1}}\prod_{m=1}^{M}\left|u^2 - a_m^2\right|^{\beta_m}\frac{\beta_j}{a_j + u}du. \tag{2.2b}$$

When $j = k$ and $j = k+1$, using Eqs. (10) and (11) in [6], we find

$$\frac{\partial I_k}{\partial a_k} = \frac{1}{a_{k+1} - a_k}\left[\alpha_{M+1}I_k + \sum_{m=1,\, m\neq k,\, k+1}^{M}\left((a_m - a_{k+1})\times\left(\frac{\partial I_k}{\partial a_m}\right)_- - (a_m + a_{k+1})\left(\frac{\partial I_k}{\partial a_m}\right)_+\right)\right], \tag{2.3a}$$

$$\frac{\partial I_k}{\partial a_{k+1}} = \frac{1}{a_k - a_{k+1}}\left[\alpha_{M+1}I_k + \sum_{m=1,\, m\neq k,\, k+1}^{M}\left((a_m - a_k)\times\left(\frac{\partial I_k}{\partial a_m}\right)_- - (a_m + a_k)\left(\frac{\partial I_k}{\partial a_m}\right)_+\right)\right]. \tag{2.3b}$$

APPENDIX 3

Conformal Mapping of an Ideal P Lens Sector onto the Upper Semiplane

For appropriate units of measure, the complex potential of an ideal *P* lens can be presented as

$$\mathcal{P}(z) = (z\exp(-i\alpha_P))^P, \tag{3.1}$$

where $\alpha_P = \pi(P - 2)/4P$. The exponential was introduced for the bisectrix of the first quadrant, $\varphi = \pi/4$, to

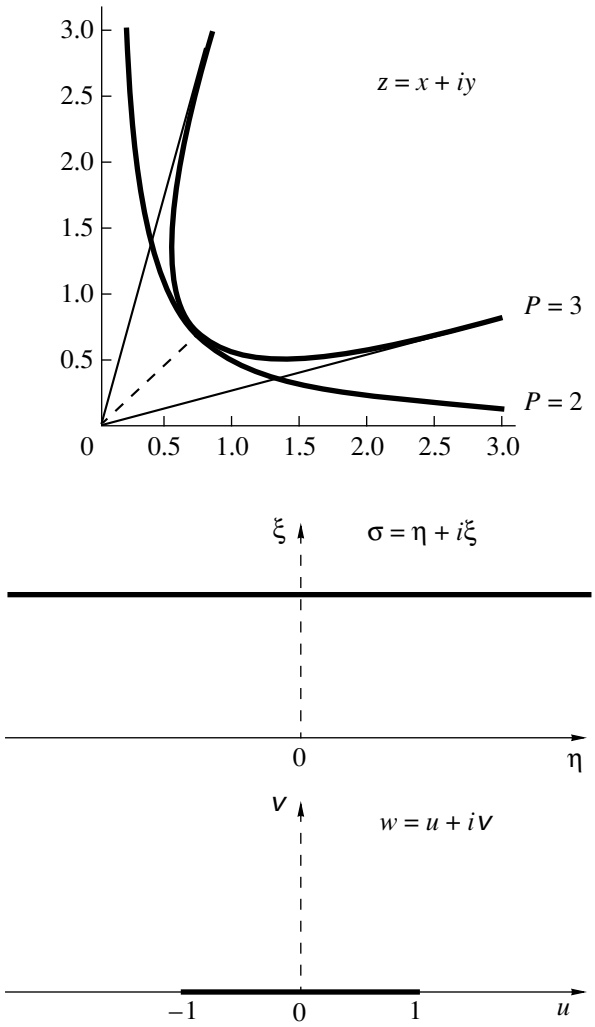


Fig. 4. Mapping of profiles of ideal *P* lenses onto the upper semiplane.

become the line of symmetry of the lens sector at any number of pole pairs *P*.

The pole profile is determined from the condition $\mathrm{Im}\mathcal{P}(z) = 1$ [2]:

$$r^P \sin[P(\varphi - \alpha_P)] = 1. \tag{3.2}$$

The function $\delta(z) = (z\exp(-i\alpha_P))^P$ maps the sector of this lens onto the strip (lune) in the plane of the intermediate variable σ. The function $w(\sigma) = -(2/\pi)\tanh^{-1}(\sigma)$ maps the upper semiplane of the *w* plane onto the same strip. Thus, one can easily find the required map of the lens sector on the upper semiplane $v \geq 0$ in the form [11]

$$w(z) = -\coth\frac{\pi}{2}(z\exp(-i\alpha_P))^P. \tag{3.3}$$

This sequence of maps is shown in Fig. 4, where thin lines (boundaries of zero-potential sectors) pass into thin lines; heavy lines (profiles of unity-potential poles) pass into heavy lines; and the dashed segment along the bisectrix of the first quadrant, within which the potential distribution is sought, passes into dashed lines.

Map (3.3) is convenient to analyze in the form

$$w(z) = -\coth\left[\left(\frac{\pi}{2}r^P\right)(\cos\tilde{\varphi} + i\sin\tilde{\varphi})\right], \qquad \tilde{\varphi} = P(\varphi - \alpha_P). \tag{3.4}$$

The imaginary semiaxis $v \geq 0$ of the *w* plane turns out to be the image of the dashed segment ($0 \leq \rho \leq 1$, $\tilde{\varphi} = \pi/2$) in the *z* plane:

$$w(\rho) = -\coth\left|i\frac{\pi}{2}\rho^P\right| = i\coth\left(\frac{\pi}{2}\rho^P\right).$$

Note that the potential on the positive imaginary semiaxis is computed by the Schwartz formula $\mathcal{P}(v) = 2/\pi\tan^{-1}(a_b/v)$ (Section 2). As seen from Fig. 4, the boundary of the image of any ideal *P* profile $a_b = 1$. Then, using Eq. (3.4), we can reconstruct the radial distribution of the potential along the bisectrix:

$$\mathcal{P}(0 \leq \rho \leq 1) = \frac{2}{\pi}\tan^{-1}\left(\frac{1}{\cot\left(\frac{\pi}{2}\rho^P\right)}\right) = \frac{2}{\pi}\tan^{-1}\left[\tan\left(\frac{\pi}{2}\rho^P\right)\right] = \rho^P. \tag{3.5}$$

Expressions (3.5) and (3.1) confirm the validity of all the constructions. For points lying on the ideal poles (heavy lines), we have

$$\mathrm{Im}\mathcal{P}(z) = r^P\sin\tilde{\varphi} = 1,$$
$$\mathrm{Re}\mathcal{P}(z) = r^P\cos\tilde{\varphi} = b \neq \mathrm{const},$$
$$w(z) = -\coth\frac{\pi}{2}(b + i)$$
$$= \frac{1 - i\coth\left(\frac{\pi}{2}b\right)\cot\left(\frac{\pi}{2}\right)}{\coth\left(\frac{\pi}{2}b\right) - i\cot\frac{\pi}{2}} = -\tanh\left(\frac{\pi}{2}r^P\cos\tilde{\varphi}\right) \tag{3.6}$$
$$= \text{real number}.$$

This real number is just the image *a* of the appropriate point of the ideal profile that is mapped into the segment $-1 \leq u \leq 1$ of the real axis of the *w* plane.

APPENDIX 4

Potential Spectrum

The potential distribution along the segment A_0A_{M+1} on the line of symmetry of the lens sector

(Fig. 3) is given by formula (10), which at $v > a_b$ can be expressed as

$$\mathcal{F}(v) = \frac{2}{\pi}\tan^{-1}\left(\frac{a_b}{v}\right) = \frac{2}{\pi}\left[\frac{a_b}{v} - \frac{1}{3}\left(\frac{a_b}{v}\right)^3 + \frac{1}{5}\left(\frac{a_b}{v}\right)^5 - \ldots\right]. \quad (4.1)$$

Since the relative contribution of upper harmonics is of major interest, it is convenient to replace v by the variable

$$\xi = \frac{1}{\alpha_0 v} \quad (4.2)$$

and represent the relative potential in the form

$$\tilde{\mathcal{F}}(\xi) = \mathcal{F}/p_0 = \xi - \frac{(a_b\alpha_0)^2}{3}\xi^3 + \frac{(a_b\alpha_0)^4}{5}\xi^5 - \ldots. \quad (4.3)$$

The quantities p_0 and α_0, appearing in Eqs. (4.3) and (4.2), are defined by formulas (17) and (18), respectively. Now we pass to the new independent variable

$$\rho = r^P \quad (4.4)$$

and will seek the dependence of ξ on ρ in the form of the series:

$$\xi = \rho[1 + \eta(\rho)], \quad \eta(\rho) = \sum_{n=1}^{N} \tilde{\alpha}_n \rho^{2n}. \quad (4.5)$$

Since the quality of the lens field is assumed to be high, the second expression in (4.5) can be considered as a small correction; i.e., $|\eta(\rho)| \ll 1$.

The factors $\tilde{\alpha}_n$, $n = 1, 2, \ldots$, appearing in Eq. (4.5), are computed as follows. From definition (4.2), we find

$$\frac{d\xi}{d\rho} = -\frac{1}{\alpha_0 P r^{P-1}}\frac{1}{v^2}\frac{dv}{dr},$$

where dv/dr is given by Eq. (13), written in the equivalent form

$$\frac{dv}{dr} = -\frac{1}{|C_0|}v^{(1+P)/P}\prod_{m=1}^{M}[1 + (a_m/v)^2]^{-\beta_m}.$$

On simple rearrangements in view of the identity $\alpha_0^{-1/P} = P|C_0|$, we obtain

$$\frac{d\xi}{d\rho} = [1 + \eta(\rho)]^{(P-1)/P}\prod_{m=1}^{M}[1 + (\alpha_0 a_m \xi)^2]^{-\beta_m}. \quad (4.6)$$

Expanding the right of (4.6) by the binomial formula and collecting terms, we come to the power series in ρ with coefficients including unknowns $\tilde{\alpha}_n$. On the other hand, the derivative $d\xi/d\rho$ can be obtained directly from fundamental definition (4.5) as the series

$$\frac{d\xi}{d\rho} = 1 + 3\tilde{\alpha}_1\rho^2 + 5\tilde{\alpha}_2\rho^4 + \ldots. \quad (4.7)$$

Comparing the coefficients at the same powers of ρ in Eqs. (4.6) and (4.7) and using the exact values of p_0 and α_0 [formulas (17) and (18), respectively], we can recurrently determine the parameters $\tilde{\alpha}_n$, $n = 1, 2, \ldots$, appearing in (4.5). This makes it possible to construct the explicit dependence $\xi(\rho)$ and, with (4.3), find the amplitudes $\tilde{p}_n$, $n = 1, 2, \ldots$, of the potential harmonics [see (7) in Section 2].

This algorithm is rather intricate and cannot be executed "manually." For this reason, a program was written in the REDUCE language [12], which allowed us to find the values of $\tilde{p}_n$, $n = 1, 2, \ldots, N$ at any N.

ACKNOWLEDGMENTS

We thank V.N. Korchuganov and E.B. Levichev for support and valuable discussions.

Special thanks to A.N. Dubrovin for putting in our disposal the MERMAID versatile computing program.

REFERENCES

1. E. D. Courant *et al.*, Phys. Rev. **88**, 1190 (1952).
2. I. M. Kapchinskiĭ, *Dynamics of Particles in Linear Resonance Accelerators* (Atomizdat, Moscow, 1966).
3. K. G. Steffen, *High-Energy Beam Optics* (Interscience, New York, 1965; Mir, Moscow, 1969).
4. V. K. Plotnikov, Prib. Tekh. Éksp., No. 2, 29 (1962).
5. M. A. Lavrent'ev and B. V. Shabat, *Methods of the Theory of Functions of Complex Variable* (Nauka, Moscow, 1973).
6. V. V. Vecheslavov and V. I. Kokoulin, Zh. Vychisl. Mat. Mat. Fiz. **13**, 865 (1973).
7. B. Z. Persov and É. M. Trakhtenberg, *Design of Experimental Physical Setups: A Textbook* (Novosibirsk. Gos. Univ., Novosibirsk, 1993).
8. F. P. Vasil'ev, *Numerical Solutions to Extremal Problems* (Nauka, Moscow, 1980).
9. L. A. Rastrigin, *Statistical Search Methods* (Nauka, Moscow, 1968).
10. W. Koppenfels and F. Stallmann, *Praxis der konformen Abbildung* (Springer-Verlag, Berlin, 1959; Inostrannaya Literatura, Moscow, 1963).
11. V. I. Lavrik and V. N. Savenkov, *Conformal Mapping: A Handbook* (Naukova Dumka, Kiev, 1970).
12. A. C. Hearn, *REDUCE User's Manual, Version 3.6.*, RAND Pub. CP78 Rev. 7/95, 1995.

Translated by N. Goryacheva

Technical Physics, Vol. 45, No. 10, 2000, pp. 1320–1324. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 90–94.

ELECTRON AND ION BEAMS, ACCELERATORS

Small-Angle Approximation for an Ion Beam in an External Electromagnetic Field

N. D. Naumov

Received December 14, 1999

Abstract—A small-angle approximation method is formulated for a curvilinear ion beam propagating in an electromagnetic field. The effects of particle velocity scatter and multiple elastic scattering on the beam path in a magnetic field, the electric field of a cylindrical capacitor, and mutually orthogonal electric and magnetic fields are considered. An analytical model for beam power compression is developed. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The need for models that simulate ion beam propagation in external electromagnetic fields is dictated by applications in the field of electron optics, mass spectroscopy, etc. [1–3]. One theoretical approach to analyzing charged beam dynamics is the paraxial theory. It is helpful in designing appropriate beam-shaping electrodes [2] and in solving the direct problem when the characteristics of an external electromagnetic field in which the beam propagates are given [4].

Another method of theoretical analysis of the beam dynamics is to solve the kinetic equation. In the framework of this approach, one can consider the effects of particle velocity spread and ion scattering by gas molecules. In this case, however, the problem becomes mathematically more complex, because it involves the solution of partial differential equations. As applied to the problem of charged beam propagation in an external field, such solutions are known for a uniform magnetic field [5–7]. This paper addresses a more general problem for a curvilinear beam propagating in an external electromagnetic field. This problem can be solved analytically for a narrow beam, when the ratio of its cross size to the radius of curvature is small.

STATEMENT OF THE PROBLEM

We assume that the beam propagates perpendicularly to the z-axis, i.e., that the beam axis is a plane curve. The beam with such a configuration can be obtained, in particular, in mutually orthogonal nonuniform electric, $\mathbf{E} = E_x\mathbf{e}_x + E_y\mathbf{e}_y$, and uniform magnetic, $\mathbf{B} = B_0\mathbf{e}_z$, fields. The analysis of a more general problem, when the beam axis is a spatial curve, involves more tedious computations. For the sake of brevity, we will study mostly the propagation of a curvilinear ribbon beam. Only in the last section, the beam limited in the z direction is considered.

Beam motion is convenient to study in a system of curvilinear coordinates s, q, ζ:

$$\mathbf{x} = \mathbf{Y}(s) + q\mathbf{n} + \zeta\mathbf{b}.$$

The curve $\mathbf{Y}(s)$ is the axis of the beam's cross section in the plane $z = 0$; the coordinate s is the length of the axial particle trajectory to the place of beam injection; and $\mathbf{t}$, $\mathbf{n}$, and $\mathbf{b} = \pm\mathbf{e}_z$ are the vectors of the Frenet trihedral related to the curve $\mathbf{Y}(s)$. The direction of the vector $\mathbf{b}$ depends on the direction of the external electric field. The axis of the beam cross section, which coincides with the axial particle trajectory, is given by a solution to the motion equation for a single particle in a given external field. We assume that this solution is known.

Substituting the expression for the particle velocity $\mathbf{v} = u\mathbf{t}$ into the equation of motion of an axial particle yields the expression for the curvature of its trajectory $k = \kappa - e\mathbf{bB}/mcu$, where $\kappa = eE_{02}/mu^2$. A change in the velocity of the axial particle in the course of propagation is determined by quantity E_{01}: $muu' = eE_{01}$. Here, E_{0i} are the components of the electric field $E(\mathbf{Y}(s)) = E_{01}\mathbf{t} + E_{02}\mathbf{n}$ on the beam axis and prime means differentiation with respect to s.

Before solving the problem, let us analyze the electric field structure near the beam axis. For brevity, we will consider the case when the electric field is independent of z. Then, in the curvilinear coordinates, the conditions $\mathrm{div}\mathbf{E} = 0$ and $\mathrm{curl}\mathbf{E} = 0$ take the form

$$kE_1 + E_2' = \rho\frac{\partial E_1}{\partial q}, \quad kE_2 - E_1' = \rho\frac{\partial E_2}{\partial q}, \qquad (1)$$

where $\rho = 1 - kq$, $E_1 = \mathbf{tE}$, and $E_2 = \mathbf{nE}$.

We expand the electric field components E_i near the curve $\mathbf{Y}(s)$ into the Taylor series $E_i = E_{0i} + g_iq + h_iq^2$ and substitute them into conditions (1). As a result, we obtain functions $g_i(s)$ that are necessary for solving the

1063-7842/00/4510-1320$20.00

problem in the general case:

$$g_1 = E'_{02} + kE_{01}, \quad g_2 = kE_{02} - E'_{01}.$$

EQUATION OF TRAJECTORY

The equation for charged particle motion in an external electromagnetic field can be derived from the Maupertuis principle

$$\delta\int\left(m\Lambda dl + \frac{e}{c}\mathbf{A}d\mathbf{x}\right) = 0.$$

Here, $\Lambda(\mathbf{x}) = \sqrt{2(W - e\Phi)/m}$, Φ and $\mathbf{A}$ are the external field potentials, and l is the distance along the trajectory. Calculations similar to those for a particle in a potential field [8] yield the trajectory equation in the presence of a magnetic field:

$$m\Lambda\frac{d}{dl}\left(\Lambda\frac{d\mathbf{x}}{dl}\right) = e\mathbf{E} + \frac{e}{c}\Lambda\left[\frac{d\mathbf{x}}{dl}\mathbf{B}\right]. \quad (2)$$

One can easily see that $\Lambda(\mathbf{x}(l))$ is the magnitude of the particle velocity. In particular, for the axial particles, $\Lambda(\mathbf{Y}(s)) = u$. In order to obtain an approximate equation of trajectory near the curve $\mathbf{Y}(s)$, the particle velocity vector should be represented in the form

$$\Lambda\frac{d\mathbf{x}}{dl} = u[(1 + \alpha)\mathbf{t} + q'\mathbf{n}]. \quad (3)$$

On a trajectory that passes near the beam axis, kq and α are small. Substituting expression (3) into Eq. (2) and neglecting the second-order terms, we come to the system of ordinary differential equations

$$q'' + \frac{u'}{u}q' + k^2q + (k + \kappa)\alpha = \frac{eg_2}{mu^2}q, \quad (4)$$

$$\alpha' + (kq + 2\alpha)\frac{u'}{u} = kq' + \frac{eg_1}{mu^2}q. \quad (5)$$

Since Eq. (5) can formally be integrated in the general form, a single integro-differential equation can be written for the trajectory of a particle that moves near the beam axis. However, in practice, it is easier to proceed from system of Eqs. (4) and (5). It is sometimes more convenient to use the time of motion of an axial particle

$$\tau = \int_o^s \frac{dx}{u(x)},$$

rather than s, as the longitudinal coordinate, i.e., to pass to the parametric representation of the trajectory. In this case, Eqs. (4) and (5) take the following form (dots mean differentiation with respect to τ):

$$\ddot{q} + k^2u^2q + \alpha(k + \kappa)u^2 = \frac{e}{m}g_2q, \quad (6)$$

$$\dot{\alpha} + (kq + 2\alpha)\frac{\dot{u}}{u} = \kappa\dot{q} + \frac{eg_1}{mu}q. \quad (7)$$

Consider particular examples. The axis of the cross section of a ribbon ion beam in a uniform magnetic field is a segment of a circle whose curvature is independent of s: $k = eB_0/mcu_0$. In this case, the magnitude of the axial particle velocity and the energy inhomogeneity of the beam do not change, because Eq. (4) gives $\alpha = \alpha_0$. Therefore, Eq. (5) acquires the much simpler form

$$q'' + k^2q + k\alpha_0 = 0.$$

Hence, the particle trajectory is

$$q = \left(q_0 + \frac{\alpha_0}{k}\right)\cos\psi + \frac{q'_0}{k}\sin\psi - \frac{\alpha_0}{k},$$

where $\psi = ks$.

Note that q'_0 characterizes the initial beam divergence. It is easy to see that, at $\alpha_0 = -kq_0$, the beam's cross sizes change only because of initial divergence $q = q_0 + (q'_0/k)\sin\psi$.

Under certain conditions, the axis of the beam propagating in the electric field of a cylindrical capacitor has the form of a segment of a circle; this kind of device is used in electrostatic analyzers employed in mass spectroscopy [3]. In this case, $k = eE_{02}/mu_0^2$, $E_{01} = 0$, and $g_2 = kE_{02}$; hence, Eqs. (4) and (5) yield the trajectory

$$q = q_0 + \frac{\alpha_0}{k}(\cos\psi - 1) + \frac{q'_0}{\sqrt{2k}}\sin\psi, \quad \psi = \sqrt{2}ks.$$

Finally, consider the beam in the mutually orthogonal uniform fields $\mathbf{E} = E_0\mathbf{e}_y$ and $\mathbf{B} = B_0\mathbf{e}_z$. For simplicity, we choose the following initial conditions for particles on the beam axis: $x_0 = 0$, $y_0 = (u_0/\Omega)\sin\phi$, $\dot{x}_0 = u_0\sin\phi$, and $\dot{y}_0 = u_0\cos\phi$, where $\Omega = eB_0/mc$ is the cyclotron frequency, ϕ is the angle between the electric field vector and direction of particle injection, and u_0 is the initial particle velocity. In this case, the curve $\mathbf{Y}(s)$ is a part of the cycloid

$$Y_x = \frac{v_0}{\Omega}(\Omega\tau + \sin2\phi - \sin2\psi),$$

$$Y_y = \frac{v_0}{\Omega}(1 - \cos2\psi).$$

Here, $v_0 = cE_0/B_0$, $\psi = \omega\tau + \phi$, $\omega = \Omega/2$, and the longitudinal coordinate is represented in the parametric form $s = 2v_0(\cos\phi - \cos\psi)/\omega$ (with $\phi \le \psi \le \pi/2$). This choice of the initial conditions leads to simple expressions for

the beam axis curvature, $k = \omega/2v_0\sin\psi$, and Frenet trihedral vectors,

$$\mathbf{t} = (\sin\psi, \cos\psi), \quad \mathbf{n} = (\cos\psi, -\sin\psi), \quad \mathbf{b} = -\mathbf{e}_z.$$

Equation (6) also becomes simple, $\ddot{q} + \omega^2 q = 0$, and easy to solve.

EFFECT OF VELOCITY SPREAD

First, consider the effect of particle velocity spread in the steady-state statement. To this end, we will use the kinetic equation in the collision-free approximation

$$L_0 f = S_0,$$
$$L_0 = \mathbf{v}\frac{\partial}{\partial \mathbf{x}} + \frac{e}{m}\left(\mathbf{E} + \frac{1}{c}[\mathbf{vB}]\right)\frac{\partial}{\partial \mathbf{v}}, \tag{8}$$

where S_0 is the source density.

Writing Eq. (8) in the curvilinear coordinates, we find the following expression for the operator L_0:

$$L_0 = \frac{v_1}{\rho}\frac{\partial}{\partial s} + \left[v_1 v_2 \frac{k}{\rho} + \frac{e}{m}\left(E_1 + \frac{v_2}{c}B_3\right)\right]\frac{\partial}{\partial v_1}$$
$$+ v_2\frac{\partial}{\partial q} + \left[\frac{e}{m}\left(E_2 - \frac{v_1}{c}B_3\right) - v_1^2\frac{k}{\rho}\right]\frac{\partial}{\partial v_2},$$

where $v_1 = \mathbf{vt}$, $v_2 = \mathbf{vn}$, and $B_3 = \mathbf{bB}$.

Assuming that $v_1 = u(1 + \alpha)$ and $v_2 = u\beta$ and leaving only the first-order terms, we eventually obtain the equation for the distribution function of a narrow beam

$$Lf(X) = Q(X), \quad Q = \frac{S_0}{u},$$
$$L = \frac{\partial}{\partial s} + \left[\kappa\beta - (2\alpha + kq)\frac{u'}{u} + \frac{eq_1}{mu^2}q\right]\frac{\partial}{\partial\alpha} \tag{9}$$
$$+ \beta\frac{\partial}{\partial q} + \left[\frac{eg_2}{mu^2}q - k^2 q - (k + \kappa)\alpha - \beta\frac{u'}{u}\right]\frac{\partial}{\partial\beta},$$

where X stands for the set of variables x, q, α, and β for simplicity.

A solution to this equation can be written in terms of the Green function:

$$f(X) = \int G(X, X_0)Q(X_0)dX_0,$$
$$LG(X, X_0) = \delta(X - X_0). \tag{10}$$

The Green function has the form

$$G(X, X_0) = \vartheta(s - s_0)\delta(q - q(s, X_0))$$
$$\times\delta(\alpha - \alpha(s, X_0))\delta(\beta - \beta(s, X_0)).$$

Here, $\vartheta(x)$ is the Heaviside step function; $q(s, X_0)$, $\alpha(s, X_0)$, and $\beta(s, X_0)$ are the solutions to the system of ordinary differential equations

$$q' = \beta, \quad \alpha' = \kappa\beta - (k + 2\alpha)\frac{u'}{u} + -\frac{eg_1}{mu^2}q, \tag{11}$$

$$\beta' = k^2 q - (k + \kappa)\alpha - \beta\frac{u'}{u} + \frac{eg_2}{mu^2}q \tag{12}$$

that satisfy the following conditions: $q(s_0, X_0) = q_0$, $\alpha(s_0, X_0) = \alpha_0$, and $\beta(s_0, X_0) = \beta_0$.

Consider the effect of particle velocity spread on ion beam propagation in the magnetic field, assuming that

$$S_0 = \frac{n_0 u}{2\pi\sigma^2}\delta(s)\delta(q)\exp\left(-\frac{\alpha^2 + \beta^2}{2\sigma^2}\right).$$

In this case, the solutions to Eqs. (11) and (12) have the form

$$q = \left(q_0 + \frac{\alpha_0}{k}\right)\cos\varphi + \frac{\beta_0}{k}\sin\varphi - \frac{\alpha_0}{k}, \tag{13}$$

$$\alpha = \alpha_0, \quad \beta = \beta_0\cos\varphi - (\alpha_0 + kq_0)\sin\varphi, \tag{14}$$

where $\varphi = k(s - s_0)$.

Substituting these expressions into (10), we obtain the beam distribution function

$$f = \frac{n_0\vartheta(s)}{2\pi\sigma^2}\delta\left[\left(q + \frac{\alpha}{k}\right)\cos\psi - \frac{\beta}{k}\sin\psi - \frac{\alpha}{k}\right]$$
$$\times\exp\left[-\frac{1}{2\sigma^2}(\alpha^2 + [\beta\cos\psi + (\alpha + kq)\sin\psi]^2)\right]$$

and then the expression for the particle number density along the beam

$$n = \int f(X)d\alpha d\beta = \frac{n_0\vartheta(s)}{\varepsilon\sqrt{2\pi}}\exp\left(-\frac{q^2}{2\varepsilon^2}\right), \tag{15}$$

where $\varepsilon = (2/k)\sigma\sin(\psi/2)$ and $\psi = ks$.

A similar result is obtained for the other two examples in the previous section. For the beam in the inhomogeneous electric field of a cylindrical capacitor, we should set in (15)

$$\varepsilon^2 = \frac{\sigma^2}{2k^2}(1 - \cos\sqrt{2}ks)(3 - \cos\sqrt{2}ks),$$

while, for the beam in the orthogonal fields, $\varepsilon = (\sigma u_0/\omega)\sin\omega\tau$. The latter problem is convenient to solve using the parametric representation of curve $\mathbf{Y}(s)$. To do this, the variable s in Eq. (9) must be replaced by τ.

NONSTATIONARY INJECTION

In the nonstationary problem, the original equation for the distribution function has the form

$$\left(\frac{\partial}{\partial t} + L_0\right)f(\xi) = S(\xi), \tag{16}$$

where ξ stands for the aggregate of X and t.

Then, (9) and (10) can be written as

$$Nf = Q, \quad N = \frac{1}{u}(\rho - \alpha)\frac{\partial}{\partial t} + L, \quad Q = \frac{S}{u},$$

$$f(\xi) = \int \Gamma(\xi, \xi_0) Q(\xi_0) d\xi_0, \quad N\Gamma(\xi, \xi_0) = \delta(\xi - \xi_0).$$

The nonstationary Green function has the form

$$\Gamma(\xi, \xi_0) = G(X, X_0)\delta(t - t(s, \xi_0)),$$

where $t(s, \xi_0)$ is a solution to the differential equation

$$t' = \frac{1}{u}(1 - kq - \alpha)$$

with $t(s_0, \xi_0) = t_0$.

In particular, setting $S = S_0\delta(t)$ and using the above results, we obtain the particle number density in the magnetic field:

$$n = \frac{n_0 \vartheta(s)}{2\pi\varepsilon^2}\exp\left(-\frac{1}{2\varepsilon^2}\left[q^2 + u^2\left(t - \frac{s}{u}\right)^2\right]\right).$$

This method can be used to construct an analytical model of beam power compression. We consider the beam in which the initial particle velocity varies as $u(t) = u_1/\mu$, where $\mu = 1 - u_1 t/R$. In this case, a particle injected at the time instant $t = 0$ will be at a distance R from the place of injection at $t = t_1 = R/u_1$. A particle emitted later, at $t = t_2$, has the initial velocity $u_2 = u_1/(1 - t_2/t_1)$ and, at the time instant $t = t_1$, will be at the same distance $u_2(t_1 - t_2) = R$ from the place of injection. Thus, the adopted variation of the initial particle velocity leads to space–time focusing of the beam in the longitudinal direction [1].

Let us solve Eq. (16) in the absence of an external field with a source that obeys the above injection law within a time interval $0 \le t \le T$; namely,

$$S = \frac{n_0 u_1}{(\sigma\sqrt{2\pi})^3}\delta(\mathbf{x})\Pi(t)\exp\left[-\frac{1}{2\sigma^2}(\mu^2 v^2 - 2\mu v_x + u_1^2)\right],$$

where $\Pi(t) = \vartheta(t) - \vartheta(t - T)$.

Here, we assume that the relative spread of the particle velocities remains constant during injection. In this case, the Green function is

$$\Gamma(t, \mathbf{x}, t_0, \mathbf{x}_0) = \vartheta(t - t_0)\delta[\mathbf{x} - \mathbf{x}_0 - \mathbf{v}(t - t_0)],$$

which implies the distribution function

$$f = \frac{n_0 u_1 v^2}{v_x(\sigma\sqrt{2\pi})^3}\Pi\left(t - \frac{x}{v_x}\right)\delta(y v_x - x v_y)\delta(z v_x - x v_z)$$

$$\times \exp\left[-\frac{1}{2\sigma^2}\left(\nu^2\frac{v^2}{v_x^2} + 2\nu u_1 + u_1^2\right)\right],$$

where $\nu = m v_x + x/t_1$.

Hence, the particle number density at the time instant $t = t_1$ is given by

$$n(\mathbf{x}, t_1) = N\left(\frac{u_1}{\sigma R\sqrt{2\pi}}\right)^3 \exp\left[-\frac{u_1^2}{\sigma^2}\left(\frac{r^2}{R^2} - 2\frac{x}{R} + 1\right)\right],$$

where N is the number of injected particles:

$$N = n_0\int_0^T u(t)dt = -n_0 R\ln\left(1 - \frac{T}{t_1}\right).$$

MULTIPLE SCATTERING

To account for the effect of multiple elastic scattering on the parameters of the beam propagating in a gaseous medium, we will use the kinetic equation with the collision integral in the small-angle approximation [9]. If the effect of the external field on the collision process is neglected, this equation for a narrow beam can be represented as

$$Mf = Q, \quad M = M_1 + M_2,$$

$$M_1 = L - \frac{\chi^2}{4}\frac{\partial^2}{\partial\beta^2}, \quad M_2 = \gamma\frac{\partial}{\partial\zeta} - \frac{\chi^2}{4}\frac{\partial^2}{\partial\gamma^2},$$

where χ^2 is the average square of the scattering angle per unit length and $\gamma = \mathbf{vb}/u$.

It is easy to check that the Green function for the operator M has the form

$$G(X_1, x_2, X_{10}, x_{20}) = \vartheta(s - s_0)G_1(X_1, X_{10})G_2(X_2, X_{20}),$$

where the functions G_n are the solutions to the equations

$$M_1 G_1 = 0, \quad \left(\frac{\partial}{\partial s} + M_2\right)G_2 = 0 \qquad (17)$$

and satisfy the conditions $G_n(s_0, x_n, X_{n0}) = \delta(x_n - x_{n0})$. Here, $X_n = \{s, x_n\}$, $x_1 = \{q, \alpha, \beta\}$, and $x_2 = \{\zeta, \gamma\}$.

Let us solve the first equation of (17) for the beam in a uniform magnetic field. In this case, one should pass from q and β to the new variables

$$\eta = (\alpha + kq)\cos\varphi - \beta\sin\varphi - \alpha - kq_0,$$

$$\Theta = \beta\cos\varphi + (\alpha + kq)\sin\varphi - \beta_0,$$

where $\varphi = k(s - s_0)$, as in expressions (13) and (14). Assuming that $G_1(X_1, X_{10}) = \delta(\alpha - \alpha_0)F_1(s, \eta, \Theta)$, we

obtain for F_1

$$\left[\frac{\partial}{\partial s} - \frac{\chi^2}{4}\left(\sin^2\varphi\frac{\partial^2}{\partial\eta^2} - \sin 2\varphi\frac{\partial^2}{\partial\eta\partial\Theta} + \cos^2\varphi\frac{\partial^2}{\partial\Theta^2}\right)\right]F_1 = 0. \quad (18)$$

To solve this equation, we will take advantage of the two-dimensional Fourier transformation with respect to η and Θ

$$F_1(s, \eta, \Theta) = \frac{1}{(2\pi)^2}\int d\lambda d\omega F(s, \lambda, \omega)\exp(i\lambda\eta + i\omega\Theta). \quad (19)$$

Substituting (18) into (17) yields the ordinary differential equation for the Fourier transform

$$F' + \frac{\chi^2}{4}(\lambda^2\sin^2\varphi - \lambda\omega\sin 2\varphi + \omega^2\cos^2\varphi)F = 0$$

with the initial condition $F(s_0, \lambda, \omega) = k$.

This equation is easy to integrate; eventually, the function F_1 can be written as

$$F_1 = \frac{k}{\pi D_1}\exp\left[-\frac{1}{D_1^2}(\eta^2 A_1 - 2\eta\Theta A_2 + \Theta^2 A_3)\right],$$

$$A_1 = \frac{\chi^2}{2k}\left(\varphi + \frac{1}{2}\sin 2\varphi\right), \quad A_2 = -\frac{\chi^2}{2k}\sin^2\varphi, \quad (20)$$

$$A_3 = \frac{\chi^2}{2k}\left(\varphi - \frac{1}{2}\sin 2\varphi\right), \quad D_1^2 = A_3A_1 - A_2^2.$$

The solution to the equation for G_2 is similar to (20) [10]:

$$G_2 = \frac{1}{\pi D_2}\exp\left(-\frac{1}{D_2^2}[p^2B_1 - 2p(\gamma - \gamma_0)B_2 + (\gamma - \gamma_0)^2B_3]\right),$$

$$p = \zeta - \zeta_0 - \gamma(s - s_0), \quad B_n = \frac{\chi^2}{n}(s - s_0)^n,$$

$$D_2^2 = B_3B_1 - B_2^2.$$

For a delta-like source $S_0 = n_0u\delta(X_1)\delta(x_2)$, the particle number density is given by

$$n = \frac{kn_0\vartheta(s)}{\pi\sqrt{a_3b_3}}\exp\left(-\frac{k^2q^2}{a_3} - \frac{z^2}{b^3}\right),$$

where a_3 and b_3 are, respectively, the values of the coefficients A_3 and B_3 at $s_0 = 0$.

These quantities characterize the beam divergence due to multiple elastic ion scattering by the gas molecules.

For the beam propagating in the field of a cylindrical capacitor, the following change of variables should be used:

$$\eta = \frac{k}{2}(q - q_0) + \alpha\left(\cos\varphi - \frac{1}{2}\right) - \frac{\beta}{\sqrt{2}}\sin\varphi - \frac{\alpha_0}{2},$$

$$\xi = k(q - q_0) - \alpha + \alpha_0,$$

$$\theta = \beta\cos\varphi + \sqrt{2}\alpha\sin\varphi - \beta_0,$$

where $\varphi = \sqrt{2}\,k(s - s_0)$.

Then, assuming that $G_1 = \delta(\xi)F_1(s, \eta, \Theta)$ and using the same way to find F_1, we arrive at the expression similar to (20) where the coefficients A_n look like

$$A_1 = \frac{\chi^2}{2\sqrt{2}k}\left(\varphi + \frac{1}{2}\sin 2\varphi\right), \quad A_2 = -\frac{\chi^2}{4k}\sin^2\varphi,$$

$$A_3 = \frac{\chi^2}{4\sqrt{2}k}\left(\varphi - \frac{1}{2}\sin 2\varphi\right).$$

The Green function for the beam in the mutually orthogonal fields can be derived in a similar manner using the parametric description of the beam axis and an appropriate change of variables. The practical value of our approach is that it gives an estimate of the parameters of a curvilinear beam propagating in an external electromagnetic field. These estimates are applicable until the beam becomes wide due to a particle velocity spread or multiple elastic scattering by the gas molecules.

REFERENCES

1. V. M. Bystritskiĭ and A. N. Didenko, *High-Power Ion Beams* (Énergoatomizdat, Moscow, 1984).
2. S. I. Molokovskiĭ and A. D. Sushkov, *High-Intensity Electron and Ion Beams* (Énergoatomizdat, Moscow, 1991).
3. R. Jayaram, *Mass Spectrometry. Theory and Applications* (New York, 1966; Mir, Moscow, 1969).
4. V. A. Syrovoĭ, Radiotekh. Élektron. (Moscow) **37**, 1692 (1993).
5. S. P. Andreev and A. V. Koshelkin, Dokl. Akad. Nauk SSSR **289**, 593 (1986) [Sov. Phys. Dokl. **31**, 570 (1986)].
6. N. D. Naumov, Zh. Tekh. Fiz. **62** (2), 178 (1992) [Sov. Phys. Tech. Phys. **37**, 212 (1992)].
7. S. D. Kovalev, A. I. Kuzovlev, and D. B. Rogozkin, Zh. Tekh. Fiz. **63** (6), 37 (1993) [Tech. Phys. **38**, 447 (1993)].
8. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 1: *Mechanics* (Nauka, Moscow, 1973; Pergamon, New York, 1988).
9. T. P. Hughes and B. B. Godfrey, Phys. Fluids **27**, 1531 (1984).
10. V. S. Remizovich, D. B. Rogozkin, and M. I. Ryazanov, *Fluctuations of Charged Particle Ranges* (Énergoatomizdat, Moscow, 1988).

Translated by A. Khzmalyan

Technical Physics, Vol. 45, No. 10, 2000, pp. 1325–1330. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 95–101.

The Microwave Photoconductivity of Electrons Flying over Natural Diamond

A. G. Zakharov, N. A. Poklonskiĭ, and V. S. Varichenko

Belarussian State University, Leningradskaya ul. 14, Minsk, 220050 Belarus

E-mail: heii@phys.bsu.unibel.by

Received July 13, 1999

Abstract—Natural diamonds of IIa type were studied by the methods of dc photoconductivity and microwave photoconductivity. Their optical absorption spectra under normal conditions and the low-pressure ($\approx 10^{-5}$ Pa) photoemission were measured. The microwave conductivity of diamond (having the negative surface electron affinity) was treated in terms of a model including both internal and external photoelectric effects. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The contactless technique for measuring the microwave photoconductivity (MWPC) of diamond has a number of advantages over the contact method of recording dc photoresponse [1]. Specifically, in the former, the photoresponse arises from the absorption of microwave energy by "free" carriers no matter whether they are inside or beyond the sample. It was found [2–4] that the hydrogen-saturated (111) diamond surface has the negative electron affinity. Thus, diamond appears to be a candidate material for microwave-biased photoreceivers that simultaneously exploit internal and external photoelectric effects. Photoemission studies suggest [5] that such devices will be able to detect quanta with energies above 2 eV.

In this work, we measured the MWPC value of natural diamond at wavelengths λ between 200 and 400 nm and constructed a model including both internal and external photoelectric effects. The main body of experimental data is reported for the first time; earlier, some of them were briefly discussed in [1, 6]. The MWPC model is applicable to type-IIa diamond under atmospheric pressure and at room temperature.

EXPERIMENTAL TECHNIQUES AND RESULTS

Samples used were 10- to 1000-μm-thick (111) polished platelets of natural (IIa) and close-to-natural diamonds. The concentration of nitrogen impurity in the *A* form was no more than 5×10^{19} cm^{-3}. Prior to measurements, the samples were chemically treated in a $K_2Cr_2O_7 + H_2SO_4 + H_2O$ mixture and then polished with corundum powder (grain size ≈14 μm) [1]. Isochronous annealing (30 min) of the samples was carried out at temperatures from 70 to 600°C in a graphite container under a pressure of $\approx 10^{-4}$ Pa.

The samples were studied by the methods of microwave photoconductivity, optical absorption (OA), and photoemission (PE). The first two were applied at $T \approx$ 300 K and atmospheric pressure ($\approx 10^5$ Pa). For microwave measurements, the samples were placed in an H_{101} rectangular microwave resonator so that they were within the antinode of the electric components of the field [7] (frequency f_e = 9.6 GHz, strength $E_0 \approx$ 113 V/cm [6, 8]). Nonequilibrium carriers were excited by radiation from a 1-kW high-pressure xenon lamp. The radiation passed through an MDR-12 monochromator ($\Delta\lambda$ = 1.0–2.4 nm). The microwave photoconductivity was detected under the synchronous mode at a photoexcitation modulation frequency of 300 Hz. Optical absorption spectra were recorded with a SPECORD-M40 spectrophotometer ($\Delta\lambda \approx$ 0.3 nm).

The photoemission current was generated by the optical system of the MWPC measuring setup without light modulation. The strength of the field "pulling" electrons out of the diamond was 400 V/cm. Diamond platelets of area 0.25–0.3 cm^2 and thickness ≈0.5 mm were placed in a vacuum chamber ($\approx 10^{-5}$ Pa) at an angle of 45° with incident light.

Our experimental findings are summarized as follows.

(1) The microwave photoconductivity at λ = 200–700 nm strongly depends on surface finish when the light penetration far exceeds (by 1 to 3 orders of magnitude) the diffusion length L_d of carriers in the diamond. Conversely, the absorption spectra are virtually independent of surface finish.

(2) After vacuum heat treatment at 950°C for 30 min, when the (111) surface is reconstructed, the MWPC value drops by more than one order in the fundamental absorption range ($\lambda \approx$ 220 nm) and significantly decreases in the impurity absorption range (236–700 nm). However, the absorption spectra and those of

dc photoconductivity do not change in this case. The initial MWPC values are regained only after the samples have been mechanically polished by diamond paste.

(3) Once the diamond surface had been treated by the corundum powder (grain size ≈14 μm) or diamond paste (grain size 0.5–5.0 μm), the MWPC signal always increased roughly by one order under illumination by light near the fundamental range. The photoemission current from the same samples was also detected only if the surface was treated in a like manner.

The available models of photoconductivity [9–13] fail to explain these facts. Below is represented a new MWPC model for diamond (which has the negative electron affinity of the surface [2]).

MWPC MODEL FOR DIAMOND WITH FLYING ELECTRONS

We will start with the following assumptions.

(1) The sum of flying electrons $\tilde{n}_f$ (i.e., those leaving the diamond due to interband illumination) and electrons in the conduction band $\tilde{n}_c$ equals the number of holes in the valence band; that is, $\tilde{n}_f + \tilde{n}_c = \tilde{p}$.

(2) Because of the negative electron affinity of the diamond surface, the emitted electrons are cold (i.e., they are emitted at $T = 300$ K).

(3) The diffusion length L_d of conduction electrons in the diamond is much greater than the thickness of the space-charge region (the length of electrostatic field screening in the near-surface region) [14].

(4) The electric component of a microwave field inside the sample and around it is distributed uniformly (the skin depth is $\approx 1.6 \times 10^3$ m for diamond with the photoconductivity $\Delta\sigma \approx 10^{-10}\ \Omega^{-1}\ \mathrm{cm}^{-1}$).

(5) The condition $\Delta\sigma \gg \sigma_0$ is satisfied (σ_0 is the dark microwave conductivity of diamond at $T \approx 300$ K).

(6) In polycrystalline (mosaic) type-IIa diamond [15], nonequilibrium holes are trapped by grain boundaries and do not contribute to the MWPC value; the lifetime of electrons inside a grain $\tau_n \gg 1/(2\pi f_e) \approx 1.6 \times 10^{-11}$ s [1].

Under assumptions 1–6, the contribution of photoemission to the absorption of microwave energy is defined by the sum of flying electrons $\tilde{n}_f$ and nonequilibrium conduction electrons $\tilde{n}_c$. It also depends on the ratio between the relaxation time of the flying electron momentum, τ_f, and that of the quasi-momentum of conduction electrons, τ_c, inside the sample (see below).

The volumetric density of the microwave power W absorbed by the sample is given by [7, 16]

$$W = \frac{\sigma_1 E_0^2}{2}, \tag{1}$$

where $\sigma_1 = \mathrm{Re}\sigma_{ac}$ is the real part of the electric conductivity $\sigma_{ac} = \sigma_1 + i\sigma_2$ in a microwave field with the electric component $E = E_0 \exp(i2\pi f_e t)$.

In diamond, the field strength E_0 is attenuated ε_r times, where $\varepsilon_r \approx 5.7$ is the relative permittivity.

When the MWPC value is measured with a resonator, the microwave power absorbed by the resonator without illumination is taken for the null signal. Then, during MWPC measurements, the power absorbed by the resonator can be written as

$$Q_f + Q_c = \int_{V_f} W_f\, dV + \int_{V_c} W_c\, dV, \tag{2}$$

where W_f is the volumetric density of the power Q_f that is absorbed in the volume V_f occupied by flying electrons, and W_c is that of the power Q_c that is absorbed in the volume V_c occupied by nonequilibrium photogenerated conduction-band electrons.

Since the microwave field is uniform, for heavily absorbable light (the light penetration is much less than the sample thickness), we have

$$Q_f + Q_c \approx W_f A L_f + W_c A/\alpha(\lambda), \tag{3}$$

where A is the illuminated area of the sample, L_f is the thickness of a layer to which flying electrons are confined, $1/\alpha(\lambda)$ is the depth from which nonequilibrium electrons are excited by light that has a wavelength λ and intensity $I_{\lambda,x} = I_\lambda \exp[-x\alpha(\lambda)]$ at a distance x from the surface.

The real part of the flying electron conductivity (hereafter flying conductivity for brevity) can be determined in terms of the Drude model [1, 16], according to which

$$\mathrm{Re}\sigma_D = \frac{\sigma_f}{1 + (\omega\tau_f)^2}. \tag{4}$$

Here, $\sigma_f = e^2 n_f \tau_f / m_0$ is the dc conductivity of flying electrons, e is the absolute value of the electron charge, m_0 is the mass of an electron, τ_f is the relaxation time of the flying momentum, and $\omega = 2\pi f_e$ is the angular microwave frequency.

The real part of the conductivity of conduction-band electrons can be found within the Lorentz model [16], since almost all natural diamonds of type IIa have the mosaic (fragmentary) structure [15] with a grain size d ranging from 1 to 100 μm. According to this model, the real part $\mathrm{Re}\sigma_L$ of the conductivity of nonequilibrium electrons has the form [1, 16]

$$\mathrm{Re}\sigma_L = \sigma_c \frac{(\omega/\tau_c)^2}{(\omega_0^2 - \omega^2)^2 + (\omega/\tau_c)^2}, \tag{5}$$

where $\sigma_c = e^2 n_c \tau_c / m_c$ is the dc conductivity of nonequilibrium electrons in the conduction band of diamond, $\tau_c = \mu_n m_c / e \approx 5 \times 10^{-13}$ s is the relaxation time of the

quasi-momentum of the nonequilibrium electrons, $\mu_n \approx 2000$ cm^2/(V s) is the mobility of the conduction-band electrons at $T \approx 300$ K [17], $m_c = 0.48m_0$ [12], $\omega_0 = (2/d)\sqrt{3k_BT/m_c}$ is the "oscillation" frequency of a nonequilibrium electron in a grain of size d [1, 16] until it recombines with a hole localized on a grain boundary, and k_B is the Boltzmann constant.

It was shown [1] that $\mathrm{Re}\sigma_L \approx \sigma_c$ if $f_e = 9.6$ GHz and $d > 7$ μm. All subsequent calculations within the elaborated MWPC model are therefore made for the case $d > 7$ μm.

As follows from (3)–(5), finding the explicit form of the power densities W_f and W_c requires that the relaxation times τ_f and τ_c, as well as the depths L_f and $1/\alpha(\lambda)$, be known.

Let us calculate the ratio of the microwave power Q_f absorbed by flying electrons to the total power $Q_f + Q_c$ absorbed by the sample in the resonator under illumination. From this ratio, one can estimate the contribution from flying electrons with a concentration n_f to the MWPC value (relative to that from the conduction-band electrons with a concentration n_c) as a function of incident light intensity $I(\lambda)$. The relative contribution of the flying electrons can be characterized by the participation factor

$$K_f \approx \frac{Q_f}{Q_f + Q_c}. \tag{6}$$

Under atmospheric pressure and $T \approx 300$ K, the time τ_f between collisions (the momentum relaxation time) of flying electrons with air molecules is given by [18]

$$\tau_f = \frac{1}{N_L v(\tau_f) S(\tau_f)}, \tag{7}$$

where $N_L \approx 2.7 \times 10^{19}$ cm^{-3} is the Loschmidt constant, $S(\tau_f)$ is the effective cross section of scattering of flying electrons by air molecules, $v(\tau_f) = v_0 - a_\rho \tau_f$ is the velocity of a flying electron to scattering, $v_0 = (8k_BT/\pi m_0)^{1/2}$ is the mean thermal velocity with which the electrons leave the surface, k_BT is thermal energy, and a_ρ is the electron deceleration due to positive charge (Coulomb attraction).

The normal-to-surface acceleration of an electron in the electrostatic field of the double electrical layer is expressed as [19]

$$a_\rho = \frac{e^2 N_p}{\varepsilon_0 m_0}, \tag{8}$$

where $N_p = \tilde{p}/A$ is the grain-boundary hole density on the (111) surface of diamond (see below) and A is the illuminated area.

When calculating τ_f by (7), we ignored the heating of flying electrons by the microwave field, since, at $T = 300$ K, we have $m_0 v(\tau_f)^2 = 3/2k_BT + (eE\tau_f)^2 \approx 3/2k_BT$.

The cross section S_N of electron scattering by a nitrogen molecule as a function of electron velocity v is linearly approximated as $S_N = k_1 v + k_2$, where $k_1 \approx 2.77 \times 10^{-23}$ cm s and $k_2 \approx 5.76 \times 10^{-7}$ cm^2 [20]. For an electron with an energy between 10 and 60 meV, the cross section of scattering by an oxygen molecule is almost constant [20]: $S_O = 3 \times 10^{-16}$ cm^{-2}. Thus, the cross section of scattering by an air molecule is $S(\tau_f) = 0.21S_O + 0.79[k_1 v(\tau_f) + k_2]$.

From formulas (7) and (8), the relaxation time $\tau_f \approx 8 \times 10^{-12}$ s depends on the concentration of positive charge N_p only slightly until $N_p \approx 6 \times 10^7$ cm^{-2} (Fig. 1). For the given value of τ_f and $f_e = 9.6$ GHz, formula (4) for the real part of the total conductivity of flying electrons yields $\mathrm{Re}\sigma_D \approx \sigma_f$.

Now we will estimate the time t_f during which an electron is beyond the sample vs. concentration of uncompensated holes $N_p = \tilde{p}/A$, where $\tilde{p}$ is the number of holes in the sample of surface area A.

It is known [21] that a flying electron completely loses the normal-to-surface component of its velocity, v_x, in one elastic collision on average. Then, the time t_f during which an electron is in the double electrical layer (beyond the surface) of thickness $L_f = v_x\tau_f$, where $v_x = v_0 - a_\rho\tau_f/2$, is found by equating the electron paths in the forward (from the surface to the point of scattering) and reverse directions:

$$v_0\tau_f - \frac{a_\rho\tau_f^2}{2} = \frac{a_\rho(t_f - \tau_f)^2}{2}. \tag{9}$$

Here, $t_f = \tau_f + (t_f - \tau_f)$, where τ_f is the time to scattering and $(t_f - \tau_f)$ is the return time, and v_0 is the initial velocity of the electron.

Figure 1 shows t_f and τ_f curves calculated from formulas (9) and (7), respectively. To $N_p = 6 \times 10^7$ cm^{-2}, the product of the momentum relaxation time τ_f and time of flight t_f exceeds the associated product $\tau_c\tau_n$ inside the material even if $\tau_n \approx 10^{-8}$ s and $\tau_c \approx 5 \times 10^{-13}$ s.

In our experiments, $\mathrm{Re}\sigma_D \approx \sigma_f$ and $\mathrm{Re}\sigma_L \approx \sigma_c$, as follows from (1)–(3) and (7)–(9); hence, the products $\tau_f t_f$ and $\tau_c\tau_n$ define the contributions of the internal and external photoelectric effects to the MWPC value. It might be expected that flying electrons make a contribution to the MWPC at $N_p < 6 \times 10^7$ cm^{-2} [formula (16)].

The model of double electrical layer produced by electrons and holes is applicable if the maximum distance L_f of flying electrons to the surface is much below the linear size of the sample $\sqrt{A} \approx 0.5$ cm. The relaxation time of the flying momentum $\tau_f \approx 10^{-11}$ s depends on the illumination intensity only weakly (Fig. 1, curve *2*). Within the time τ_f, the normal-to-surface component of

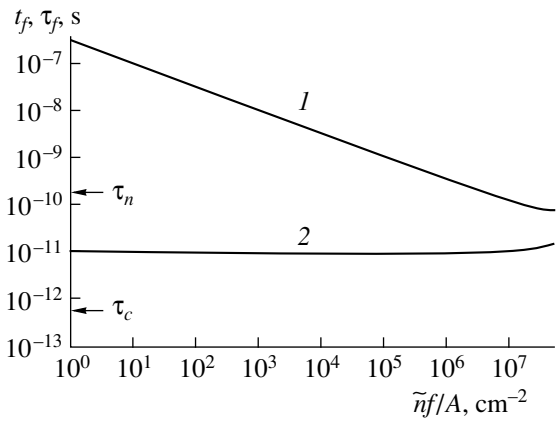


Fig. 1. Lifetime t_f (1) and momentum relaxation time τ_f (2) for flying electrons vs. density of uncompensated holes in type-IIa natural diamond.

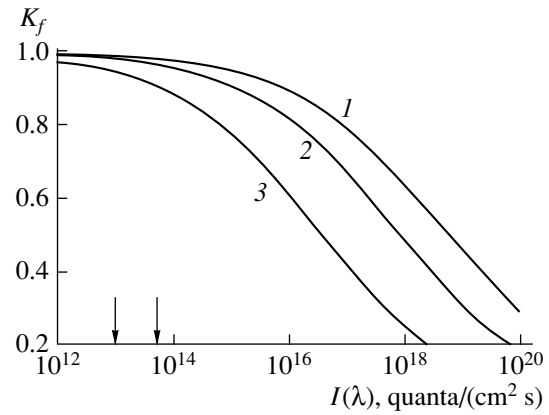


Fig. 2. Participation factor K_f (16) vs. light intensity for diamond with the characteristic lifetime of conduction-band electrons $\tau_n = 10^{-10}$ s for λ = (*1*) 220, (*2*) 225, and (*3*) 230 nm. The intensity range $I(\lambda) = (1–6) \times 10^{13}$ quanta/(cm^2 s) at λ = 220–230 nm (with the 4-mm-wide slits of the MDR-12 monochromator fully opened) is marked by vertical arrows. The upper range of $I(\lambda)$ does not exceed that of $\tilde{n}_f$ in Fig. 1. The correspondence between $I(\lambda)$ and $\tilde{n}_f/A$ is established by formula (14).

the mean electron velocity is $v_x = v_0 - a_\rho \tau_f/2$. Then, the thickness of the double electrical layer is expressed as

$$L_f = v_x \tau_f = (v_0 - a_\rho \tau_f/2)\tau_f = a_\rho (t_f - \tau_f)^2/2 \approx 0.5\ \mu\text{m}, \quad (10)$$

that is, L_f is much less than the sample dimension $\sqrt{A}$.

Having found τ_f and L_f, and taking into account that $\mathrm{Re}\sigma_D \approx \sigma_f$ and $\mathrm{Re}\sigma_L \approx \sigma_c$, the total power absorbed in the resonator is [formula (3)]

$$Q_f + Q_c = \frac{A}{2} E_0^2 \left\{ \sigma_f L_f + \frac{\sigma_c}{\varepsilon_r^2 \alpha(\lambda)} \right\}. \quad (11)$$

In view of (11), expression (6) takes the form

$$K_f \approx \left(1 + \frac{\tilde{n}_c \tau_c}{15.6 \tilde{n}_f \tau_f} \right)^{-1}, \quad (12)$$

where $\tilde{n}_f = n_f A L_f$ and $\tilde{n} = n_c A/\alpha(\lambda)$ are the numbers of flying electrons and nonequilibrium electrons in the conduction band, respectively; $\varepsilon_r^2 m_c/m_0 \approx 15.6$; A is the illuminated area; L_f is the mean distance of the electrons to the surface; and $\alpha(\lambda)$ is the light absorption coefficient at a wavelength λ.

To determine the contribution of flying electrons to the MWPC, we will first find the dependences of N_p, $\tilde{n}_f$, and $\tilde{n}_c$ on light intensity $I(\lambda)$. It is assumed that the intensity of light with a wavelength λ at a distance x inward to diamond is expressed as $I(\lambda, x) = I(\lambda) \exp[-\alpha(\lambda)x]$. Also, the current density J_e of emitted electrons is assumed to be equal to the diffusion component J_d of the current density (flat-band approximation); $J_e \approx J_d = eD_n dn_c/dx$. Under the steady-state illumination of the diamond surface, the current density of emitted electrons J_e equals that of returning electrons $J_r = en_f v_f$, where $v_f = a_\rho(t_f - \tau_f)$ is the velocity gained by a flying electron due to Coulomb attraction for the time $t_f - \tau_f$ (recall that a flying electron is assumed to lose its normal-to-surface velocity in one elastic collision with an air molecule [21]).

Since the motion of electrons in diamond is random, 1/6 of them tend to the surface and may leave the crystal. The electrons are emitted from a layer of thickness $L_d = \sqrt{D_n \tau_n}$, where $D_n = k_B T \mu_n/e$ is the diffusion coefficient and τ_n is the mean lifetime of conduction-band electrons to recombination with holes localized at grain boundaries. The probability of a nonequilibrium electron reaching the surface depends on the distance x to the surface and diffusion length L_d as $\exp(-x/L_d)$. The depth of generation of electrons and holes in the sample is $\approx 1/\alpha(\lambda)$. A layer of thickness dx absorbs $dI(\lambda, x) = \alpha(\lambda) I(\lambda, x) dx$ photons [21]. Then, at steady-state conditions and $H \gg L_d$ and $1/\alpha(\lambda)$ (H is the sample thickness), the total number N_f of leaving and returning electrons per unit time takes the form

$$N_f \approx \frac{1}{6}(1-r)\eta p_e \alpha(\lambda) I(\lambda) A \int_0^H \exp[-(\alpha(\lambda) + 1/L_d)x]\,dx \approx \frac{1}{6}(1-r) I(\lambda) \eta p_e \frac{\alpha(\lambda) L_d}{[1 + \alpha(\lambda) L_d]} A. \quad (13)$$

Here, r is the diamond reflectivity (in the fundamental absorption range, $(1 - r) \approx 0.7$ [15]), η is the internal photoelectric yield for illumination of intensity $I(\lambda)$, $\alpha(\lambda)$ is the absorption coefficient at a wavelength λ, p_e

is the mean probability of a conduction electron leaving the sample without reflecting back to the sample and recombining with holes at grain boundaries, and A is the illuminated area.

The number of electrons beyond the sample $\tilde{n}_f$ is equal to the number of electrons leaving diamond per unit time N_f multiplied by the mean time of flight t_f:

$$\tilde{n}_f = N_f t_f. \quad (14)$$

Expressions (14), (13), and (7)–(9) specify the dependence of the surface density of leaving electrons $N_p = \tilde{n}_f / A$ on illumination intensity $I(\lambda)$.

By analogy with (14), the number of nonequilibrium electrons in the conduction band $\tilde{n}_c$ depends on the intensity of heavily absorbable light $I(\lambda)$ as [11, 12]

$$\tilde{n}_c = (1 - r)\eta I(\lambda)\tau_n A. \quad (15)$$

From (6), (7), and (12)–(15), the participation factor of flying electrons can be written in the form

$$K_f = \left[1 + \frac{\tau_n \tau_c (1 + \alpha(\lambda)\sqrt{D_n \tau_n})}{2.6 p_e t_f \tau_f \alpha(\lambda)\sqrt{D_n \tau_n}}\right]^{-1}, \quad (16)$$

where $\sqrt{D_n \tau_n} = L_d$ is the diffusion length of electrons in the conduction band.

Figure 2 plots K_f against intensity $I(\lambda)$ for λ = 220, 225, and 230 nm. The lifetime of nonequilibrium electrons is set equal to $\tau_n = 10^{-10}$ s for $p = 0.1$ (the approximate quantum yield of photoemission [14] when the penetration depth $1/\alpha(\lambda)$ is shorter than or equal to the diffusion length L_d). The values of $\alpha(\lambda)$ for λ = 220, 225, and 230 nm were found from experimental data in [22]. A decrease in K_f with growing intensity is due to a decrease in the time of flight t_f of electrons.

Formula (16) for K_f implies that any surface modification that changes its electron affinity (contamination, reconstruction, oxidation, etc.) but does not "noticeably" affect the optical absorption spectra may lead to the following.

(1) In the interband excitation range (λ = 200–236 nm), the MWPC long- and short-wavelength parts will be changed to the least and greatest extent, respectively, since K_f drops with increasing λ. A decrease in K_f means a decreased contribution from the external photoelectric effect to the MWPC value (this contribution depends on the surface electron affinity).

(2) At λ = 220 nm, the MWPC signal must drop by one order of magnitude or more, since K_f lies in the 0.88–0.97 interval for λ between 220 and 230 nm. In the calculation, we used the intensity value λ = 220 nm $I(220) \approx 5 \times 10^{13}$ quanta/(cm^2 s) and the mean lifetime of electrons in the conduction band of natural diamond crystals τ_n = 150 ps [14].

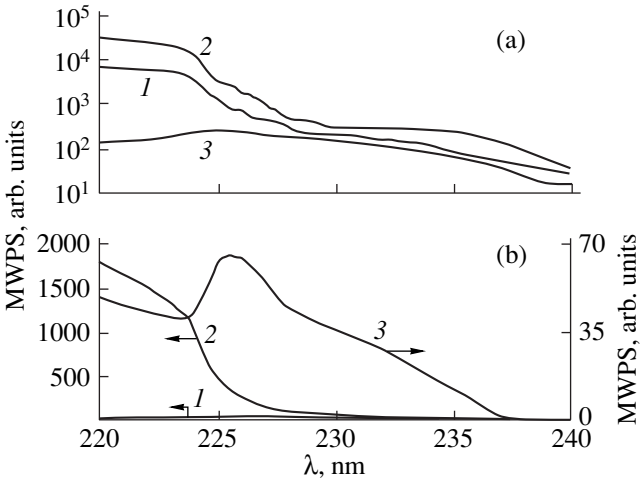


Fig. 3. (a) MWPC spectra for the isochronous (30 min) annealing temperatures (*1*) 100, (*2*) 900, and (*3*) 1200°C. The spectral width of the slit is (*1*) 0.48, (*2*) 0.36, and (*3*) 2.4. The slit width was selected such that noise did not exceed 5% of the valid signal peak. (b) MWPC spectra after (*1*, *3*) chemical cleaning in the $K_2Cr_2O_7 + H_2SO_4 + H_2O$ mixture and (*2*) powder treatment. Curve *3* is the more detailed representation of curve *1*.

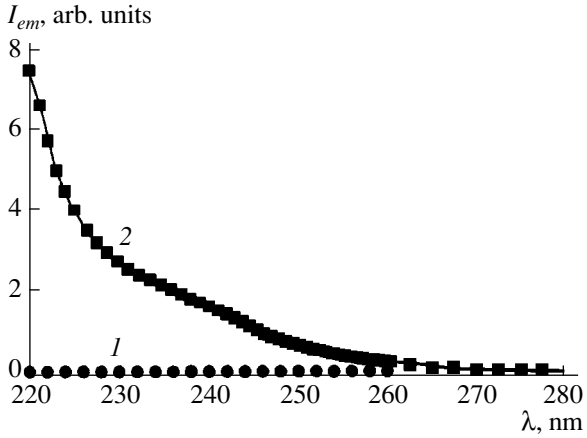


Fig. 4. Spectral distribution of the photoemission current I_{em} in a 5 × 7 × 0.15-mm type-IIa diamond sample. (*1*) As-prepared sample and (*2*) the sample after powder treatment (grain size 14 μm). The spectral width of the slit is $\Delta\lambda \approx$ 10 nm.

If an MWPC model includes internal photoelectric effect alone [9–13], the MWPC spectrum must not change (at least for the majority of natural diamonds of type IIa) at $\lambda > 220$ nm (the change is related to that in the wavelength-dependent surface recombination rate). In fact, the depth of surface influence roughly equals the diffusion length L_d. For the most perfect samples ($\tau_n \approx 10^{-8}$ s, $\mu_n \approx 2400$ cm^2/(V s), $L_d \leq 7$ μm, while in "conventional" type-IIa crystals ($\tau_n \approx 1.5 \times 10^{-10}$ s, $\mu_n \approx$ 2000 cm^2/(V s), $L_d \leq 1$ μm [14]. At the same time, the penetration depths are $1/\alpha(220) \approx 8$, $1/\alpha(225) \approx 20$, and $1/\alpha(230) \approx 100$ μm [22].

DISCUSSION

The results obtained in [1, 6, 8] will be treated in terms of our model, which takes into account the contribution of the external photoelectric effect to MWPC.

From Fig. 3a, one can see that vacuum heat treatment at temperatures above 1000°C (the reconstruction of the (111) diamond surface is observed at 950°C [17]) changes the long-wavelength part of the MWPC spectrum to the least extent. On the contrary, in the fundamental absorption range ($\lambda < 225$ nm), the MWPC drops by one order. Chemical cleaning [1] used to prepare the samples for MWPC contact measurements leads to similar results (Fig. 3b). Both heat treatment at 900–1000°C and chemical treatment in the chromium mixture cause the MWPC value to decrease in the "impurity illumination" range. However, in both cases, the optical absorption spectra remain unchanged within the error of $\alpha(\lambda)$ measurement. The initial MWPC signals are regained once a ~10-μm-thick layer has been removed by grinding and polishing the samples with the corundum powder and diamond paste. Similar changes in the MWPC spectrum were also observed when the diamond surface was bombarded by Ar^+ ions.

The validity of our MWPC model is also confirmed by the photoemission studies. In the range $280 > \lambda > 200$ nm, we failed to detect the photoemission at a level of 1 pA (the detectability limit of photoemission current) after chemical cleaning of the surface (as in [2, 23]). The photoemission becomes detectable in this spectral range if the samples are treated either by the powder or by the paste prior to measurements (Fig. 4). As was noted, the MWPC signal from the powder-treated diamond-IIa surface (Fig. 3) increases roughly tenfold [1].

CONCLUSION

The MWPC of natural type-IIa diamond was first studied by a contactless method. The MWPC results are compared with those obtained by the methods of dc photoconductivity, optical absorption under normal conditions, and vacuum photoemission. The MWPC model for natural diamond was elaborated. It takes into account both the optical generation of nonequilibrium electrons and grain-boundary-localized holes and the emission of electrons from the illuminated diamond surface, having the negative electron affinity.

REFERENCES

1. A. G. Zakharov, N. A. Poklonskiĭ, V. S. Varichenko, and A. G. Gontar', Fiz. Tverd. Tela (St. Petersburg) **42**, 647 (2000) [Phys. Solid State **42**, 664 (2000)].
2. F. J. Himpsel, J. A. Knapp, van Vechten, and D. E. Eastman, Phys. Rev. B **20**, 624 (1979).
3. O. S. Moryakov, V. B. Kvaskov, V. V. Gorbachev, *et al.*, Obz. Élektron. Tekh., Ser. 2: Poluprovodn. Prib. **4** (1342) (1988).
4. J. van der Weide, Z. Zhang, P. K. Baumann, *et al.*, Phys. Rev. B **50**, 5803 (1994).
5. J. Ristein, W. Stein, and L. Ley, Phys. Rev. Lett. **78**, 1803 (1997).
6. V. M. Kulakov, N. A. Tat'yanina, R. B. Zezin, *et al.*, in *Noble Metals and Diamond in Advanced Domains of Technology: Proceedings of the State Institute for Diamond and Gold* (Moscow, 1991), pp. 181–188.
7. V. A. Milyaev and V. A. Sanina, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **23**, 407 (1980).
8. A. G. Zakharov, V. S. Varichenko, and A. G. Gontar', Sverkhtverd. Mater., No. 2, 7 (1997).
9. A. N. Vasil'ev and V. V. Mikhaĭlov, *The Introduction to Solid State Spectroscopy* (Mosk. Gos. Univ., Moscow, 1987).
10. S. M. Ryvkin, *Photoelectric Effects in Semiconductors* (Fizmatgiz, Leningrad, 1963; Consultants Bureau, New York, 1964).
11. R. H. Bube, *Photoconductivity of Solids* (Wiley, New York, 1960; Inostrannaya Literatura, Moscow, 1962).
12. V. L. Bonch Bruevich and S. G. Kalashnikov, *The Physics of Semiconductors* (Nauka, Moscow, 1990).
13. P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors* (Springer-Verlag, Berlin, 1999).
14. *Natural Diamonds of Russia*, Ed. by V. B. Kvaskov (Polyaron, Moscow, 1997).
15. G. B. Bokiĭ, G. N. Bezrukov, Yu. A. Klyuev, *et al.*, *Natural Diamonds and Synthetic Diamonds* (Nauka, Moscow, 1986).
16. P. Grosse, *Freie Electronen in Festköpern* (Springer-Verlag, Heidelberg, 1979; Mir, Moscow, 1982).
17. *Physical Properties of Diamonds: Reference Book*, Ed. by N. V. Novikov (Naukova Dumka, Kiev, 1987).
18. *CRC Handbook of Chemistry and Physics* (Springer-Verlag, Berlin, 1999).
19. S. G. Kalashnikov, *Electricity* (Nauka, Moscow, 1977).
20. *Physical Magnitudes: Reference Book*, I. S. Grigor'ev and E. Z. Meĭlikhov (Énergoatomizdat, Moscow, 1991).
21. P. S. Kireev, *The Physics of Semiconductors* (Vysshaya Shkola, Moscow, 1975).
22. D. R. Kania, L. S. Pan, P. Bell, *et al.*, J. Appl. Phys. **68**, 124 (1990).
23. B. B. Pate, Surf. Sci. **165**, 83 (1986).

Translated by V. Isaakyan

Technical Physics, Vol. 45, No. 10, 2000, pp. 1331–1334. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 102–105.

Phase Transformations and Changes in the Electrophysical Properties of Ti–Si Contacts Irradiated by a Nitrogen–Hydrogen Plasma

A. M. Chaplanov and E. N. Shcherbakova

Institute of Electronics, Belarussian Academy of Sciences, Minsk, 220841 Belarus

E-mail: chap@inel.bas-net.by

Received June 22, 1999; in final form, February 3, 2000

Abstract—Phase transformations in a Ti–Si system irradiated by a nitrogen–hydrogen plasma were investigated with electron diffraction and Auger spectroscopy. Optimum conditions for the formation of TiN thin films on a silicon wafer were found. Contacts irradiated by a nitrogen–hydrogen plasma exhibit the lower Schottky barrier, and their breakdown voltage is almost one order of magnitude higher than in as-prepared samples. Relationships between the electrophysical properties of the contacts and the parameters of plasma processing were set. © *2000 MAIK "Nauka/Interperiodica".*

The extensive application of a low-pressure low-temperature gas-discharge plasma (including chemically active) is a distinct feature of submicron IC technology. Of particular interest is the processing of thin-film systems in a nitrogen-containing plasma to form a nitride-containing (in particular, TiN) barrier layer. It was reported [1–3] that titanium nitride meets the requirements for barrier contacts to active IC elements and is compatible with both conventional aluminum metallization and young copper-based metallization.

In this paper, phase transformations in the Ti–Si system due to nitrogen–hydrogen plasma processing were investigated with an EMR-102 electron diffractometer and a PHI-660 scanning Auger spectrometer. The electrophysical parameters of Ti–Si contacts vs. plasma processing conditions were determined from the current–voltage characteristics (CVCs) or by calculations outlined in [4, 5]. 100-nm-thick polycrystalline titanium films were grown on *p*-Si(111) wafers by electron-beam deposition. Resulting Ti–Si contacts were processed by an arc plasma containing nitrogen and hydrogen ions. The sample temperature was 500, 600, or 700°C, depending on the cathode current. The pressure of the nitrogen–hydrogen mixture was kept at a level of 5–6 Pa. The hydrogen content was 5% [6]. The ion current was measured by a Langmuir probe; the ion current density was $J_I = 4.0 \pm 0.2$ mA/cm^2. The accumulated dose N_S was calculated as [7]

$$N_S = J_I t/q,$$

(where t is the processing time and q is the ion charge) and ranged from 5×10^{18} to 5×10^{19} cm^{-2} for the processing times within 3–30 min.

The electron diffraction investigations showed that the titanium films were fine-grain polycrystals with an average grain size of 5–10 nm. To obtain the temperature dependence of the phase transformations, the dose was kept constant: $N_S = 1.5 \times 10^{19}$ cm^{-2}. It follows from the electron diffraction patterns that, when the Ti–Si system is plasma-processed at 500°C, TiN and Ti_2N form on its surface. In increasing the processing temperature to 600°C, the low-nitrogen nitride phase disappears and a gold-colored TiN film covers the silicon surface. As the processing temperature increases to 700°C, the phase composition of the Ti–Si system surface remains the same. The results obtained from the electron diffraction patterns are presented in Table 1.

To study the dose dependence of the Ti–Si phase composition, the films were processed at a constant temperature (600°C) for 3, 10, and 30 min. After irradiating the Ti films at $N_S = 3 \times 10^{18}$ cm^{-2}, the diffraction rings of titanium totally disappear and the diffraction patterns consist of the rings typical of titanium nitride TiN. As the irradiation dose rises to 1.5×10^{19} and 5×10^{19} cm^{-2}, the phase composition of the Ti–Si system does not change. Thus, irrespective of the irradiation dose, the plasma processing of the titanium–silicon contact at 600°C results in the formation of titanium nitride TiN on its surface.

Depth profiling of chemical elements was performed with Auger spectroscopy combined with argon ion etching. It was found that the as-prepared Ti films contain a large amount of oxygen (~55%) and carbon (~15%) impurities on their surface. This is due to adsorption of residual gases during film deposition owing to gettering properties of titanium. At a depth of ~30 nm, the films become free of carbon and the oxygen concentration is greatly reduced to 5% throughout the film (Fig. 1a). At the Ti/Si interface, the oxygen concentration slightly grows (suggesting that an oxide

1063-7842/00/4510-1331$20.00

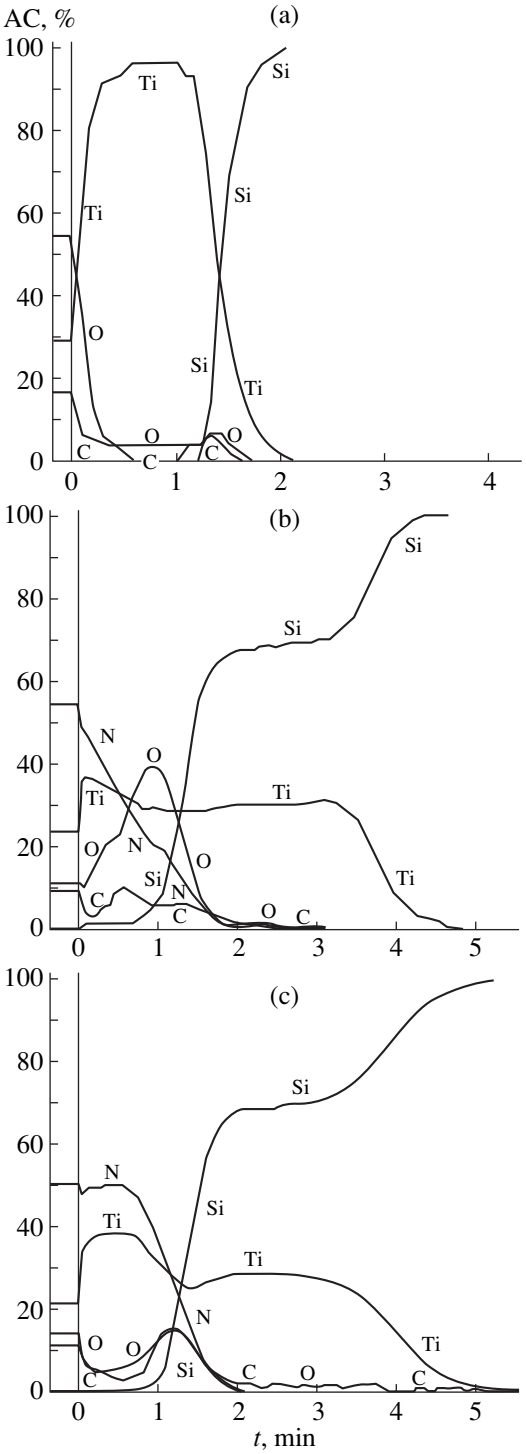


Fig. 1. Auger profiles in the Ti–Si system for (a) as-prepared samples and those processed at (b) 500 and (c) 600 °C. *t* is the time of argon ion sputtering and AC is the Auger concentration.

layer is present on the silicon) and the carbon concentration reaches ~8%.

After irradiation at $T = 500$°C, the oxygen and carbon contents at the film surface significantly decrease as compared to the as-prepared samples. This may result from the spattering effect of plasma ions and neutralization of oxygen by hydrogen ions. The oxygen rapidly diffuses from the surface inward to the film, and its concentration peaks at a depth of ~60 nm (Fig. 1b). In addition, silicon diffuses from the wafer into the Ti film, which is associated with the thermal effect of the plasma. The silicon concentration in the film is 5–10%; according to the constitutional diagram, this is insufficient for the formation of a Ti–Si compound. The concentration of nitrogen, diffusing into the titanium film from the plasma, linearly declines from 55% at the Ti surface to its minimum at the Ti/Si interface. The concentrations of nitrogen and titanium become equal at depths of 10–15 nm, which means that titanium nitride TiN is present in the system. As the interface is approached, the nitrogen concentration drops to its stoichiometric concentration in titanium nitride Ti_2N. Titanium diffuses into the silicon wafer to a depth of nearly 120 nm under the thermal action of the plasma. Its concentration corresponds to stoichiometric $TiSi_2$. At larger depths, the titanium concentration decreases, reaching zero at depths of 250–300 nm.

For the Ti–Si system processed at 600°C, the depth profiles somewhat change. The nitrogen concentration (~50%) remains constant to a depth of ~50 nm (Fig. 1c). The Ti/N concentration ratio corresponds to titanium nitride TiN with excessive nitrogen within the homogeneity region. This indicates that, under the given treatment conditions, the resulting films consist totally of titanium nitride. As compared with the samples processed at $T = 500$°C, the titanium concentration in the silicon wafer slightly rises because of an increase in the diffusion rate at the elevated temperature. The silicon profile remains intact; its diffusion into the metal film is insignificant as before. The oxygen and carbon concentrations peak (15%) on the surface and at the interface. In the interior of the film, they are 5%. The total oxygen concentration in the Ti–Si system decreases as compared to the samples processed at 500°C. It seems that, at $T > 600$°C, the oxygen adsorbed on the film surface extensively interacts with hydrogen ions of the plasma and then is removed from the surface of growing titanium nitride. The steps in the oxygen and carbon profiles at the Ti/Si interface are due to the gettering effect and also to the fact that the interfacial concentrations of these elements in the as-prepared compositions were maximum. At the processing temperature 700°C, the depth profiles of the elements (with the exception of nitrogen) are almost the same; the nitrogen concentration rises only slightly at the surface and becomes less uniform in depth.

The Auger data are consistent with those of electron diffraction. During the plasma processing of the Ti–Si system, it undergoes temperature-dependent phase transformations according to the following scheme:

$$\mathrm{Ti/Si} \xrightarrow{500^\circ\mathrm{C}} \mathrm{TiN} + \mathrm{Ti_2N/TiSi_2/Si} \xrightarrow{600\text{–}700^\circ\mathrm{C}} \mathrm{TiN/TiSi_2/Si}.$$

Table 1. Phase compositions of the Ti–Si system surface during nitrogen–hydrogen plasma processing

d_{exp}, nm	d_{theor}, nm	*hkl*	Before processing	Processing temperature, °C		
				500	600	700
0.260	0.259	101	–	Ti_2N	–	–
0.258	0.256	010	Ti	–	–	–
0.245	0.247	200	–	Ti_2N	–	–
0.243	0.244	111	–	–	TiN	TiN
0.226	0.229	111	–	Ti_2N	–	–
0.222	0.224	011	Ti	–	–	–
0.210	0.212	200	–	TiN	TiN	TiN
0.178	0.179	211	–	Ti_2N	–	–
0.152	0.151	002	–	Ti_2N	–	–
0.148	0.1496	220	–	TiN	TiN	TiN
0.131	0.133	103	Ti	–	–	–
0.127	0.1277	311	–	TiN	TiN	TiN
0.126	0.125	213	–	Ti_2N	–	–
0.122	0.1233	201	Ti	–	–	–
0.120	0.122	222	–	TiN	TiN	TiN
0.094	0.0946	211	Ti	–	–	–
0.092	0.094	420	–	TiN	TiN	TiN

Titanium disilicide at the Ti/Si interface results from a high-temperature metal–silicon reaction, which is enhanced by the plasma thermal effect. To trace the phase composition with depth, the electron diffraction studies were performed after etching off a 100- to 150-nm-thick surface layer. The associated patterns from the samples processed by the nitrogen–hydrogen plasma with $N_S = 1.5 \times 10^{19}$ cm^{-2} at 600 and 700°C show that titanium disilicide of S-54 modification arises at the Ti/Si interface.

The Auger data suggest that the surface film of titanium nitride is a reliable barrier to silicon diffusion. As follows from Figs. 1b and 1c, the region of titanium disilicide formation extends from the interface inward to the silicon wafer to a depth of 120–140 nm. The formation of silicides on the contact surface, which usually takes place on heating to 500–700°C [8, 9], was not observed. Thus, during nitrogen–hydrogen plasma processing of the titanium–silicon system, titanium is a dominant diffusant. It penetrates deeply into the silicon wafer and forms S-54 titanium disilicide as a result of the high-temperature reaction, enhanced by the plasma thermal effect.

The phase transformations in the Ti–Si system change the electrophysical parameters of the contact. The as-prepared samples had an asymmetric CVC, and their breakdown voltage was 8 V. After plasma processing at $N_S = 1.5 \times 10^{19}$ cm^{-2}, the breakdown voltage becomes as high as 50, 70, and 90 V for the processing temperatures 500, 600, and 700°C, respectively (Table 2). A rise in the breakdown voltage was also observed after the processing of the contacts at a fixed (600°C) temperature with irradiation doses of 5×10^{18}–5×10^{19} cm^{-2}. It seems likely that the postprocessing increase in the breakdown voltage of the contacts is a result of the phase transformations at the Ti/Si interface, since they substantially change its electronic structure.

In parallel with the change in the breakdown voltage, there also are changes in the ascending branches of the CVCs, as demonstrated by the Schottky barrier heights φ_b. Calculations showed that φ_b decreases from 0.62 eV for the as-prepared contacts to 0.58–0.60 eV after plasma processing. Presumably, this is related to the formation of $TiSi_2$ at the metal–semiconductor interface, which alters the surface state density.

To conclude, our investigations showed that nitrogen–hydrogen plasma processing of the Ti–Si system causes titanium nitride to form on the silicon surface and S-54 titanium disilicide to grow at the interface. It was found that these phase transformations govern the electrophysical characteristics of the contacts. When compared to the as-prepared Ti–Si contacts, the processed ones offer increased breakdown voltages and decreased Schottky-barrier heights. Thus, nitrogen–hydrogen plasma processing makes it possible to form barrier contacts based on titanium nitride and produce Schottky barriers with electrophysical characteristics depending on treatment conditions. The advantages of this technique are highly local energy input, selectivity, as well as no need for additional annealing, which is necessary when titanium nitride is deposited by conventional means. The use of a low-temperature arc plasma makes it possible to form thin films of refrac-

Table 2. Changes in the electrophysical parameters of the Ti–Si contacts during nitrogen–hydrogen plasma processing

Parameter	As-prepared sample	$N_S = 1.5 \times 10^{19}$ cm^{-2}			600°C		
		500°C	600°C	700°C	5.0×10^{18} cm^{-2}	1.5×10^{19} cm^{-2}	5.0×10^{19} cm^{-2}
φ_b, eV	0.62	0.60	0.60	0.59	0.60	0.60	0.58
U, V	8	50	70	90	55	70	70

Note: φ_b, Schottky barrier height; U, breakdown voltage.

tory compounds at lower energy and reagent consumptions.

REFERENCES

1. Shi Qing Wang, Ivo Raaijmakers, and Brad J. Burrow, J. Appl. Phys. **68**, 5176 (1990).
2. Do-Heyoung Kim and Ki Bum Kim, Appl. Phys. Lett. **69**, 4182 (1996).
3. J. O. Olowolafe, Jian Li, and F. G. Colgan, Appl. Phys. Lett. **58**, 469 (1991).
4. S. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1981; Mir, Moscow, 1984).
5. E. H. Rhoderick, *Metal–Semiconductor Contacts* (Claredon, Oxford, 1988, 2nd ed.; Radio i Svyaz', Moscow, 1982).
6. B. N. Arzamasov, *Chemicothermal Metal Treatment in Activated Gas Media* (Mashinostroenie, Moscow, 1979).
7. W. Till and J. Luxon, *Integrated Circuits* (Prentice-Hall, Englewood Cliffs, 1982; Mir, Moscow, 1985).
8. *Thin Films: Interdiffusion and Reactions*, Ed. by J. M. Poate, K. Tu, and J. Meier (Wiley, New York, 1978; Mir, Moscow, 1982).
9. S. Murarka, *Silicides for VLSI Applications* (Academic, New York, 1983; Mir, Moscow, 1986).

Translated by B. Malyukov

Technical Physics, Vol. 45, No. 10, 2000, pp. 1335–1338. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 106–109.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Theoretical and Experimental Study of an Absorbable Gas Diffusion through a Fine Filter

V. A. Klimenko[1], S. A. Mulenko[2], and V. S. Ovechko[3]

[1] *National Technical University (Kiev Polytechnical Institute), Kiev, Ukraine*

[2] *Institute of Metal Physics, National Academy of Sciences of Ukraine, Kiev, 252680 Ukraine*

[3] *Kiev State University, Kiev, 252601 Ukraine*

E-mail: ovs@rpd.univ.kiev.ua

Received November 9, 1999

Abstract—A mathematical model of the diffusion efflux of an absorbable gas into vacuum through a fine membrane with a distance between pores comparable with their diameter is proposed. The model is analyzed for a steady-state regime. The results of the theoretical analysis are compared to experiment. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

A wide use of fine filters (membranes) with a distance between pores comparable with their diameter (for example, porous glasses with a porosity of $\varepsilon = 0.2$–0.4) in physical investigations and technological processes accounts for the importance of the mathematical modeling of the diffusion of absorbable gas molecules through such filters. A distinctive feature of this process is that molecules of the absorbable gas penetrate into pores as a result of diffusion over the filter surface, as well as directly from the gas phase. Therefore, the boundary conditions must take into account the adsorption of gas molecules and their evaporation back into the gas phase after some delay time τ_s (adsorption time [1]) from the irregular surface comprising pore entrance holes and areas between them (conventionally considered as flat).

Available models do not provide for an adequate theoretical analysis of such structures. Karlov *et al.* [2] modeled the process of the absorbable gas efflux into vacuum via a membrane with a narrow single through channel (pore) under the condition that

$$l \gg \lambda \gg r, \tag{1}$$

where l is the free path of molecules in the gas phase, λ is the surface diffusion length, and r is the pore radius. In this case, no boundary condition for the pore entrance hole is formulated, since direct penetration of molecules into a single pore can be neglected if condition (1) is fulfilled. Obviously, this model is insufficient for the analysis of a multiporous structure in which the area occupied by pores is comparable with that free of the pores. Similar considerations can be applied to model [3], where the mechanism of transfer of absorbable molecules through a capillary is analyzed without taking into account the diffusion of molecules from the surface adjacent to the channel entrance.

The aim of this work was to generalize the previous models to the case of a multiporous membrane, retaining the condition $l \gg \lambda$ that provides for the flow of gas molecules through pores in the Knudsen regime, whereby the number of collisions between molecules is smaller than that with the pore walls.

PROBLEM FORMULATION

Let us consider the diffusion efflux of a gas into vacuum through a plate (membrane) with fine pores. We introduce the Cartesian coordinates (x, y, z) and place the membrane at the boundary of the half-space $z > 0$. The membrane has the thickness L and a regular system of through cylindrical pores with the radius r and the distance $2R$ between their centers (Fig. 1).

In the halfspace $z < 0$, a constant pressure is kept so as to satisfy the condition (1), and at $z > L$ pressure is close to zero (vacuum). Molecules penetrate into the pores both directly from the gas phase and as a result of the surface migration with the diffusion coefficient [4]

$$D_s = \frac{\lambda_0^2}{2\tau_i^*}\exp(-\alpha U), \tag{2}$$

where λ_0 is the mean jump length of a molecule on the

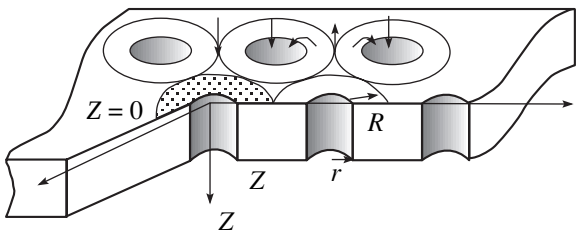


Fig. 1. Schematic diagram of a porous membrane and the motion of molecules: L is the membrane thickness, r is the pore radius, R is the radius of "collection" of molecules.

1063-7842/00/4510-1335$20.00

surface, U is the adsorption potential, τ_i^* is a constant associated with the molecule oscillation frequency in the adsorption potential, and α is a constant factor equal to 1/3–1/4.

Let us introduce a polar coordinate system (ρ, φ) with the origin at the pore center and assume that molecules moving on the surface reach the entrance of a given pore only from a ring area adjacent to the pore

$$r \le \rho \le R. \tag{3}$$

This is the main simplification that allows one to carry out the further analysis. Let us denote the surface density of molecules at $z = 0$ by $a(\rho, t)$ and at $z = L$ by $c(\rho, t)$, respectively. The transfer of molecules inside the pore proceeds both in the volume and on the surface. Assuming the total concentration of molecules $b(z)$ at a point with the coordinate of z $(0 \le z \le L)$ to be independent of ρ, we obtain

$$b(z, t) = \frac{2}{r} b_s(z, t) + b_v(z, t). \tag{4}$$

Here, the surface b_s and volume b_v concentrations are related as

$$b_s = \frac{r}{2}\frac{\tau_s}{\tau_v} b_v(z, t), \tag{5}$$

where $\tau_s = \tau_s^* \exp(U/kT)$, $\tau_s^* \approx \tau_i^*$ is the adsorption time, $\tau_v = 2r/V$ is the mean time for a molecule with the mean thermal velocity V to flight between the pore walls.

Thus, diffusion of a mixed type takes place inside the pore with a diffusion coefficient given by the formula [3]

$$D = \frac{4r^2 + 3D_s\tau_s}{3(\tau_s + \tau_v)}.$$

Let us denote by I the mean number of molecules incident onto a unit area of the membrane per unit time from the gas phase:

$$I = \frac{1}{4} nV = \frac{P}{2\sqrt{2mkT}},$$

where P is the gas pressure. Then, the process of diffusion with the uniformly distributed intensity of positive sources

$$I - \frac{a(\rho, t)}{\tau_s} > 0 \quad \text{at} \quad z = 0 \tag{7}$$

and negative sources

$$-\frac{c(\rho, t)}{\tau_s} > 0 \quad \text{at} \quad z = L \tag{8}$$

proceeds in the areas between pores.

To derive an equation of diffusion in the pore, let us assume that the pore entrance hole is characterized by the "surface" concentration of molecules

$$b_s(0, t) = a(r, t), \quad 0 \le \rho \le r, \tag{9}$$

which determines the intensity of positive sources at the pore entrance:

$$I - \frac{2}{r}\left(1 + \frac{\tau_v}{\tau_s}\right) D_s \frac{\partial a(r, t)}{\partial \rho} - \frac{b_s(\rho, t)}{\tau_s}, \quad 0 \le \rho \le r \tag{10}$$

The intensity of negative sources at the pore exit is

$$\frac{b_s(L, t)}{\tau_s} - \frac{2}{r}\left(1 + \frac{\tau_s}{\tau_v}\right) D_s \frac{\partial c(r, t)}{\partial \rho}, \quad 0 \le \rho \le r, \tag{11}$$

where a similar assumption is taken for

$$b_s(L, t) = c(r, t), \quad 0 \le \rho \le r. \tag{12}$$

As a result, we obtain the following set of equations:

$$\frac{\partial}{\partial t} a(\rho, t) = I - \frac{a(\rho, t)}{\tau_s} + D_s\left(\frac{\partial^2 a(\rho, t)}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial a(\rho, t)}{\partial \rho}\right),$$

$$\frac{\partial}{\partial t} b(z, t) = -D\frac{\partial^2}{\partial z^2} b(z, t), \tag{13}$$

$$\frac{\partial}{\partial z} c(\rho, t) = -\frac{c(\rho, t)}{\tau_s} + D_s\left(\frac{\partial^2}{\partial \rho^2} c(\rho, t) + \frac{1}{\rho}\frac{\partial}{\partial \rho} c(\rho, t)\right),$$

supplemented by the initial conditions at $t = 0$

$$\begin{aligned} a(0, \rho) &= I\tau_s \equiv a_1, \quad r \le \rho \le R, \\ b(0, z) &= 0, \quad 0 \le z \le L, \\ c(0, \rho) &= 0, \quad r \le \rho \le \infty, \end{aligned} \tag{14}$$

and by the boundary conditions at $z = 0$

$$\begin{aligned} a(R, t) &= a_1, \\ a(r, t) &= b_s(0, t), \quad 0 \le \rho \le r, \\ I - \frac{2}{r}\left(1 + \frac{\tau_v}{\tau_s}\right) D_s \frac{\partial}{\partial \rho} a(r, t) + \frac{1}{\tau_s} b_s(0, t) &= -D\frac{\partial}{\partial z} b(0, t), \\ 0 \le \rho \le r, & \end{aligned} \tag{15}$$

and at $z = L$

$$\begin{aligned} c(\rho, t) &\longrightarrow 0, \quad \rho \longrightarrow R, \\ c(r, t) &= b_s(L, t), \quad 0 \le \rho \le r, \\ -\frac{2}{r}\left(1 + \frac{\tau_v}{\tau_s}\right) D_s \frac{\partial}{\partial \rho} c(r, t) + \frac{1}{\tau_s} b_s(L, t) &= -D\frac{\partial}{\partial z} b(L, t), \\ 0 \le \rho \le r. & \end{aligned} \tag{16}$$

Now, the mixed boundary-value problem (13)–(16) for the system of differential equations of the parabolic type is correctly formulated [5] that describes the diffusion of gas molecules into vacuum through a fine filter.

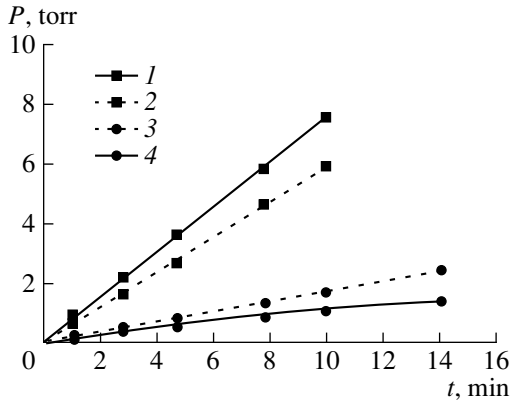


Fig. 2. Time variation of the pressure P at the membrane exit for $d = 120$ Å, $K_L = 1$, $\varepsilon = 0.4$, $P_0 = (1, 2)$ 456, $(3, 4)$ 152 torr: $(1, 4)$ experiment; $(2, 3)$ theory.

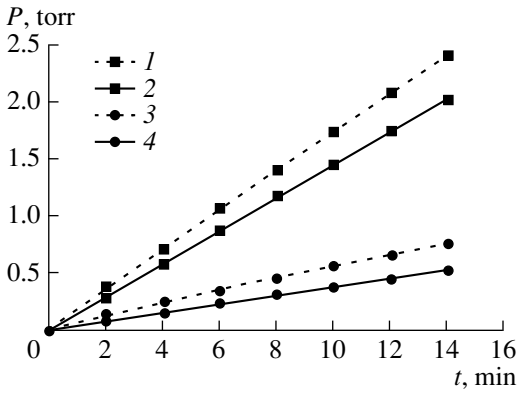


Fig. 3. The same as in Fig. 2 for $d = 67$ Å, $K_L = 2$, $\varepsilon = 0.3$ and at the same values of P_0: $(2, 4)$ experiment; $(1, 3)$ theory.

Below, we shall analyze a steady-state solution to this problem and compare this with experimental data.

CALCULATION OF A STEADY-STATE REGIME

Let us analyze a steady-state regime for the gas flow through a membrane when the concentrations of gas contained in the pores and adsorbed on the surface does not depend on time. Then, time derivatives disappear and

$$-D\frac{db(z)}{dz} = \frac{q}{\pi r^2}, \quad 0 \le z \le L,$$

where q is molecule flow through a pore.

In this case, the solution of system (13) supplemented by the boundary conditions (15) and (16) leads to the following expression for a molecular flow:

$$q = \frac{\pi r^2 I}{(1 + K_{10})^{-1} + F_{10} + \dfrac{3L}{8r}\left(1 + \dfrac{3R_1^2}{4r^2}\right)^{-1}}, \qquad (17)$$

where

$$K_{10} = 2\frac{D_s(\tau_s + \tau_v)}{rR_1}\frac{K_1(r/R_1)}{K_0(r/R_1)}, \quad R_1^2 = D_s\tau_s,$$

$$F_{10} = \left(1 + 2f_{10}\frac{D_s(\tau_s + \tau_v)}{rR_1}\right)^{-1},$$

$$f_{10} = \frac{I_1\left(\dfrac{r}{R_1}\right)K_0\left(\dfrac{R}{R_1}\right) + I_0\left(\dfrac{R}{R_1}\right)K_1\left(\dfrac{r}{R_1}\right)}{I_0\left(\dfrac{r}{R_1}\right)K_0\left(\dfrac{R}{R_1}\right) - I_0\left(\dfrac{R}{R_1}\right)K_0\left(\dfrac{r}{R_1}\right)}.$$

I_0 and K_0 are the modified zero-order Bessel functions of the 1st and 2nd kind, respectively; I_1 and K_1 are the corresponding first-order functions, respectively [6]; and R_1 is a parameter having the sense of the surface diffusion length.

Expression (17) can be used to obtain a more "rough" formula quite adequate for estimations. To make the approximation more realistic, let us replace the parameter L (pore length) in Eq. (17) with $K_L L$, where K_L is a phenomenological coefficient taking into account the so-called "sinuosity" of pores. Then we obtain from (17) the following approximate expression for the total flow Q of gas molecules through a porous membrane with the area S and the porosity ε:

$$Q \approx I\varepsilon S\frac{8r}{3Lk_L}\left(1 + \frac{3}{4}\left(\frac{R_1}{r}\right)^2\right), \quad L \gg r. \qquad (18)$$

Thus, the membrane acts as a filter passing a gas flow $Q = IS$ coming to the membrane with a transmission factor equal to $K_t = (1 + 3/4(R_1/r)^2)8r\varepsilon/3LK_L$.

EXPERIMENTAL RESULTS

To verify experimentally the above expression obtained for the transmitted gas flow Q, we studied the steady-state diffusion of argon atoms into vacuum through a porous glass membrane. The membrane was mounted (glued) in the channel connecting two cells, a pressure in one of them (low-pressure cell) was much lower than that in the other (high-pressure cell). Membranes with a diameter of 6 mm and a thickness of 2 mm, made of the porous glass plates of two types with a mean square pore diameters of ~67 ($\varepsilon \sim 0.3$) and ~120 Å ($\varepsilon \sim 0.4$), were used in the experiments.

We studied the steady-state diffusion of argon through the porous membrane at room temperature ($T = 290$ K) by measuring the gas pressure in the low-pressure cell as a function of time t at a constant pressure in the high-pressure cell equal to 152 or 456 torr.

The theoretical value of the pressure of diffusing argon as a function of time was determined as

$$P = \frac{qt}{V}kT, \tag{19}$$

where V is the volume of the low-pressure cell.

In the calculations, the potential of argon adsorption on the porous glass surface was taken to be equal to [1, p. 131]

$$U \cong 1773 \text{ K } (T = 290 \text{ K}).$$

The other parameters were as follows: $\lambda_0 = 5$ Å, $\alpha = 0.25$, $V \cong 1000$ cm^3, $\tau_s^* \approx 10^{-13}$ s, and $K_L = 1$ (for a porous specimen of the second type). The results of calculations and experiments are shown in Figs. 2 and 3.

The dependencies presented show a quite satisfactory agreement between experimental results and estimations according to formula (18). The model proposed can also be used for the numerical modeling of unsteady-state regimes of gas flow through a membrane.

REFERENCES

1. D. A. Fridrikhsberg, *A Treatise on Colloid Chemistry* (Khimiya, Moscow, 1974), p. 129.
2. N. V. Karlov, A. N. Orlov, Yu. N. Petrov, and A. M. Prokhorov, Izv. Akad. Nauk SSSR, Ser. Fiz. **49**, 500 (1985).
3. V. A. Kravchenko, A. N. Orlov, Yu. N. Petrov, and A. M. Prokhorov, Tr. Inst. Obshch. Fiz. Akad. Nauk SSSR **11**, 113 (1988).
4. F. O. Boodman, Surf. Sci. **26**, 327 (1971).
5. V. S. Vladimirov, *Equations of Mathematical Physics* (Nauka, Moscow, 1971; Dekker, New York, 1971).
6. A. N. Orlov, R. P. Petrov, and Yu. N. Petrov, Zh. Tekh. Fiz. **51**, 1681 (1981) [Sov. Phys. Tech. Phys. **26**, 966 (1981)].

Translated by M. Astrov

Technical Physics, Vol. 45, No. 10, 2000, pp. 1339–1345. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 110–116.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Construction of a Synthetic Aperture for Multiskan Photodetectors

B. G. Podlaskin, A. V. Vasil'ev, E. G. Guk, and N. A. Tokranova

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

Received December 27, 1999

Abstract—A two-aperture synthesis as applied to Multiskan photodetectors is suggested. It is intended for improving device sensitivity and reducing location error due to high-intensity background illumination. Different approaches to constructing the synthetic aperture, which performs light conversion near the saturation region of the Multiskan current–voltage characteristic, are analyzed. © *2000 MAIK "Nauka/Interperiodica".*

Current techniques for precisely determining light signal coordinates usually use the position of the median of input light distribution. The value of the median is found by comparing the weighted values of photocurrents to the left and to the right of a sought coordinate.

The potentially high accuracy of such methods (to 10^{-5} of the field of vision) can be realized only if the photodetector area is uniformly sensitive in a wide dynamic range of light intensity. In this case, the weak varying signal to be located can be separated from a high-intensity focused background.

Such an integral approach to tracing a single modulated light signal is successfully implemented with a Multiskan linear photodetector [1]. Without background illumination, this device is capable of locating the light distribution median with an accuracy to 10^{-5}–10^{-4} of the photosensitive area linear size, i.e., to 0.2–2.0 μm. However, for particular mutual arrangements of the modulated signal and a background spot, the precise location of the median becomes very difficult.

In this work, we describe the synthesis of a complex Multiskan aperture that reduces the effect of a focused background on the median location accuracy.

As is known [1], a Multiskan comprises a set of back-to-back-connected silicon photodiodes. Some of their electrodes are connected to a distributed voltage divider, and the others, to a low-resistance signal line. When a constant voltage is applied to the divider (Fig. 1), the potential is uniformly distributed over pairs of back-to-back-connected *p*–*n* junctions. The directions and values of their photocurrents make up the current–voltage characteristic (CVC) of the device. The characteristic has positive and negative saturation regions, as well as a transition region whose width (aperture) depends on the voltage applied to the resistive line. The position of the zero is defined by the zero potential point on the resistive line. Due to the feedback loop, the voltage across the resistive line is adjusted so that the total current of the Multiskan output signal is set to zero: $\sum I_s = 0$.

The feedback voltage U_{control}, which shifts the CVC zero, corresponds to the coordinate of the median in a voltage scale with great accuracy. It serves as the output signal of the measuring system.

Figure 1a shows the block diagram of a Multiskan and a device that filters, amplifies, and integrates the signal, as well as generates the control voltage U_{control}, corresponding to the median position U_s of a signal. The Multiskan is under an applied voltage $2E$. In Fig. 1b, the position of a modulated light signal with the median at a point U_s is shown. Also given is the CVC whose zero U_0 (ξ_0 in the x-axis) corresponds to U_s (ξ_s in the x-axis) in the $2E$ scale. The correspondence between the CVC zero and the light distribution median follows from the solution of the current balance equation in both the coordinate and voltage spaces:

$$I_s = \alpha\int_0^L f(x)\tanh\frac{x-\xi_0}{A_x}dx = 0$$

and

$$I_s = \beta\int_{-E}^{E} f(u)\tanh\frac{u-U_0}{A_u}du = 0, \qquad (1)$$

where L is the Multiskan length, μm; E is the voltage applied to the resistive line, V; A_x is the space width of the aperture, μm; A_u is the width of the aperture in the voltage scale, V; and α and β are proportionality factors (hereafter, we assume that $\alpha = \beta = 1$).

1063-7842/00/4510-1339$20.00

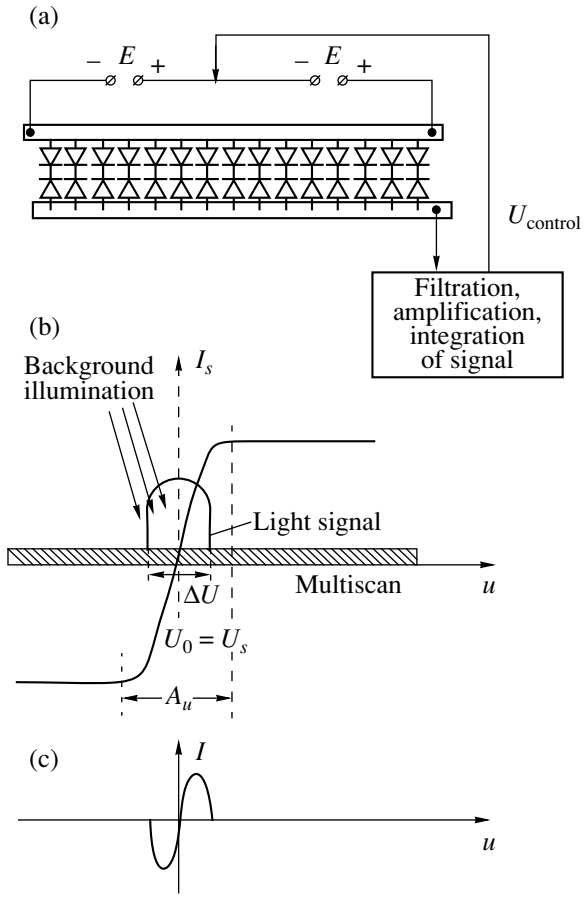


Fig. 1. Principle of Multiskan operation. (a) Regime of modulated signal location, (b) mutual arrangement of light signal and Multiskan current–voltage characteristic, and (c) positive and negative components of photocurrent generated by the modulated signal.

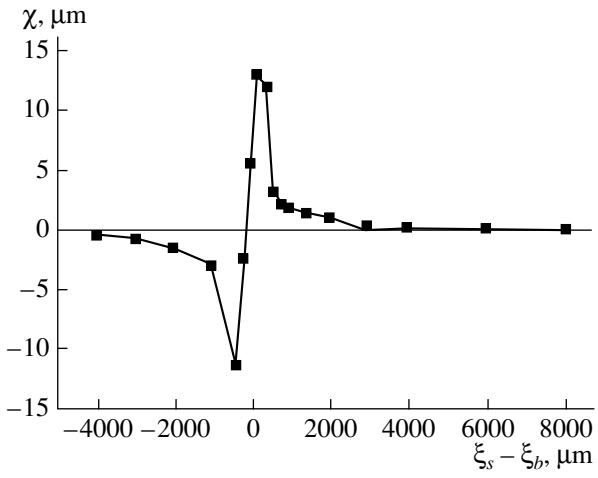


Fig. 2. Experimental dependence of error in the signal coordinate on the mutual arrangement of signal and background.

Let the signal $f(x)$ be near-rectangular; that is,

$$f(x) = \begin{cases} 1 & \xi_1 \le x \le \xi_2 \\ 0 & \xi_2 < x < \xi_1, \end{cases}$$
$$f(u) = \begin{cases} 1 & U_{\xi_1} \le u \le U_{\xi_2} \\ 0 & U_{\xi_2} < u < U_{\xi_1}, \end{cases} \quad (2)$$

where ξ_1, ξ_2 and U_{ξ_1}, U_{ξ_2} are the boundaries of the rectangular optical signal in the coordinate and voltage scales, respectively.

Subsequently, the width of the optical signal will be designated as $\delta\xi$ and ΔU, respectively. Then, the current balance equation is

$$\int_{U_{\xi_1}}^{U_{\xi_2}} \tanh\frac{u - U_0}{A_u} du = 0.$$

Hence, $U_0 = U_{\xi_1} + U_{\xi_2}/2 = U_s$. In other words, if $U_{control} = U_0$, the CVC zero coincides with the median U_s of the distribution $f(u)$.[1]

When a valid modulated signal and high-intensity focused background are simultaneously applied to the detector, the median of the background distribution in the voltage space will be designated as U_b (ξ_b in the x-axis). The mutual arrangement of the signal $f(u)$ and background $B(u)$ may be arbitrary. If both pass through the same diaphragm, the widths of the signal $\Delta\xi_s$ and background $\Delta\xi_b$ can be considered to be equal.

Experiments have shown that a modulated signal against a high-intensity focused background can, in general, be located with a high accuracy (about 2 μm). However, in the region where the signal and background distributions partially superpose, the error drastically increases. Figure 2 depicts an experimental curve that plots the measurement error ($\chi = \xi_s - \xi_0$, μm) for a stationary signal vs. the mutual arrangement of the signal and background. The Multiskan photosensitive area was scanned by a background spot whose intensity exceeded that of the signal by two orders of magnitude. The slight broadening of the curve where the background and the signal came closer together is most probably due to a spread of generated carriers within their diffusion length.

The shape of the curve and its passage through the zero $\xi_b = \xi_s$ indicate that the error arises when the symmetry of the photocurrent distribution breaks. This takes place when the signal $f(x)$ and background $B(x)$

[1] Hereafter, solutions with negative $(U_{\xi_1} + U_{\xi_2})/2$ are rejected, since the conditions for negative feedback and system stability are violated in this case.

superpose. In other words, in regions where the intensities of the weak modulated signal and background add up, the sensitivity of the photodetector changes.

Thus, the location error arises when the luminous characteristic becomes nonlinear; this, in turn, distorts the CVC of the photodetector under high-intensity illumination.

The linearity of the luminous characteristic of silicon photodiodes in a wide dynamic range is vital for photometry. The luminous characteristic of photodiodes with shallow p^+–n junctions is known to be very close to linear in the visible and near-UV ranges [2, 3]. However, it deviates from linearity in the red and near-IR ranges, which are used in applications similar to those considered in this paper. In [4], the linearity of the luminous characteristics of photodiodes for the light intensity varying within two orders of magnitude was studied. The nonlinearity was shown to be as large as 20%. Similar studies performed on Multiskan devices gave like results: when the illumination intensity was increased 100-fold, the current due to modulated light changed within 15%.

Let a modulated signal be approximated by a rectangular pulse. Then, as follows from calculations, a location error due to a signal current change in one-half of the illuminated area can be roughly estimated as

$$\chi = \frac{\Delta\xi}{4}\left(1 - \frac{1}{k}\right), \quad (3)$$

where $\Delta\xi$ is the modulated-signal width, k is the coefficient that takes into account an increase in the signal current under background illumination, and $\chi = \xi_s - \xi_0$ is the location error. (Note that this formula ignores the finite width of the CVC transition region and light-to-photocurrent conversion near the zero bias voltage.)

If, for example, the light spot size is 0.3 mm and k = 1.1, χ = 10 μm, which agrees with Fig. 2. At k = 1.05, χ = 4 μm. The use of focusing optics and sharper distributions, such as Gaussian, will further decrease the error. Yet, according to present-day requirements for location accuracy under high-intensity illumination, the error should be decreased down to 1 or 2 μm. Two approaches to linearization are possible: (1) by technological means, when the characteristic is linearized to three-place accuracy for a dynamic range of background intensity covering three orders of magnitudes and (2) using circuit-design methods that compensate for nonlinearity by means of multiaperture synthesis. Either of the approaches has disadvantages and restrictions of its own.

The undesired nonlinearity of the luminous characteristic is usually related to recombination centers in n-Si, which act as traps [3, 5]. Traps are known to be localized states in the energy gap which capture free carriers. These states may be associated with impurities and defects or may arise at the Si/SiO_2 interface. Current semiconductor technology allows the fabrication of diodes with high electrophysical parameters; however, their active area is limited. Large-area photodetectors with a low parameter spread are almost impossible to produce, since the concentration of the traps and their distribution in the silicon wafer are defined by the wafer history and previous process steps. This is especially true for multielement Multiskan devices, which have a large linear size. It was suggested [4] to apply a higher reverse bias voltage to the diodes so that each photogenerated carrier contributes to the photocurrent, thus decreasing the nonlinearity of the luminous curve to fractions of a percent. In the case of Multiskan, however, such an approach fails, since the position of the optical signal median is coincident with the position of the CVC zero ($U_0 = U_s$), so that valid and background signals interfere at low negative biases. This does not allow us to use recombinations described in [4]. It is nevertheless hoped that gettering and special anneals will make it possible to obtain devices with the less-than-1% nonlinearity of the luminous characteristic.

The essence of two-aperture synthesis, which is a circuit-design solution to the problem, is as follows. In the above method of signal location, where Multiskan photocurrents to the left and to the right of the CVC inflection point U_0 are compared, two photocurrent areas form. These areas may be illuminated by background light to various degrees (asymmetrically) (Figs. 1b, 1c). Due to the nonlinear luminous characteristic, this may break the current balance.

Let us try to create the situation where a background spot incident on the Multiskan photosensitive layer near the modulated signal is symmetric (that is, it evenly affects the positive and negative photocurrent branches). In this situation, a location error can be avoided. We will show that such a situation can be realized by means of a synthetic aperture made up of two Multiskans whose illuminances due to the modulated signal and background light are spatially identical.

Consider two Multiskans with voltages 2E applied to the resistive layers. Their signal lines are connected to each other. Let us apply a small bias ε so that all the potentials on one device are biased by $+\varepsilon$ and on the other, by $-\varepsilon$ (Fig. 3). Consider the operation of this combined device as a single photodetector. Here, the values and directions of the photocurrents appearing in the Multiskans depend on the complex aperture. It is composed of two signal coordinates shifted relative to each other along the voltage axis and two appropriately shifted CVCs.

A voltage axis forms on either Multiskan. We designate them as z-axis for one Multiskan and ν-axis for the other. Accordingly, the devices will be referred to as z- and ν-Multiskans. The positions of the signal potentials on these axes are related as $z = \nu - 2\varepsilon$. The expres-

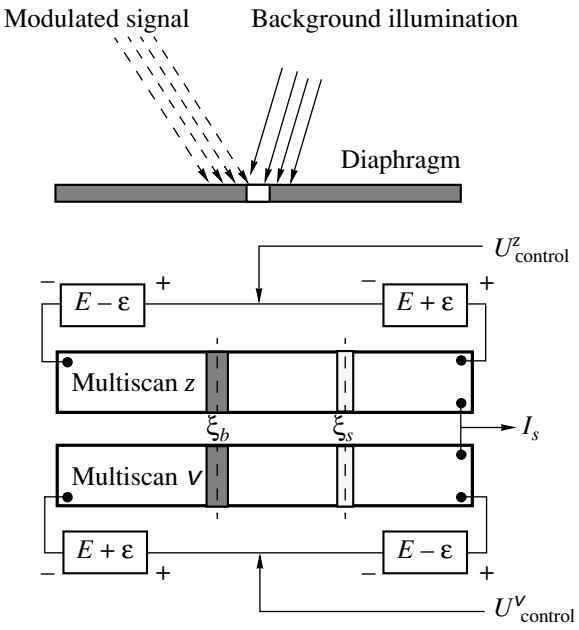


Fig. 3. Electrical connection of two Multiskans to create a synthetic aperture.

sions for currents in the z- and ν-Multiskans are

$$I^z = \int_{z1}^{z2} f(z)\tanh\frac{z-Z_0}{A_u}dz;$$
$$I^\nu = \int_{\nu1}^{\nu2} f(\nu)\tanh\frac{\nu-V_0}{A_u}d\nu. \quad (4)$$

Now we introduce the axis u such that $z = u + \varepsilon$ and $\nu = u - \varepsilon$ and write the current balance equation relative to this axis. First, we transfer the z and ν coordinates for the signals and CVCs to the u-axis. To do this, we introduce the feedback voltages U^z_{control} and U^ν_{control} into Eq. (4). These voltages are internally applied voltages that set the current balance. Then, we have

$$\int_{-E+\varepsilon}^{E+\varepsilon} f(U_{\xi1}+\varepsilon, U_{\xi2}+\varepsilon)\tanh\frac{u+\varepsilon-U^z_{\text{control}}}{A_u}du$$
$$+\int_{-E-\varepsilon}^{E-\varepsilon} f(U_{\varepsilon1}-\varepsilon, U_{\varepsilon2}-\varepsilon)\tanh\frac{u-\varepsilon-U^\nu_{\text{control}}}{A_u}du = 0. \quad (5)$$

For rectangular signals [see Eq. (2)], Eq. (5) takes the form

$$\int_{U_{\xi1}+\varepsilon}^{U_{\xi2}+\varepsilon}\tanh\frac{u+\varepsilon-U^z_{\text{control}}}{A_u}du$$
$$+\int_{U_{\xi1}-\varepsilon}^{U_{\xi2}-\varepsilon}\tanh\frac{u-\varepsilon-U^\nu_{\text{control}}}{A_u}du = 0. \quad (6)$$

Integration yields

$$\ln\frac{\cosh(U_{\xi2}+2\varepsilon-U^z_{\text{control}})}{\cosh(U_{\xi1}+2\varepsilon-U^z_{\text{control}})}$$
$$\times\frac{\cosh(U_{\xi2}-2\varepsilon-U^\nu_{\text{control}})}{\cosh(U_{\xi1}-2\varepsilon-U^\nu_{\text{control}})} = 0;$$

that is,

$$\frac{\cosh(U_{\xi2}+2\varepsilon-U^z_{\text{control}})}{\cosh(U_{\xi1}+2\varepsilon-U^z_{\text{control}})}$$
$$\times\frac{\cosh(U_{\xi2}-2\varepsilon-U^\nu_{\text{control}})}{\cosh(U_{\xi1}-2\varepsilon-U^\nu_{\text{control}})} = 1. \quad (7)$$

Solutions of Eq. (7) can be found for two cases: (1) $U^z_{\text{control}} = U^\nu_{\text{control}}$ and (2) $U^z_{\text{control}} \neq U^\nu_{\text{control}}$. In the former case, we use the equalities

$$\cosh(U_{\xi2}+2\varepsilon-U^z_{\text{control}}) = \cosh(U_{\xi1}-2\varepsilon-U^\nu_{\text{control}})$$

or

$$\cosh(U_{\xi1}+2\varepsilon-U^z_{\text{control}}) = \cosh(U_{\xi2}-2\varepsilon-U^\nu_{\text{control}}).$$

It is easy to check that, due to the evenness of the hyperbolic cosine, both equalities yield the same solution:

$$U^z_{\text{control}} = U^\nu_{\text{control}} = U_{\text{control}} = \frac{U_{\varepsilon1}+U_{\varepsilon2}}{2} = U_s. \quad (8)$$

Substituting (8) into (6) gives

$$\int_{\varepsilon-\frac{U_{\xi2}-U_{\xi1}}{2}}^{\varepsilon+\frac{U_{\xi2}-U_{\xi1}}{2}}\tanh\frac{u+\varepsilon-U_s}{A_u}du$$
$$+\int_{-\varepsilon-\frac{U_{\xi2}-U_{\xi1}}{2}}^{-\varepsilon+\frac{U_{\xi2}-U_{\xi1}}{2}}\tanh\frac{u-\varepsilon-U_s}{A_U}du = 0. \quad (9)$$

In the latter case, we take advantage of the equalities

$$\cosh(U_{\xi2}+2\varepsilon-U^z_{\text{control}}) = -\cosh(U_{\xi1}-2\varepsilon-U^z_{\text{control}})$$

or

$$\cosh(U_{\xi1}+2\varepsilon-U^\nu_{\text{control}}) = -\cosh(U_{\xi2}-2\varepsilon-U^\nu_{\text{control}}),$$

to obtain the solutions

$$U^z_{\text{control}} = U_s + 2\varepsilon \quad \text{and} \quad U^\nu_{\text{control}} = U_s - 2\varepsilon. \quad (10)$$

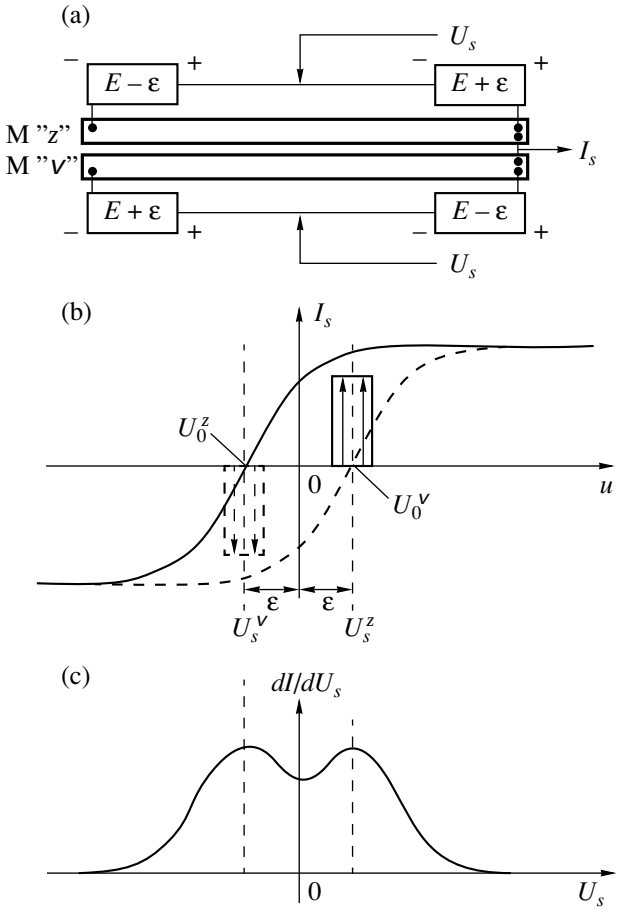


Fig. 4. Synthetic aperture for $U_{\text{control}} = U_s$. (a) Electrical connection of Multiskans, (b) spatial arrangement of the CVC for two electrically connected devices and currents generated by a modulated signal, (c) derivative of total current with respect to bias voltage.

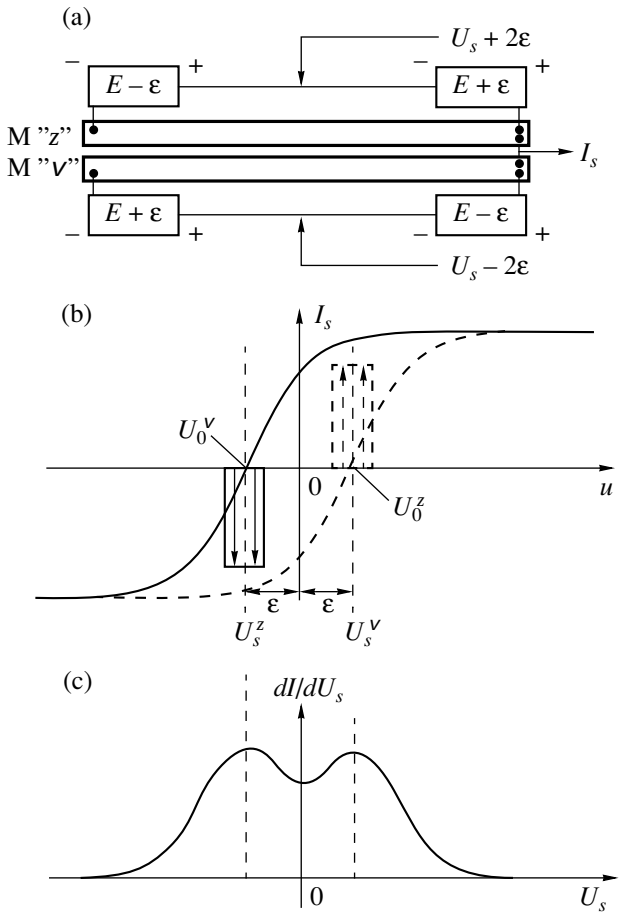


Fig. 5. Synthetic aperture for $U^z_{\text{control}} = U_s + 2\varepsilon$ and $U^v_{\text{control}} = U_s + 2\varepsilon$. (a)–(c) the same as in Fig. 4.

After substitution, we have

$$\int\limits_{-\varepsilon-\frac{U_{\xi2}-U_{\xi1}}{2}}^{-\varepsilon+\frac{U_{\xi2}-U_{\xi1}}{2}} \tanh\frac{u-\varepsilon-U_s}{A_u}du + \int\limits_{\varepsilon-\frac{U_{\xi2}-U_{\xi1}}{2}}^{\varepsilon+\frac{U_{\xi2}-U_{\xi1}}{2}} \tanh\frac{u+\varepsilon-U_s}{A_u}du = 0. \quad (11)$$

Thus, in both cases, the solutions are symmetric and, hence (in terms of the posed problem), identical (Figs. 4, 5). The solution of case 2 suggests that the synthetic aperture can be modified further by inserting additional voltages in parallel with U_{control}.

Let us assume that the sources $+\varepsilon$ and $-\varepsilon$ are connected in series with U_s, so that $U^z_{\text{control}} = U_s - \varepsilon$ and $U^v_{\text{control}} = U_s + \varepsilon$. Then, substitution into (6) yields

$$\int\limits_{2\varepsilon-\frac{U_{\xi2}-U_{\xi1}}{2}}^{2\varepsilon+\frac{U_{\xi2}-U_{\xi1}}{2}} \tanh\frac{u-U_s}{A_u}du + \int\limits_{-2\varepsilon-\frac{U_{\xi2}-U_{\xi1}}{2}}^{-2\varepsilon+\frac{U_{\xi2}-U_{\xi1}}{2}} \tanh\frac{u-U_s}{A_u}du = 0, \quad (12)$$

which means a new aperture configuration (Fig. 6).

It follows from Figs. 4b–6b that each of the synthetic aperture versions obeys the principle underlying the idea of symmetrizing the effect of background on the photocurrents of both Multiskans. Namely, the inflection points of the CVCs of both Multiskans are shifted relative to the medians of light signals incident on the photodetector, so that the photocurrent of one device is formed only with the positive CVC branch,

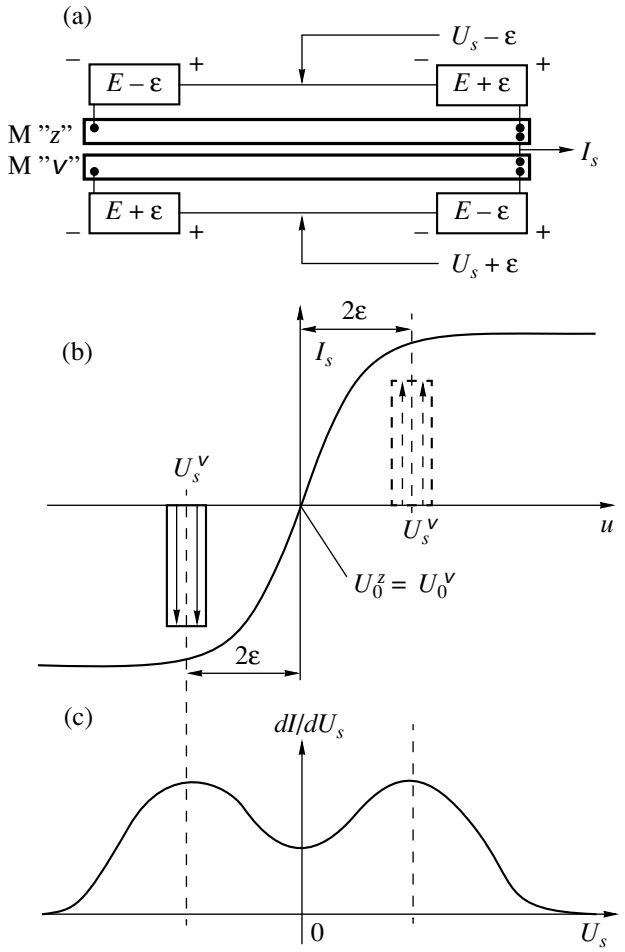


Fig. 6. Synthetic aperture for $U^z_{\text{control}} = U_s - \varepsilon$ and $U^v_{\text{control}} = U_s + \varepsilon$. (a)–(c) the same as in Fig. 4.

while that of the other is formed only with the negative one. As a result, for any position of background illumination, changes in the symmetry of photocurrent distribution take place in both the positive and negative photocurrent components, which are involved in the current balance equation and to a great extent compensate each other.

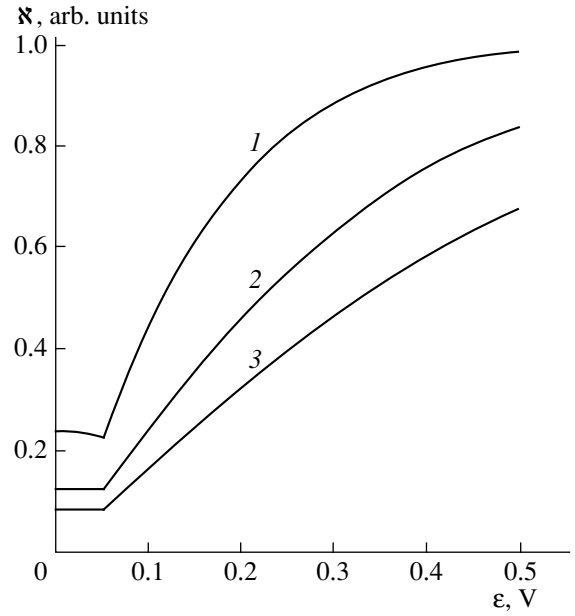


Fig. 7. Conversion efficiency vs. ε for A_u = (*1*) 0.1, (*2*) 0.2, and (*3*) 0.3 V.

Figures 4c–6c show the derivatives of the total output current with respect to voltage that shifts the signal position relative to that of the synthetic aperture. It is seen that the peaks of the derivatives coincide with the signal medians U_s. Hence, the synthetic aperture retains the property of keeping the dI/dU_s peaks at the points of the signal medians.

It should be noted that such a synthetic aperture seems to be the only practical way to increase the negative bias of the *p*–*n* junctions for valid signals, which are responsible for light-to-photocurrent conversion, in either Multiskan. This, according to [4], must improve the linearity of the luminous characteristic. Moreover, under such conditions, the conversion efficiency must greatly increase, which is tantamount to improving the system sensitivity.

The conversion efficiency for the given system is given by

$$\aleph = \frac{\left|\int\limits_{U_{\xi 1}+\varepsilon}^{U_s+\varepsilon} f(u)\tanh\frac{u-U_s-\varepsilon}{A_u}du\right| + \left|\int\limits_{U_s+\varepsilon}^{U_{\xi 2}+\varepsilon} f(u)\tanh\frac{u-U_s-\varepsilon}{A_u}du\right|}{\int\limits_{U_{\xi 1}+\varepsilon}^{U_{\xi 2}+\varepsilon} f(u)du}. \quad (13)$$

Figure 7 depicts ℵ vs. ε curves for different aperture values.

CONCLUSIONS

Thus, Multiskan devices allow the use of a synthetic aperture. With this aperture, the photocurrents from either device are formed in the positive or negative CVC branches. Moreover, the background affects both photocurrent components and their changes compensate each other. The degree of compensation depends on the differences in the sensitivities of the devices and

changes in the CVC slopes: the less the differences, the greater the compensation.

It should be noted, however, that the solutions presented here differ either by the axis of symmetry (versions 1 and 2) or by the value of ε (in version 3, ε can readily be changed to $\varepsilon/2$). One can therefore argue that any of the constructed models are suitable for the analysis of the error χ introduced by background. The aim of this analysis is to derive a dependence of χ on bias ε, aperture width A_u, signal width ΔU, and a background-induced change in the CVC shape.

This issue will be addressed in the second part of the work.

REFERENCES

1. K. F. Berkovskaya, N. V. Kirillova, and B. G. Podlaskin, *Advances in Science and Technology* (VIMI, Moscow, 1992), Vol. 2, pp. 22–25.
2. W. Budde, Appl. Opt. **18**, 1555 (1979).
3. A. R. Shaefer, E. F. Zalewski, M. A. Lind, and J. Geist, in *Proceedings Electro-Optics/Laser 77 Conference, Chicago, 1977*, p. 459.
4. A. R. Shaefer, E. F. Zalewski, and J. Geist, Appl. Opt. **22**, 1232 (1983).
5. *Solid State Imaging*, Ed. by P. Jespers, F. van de Wiele, and M. White (Noordhoff, Leyden, 1976; Mir, Moscow, 1979).

Translated by V. Isaakyan

Technical Physics, Vol. 45, No. 10, 2000, pp. 1346–1348. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 117–119.

BRIEF COMMUNICATIONS

Optical Characteristics of the Plasma of a Transverse Volume Discharge in Cl_2 and He/Cl_2 Mixtures

A. K. Shuaibov

Uzhgorod State University, ul. Podgornaya 46, Uzhgorod, 294000 Ukraine

Received November 25, 1999

Abstract—Results are presented from a study of UV and VUV emission from the plasma of a transverse volume discharge in chlorine and a He/Cl_2 mixture. In the wavelength range $\Delta\lambda$ = 140–300 nm, the $Cl_2(D'–A')$ band with an edge at 258 nm and the Cl_2^* band with edge at λ = 195 nm are found to be dominant. It is shown that, in the pressure range [Cl_2] = 0.1–2.0 kPa, the intensity of emission with $\lambda \le 195$ nm is higher than the intensity of the $Cl_2(D'–A')$ band. At [Cl_2] ≥ 2 kPa, emission in the 258-nm band is dominant. © *2000 MAIK "Nauka/Interperiodica".*

Low-pressure electric discharges in chlorine and its mixtures with inert gases are widely used in plasmochemistry to etch thin films [1]. Recently, various types of longitudinal discharges in Kr(Xe)/Cl_2 mixtures have been applied in high-power low-pressure excimer lamps [2, 3]. Chlorine molecules are an important component of the working medium of lasers operating at the ClF* 285-nm band [4] with pumping by a transverse volume discharge (TVD). In some cases, Cl_2 molecules are more preferable for pulsed TVD-based excimer lamps. Thus, in [5], it was shown that the efficiency of TVD-based spontaneous emitters operating at the ArCl(*B*–*X*) 175-nm band increases significantly if, instead of HCl molecules, Cl_2 molecules are used as chlorine carriers. Chlorine molecules also enter into the composition of the working medium of infrared HCl(v) chemical lasers with pumping by a TVD [6]. In most cases, studies of the plasma of a longitudinal glow discharge in Cl_2 were conducted at reduced pressures, because such discharges contract in strongly electronegative working media. The characteristics of pulsed lamps based on Cl_2 and He/Cl_2 mixtures (in particular, their emission at $\lambda \le 200$ nm) are poorly investigated. Such studies are important for developing high-power spontaneous UV–VUV sources and pulsed chlorine-based plasmochemical reactors operating at elevated pressures of the working medium.

In this paper, we present the results of studying the optical characteristics of a plasma produced in a pulsed TVD in Cl_2 and He/Cl_2 mixtures in the 140- to 300-nm wavelength range.

Gas molecules were excited by a TVD with automatic spark preionization in a volume of 18 × 2.2 × 1.0 cm, where 2.2 cm is the distance between the discharge electrodes. The TVD emitter circuit is described in [7]. A 30-nF main storage capacitor and a 9.4-nF sharpening capacitor were used in the two-contour *LC* circuit to ignite a TVD. A TGI-1-1000/25 thyratron was used as a switch. The sharpening ceramic capacitors KVI-3 (470 pF, 20 kV) were filled with an insulating compound and were mounted as two blocks inside the discharge chamber near the TVD electrodes. The minimum inductance of the main discharge-ignition circuit was $L \le 10$ nH). The presence of insulators between the TVD plasma and grounded screens makes this excitation system similar to circuits used to generate fast ionization waves in longitudinal pulsed discharges [8–10]. The FWHM duration of the TVD current was ≤30 ns. Note that the same value of the parameter *E*/*N* at the front of a fast ionization wave can be achieved at lower voltages of the power source; in this case, the plasma volume can also be increased substantially.

Emission from the TVD plasma was analyzed with a half-meter vacuum monochromator based on the Seia–Namioki scheme and equipped with a 1200-line/mm grating. The TVD chamber was hermetically connected to the vacuum monochromator through a CaF_2 window. An FEU-142 photomultiplier with a LiF window was used as a photodetector. The compartment where the FEU-142 was situated and the grating chamber were evacuated to a residual pressure of $\le 10^{-3}$ Pa. The "vacuum monochromator + FEU-142" system was calibrated relative to the H_2 continuum in the 165- to 350-nm region.

Figures 1 and 2 show the emission spectra of the plasma of a TVD in pure chlorine and He/Cl_2 mixtures. The spectra are drawn without taking into account the relative spectral sensitivity of the vacuum spectrometer (k_λ). The identified spectral bands and the relative intensities of the main band maximums (taking into account k_λ) are listed in the table. The chlorine bands were recognized according to the data presented in [11–14]. The dominant bands in the spectral region being studied were the $Cl_2(D'–A')$ band with the long-

1063-7842/00/4510-1346$20.00 © 2000 MAIK "Nauka/Interperiodica"

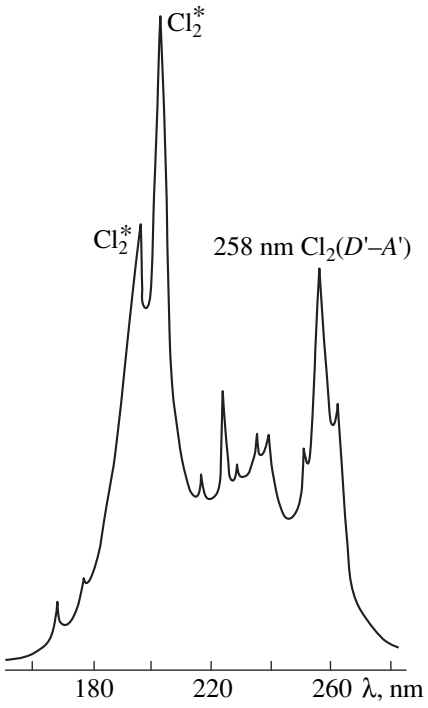


Fig. 1. Plasma emission spectrum of a TVD in chlorine at a pressure of 1.2 kPa and U_{ch} = 10 kV.

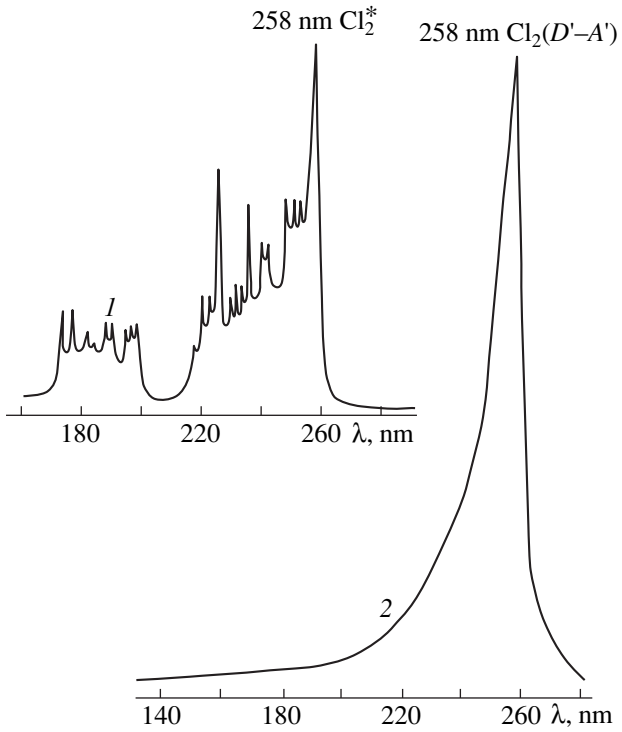


Fig. 2. Plasma emission spectra of TVDs in He/Cl_2 = (*1*) 2/0.4 and (*2*) 50/0.04-kPa mixtures.

wavelength edge at λ = 258 nm and the bands with edges at 236 and 195 nm. Note that the upper emitting state for the latter two bands are unknown, whereas the lower states are rather high-lying $A^3\Pi(O_0^+)$ [13] and $B(^3\Pi_0^+)$ [14] states. The upper states for the band with an edge at λ = 195 nm may be low vibrational levels of the $Cl_2(^1\Sigma_u^{+*})$ states that are formed from the $Cl^+(^1D)$ and $Cl^-(^1S)$ states, and the lower state is $A'^3\Pi(2u)$ [15]. In TVDs ignited in He/Cl_2 mixtures (especially in those with a substantial helium content), the singlet–triplet relaxation of Cl_2^{**} into the $D^3\Pi_{2q}$ state occurs, as a result of which only the $Cl_2(D'–A')$ 258-nm band remains in the plasma emission (Fig. 2).

Figure 3 shows the total intensities of bands with edges at λ = 195 and 258 nm as functions of Cl_2 pressure for the high-current phase of a TVD in pure chlorine. It is seen from the figure that, at $[Cl_2]$ = 0.4–0.2 kPa, the 195-nm band is dominant. As the chlorine pressure increases further, the $Cl_2(D'–A')$ 258-nm emission becomes dominant. A comparison between the decay rates of the intensity of the 195-nm band for different Cl_2 and He pressures in the He/Cl_2 mixture showed that the rate of the $Cl_2^{**}–Cl_2(D')$ singlet–triplet transition (i.e., the quenching rate for this band) under the action of Cl_2 molecules is more than one order of magnitude higher than that under the action of He atoms.

Hence, we have shown that, for chlorine at pressures of (0.4–3.0) kPa and for He/Cl_2 mixtures at pressures of ≤30 kPa, a steady-state phase can occur in high-current TVDs, which is of interest for plasmochemical applications. For a low Cl_2 content (P_{opt} = 0.8–1.6 kPa), this

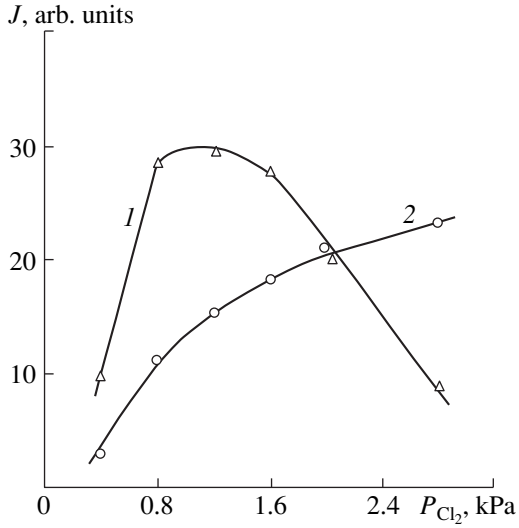


Fig. 3. Intensities of the bands with edges at λ = (*1*) 195 and (*2*) 258 nm in a high-current TVD in chlorine as a function of pressure at U_{ch} = 10 kV.

Intensity distribution in the emission spectrum of a TVD in chlorine for $[Cl_2]$ = 2.8 kPa and U = 10 kV

λ, nm	I/k_λ, au	Identification
168	0.73	$Cl_2[H–X^1\Sigma_g^+]$
177	0.70	$Cl_2[^1\Sigma–^1\Pi_4]$ [14]
184	0.79	180- to 195-nm Cl_2^* band
189	0.87	The same
193	1.06	"
195	1.87	"
200	0.95	Maximum in the chlorine continuum at λ = 200 nm [11, 12]
212	0.63	
223	0.88	223- to 236-nm $Cl_2(I–B^3\Pi_0^+)$ band [14]
228	0.70	The same
231	0.78	"
235	0.77	"
239	0.97	239- to 260-nm $Cl_2(D'–A')$ band [15]
247	1.53	The same
252	1.44	"
258	2.65	"
276	0.82	Maximum in the chlorine continuum at λ = 276.1 nm [11, 12]
291	0.70	
295	1.03	Maximum in the chlorine continuum at λ = 295.7 nm [11, 12]

discharge is an intense source of radiation in the band with an edge at λ = 195 nm. When helium ([He] ≥ 10–30 kPa) is mixed with chlorine, UV–VUV emission from the TVD plasma is concentrated in the $Cl_2(D'–A')$ 258-nm band.

REFERENCES

1. A. P. Kupriyanovskaya, V. I. Svettsov, and I. A. Lukashina, Izv. Vyssh. Uchebn. Zaved., Khim. Khim. Tekhnol. **30** (11), 52.
2. A. N. Panchenko, É. A. Sosnin, and V. F. Tarasenko, Zh. Tekh. Fiz. **67** (1), 78 (1997) [Tech. Phys. **42**, 68 (1997)].
3. A. P. Golovitskiĭ, Pis'ma Zh. Tekh. Fiz. **24** (6), 63 (1998) [Tech. Phys. Lett. **24**, 233 (1998)].
4. M. Diegelmann, K. Honla, and K. L. Kompa, Opt. Commun. **29**, 334 (1979).
5. A. K. Shuaibov, L. L. Shimon, A. I. Dashchenko, *et al.*, Pis'ma Zh. Tekh. Fiz. **25** (11), 29 (1999) [Tech. Phys. Lett. **25**, 433 (1999)].
6. E. B. Gordon, V. I. Matyushenko, R. B. Repin, and V. D. Sizov, Kvantovaya Élektron. (Moscow) **17**, 554 (1990).
7. A. K. Shuaibov, Kvantovaya Élektron. (Moscow) **26**, 127 (1999).
8. É. I. Asinovskiĭ, L. M. Vasilyak, and V. V. Markovets, Teplofiz. Vys. Temp. **21**, 371 (1983).
9. A. G. Abramov, É. I. Asinovskiĭ, and L. M. Vasilyak, Fiz. Plazmy **14**, 979 (1988) [Sov. J. Plasma Phys. **14**, 575 (1988)].
10. D. N. Gerasimov and O. A. Sinkevich, Fiz. Plazmy **25**, 376 (1999) [Plasma Phys. Rep. **25**, 339 (1999)].
11. R. W. B. Pears and A. G. Gaydon, *The Identification of Molecular Spectra* (Chapman Hall, London, 1963), p. 142.
12. *Spectroscopic Data*, Vol. 2: *Homonuclear Diatomic Molecules*, Ed. by S. N. Suchard and J. E. Melzer (Plenum, New York, 1976), p. 139.
13. K.-P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure* (Van Nostrand, New York, 1979; Mir, Moscow, 1984), Part 1.
14. B. M. Smirnov and A. S. Yatsenko, *Plasma Chemistry*, Ed. by B. M. Smirnov (Énergoatomizdat, Moscow, 1989).
15. V. S. Zuev, A. V. Kanaev, and L. D. Mikheev, Kvantovaya Élektron. (Moscow) **11**, 354 (1984).

Translated by N. Larionova

Technical Physics, Vol. 45, No. 10, 2000, pp. 1349–1350. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 120–121.

BRIEF COMMUNICATIONS

On the Contactless Effect of an Electric Field on Liquid Systems

N. N. Krasikov

Kovrov State Academy of Technology, Kovrov, 601910 Russia

Received December 1, 1999

Abstract—An effect of an electric field on the evaporation of polar and nonpolar liquids (water and a hydrocarbon, respectively) is established. The effect is stronger in the former case. © *2000 MAIK "Nauka/Interperiodica".*

Generally, an electric field is applied to a liquid by means of electrodes connected to a voltage source and submerged in the liquid, giving rise to an ion current and near-electrode phenomena. Processes of this kind are analyzed in electrochemistry. They have been utilized in a number of technologies.

The effect of an electric field on molecular ordering in liquids, including surface layers, was systematically studied at the Ioffe Physicotechnical Institute in the prewar years [1].

We have proposed to apply an electric field to a liquid by setting the electrodes in a contactless (remote) position. The setup used to implement the proposed method is schematized in Fig. 1. Here, a high-voltage source with output U = 15–20 kV is used and the current is blocked, which makes the scheme safe in operation. The liquid is poured into a dielectric cell placed between the electrodes. In this system, electric contact of the upper and lower electrodes with the liquid is prevented by a vapor–gas layer and the dielectric (polytetrafluoroethylene) bottom of the vessel, respectively. As a consequence, only a negligible leakage current is present in the circuit and Joule losses are absent. This setup can be used to generate fields of intensity E directed both downwards (with the "+" electrode at the top) and upwards (with the "–" electrode at the top).

In previous studies [2, 3], it was demonstrated that the growth of plantlike organisms at the early stages of their development is enhanced in water exposed to an electric field for 2–10 min. The exposed water produces a positive effect when used in treating respiratory diseases and allergies. The biological activity of the water depends on the direction of E: it is always higher when the positive electrode is placed above the liquid, so that the generated field is directionally similar to the atmospheric electric field [4].

Since the effects of electric and other physical fields on water inferred from the reactions of living organisms are of a subjective nature, additional evidence is required. To explore the field effect on a phase transition (evaporation), we took two liquids that have similar thermodynamic characteristics but widely different polarities: distilled water as a polar liquid with $\varepsilon = 80$, conductivity $\gamma = 2 \times 10^{-4}\ \Omega\ m^{-1}$, and pH = 6.4 and chemically pure dodecane $C_{12}H_{26}$ (a hydrocarbon of the paraffin series) as a nonpolar liquid with $\varepsilon = 1.95$ and conductivity lower by eight orders of magnitude. Our experiments were conducted under standard conditions ($t° = 20 \pm 0.5$°C, p = 752–761 torr). A quasi-uniform electric field was generated by plane electrodes connected with a source of U = 15 kV. The thicknesses of the air layer, water layer, and dielectric bottom were 8 mm, 1.5 mm, and 1.7 mm, respectively. Reference samples were kept under identical conditions, but with the voltage source switched off. The dielectric cell material was not wetted by water, and preferential evaporation from the meniscus was therefore ruled out. The electric field was applied continually and was switched off only at the moment when the cell was weighed on a VLR-200 analytical balance to within an error of ±0.2 mg. To preclude the formation of trapped-

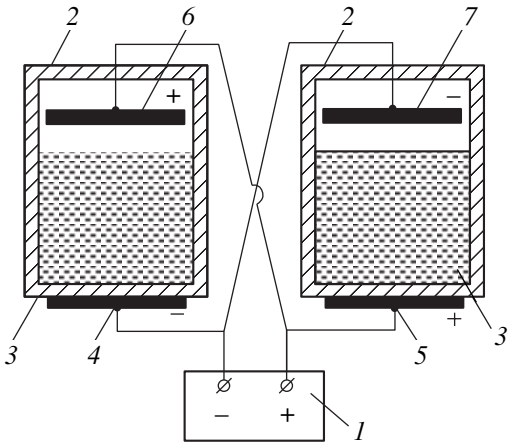


Fig. 1. Schematic of a setup for contactless treatment of liquids: (*1*) high-voltage source; (*2*) dielectric vessel; (*3*) treated liquid; (*4–7*) electrodes.

1063-7842/00/4510-1349$20.00 © 2000 MAIK "Nauka/Interperiodica"

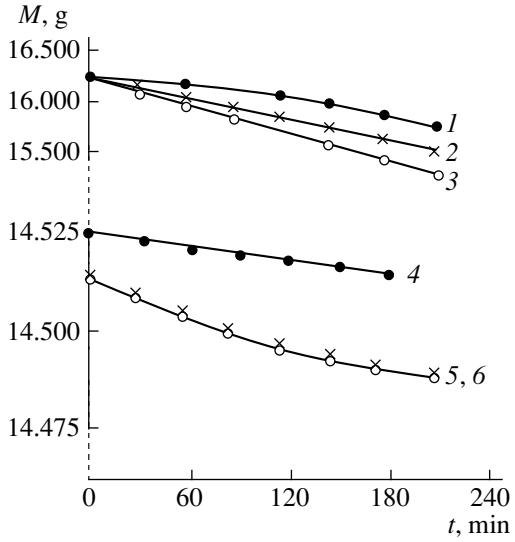


Fig. 2. Dependence of change in the liquid mass M [g] on the duration of electrical treatment t [min] for distilled water (*1–3*) and chemically pure $C_{12}H_{26}$ (*4–6*): (*1*, *4*) reference samples; (*2*, *5*) with E directed upwards; (*3*, *6*) with E directed downwards. The ordinate scales are different for different liquids.

gas zones over the liquids, the gas was blown by using the same blowing scheme for all samples.

The experimental results shown in Fig. 2 demonstrate the substantially higher rate of evaporation of water compared to the hydrocarbon. Note the stronger effect characteristic of water placed in the electric field directed downwards (like the atmospheric field). The mass of a nonpolar liquid exposed to an electric field decreases; but the decrease is not only substantially weaker, but is also independent of the E direction. Moreover, note that the rate of water charging by means of a needlelike upper electrode is higher when the needle is at positive potential. This observation confirms the ideas of Frumkin and Rusanov about the preferential orientation of H_2O molecules, with oxygen in the outward position in the surface layer [5].

An explanation of the phenomenon demonstrated here should be sought in the specific properties of liquid surface layers, which differ from the bulk properties of liquids [6]. The double electric layer at the surface changes when placed in an electric field of the type considered here: conversion of molecules into ions in it is intensified. Ions are forced out of the liquid by the electric field, and this effect is more pronounced for the OH^- groups associated with neutral water molecules when the field is directed downwards [5].

Molecules of a nonpolar liquid have zero dipole moments and are hardly ionized. They are not involved in electrostatic (ion or dipole) interactions that could hold them in the liquid. The orientational ordering of elongated molecules in electric fields, which was discussed in a classical monograph [1], can be specifically manifested in surface layers by an increase or decrease in the loss of molecules by liquids. In the case under study, an increase was observed and found to be independent of the E direction.

The effect of an electric field on a phase transition (evaporation of liquids) appears to be weak, but it must be involved in natural phenomena, including biology, and can be utilized in technologies with low heat losses.

REFERENCES

1. I. Ya. Frenkel, *Kinetic Theory of Liquids* (Nauka, Leningrad, 1975; Clarendon Press, Oxford, 1946).
2. N. N. Krasikov, V. K. Koekin, and I. V. Slyusar', Biofizika **39**, 923 (1994).
3. N. N. Krasikov, Élektrotekhnika, No. 4, 57 (1996).
4. R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics* (Addison-Wesley, Reading, Mass., 1963; Mir, Moscow, 1966), Vol. 5.
5. N. N. Krasikov and T. M. Marenkova, Kolloidn. Zh. **59**, 855 (1997).
6. S. S. Dukhin and B. V. Deryagin, *Electrophoresis* (Nauka, Moscow, 1976).

Translated by A. Betev

Technical Physics, Vol. 45, No. 10, 2000, pp. 1351–1353. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 122–124.

BRIEF COMMUNICATIONS

The Effect of Ion Emission on the Characteristics of a Glow Discharge with a Hollow Cathode

S. P. Nikulin

Institute of Electrophysics, Ural Division, Russian Academy of Sciences, Komsomol'skaya ul. 34, Yekaterinburg, 620049 Russia

E-mail: nikulin@ief.uran.ru

Received December 20, 1999

Abstract—The effect of extraction of the ion current on the conditions for sustaining a low-pressure glow discharge with a hollow cathode, which is a topical problem for developing ion sources, is studied. It is shown that there exists a certain optimal level of the ion emission at which the energy efficiency reaches its maximum.

Interest in studying the characteristics of low-pressure glow discharges is mainly explained by the possibility of applying them in the development of sources of charged particle beams. Ion sources based on gas-discharge systems with cold cathodes, in particular, those based on a glow discharge with a hollow cathode (GDHC) have a considerably longer service life and greater reliability than hot cathode systems under the technical vacuum conditions typical of industrial systems. In addition, in a pulsed or pulsed-periodic operating regime, cold cathode systems are more energy-efficient, as there is no need to sustain the heater current [1]. At the same time, in a continuous operating regime, the energy efficiency of the glow discharge ion sources is appreciably lower compared with arc discharge sources with a hot cathode. This is explained by the fact that the coefficient of the potential ion–electron emission is small, and, accordingly, the burning voltage of the discharge is high. However, in contrast to arc discharges where the ion component of the discharge current is rather small, in a glow discharge, the current to the cathode is almost entirely carried by ions. Therefore, the source energy efficiency can be enhanced if the extracted fraction of the ion current is increased by increasing the total area of the emission orifices in the cathode; however, such an increase in the ion emission may interfere with sustaining the discharge, inasmuch as the particles collected into the beam miss the cathode and do not participate in γ-processes at its surface. In this connection, it is important to study the effect of ion extraction on the GDHC characteristics and determine an optimum level of ion emission at which one can obtain a sufficiently high emission current without significant difficulties in sustaining the discharge.

It is generally accepted [2–4] that in GDHCs the main contribution to the ionization process is made by the so-called fast electrons originating at the cathode and acquiring high energy as they pass through the cathode potential drop region across which almost the entire discharge voltage U drops. As a result of an abrupt decrease in the Coulomb interaction cross section at high energies, the fast electrons mainly interact with neutral atoms [4, 5], and the energy distribution function $f(\varepsilon)$ for fast electrons can be written as

$$\gamma(1-\alpha)\frac{I_i}{e}\delta(\varepsilon-\varepsilon_0) - N\sigma(\varepsilon)\sqrt{\frac{2\varepsilon}{m}}f(\varepsilon) + \int N\sigma(\varepsilon+\omega,\omega)\sqrt{\frac{2(\varepsilon+\omega)}{m}}f(\varepsilon+\omega)d\omega - \frac{f(\varepsilon)S_a}{4V}\sqrt{\frac{2\varepsilon}{m}} = 0, \quad (1)$$

where I_i is the total ion current in the discharge; α is the ion current fraction collected into the beam; $(I-\alpha)I_i$ is the ion current at the cathode; γ is the ion–electron emission coefficient; e is the electron charge; $\varepsilon_0 = eU$ is the energy acquired by an electron in the cathode potential drop region; δ is the Dirac δ-function; N is the concentration of neutral atoms in the cathode cavity; m is the electron mass; $\sigma(\varepsilon)$ is the total cross section of electron–atom collisions; $\sigma(\varepsilon, \omega)$ is the cross section corresponding to the loss of energy ω in a collision by a colliding electron with energy ε; S_a is the anode area; and V is the cathode cavity volume.

The first term in (1) describes the generation of new fast particles as a result of ion–electron emission from the cathode. The second and the third terms characterize the disappearance and emergence of the electrons, respectively, in a narrow energy range around ε as a result of collisions (integration in the third term is carried out over all possible ω). The last term allows for withdrawal of fast electrons to the anode from the discharge, according to an expression for the number of particle–wall collisions known from molecular-kinetic theory. In deriving it, we assumed that the fast particles

1063-7842/00/4510-1351$20.00

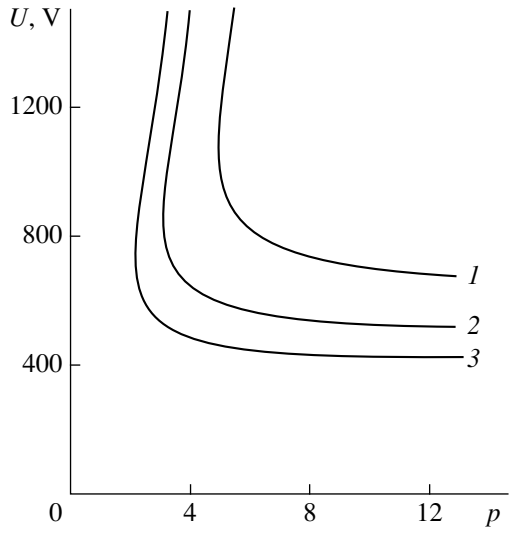


Fig. 1. The burning voltage of a discharge as a function of pressure for $\tilde{\gamma}$ = (*1*) 0.06; (*2*) 0.08; (*3*) 0.1.

are uniformly distributed throughout the cathode cavity volume.

Taking into account that the energy lost in a single collision is considerably less than the energy acquired in passing the cathode potential drop region, it is possible to simplify (1) by using the Fokker–Planck approximation. Expanding the integrand into a series and permuting the differentiation and integration, after some transformations, we obtain

$$\gamma(1-\alpha)\frac{I_i}{e}\delta(\varepsilon-\varepsilon_0)+\frac{\partial}{\partial\varepsilon}\left(NQ(\varepsilon)\sqrt{\frac{2\varepsilon}{m}}f(\varepsilon)\right) - \frac{f(\varepsilon)S_a}{4V}\sqrt{\frac{2\varepsilon}{m}} = 0, \quad (2)$$

where $Q(\varepsilon)$ is the so-called stopping power defined by the following expression:

$$Q(\varepsilon) = \int\sigma(\varepsilon,\omega)\omega d\omega. \quad (3)$$

Equation (2) can be solved analytically to obtain the following expression for the energy distribution function of the fast electrons:

$$f(\varepsilon) = \frac{\gamma(1-\alpha)I_i}{eNQ(\varepsilon)}\sqrt{\frac{m}{2\varepsilon}}\exp\left(-\frac{S_a}{4NV}\int_{\varepsilon}^{\varepsilon_0}\frac{d\varepsilon'}{Q(\varepsilon')}\right). \quad (4)$$

At the same time, the ion current is related to the distribution function of the fast electrons by the following expression:

$$I_i = e\int_{\varepsilon_i}^{\varepsilon_0} N\sqrt{\frac{2\varepsilon}{m}}\sigma_i(\varepsilon)f(\varepsilon)d\varepsilon, \quad (5)$$

where ε_i and σ_i are the threshold energy and the ionization cross section, respectively.

Combining (4) and (5), a condition for self-maintained discharge in the ion emission regime can be obtained:

$$\gamma(1-\alpha)\int_{\varepsilon_i}^{\varepsilon_0}\frac{\sigma_i(\varepsilon)}{Q(\varepsilon)}\exp\left(-\frac{S_a}{4NV}\int_{\varepsilon}^{\varepsilon_0}\frac{d\varepsilon'}{Q(\varepsilon')}\right)d\varepsilon = 1. \quad (6)$$

Expression (6) gives an implicit dependence of the burning voltage of a discharge on the neutral gas concentration or on the pressure $P = NkT$, where k is Boltzmann's constant and T is the gas temperature.

Figure 1 presents calculated plots of the burning voltage as a function of the dimensionless pressure

$$p = \frac{PV\sigma_0}{kTS_a} = \frac{NV\sigma_0}{S_a}, \quad (7)$$

where $\sigma_0 = \pi a_0^2$ is the area of the Bohr orbit, and a_0 is the Bohr radius. The calculations were carried out for various values of the parameter $\tilde{\gamma} = \gamma(1-\alpha)$.

In the calculation, argon was adopted as a working gas and the data on the ionization cross sections were taken from [6]. As for the stopping power, which, at energies exceeding the ionization threshold, is mainly governed by inelastic collisions, it was described by the expression derived for the classical Thomson model

$$Q(\varepsilon) = \frac{Z\pi e^4}{\varepsilon}\ln\left(\frac{\varepsilon}{\varepsilon_1}\right), \quad (8)$$

where Z is the number of electrons in the outermost shell and ε_1 is the energy of the first excitation level. Expression (8) agrees with the results of quantum-mechanical calculations for fast electrons [7].

As seen from Fig. 1, sustaining the discharge is impossible at any voltage if the pressure is lower than a certain critical value, and at pressures above critical the functions obtained are ambiguous. Similar results were obtained earlier for the dependence of burning voltage on the magnetic induction in a discharge with electrons oscillating in a magnetic field [5]. The states corresponding to the lower branches of the curves in Fig. 1 are stable and observable in an experiment.

As the emission increases and, correspondingly, the parameter $\tilde{\gamma}$ decreases, the lower boundary of the operating pressure range gradually rises. This imposes certain constraints on the possible increase in the ion extraction efficiency in ion sources, because, at a higher pressure, the electrical strength of the accelerating gap may be exceeded. However, even if this has not occurred, the rise in the voltage with decreasing $\tilde{\gamma}$ may make the increase in the extracted current fraction worthless because of a decrease in the source energy

efficiency:

$$\alpha_e = \frac{I_b}{W} = \frac{\alpha}{U}, \quad (9)$$

where I_b is the ion beam current and $W \approx I_i U$ is the power consumed in sustaining the discharge.

Figure 2 shows $\alpha_e(\alpha)$ plots for a number of pressure values. It is seen that all the curves have the same qualitative character: initial growth gradually slows down, the curves reach their maxima at some point, and then the energy efficiency decreases. Thus, there exists a certain optimal level of the extracted ion current fraction depending on the pressure at which the energy efficiency reaches its maximum.

At higher pressures, as is also seen in Fig. 2, it is possible to achieve a higher energy efficiency. Therefore, in order to increase α_e, one should increase the gas flow through the source, either as much as the electrical strength of the gap will allow or until the gas flow (usually expressed in cm^3at/h) exceeds a certain limit set by the characteristics of the technological process in which an ion source is used. However, increasing the gas flow degrades another important characteristic of the source, namely, the gas efficiency

$$\alpha_g = \frac{I_b}{G}. \quad (10)$$

This may make the source operation in a regime with an increased gas flow rate economically inefficient in spite of the high energy efficiency.

A search for the most profitable operating regime for the source can be performed on the basis of the following model. Assuming that the basic expenses for maintaining the source operation are connected with both the electric power consumption in its gas-discharge system and gas consumption, let us introduce a quantity C, which characterizes the cost of one ampere-hour of source operation:

$$C = \frac{C_e}{\alpha_e} + \frac{C_g}{\alpha_g}, \quad (11)$$

where C_e is the cost of one kilowatt-hour of electrical energy (the energy efficiency in expression (11) is assumed to be expressed in A/kW, as is usually done) and C_g is the cost of one cubic centimeter of the working gas at atmospheric pressure.

For given costs of the electrical energy and the working gas and assuming that the relation between the pressure in the gas-discharge chamber and the gas flow for various values of α (that is, for practical purposes, for various magnitudes of the total area of the emission orifices) is known, finding a minimum of expression (11) is not a problem and can easily be carried out for any technological installation.

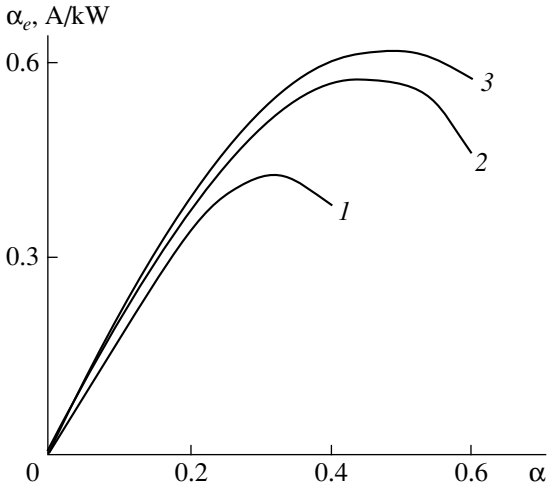


Fig. 2. The energy efficiency of an ion source as a function of the relative magnitude of the extracted ion current for p = (*1*) 5; (*2*) 10; (*3*) 15.

REFERENCES

1. M. A. Zav'yalov, Yu. E. Kreĭndel', A. A. Novikov, and L. P. Shanturin, *Plasma Processes in Electron Guns* (Énergoatomizdat, Moscow, 1989).
2. V. I. Kirichenko, V. M. Tkachenko, and V. B. Tyutyunnik, Zh. Tekh. Fiz. **46**, 1857 (1976) [Sov. Phys. Tech. Phys. **21**, 1080 (1976)].
3. A. S. Metel', Zh. Tekh. Fiz. **54**, 241 (1984) [Sov. Phys. Tech. Phys. **29**, 141 (1984)].
4. Yu. M. Kagan, R. I. Lyagushchenko, A. S. Taroyan, and S. N. Khvorostovskiĭ, Zh. Tekh. Fiz. **43** (7), 1488 (1973) [Sov. Phys. Tech. Phys. **18**, 943 (1973)].
5. S. P. Nikulin, Zh. Tekh. Fiz. **68** (7), 56 (1998) [Tech. Phys. **43**, 795 (1998)].
6. M. A. Lennon, K. L. Bell, H. V. Gilbody, *et al.*, J. Phys. Chem. Ref. Data **17**, 1285 (1988).
7. L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory* (Nauka, Moscow, 1989, 4th ed.; Pergamon, Oxford, 1977, 3rd ed.), Vol. 1.

Translated by N. Mende

Technical Physics, Vol. 45, No. 10, 2000, pp. 1354–1356. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 125–126.

BRIEF COMMUNICATIONS

A Quasi-One-Dimensional Ballistic Ring in a Constant Electric Field

G. M. Shmelev, É. M. Épshteĭn, and G. A. Syrodoev

Volgograd State Pedagogical University, Volgograd, 400013 Russia

E-mail: belgen@avtlg.ru

Received January 25, 2000

Abstract—For an electron placed in a quasi-one-dimensional ballistic ring and subjected to a constant electric field, the mean dipole moment and the intensity of dipole radiation were calculated. As the field intensity decreases, the initially paraelectric ring turns dielectric. © *2000 MAIK "Nauka/Interperiodica".*

The advances in single-electron device technology have stimulated investigation into quasi-one-dimensional rings. In the literature, the ring is most often considered in terms of quantum effects: size quantization, interference phenomena like the Aharonov–Bohm effect, Coulomb blockade, etc. In this work, we show that intriguing classical effects may occur in quasi-one-dimensional ballistic rings subjected to external electric fields. Specifically, we will deal with a planar quasi-one-dimensional ring whose width is much less than the radius R. The ring is essentially a quantum well between two concentric circular potential barriers. The well is only width-quantized, and angular motion is classical and ballistic (the mean free path of an electron greatly exceeds $2\pi R$). We are interested in the response of the ring to an external constant electric field F parallel to its plane (the field is produced by a parallel-plate capacitor).

The position of an electron in the ring is specified by the angular variable φ, which is reckoned from the "lowest" point, i.e., the point nearest to the positively charged plate of the capacitor. The equation of motion for an electron is similar to that of a pendulum:

$$\frac{d^2\varphi}{dt^2} + \omega_0^2 \sin\varphi = 0, \qquad (1)$$

where $\omega_0^2 = eF/mR$ (e and m are the electron charge and mass, respectively).

The exact solution of this equation (see, e.g., [1]) depends on F. If $F > F_0 \equiv E/2eR$, a pendulum (electron) with an energy E vibrates about the equilibrium point. At $F < F_0$, it executes angular motion with monotonically increasing φ. In the transition field $F = F_0$, the motion is aperiodic.

(1) The dipole moment of a ring that has one electron and a distributed neutralizing charge (jelly model) is $p = \{eR\cos\varphi, eR\sin\varphi\}$. Using the exact solution of (1) with the boundary conditions $\varphi(0) = 0$, $\dot{\varphi}(0) = \sqrt{2E/mR^2} \equiv \Omega$ and averaging $p(t)$ over time, we find the steady-state component of the dipole moment $\bar{p} = \{p_0, 0\}$:

$$p_0 = eR\left[-1 + 2\frac{\mathbf{E}(\alpha)}{\mathbf{K}(\alpha)}\right], \quad (F > F_0), \qquad (2)$$

$$p_0 = eR\left[1 - 2\alpha^2\left(1 - \mathbf{E}\left(\frac{1}{\alpha}\right)/\mathbf{K}\left(\frac{1}{\alpha}\right)\right)\right], \quad (F < F_0). \qquad (3)$$

Here, $K(k)$ and $E(k)$ are the complete elliptic integrals of the first and second kind and $\alpha = \sqrt{F_0/F}$. At $F \gg F_0$, we have from (2) $p_0 = eR$ (the field presses the electron against the lowest point of the ring). In weak fields ($F \ll F_0$), $p_0 = -5eR/16\alpha^2$ from (3); in other words, the permittivity of the ring is negative. To better appreciate this result, we note that the dipole moment in the intermediate case $F \longrightarrow F_0 \pm 0$ is $p_0 = -eR$. In this limit, the period of motion becomes infinite; that is, the electron stays at the "highest" point throughout the entire time. Then, the previous result also becomes clear: having passed through the top point, the electron spends more time in the upper semicircle than in the lower one. This corresponds to the negative (in opposition to the field) dipole moment. Hence, the dipole moment vanishes between $\alpha = 0$ and $\alpha = 1$ (more strictly, $\alpha = \alpha_0 \approx 0.91$); that is, the paraelectric behavior changes to dielectric. The field dependence of p_0 is shown in Fig. 1. To avoid misunderstanding, it should be stressed that the occurrence of the negative component of the dipole moment is not contradictory to the well-known theorem on the positiveness of static permittivity [2]. This theorem applies to thermodynamically equilibrium systems and, accordingly, is proved in terms of thermodynamics.

1063-7842/00/4510-1354$20.00

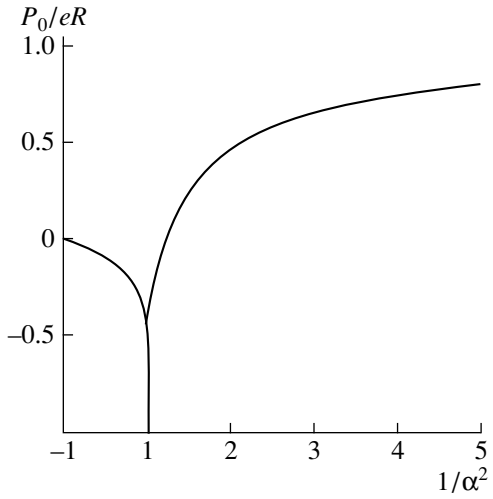


Fig. 1. Mean dipole moment of the ring vs. $\alpha^{-2} = 2eRF/E$.

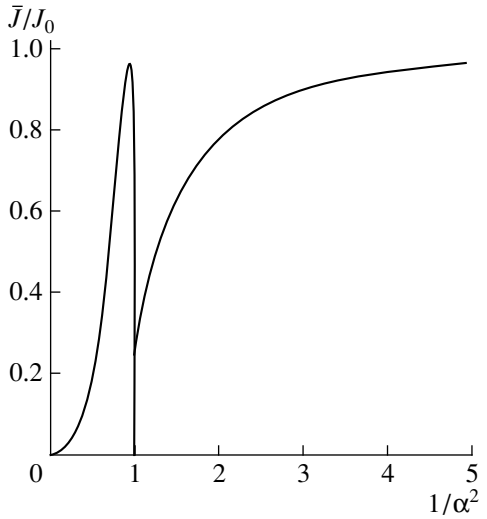


Fig. 2. Mean radiation intensity vs. α^{-2}.

Our system is not in equilibrium, in particular, because of ballistic effects taking place between electron collisions with scatterers; i.e., before equilibrium is established.

(2) Let us consider dipole radiation of electromagnetic waves by the ring. Averaging the instantaneous radiation intensity $J(t) = 2\dot{p}^2/3c^2$ [3] over time yields

$$\bar{J} = \frac{J_0}{4\alpha^2}\left[3 - 7\alpha^2 + 4\alpha^2 - 3(1-2\alpha^2)\frac{\mathbf{E}(\alpha)}{\mathbf{K}(\alpha)}\right], \quad (F > F_0), \tag{4}$$

$$\bar{J} = J_0\left[-\frac{1}{2} + \frac{1}{2\alpha^2} + 3\left(\frac{1}{2} - \frac{1}{4\alpha^2}\right)\mathbf{E}\left(\frac{1}{\alpha}\right)/\mathbf{K}\left(\frac{1}{\alpha}\right)\right], \quad (F < F_0), \tag{5}$$

where $J_0 = 8e^2E^2/(3m^2c^3R^2)$.

As follows from (4) and (5), at $F \longrightarrow F_0 \pm 0$, $\bar{J} \longrightarrow 0$ because of an infinite increase in the motion period; in other words, we are dealing with a sort of soft mode that corresponds to "resonance," $\omega_0 = \Omega/2$. With this condition fulfilled, it makes sense to consider the energy radiated over an infinite period of time, $U = 3J_0/2\omega_0$. In this case, we have a continuous radiation spectrum. Figure 2 plots $\bar{J}$ against applied field F.

Fourier analysis of the function $J(t)$ implies that, at $F > F_0$, the fundamental frequency is $\omega = \pi\omega_0/2\mathbf{K}(\alpha)$, while at $F < F_0$, it is given by

$$\omega_1 = \pi\Omega/2\mathbf{K}\left(\frac{1}{\alpha}\right)$$

(in the latter case, the formulas for intensities of even and odd harmonics of frequency ω differ somewhat). As the field grows from zero to F_0, the spectrum crowds (ω decreases from Ω to zero). At $F = F_0$, the continuous spectrum is described by a bell-shaped spectral density function $S(\omega)$ with a peak of $S_{\max} \approx S(2\omega_0) \approx 4\pi e^3F_0R/3mc^3$. With a further increase in F ($F > F_0$), the spectrum becomes sparse (the fundamental frequency ω grows from zero to ω_0) and the intensity of higher harmonics drops. At $F \longrightarrow \infty$, the radiation becomes monochromatic with a frequency ω_0 varying as $\sqrt{F}$. Such an evolution of the spectrum may be viewed as the delocalization-to-localization transition. In fact, at $F < F_0$, the motion is infinite (the angular coordinate φ monotonically increases with time), while at $F > F_0$, φ varies in certain limits.

(3) Above, the ring was assumed to have a single electron. If there appear N electrons due to point injection, the factor N^2 should be introduced into formulas (4) and (5) and N, into (2) and (3). For a degenerate electron gas in the ring, E has the meaning of Fermi energy. Coulomb interaction between the electrons can be neglected if $\omega_0 \gg \omega_p$, where ω_p is the frequency of plasma oscillations in the ring (the calculation of ω_p is the subject of a separate paper).

The following contactless procedure can be used instead of point injection. First, with the same parallel-plate capacitor, a strong ($F \gg F_0$) electric field is applied in order that all free electrons pass to the lowest point. The initial condition $\varphi(0) = 0$ is provided in this way, which prevents a random scatter in the phases of oscillations between the motions of different electrons.

Then, the field is brought to a desired value. We make several estimates in conclusion. At $R = 10^{-5}$ cm, $m = 0.1m_e$, $E = 0.1$ eV, we obtain $\Omega = 6 \times 10^{12}$ s^{-1}. The bifurcation point $\omega_0 = \Omega/2$ corresponds to the field $F = 5 \times 10^3$ V/cm.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 97-02-16321).

REFERENCES

1. G. M. Zaslavskiĭ and R. Z. Sagdeev, *Introduction to Nonlinear Physics* (Nauka, Moscow, 1988).
2. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1984).
3. L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (Nauka, Moscow, 1973; Pergamon, Oxford, 1975).

Translated by V. Isaakyan

Technical Physics, Vol. 45, No. 10, 2000, pp. 1357–1361. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 127–131.

BRIEF COMMUNICATIONS

Spatial–Frequency Characteristics of Radio Emission Induced by Interaction of an Extensive Air Shower with the Earth's Surface

A. D. Filonenko

East Ukrainian State University, Lugansk, 348034 Ukraine

E-mail: uni.@vugu.lugansk.ua

Received February 16, 2000

Abstract—The spatial directional pattern of the radio emission induced by interaction of the excess electrons of an extensive air shower with the Earth's surface was found in the approximation of perfect soil conductance. The results made it possible to determine the frequency range (≈0.2–1 MHz), within which the radio emission of the shower is maximum, provided that this emission mechanism is valid. In addition, it became possible to explain the previously observed lack of direct proportionality between the shower energy and the radio signal amplitude. It was found that this disproportionality is caused mainly by the high directivity of low-frequency (0.2–3 MHz) radiation patterns within a wide range of the values of the axial tilt. © *2000 MAIK "Nauka/Interperiodica".*

The most relevant problems of cosmic-ray physics concern particles with energies higher than 10^{19} eV [1]. For example, despite theoretical predictions of the energy spectrum cutoff by the microwave cosmic background at energies $\approx 5 \times 10^{19}$ eV, more than ten events with energies higher than 10^{20} eV have already been detected. The results reported by the Akeno group are also very interesting [2]. A total of 36 events with energies higher than 4×10^{19} eV were observed over a period of 6 years using the Akeno array. The analysis of these events revealed the possibility of the existence of point sources of ultrahigh-energy cosmic rays (CR). The current interest in ultrahigh-energy CR was inspired by the construction of giant detectors, such as ShAL-1000 (Extensive Air Shower, EAS-1000) or Pierre Auger. However, even these giant detectors would not be able to detect particles with energies of 10^{22}–10^{23} eV. In my opinion, there is hope of detecting such superparticles (if they actually exist) in the foreseeable future using relatively inexpensive detectors. Many already existing detectors are able to solve this problem. For this purpose, the method of radio detection of cosmic rays suggested in [3, 4] should be used. It was shown earlier [5–7] that this method allows the effective area of the detector to be increased by approximately 6 orders of magnitude (i.e., to 10^7 km^2). The mechanism of electron–photon avalanche radio emission used in this method was studied in [8–10]. The electromagnetic field intensity within the long-wave range (shower propagation in the Earth's atmosphere) and the decameter range (shower propagation at the surface of the Moon) was estimated [3, 4].

However, this is not the only radio emission mechanism that does not require a magnetic field. In recent years, the mechanism of transient radiation induced by the interaction of excess electrons of an extensive air shower with the Earth's surface have been investigated by many groups of researchers worldwide (see, for example, [11–18]). In my opinion [19], the method of observation used in [11–18] should be radically modified. However, this method may prove very effective for detecting ultrahigh-energy particles. The goal of this work was to determine the directional pattern of this type of transient radiation. This should be done, because the potential of this mechanism of EAS emission is difficult to study without knowing its spatial characteristics.

Let us first determine the intensity of radio emission of a single excess electron of EAS. It should be noted that avalanche positron annihilation is the main cause of the disturbance of the shower neutrality that results in radio emission. The moment of annihilation should be considered as the outbreak of excess electron motion. Obviously, the excess electron emits not only while crossing the Earth's surface, but also at the initial stage of motion. Strictly speaking, it is the stopping positron that is emitted at this stage. However, this process is equivalent to excess-electron acceleration. Therefore, it may be assumed that, at the early stages of the motion (i.e., when the positron disappears), the excess electron attains the velocity $\beta = v/c \approx 1$ almost immediately within a time interval τ. Then, it moves uniformly for a time t_0. When the electron reaches the Earth's surface, its velocity decays to zero within a time interval τ (i.e., as quickly as it was accelerated). The results of calculations based on this model of transient radiation for $\varepsilon \longrightarrow \infty$ agree well with the experimental data (see, for example, [20]). The time τ is selected so

1063-7842/00/4510-1357$20.00

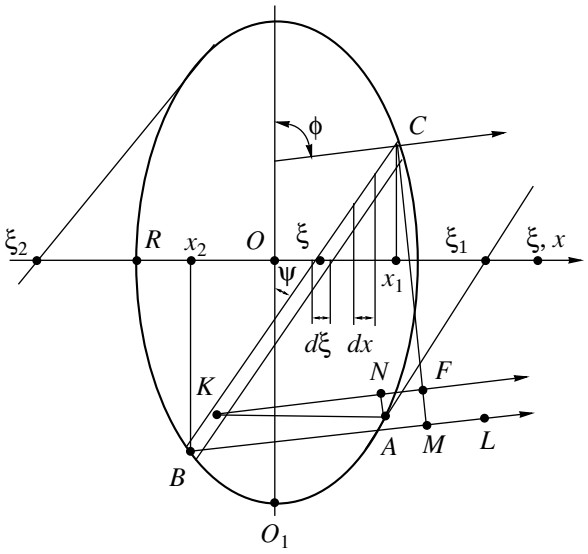


Fig. 1. The ellipse obtained by projection of the shower disk onto the Earth's surface; *CM* is the wave surface.

that $\omega\tau \ll 1$ for all radio frequencies. The spectral intensity of the field produced by the electron in the wave zone can be determined from the following equation (see, e.g., [21]):

$$dE = \frac{1}{2\pi}\frac{e}{4\pi\varepsilon_0 R_0 c^2}\int_{-\infty}^{\infty}\frac{a(t')\sin\Theta e^{i\omega t}dt}{(1-\vartheta(t')c^{-1}\cos\Theta)^3}, \qquad (1)$$

where $dt = (1-\vartheta(t')c^{-1}\cos\Theta)dt'$ and $t = t' + R_0/c - (\mathbf{n}, \mathbf{r})$.

Integration (1) can be conveniently performed on three successive intervals $(0, \tau)$, $(\tau, t_0 + \tau)$, and $(t_0 + \tau, t_0 + 2\tau)$. On the second interval, $a(t') \approx 0$. For $\tau \longrightarrow 0$, we obtain

$$dE_1(\omega, z) = \frac{e}{4\pi\varepsilon_0}\frac{\sin\Theta\exp(-i\omega R_0/c)}{2\pi R_0(1-\beta\cos\Theta)c} \times\{1-\exp[-i\omega z(1-\beta\cos\Theta)/c]\}. \qquad (2)$$

The angle Θ in Eq. (1) is taken to be constant, because the distance to the observer R_0 greatly exceeds the excess-electron range z.

Because soil is a good conductor ($\varepsilon \longrightarrow 0$), the resulting pattern of radio emission is determined by both the excess electron motion and its electric image. Therefore, vector addition of the fields should be performed. We denote the angle between the vertical and the shower axis as Θ_0. Thus, $\cos\Theta = -\cos\varphi\sin\Theta_0$, where φ is the azimuth angle of an observer positioned on the Earth's surface. The azimuth angle is measured from the line of intersection of the plane containing the vertical and the velocity vector **v** with the Earth's surface. The absolute value of the resulting intensity *dE* of the electron field and its image is related to equation (2) by the expression

$$dE = 2dE_1(\omega, z)\cot\Theta_0(\cot^2\Theta_0 + \sin^2\Theta_0)^{-1/2}.$$

If the travel direction of the shower disk $\Theta_0 \neq 0$, the region of interaction between the shower electrons and the Earth's surface is limited by an ellipse. This ellipse is formed by the projection of the shower disk onto the Earth's surface at the angle Θ_0. The ellipse semiaxes are equal to r and $r/\cos\Theta_0$, where r is the shower disk radius (i.e., the radius outside of which the charged particle density is negligible).

The electromagnetic waves emitted by the excess electrons reaching the line *BC* (Fig. 1) reach the observer simultaneously if the line *BC* makes an angle Ψ with the major semiaxis of the ellipse, from which the azimuth angles φ are measured. It is easily seen that this condition is met if $BC\cos\Psi/\vartheta_0 = BC\cos(\varphi + \Psi)/c$, where $\vartheta_0 = c/\sin\Theta_0$ is the travel velocity of the line of interaction between the disk and the Earth's surface. On substitution, we obtain $\tan\Psi = (\sin\Theta_0 - \cos\varphi)/\sin\varphi$. Let the straight line $BCy = (x - \xi)\cot\Psi$ intersect the *x*-axis at a given point ξ. Translate the line *BC* to the right by $d\xi (d\xi \ll \lambda)$. Let us determine the charge *dq* of the excess electrons falling within the strip *BC*. The charged particle distribution $\rho(r_d)$ over the shower disk is assumed to be known (see, for example, [22]). An element of the strip *BC* falling within the interval *dx* has the area $dS_e = dl dx/\sin\Psi$, where $dl = d\xi\cos\Psi$ is the width of the strip. Therefore, the charge *dq* is equal to the sum of all elements $\rho_e(x, y, \Theta_0)dS_e$, where *x* and *y* are the coordinates of the strip element dS_e. Let the charge distribution over the shower disk be described by the following function:

$$\rho_d(r_d) = \rho_d(\sqrt{x_d^2 + y_d^2}),$$

where x_d and y_d are the coordinates of an element dS_d of the disk area. This function can be used to calculate the charge falling within the ellipse area element dS_e. The element dS_e is the projection of the disk element dS_d onto the Earth's surface. On projection, we obtain $x_d = x_e$ and $y_d \longrightarrow y_e = y_d/\cos\Theta_0$. Thus,

$$\rho_d(r_d) = \rho_d(\sqrt{x_d^2 + y_d^2}) = (\sqrt{x_e^2 + y_e^2\cos^2\Theta_0}).$$

The projection procedure meets the condition $dS_e = dS_d/\cos\Theta_0$ (this actually means that $\rho_e = \rho_d\cos\Theta_0$. Therefore, the charge *dq* falling within the strip *BC* is

$$dq = \cot\Psi\cos\Theta_0 d\xi\int_{x_1}^{x_2}\rho_d(\sqrt{x_e^2 + y_e^2\cos^2\Theta_0})dx, \qquad (3)$$

where $y_e = (x - \xi)\cot\Psi$; x_1 and x_2 are the abscissas of the points of intersection of the straight line *BC* with the

ellipse boundary $x_1^2 + y_1^2\cos^2\Theta_0 = r^2$,

$$x_{1,2} = \{\xi\cot^2\Psi\cos^2\Theta_0 \pm [\cot^2\Psi\cos^2\Theta_0(r^2 - \xi^2) + r^2]^{1/2}\} \times (1 + \cot^2\Psi\cos^2\Theta_0)^{-1}. \quad (4)$$

To find the total amplitude at the observation point, the phase in Eq. (2) should be expressed as a function of the position of the charge dq of an arbitrary strip BC. The phase value for each of these parallel strips, into which the ellipse is divided, depends on the distance R_0 from the strip to the observer. Obviously, the signal from the point A will reach the observer first. Point A corresponds to a strip coinciding with the tangent (H, ξ_1) to the ellipse. This tangent intersects the x-axis at the point ξ_1. The signal from point A_1 positioned on the opposite side of the ellipse (Fig. 1) will be the last to reach the observer. The coordinates ξ_1 and ξ_2 correspond to the zero value of the discriminant of Eq. (4):

$$\xi_{1,2} = \pm r(\tan^2\Psi + \cos^2\Theta_0)^{1/2}/\cos\Theta_0. \quad (5)$$

On approaching the Earth, the disk touches it at the point O_1. Then, as the ellipse is formed, the boundary of interaction with the Earth moves along the x-axis. When it reaches the position AK, the signals from the points K and A will be emitted simultaneously. Meanwhile, as the boundary reaches the level of point C, the signal emitted from point A will have traveled a distance AL, whereas the signal emitted from point K will reach the straight line CF (at this instant, the signals from the other points on strip BC will reach the straight line CF). Obviously, $KF = AL$, so that the time interval t between the signals emitted from points K and A is determined by the equation $t = KN/c$. Given that $K\xi\xi_1A$ is a parallelogram and $\angle NAK = \varphi$, we obtain $t = (\xi_1\xi)c^{-1}\sin\varphi$.

The sum of all amplitudes $dE(\omega, z)$ can be found from the values obtained above by taking the phases into account. It should be noted, however, that all excess electrons in the shower travel a distance z before reaching the Earth's surface. The fraction of excess electrons in the shower maximum is estimated at about $k = 0.1$ (see, for example, [3]). For the showers under consideration (i.e., showers with energies higher than 10^{19} eV), the number of particles is decreased by a factor e over a distance of 15–18 radiation units (5–6 km in the normal atmosphere). If such a shower falls vertically on the Earth, the distance from the shower maximum to the Earth's surface is about 1 km. In addition, it is well known that the energy of most of the electrons in the shower maximum is about 100 MeV. Thus, if the ionization losses are about 2 MeV/(g cm), the electron range is $z_0 \approx 500$ m. In addition, the probability of positron annihilation within the active interval z_0 is significantly lower than the probability of complete loss of energy through ionization. Thus, we suggest that the distribution of the excess-electron ranges z within the range from 0 to z_0 is nearly uniform.

Thus, the resulting field $dE(\omega)$ produced by the strip BC can be determined by integrating $dE(x, \omega)dz/z_0$ between 0 and z_0. Obviously, this is valid only for the multiplier $f(z)$ in Eq. (2) (the multiplier in square brackets). The calculations give $f_1(z) \longrightarrow f_2(y) = 1 - y^{-1}\sin y \times \exp(-iy)$, where $y = \omega z_0(1 + \beta\cos\varphi\sin\Theta_0)/2c$. Therefore, for $z_0 \approx 500$ m and the frequency $\nu \geq 200$ kHz, we obtain $y^{-1} < 1$ and $f_2(y) \approx 1$. The final expression for the resulting field at the observation point is obtained by integration of $dE(\omega)$ with respect to ξ over the interval (ξ_1, ξ_2):

$$E(\omega, \Theta_0, \varphi) = A(\varphi, \Theta_0)\int_{\xi_1}^{\xi_2}\exp\left(-i\frac{\omega(\xi_1 - \xi)}{c}\sin\varphi\right) \times \int_{x_1}^{x_2}\rho_d(r)dxd\xi\ (\text{V/m Hz}), \quad (6)$$

where

$$A(\varphi, \Theta_0) = \frac{2\cot\Psi\cos\Theta_0(1 - \cos^2\varphi\sin^2\Theta_0)^{1/2}\exp(-i\omega R_0/c)}{4\pi\varepsilon_0 2\pi c R_0(1 + \beta\cos\varphi\sin\Theta_0)(1 + \tan^2\Theta_0\sin^2\varphi)^{1/2}}, \quad (7)$$

$$r = [x^2 + (x - \xi)^2\cot^2\Psi\cos^2\Theta_0]^{1/2}; \quad \cot\Psi = \sin\varphi(\sin\Theta_0 - \cos\varphi)^{-1}. \quad (8)$$

To obtain the radiation diagram, it is necessary to determine the value of the function $F(\Theta_0, \varphi, \omega)$ on the interval $0 \leq \varphi \leq 2\pi$:

$$F(\Theta_0, \varphi, \omega) = |E(\Theta_0, \varphi, \omega)|^2\Big/\int_0^{2\pi}|E(\Theta_0, \varphi, \omega)|^2 d\varphi. \quad (9)$$

The measurements performed by the Yakutsk group of researchers [22, 23] have shown that the spatial distribution of charged particles within the range of distances of 50–500 m from the shower axis can be approximated by the function $\rho(r/r_0)^{1-s}(1 + r/r_0)^{s-4}C(s)$, where s is a parameter of the shower age and $f_0 \approx 100$ m. It is also well known that, at the distance $r \approx 3r_0$, the density reduces almost to zero. Therefore, the expression for the density normalized to the number of particles N_0 in the shower at its maximum can be recast as

$$\rho(r) = (16N_0/15\pi r_0^2)(r/r_0)^{-1}(1 + r/r_0)^{-3}. \quad (10)$$

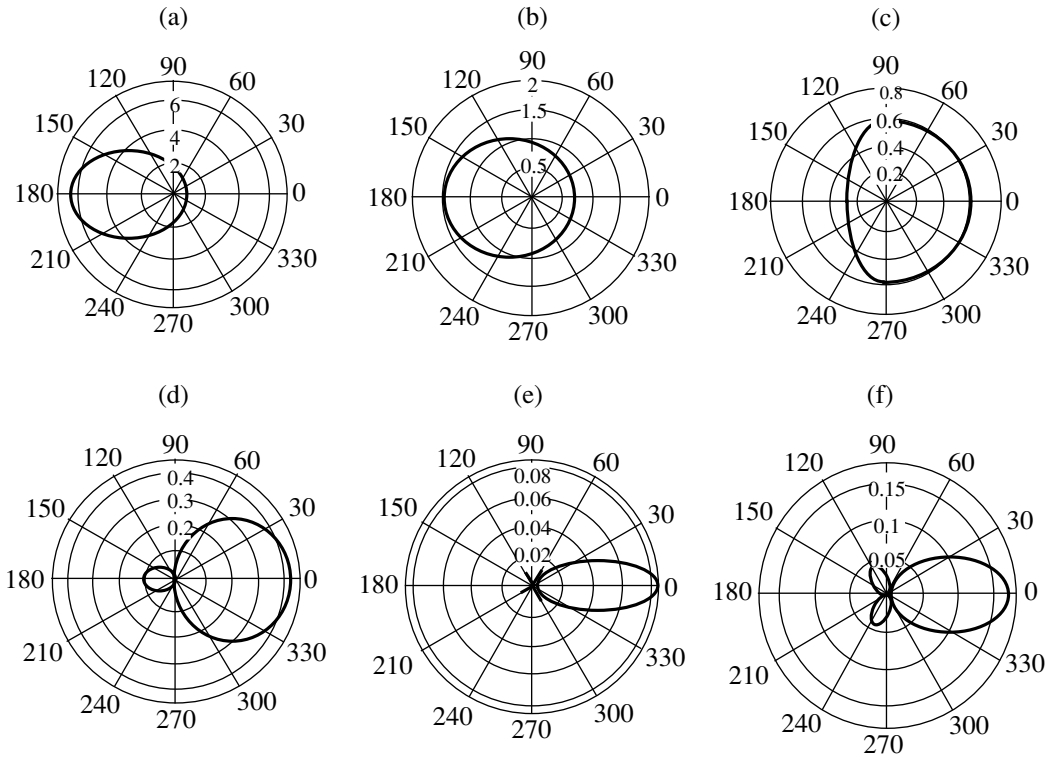


Fig. 2. Typical patterns of radiation at different frequencies and angles Θ.

On replacing the variables $x/r_0 = z$ and $\xi/r_0 = u$, Eq. (6) gives

$$E(\omega, \varphi, \Theta_0) = A(\varphi, \Theta_0)\frac{16N_0ke}{15\pi}\int e^{-i\omega_0(u_1-u)\sin\varphi} \times \left(\int I(z, \varphi, \Theta_0, u)dz\right)du, \quad (11)$$

where

$$I = (z, \varphi, \Theta_0, u) = [z^2 + (z-u)^2a^2]^{-1/2} \times \{[z^2 + (z-u)^2a^2]^{1/2} + 1\}^{-3},$$

$$a = \cot\Psi\cos\Theta_0,$$

$$z_{1,2}(u) = (ua^2 \pm a\sqrt{9(1+a^2) - u^2})/(1+a^2),$$

$$\cot\Psi = \sin\varphi/(\sin\Theta_0 - \cos\varphi),$$

$$u_{1,2} = \pm 3(1 + a^{-2}), \quad \omega_0 = \omega r_0/c, \quad k = 0.1,$$

$$c = 3\times 10^8 \text{ m/s}, \quad e = 1.6\times 10^{-19} \text{ C}.$$

The directional diagram was calculated under the assumption of a uniform distribution of the charged particle density ρ = const in a disk of radius r_0 = 100 m. In this case, the calculations can be significantly facilitated by changing the variables in Eq. (11): $v = u\sin\varphi = \xi/r_0\sin\varphi$. We obtain

$$E_1 = A(\varphi, \Theta_0, \omega_0) \times \int_{v_1}^{v_2} \cos[\omega_0(v_1 - v)][v_1^2 - v^2]^{1/2}dv, \quad (12)$$

$$E_2 = A(\varphi, \Theta_0, \omega_0) \times \int_{v_1}^{v_2} \sin[\omega_0(v_1 - v)][v_1^2 - v^2]^{1/2}dv, \quad (13)$$

where

$$A(\varphi, \Theta_0) = \cos^2\Theta_0 \times (1 - \cos^2\varphi\sin^2\Theta_0)^{1/2}/(1 + \beta\cos\varphi\sin\Theta_0) \times (1 + \tan^2\Theta_0\sin^2\varphi)^{1/2}(1 - \sin\Theta_0\cos\varphi)^2,$$

$v_{1,2} = \pm(1 - \sin\Theta_0\cos\varphi)/\cos\Theta_0$; E_1 and E_2 are the real and imaginary parts of the field intensity, respectively.

Typical radiation patterns for $\Theta = \pi/6$ and $\omega_0 = \pi\sqrt{3}/12$, $3\pi\sqrt{3}/12$, $4\pi\sqrt{3}/12$, $6\pi\sqrt{3}/12$, and $9\pi\sqrt{3}/12$ are shown in Figs. 2a–2e. The radiation pattern for $\Theta = \pi/4$ and $\omega_0 = 9\pi(\sqrt{2} + 1)/12$ is shown in Fig. 2f. Numerical analysis shows that, at sufficiently low frequencies ($\nu \approx 200$ kHz), the radiation maximum is

directed backward, i.e., opposite to the shower motion. As the frequency increases, the ratio of energies radiated forward and backward decreases (Fig. 2a, 2b). At $\nu \approx 0.8$ MHz, the energy radiated forward becomes greater than that radiated backward (Fig. 2c). At $\omega = 1.95$ and 2.12 MHz, the energy radiated forward is several orders of magnitude greater than the energy radiated backward (Fig. 2e, 2f). This provides an explanation for the experimentally observed lack of direct proportionality between the radio signal amplitude and the number of particles N_0 in the shower [11–18].

Analysis of the radiation patterns and Eqs. (12) and (13) shows that the signal intensity decreases with increasing frequency. This can be explained by mutual absorption of the wave zones (known in optics as Fresnel zones). An unexpected maximum of the radiating capacity of the shower was observed within a range of ≈0.2–1 MHz. This finding provides new insight into the results reported in [11–18] and earlier publications (1965–1980). In these papers, the experimentally observed value of the low-frequency (2–3 MHz) radio signal was found to be many times higher than the value calculated theoretically on the basis of the geomagnetic mechanism of radiation.

It should be noted in conclusion that several diagrams of the density distribution were obtained on the basis of Eq. (10). The time required for machine computation of such diagrams is several orders of magnitude longer than the time required for computing the radiation patterns. However, the difference between these results and the results obtained using Eq. (12) was found to be only moderate. It should be emphasized that the results obtained in this study are new and rather important, so that they should be taken into account in designing experiments. It is hoped that these results will stimulate in-depth studies of the possibilities of radio detection of ultrahigh-energy extensive air showers.

REFERENCES

1. V. L. Ginzburg, Usp. Fiz. Nauk **166**, 169 (1996) [Phys. Usp. **39**, 155 (1996)].
2. N. Nayashida *et al.*, Phys. Rev. Lett. **77**, 1000 (1996).
3. G. A. Askar'yan, Zh. Éksp. Teor. Fiz. **41**, 616 (1961) [Sov. Phys. JETP **14**, 441 (1962)].
4. G. A. Askar'yan, Zh. Éksp. Teor. Fiz. **48**, 988 (1965) [Sov. Phys. JETP **21**, 658 (1965)].
5. A. D. Filonenko, Izv. Ross. Akad. Nauk, Ser. Fiz. **61**, 543 (1997).
6. A. D. Filonenko, Pis'ma Zh. Tekh. Fiz. **23** (10), 57 (1997) [Tech. Phys. Lett. **23**, 399 (1997)].
7. A. D. Filonenko, Pis'ma Zh. Tekh. Fiz. **24** (24), 65 (1998) [Tech. Phys. Lett. **24**, 975 (1998)].
8. P. I. Golubnichiĭ and A. D. Filonenko, Pis'ma Zh. Tekh. Fiz. **20** (12), 57 (1994) [Tech. Phys. Lett. **20**, 499 (1994)].
9. P. I. Golubnichiĭ and A. D. Filonenko, Pis'ma Zh. Tekh. Fiz. **20** (23), 59 (1994) [Tech. Phys. Lett. **20**, 960 (1994)].
10. P. I. Golubnichiĭ, A. D. Filonenko, and V. I. Yakovlev, Izv. Ross. Akad. Nauk, Ser. Fiz. **58** (12), 115 (1994).
11. T. Kaneko, F. Kakimoto, K. Suga, *et al.*, in *Proceedings of the 18th Int. Cosmic Ray Conference, 1985*, Vol. 11, p. 428.
12. K. Suga, F. Kakimotro, and K. Nishi, in *Proceedings of the 19th Int. Cosmic Ray Conference, 1985*, Vol. 11, p. 268.
13. J. Nishimura, in *Proceedings of the 19th Int. Cosmic Ray Conference, 1985*, Vol. 7, p. 308.
14. K. Suga and K. Nishi, in *Proceedings of the 20th Int. Cosmic Ray Conference, 1987*, Vol. 6, p. 125.
15. F. Kakimoto, T. Umezawa, *et al.*, in *Proceedings of the 21st Int. Cosmic Ray Conference, 1990*, Vol. 9, p. 213.
16. P. Datta and K. M. Pathak, in *Proceedings of the 21st Int. Cosmic Ray Conference, 1990*, Vol. 9, p. 218.
17. P. Datta, R. Baishya, *et al.*, in *Proceedings of the 22st Int. Cosmic Ray Conference, 1991*, Vol. 4, p. 213.
18. R. Baishya, M. Deka, *et al.*, in *Proceedings of the 25th Int. Cosmic Ray Conference, 1997*, Vol. 6, p. 237.
19. A. D. Filonenko, Izv. Ross. Akad. Nauk, Ser. Fiz. **63**, 567 (1999).
20. V. L. Ginzburg, *Theoretical Physics and Astrophysics* (Nauka, Moscow, 1981; Pergamon, Oxford, 1979).
21. L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (Nauka, Moscow, 1967; Pergamon, Oxford, 1975).
22. V. S. Murzin, *An Introduction to the Physics of Cosmic Rays* (Mosk. Gos. Univ., Moscow, 1988).
23. *Problems of Cosmic Ray Physics* (Nauka, Moscow, 1987).

Translated by S. Chamorovskiĭ

Technical Physics, Vol. 45, No. 10, 2000, pp. 1362–1364. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 132–134.

BRIEF COMMUNICATIONS

Low-Frequency Radio Emission Induced by the Coherent Magneto-Bremsstrahlung of Charged Particles of an Extensive Air Shower

A. D. Filonenko

East Ukrainian State University, Lugansk, 348034 Ukraine

E-mail: uni@vugu.lugansk.ua

Received February 16, 2000

Abstract—A previously unknown coherent low-frequency ($0 < \nu < 50$ kHz) component of radio emission of an extensive air shower in the geomagnetic field is reported. This type of emission induces a high-intensity field (~200 μV/m MHz at a distance of 100 km from the shower axis). This emission mechanism is believed to be responsible for the previously observed high-intensity radio emission of the shower at low frequencies. The possibility of radio detection of ultrahigh-energy ($W_0 > 10^{21}$ eV) cosmic rays is considered. © *2000 MAIK "Nauka/Interperiodica".*

The fundamental possibility of enlarging the effective area of a detector composed of a relatively small number of sensitive elements is the main feature of the method of radio detection of cosmic rays. The number of scintillation counters used in conventional detectors is rather large [e.g., for the ShAL-1000 [1] (Extensive Air Shower, EAS-1000) Project, it is several tens of thousands], whereas a radio detector containing only several antennas may prove to be sufficient.

Experience shows that the advantages of the conventional method of detecting cosmic rays with relatively low energies ($W_0 < 10^{20}$ eV) are indisputable. On the other hand, to detect cosmic rays with energies of 10^{22}–10^{23} eV, the effective area of the detector should be about 10^6–10^7 km^2 (the incidence of events is taken to be the minimum permissible, i.e., ~10 year^{-1}). Obviously, it will be impossible to implement such a project in the foreseeable future.

Previous experimental studies of extensive air showers showed that the intensity of the high-frequency (30–100 MHz) radio emission induced by the Cherenkov [2] and geomagnetic [3] radiation mechanisms is rather low (5–10 μV/m MHz at a distance of 100 m from the shower axis [4]). In addition, the theoretically predicted direct proportionality between the radio signal amplitude and the shower energy was not observed in the overwhelming majority of events. At the frequencies of interest, the equivalent temperature of the celestial sphere is about 10^4–10^5 K. Therefore, the useful signal amplitude is comparable to galactic radio noise. For the frequency range under consideration, the significant discrepancy between the predicted and experimental values is presumably caused by the idealization of the radiation mechanisms. In particular, quantitative estimates were obtained without regard to the coherence violation caused by considerable scattering of the avalanche electrons during the longitudinal advancement of the shower. On the other hand, it was found that, for the radiation mechanisms mentioned above, the intensity of the low-frequency radio emission is several orders of magnitude greater than the theoretically predicted values. Presumably, this explains the growing interest of researchers in the nature of this type of radio emission (see [5] and references therein). The wavelengths of the emission of interest fall within the range $2L < \lambda < \infty$, where L is the longitudinal size of the shower. Shower δ-electron bremsstrahlung is a possible mechanism for the low-frequency radio emission. This mechanism does not involve either geomagnetic or geoelectric fields [6–9]. This mechanism was most effectively used to substantiate the method of radio detection of cosmic rays with energies of 10^{21}–10^{23} eV outside circumterrestial space [10]. Several variants of testing this method were suggested [11, 12].

The goal of this work was to study a previously unknown possibility of detecting high-energy ($>10^{21}$ eV) cosmic rays using the coherent low-frequency magneto-bremsstrahlung of an extensive air shower. It is shown below that the intensity of the field induced by this type of radio emission at a distance of 100 km from the shower axis is approximately 200 μV/m MHz for a particle with an initial energy of $W_0 = 10^{21}$ eV. This value is much higher than the field estimates for all previously known radio emission mechanisms of the shower. In addition, this type of radio emission has an extensive radiation pattern, which is also very important.

An avalanche electron with the mean energy $W_1 \sim 25$ MeV ($\gamma \approx 50$) describes an arc of radius $a = mV\gamma/eB \approx 1700$ m and length $S \approx W_1/(\partial W/\partial x) \approx 100$ m (taking into account the ionization losses, we obtain

1063-7842/00/4510-1362$20.00

$\partial W/\partial x \approx 2$ MeV/g cm^{-2}). For the sake of convenience, we assume that the shower falls vertically on the Earth's surface. In this case, we are concerned only with the electromagnetic energy emitted in the direction perpendicular to the shower axis. The problem of the radiation field of a point charge moving in an arc of arbitrary length has been already solved [13]. In the case under consideration, however, the results can be presented in a simpler and more convenient form. This can be done because we are concerned only with the low-frequency spectrum region. The values (**n**, **a**) are almost unchanged by integration, because $S \ll a \ll R_0$.

Let us express the electron (positron) velocity by a symmetrical step function $u(t)$: $\mathbf{v} = \mathbf{v}_1(t) + \mathbf{v}_2(t) = \mathbf{p}v[1 - V(t)] + \mathbf{q}vU(t)$, where v is the absolute value of the particle velocity and **p** and **q** are the unit vectors for $t < 0$ and $t > 0$, respectively. Assume that the electron moves uniformly and rectilinearly at $-\infty < t < \infty$ (except for the instant $t = 0$), i.e., $|\mathbf{v}_1| = |\mathbf{v}_2| = v$, whereas at the instant $t = 0$, the electron is in the middle of an infinitely small arc S (see figure). In this case, the electron acceleration is $\dot{\mathbf{v}} = v(\mathbf{q} - \mathbf{p})\delta(t)$, where $\delta(t)$ is the delta function. Such an idealization of the electron motion is permissible within the estimation accuracy, because the travel time S/c is far less than the period of the oscillatory circuit of the radio receiver ($S/c \lhd\lhd 2L/c \le T$) if the condition for coherence is met. In this case, the Fourier component of the field intensity is

$$\mathbf{E}(\omega, \mathbf{n}) \approx \frac{ee^{ikR_0}}{4\pi\varepsilon_0 cR_0} \times \int_{-\infty}^{\infty} \frac{[\mathbf{n}[(\mathbf{n}-\boldsymbol{\beta})\dot{\boldsymbol{\beta}}]]}{(1-\boldsymbol{\beta}\mathbf{n})^2} e^{i\omega\left(\tau - \frac{\mathbf{na}}{c}\right)} d\tau, \quad (1)$$

where all quantities are taken at the instant $\tau = t - R/c$, $\dot{\boldsymbol{\beta}} = \beta\delta(t)(\mathbf{q} - \mathbf{p})$, $\boldsymbol{\beta} = \mathbf{v}/c$, $\mathbf{n} = \mathbf{R}_0/R_0$, and **a** is the radius vector of the particle (the origin of the reference frame is placed at the center of a circle containing the arc S).

In the calculations, the Fourier component is determined to the accuracy of an insignificant common phase factor:

$$\mathbf{E}(\omega, \mathbf{n}) = \frac{e\beta}{4\pi\varepsilon_0 cR_0} \times \frac{(\mathbf{n} - \beta\mathbf{v}_0)(\mathbf{nW}_\perp) - \mathbf{W}_\perp[1 - \beta(\mathbf{v}_0\mathbf{n})]}{[1 - \beta(\mathbf{v}_0\mathbf{n})]^2}, \quad (2)$$

where, by the definition of $U(t)$, $\boldsymbol{\beta}(0) = \boldsymbol{\beta} = \beta\mathbf{v}_0$ and $\mathbf{v}_0$ is the unit vector directed along the vertical (see figure); the vector $\mathbf{W}_\perp = \beta(\mathbf{q} - \mathbf{p})$ is perpendicular to $\mathbf{v}_0$ and lies in the plane of the arc ($|\mathbf{W}_\perp| = \beta S/a$).

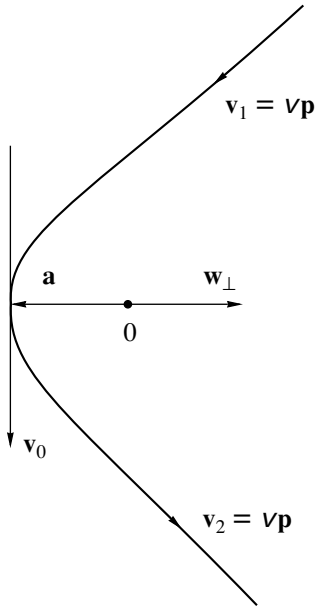


Figure.

In the direction of observation $\mathbf{n} \parallel \mathbf{W}_\perp$ and $\mathbf{n}_\perp\mathbf{v}_0$, we obtain

$$\mathbf{E}_\parallel = -\mathbf{v}_0 \frac{e|\mathbf{W}_\parallel|\beta}{4\pi\varepsilon_0 cR_0}. \quad (3)$$

It follows from Eq. (3) that the vector $\mathbf{E}_\parallel$ changes its sign if electrons are replaced with positrons. Therefore, the total contribution of the shower particles in this direction is close to zero. In the direction $\mathbf{n}_\perp\mathbf{W}_\perp$ and $\mathbf{n}_\perp\mathbf{v}_0$, we obtain

$$\mathbf{E}_\perp = -\mathbf{W}_\perp \frac{e\beta}{4\pi\varepsilon_0 cR_0}. \quad (4)$$

It follows from Eq. (4) that this component does not change its sign if e^+ replaces e^-, because the direction of the vector $\mathbf{W}_\perp$ simultaneously reverses. In other words, the radiation intensity is maximum in the direction perpendicular to the plane of the arc.

If N_0 is the total number of particles in the shower and L is the path length, the number of particles contributing to the field amplitude given by Eq. (4) is N_0L/S. Thus, we obtain

$$\left|\mathbf{E}_\perp^{\max}\right| = \frac{N_0 Le^2\mu_0 H}{4\pi\varepsilon_0 c^2 R_0 m\gamma}\left(\frac{\text{V}}{\text{MHz}}\right). \quad (5)$$

Substituting typical values $L = 6 \times 10^3$ m, $N_0 = 10^{12}$ (for $W_0 = 10^{21}$ eV), $e = 1.6 \times 10^{-19}$ C, $\mu_0 = 4\pi \times 10^{-7}$ H/m, $H = 40$ A/m, $4\pi\varepsilon_0 \approx 10^{-10}$ F/m, $c = 3 \times 10^8$ m/s, $R_0 = 10^5$ m, $m = 0.9 \times 10^{-30}$ kg, and $\gamma = 50$ into Eq. (5), we obtain $|\mathbf{E}_\perp| = 200$ μV/m MHz. This value of the field intensity at a distance of 100 km from the shower axis

is very high. However, it is difficult to obtain emission-induced fields of such intensity in practice. It is well known that the intensity of atmospheric disturbances increases with decreasing frequency. Within the range under consideration, the pulse amplitude is thought to be comparable to the desired signal amplitude. The experimental conditions can be radically changed if the measurements are taken in the circumpolar zone, where the intensity of atmospheric disturbances is 1.5 to 2 orders of magnitude lower than in middle latitudes.

The effective area of the Pierre Auger detector, which is presently under construction in the USA, will be 5×10^3 km^2. There is good reason to consider the possibility of radio detection of particles with energies 10^{22} eV and above. There is evidence that the energy spectrum of cosmic rays has a second bend at $W_0 = 10^{19}$ eV, so that the spectrum coefficient γ is reduced to $\gamma = 2.7$ [14, 15]. In this case, the particle flux intensity is

$$J = (>W_0) = 0.2 \times 10^4 W_0^{-1.7} \ (\text{m}^2\text{s sr})^{-1},$$

where W_0 is expressed in GeV. For $W_0 > 10^{22}$ eV, the incidence of showers over an area of 10^6 km^2 is found to be approximately 20 (years sr)$^{-1}$. Four antennas can be arranged over this area in such a way that any point of the area would be no more than 350 km away from one of the antennas. Taking into account that $W_0 \sim N_0$ and $E(\omega) \sim N_0$, we find, with regard to the estimations made above, that the value of $E(\omega)$ for this segment of the surface is $E(\omega) \geq 0.57$ mV/m MHz. This value of the field intensity is very high. Thus, a radio detector of this type can be expected to be quite effective.

Studies of radio-frequency pulses accompanying an extensive air shower are performed using high-speed electronic equipment. Synchronization of this type of equipment is performed using a master signal. One of the disadvantages of the experiment under consideration is the absence of a synchronizing pulse (in conventional detectors such pulses are generated by scintillation counters). The use of synchronizing pulses would significantly facilitate identification of the desired signal. However, except for technogenic noise, lightning discharges are the only source of high-power noise signals. Lightning discharge is a discontinuous current pulse with an overall duration of tens or even hundreds of milliseconds. On the other hand, the signal induced by an extensive air shower consists of a single pulse with a simple shape. According to the estimations made above, its duration is 10–20 μs. These are not the only specific features of the desired radio signal. It is also well known that a system of crossed antennas allows the location of the signal source or the direction toward the source to be determined. All this, as well as the fact that the radio emission mechanism considered in this work determines the horizontal polarization of waves (in the case of a vertical shower), is sufficient for unique identification of a radio signal.

REFERENCES

1. G. B. Khristiansen, V. V. Vashkevich, O. V. Vedeneev, *et al.*, Izv. Akad. Nauk SSSR, Ser. Fiz. **53** (2), 286 (1989).
2. G. A. Askar'yan, Zh. Éksp. Teor. Fiz. **48**, 988 (1965) [Sov. Phys. JETP **21**, 658 (1965)].
3. F. D. Kahn and I. Lerche, Proc. Phys. Soc. London, Sect. A **289**, 206 (1966).
4. V. B. Atrashkevich, O. V. Videneev, H. R. Allan, *et al.*, Yad. Fiz. **28**, 712 (1978).
5. A. D. Filonenko, Izv. Ross. Akad. Nauk, Ser. Fiz. **63**, 565 (1999).
6. P. I. Golubnichiĭ and A. D. Filonenko, Pis'ma Zh. Tekh. Fiz. **20** (12), 57 (1994) [Tech. Phys. Lett. **20**, 499 (1994)].
7. P. I. Golubnichiĭ and A. D. Filonenko, Pis'ma Zh. Tekh. Fiz. **20** (23), 59 (1994) [Tech. Phys. Lett. **20**, 960 (1994)].
8. P. I. Golubnichiĭ, A. D. Filonenko, and V. I. Yakovlev, Izv. Ross. Akad. Nauk, Ser. Fiz. **58** (12), 115 (1994).
9. P. I. Golubnichiĭ and A. D. Filonenko, Ukr. Fiz. Zh. **41**, 696 (1996).
10. A. D. Filonenko, Pis'ma Zh. Tekh. Fiz. **23** (10), 57 (1997) [Tech. Phys. Lett. **23**, 399 (1997)].
11. A. D. Filonenko, Izv. Ross. Akad. Nauk, Ser. Fiz. **61** (3), 543 (1997).
12. A. D. Filonenko, Pis'ma Zh. Tekh. Fiz. **24** (1), 65 (1998) [Tech. Phys. Lett. **24**, 975 (1998)].
13. V. G. Bagrov, I. M. Ternov, and N. I. Fedosov, Zh. Éksp. Teor. Fiz. **82** (5), 1442 (1982) [Sov. Phys. JETP **55**, 835 (1982)].
14. V. I. Ginzburg, Usp. Fiz. Nauk **166** (2), 169 (1996) [Phys. Usp. **39**, 155 (1996)].
15. V. S. Murzin, *An Introduction to the Physics of Cosmic Rays* (Mosk. Gos. Univ., Moscow, 1988).

Translated by S. Chamorovskiĭ

Technical Physics, Vol. 45, No. 10, 2000, pp. 1365–1367. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 135–137.

BRIEF COMMUNICATIONS

Interaction between Oscillations in a Self-stochastic System

É. V. Kal'yanov

Institute of Radio Engineering and Electronics, Russian Academy of Sciences, pl. Vvedenskogo 1, Fryazino, Moscow oblast, 141120 Russia

Received February 16, 2000

Abstract—A system of two bistable circuits connected by a low-pass filter is studied theoretically and experimentally. It is demonstrated that the system may produce self-excited chaotic oscillations. To examine the interaction between respective chaotic oscillations in bistable circuits, motion equations are derived and a numerical analysis of the interaction is carried out. Results of an experiment on a breadboard model are presented. © *2000 MAIK "Nauka/Interperiodica".*

The extensive use of bistable circuits, also known as flip-flops, and their expansion into the HF and microwave bands call for research into their interaction under different types of coupling. At higher frequencies, flip-flop ensembles are liable to spurious coupling, which in turn makes them operate in a complicated fashion. Remarkably, as few as two coupled flip-flops may produce chaotic oscillations under capacitive coupling [1]. This study examines interacting chaotic oscillations in a system of two flip-flops coupled via a low-pass filter.

The system is schematically shown in Fig. 1. Each flip-flop FFi (i = 1, 2) consists of an inductor L_i, a resistor R_i, a capacitor C_i, and a nonlinear element g_i (Fig. 1a). The low-pass filter consists of an inductor L_0 and capacitors C_{01} and C_{02}. It is attached at flip-flop terminals a_i and b_i (Fig. 1b).

As is known [2], a flip-flop has two stable states and goes from one of them to the other only on receiving an external trigger pulse. In contrast, coupled flip-flops such as those considered here may exhibit self-excited and self-sustaining oscillations with two basins of attraction. This is due to the fact that flip-flops FF_1 and FF_2, together with capacitors C_{01} and C_{02}, respectively, make up two chaotic oscillators coupled via inductor L_0; the oscillators are known as Chua's circuits [3]. If the flip-flops are identical, the inductor influences the interaction of self-excited oscillations in much the same way as a resistor connecting two identical Chua's circuits [4]. However, systems with differently adjusted oscillators are more relevant for practical purposes, because it appears almost impossible to make identical chaotic oscillators. Recall that chaotic systems are extremely sensitive to the initial conditions and parameter values.

With arbitrarily adjusted flip-flops, the system in Fig. 1b obeys the following normalized equations:

$$\begin{aligned}
\dot{x}_1 &= y_1 - x_1 - h_1(x_1),\\
\dot{y} &= (x_1 - y_1 + z_1 - u)\alpha_1^{-1},\\
\dot{z}_1 &= -(\beta_1/\alpha_1)y_1,\\
\dot{x}_2 &= \gamma_R(\alpha_2/\alpha_1)[y_2 - x_2 - h_2(x_2)],\\
\dot{y}_2 &= (\alpha_0/\alpha_1)[\gamma_R(x_2 - y_2 + z_2) + u],\\
\dot{z}_2 &= -(\beta_2/\alpha_1\gamma_R)y_2,\\
\dot{u} &= (\beta_0/\alpha_1)(y_1 - y_2).
\end{aligned} \tag{1}$$

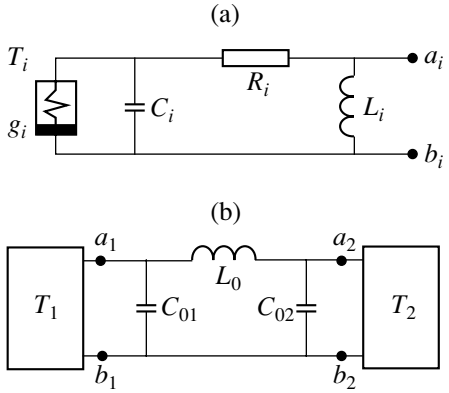


Fig. 1. System of two flip-flops connected by a low-pass filter: (a) the configuration of an individual flip-flop and (b) the circuit diagram of the connection.

1063-7842/00/4510-1365$20.00 © 2000 MAIK "Nauka/Interperiodica"

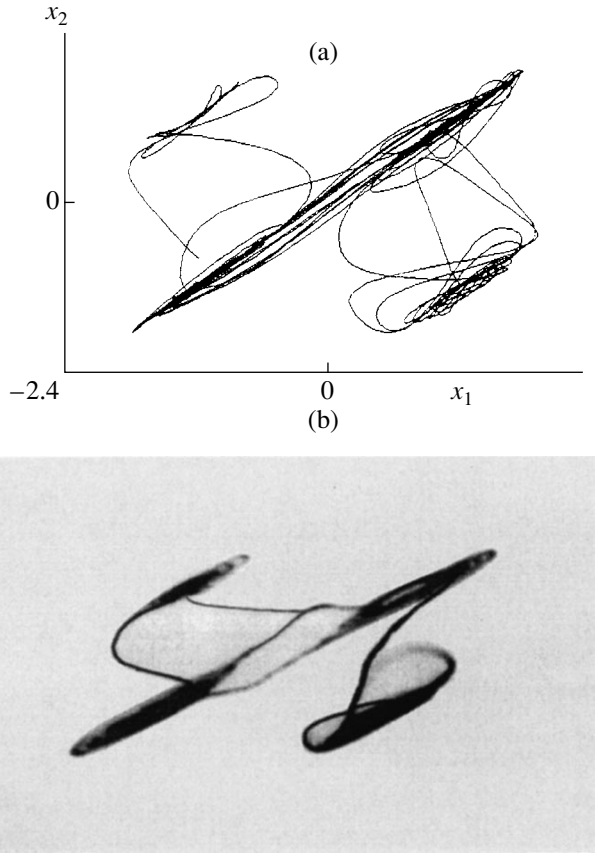


Fig. 2. Trajectory in the phase space of the system: results of (a) the numerical analysis and (b) the experiment.

The variables x_i and y_i stand for the voltages across capacitors C_i and C_{0i}, respectively; and z_i and u refer to the respective currents through inductors L_i and L_0. The dot denotes differentiation with respect to dimensionless time t. The constant coefficients are expressed in terms of circuit parameters as follows:

$$\alpha_i = C_{01}/C_i, \quad \alpha_0 = C_{01}/C_{02}, \quad \beta_i = C_{01}/L_i G_i^2,$$

$$\beta_0 = C_{01}/L_0 G_1^2, \quad \gamma_R = G_2/G_1, \quad G_i = 1/R_i.$$

There are different ways of approximating the respective nonlinearities $h_i(x_i)$ of active elements g_i. Here we adopt the approximation

$$h_i(x_i) = -\delta_i \tanh(x_i), \tag{2}$$

where δ_i are constants.

The system of nonlinear differential equations (1) and (2) can only be solved numerically. We carried out a numerical analysis of (1) and (2) for the cases where the flip-flops were adjusted identically and differently, respectively. In the former case, we set $\alpha_1 = \alpha_2 = 10$, $\alpha_0 = \gamma_R = 1$, $\beta_1 = \beta_2 = 15$, and $\delta_1 = \delta_2 = 1.5$. The amount of coupling between the flip-flops is evaluated by β_0. This parameter was varied in a wide range. For the differently adjusted flip-flops, we set $\beta_2 = 14$; the other parameter values were the same as in the above.

In each of the two cases, we found that chaotic oscillations are excited and that the oscillations $x_1(t)$ and $x_2(t)$ have chaotic attractors of the double-scroll type, as with Chua's circuit. Indeed, chaotic oscillations interact with each other in a more complicated fashion if the flip-flops are adjusted differently.

Note that, with misadjusted flip-flops, the oscillations $x_1(t)$ and $x_2(t)$ differ under any initial conditions if the coupling is weak or absent. Accordingly, there are four dense regions in the x_1–x_2-plane. For small β_0, the motion is directed randomly in the dense regions. As β_0 increases, one can see a growing mutual synchronism in intrabasin oscillations: in the same regions, the representative point now moves predominantly parallel to the set $x_1 = x_2$. On the other hand, the presence of *four* dense regions indicates that the flip-flop transitions are not fully synchronous with each other. If total synchronism is reached, $x_1(t)$ and $x_2(t)$ share one of two basins of attraction (a positive or a negative one). In contrast to identical flip-flops, the representative point in the x_1–x_2-plane moves outside the set $x_1 = x_2$ (but not far from it).

We also performed a physical experiment on a breadboard model with $R_1 = R_2 = 0.4$ kΩ, $C_1 = C_2 = 6$ μF, $L_1 = 17$ mH, $L_2 = 18$ mH, and $C_{01} = C_{02} = 4.7$ μF. Nonlinear elements g_i were built around operational amplifiers using LM324N integrated circuits.

Figure 2 presents the most important results of the numerical analysis and the physical experiment. Figure 2a shows a computed trajectory projected onto the x_1–x_2-plane for $t \in [1500, 3000]$ at $\beta_0 = 15$, with initial conditions $x_1(0) = 0.1$, $x_2(0) = -0.1$, and $y_i(0) = z_i(0) = u(0) = 0$. Figure 2b shows a measured trajectory for a unit time interval at $L_0 = 38$ mH. It is seen that the experimental results agree with the numerical ones.

It is important to note that both the computation and the experiment revealed the existence of a synchronization threshold for chaotic oscillations. Recall that no such threshold exists in regular systems [5, 6].

To sum up, the computation and the experiment have demonstrated that interacting chaotic oscillations may arise in a system of flip-flops coupled via a low-pass filter. The results appear to give a better understanding of how spurious coupling arises in flip-flop ensembles and how differing Chua's circuits interact with each other via inductive coupling.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (grant no. 98-02-16622).

REFERENCES

1. É. V. Kal'yanov, Pis'ma Zh. Tekh. Fiz. **25** (6), 1 (1999) [Tech. Phys. Lett. **25**, 207 (1999)].
2. W. J. Cunninghem, *Introduction to Nonlinear Analysis* (McGraw-Hill, New York, 1958; Gosénergoizdat, Moscow, 1962).
3. C. W. Wu, T. Yang, and L. O. Chua, Int. J. Bifurcation Chaos **6**, 455 (1996).
4. V. V. Astakhov, A. V. Shabunin, and V. S. Anishchenko, Radiotekh. Élektron. (Moscow), No. 8, 974 (1997).
5. P. S. Landa, *Self-oscillations in Systems with a Finite Number of Degrees of Freedom* (Nauka, Moscow, 1980).
6. B. E. Zhelezovskiĭ and É. V. Kal'yanov, *Multifrequency Operating Modes in Microwave Devices* (Svyaz', Moscow, 1978).

Translated by A. Sharshakov

Technical Physics, Vol. 45, No. 10, 2000, pp. 1368–1370. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 138–140.

BRIEF COMMUNICATIONS

Localization of Stretching Strain in Doped Carbon γ-Fe Single Crystals

S. A. Barannikova

Institute of Strength Physics and Materials Science, Siberian Division, Russian Academy of Sciences, Akademicheskiĭ pr. 2/1, Tomsk, 634021 Russia

E-mail: lev@zuevs.tomsk.ru

Received February 16, 2000

Abstract—The evolution of local strain during stretching of high-manganese carbon austenite was studied. The ordered patterns of strain localization proved to be closely related to the stages in the stress–strain curve. The results of this study are compared with analogous data for chromium–nickel nitrogen austenite single crystals. The velocity of self-consistent motion of the sites where plastic strain during stretching of γ-Fe single crystals is nonuniform was determined as a function of the strain hardening coefficient and deformation mechanism. © *2000 MAIK "Nauka/Interperiodica".*

Speckle interferometry data for single-crystal and polycrystalline metals and alloys [1–3] indicate that strain localization may be of three types, each corresponding to a particular stage of plastic flow. The solitary strain front, as well as mobile and stationary dissipative strain structures, were interpreted [1, 2] as various self-sustained processes in a nonlinear medium [4].

Samples used in the experiments were single crystals of Gadfield (Fe–13% Mn–0.93% C) steel grown by the Bridgman method in a helium atmosphere. An electrical discharge machine was used to cut out samples in the form of a double spade with a cross section of 1.5×5.0 mm and an operating length of 28 mm. The single crystals were homogenized for 24 h at 1373 K in an inert gas, kept at 1373 K for 1 h, and quenched in water. The samples were stretched on an Instron-1185 testing machine at room temperature at a rate of 1.2×10^{-4} s^{-1}. The field of displacement vectors on the sample front surface was determined by speckle interferometry [3] in the range between the yield point and tensile strength at 15-s intervals (in 0.2% increments of the total strain). By numerical differentiation with respect to coordinates, the distributions of the longitudinal, transverse, shear, and rotational components of the plastic distortion tensor were calculated at each point of the surface under study. Then, for different time instants or various values of the total strain, the local strain distribution over the whole sample or along its axial line was plotted.

In carbon γ-Fe single crystals oriented in the [012] direction, gliding is the main mechanism of stretching-induced plastic flow [5]. Beyond the yield point, the stress–strain curve has two regions where the strain hardening coefficient is constant (Fig. 1a).

The distributions of the distortion tensor components were recorded for the total strain ε_{tot} in the range of 1–12%. As follows from the distribution of the local elongations ε_{xx} (the other components behave similarly), at the initial stage of easy glide beyond the yield point, the solitary strain front shifts from one grip of the testing machine to the other. As the strain front passes through the sample (section *EF* in the stress–strain curve), its direction reverses and its velocity changes (section *FK*) (Fig. 1a). The strain front velocity *V* was evaluated from the slope of the linear distribution of the

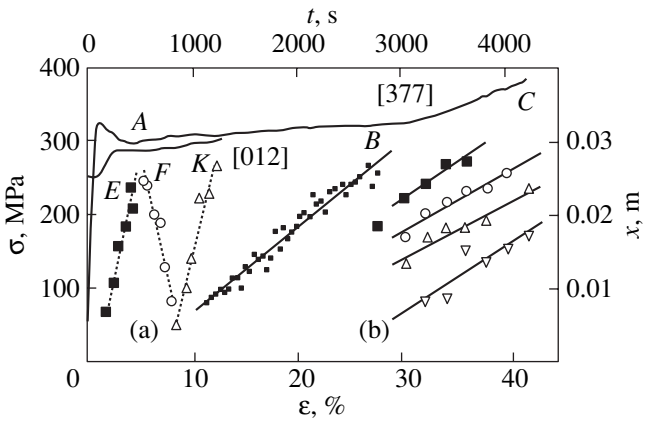


Fig. 1. Distributions of the local elongation peaks vs. time of stretching in accordance with the stress–strain curves for Gadfield steel. (a) Easy glide in single crystals [012] and (b) easy glide and linear hardening in single crystals [377].

1063-7842/00/4510-1368$20.00

Strain front velocity V vs. strain hardening coefficient Θ and deformation mechanism at various stages of the $\sigma(\varepsilon)$ curve for Gadfield steel single crystals

Orientation	Deformation mechanism	Stages of $\sigma(\varepsilon)$ curve	Θ, MPa	V, m/s
[012]	Dislocation gliding	Easy glide (I) EF	130	5×10^{-5}
		Easy glide (II) FK	170	7×10^{-5}
[377]	Twinning	Easy glide (I) AB	80	9×10^{-6}
		Linear hardening (II) BC	550	1.25×10^{-4}

coordinates x of the ε_{xx} peaks with time of stretching t (the distribution of the local elongations is shown in Fig. 2a). We obtained $V_I = V_{II} = 5 \times 10^{-5}$ m/s and $V_{III} = 7 \times 10^{-5}$ m/s.

It is worth noting that, in the strained Gadfield steel crystals with the [012] orientation, plastic zones moved with a velocity of $\sim(5–7) \times 10^{-5}$ m/s. In glide-strained nitrogen chromium–nickel austenite crystals, the deformation zones at the stage of easy glide moved with a velocity of $\sim 3.5 \times 10^{-5}$ m/s [2].

According to [5], in carbon γ-Fe single crystals oriented in the [377] direction, the greatest contribution to the plastic strain at room temperature is from twinning, which arises at the very beginning of plastic deformation. The easy glide stage, the sharp yield point included, extends to $\varepsilon \approx 30\%$ (Fig. 1b). At this stage, the Lüders band propagates throughout the crystal. Further deformation of the [377] oriented crystals leads to linear hardening up to 45% with subsequent fracture. The stage of parabolic hardening was observed neither in the [377] nor in the [012] oriented crystals.

We also studied the distribution of local-strain zones. Throughout the stage of easy glide (section *AB* in the stress–strain curve), the solitary front of plastic strain moves with a velocity of $V_I = 9 \times 10^{-6}$ m/s (Fig. 1b). At the stage of linear hardening (*BC*), four or five local-strain peaks move along the sample axis with a constant velocity of $V_{II} = 1.25 \times 10^{-4}$ m/s. The spatial period of the process at its linear stage (Fig. 1b) was found to be $\lambda = 5 \pm 1$ mm.

In strained Gadfield steel samples oriented along [377], the solitary plastic front moves with the velocity $\sim 9 \times 10^{-6}$ m/s at the stage of easy glide. This value is less than that in the [012] case and those reported in early works [1, 2] for Cu–10% Ni–6% Sn single crystals ($V \sim 6.5 \times 10^{-5}$ m/s), titanium nickelide ($V \sim 1.5 \times 10^{-5}$ m/s), high-nitrogen steels ($V \sim 3.5 \times 10^{-5}$ m/s), etc. This lower velocity may be related to twinning, unlike the previous cases.

At the stage of linear hardening, the set of plastic strain fronts moves along the sample with a velocity of $\sim 1.25 \times 10^{-4}$ m/s, which is higher than in nitrogen austenite ($V \sim 2 \times 8 \times 10^{-5}$ m/s). In our previous study on high-nitrogen chromium–nickel γ-Fe [6], the dependence of the strain front velocity on the coefficient of linear strain hardening was obtained. Specifically, during linear hardening of carbon high-manganese γ-Fe single crystals with the [377] orientation, the sites of nonuniform plastic strain (a site of nonuniformity is shown in Fig. 2a) propagate with a velocity which is inversely proportional to the hardening coefficient: $V \sim \Theta^{-1}$ (Fig. 2b).

The results of this study confirm the conclusions that, as the total strain increases, the local strain distributions and their evolution depend on the strain hardening coefficient [1, 2]. However, for plastic flow in doped carbon γ-Fe single crystals, the kinetics of the periodic process has several particular features. From the experimental data (see table), it follows that the velocity of the zones of plastic flow localization is defined by the strain hardening coefficient and deformation mechanisms, which are gliding and twinning.

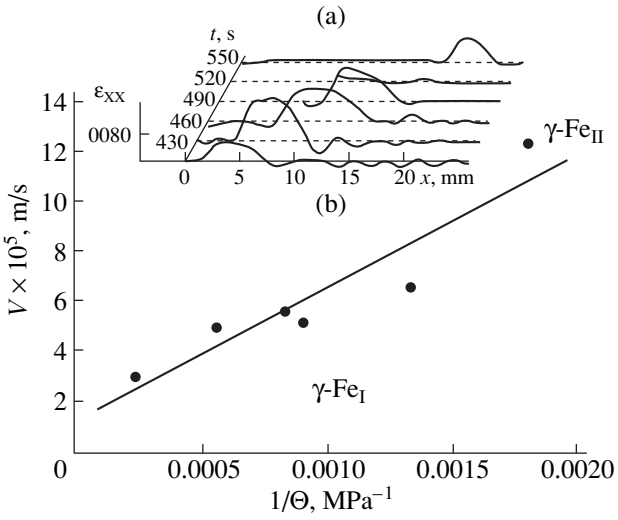


Fig. 2. (a) Distribution of the local elongations that illustrates the motion of the localized strain zone and (b) velocity of the localized strain zones vs. strain hardening coefficient for γ-Fe_I (nitrogen chromium–nickel austenite) and γ-Fe_{II} (carbon high-manganese austenite) single crystals.

The elucidation of the $V(\Theta)$ dependence can provide a better insight into the physics of the periodic processes observed in strained crystals.

REFERENCES

1. L. B. Zuev, V. I. Danilov, and N. V. Kartashova, Mater. Sci. Eng., A **234–236**, 699 (1997).
2. L. B. Zuev, S. A. Barannikova, and V. I. Danilov, Zh. Tekh. Fiz. **69** (10), 56 (1999) [Tech. Phys. **44**, 1179 (1999)].
3. R. Jones and C. Wykes, *Holographic and Speckle Interferometry* (Cambridge Univ. Press, Cambridge, 1983; Mir, Moscow, 1986).
4. V. A. Vasil'ev, Yu. M. Romanovskiĭ, and V. G. Yakhno, *Self-sustained Processes* (Nauka, Moscow, 1987).
5. E. I. Litvinova, I. V. Kireeva, E. G. Zakharova, *et al.*, Fiz. Mezomekh. **2** (1–2), 115 (1999).
6. S. A. Barannikova, L. B. Zuev, and V. I. Danilov, Fiz. Tverd. Tela (St. Petersburg) **41**, 1222 (1999) [Phys. Solid State **41**, 1112 (1999)].

Translated by A. Sidorova

Technical Physics, Vol. 45, No. 10, 2000, pp. 1371–1373. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 10, 2000, pp. 141–143.

BRIEF COMMUNICATIONS

The Properties of $CuInSe_2$ Films Obtained by Selenation of Cu–In Layers

E. P. Zaretskaya, V. F. Gremenyuk, V. B. Zalesskiĭ, V. A. Ivanov, I. V. Viktorov, V. I. Kovalevskiĭ, O. V. Ermakov, and T. P. Leonova

Institute of Solid-State and Semiconductor Physics, Belarussian Academy of Sciences, ul. Brovki 17, Minsk, 220076 Belarus

Institute of Electronics, Belarussian Academy of Sciences, Minsk, 220076 Belarus

E-mail: gremenok@Ifttp.bas-net.by

Received February 26, 2000

Abstract—$CuInSe_2$(112) films with the chalcopyrite structure were obtained by the selenation of Cu–In layers (Cu/In = 0.5–2.9) in a closed-type reactor. The surface morphology, phase composition, electrical performance, and low-temperature photoluminescence (PL) of the films depend on the Cu/In ratio. © *2000 MAIK "Nauka/Interperiodica".*

In the last decade, three-component copper-based compounds of class $CuB^{III}C_2^{VI}$, the energy gap of which is best suited for converting solar radiation, have received much attention. Among these materials, $CuInSe_2$ (CIS) semiconductors and $CuIn_xGa_{1-x}Se_2$ (CIGS) solid solutions are considered to be the most promising for thin-film solar cells (SCs) [1–4]. In this work, we studied the physical properties of $CuInSe_2$ films obtained by two-stage synthesis in a closed-type system.

0.4- to 0.8-μm-thick Cu–In layers (Cu/In = 0.5–2.9) used for film synthesis were applied on pure glass substrates and on glass substrates coated with molybdenum in two ways: (1) by thermal evaporation at a pressure of 5×10^{-6} torr and substrate temperature T = 100°C (metal charges were simultaneously evaporated from a molybdenum boat at the mean rate 0.5 μm/min) or (2) by ion–plasma sputtering of a composite (two-component) target in a triode-type system (argon atmosphere, pressure 2×10^{-4} torr, deposition rate 30 Å/min).

Selenation was accomplished by two-stage synthesis. At the first stage, the Cu–In alloy reacts with selenium at 280–300°C for 20–50 min. At the second stage, the formed layer is recrystallized at 500°C for 20–90 min. The structure and phase composition of the Cu–In layers before and after selenation were studied by X-ray diffraction analysis (CuK_α radiation, 2θ = 15–100°). Phases were identified by contrasting experimentally found interplanar spacings d with those listed in the JCPDS tables. It turned out that the thermally evaporated Cu–In films represent a mixture of binary components Cu_xIn and elementary In and Cu, while the layers produced by ion–plasma sputtering have the $Cu_{11}In_9$ phase alone. Quantitative analysis data for the $CuInSe_2$ (±5%) layers are listed in the table.

Selenation produces polycrystalline $CuInSe_2$ films. Their phase composition and structure depend on the Cu/In ratio in the starting layer [5–8]. When copper-enriched (Cu/In > 1.2) layers are selenated, the chalcopyrite-like $CuInSe_2$ compound is formed. It contains Cu_xSe traces (Fig. 1), which agrees well with the Cu_2Se–In_2Se_3 pseudobinary phase diagram [6]. In diffractograms taken from these films, the (101) and (103) superlattice reflections are absent. This is indicative of cation sublattice disorder. Indium-enriched (Cu/In < 0.69) films also contain the disordered chalcopyrite-like $CuInSe_2$ phase and, in addition, a small amount of In_2Se. In the latter films, the (116)/(312) doublet remains unresolved (Fig. 1, curve *a* in the insert). $CuInSe_2$ layers with Cu/In = 0.83–1.04 are single-phase and have the chalcopyrite structure. Along with the usual reflections, (112), (220)/(204), (312)/(116), etc., diffractograms taken from these films also show the superlattice reflections (101), (103), and (211); moreover, tetragonal splitting of the (116)/(312) doublets (Fig. 1, curve *b* in the insert) is observed. All these observations are typical of the ordered chalcopyrite structure.

Table

Ratio of film elements		at. %		
Cu/In	Se/(Cu + In)	Cu	In	Se
0.52	1.18	16.09	29.80	54.11
0.69	1.07	19.76	28.59	51.65
0.83	1.06	22.66	25.81	51.53
1.29	0.99	28.15	21.96	49.89

1063-7842/00/4510-1371$20.00 © 2000 MAIK "Nauka/Interperiodica"

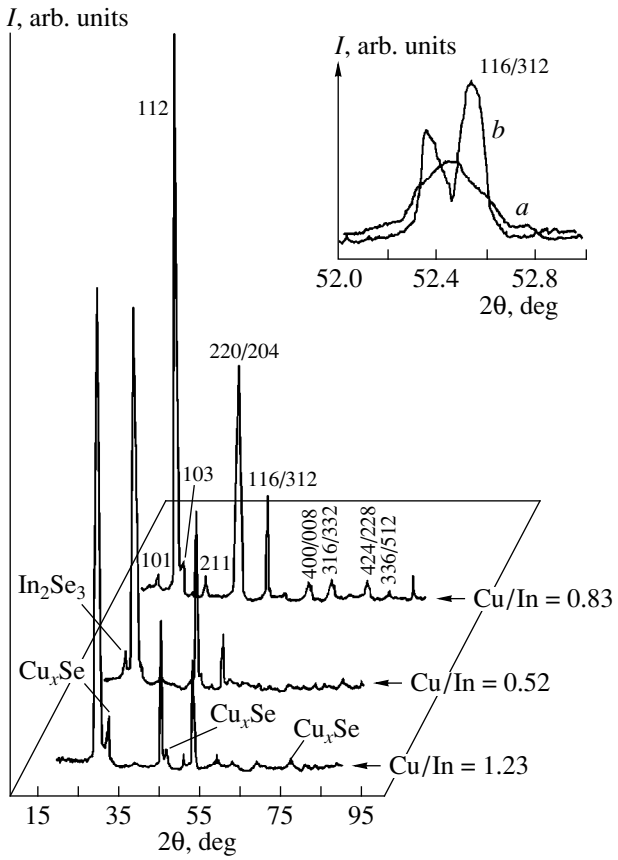


Fig. 1. Diffractograms taken from $CuInSe_2$ films with different Cu/In ratios. The insert shows the resolution of the (116)/(312) doublet: Cu/In = (*a*) 0.52 and (*b*) 0.83.

Micrographs of the film surface show that the growth mechanism of the films is related to the Cu/In ratio, which controls phase transformations in Cu_2Se–In_2Se_3 solid solutions [6]. The indium-enriched (Cu/In ≤ 0.69) films consist of compact fine-grained unfaceted agglomerates (Fig. 2a). The surface of the films with Cu/In = 1.29 is made up of tetrahedral grains 2–4 μm in size (Fig. 2b). Their composition corresponds to a Cu_xSe binary compound [6, 7].

The films were of *p*-type conduction, and their resistivity varied from 10^{-2} to 10^3 Ω cm with increasing indium content. From the exponential temperature dependence of the resistivity of the $CuInSe_2$ thin films, $\delta = \delta_0 \exp(\Delta E/kT)$, the activation energies for levels in the energy gap were estimated. In the 80–400 K temperature range, these are ΔE_1 = 42–44 meV and ΔE_2 = 71–73 (or 92–94) meV. Note that the level with ΔE_1 was observed for films with different Cu/In ratios. The presence of the levels with ΔE_2 is essentially dependent on the film composition and is in good agreement with results of other works [9, 10]. The activation energy ΔE_1 can be assigned to copper vacancies V_{Cu}, which produce acceptor levels, and ΔE_2, to selenium vacancies V_{Se} or substitutional defects In_{Cu}. The last two produce donor levels in $CuInSe_2$ [1, 9]. The optical properties of the films vs. composition were determined from PL studies at 77 K. The PL spectra were found to have a single line, whose half-width and peak position depend on the Cu/In ratio. As the In content grows, the PL line broadens and its peak shifts from 0.975 eV (Cu/In = 1.29) to 0.860 eV (Cu/In = 0.69). If the indium content is very high (Cu/In = 0.52), the peak is situated at 0.943 eV. $CuInSe_2$ films used to fabricate SC modules with an efficiency η > 12% [4, 9] exhibited a similar PL spectrum. The PL spectral lines in $CuInSe_2$ films can be associated with donor-to-acceptor transi-

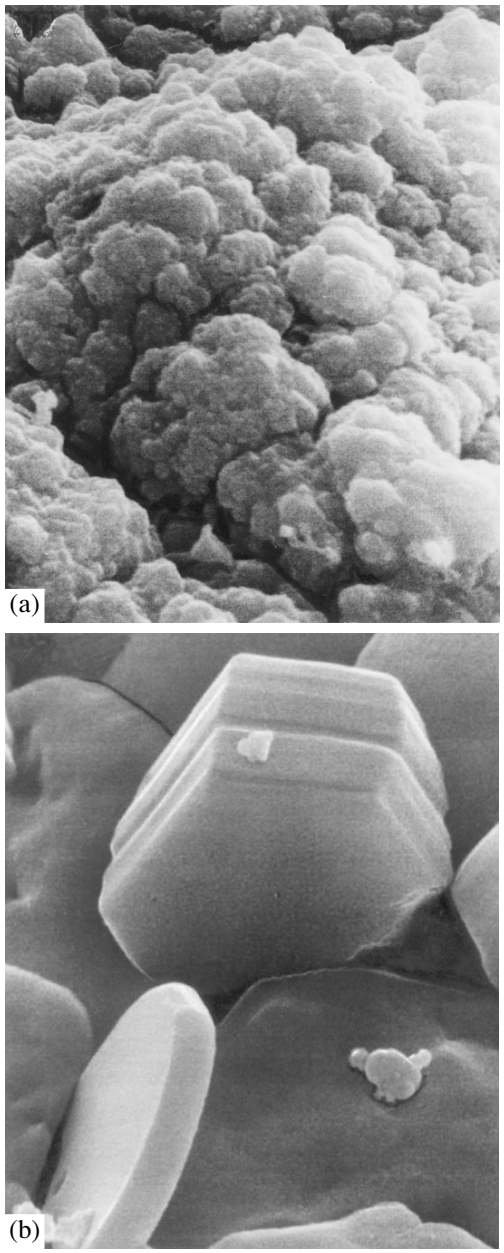


Fig. 2. Surface micrographs for $CuInSe_2$ films with Cu/In = (a) 0.69 and (b) 1.29.

tions involving an acceptor level V_{Cu} and two donor levels V_{Se} and In_{Cu} [9, 11].

Thus, the physical properties of $CuInSe_2$ films obtained by two-stage selenation in a closed volume were shown to depend on the composition of the starting Cu–In layers. The resulting $CuInSe_2$ films are suitable for thin-film SCs.

ACKNOWLEDGMENTS

This work was supported by the Belarussian Republican Foundation for Basic Research and Joint Belarussian–Russian Program "Kosmos-BR."

REFERENCES

1. A. Rockett and R. W. Birkmire, J. Appl. Phys. **70**, R81 (1991).
2. J. Hedstron, H. Ohlsen, M. Bodegart, *et al.*, in *Proceedings 23rd Photovoltaic Special Conference, New York, 1993*, p. 364.
3. S. K. Deb, in *Proceedings of the International Physical Conference, 1998*, Ser. No. 152, p. 923.
4. V. Alberts, J. H. Schon, J. W. Witcomb, *et al.*, J. Phys. D **31**, 2869 (1998).
5. J. H. Schon, V. Alberts, and E. Bucher, Thin Solid Films **301** (3), 115 (1997).
6. J. Kessler, H. Dittrich, F. Grunvald, and H. W. Shock, in *Proceedings 12th EC Photovoltaic Energy Conference, Amsterdam, 1994*, p. 879.
7. F. O. Adurodija, J. Forbes, M. J. Carter, and R. Hill, J. Mater. Sci. Lett. **11**, 478 (1996).
8. F. O. Adurodija, J. Song, S. D. Kim, *et al.*, Thin Solid Films **9**, 13 (1999).
9. J. H. Schon, V. Alberts, and E. Bucher, Thin Solid Films **301** (3), 115 (1997).
10. M. Tanda, S. Manaka, and A. Yamada, Jpn. J. Appl. Phys. **32**, 1913 (1993).
11. S. Zott, K. Leo, M. Ruckh, and H. M. Schock, Cryst. Res. Technol. **31**, 729 (1996).

Translated by V. Isaakyan