Variational principles for constructing few-mode models of the laminar-to-turbulent transition

N. M. Zubarev

Institute of Electrophysics, Urals Branch of the Russian Academy of Sciences, 620049 Ekaterinburg, Russia
(Submitted September 18, 1995; resubmitted January 22, 1996)
Zh. Tekh. Fiz. **67**, 1–5 (May 1997)

A method based on variational principles is proposed for constructing few-mode models of distributed nonlinear systems to a higher accuracy of approximation than the Galerkin method for a prespecified basis of trial functions. This method may be useful for studying the initial stages of the laminar-to-turbulent transition, when the increase in the number of trial functions sharply increases the stiffness of the reduced system, so that it is advisable to use optimal few-mode models. The proposed method is used to construct and analyze a modification of the Lorenz model for a Saltzman system. © *1997 American Institute of Physics.*
[S1063-7842(97)00105-0]

INTRODUCTION

The Ruelle–Takens theorem,[1] which states that the transition to chaos in dynamical systems can occur through a finite number of bifurcations, raises interest in finite-dimensional models for describing the laminar-to-turbulent transition in systems with distributed parameters. Such models are ordinarily constructed by the Galerkin procedure,[2] which gives a system consisting of a finite number of ordinary differential equations for the amplitudes of the perturbations. The choice of a basis of trial functions in the procedure is based on *a priori* information about the behavior of the system and the character of its instabilities, and the minimum number of these functions that can provide sufficiently complex behavior is 3. It is clear, however, that the use of a small number of modes in the Galerkin method leads to a significant inaccuracy of approximation. Of course, the necessary accuracy can always be achieved by keeping a large number of modes of the perturbations. However, enlarging the basis of trial functions without changing the number of order parameters in the dynamics of a system near the threshold of its instability (see, e.g., Ref. 3) will significantly increase the stiffness of the system and require a substantial decrease in the size of the numerical integration step.[4]

In this situation it seems advisable to proceed with the analysis of the initial stages of the laminar-to-turbulent transition by constructing an optimal model in the framework of a prespecified basis of a small number of trial functions. This paper gives a further development of the method proposed by the author in Ref. 5 for constructing such a model from considerations of minimizing the integral of the square of the residual in the neighborhood of the moving point of the dynamical system, which permits a significant improvement in accuracy as compared to the Galerkin approximation.

The use of the method is demonstrated on the Saltzman system, which arises in a number of problems in the description of the initial stages of the laminar-to-turbulent transition (the onset of Bénard cells[6] or stratification of a current-carrying liquid-metal conductor[7,8]). For this system we shall construct the simplest three-mode model, which turns out to be substantially different from the Lorenz model obtained by the Galerkin procedure.[9]

BRIEF EXPOSITION OF THE METHOD

We consider a partial differential equation of the form

$$\frac{\partial u}{\partial t}=F\left(\frac{\partial}{\partial x_1},\ldots,\frac{\partial}{\partial x_m},\mathbf{r},u\right), \tag{1}$$

where $\mathbf{r}=\{x_1,\ldots,x_m\}$ is the spatial coordinate, the function u on its domain of definition V is subject to the initial condition $u(\mathbf{r},t)|_{t=0}=u^0(\mathbf{r})$, and on the boundary of this domain the function u satisfies the appropriate boundary conditions for the physics of the phenomenon under study.

Let us now define the scalar product on V as $(f(\mathbf{r}),g(\mathbf{r}))=\int_V f(\mathbf{r})g(\mathbf{r})dV$. We will seek an approximate solution of equation (1) in the form

$$u=\tilde{u}=\sum_{i=1}^{n}A_i(t)H_i(\mathbf{r}), \tag{2}$$

where the trial functions H_i satisfy the necessary boundary conditions and, in addition, obey the condition that $(H_iH_j)\equiv 0$ only if $i=j$, i.e., they are orthogonal. We then introduce the residual

$$N(\dot{\mathbf{A}},\mathbf{A},\mathbf{r})=\frac{\partial\tilde{u}}{\partial t}-F\left(\frac{\partial}{\partial x_1},\ldots,\frac{\partial}{\partial x_n},\mathbf{r},\tilde{u}\right),$$

where $\mathbf{A}=\{A_1,\ldots,A_n\}$.

In accordance with the Galerkin method (see, e.g., Ref. 2), a system of n ordinary differential equations (ODEs) for the amplitudes $A_1,\ldots,A_n$ can be obtained from the requirements of orthogonality of the residual of the trial functions:

$$(H_i(\mathbf{r}),N(\dot{\mathbf{A}},\mathbf{A},\mathbf{r}))=0,\quad i=1,\ldots,n. \tag{3}$$

Ultimately we obtain the system of equations

$$\dot{\mathbf{A}}=\mathbf{f}(\mathbf{A}) \tag{4}$$

with the initial conditions

$$A_i(0)=(u^0,H_i)/(H_i,H_i),\quad i=1,\ldots,n, \tag{5}$$

 1063-7842/97/050455-05$10.00

where $\mathbf{f}=\{f_1, \ldots, f_n\}$ are nonlinear (in general) functions of the amplitudes.

In the present paper we propose to construct a system of ODEs on considerations that the following functional be minimal on the trajectory:

$$S=\int_{t_0-\varepsilon}^{t_0+\varepsilon} L(\dot{\mathbf{A}},\mathbf{A})dt \equiv \int_{t_0-\varepsilon}^{t_0+\varepsilon} (N,N)dt, \tag{6}$$

where t_0 is the moving time variable and ε is the time interval on which the minimization is carried out.

This means that we launch the trajectory from the moving point in such a way that is optimally approximates the initial equation (1) in a temporal neighborhood ε of the moving point. It is easy to see that this approach reduces to the Galerkin method for $\varepsilon \equiv 0$.

We denote the corrections to equations (4) by $\mathbf{B}=\{B_1, \ldots, B_n\}$:

$$\mathbf{B}=\dot{\mathbf{A}}-\mathbf{f}(\mathbf{A});$$

then the function L may be written in the form

$$L(\dot{\mathbf{A}},\mathbf{A}) \equiv (N,N)=C(\mathbf{A})+\sum_{i=1}^{n} \alpha_i B_i^2, \tag{7}$$

where $\alpha_1, \ldots, \alpha_n$ are positive coefficients, and the function $C(\mathbf{A}) \geqslant 0$ reflects the inaccuracy of approximation (4).

Without loss of generality one can assume that $\alpha_1=\ldots=\alpha_n=1$. Indeed, if this condition is not met, then one can take the trial functions in (2) in the form $\widetilde{H}_i=H_i/\sqrt{\alpha_i}$ $(i=1, \ldots, n)$, which brings Eq. (7) to the required form.

It is helpful to separate the problem of minimizing the functional S into two separate problems: that of minimizing the functional $S_1=\int_{t_0-\varepsilon}^{t_0}(N,N)dt$ with the first end free and the second end fixed; and the problem of minimizing the function $S_2=\int_{t_0}^{t_0+\varepsilon}(N,N)dt$, with the first end fixed and the second end free. Under the condition that the parameter ε is small, the solution of such problems[5] leads to systems of ODEs:

$$\frac{d}{dt}A_i=f_i(\mathbf{A})\pm\frac{\varepsilon}{2}\frac{\partial C(\mathbf{A})}{\partial A_i}-\frac{\varepsilon^2}{4}\frac{\partial}{\partial A_i} \times\left(\sum_{j=1}^{n}\frac{\partial C}{\partial A_j}f_j(\mathbf{A})\right)+O(\varepsilon^3), \quad i=1,\ldots,n, \tag{8}$$

where the plus and minus signs correspond to the first and second problems, respectively.

The difference in signs arises because we have broken the invariance with respect to the operation $t \to -t$ by minimizing the functionals on asymmetric time intervals $[t_0, t_0+\varepsilon]$ and $[t_0-\varepsilon, t_0]$. The solution of the original problem of minimizing the functional (6) on the symmetric interval $[t_0-\varepsilon, t_0+\varepsilon]$ is obtained by dropping all the terms in (8) which are odd in ε. We obtain the few-mode model

$$\frac{d}{dt}A_i=f_i(\mathbf{A})-\frac{\varepsilon^2}{4}\frac{\partial}{\partial A_i} \times\left(\sum_{j=1}^{n}\frac{\partial C}{\partial A_j}f_j(\mathbf{A})\right)+O(\varepsilon^4), \quad i=1,\ldots,n \tag{9}$$

with the initial conditions (5); this model differs from the model obtained by the Galerkin method and contains an undetermined parameter ε, the optimum value of which should correspond to a minimum of some figure of merit of the model, such as $I(\varepsilon)=\int_{T_0}^{T_0+T} L d\tau$ (the numbers T_0 and T are chosen appreciably longer than the characteristic times for system (9)). It is easy to see that one can to good accuracy take the optimum value of ε equal to the shortest characteristic time of the stability, as that time is what determines the radius of convergence of the expansion in ε in Eq. (8).

GENERALIZATION OF THE METHOD TO THE CASE OF A SYSTEM OF EQUATIONS

Let us consider a system of k partial differential equations

$$\frac{\partial u_j}{\partial t}=F\left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_m}, \mathbf{r}, u_1, \ldots, u_k\right), \quad j=1,\ldots,k \tag{10}$$

such that $u_1=u_2=\ldots=u_k=0$ is a particular solution of this system. We seek an approximate solution of the system by approximating each function u_j $(j=1, \ldots, k)$ by means of n_j trial functions:

$$u_j=\widetilde{u}_j=\sum_{i=d_j+1}^{d_j+1} A_i(t)H_i(\mathbf{r}), \quad j=1,\ldots,k, \tag{11}$$

where $d_j=\Sigma_{i=1}^{j} n_i$ and, hence, $d_k \equiv n$ is the total number of trial functions.

Let us now introduce the residual for each equation of system (10):

$$N_j(\dot{\mathbf{A}},\mathbf{A},\mathbf{r})=\frac{\partial \widetilde{u}_j}{\partial t}-F\left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}, \mathbf{r}, \widetilde{u}_1, \ldots, \widetilde{u}_k\right),$$

$$j=1,\ldots,k.$$

As before, the requirement that the residual be orthogonal to the trial functions yields a system of n ODEs (4) for the amplitudes $A_1, \ldots, A_n$. Further, in distinction with Eq. (6), here we we take the integrand L in the form

$$L(\dot{\mathbf{A}},\mathbf{A})=\sum_{j=1}^{k} (N_j,N_j)\beta_j, \tag{12}$$

where $\beta_1, \ldots, \beta_k$ are non-negative coefficients which determine the contribution of the residuals of the various equations to the function L.

The values of these coefficients will be chosen below. Each of the scalar products in (12) can be represented in the form

$$(N_j, N_j) + C_j(\mathbf{A}) + \sum_{i=d_j+1}^{d_j+1} \gamma_i B_i^2,$$

where, as before, $B_1, \ldots, B_n$ are corrections to equations (4).

From the fact that equation (12) is satisfied by $u_1 = u_2 = \ldots = u_k = 0$, it is clear that $\mathbf{A} = 0$ will be a stationary solution of the system of equations (4). We now require that the values of the coefficients $\beta_1, \ldots, \beta_k$ be such that near this point in the phase space of the system (4) the distribution of the various terms in (12) will be equivalent. Since the corrections to equations (4) in the proposed method are substantially nonlinear, it is sufficient to use the Galerkin approximation (4) to determine the values of the coefficients. We rewrite this system in the form

$$\dot{\mathbf{A}} = D\mathbf{A} + O(A^2), \tag{13}$$

where D is a square matrix with the elements

$$D_{ij} \equiv \left.\frac{\partial f_i}{\partial A_j}\right|_{\mathbf{A}=0}.$$

In the general case, system (13) can be reduced to the following form with the aid of a linear coordinate transformation $\mathbf{A} = Va$, where V is a square matrix having the property $V^{-1}DV = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$:

$$\frac{d}{dt}\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & & \\ \vdots & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + O(a^2).$$

Without loss of generality, it can be assumed that $\mathrm{Re}(\lambda_1) \geqslant \mathrm{Re}(\lambda_2) \geqslant \ldots \geqslant \mathrm{Re}(\lambda_n)$.

It is clear that the behavior of the dynamical system and, hence, the value of the terms in Eq. (12) will be determined by the fastest amplitudes a, i.e., by the amplitudes with the largest real part of the index λ.

There are two most important cases: the first is when λ_1 is real, and the second is when λ_1 and λ_2 are a complex-conjugate pair (we assume that these eigenvalues are nondegenerate).

In the first case, we rewrite expression (12) for L in the variables $a_1, \ldots, a_n$; then, assuming that $a_2 = a_3 = \ldots = a_n = 0$, and keeping only the terms quadratic in a_1, we get

$$L \approx \left(\frac{d}{dt}a_1 - \lambda_1 a_1\right)^2 \sum_{i=1}^{n} \gamma_i V_{1i}^2,$$

where V_{1i} $(i = 1, \ldots, n)$ are the elements of the first column of the matrix V.

We see immediately that for the coefficients $\beta_1, \ldots \beta_k$ we should set

$$\beta_j = \left(\sum_{i=d_j+1}^{d_j+1} \gamma_i V_{1i}^2\right)^{-1}, \quad j = 1, \ldots, k.$$

Of course, all this is valid if the sums on the right-hand sides are nonzero. The weight coefficients β for which this is not the case are determined in an analogous way by comparing the behavior of the slower amplitudes (i.e., those with smaller indices λ).

In the second case the equivalence of the terms in Eq. (12) will be understood to mean equivalence of their average values over a time $T = 2\pi/\mathrm{Im}(\lambda_1)$. Rewriting the expression for L in the variables $a_1, \ldots, a_n$, then assuming that $a_3 = \ldots = a_n = 0$, and keeping only the terms quadratic in a_1 and a_2, we obtain

$$\frac{1}{T}\int_t^{t+T} L\,dt \approx 2\left(\frac{d}{dt}\sqrt{a_1 a_2} - \mathrm{Re}(\lambda_1)\sqrt{a_1 a_2}\right)^2 \sum_{i=1}^{n} \gamma_i V_{1i} V_{2i},$$

where V_{1i} and V_{2i} $(i = 1, \ldots, n)$ are the elements of the first and second columns of the matrix V.

We must therefore set

$$\beta_j = \left(\sum_{i=d_j+1}^{d_j+1} \gamma_i V_{1i} V_{2i}\right)^{-1}, \quad j = 1, \ldots, k.$$

We have thus found the values of the coefficients $\beta_1, \ldots, \beta_k$. If we now rewrite the function L in the form (7), where we will have

$$C(\mathbf{A}) = \sum_{j=1}^{k} \beta_j C_j, \quad \alpha_i = \beta_j \gamma_i, \quad i = 1, \ldots, n,$$

and the subscript j is determined by the inequalities

$$j = \begin{cases} 1, & 0 < i \leqslant d_1, \\ 2, & d_1 < i \leqslant d_2, \\ \vdots & \vdots \\ k, & d_{k-1} < i \leqslant n, \end{cases}$$

we will again get the problem of optimizing the functional (6), and from there on the problem of constructing the system of ODEs can be treated in analogy with the previous Section.

EXAMPLE OF THE USE OF THE METHOD

In this Section we will consider as the initial system of partial differential equations the Saltzman system, which arises in a number of problems in the description of the initial stages of the laminar-to-turbulent transition (the onset of Bénard cells,[6] the stratification of a current-carrying liquid-metal conductor,[7,8] etc.):

$$\frac{\partial \nabla^2 \psi}{\partial t} = \frac{\partial \nabla^2 \psi}{\partial x}\frac{\partial \psi}{\partial z} - \frac{\partial \nabla^2 \psi}{\partial z}\frac{\partial \psi}{\partial x} + \alpha\frac{\partial \theta}{\partial x} + \gamma \nabla^4 \psi, \tag{14}$$

$$\frac{\partial \theta}{\partial t} = \frac{\partial \theta}{\partial x}\frac{\partial \psi}{\partial z} - \frac{\partial \theta}{\partial z}\frac{\partial \psi}{\partial x} + \beta\frac{\partial \psi}{\partial x} + \zeta \nabla^2 \theta, \tag{15}$$

where α, β, γ, and ζ are positive coefficients: $\psi(x,z,t)$, $\theta(x,z,t)$ are some scalar fields satisfying the boundary conditions

$$\theta|_{z=0} = \theta|_{z=z_0} = \psi|_{z=0} = \psi|_{z=z_0} = \nabla^2\psi|_{z=0} = \nabla^2\psi|_{z=z_0} = 0.$$

For the problem of two-dimensional thermal convection in a plane layer of liquid of height z_0 as a result of its heating from below, the coefficient γ is the viscosity, ζ is the thermal diffusivity, α is the product of the acceleration of gravity

and the coefficient of thermal expansion, and β is the ratio of the difference between the temperatures of the lower and upper layers of liquid to the height z_0. In this problem ψ is the stream function of the liquid and θ is the deviation of the temperature from the equilibrium (linear in z) profile.

For the problem of stratification of a liquid conductor, the coefficient γ is the viscosity, ζ is the magnetic viscosity, and α and β are proportional to the total electrical current through the conductor. As before, ψ is the stream function, but now θ is the perturbation of the equilibrium profile of the magnetic field distribution over the cross section of the conductor.[7,8]

This problem is attracting interest because it is the underlying basis of one of the best-known few-mode models that have been proposed for the laminar-to-turbulent transition, viz., the Lorenz model[9]

$$\dot{X}=\sigma Y-\sigma X, \tag{16}$$

$$\dot{Y}=rX-XZ-Y, \tag{17}$$

$$\dot{Z}=XY-bZ, \tag{18}$$

where $\sigma=\gamma/\zeta$, $b=8/3$, and $r=4\alpha\beta a^4(27\gamma\zeta)^{-1}$ (here $a=z_0/\pi$), and the a dot denotes differentiation with respect to the dimensionless time $\tau=3a^{-2}\zeta t/2$.

This model is obtained by substituting

$$\psi=3\zeta X(t)\sin\left(\frac{z}{a}\right)\sin\left(\frac{x}{\sqrt{2a}}\right), \tag{19}$$

$$\theta=\frac{\beta a}{r}\left(\sqrt{2}Y(t)\cos\left(\frac{z}{a}\right)\sin\left(\frac{x}{\sqrt{2a}}\right)-Z(t)\sin\left(\frac{2z}{a}\right)\right) \tag{20}$$

into the Saltzman system and then eliminating the modes that do not appear here by means of the standard procedure.[3]

It is known from Eq. (6) that as the number of modes in the Galerkin approximation increases near the threshold of instability, these three modes give the largest contribution. It is also clear that as the number of modes n increases, the number of order parameters remains unchanged,[3] and the remaining modes will be adiabatically tuned to them. Then for the minimum relaxation time $\Delta\tau$ at sufficiently large n we have $\Delta\tau\simeq n^{-2}$. For example, it is clear from Ref. 4 that the presence of such rapidly damped modes, which are not important in a qualitative analysis of the character of the phase transition, does, however, force one to choose the step for the numerical integration much smaller than $\Delta\tau$, on account of the stiffness of such systems. Thus analysis of the laminar-to-turbulent transition provides the clearest example of the necessary of increasing the accuracy of the model not through an enlargement of the basis but by using more-complicated expressions on the right-hand sides of the equations of the few-mode model.

Let us now use the method proposed in this paper to construct a modification of the Lorenz model. We take the function L in the form

$$L(\dot{X},\dot{Y},\dot{Z},X,Y,Z)=\beta_1\int_V N_1^2dV+\beta_2\int_V N_2^2dV,$$

where N_1 and N_2 are the residuals for Eqs. (14) and (15), respectively, and the positive coefficients β_1 and β_2 are chosen in accordance with the previous Section, such that at small X,Y,Z the contributions of the residuals of these equations to the function L will be equivalent.

For example, in view of the fact that the dynamics of the system will be governed by the fastest mode, with index

$$\lambda=\frac{-(\sigma+1)+\sqrt{(\sigma-1)^2+4r\sigma}}{2},$$

we obtain

$$L(\dot{X},\dot{Y},\dot{Z},X,Y,Z)=(\dot{X}-\sigma Y+\sigma X)^2/c(\sigma,r)+(\dot{Y}-rX+XZ+Y)^2+(\dot{Z}-XY+bZ)^2+X^2Z^2,$$

where $c=\sigma^2/(\sigma+\lambda)^2$.

At large values of the control parameter r the function $c(\sigma,r)$ is given approximately by $c\approx\sigma/r$.

Using Eq. (9), we obtain to terms of second order in ε

$$\dot{X}=\sigma(Y-X)-\varepsilon^2(YZ(c\sigma Z+(2c+1)X^2)/2-XZ^2(c\sigma+cb)), \tag{21}$$

$$\dot{Y}=rX-XZ-Y-\varepsilon^2XZ(X^2+c\sigma Z)/2, \tag{22}$$

$$\dot{Z}=XY-bZ-\varepsilon^2(YX(X^2+2\sigma Z)/2-ZX^2(\sigma+b)). \tag{23}$$

Thus for the Saltzman system the Lorenz substitution (19), (20) yields a model which is different from the known model (16)–(18). It contains an additional parameter ε, which determines the time interval over which the minimization of the functional $S=\int Ld\tau$ is carried out. As we have said above, the optimal value of this parameter corresponds to the minimum of the integral $I(\varepsilon)+\int_{T_0}^{T_0+T}ld\tau$. It is easy to see that below the boundary for loss of stability of the trivial stationary solution of the system (i.e., for $r\lesssim 1$) one must set $\varepsilon=0$, and this model goes over to the Lorenz model, i.e., in this particular case the proposed method gives the same model as does the Galerkin method.

In the case $r>1$ it is clear that as ε is increased from zero the value of the integral initially decreases (i.e., our model is more exact than the Lorenz model), reaches a minimum at some value of ε, and then begins to increase. This is because at large ε it is insufficient to stop at the terms quadratic in ε in Eq. (9).

For example, for the model (21)–(23) with the parameters $\sigma=10$ and $r=28$, the function $I(\varepsilon)$ reaches a minimum at $\varepsilon=0.035$. Its value at that point, obtained as a result of a numerical experiment, is $0.77I(0)$, i.e., it is substantially less than in the Lorenz model. Interestingly, the case under discussion corresponds to a strange attractor of the Lorenz type.

In general, at large supercriticalities (i.e., for $\sigma\ll r$ and $b\ll r$) the shortest characteristic time of the system is $1/r$, and one can take $\varepsilon=1/r$. For example, for $r=28$ one will have $\varepsilon\approx 0.036$, in agreement with the numerical integration results reported above.

Of course the use of the model (21)–(23) instead of (16)–(18) gives significant corrections to the value of the threshold of instability of the nontrivial stationary states of

the initial system (14), (15). For example, whereas in the Lorenz model all the stationary solutions lose stability and the Lorenz attractor with $r \approx 24.74$ becomes the only stable limit set, for the model constructed in this paper the dynamics of the system will become chaotic by the point $r = 19$, $\varepsilon = 0.045$, for example, and the relation $I(\varepsilon) = 0.78I(0)$ holds.

CONCLUSION

In this paper a method of constructing few-mode models is proposed which generalizes the Galerkin method; this method was used to construct a modification of the Lorenz model for a Saltzman system, and it was shown that this model gives a significantly smaller residual than the Lorenz model. It is hoped that the simplicity of analytical calculations in the construction of the model, which, however, leads to a substantial improvement in the accuracy of approximation in comparison with the Galerkin method, may prove useful for a qualitative analysis of the character of the loss of stability of stationary states and in the construction of models for the laminar-to-turbulent transition in nonlinear systems with distributed parameters.

The author is grateful to N. B. Volkov and A. M. Iskol'dskiĭ for some stimulating discussions and to O. V. Zubareva for assistance in doing the calculations.

This study was done as part of Project No. 94-02-06654-a.

[1] D. Ruelle and F. Takens, Commun. Math. Phys. **20**, 167 (1971).
[2] C. A. J. Fletcher, *Computational Galerkin Methods*, Springer-Verlag, New York, 1984) [Russ. transl., Mir, Moscow, 1988].
[3] H. Haken, *Synergetics: an Introduction*, Springer-Verlag, Berlin (1977) [Russ. transl., Mir, Moscow (1980)].
[4] K. Dekker and J. G. Verwer, *Stability of Runge–Kutta Methods for Stiff Nonlinear Differential Equations*, North-Holland, New York (1984) [Russ. transl., Mir, Moscow (1988)].
[5] N. M. Zubarev, Pis'ma Zh. Tekh. Fiz. **21**(15), 83 (1995) [Tech. Phys. Lett. **21**, 623 (1995)].
[6] B. Saltzman, J. Atmos. Sci. **19**, 329 (1962).
[7] N. B. Volkov and A. M. Iskol'dskiĭ, JETP Lett. **51**, 634 (1990).
[8] N. B. Volkov and N. M. Zubarev, Zh. Éksp. Teor. Fiz. **107**, 1868 (1995) [JETP **80**, 1037 (1995)].
[9] E. N. Lorenz, J. Atmos. Sci. **20**, 130 (1963).

Translated by Steve Torstveit

Moving inhomogeneity in a magnetohydrodynamic medium

A. A. Aleksandrova and Yu. N. Aleksandrov

Kharkov Military University, 310043 Kharkov, Ukraine;
Kharkov Technical University of Radio Engineering, 310726 Kharkov, Ukraine
(Submitted November 2, 1995; resubmitted March 4, 1996)
Zh. Tekh. Fiz. **67**, 6–11 (May 1997)

The integral equations of ideal magnetohydrodynamics are derived by the Green's function method, taking into account the velocity of the medium inside the inhomogeneity before and after its disturbance by the incident field. The extinction principle of of magnetohydrodynamics is demonstrated for a moving half space. A comparative analysis is made with the results of an analogous problem in which only the velocity of the interface between the two media is taken into account. © *1997 American Institute of Physics.* [S1063-7842(97)00205-5]

As a rule, in studying the propagation of small disturbances in a conducting medium in a uniform static magnetic field, it is assumed that $\mathbf{B}=\mathbf{B}_0+\mathbf{b}$, $\rho=\rho_0+\tilde{\rho}$, and $P=P_0+\tilde{p}$, i.e., that the total magnetic field $\mathbf{B}$ is a superposition of the unperturbed field $\mathbf{B}_0$, which is specified and produced by external currents (and analogously for the unperturbed density ρ_0 and pressure P_0), and an induced field $\mathbf{b}$ $(\tilde{\rho},\tilde{p})$, which is produced by currents caused by the disturbance, i.e., which is due to wave motion. The assumption of small amplitudes implies that $b\ll B_0$, $\tilde{\rho}\ll\rho_0$, and $\tilde{p}\ll P_0$. Also of the same order of smallness is the velocity $\mathbf{u}$, but which has an equilibrium value of zero.

For a number of problems, particularly in stability analysis, one is interested in the case of nonzero initial velocity of the medium. In the present paper this will mean taking into account a velocity U_0 of the medium of the inhomogeneity prior to its disturbance by the incident field. Here it is assumed that the inhomogeneity is located in an unbounded magnetohydrodynamic (MHD) medium characterized by the parameters $\mathbf{B}_1$ (the unperturbed magnetic field), V_{A1} (the Alfvén velocity), V_{S1} (the speed of sound), and ρ_2 (the density of the medium).

We will also assume that there is a certain inhomogeneity (a geometrically uniform region), which is characterized by the parameters $\mathbf{B}_2$, V_{A2}, V_{S2}, and ρ_2 and has a volume $V(t)$ and a boundary $S(t)$ that in general depend on time. We will investigate how the deviation $\mathbf{u}$ from the equilibrium velocity U_0 is affected by the velocity of the medium in the inhomogeneity, the velocity of its boundaries, and the rate of deformation of the boundary of the inhomogeneity under the influence of an incident disturbance (the last being a mutual effect).

The approach used in this paper is to generalize further the method of integral equations in MHD,[1] with the introduction of discontinuous functions describing the medium both inside and outside the inhomogeneity. Let us briefly consider the derivation of the integral equations for the given case of boundary-value problems under general assumptions about the parameters of the medium, the shape of the inhomogeneity, the time dependence of the properties of the medium inside the inhomogeneity, and the motion of it boundary. However, we will make the simplifying assumption that the disturbance is adiabatically turned on at infinity, i.e., the time dependence of the problem begins in the infinitely remote past. As we shall show, the form of the integral equation depends on the form of the background medium surrounding the inhomogeneity, which we will call the exterior medium in distinction with the internal medium of the inhomogeneity (the interior medium).

As we know, the equations of linearized ideal MHD for the deviations $\mathbf{u}$ and $\mathbf{b}$ of the velocity and magnetic field from the equilibrium values in an exterior medium with parameters described by continuous functions are of the form [note: the square brackets denote the vector (cross) product]

$$V_{S1}^2\,\mathrm{grad\ div}\,\mathbf{u}-\frac{\partial^2\mathbf{u}}{\partial t^2}+\left[\mathrm{curl}\frac{\partial\mathbf{b}}{\partial t},\frac{\mathbf{B}_1}{4\pi\rho_1}\right]=0,$$

$$\mathrm{curl}[\mathbf{u},\ \mathbf{B}_1]-\frac{\partial\mathbf{b}}{\partial t}=0,\quad \mathbf{r}\ni V(t). \tag{1}$$

The corresponding equations in the interior medium are

$$V_{S2}^2\,\mathrm{grad\ div}\mathbf{u}-\frac{\partial^2\mathbf{u}}{\partial t^2}+\mathrm{curl}\left[\frac{\partial\mathbf{b}}{\partial t},\frac{\mathbf{B}_2}{4\pi\rho_2}\right]+\mathrm{grad\ div}(\rho\mathbf{U}_0)=0,$$

$$\mathrm{curl}[\mathbf{u},\mathbf{B}_2]-\frac{\partial\mathbf{b}}{\partial t}+\mathrm{curl}[\mathbf{U}_0,\mathbf{b}]=0,\quad \mathbf{r}\in V(t). \tag{2}$$

On the surface of discontinuity $S(t)$ the field functions must satisfy boundary conditions which follow from the continuity of the mass flux $\{\rho u_n\}_S=0$, momentum $\{\pi_{in}\}_S=0$, energy $\{W_n\}_S=0$, tangential component of the electric field $\{\mathbf{E}_\tau\}_S=0$, and normal component of the magnetic field $\{\mathbf{B}_n\}_S=0$:

$$\{\rho U_z\}=0,\quad \{\rho U_xU_z-B_xB_z/4\pi\}=0,$$

$$\{\rho U_yU_z-B_yB_z/4\pi\}=0,$$

$$\{\rho U_z^2+p+(B_x^2+B_y^2)/8\pi\}=0,$$

$$\{U_xB_z-U_zB_x\}=0,\quad \{U_yB_z-U_zB_y\}=0,\quad \{U_z\}=0,$$

$$\left\{\rho U_z\left(\frac{U^2}{2}+\varepsilon\right)+U_zp+(B^2U_z-B_z(\mathbf{U}\cdot\mathbf{B}))/4\pi\right\}=0;$$

$$\mathbf{U}=\begin{cases}\mathbf{U}_0+\mathbf{u}, & \mathbf{r}\in V(t),\\ \mathbf{u}, & \mathbf{r}\ni V(t),\end{cases}\qquad \mathbf{B}=\begin{cases}\mathbf{B}_2+\mathbf{b}, & \mathbf{r}\in V(t),\\ \mathbf{B}_1+\mathbf{b}, & \mathbf{r}\ni V(t),\end{cases} \tag{3}$$

 1063-7842/97/050460-06$10.00

where $\{\mathbf{a}\}$ is the jump in the quantity $\mathbf{a}$ on passage through the surface $S(t)$.

For specificity, we take the $z=0$ plane to be the plane tangent to the surface of discontinuity, i.e., the z axis is directed along the normal to the surface of discontinuity.

When considering the boundary conditions (3) in the laboratory reference frame, in which the surface of discontinuity is perpendicular to the z axis and moves along that axis with a velocity u_S, one must replace u_z by u_z-u_S everywhere in Eq. (3) (see Ref. 2 for details).

In seeking solutions in integral form in the class of piecewise-smooth functions, with allowance for certain additional conditions, the discontinuous solutions are, generally speaking, obtained automatically from the statement of the problem. For this, with the aid of the characteristic function

$$\chi(\mathbf{r},t)=\begin{cases}1, & \mathbf{r}\in V(t),\\ 0, & \mathbf{r}\ni V(t),\end{cases}$$

Eqs. (1) are propagated to the entire space under consideration in the following manner:

$$(V_{A1}^2+V_{S1}^2)\text{grad div }\mathbf{u}-\frac{\partial^2\mathbf{u}}{\partial t^2}-V_{A1}^2\mathbf{s}_1(\mathbf{s}_1\cdot\text{grad div }\mathbf{u})$$
$$-V_{A1}^2[\mathbf{s}_1,(\mathbf{s}_1\cdot\nabla)\text{curl }\mathbf{u}]=\mathbf{W}, \tag{4}$$

where

$$\mathbf{W}=\chi(V_{S1}^2-V_{S2}^2)\text{grad div }\mathbf{u}-\chi\left[\frac{V_{A1}^2}{B_1}\mathbf{s}_1-\frac{V_{A2}^2}{B_2}\mathbf{s}_2,\text{curl}\frac{\partial\mathbf{b}}{\partial t}\right]$$
$$+\chi\frac{V_{A1}^2}{B_1}\left[\mathbf{s}_1,\text{curlcurl}\left[\mathbf{s}_1-\frac{B_2}{B_1}\mathbf{s}_2,\mathbf{u}\right]\right]$$
$$-\chi\frac{V_{S2}^2}{\rho_2}\text{grad}(\text{grad}\rho\cdot\mathbf{U}_0)+\chi\left[\frac{\mathbf{B}_1}{4\pi\rho_1},\text{curlcurl}[\mathbf{U}_0,\mathbf{b}]\right]$$
$$+\left[\frac{\mathbf{B}_1}{4\pi\rho_1},[\mathbf{n},\{\text{curl}[\mathbf{u},\mathbf{B}_1]\}_S]\right]\delta S(t)$$
$$+\left[\frac{\mathbf{B}_1}{4\pi\rho_1},[\mathbf{n},\{\text{curl}[\mathbf{U}_0,\mathbf{b}]\}_S]\right]\delta S(t).$$

In Eq. (4) the time derivative is a derivative in the generalized sense of the word,[3] i.e.,

$$\frac{\partial\mathbf{u}}{\partial t}=\left(\frac{\partial\mathbf{u}}{\partial t}\right)+\delta(t)\langle\mathbf{u}(0)\rangle,$$

where $(\partial\mathbf{u}/\partial t)$ is the ordinary derivative, and $\langle\mathbf{u}(0)\rangle=\mathbf{u}(+0)-\mathbf{u}(-0)=0$, since the parameters of the medium do not undergo jumps at a finite time.

The fact that all the quantities suffer a discontinuity at the boundary $S(t)$ of the region $V(t)$ is taken into account by replacing the classical derivatives by the generalized derivatives according to the rules[3]

$$\text{curl }\mathbf{a}=(\text{curl }\mathbf{a})+\mathbf{n}\times\{\mathbf{a}\}_S\delta S(t),$$

$$\text{div }\mathbf{a}=(\text{div }\mathbf{a})+\mathbf{n}\cdot\{\mathbf{a}\}_S\delta S(t),$$

$$\text{grad }\varphi=(\text{grad}\varphi)+\mathbf{n}\{\varphi\}_S\delta S(t),$$

$$\frac{\partial\mathbf{a}}{\partial t}=\left(\frac{\partial\mathbf{a}}{\partial t}\right)-\{\mathbf{a}\}_S u_n\delta S(t).$$

Here the curly brackets denote the classical derivative, $\{\mathbf{a}\}_S$ is the jump in the function $\mathbf{a}$ on passage through the surface $S(t)$ from the outside and satisfies conditions (3), and $\delta S(t)$ is the surface δ function. Equation (4) describes the field throughout the entire space under study, since the boundary conditions at the surfaces of discontinuity of the field functions have already been taken into account. Here surface terms arise which are obviously due to surface currents. Requiring that the right-hand side of Eq. (4) be finite over the spatial and temporal coordinates, we write the general solution of equation (4) in the form of a convolution:

$$\mathbf{u}=\mathbf{u}_0+\hat{G}*\mathbf{W},$$

where $\mathbf{u}_0$ is the general solution of the corresponding homogeneous equation, i.e., the incident field, and $\hat{G}$ is the Green's function, which is the fundamental solution of equation (4), i.e., it satisfies the equation

$$(V_{A1}^2+V_{S1}^2)\text{grad div }\hat{G}-\frac{\partial^2 G}{\partial t^2}-V_{A1}^2\mathbf{s}_1(\mathbf{s}_1\cdot\text{grad div }\hat{G})$$
$$-V_{A1}^2[\mathbf{s}_1,(\mathbf{s}_1\cdot\nabla)\text{curl}\hat{G}]=\varepsilon\,\delta(t-t')\,\delta(\mathbf{r}-\mathbf{r}').$$

This function, which has the form $\hat{G}(\mathbf{r}-\mathbf{r}',t-t')=\hat{G}\cdot I$, is found and described in detail in Ref. 1; here $\hat{G}$ is a differential operator written in the basis $\langle\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\rangle$, where $\mathbf{e}_2=\mathbf{s}_1=\mathbf{B}_1/B_1$, and $I=I(\mathbf{r}-\mathbf{r}',t-t')$ is written in the form of a Fourier–Laplace integral. Knowing the fundamental solution $\hat{G}(\mathbf{r}-\mathbf{r}',t-t')$, and taking into account the properties of the convolution and the presence of the characteristic function $\chi(\mathbf{r},t)$, we can rewrite Eq. (4) as an integral equation of MHD:

$$\mathbf{u}(\mathbf{r},t)=\mathbf{u}_0(\mathbf{r},t)+(V_{S1}^2-V_{S2}^2)\hat{G}\text{grad div}\int_{-\infty}^{\infty}dt'\int_{V(t')}\mathbf{u}Id\mathbf{r}'$$
$$+V_{A1}^2\hat{G}\left[\mathbf{s}_1\text{curlcurl}\left[\mathbf{s}_1-\frac{B_2}{B_1}\mathbf{s}_2,\int_{-\infty}^{\infty}dt'\int_{V(t')}\mathbf{u}Id\mathbf{r}'\right]\right]$$
$$-\frac{\hat{G}}{B_1}\left[V_{A1}^2\mathbf{s}_1-\frac{B_1}{B_2}\mathbf{s}_2,\ \text{curl}\frac{\partial}{\partial t}\int_{-\infty}^{\infty}dt'\int_{V(t')}\mathbf{b}Id\mathbf{r}'\right]$$
$$-\hat{G}\frac{V_{S2}^2}{\rho_2}\text{grad}\left(\mathbf{U}_0\cdot\text{grad}\int_{-\infty}^{\infty}dt'\int_{V(t')}\rho Id\mathbf{r}'\right)$$
$$-\hat{G}\left[\frac{\mathbf{B}_1}{4\pi\rho_1},\text{curl curl}\left[\mathbf{U}_0,\int_{-\infty}^{\infty}dt'\int_{V(t')}\mathbf{b}Id\mathbf{r}'\right]\right]$$
$$+\int_{-\infty}^{\infty}dt'\int_{S(t')}\hat{G}(\mathbf{r}-\mathbf{r}',t-t')\left[\frac{\mathbf{B}_1}{4\pi\rho_1},\right.$$
$$\left.\times[\mathbf{n},\{\text{curl}[\mathbf{u}^{ex}-\mathbf{u}^{in},\mathbf{B}_2]+\text{curl}[\mathbf{U}_0,\mathbf{b}^{ex}-\mathbf{b}^{in}]\}]\right]dS, \tag{5}$$

where the exterior components of the fields $\mathbf{u}^{ex}$ and $\mathbf{b}^{ex}$ at the boundary $S(t)$ are expressed in terms of the interior components $\mathbf{u}^{in}$ and $\mathbf{b}^{in}$ with the aid of Eq. (3).

It must be emphasized that the method used to solve the boundary-value problems on the basis of the integral equations (5) presupposes not only that the initial differential equations (1) and (2) are represented in integral form, which can always be done by constructing the appropriate Green's function, but also the use of an additional assertion — the extinction principle. The physical meaning of this principle is known from electrodynamics, but here it has a broader content.[1] This principle was used to develop an algorithm for solving diffraction problems. Outside the regions of inhomogeneities Eqs. (5) can be written in the form of integrals of the known internal fields of the inhomogeneity, i.e., the total field outside an inhomogeneity (the diffracted field) is written as a sum of an incident and a scattered wave. The scattered field is expressed in terms of MHD potentials of the type

$$\begin{Bmatrix} \mathbf{\Pi}_u \\ \mathbf{\Pi}_b \end{Bmatrix} = \int_V \begin{Bmatrix} \mathbf{u}(\mathbf{r}',t') \\ \mathbf{b}(\mathbf{r}',t') \end{Bmatrix} I(\mathbf{r}-\mathbf{r}',t-t')d\mathbf{r}'.$$

In the case $\mathbf{r} \in V$ the field functions on the left-hand side of Eq. (5) are treated as the total internal fields at a point $\mathbf{r}$ inside volume V. The integral equation formalism permits one to decouple the initial group of equations into equations for finding the propagation constants in the inhomogeneity and the equations for the amplitudes of the waves in it.

As an example, let us demonstrate the extinction principle for the diffraction of MHD waves on a moving half space. As we solve the problem we will make reference to the analogous problem of the diffraction of MHD waves on a uniformly moving plane boundary between two MHD media.[4]

Suppose there is an MHD inhomogeneity characterized by a density ρ_2, a static magnetic field B_2, and sound and Alfvén velocities V_{S2} and V_{A2} and located in a space xyz bounded by a plane parallel to the xy plane and moving uniformly with a velocity U_0 perpendicular to the plane of the interface. The medium outside the inhomogeneity has parameters β_1, $\mathbf{B}_1$, V_{S1}, and V_{A1}; the vectors $\mathbf{B}_i$ lie in the yz plane. Incident on the inhomogeneity is a packet of MHD waves: an Alfvén wave, having components u_x and b_x, and accelerated and slowed magnetosonic waves $u_2^{\pm}$, $u_3^{\pm}$, $b_2^{\pm}$, $b_3^{\pm}$ of the form

$$u_{0x}(\mathbf{r},t) = u_{0x}\exp(-i\mathbf{k}_{A0}\cdot\mathbf{r}+i\omega_0 t),$$

$$\mathbf{k}_{A0} = \omega_0\mathbf{n}_{A0}/V_{A1}(\mathbf{n}_{A0}\cdot\mathbf{s}_1),$$

$$u_{0i}(\mathbf{r},t) = u_{0i}^{+}\exp(-i\mathbf{k}_0^{+}\cdot\mathbf{r}+i\omega_0^{+}t) + u_{0i}^{-}\exp(-i\mathbf{k}_0^{-}\cdot\mathbf{r}+i\omega_0^{-}t),$$

where $i=2,3$; the wave vectors $\mathbf{k}_0^{\pm}$ satisfy the dispersion relation

$$\omega_0^4-\omega_0^2(V_{A1}^2+V_{S1}^2)k_0^2+V_{A1}^2V_{S1}^2k_0^2(\mathbf{k}_0\cdot\mathbf{s}_1)^2=0.$$

The components u_{0i} are considered in the basis $\langle \mathbf{e}_i,\mathbf{e}_2,\mathbf{e}_3\rangle$, where $\mathbf{e}_2=\mathbf{s}_1$, and the unit vector $\mathbf{e}_1$ is coincident with the unit vector $\mathbf{i}$. For brevity and convenience in what follows we introduce the following notation for the incident field:

$$\mathbf{u}_0(\mathbf{r},t)=\sum_{j=1}^{3}\mathbf{u}_0^j\exp(-i\mathbf{k}_0^j\cdot\mathbf{r}+i\omega_0^j t), \tag{6}$$

where $j=1$ corresponds to the Alfvén wave, $j=2$ to the accelerated magnetosonic wave, and $j=3$ to the slowed magnetosonic wave.

We choose as a trial solution for the transmitted field in the half space $z>0$ a superposition of the following waves:

$$\mathbf{u}(\mathbf{r},t)=\sum_{j=1}^{3}\mathbf{u}^j\exp(-i\mathbf{k}^j\cdot\mathbf{r}+i\omega^j t) \tag{7}$$

with as-yet unknown amplitudes $\mathbf{u}^j$, wave numbers $\mathbf{k}^j$, and frequencies ω^j. By virtue of the linearity of the initial differential equations and of the boundary conditions, the velocity u_S of the interfacial surface will obviously also be expressed as a superposition of waves:

$$u_S(\mathbf{r},t)=\sum_m [u_S^m\exp(i\mathbf{k}^m\cdot\mathbf{r}+i\omega^m t) + \tilde{u}_S^m\exp(-i\mathbf{k}_0^m\cdot\mathbf{r}+i\omega_0^m t)].$$

After substituting the trial solution (7) into the integrand of the volume integrals, we do the integration over the volume $V(t')$ ($\mathbf{r}\in V(t')$) and over time. As a result, we obtain an integral expression for the Alfvén wave

$$I_x=\int_{-\infty}^{\infty}dt'\int_{V(t')}u_x(\mathbf{r},'t')I(\mathbf{r}-\mathbf{r}',t-t')d\mathbf{r}'$$
$$=u_x\left\{\frac{\exp(-i\mathbf{k}_A\cdot\mathbf{r}+i\omega_A t)}{V_{A1}^2(\mathbf{k}_A\cdot\mathbf{s}_1)^2-\omega_A^2} + \frac{\exp\{-i[k_y y+z(\omega'/V_{A1}s_{1z}-k_y s_{1z}/s_{1z})]+i\omega' t\}}{2\omega'[V_{A1}(\mathbf{k}_A\cdot\mathbf{s}_1)-\omega']}\right\}, \tag{8}$$

where

$$\omega'=\frac{\omega_A-U_0(\mathbf{k}_A\cdot\mathbf{s}_1)/s_{1z}}{1-U_0/V_{A1}s_{1z}},$$

and, analogously, for the magnetosonic waves

$$I_{i=2,3}^{\ i}=\int_{-\infty}^{\infty}dt'\int_{V(t')}u_i(\mathbf{r}',t')I(\mathbf{r}-\mathbf{r}',t-t')d\mathbf{r}'$$
$$=u_i\left\{\frac{\exp(-i\mathbf{k}_s\cdot\mathbf{r}+i\omega_S t)}{\Delta(\omega,k_s)}+\frac{\exp(-i\mathbf{k}_0^{+}\cdot\mathbf{r}+i\omega_0^{+}t)}{\Omega(k_0^{+})(k_z-k_{0z}^{+})} + \frac{\exp(-i\mathbf{k}_0^{-}\cdot\mathbf{r}+i\omega_0^{-}t)}{\Omega(k_0^{-})(k_z-k_{0z}^{-})}\right\}, \tag{9}$$

where

$$\Delta(\omega,k_s)=\omega^4-(V_{A1}^2+V_{S1}^2)\omega^2k_S^2+V_{A1}^2V_{S1}^2k_S^2(\mathbf{k}_S\mathbf{s}_1)^2,$$

$$\Omega(k_0)=2k_{0z}\{(V_{A1}^2+V_{S1}^2)\omega_0^2 - V_{A1}^2V_{S1}^2k_0^2[\cos^2\vartheta+\cos\vartheta s_{1z}/n_{0z}]\}.$$

Accordingly, doing the surface integral gives

$$\tilde{\mathbf{I}}_{\overline{j=1,3}}^{j}=\int_{-\infty}^{\infty}dt'\int_{S(t')}\mathbf{u}^j(\mathbf{r}',t')I(\mathbf{r}-\mathbf{r}',t-t')dS = i\mathbf{u}^j\frac{\exp(-i\mathbf{k}_0^j\cdot\mathbf{r}+i\omega_0^jt)}{\Omega(k_{0x}^j,k_{0y}^j,k_{0z}^j)},$$

where

$$\Omega(k_{0x}^1,k_{0y}^1,k_{0z}^1)=2\omega_0^1V_{A1}s_{1z}[(\omega_0^1)^4-(V_{A1}^2+V_{S1}^2)\times(\omega_0^1)^2(k_0^1)^2+V_{A1}^2V_{S1}^2(k_0^1)^2(\mathbf{k}_0^1\cdot\mathbf{s}_1)^2],$$

$$\Omega_{j=2,3}(k_{0x}^j,k_{0y}^j,k_{0z}^j)=2[(\omega_0^j)^2-V_{A2}^2(\mathbf{k}_0^j\cdot\mathbf{s}_1)^2]\times\{(V_{A1}^2+V_{S1}^2)(\omega_0^j)^2k_{0z}^j+V_{A1}^2V_{S1}^2\times[k_{0z}^j-(k_{0z}^j)^2s_{1z}(\mathbf{k}_0^j\cdot\mathbf{s}_1)]\}.$$

Analyzing Eqs. (8) and (9), we can say that the integral terms describing the diffracted field decompose into a number of terms: terms of the type

$$\exp(-i\mathbf{k}_A\cdot\mathbf{r}+i\omega_At),\exp(-i\mathbf{k}^{\pm}\cdot\mathbf{r}+i\omega^{\pm}t),$$

describing the refracted waves, and terms of the type

$$\exp\{-i[k_yy+z(\omega'/V_{A1}s_{1z}-k_ys_{1y}/s_{1z})]+i\omega't\},$$

$$\exp(-i\mathbf{k}_0^{\pm}\cdot\underline{r}+i\omega_0^{\pm}t).$$

While the last exponential function describes the same wave as the incident magnetosonic field, the next-to-last exponential function coincides with the incident Alfvén wave if we set $\omega'=\omega_0$; then

$$\frac{\omega_0}{V_{A1}s_{1z}}-\frac{k_ys_{1y}}{s_{1z}}=k_{A0z},$$

and, owing to the translational symmetry of the problem, we have $k_y=k_{A0y}$. From the requirements imposed, we find the frequency of the refracted Alfvén wave:

$$\omega_A=\omega_0\frac{2(V_{A1}s_{1z}-U_0)V_{A2}^2(\mathbf{s}_2\cdot\mathbf{n}_A)^2}{V_{A1}\{2s_{1z}V_{A2}^2(\mathbf{s}_2\cdot\mathbf{n}_A)^2-U_0(\mathbf{n}_A\cdot s_1)\sqrt{(\mathbf{n}_A\cdot\mathbf{U}_0)^2+4V_{A2}^2(\mathbf{n}_A\cdot\mathbf{s}_2)^2-(\mathbf{n}_A\cdot\mathbf{U}_0)}\}} \tag{10}$$

and the law of refraction

$$(V_{A1}s_{12}-U_0)\frac{s_{1y}+s_{1z}\cot\alpha}{s_{1y}+s_{1z}\cot\beta}=\frac{2s_{1z}V_{A2}^2\dfrac{(s_{2y}+s_{2z}\cot\beta)^2}{(s_{1y}+s_{1z}\cot\beta)U_0\cot\beta}-U_0\left|\sqrt{1+4V_{A2}^2\,s_{2y}+s_{2z}\cot\beta}\right)^2}{U_0^2\cot^2\beta} - 1\sqrt{1+4V_{A2}^2\frac{(s_{2y}+s_{2z}\cot\beta)^2}{U_0^2\cot^2\beta}}-1 \tag{11}$$

where α and β are, respectively, the angles of incidence and refraction of the Alfvén wave.

Analogously, for the magnetosonic waves we have the frequency transformation

$$\omega^{\pm}=\omega_0^{\pm}\frac{(Q_1^{\pm}-U_0\cos\alpha_{pm})Q_2^{\pm}}{(Q_2^{\pm}-U_0\cos\beta_{pm})Q_1^{\pm}}, \tag{12}$$

and, in addition,

$$\frac{\omega^+-\omega_0^-}{\omega^--\omega_0^+}=\frac{k_z^+-k_{0z}^-}{k_z^--k_{0z}^+},$$

and the law of refraction is

$$(Q_1^{\pm}-U_0\cos\alpha_{\pm})\sin\beta_{\pm}=(Q_2^{\pm}-U_0\cos\beta_{\pm})\sin\alpha_{\pm}, \tag{13}$$

where $Q_{1,2}^{\pm}$ will be discussed below.

Thus after substitution of the integral terms I_x, I_i, $\tilde{I}_j$ ($i=2,3$; $j=1,3$) into the initial integral equation, we obtain an identity which is valid for all interior points of the region:

$$\sum_{j=1}^{3}\mathbf{u}\exp(-i\mathbf{k}^j\cdot\mathbf{r}+i\omega^jt) = \sum_{j=1}^{3}\mathbf{u}^j\exp(-i\mathbf{k}_0^j\cdot\mathbf{r}+i\omega_0^jt) + \sum_{j=1}^{3}\hat{A}_j\mathbf{u}^j\exp(-i\mathbf{k}^j\cdot\mathbf{r}+i\omega^jt) + \sum_{j=1}^{3}\hat{B}_j\mathbf{u}^j\exp(-i\mathbf{k}_0^j\cdot\mathbf{r}+i\omega_0^jt), \tag{14}$$

where $\hat{A}_j$ and $\hat{B}_j$ are the result of the operation of the differential operators on the corresponding exponential functions in the initial integral equation.

It should be noted that for clarity of notation the integrals I_x and I_i are evaluated under the condition $k_{0x}^1\neq k_{0x}^2\neq k_{0x}^3$, $k_{0y}^1\neq k_{0y}^2\neq k_{0y}^3$, i.e., we have ruled out the conversion of one type of MHD wave into a wave of another type; this is described mathematically in Ref. 5 and occurs at points where the phase velocities of two or more waves coincide in their transformation on a flat surface.

In the analysis of identity (14) an extinction principle comes into play, according to which the MHD field of dipoles can be represented as a sum of two groups of terms, one of which satisfies the field equation in the exterior medium, i.e., terms of that group have the character of an incident field. Thus the the incident wave is exactly extinguished at any point inside the medium as a result of interference of the field it creates with the field of the dipoles, and a new wave appears with a different propagation velocity (k^j). It follows that the terms of identity (14) which vary according to $\exp(-i\mathbf{k}^j\cdot r+i\omega^j t)$, form the relation

$$\mathbf{u}^j\exp(-i\mathbf{k}^j\cdot\mathbf{r}+i\omega^j t)=\hat{A}_j\mathbf{u}^j\exp(-i\mathbf{k}^j\cdot\mathbf{r}+i\omega^j t)$$

and cancel each other out, if it is assumed that the $\mathbf{k}^j$ satisfy the following dispersion relations:

$$\omega_A^2-\omega_A(\mathbf{k}_A U_0)-V_{A2}^2(\mathbf{s}_2\cdot\mathbf{k}_A)^2=0 \tag{15}$$

for the Alfvén waves and

$$(\omega^{\pm})^4-2(\omega^{\pm})^3(\mathbf{k}^{\pm}\cdot\mathbf{U}_0)-(\omega^{\pm})^2$$
$$\times\{(V_{S2}^2+V_{A2}^2)(k^{\pm})^2[1-(\mathbf{k}^{\pm}\cdot\mathbf{U}_0)-(\mathbf{k}^{\pm}\cdot\mathbf{U}_0)^2]\}$$
$$+V_{A2}^2V_{S2}^2(k^{\pm})^2(\mathbf{k}^{\pm}\cdot\mathbf{s}_2)^2=0 \tag{16}$$

for the magnetosonic waves. The phase velocities $Q_{1,2}^{\pm}=\omega^{+}/k_{1,2}^{+}$, which were discussed above, are found from relation (16). It should be noted that the wave numbers of the forward and backward waves have different absolute values both in the case of the Alfvén wave and the magnetosonic waves. For the Alfvén wave they are given by

$$k_A=\omega_A\frac{-(\mathbf{n}_A\cdot\mathbf{U}_0)\pm\sqrt{(\mathbf{n}_A\cdot\mathbf{U}_0)^2+4V_{A2}^2(\mathbf{s}_2\cdot\mathbf{n}_A)^2}}{2V_{A2}^2(\mathbf{s}_2\cdot\mathbf{n}_A)^2}.$$

The incident field is extinguished by the field of the secondary waves. Hence we obtain equations for finding the amplitudes of the secondary waves (these have not been written out for the sake of brevity)

$$\mathbf{u}_0^j(\mathbf{r},t)+\hat{B}_j\mathbf{u}^j\exp(-i\mathbf{k}_0^j\cdot\mathbf{r}+i\omega_{0j=\overline{1,3}}^j t)=0.$$

In the case when the second medium is initially at rest and only the interface is moving, the structure of the internal field will be the same, but the wave numbers in this case will satisfy the following equations:

$$\omega_A^2-V_{A2}^2(\mathbf{k}_A\cdot\mathbf{s}_2)^2=0$$

for the Alfvén waves, and

$$\omega^4-(V_{A2}^2+V_{S2}^2)\omega^2k^2+V_{A2}^2V_{S2}^2k^2(\mathbf{k}\cdot\mathbf{s}_2)^2=0$$

for the magnetosonic waves.

The laws of refraction and of the change of frequency will be of an absolutely different character. For example, the law of refraction for an Alfvén wave has the form

$$V_{A1}(s_{1y}+s_{1z}\cot\alpha)-U_0\cot\alpha=V_{A2}(s_{2y}+s_{2z}\cot\beta)-U_0\cot\beta$$

while the frequency change is given by the relation

$$\omega_A=\omega_0\frac{(V_{A1}s_{1z}-U_0)V_{A2}(\mathbf{s}_2\cdot\mathbf{n}_A)}{V_{A1}[V_{A2}s_{1z}(\mathbf{s}_2\cdot\mathbf{n}_A)-U_0(\mathbf{s}_1\cdot\mathbf{n}_A)]}.$$

In the case of magnetosonic waves the laws of refraction and of the frequency change are described formally by formulas (13) and (12), but $Q_i^{\pm}$ has a simpler form:

$$Q_i^{\pm}=\sqrt{\frac{1}{2}[(V_{Ai}^2+V_{Si}^2)\pm\sqrt{(V_{Ai}^2+V_{Si}^2)^2-4V_{Ai}^2V_{Si}^2\cos^2\vartheta_i}]}.$$

Thus in the boundary-value problem of MHD we observe a frequency change due to motion of the medium in the transformation of the waves; this is called the Doppler effect (relations (10) and (12)). However, unlike the case of acoustic and electrodynamics, where the frequency is affected solely by the velocity of the source, in MHD the frequency change is influenced not only by the velocity of the medium but also (in the case of Alfvén waves, for example) by the Alfvén velocities V_{A1} and V_{A2} in the exterior and interior regions and by the direction of propagation of the Alfvén wave relative to the unperturbed magnetic fields $\mathbf{B}_1$ and $\mathbf{B}_2$. Thus, for example, in the propagation of an Alfvén wave along the field $\mathbf{s}_1$ we have

$$\omega_A=\frac{\omega_0(1-U_0/V_{A1}s_{1z})}{1-U_0/V_{Az}(\mathbf{n}_A\cdot\mathbf{s}_2)s_{1z}},$$

and perpendicular to it

$$\omega_A=\omega_0(1-U_0/V_{A1}s_{1z}).$$

Along the magnetic field $\mathbf{s}_2$ we have

$$\omega_A=\frac{\omega_0(1-U_0/V_{A1}s_{1z})}{1-U_0(\mathbf{n}_A\cdot\mathbf{s}_1)/V_{A2}s_{1z}}$$

and for $(\mathbf{n}_A\cdot\mathbf{s}_2)\to 0\Rightarrow\omega_{A\to 0}$.

It is important to note that the frequencies of the refracted magnetosonic waves are interrelated not only with the corresponding frequencies of the incident unperturbed waves but also with one another. In the case when the boundary moves with a velocity $U_0>Q_i^{\pm}/\cos\alpha_{\pm}$, the refracted waves (the accelerated and slowed waves) are not excited in the interior region, since the plane moves away faster than the unperturbed wave reaches it. If $Q_1^{-}/\cos\alpha_{-}<U_0<Q_1^{+}/\cos\alpha_{+}$, then we will have only the accelerated refracted wave in the interior medium.

If the unperturbed Alfvén wave is incident at an angle satisfying the relation

$$\tan\alpha_k=\frac{U_0-V_{A1}s_{1z}}{V_{A1}s_{1y}-V_{A2}s_{2y}},$$

then ''total internal reflection'' of the Alfvén waves will occur, and the corresponding angle α_k by analogy with electromagnetic waves can be called the critical angle of incidence.

The scattered field is found by a simple integration of the internal field, since for points outside the inhomogeneity ($\mathbf{r}\in V(t)$) the integrands do not contain singularities, and Eq. (5) becomes an equivalency relationship. Substituting the internal field found for the Alfvén component of the velocity into Eq. (5) and taking into account that $\mathbf{r}\ni V(t)$, we find the integrand outside the inhomogeneity:

$$I_x^{\mathrm{ref}} = \int_{-\infty}^{\infty} dt' \int_{V(t')} u_x(\mathbf{r}', t') I(\mathbf{r}-\mathbf{r}', t-t')\, d\mathbf{r}'$$

$$= u_x \frac{\exp\{i\omega'' t - ik_{0y} y + iz[\omega''/V_{A1} s_{1z} + k_{0y} s_{1y}/s_{1z}]\}}{2\omega''[V_{A1}(\mathbf{k}_A \cdot \mathbf{s}_1) + \omega'']},$$

where

$$\omega'' = \frac{\omega_A - U_0(\mathbf{k}_A \cdot \mathbf{s}_1)/s_{1z}}{1 + U_0/V_{A1} s_{1z}}.$$

The final reflected field of the Alfvén wave is

$$u_x^{\mathrm{ref}}(\mathbf{r}, t) = u_s^{\mathrm{ref}} \exp(-i\mathbf{k}_A^{\mathrm{ref}} \cdot \mathbf{r} + i\omega_A^{\mathrm{ref}}),$$

where

$$\omega_A^{\mathrm{ref}} = \omega_0 \frac{U_0 - V_{A1} s_{1z}}{U_0 + V_{A1} s_{1z}},$$

$$(\mathbf{k}_A^{\mathrm{ref}} \cdot \mathbf{r}) = k_{A0y} y - [\omega_A^{\mathrm{ref}}/V_{A1} s_{1z} + k_{A0y} s_{1y}/s_{1z}] z,$$

and the field structure is conserved both in the motion of the medium and in the case when the medium is at rest and only the interface is moving. Motion of the medium has a substantial influence on the amplitude of the reflected wave. The magnetosonic reflected waves are treated in an analogous way.

[1] A. A. Aleksandrova and N. A. Khizhnyak, *Boundary-Value Problems in Magnetohydrodynamics* [in Russian], Test-Radio LTD, Kharkov (1993).
[2] A. A. Alexsandrova and N. A. Khiznjak, *MAHYD-95. Proceedings of the 14th International Conference*, Riga (1995), p. 17.
[3] V. S. Vladimirov, *Equations of Mathematical Physics* [in Russian], Nauka, Moscow (1971).
[4] A. A. Aleksandrova and N. A. Khizhnyak, Magnitnaya Gidrodinamika, No. 4, pp. 15–20 (1987).
[5] A. A. Aleksandrova, Magnitnaya Gidrodinamika, No. 2, pp. 21–28 (1993).

Translated by Steve Torstveit

Evaluation of the rate constants of dissociative and ternary recombination reactions of argon ions on the basis of the results of ballistic experiments

N. N. Pilyugin

Scientific-Research Institute of Mechanics, M. V. Lomonosov Moscow State University, 119899 Moscow, Russia

(Submitted May 30, 1995; resubmitted November 4, 1995)
Zh. Tekh. Fiz. **67**, 12–18 (May 1997)

Experimental data on the change in the electron density in the wake of a ballistic object traveling at velocities $V_\infty = 3.4$–4.9 km/s in argon at pressures $p_\infty = 30$–100 Torr are processed and analyzed. A reaction scheme is proposed which takes into account the recombination of charged particles, processes of ionic conversion, and the excited states of the atom. The solution of the equations of a nonequilibrium boundary layer for flow in the wake is used to formulate the inverse problem of determining the rate constants for dissociative recombination $Ar_2^+ + e \rightarrow Ar + Ar$ and ternary recombination $Ar^+ + e + Ar \rightarrow Ar + Ar$. The ''nearest-neighbor'' approximation is used to obtain theoretically an expression for the ternary recombination coefficient as a function of temperature and pressure. Numerous solutions of inverse problems and a comparison with experiments demonstrates the validity of the expression obtained for the ternary recombination coefficient. It is shown that this expression is valid for moderate pressures and complements the Pitaevskiĭ result for low pressures and the Langevin result for high pressures. © *1997 American Institute of Physics.* [S1063-7842(97)00305-X]

INTRODUCTION

The nonequilibrium flow of a partially ionized gas in the wake behind a ballistic object traveling at hypersonic velocity is characterized by a change in pressures from several atmospheres to a few torr, temperatures from 8000 to 300 K, and a reaction zone that extends for more than 10^4 diameters of the object. Therefore, in just a single ballistic experiment one can determine the nonequilibrium physico-chemical processes over a wide range of variation of the thermodynamic parameters.[1] Thanks to the present-day achievements in the mathematical theory of information processing and modeling of hypersonic wake flows, it has become possible to make use of the vast information obtained in ballistic experiments.[2–4] In Ref. 5 the electrodynamic method was used to measure the conductivity and electron density n_e in the wake behind models traveling at velocities $V_\infty = 3.3$–4.9 km/s in argon at pressures $p_\infty = 30$–100 Torr. With the results of the measurements of Ref. 5, the solution of the inverse problem was used in Ref. 2 to find the rate constant $k_r(T)$ for dissociative recombination of $Ar_2^+ + e$ and also the Schmidt and Prandtl numbers. The data processing in Ref. 2 made use of the theory of a laminar boundary layer for the only 3 of the experiments in Ref. 5 that even approximately satisfied this flow regime (according to the characteristic Reynolds number). In the present paper all seven of the measurements in Ref. 5 are used on the basis of a turbulent model of wake flow[3] (and according to the authors of Ref. 5, this model is in better agreement with the observations). An expression for the effective recombination rate constant is obtained and used to find the true rate constants for dissociative and ternary recombination. The temperature dependence derived for the rate constant for ternary recombination is confirmed by an additional processing of the experimental data on the reaction $Al^+ + e + N_2 \rightarrow Al + N_2$. It is shown that the proposed model for the ternary recombination coefficient is well satisfied for this reaction also.

MODEL OF THE KINETICS WITH THE PARTICIPATION OF CHARGED PARTICLES IN THE WAKE OF A BALLISTIC OBJECT IN ARGON

Let us consider the following most probable processes in a medium around an object moving in argon:[6–10]

1) dissociative recombination

$$Ar_2^+ + e \xrightarrow{k_1} Ar^* + Ar,$$

2) conversion of atomic ions

$$Ar^+ + 2Ar \underset{k_2 f}{\overset{k_2 b}{\rightleftarrows}} Ar_2^+ + Ar,$$

3) Ternary recombination

$$Ar^+ + e + Ar \xrightarrow{k_3} Ar^* + Ar,$$

4) capture reaction

$$Ar^+ + e \xrightarrow{k_{capt}} Ar^*,$$

5) quenching collision

$$Ar^* + Ar \xrightarrow{k_{quench}} Ar + Ar,$$

6) decay of an atom

$$Ar^* \xrightarrow{\tau} Ar^+ + e,$$

7) de-excitation of an excited state

$$Ar^* \xrightarrow{\tau_1} Ar + h\nu.$$

 1063-7842/97/050466-07$10.00

Here A^* is an atom in an excited state. We denote the number densities of Ar, Ar^+, $Ar_2{}^+$, and Ar^+ by n_a, n_1, n_2, and n_*, respectively. Then in the simplest case (without diffusion) the kinetic equations in the medium take the form

$$\frac{dn_e}{dt} = -k_1 n_2 n_e - k_3 n_e n_a n_1 - k_{\text{capt}} n_1 n_e + \frac{n_*}{\tau}, \tag{1}$$

$$\frac{dn_*}{dt} = k_1 n_2 n_e + k_3 n_e n_a n_1 - k_{\text{capt}} n_1 n_e + \frac{n_*}{\tau} - k_{\text{quench}} n_* n_a - \frac{n_*}{\tau_1}, \tag{2}$$

$$\frac{dn_2}{dt} = -k_1 n_2 n_e + [k_{2b} n_1 n_a^2 - k_{2f} n_2 n_a]. \tag{3}$$

These equations are supplemented by the condition of quasineutrality

$$n_e = n_1 + n_2 \tag{4}$$

and the expression for the total number of particles

$$\frac{p}{kT} = n_1 + n_2 + n_e + n_* + n_a. \tag{5}$$

On the basis of the available constants for similar processes[7–12] let us estimate the right-hand sides of Eqs. (2) and (3) for the given experiments. For $n_e \approx 10^{11}$ cm^{-3}, $n_1 \approx n_2 \approx 5\times 10^{10}$ cm^{-3}, $T = 10^3$ K, and $n_a = 3\times 10^{17}$ cm^{-3} we have $k_1 \approx 10^{-8}$ cm^3/s, $k_{2b} = 10^{-31}$ cm^6/s, $k_{2f} = 10^{-14}$ cm^3/s, $k_{\text{quench}} = 10^{-15}$ cm^3/s, $\tau = 10^{-6}$ s, and $\tau_1 \cong 10^{-8}$ s. Estimates of the terms in Eq. (3) show that the ion conversion reaction takes place in a quasi-equilibrium manner, i.e., the following relation holds:

$$\frac{n_1 n_a}{n_2} = \frac{k_{2f}}{k_{2b}} = K_p(T), \tag{6}$$

where $K_p(T)$ is the equilibrium constant for ion conversion processes, which is given by[11]

$$K_p(T) = \frac{g_0 g_1}{g_2} \frac{1}{2\pi r_i^2} \sqrt{\frac{MkT}{4\pi h^2}} e^{-\frac{Ed}{kT}} \left(1 - e^{-\frac{\hbar\omega_e}{kT}}\right), \tag{7}$$

where g_0, g_1, and g_2 are the statistical weights of the atom, atomic ion, and molecular ion, respectively, M is the mass of an atom, $r_i = 2.434$ Å is the distance between nuclei;[8,13] $\omega_e = 1.5807\times 10^{13}$ s^{-1} is the frequency of a vibrational quantum, and $E_d = 0.5$ eV is the energy of dissociation of the molecular ion (Ar_2^+).

Equations (4) and (6) imply the relations

$$n_2 = \frac{n_e}{1 + \dfrac{K_p}{n_a}}, \quad n_1 = \frac{\dfrac{K_p}{n_a} n_e}{1 + \dfrac{K_p}{n_a}}. \tag{8}$$

It follows from estimates of the terms in Eq. (2) that $dn_*/dt \approx 0$, which leads to the following expression for the quasisteady density of excited atoms:

$$n_* = \frac{k_{\text{capt}} n_1 n_e + k_1 n_2 n_e + k_3 n_e n_1 n_a}{k_{\text{quench}} n_a + \tau^{-1} + \tau_1^{-1}}. \tag{9}$$

If Eqs. (8) and (9) are substituted into (1), then the kinetic equation for the electrons becomes

$$\frac{dn_e}{dt} = -k_{\text{eff}}(T) n_e^2, \tag{10}$$

$$k_{\text{eff}}(T) = \frac{\left[1 - \left(k_{\text{quench}} n_a \tau + \dfrac{\tau}{\tau_1} + 1\right)^{-1}\right]\left[k_1 + k_3 K_p + \dfrac{k_{\text{capt}} K_p}{n_a}\right]}{1 + \dfrac{K_p}{n_a}}. \tag{11}$$

Estimates of the last two terms in Eq. (11) give, in order of magnitude, $k_{\text{quench}} n_a \tau \approx 3\times 10^1 \gg 1$, $\tau/\tau_1 \sim 10^2 \gg 1$, $k_{\text{capt}}(K_p/n_a) \approx 10^{-9} \ll k_1$. Thus it follows from these estimates and Eq. (11) that

$$k_{\text{eff}}(T) \approx \frac{k_1 + k_3 K_p(T)}{1 + \dfrac{K_p(T)}{n_a}}. \tag{12}$$

Expression (12) contains the rate constants of dissociative recombination $k_1(T)$ and ternary recombination $k_3(T)$, where the third particle is an Ar atom. Using expression (12) and the temperature dependence $k_{\text{eff}}(T)$ found previously from experiments, one can in principle find $k_1(T)$ and $k_3(T)$. The form of these functions was found in accordance with the theoretical models of Refs. 7, 9, and 12:

$$k_1(T) = \frac{k_{10}}{T^m}, \quad k_3(T) = \frac{k_{30}}{T^n}. \tag{13}$$

CALCULATION OF THE EFFECTIVE RECOMBINATION RATE K_r^{eff} OF ELECTRONS WITH ARGON IONS ACCORDING TO THE RESULTS OF WAKE MEASUREMENTS

Let us first solve the inverse problem of determining the effective rate constants of a reaction from the measurements[5] and Eqs. (10) and (11) under the following rather general basic assumptions.

1. The main reaction in the medium behind the object at $x/D = 50$–500 for the electrons, according to the relations given above, is the effective reaction

$$Ar_2^+ + e \rightarrow Ar + Ar. \tag{14}$$

2. The rate constant of the reaction up to a constant factor k_0 has the form

$$k_{\text{eff}}(T) = k_0 \tilde{k}(T)\,(\text{cm}^3/\text{s}),$$

where the function $k(\tilde{T})$ is given below.

3. One can use the solution of the equations of a nonequilibrium boundary layer, which describe the flow in the far turbulent isobaric wake.[6] This yields the following formula for the temperature distribution along the axis of the wake:[6]

$$\frac{T}{T_\infty}=1+(\gamma-1)M_\infty^2\left(1-\frac{C_H}{C_x}\right)C_x\mathrm{Pr}U_0(x),$$

$$U_0(x)=2.72\left(\frac{x}{\sqrt{C_xS}}\right)^{-2/3}, \tag{15}$$

where x is the coordinate along the axis of the wake, C_x and C_H are the drag and heat exchange coefficients of the object, S is the area of the midsection of the object, $\mathrm{Pr}=0.7$ is the turbulent Prandtl number, and γ is the ratio of specific heats.

The coefficient of convective heat transfer in the flow of an inert gas around a sphere was found in accordance with the recommendations of Ref. 11.

With allowance for quasi-neutrality and the assumptions made above, we obtain for the reaction (14) by the technique of Refs. 4 and 6 the following kinetic equation for the molar mass density Ψ of electrons on the axis of the wake:

$$\frac{d\Psi}{dx}+\frac{2}{3}\frac{\Psi}{x}=-\Gamma(x)\Psi^2,$$

$$\Psi=\Psi_H, \quad x=x_H,$$

$$\Gamma=\frac{\rho_\infty\rho_0 Dk_{\mathrm{eff}}(T)}{V_\infty\mu_\infty}, \quad \rho_0=\frac{T_\infty}{T}, \quad \Psi=\frac{c_e\mu_\infty}{\mu_e}, \tag{16}$$

where ρ_0 is the dimensionless density of the gas on the axis of the wake, μ_∞ and μ_e are the mean molecular weight of the gas and the mass of an electron, D is the diameter of the model, and V_∞ is the velocity of the model.

Equation (16) is solved in quadratures in the dimensionless variables

$$\theta_e=\frac{\Psi}{\Psi_H}, \quad y=\frac{x}{x_H}, \quad \frac{T}{T_\infty}=t=1+\frac{A}{y^{2/3}}. \tag{17}$$

the solution of Eq. (16) is written in the form

$$\theta_e=y^{-2/3}(1+B_1J_1(y))^{-1},$$

$$B_1=3\sqrt{A}\frac{\rho_\infty Dk_0\Psi_H x_H}{V_\infty\mu_\infty},$$

$$J_1=\int_{\sqrt{A}y^{-1/3}}^{\sqrt{A}}\frac{dv}{v^2}\frac{\tilde{k}(T_\infty(1+v^2))}{(1+v^2)}. \tag{18}$$

Taking into account the relation $n_e=\Psi\rho/\mu_\infty$ between the molar mass density Ψ and the volume density n_e, we obtain from Eq. (18) the expressions

TABLE I.

Experiment No.	V_∞, km/s	p_∞, Torr	k_0, $\frac{\mathrm{cm}^3}{\mathrm{s}}\cdot K^m$	m	$n_{eH}\cdot 10^{-11}$ cm^{-3}
1	4.9	100	$1.194\cdot10^{-2}$	1.5	8.03
2	4.2	100	$4.896\cdot10^{-2}$	1.7	3.73
3	3.5	100	$1.999\cdot10^{-2}$	1.6	1.83
4	3.5	80	$5.411\cdot10^{-3}$	1.4	1.39
5	3.6	60	$1.595\cdot10^{-3}$	1.2	1.00
6	3.5	40	$1.838\cdot10^{-3}$	1.2	0.645
7	3.3	30	$1.316\cdot10^{-4}$	1.0	0.464

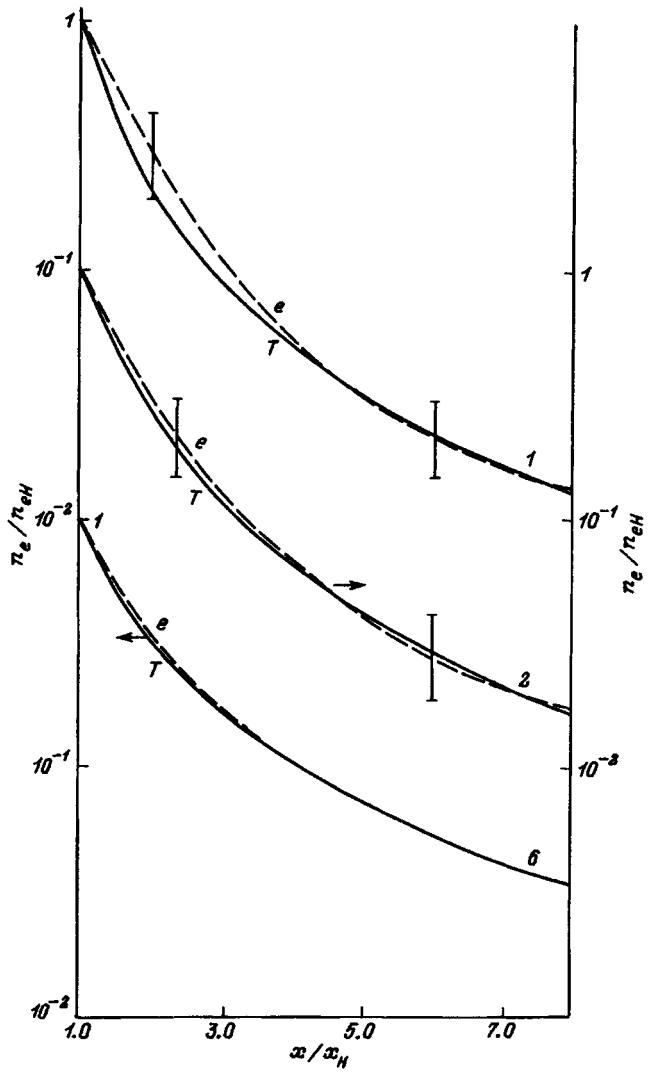

FIG. 1.

$$F(y)=\frac{n_e(y)}{n_{eH}}=\frac{t(1)}{t(y)}\frac{1}{y^{2/3}}\frac{1}{1+B_1J_1(y)}, \tag{19}$$

$$U(y)=\frac{n_e(y)}{n_{eH}}\frac{t(y)}{t(1)}y^{2/3}=\frac{1}{1+B_1J_1(y)}. \tag{20}$$

From the measured values of n_e and n_{eH}, relation (19) is used to find k_0 and the other constants which govern the temperature dependence of $k_{\mathrm{eff}}(T)$. In the case $k_{\mathrm{eff}}=k_0/T^m$ the technique of solving the inverse problem is identical to that set forth in Ref. 4.

A numerical solution for all 7 experiments gave the values of k_0 and m presented in Table I, which gives the velocity V_∞, the gas pressure p_∞, and the measured initial values of n_{eH} (at $x_H/D=50$) for each experiment. Figure 1 shows a comparison of the measured (curves labeled *e*) and theoretical (curve labeled *T*) distributions $F(y)$ for experiments *1*, *2*, and *6*, respectively, and the vertical bars indicate the confidence interval. Interval estimates of k_0 and $F(y)$ were constructed using relations similar to those considered in Refs. 3 and 4.

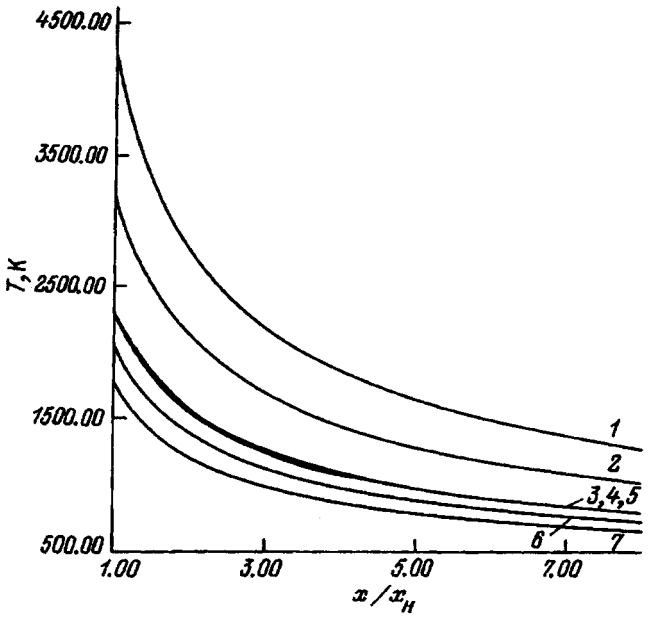

FIG. 2.

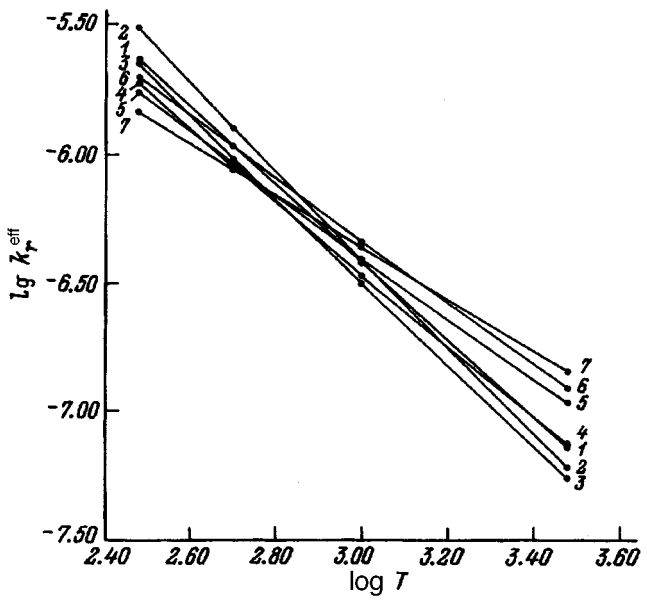

FIG. 3.

Figure 2 shows distributions of the temperature along the axis of the wake, where the numbers on the curves correspond to the experiment numbers in Table I, and Fig. 3 shows the temperature dependence of the effective coefficient of dissociative recombination $k_{\text{eff}}=k_0/T^m$. Curves *1–7* correspond to the experiment numbers. An analysis of the calculations showed that the initial conditions (V_∞, p_∞) have an appreciable influence on the temperature dependence of $k_{\text{eff}}(T)$. The inferred dependence of k_{eff} on V_∞ and p_∞ indicates that it is necessary to use formula (12).

METHOD OF SOLVING THE INVERSE PROBLEM AND THE RESULTS OF A NUMERICAL SOLUTION

To find the four parameters k_{10}, m, k_{30}, and n, we use Eq. (12) to determine the minimum of the following sum for each experiment:

$$S=\sum_{i=1}^{N}(y_i^e-y_i^T)^2, \tag{21}$$

where y_i^e is a coefficient found from experiment (see Table I); y_i^T is the theoretical expression (12), which contains the constants to be determined:

$$y_i^T=\frac{k_{10}T_i^{-m}+k_{30}T_i^{-n}K_p(T_i)}{1+K_p(T_i)n_a^{-1}(T_i)}, \tag{22}$$

where N is the number of points into which each temperature interval is divided.

By virtue of the nonlinearity of $S(k_{10},m,k_{30},n)$ the minimum of the function was found by direct coordinate descent.[14] The parameters to be determined were sought in the intervals $0.5\leqslant m\leqslant 2.0$, $0.3\leqslant n\leqslant 1.6$. The number of mesh points taken was $N=11$. The search for the minimum was terminated when the following condition was met:

$$|\boldsymbol{\alpha}^{(k+1)}-\boldsymbol{\alpha}^{(k)}|<\varepsilon\,\boldsymbol{\alpha}^{(k)},$$

where $\boldsymbol{\alpha}=\{k_{10},m,k_{30},n\}$, k is the number of the iteration, and $\varepsilon=0.01$ is the specified accuracy.

The calculations imply a very good ($\lesssim 1\%$) agreement between $k_{\text{eff}}(T)$ and the function (22) at the optimal values of $\boldsymbol{\alpha}$ found. Analysis of the data shows that the temperature dependence of k_1 is close to $T^{-1.5}$, in accordance with the theoretical ideas.[8] The results of the calculations of the constants $\boldsymbol{\alpha}$ for all the experiments were subjected to further statistical processing. The sampling mean $\overline{a}$, the true value $\hat{a}$, and the sampling standard deviation $\hat{\sigma}$ for the coefficients found were computed according to the formulas from Refs. 15 and 16. Values adopted for the data processing included significance level $\alpha=0.01$, $N=7$, and Student's coefficient for $(N-1)$ degrees of freedom $t_\alpha=3.71$. The result was

$$\overline{k_{10}}=8.83\times 10^{-3}\frac{\text{cm}^3}{\text{s}}K^m, \quad 0.553\leqslant\frac{\hat{k}_{10}}{10^{-2}}\leqslant 1.149, \tag{23}$$

$$m=1.458, \quad 1.386\leqslant m\leqslant 1.530. \tag{24}$$

Similarly, for the rate coefficient of ternary recombination we got

$$\overline{k_{30}}=1.63\times 10^{-23}\frac{\text{cm}^6}{\text{s}}K^n, \quad 0.855\leqslant\frac{\hat{k}_{30}}{10^{-23}}\leqslant 3.11, \tag{25}$$

$$\overline{n}=0.468, \quad 0.383\leqslant\hat{n}\leqslant 0.552. \tag{26}$$

We note that the relations for the coefficient of dissociative recombination which were obtained in Ref. 2 and used in Ref. 5 are statistically indistinguishable from formula (23) according to the Fisher test .

The ternary recombination coefficient obtained here (formula (25)) has the following noteworthy features: 1) the value of $k_3(T)$ at $T\sim 1000$ K is significantly higher (by several orders of magnitude) than the values given by the theories of Thomson and Pitaevskiĭ;[8,17] 2) the value of $k_3(T)$ for the reaction $Ar^+ + e + Ar \rightarrow Ar + Ar$ is close to that obtained

in Ref. 18 for the analogous coefficient in the reaction $Ar^+ + e + N_2 \rightarrow Ar + N_2$; 3) the dependence of k_3 on T is close to $T^{-0.5}$, which differs from the depends obtained by Thomson and Pitaevskiĭ.[8,17]

These facts require a theoretical explanation, which is set forth below. Based on the above processing and formulas (25) and (26), the data were recalculated for the theoretically more trustworthy relations

$$k_1 = \frac{\tilde{k}_{10}}{T^{1.5}}, \quad k_3 = \frac{\tilde{k}_{30}}{T^{0.5}}. \tag{27}$$

The calculations showed that $\tilde{k}_{10}$ has approximately the same value for all the experiments and is independent of p_∞, while the value of $\tilde{k}_{30}$ varies appreciably with pressure. This necessitates a more detailed analysis of the ternary recombination process $e + A^+ + A \rightarrow A + A$.

DERIVATION OF AN EXPRESSION FOR THE COEFFICIENT OF TERNARY RECOMBINATION

Let us first estimate the mean free path in the wake plasma for the characteristic values $T \approx 2 \times 10^3$ K, $n_a = 3 \times 10^{17}$ cm^{-3}, $n_e = 10^{11}$ cm^{-3}, $\sigma_{ea} = 10^{-15}$ cm^2. The mean free path of an electron for the Coulomb interaction with an ion[10] is $l_{ei} \approx 1.4$ cm. The mean free path of an electron for an interaction with an atom is $\lambda_{ea} \approx 3 \times 10^{-3}$ cm. The average distance between atoms

$$R_0 = \left(\frac{4\pi}{3} n_a \right)^{-1/3} \approx 2.3 \times 10^{-6} \text{ cm} \ll \lambda_{ea} \ll l_{ei}. \tag{28}$$

Let us estimate to within a numerical coefficient of order unity the rate of ternary recombination in the presence of an atom as a third particle, employing arguments analogous to those[8,10] for $A^+ + e + e \rightarrow A + e$. We assume that an electron can be captured by an ion A^+ to a closed orbit and can recombine, provided that it flies past the ion at an impact parameter r_0 such that the potential energy of attraction to the ion e^2/r_0 is greater than or of the order of the average kinetic energy of the electron $\sim (3/2)kT$. Consequently, $e^2/r_0 \approx (3/2)kT$, from which we find the effective radius of the Coulomb interaction $(e + A^+)$ or the Thomson critical radius

$$r_0 \approx \frac{2e^2}{3kT}. \tag{29}$$

The number of such collisions per cubic centimeter per second is equal to

$$n_e \overline{v}_e \pi r_0^2 n_i, \tag{30}$$

where $\overline{v}_e$ is the average thermal speed of the electron and n_i is the ion density.

In order for capture to occur it is necessary that in flying past the ion the electron interact with an atom, to which it can transfer the potential energy released in the capture. Then the probability that at the time of the collision of the electron with the ion there will be a third particle in the neighborhood, i.e., at a distance R_0, is approximately equal to the average number of particles in a volume $(4/3)\pi R_0^3$, i.e., $(4/3)\pi R_0^3 n_a$. Using this and expression (30), we obtain for the number of recombination events in the so-called ''nearest-neighbor'' approximation

$$z_{\text{rec}} = n_e \overline{v}_e \pi r_0^2 n_i \frac{4}{3} \pi R_0^3 n_a = n_e n_i n_a \pi \overline{v}_e \left(\frac{2e^2}{3kT} \right)^2 \frac{1}{n_a}. \tag{31}$$

Hence the ternary recombination coefficient which we are looking for is $(\beta_a \equiv k_3)$

$$\beta_a = \pi \sqrt{\frac{8kT}{\pi m_e}} \frac{4}{9} \frac{e^4}{kTp} \equiv \frac{\beta_0}{T\sqrt{T} n_a}. \tag{32}$$

The temperature dependence of k_3 in (32) differs from the Pitaevskiĭ formula[17] for the process $A^+ + e + A \rightarrow A + A$:

$$\beta_a^{\text{P}} = \frac{2^5 \sqrt{2\pi} \sqrt{m_e} e^6 \sigma^6 \sigma_{ea}}{3(kT)^{5/2} M_a}. \tag{33}$$

Analysis of the derivation[17] of formula (33) shows that it was obtained under the assumption that the mean free path obeys the relation $l_{ei} \lesssim \lambda_{ea}$. Using the relations given above, one can obtain from this a restriction on the pressure:

$$p(\text{Torr}) \lesssim 3.05 \times 10^{-24} \frac{n_e}{T\sigma_{ea}}. \tag{34}$$

For $n_e \approx 10^{12}$ cm^{-3}, $T = 2 \times 10^3$ K, $\sigma_{ea} = 10^{-15}$ cm^2, it follows from this that $p < 2 \times 10^{-3}$ atm ≈ 1.5 Torr. Thus formula (33) is inapplicable for wake experiments on account of the pressure limitation.

RESULTS OF CONSISTENT CALCULATIONS

It follows from Eq. (32) that the quantity

$$\beta_a p \sqrt{kT} \approx \text{const}. \tag{35}$$

Direct calculations have been carried out to search for the optimum parameters k_{10}, n, β_0, and n, with the following expression used as the theoretical recombination coefficient instead of (22), in accordance with Eq. (35):

$$y_i^T = \frac{k_{10} T_i^{-m} + \beta_0 T_i^{-(n+1)} K_p(T_i) n_a^{-1}(T_i)}{1 + K_p(T_i) n_a^{-1}(T_i)}, \tag{36}$$

where the parameters were sought in the interval $1.45 \leqslant m \leqslant 1.55$, $0.45 \leqslant n \leqslant 0.55$.

If it is assumed that the measurements are equally accurate, when the data are processed according to the formulas of Ref. 16 one finds for the coefficient of dissociative recombination

$$\overline{k}_{10} = 1.17 \times 10^{-2} \frac{\text{cm}^3}{\text{s}} K^m, \quad 1.06 \leqslant \frac{\hat{k}_{10}}{10^{-2}} \leqslant 1.29,$$

$$\overline{m} = 1.500, \quad 1.498 \leqslant \hat{m} \leqslant 1.502, \tag{37}$$

which is in good agreement with other data.[7,8] In a similar way we obtain for the rate coefficient of ternary recombination

$$k_3 = \frac{k_{30}}{\sqrt{T}} \frac{100}{p(\text{Torr})},$$

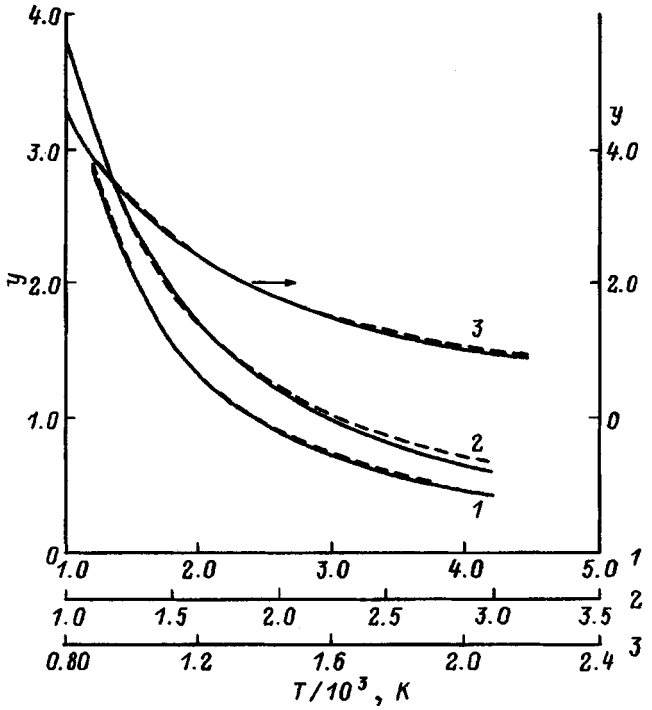

FIG. 4.

$$\overline{k_{30}} = 1.36\times 10^{-23}\frac{\mathrm{cm}^6}{\mathrm{s}}K^n, \quad \overline{n} = 0.500,$$

$$1.03 \leqslant \frac{\hat{k}_{30}}{10^{-23}} \leqslant 1.78, \quad 0.499 \leqslant \hat{n} \leqslant 0.501. \tag{38}$$

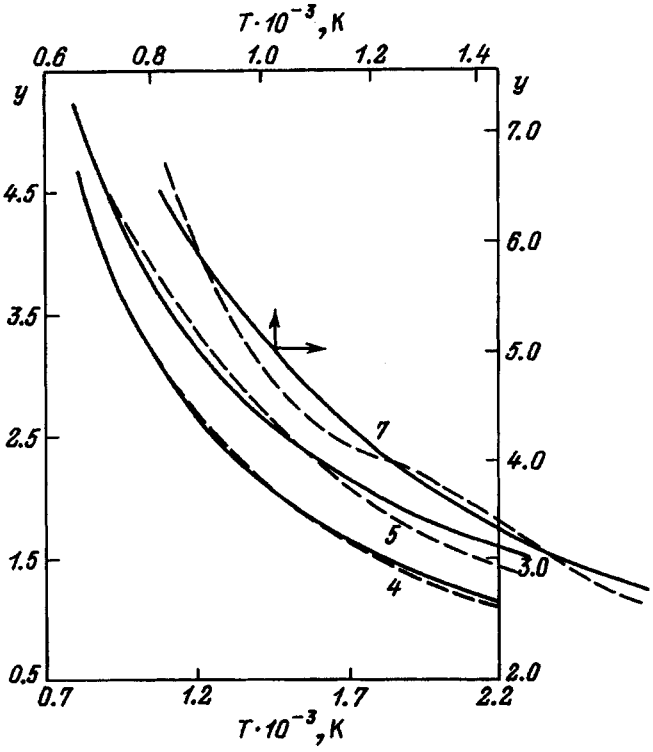

FIG. 5.

The results of calculations using these constants are presented in Fig. 4 and 5, where the solid curves are $y^e(T)$, the dashed curves were calculated according to formula (36) and the optimum constants (37) and (38) found here.

To ascertain the uncertainty in the initial numerical parameters and to establish how well the theoretical model fits the experimental data, we carried out special mathematical experiments with a variation of $K_p(T)$. It was found that varying K_p over a range of 4 orders of magnitude had little effect on the results.

At first glance it appears that the pressure dependence obtained for the ternary recombination coefficient can be explained by the phenomenon of ion mobility in a gas.[8] In that case the mean free path l_{ia} of the ions is less than the Thomson kinetic radius r_0 of Eq. (29), which leads to a restriction on the gas pressure $p > p_*$, where $p_* = 10^{-3}T^2/\sigma_0$ [atm], $[\sigma_0] = \text{Å}^2$. For our experiments $p \sim 0.1\text{atm} \ll p_* \approx 40$ atm, i.e., the inequality $p > p_*$ does not hold. Direct calculations have shown that the model for the recombination coefficient based on the Langevin theory agrees poorly with the experimental data.

COMPARISON WITH ANOTHER TERNARY RECOMBINATION REACTION

It is of interest to do an independent comparison of the proposed temperature dependence of the ternary recombination coefficient $k_3(T)$ for another gaseous medium. For this purpose we did additional calculations for the problem of Ref. 18, where the rate constant of the following reaction was investigated:

$$\mathrm{Al}^+ + e + \mathrm{N}_2 \rightarrow \mathrm{Al} + \mathrm{N}_2. \tag{39}$$

For this reaction, the rate constant was specified in Ref. 18 to be of the form $k_3 = k_{30}/T^m$, where the exponent was sought in the interval $2.3 \leqslant m \leqslant 6.0$, and it was found that

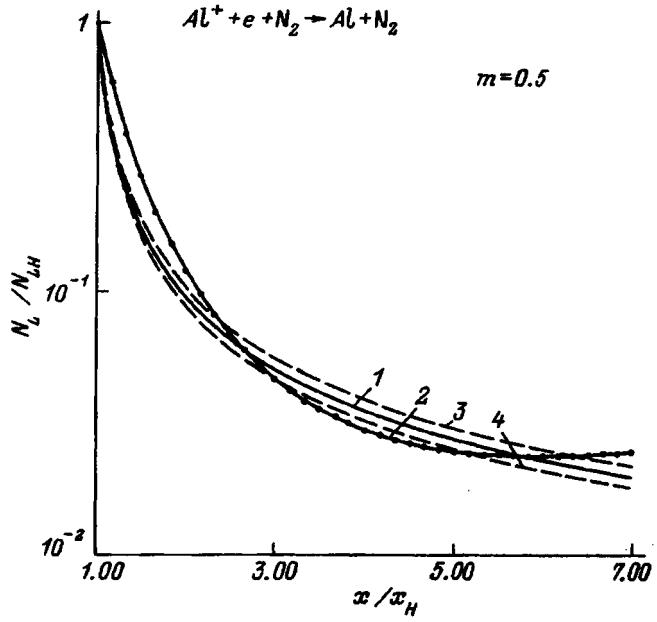

FIG. 6.

$$k_3 = \frac{(1.27 \pm 0.15) \times 10^{-17}}{T^{5/2}}, \ \frac{\text{cm}^6}{\text{s}}. \quad (40)$$

In the present paper we calculated the inverse problem[18] for $m \sim 0.5$ and found that

$$k_3 = \frac{(7.35 \pm 0.69)}{\sqrt{T}} 10^{-23}, \ \frac{\text{cm}^6}{\text{s}}. \quad (41)$$

The result of a calculation of the linear electron density N_L/N_{LH} along the axis of the wake with the use of this constant is shown in Fig. 6 (curve *1* is the theory, curve *2* experiment). The dashed curves *3* and *4* in Fig. 5 show the confidence interval. It should be noted that the variance of N_L differs little in the two cases (40) and (41). From a comparison of N_L for Ar^+ (formula (38)) and Al^+ (formula (41)) it is seen that the rate constants for ternary recombination are in good agreement. This serves as additional evidence in favor of the adopted theoretical model for $k_3(T)$.

The author is grateful to S. A Losev and G. D. Smekhov for a discussion of the results and to A. N. Pilyugin for assistance in doing the calculations.

This study was done with the financial support of the Russian Fund for Fundamental Research (Project Code No. 95-01-01611).

[1] E. P. Andreev, N. N. Pilyugin, O. K. Taganov, and S. G. Tikhomirov, *Investigation of the Radiation Emitted from Gases in Aeroballistic Experiments* [in Russian], Izd. MGU, Moscow (1988).
[2] S. Yu. Menzhunskiĭ and N. N. Pilyugin, Teplofiz. Vys. Temp. **31**, 787 (1993).
[3] A. N. Pilyugin, N. N. Pilyugin, and S. G. Tikhomirov, Teplofiz. Vys. Temp. **31**, 517 (1993).
[4] N. N. Pilyugin and A. N. Pilyugin, Teplofiz. Vys. Temp. **33**, 351 (1995).
[5] V. N. Gadion, V. G. Ivanov, G. I. Mishin *et al.*, Zh. Tekh. Fiz. **42**, 1049 (1972) [Tech. Phys. **17**, 832 (1972)].
[6] N. N. Pilyugin, S. G. Tikhomirov, and S. Yu. Chernyavskiĭ, Izv. Akad. Nauk SSSR Mekh. Zhidk. Gaz., No. 6, pp. 165–175 (1980).
[7] A. V. Eletskiĭ and B. M. Smirnov, Usp. Fiz. Nauk **136**, 25 (1982) [Sov. Phys. Usp. **25**, 13 (1982)].
[8] B. M. Smirnov, *Atomic Collisions and Elementary Processes in Plasmas* [in Russian], Atomizdat, Moscow (1984).
[9] A. Kimura, K. Teshima, and M. Nishida, Trans. Jpn. Soc. Aero Space Sci. **18**, No. 41, pp. 103–113 (1975).
[10] Ya. B. Zeldovich and Yu. P. Raĭzer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena* [in Russian], Nauka, Moscow (1966).
[11] M. G. Vasil'ev, T. V. Zhikhareva, and G. K. Tumakaev, Zh. Tekh. Fiz. **49**, 541 (1979) [Sov. Phys. Tech. Phys. **24**, 311 (1979)].
[12] D. I. Slovetskiĭ, *Mechanisms of Chemical Reactions in Nonequilibrium Plasmas* [in Russian], Nauka, Moscow (1980).
[13] R. S. Devoto, PTK **7**, No. 2, pp. 10–17 (1969).
[14] A. E. Mudrov, *Numerical Methods for the PC in BASIC, Pascal, and FORTRAN* [in Russian], MP Rasko, Tomsk (1991).
[15] S. A. Aĭvazyan, I. S. Enyukov, and L. D. Meshalkin, *Applied Statistics. Investigation of Dependencies* [in Russian], Finansy i Statistika, Moscow (1995).
[16] V. A. Kolemaev, O. V. Staroverov, and V. B. Turundaevskiĭ, *Theory of Probabilities and Mathematical Statistics* [in Russian], Vysshaya Shkola, Moscow (1991).
[17] L. P. Pitaevskiĭ, Zh. Éksp. Teor. Fiz. **42**, 1326 (1962) [Sov. Phys. JETP **15**, 919 (1962)].
[18] N. N. Pilyugin, S. Yu. Menzhinskiĭ, and A. N. Pilyugin, Teplofiz. Vys. Temp. **32**, 656 (1994).
[19] N. N. Pilyugin and G. A. Tirskiĭ, *Dynamics of Radiation-Emitting Ionized Gases* [in Russian], Izd. MGU, Moscow (1989).

Translated by Steve Torstveit

Interaction of an expanding gas cloud with a perforated screen

A. V. Vinogradov, V. A. Volkov, V. Yu. Gidaspov, and P. V. Rozovskiĭ

Moscow State Aviation Institute, 125871 Moscow, Russia
(Submitted September 29, 1995; resubmitted March 18, 1996)
Zh. Tekh. Fiz. **67**, 19–22 (May 1996)

The expansion of a high-pressure gas cloud in a vacuum chamber in the presence of a low-pressure background gas is considered. The interaction of the expanding cloud with the background gas and with a perforated screen placed in the path of the flow is modeled numerically. The influence of the permeability of the screen on the overall picture of the interaction and on the parameters of the resulting flow is studied. © *1997 American Institute of Physics.*
[S1063-7842(97)00405-4]

In this paper we consider unsteady flows arising in the expansion of a high-pressure gas cloud in a low-pressure background gas. Such a situation arises in the vacuum chambers of certain experimental apparatus[1,2] on account of an insufficiently deep evacuation of the chamber. In the corresponding experiments the high-pressure gas cloud was formed either as a result of the electrical explosion of a foil or thin wire when a high discharge current was passed through it from a capacitor bank,[1] or as a result of evaporation of the sample material by means of a high-power laser pulse.[2]

The chief goal of the studies done on such apparatus is to create optimum conditions for the formation of condensation clusters in the expanding cloud and their subsequent deposition on special substrates or on the walls of the chamber. However, studies by the present authors have shown[3–6] that the bow shock wave that is reflected off the wall, upon reaching the boundary of the layer of compressed background gas, decomposes into refracted and reflected waves, after which there occur multiple reflections of the newly formed shock waves from the walls and from the boundary of the layer of compressed background gas. This causes additional adiabatic compression and a corresponding increase in the temperature of this layer, and that may lead to evaporation of the clusters, which have to cross it in order to reach the wall or target. From the standpoint of theoretical applications and the mounting of experiments it is important to have at one's disposal various means by which one can influence the parameters of the unsteady flow in the apparatus, thereby altering the conditions of formation and precipitation of the clusters. One such means is a perforated screen mounted in the path of the expanding cloud. The present study is a natural continuation of Refs. 3–5. Here a numerical simulation is used to investigate the influence of the permeability of the perforated screen on the overall pattern and parameters of the flow.

We considered flows with planar, cylindrical, or spherical symmetry, which were described by time-dependent quasi-one-dimensional Euler equations. The field of the parameters at the initial time were specified as follows. It was assumed that the cloud of evaporated material occupies the spatial region $0<r<r_0$, where r is a generalized coordinate equal to the distance from the point in question to the plane, axis, or center of symmetry. All the remaining volume of the chamber is assumed to be filled with a background gas. The velocity is assumed to be zero everywhere, and the temperature and pressure in the vapor cloud and in the background gas are assumed to be uniform and equal to T_0, p_0 and T_H, p_H, respectively. In addition, it is assumed that $p_0 \gg p_H$ and that both gases are ideal, with specific heats that do no depend on temperature.

This initial field is a rather crude approximation to the real flow fields of expanding clouds in experimental apparatus. Nevertheless, in the course of the subsequent temporal evolution of such an initial field the flow that will form is qualitatively and in some ways quantitatively similar to the real flow fields.

The influence of the permeability of the perforated screen on the flow was modeled in analogy with Refs. 7 and 8 with suitable boundary conditions at the screen, obtained in the framework of the following assumptions about the gas flow in the perforation channels: 1) At any instant in time the flow in the perforation channels and in some neighborhood of them is steady, and the influence of the viscosity and thermal conductivity of the gas can be neglected. 2) The gas flow prior to entering the perforation channels is continuous and subsonic. 3) On the whole, in each channel and in a certain neighborhood of it there is blocked flow with a transition through the speed of sound.

Under these assumptions the flow field in front of the screen, i.e., for $0<r<r_W$, where r_W is the coordinate of the screen, is completely independent of the character of the flow behind the screen, which can be ignored altogether. The flow field in front of the screen can be calculated by solving numerically on the given interval a suitable initial/boundary-value problem for the time-dependent quasi-one-dimensional Euler equations with the boundary conditions described below, which take into account the specifics of the problem under study.

At the coordinate origin ($r=0$) the obvious symmetry boundary condition $u=0$ is imposed. At the perforated screen ($r=r_W$) the form of the boundary condition depends on time. In particular, up until the time t_A corresponding to the arrival of the bow shock wave at the screen, one has on both sides of the screen a nonmoving background gas at equal pressures. Therefore, for $t_0<t<t_A$ the condition $u_W=0$ is imposed.

The boundary condition at the screen after reflection of

 1063-7842/97/050473-04$10.00

the bow shock wave from it and at all subsequent times can be written, under assumptions 1–3, in the form

$$\rho_W u_W = \rho^* a^* \varepsilon \mu, \tag{1}$$

where the asterisks indicate the critical sound velocity and density of the gas flowing to the screen, μ is the flow-rate coefficient, ε is the ratio of the total area of the minimum cross sections of the perforation channels to the total area of the entire screen, and the product $\varepsilon\mu = \eta$ is called the permeability of the perforated screen.[7]

The difference of the flow-rate coefficient μ from unity can be neglected, and one can assume approximately that $\eta \approx \varepsilon$ (Refs. 7 and 8). Using assumptions 1–3 and the known gasdynamic relations for isentropic flows, one can reduce Eq. (1) to the form

$$M_W \left(\frac{k+1}{2+(k-1)M_W^2} \right)^{(k+1)/2(k-1)} = \eta, \tag{2}$$

where the left-hand side is the known gasdynamic function $q(M,k)$, with k the adiabatic exponent.

For $\eta < 1$ this equation, as we know, has two roots: supersonic and subsonic. For the latter we introduce the notation $M_W = M_{\text{sub}}(\eta,k)$. Let us also take t_B to be the time at which the outer boundary of the expanding cloud reaches the screen. Then the boundary condition at the screen as a function of time can be written as

$$M_W = \begin{cases} 0 & \text{for} \quad t_0 < t < t_A, \\ M_{\text{sub}}(\eta, k_H) & \text{for} \quad t_A < t < t_B, \\ M_{\text{sub}}(\eta, k_0) & \text{for} \quad t_B < t, \end{cases} \tag{3}$$

where k_H and k_0 are the isentropic exponents of the classical gas and vapor, respectively.

The boundary condition (3) together with the characteristic relations satisfied along the characteristics $r_t' = u$ and $r_t' = u + a$ enable one to find all the remaining parameters of the gas flowing into the screen and, hence, they completely predetermine the initial/boundary-value problem under study.

For numerical solution of the problem we used the same algorithm[6] as in Refs. 4 and 5. However, since a perforated screen, unlike a wall, is permeable for the gas, it can be reached not only by shock waves but also by contact discontinuities. In both cases, as a rule, discontinuities of the gasdynamic parameters will arise at the corresponding points in time. The problem of the decay of an arbitrary discontinuity at a perforated screen has been considered in a rather general formulation in Ref. 7. There, however, only the most fundamental aspects of the construction of the solution are discussed, and the final form of the computational formulas is not given. In the particular case in question, i.e., in the framework of assumptions 1–3, a complete solution of the given problem can be obtained in closed form.

In the formulas that follow, the subscript S indicates parameters of the flow behind the shock wave or contact discontinuity at the time when they have just reached the screen, and the subscript W indicates the parameters at the screen immediately after the decay of the corresponding discontinuity. We also let M_W be equal to the subsonic root of equation (2) at values of the permeability η and isentropic exponent k for the gas behind the strong discontinuity that has arrived at the screen, i.e., for the gas flow to the screen.

The complete solution for the decay of an arbitrary strong discontinuity on a screen at which assumptions 1–3 hold depends on two parameters, viz., M_W and the Mach number M_S of the oncoming flow, the first of which, according to the assumptions adopted, is clearly less than 1, while the second can be arbitrary. Depending on the particular values of M_W and M_S, the complete solution of the given problem admits five subcases, which we shall henceforth denote by a pair of numbers separated by a decimal point and enclosed in parentheses.

1. If M_W satisfies the inequality

$$1 > M_W > \sqrt{(k-1)/2k}, \tag{4}$$

then, depending on the Mach number M_S of the oncoming flow, three different types of flow can be realized after the decay of the discontinuity: these will be designated as (1.1), (1.2), and (1.3). The first of these, (1.1), is realized if the condition $M_S < M_W$ holds in addition to the condition (4). In this case a fan of rarefaction waves propagates counter to the oncoming flow, and the flow parameters immediately in front of the screen are found from the formula

$$\frac{a_W}{a_S} = \frac{(k-1)M_S + 2}{(k-1)M_W + 2}, \quad \frac{\rho_W}{\rho_S} = \left(\frac{a_W}{a_S} \right)^{2/(k-1)},$$
$$\frac{p_W}{p_S} = \left(\frac{a_W}{a_S} \right)^{2k/(k-1)}. \tag{5}$$

The second type of flow, (1.2), comes about as a result of the decay of the discontinuity when the following condition holds in addition to (4):

$$M_W < M_S < M_{SW},$$
$$M_{SW} = \sqrt{1 + \frac{k-1}{2} M_W^2} \Big/ \sqrt{kM_W^2 - \frac{k-1}{2}}. \tag{6}$$

In this case the disturbance moving counter to the oncoming flow is a shock wave with a Mach number in relation to the gas $M_{RS} > 1$ which is found by solving an irrational equation of the form

$$M_W = \frac{\left[M_S - \frac{2}{k+1} \frac{M_{RS}^2 - 1}{M_{RS}} \right]}{\sqrt{1 + \frac{2(k-1)}{k+1} \frac{M_{RS}^2}{M_{RS}^2} \left(1 + \frac{k}{k+1} (M_{RS}^2 - 1) \right)}}. \tag{7}$$

Here the velocity D_{RS} of the reflected shock wave relative to the screen and all the remaining parameters of the flow immediately in front of the screen can be expressed in terms of the parameters of the oncoming flow and the value obtained for M_{RS} in accordance with the formulas

$$D_{RS} = a_S (M_S - M_{RS}),$$
$$u_W = u_S - 2a_S (M_{RS}^2 - 1)/(k+1) M_{RS},$$
$$\frac{p_W}{p_S} = 1 + \frac{2k}{k+1} (M_{RS}^2 - 1), \quad \frac{\rho_W}{\rho_S} = \frac{M_{RS}^2 (k+1)}{(k-1)M_{RS}^2 + 2}. \tag{8}$$

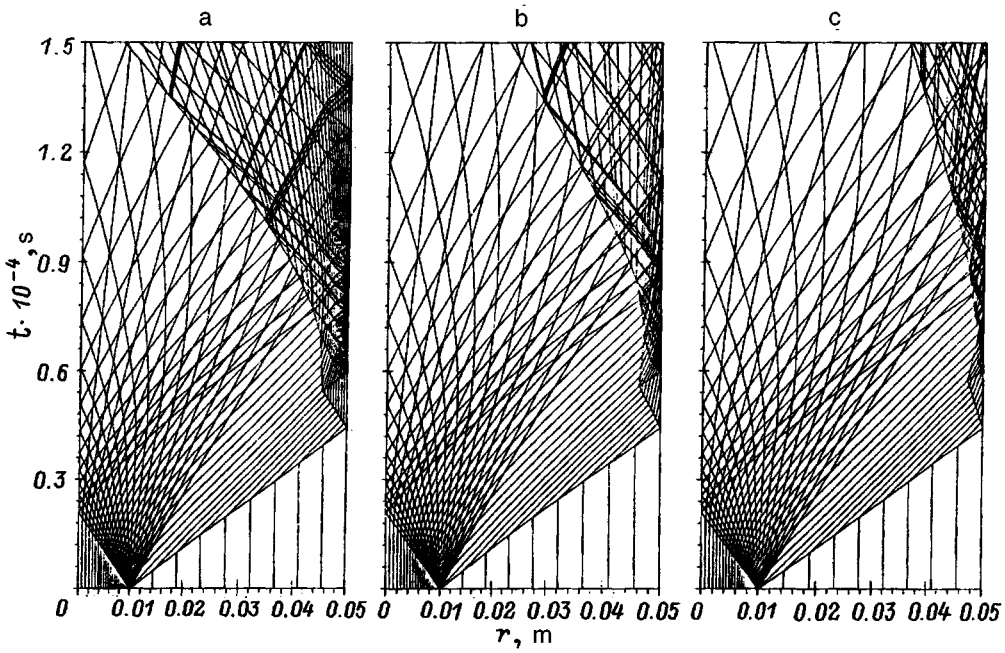

FIG. 1.

It can be shown that fulfillment of condition (6) will guarantee that the value of D_{RS} found according to formula (8) will be negative. However, as M_S approaches M_{SW}, the corresponding value of D_{RS} approaches zero, so that in the limiting case when M_S is exactly equal to M_{SW}, a shock wave that does not move relative to the screen is formed in front of it.

The third possible type of flow, (1.3), comes about as a result of the decay of a discontinuity at the screen when the condition $M_S > M_{SW}$ holds in addition to (4). In this case a single shock wave, even one that does not move relative to the screen, is no longer sufficient for braking the oncoming gas the required amount, i.e., to M_W. Therefore, in this case two shock waves arise, one of which does not move relative to the screen and one of which propagates counter to the oncoming gas. The flow parameters between these two shock waves are indicated by the subscript SW. To determine the parameters in front of the screen, let us first solve numerically the equation analogous to (7) but with M_{SW} in place of M_W to find the Mach number M_{RS} of the shock wave moving counter to the oncoming gas flow. Then we use the value obtained for M_{RS} and the known parameters of the oncoming flow in formulas analogous to (8) but with W replaced by WS to find the parameters of the flow between the two shock waves. Then, using formulas analogous to (8) but with M_{RS} replaced by M_{SW} and with the subscript S replaced by WS, we find the parameters of the flow directly in front of the screen.

2. If $M_W < 1$ then the inequality (4) does not hold, and, depending on the particular value of M_S, there are just two types of flows that can be realized after the decay of the discontinuity, namely (2.1) and (2.2). Type (2.1) comes about for $M_S < M_W$ and type (2.2) for $M_W < M_S < \infty$. These two types are completely analogous to types (1.1) and (1.2), including the use of the computational formulas (6) and (7), (8), respectively.

It is easy to see that in the limiting case for $\eta \to 0$ the solution given above goes over continuously to the known solution for the reflection of a shock wave from an impermeable wall.

The main purpose of an experiment is ordinarily to achieve the maximum number of clusters reaching the wall or substrate; this number is determined, in particular, by the parameters of the reflected secondary shock waves and of the layer of shock-compressed gas. Therefore, one is most interested in quantitative data concerning the influence of the permeability of the screen on the intensity of the reflected shock waves and on the parameters of the flow immediately in front of the screen. Computational results illustrating this influence in the case of a flow with planar symmetry are presented in Figs. 1 and 2. For the series of calculations shown only the permeability was varied, while the rest of the initial data were identical and corresponded to typical conditions for experiments with silver vapor: $r_0 = 1$ cm, $r_W = 5$ cm, $p_0 = 5.27$ kPa, $T_0 = 1792$ K. The background gas was

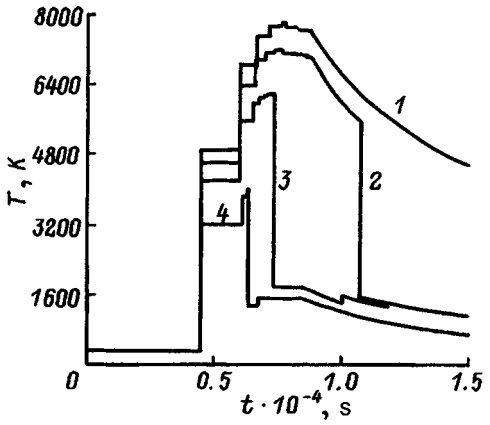

FIG. 2.

also taken to be silver vapor at $p_H = 10$ Pa and $T_H = 300$ K, which corresponds to the case when the latter remains in the chamber on account of insufficiently deep evacuation of the chamber after a previous experiment.

Figure 1 illustrates the influence of the permeability of the screen on the overall pattern of the flow in front of it. This figure was made by constructing on the plane of independent variables r,t the streamlines, the trajectories of the strong and weak discontinuities, and the sonic characteristics of the fans of rarefaction waves, which were captured by means of additional meshes. The pattern of the flow in the case of an impermeable wall, i.e., for $\eta = 0$, is shown in Fig. 1a, while the patterns in the case of perforated screens with $\eta = 0.3$ and 0.6, respectively, are shown in Figs. 1b and 1c. It is seen from Fig. 1 that qualitatively similar flow patterns are observed in all three cases. On the other hand, it is also seen that there are significant quantitative differences, which are due to the leaking of the compressed background gas through the perforated screen. In particular, in the last two cases the thickness of the layer of compressed background gas decreases monotonically and ultimately vanishes completely. At this time t_B the directly expanding vapor begins to reach the screen. Therefore, the clusters formed in the vapor can now reach the screen unimpeded, provided that they have not evaporated sooner, i.e., after after the increase in temperature and pressure caused by the passage through the expanding vapor cloud of the shock waves reflected from the screen.

In connection with this last circumstance it is useful to examine some plots which will illustrate the influence of the permeability on the character of the time dependence of the temperature of the gas flowing into the screen. These plots are shown in Fig. 2 for four values of the permeability, 0, 0.1, 0.3, and 0.8, in increasing order (curves *1–4*, respectively). The last three plots pertain to the case of a permeable screen. They differ qualitatively from the first plot, which is for an impermeable wall, in that they exhibit a downward jump of the temperature. This jump occurs when all of the background gas contained between the outer boundary of the expanding cloud and the screen has flowed through the screen and the expanding vapor begins to come through. In principle the temperature and pressure of this vapor should depend on the permeability of the screen, since prior to this these parameters could have increased in the reflected shock waves propagating counter to the vapor flow, the intensities of which, in turn, depend on the permeability of the screen. However, in comparing the temperatures of the vapor flowing into the screen, we note that they are just the same for $\eta = 0.1$ and 0.3 and only slightly different, much less than the difference in the corresponding temperatures of the background gas, for $\eta = 0.1$ and 0.8.

This leads us to the preliminary conclusion that varying the permeability of the screen over rather wide limits will apparently not enable one to influence substantially the conditions of cluster formation in the expanding vapor cloud, in spite of the fact that the intensities of the shock waves reflected from the screen depend on its permeability. However, the final justification of this conclusion will require additional results from parametric calculations and also experimental confirmation. Meanwhile, another conclusion is obvious from the already available results of the calculations: a permeable screen or target will pass all the compressed and (accordingly) heated background gas, after which the conditions for the clusters formed in the expanding vapor cloud to reach the surface become more favorable than in the case of an impermeable wall or target.

[1] W. G. Chace and H. K. Moore (eds.), *Exploding Wires*, Plenum Press, New York (1962) [Russ. transl., IL, Moscow (1963)].
[2] A. V. Bulgakov, B. N. Kozlov, A. P. Maĭorov *et al.*, *Abstracts of the Eleventh All-Union Conference*, Leningrad, (1991), p. 131.
[3] A. L. Itkin and P. V. Rozovskiĭ, Zh. Tekh. Fiz. **55**, 1182 (1985) [Sov. Phys. Tech. Phys. **30**, 675 (1985)].
[4] A. V. Vinogradov, V. A. Volkov, V. Yu. Gidaspov, and P. V. Rozovskiĭ, Zh. Tekh. Fiz. **63**(11), 20 (1993) [Tech. Phys. **38**, 946 (1993)].
[5] A. V. Vinogradov, V. A. Volkov, V. Yu. Gidaspov *et al.*, Vychislitel. Tekhnol. **4**(10), 89 (1995).
[6] A. V. Vinogradov, V. A. Volkov, V. Yu. Gidaspov *et al.*, in *Mathematical Modeling* [in Russian] (in press, 1995).
[7] V. T. Grin', A. N. Krai'ko, and L. G. Miller, Zh. Prikl. Mekh. Tekh. Fiz., No. 3, pp. 95–102 (1981).
[8] S. Yu. Tonchak, Candidate's Dissertation [in Russian], Moscow (1988), 148 pp.

Translated by Steve Torstveit

Role of different polarization mechanisms in the self-organization of the director of a thin layer of nematic liquid crystal

Yu. K. Kornienko and A. P. Fedchuk

I. I. Mechnikov Odessa State University, 270100 Odessa, Ukraine
(Submitted November 28, 1995)
Zh. Tekh. Fiz. **67**, 23–28 (May 1997)

A theoretical analysis of the processes occurring in a symmetric metal–liquid-crystal–metal structure is carried out to obtain the dependence of the anisotropic component of the surface free energy on the angle of deviation of the director from the direction of easy orientation. The differences between the angular dependence of the surface energy obtained in this paper from the form described by the Rapini potential are discussed. The coordinate dependence of the ionic and dipolar polarization are calculated, as is the order parameter characterizing the ordering of the dipole moments of the liquid-crystal molecules. The interrelationships of the the boundary polarization with the equilibrium orientation of the director near the surface bounding the liquid crystal and with the value of the internal electric field in the volume of the mesophase are discussed. [S1063-7842(97)00505-9]

INTRODUCTION

Liquid crystals (LCs) are unique objects for investigating collective effects, particularly self-organization processes,[1] which are currently a focus of research attention. The long-range correlations, as a rule, are determined by the properties of the LC volume in the absence of outside influences, and the orientation fluctuations propagating into the interior of the LC layer are specified by the boundary conditions at the surfaces bounding the sample.[2] However, as was shown in Ref. 3, in a thin gap filled with a dielectric liquid between conducting walls, the long-range correlations fall off with distance much more slowly than in the case of dielectric walls. This leads to a number of interesting effects. For example, in such a system one observes an orientational phase transition analogous to the Fréedericksz effect but occurring in the absence of crystal electric fields; this was first predicted by the authors in Ref. 4. If the bounding surface also has a definite translational symmetry, then the propagation of long-range correlations of the director orientation into the interior of the sample depends on the character of the packing of the surface atoms, as we showed in Ref. 5. Further, as was shown in Ref. 6, when the uniform orientation of the nematic LC is disrupted the medium loses local inversion symmetry, which in the unperturbed case is ensured by antiparallel packing. This leads, in particular, to a polarized surface layer,[7,8] the experimental characteristics of which are directly related to the orientation of the director near the bounding surface.[9] Consequently, there is an obvious need for a study of the characteristics of this layer in connection with research on self-organization processes in such systems. To determine the boundary orientation of the molecules, as we know,[10] one must minimize the surface energy of the LC layer. Despite the fact that there have been a number of papers on this topic (e.g., Ref. 11), this problem remains unsolved from a theoretical standpoint for systems including nematic LCs with conducting solid substrates.

It is known[10] that the surface free energy F_0 assumes its minimum value along the direction of easy orientation. The surface energy includes an isotropic component F_{0i} and an anisotropic component F_{0a}. As was noted in Ref. 10, the specific form of F_{0a} is a key question in the physics of LC surfaces. In the simplest case of small angles of deviation α from the direction of easy orientation[12,13] the anisotropic surface energy is of the form $F_{0a}=(1/2)W\sin^2\alpha$. In all other cases one uses for F_{0a} a potential of more complex form, e.g., in the form of an elliptic sine[14] or Legendre polynomials of even powers, with an argument of $\cos\alpha$ or $\sin\alpha$ (Ref. 15). However, all these approximations, in our opinion, have a common shortcoming: the lack of a consistent physical model. Furthermore, there has been no conclusive theoretical analysis of the anisotropic surface energy. There is just one paper in the literature[16] in which a mechanism was proposed for making the surface energy different from the Rapini potential: the flexoelectric effect. The authors arrive at an approximation of the form

$$F_{0a}=-\left(\frac{W}{2}+\frac{2}{3}U_4\right)\cos^2\theta+U_4\cos^4\theta+\text{const},$$

where U_4 is a quantity that depends on the difference of the order parameter near the surface S_s and in the volume of the mesophase S_v and which is related in a definite way with the flexoelectric coefficients θ_{11} and θ_{33}.

Even in that paper, however, the anchoring energy W enters only as an adjustable parameter and is not tied in with the properties of the structure components. Therefore, the given model, in particular, cannot be used for an analysis of the aforementioned structure in which the LC layer is bounded by solid conducting surfaces.

It is certainly of interest to analyze such a system theoretically in order to find the anisotropic surface energy. In addition, as we mentioned earlier, near the solid surface bounding the LC there arises a polar layer of molecules. There are no reported studies of the contribution F_{0a} of this layer to the surface energy. We thought it was very important to fill this gap in knowledge by considering in the framework of a unified model the processes occurring in a structure

consisting of a thin layer of nematic LC bounded by solid substrates with conducting coatings, including processes of near-electrode polarization, and on the basis of this analysis to determine the anisotropic component of the surface energy of the LC molecules.

MODEL CONCEPTS

We consider a symmetric structure consisting of a metal (M), a nematic liquid crystal (NLC), and a metal (as necessary the results of this study can be generalized to the case of an asymmetric metal–LC–semiconductor (S) structure; an incomplete step in this direction this was taken by the authors in a previous paper[4]). We consider the simplest one-dimensional case: the x axis is directed perpendicular to the two electrodes, one of which is at the coordinate origin and the other at d. The LC is assumed ideal, i.e., devoid of intrinsic charge carriers, which might be formed as a result of dissociation of the molecules or which might be impurity ions. Thus charges can appear in the NLC only on account of the emission of charge carriers from the metal electrodes; in the simplest case one may consider only thermionic emission. The thermionic current through the M–LC barrier is determined by the parameters of both the metal electrode and the electrical double layer (EDL) of the liquid crystal, and the properties of the EDL in turn depend on this current. As a result, in the immediate proximity of the contact a definite density of emitted charge carriers n_{m0} is established in the NLC; according to Refs. 4 and 17, this density is given by

$$n_{m0}=\frac{AT^2}{q}\left(\frac{2\pi m}{kT}\right)^{1/2}\exp\left(\frac{\phi-\varphi_m}{kT}\right)\exp\left(-\frac{q\Delta\Psi_{LCM}}{kT}\right), \tag{1}$$

where ϕ and φ_m are the work function of an electron in the LC and metal, respectively, $\Delta\Psi_{LCM}$ is the potential drop across the EDL of the LC, which is determined in turn by the boundary orientation of the director and the packing of the NLC molecules on the surface.[4,18]

The presence of charge carriers in the boundary regions of the NLC causes their diffusional transport, which is directed toward the center of the LC layer, leading to a nonuniform distribution over x of the charges in the LC layer; this is equivalent to the appearance of a nonuniform electric field $E_{LC}(x)$ in the structure. To find the functional dependence of the electric field strength and of the density of charge carriers, we solved jointly Poisson's equation and the equation of diffusion–drift equilibrium with the boundary density (1) substituted in. Previously in Ref. 4 we obtained the following functions:

$$n(x)=n_0/\cos^2\left((Cn_0/2)^{1/2}\left(x-\frac{d}{2}\right)\right), \tag{2}$$

$$E_{LC}(x)=\frac{kT}{q}(2Cn_0)^{1/2}\tan\left((Cn_0/2)^{1/2}\left(x-\frac{d}{2}\right)\right), \tag{3}$$

where $C=q^2/(kT\varepsilon_{LC}\varepsilon_0)$, n_0 is the minimum density of charge carriers in the LC (at $x=d/2$), and ε_{LC} is the dielectric constant of the LC.

As we have shown in Ref. 4, ε_{LC} depends in a definite way on the angle θ between the director and the plane of the electrode, specifically

$$\varepsilon_{LC}=\varepsilon_\perp+\varepsilon_a\sin^2\theta, \tag{4}$$

where $\varepsilon_\perp$ is the dielectric constant corresponding to the planar orientation of the LC, and ε_a is the dielectric anisotropy.

The boundary values of the electric field and carrier density at $x=0$ and d will be, respectively,

$$E_{LC}|_{x=0,d}=\mp\frac{kT}{q}(2Cn_0)^{1/2}\tan\left((Cn_0/2)^{1/2}\frac{d}{2}\right), \tag{5}$$

$$n_{m0}=n|_{x=0,d}=n_0/\cos^2\left((Cn_0/2)^{1/2}\frac{d}{2}\right). \tag{6}$$

We must also use the following boundary condition, which comes from applying Gauss' theorem of electrostatics for the metal–LC interface:

$$\varepsilon_{LCM}\varepsilon_0E_{LC}|_{x=0}=q(n_{m0})^{2/3}. \tag{7}$$

In writing this theorem we took into account, first, that the field is zero inside the metal electrode and, second, that the surface charge density is equal to the product of the elementary charge and the boundary surface density of carriers (which is equal to the boundary volume density n_{m0} raised to the 2/3 power).

RESULTS OF THE CALCULATION AND DISCUSSION

As we know, the surface energy enters as a term in the free energy functional:

$$\Phi(\theta)-\int f\left(\theta,\frac{df}{dx}\right)dx+F_{01}+F_{02}, \tag{8}$$

where F_{01} and F_{02} are the unknown values of the surface energy for each of the boundaries, and $f(\theta,d\theta/dx)$ is the volume free energy density, which in the general case is of the form[19]

$$f=\frac{1}{2}[K_1(\nabla\cdot\mathbf{n})^2+K_2(\mathbf{n}\cdot(\nabla\times\mathbf{n}))^2 + K_3(\mathbf{n}\cdot(\nabla\times\mathbf{n}))^2-\varepsilon_a\varepsilon_0E^2\sin^2\theta]. \tag{9}$$

The problem of minimizing the functional (8) is equivalent[10] to solving the Euler–Lagrange equation, which in general has the form

$$\frac{df}{dx}-\frac{d}{dx}\left(\frac{df}{d(d\theta/dx)}\right)=0 \tag{10}$$

with the boundary conditions

$$\left.\frac{df}{d(d\theta/dx)}\right|_{x=0}+\left.\frac{dF_0}{d\theta}\right|_{x=0}=0, \tag{11}$$

$$\left.\frac{df}{d(d\theta/dx)}\right|_{x=d}-\left.\frac{dF_0}{d\theta}\right|_{x=d}=0. \tag{12}$$

Let us restrict discussion to the case of a transverse bending deformation of the LC director, with a characteristic elastic constant K_1. As we have shown previously,[18] this

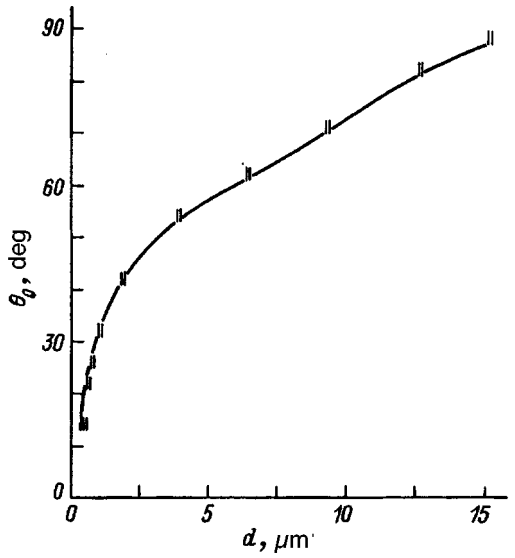

FIG. 1. Boundary angle θ_0 versus the thickness of the LC layer.

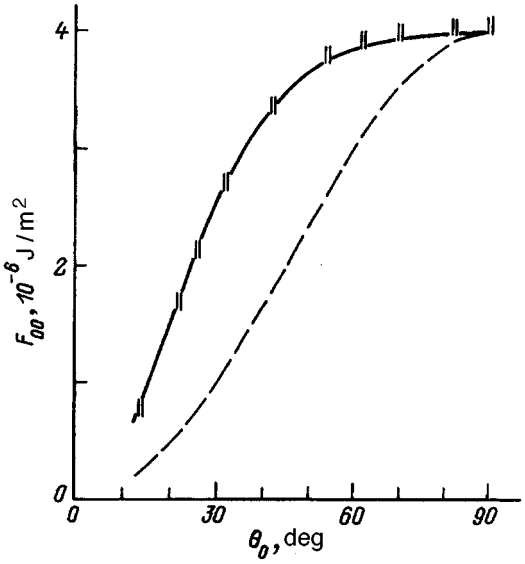

FIG. 2. Surface energy versus the boundary angle θ_0.

situation arises when an initial planar orientation of the molecules (which in practice can come about in the capillary filling of a cell) is then subjected to the influence of the internal electric field, which is directed, as we have said, perpendicular to the planes of the electrodes of the structure. This gives a volume free energy density of the form

$$f=\frac{1}{2}K_1\cos^2\theta\left(\frac{d\theta}{dx}\right)^2-\frac{1}{2}\varepsilon_a\varepsilon_0E_{LC}^2\sin^2\theta. \tag{13}$$

With the given simplifications, the Euler–Lagrange equation becomes

$$\frac{d^2\theta}{dx^2}-\left(\frac{d\theta}{dx}\right)^2\tan\theta+\frac{\varepsilon_a\varepsilon_0E_{LC}^2}{K_1}\tan\theta=0. \tag{14}$$

Numerical estimates show that in our case the third term can be neglected in comparison with the first two. The solution of the resulting differential equation is

$$\theta(x)=\theta_0\left(1-\frac{2x}{d}\right), \tag{15}$$

where $\theta_0=\theta|_{x=0}$.

To determine the boundary value θ_0 we solved Eqs. (5)–(7) jointly. This gave us the dependence of θ_0 on the LC layer thickness d (Fig. 1). The following parameters of the structure components were used in the calculations: $\varphi_m=4.4$ eV (SnO_2), $\varepsilon_\perp=7$, $\varepsilon_a=10$, $\Phi=4.2$ eV (Ref. 20). As we see from the plotted curve, at small LC layer thicknesses ($d<1$ μm) the director orientation is nearly planar, but when the thickness is increased to tens of microns a homeotropic orientation is observed. We are in fact talking about an orientational phase transition analogous to the Fréedericksz transition, but in the absence of external field (we first predicted this effect for an M–LC–S structure[7]). The function $\theta(x)$ obtained above can now be used to determine the surface energy F_0. Using the boundary conditions (11) and (12) (in view of the symmetry of the structure these are the same except for a sign), we find from Eq. (15) that

$$F_0=K_1\int\frac{\sin(2\theta_0)}{d}d\theta_0. \tag{16}$$

Since the boundary angle θ_0 depends on the thickness of the layer, the integration cannot be done analytically. Figure 2 shows the function $F_0(\theta_0)$ which we obtained by numerical methods. It is in satisfactory agreement with the experimental curves obtained by other authors.[21,22] Let us discuss the result. It is clear from the Fig. 2 that the function $\Phi_0(\theta_0)$ cannot be approximated by a Rapini potential (the dashed curve in Fig. 2). It seems reasonable to explain the differences as follows: it has been noted in several papers[7,23] that a polar layer with a nonzero polarization vector P, which in general depends on the coordinate, arises in an LC near a solid boundary surface. There are two possible causes for a polarized state: ionic polarization and dipole polarization. The presence of a polar layer leads to the following contribution to the free energy of the LC:

$$f_E(x)=\frac{1}{2}P(x)E_{LC}(x), \tag{17}$$

where $E_{LC}(x)$ is the electric field calculated earlier.

With the use of (17) the surface energy due to the mechanism under discussion is

$$f_s=\int f_E(x)dx. \tag{18}$$

The resulting surface energy can be written as

$$F_0=\frac{1}{2}W\sin^2\theta+f_s. \tag{19}$$

To find the contribution f_s to the surface energy we need to know the coordinate dependence of the polarization $P(x)$.

Let us first consider the dipole polarization mechanism, which is more characteristic of an LC. Suppose that the LC molecule has an intrinsic dipole moment p and that the den-

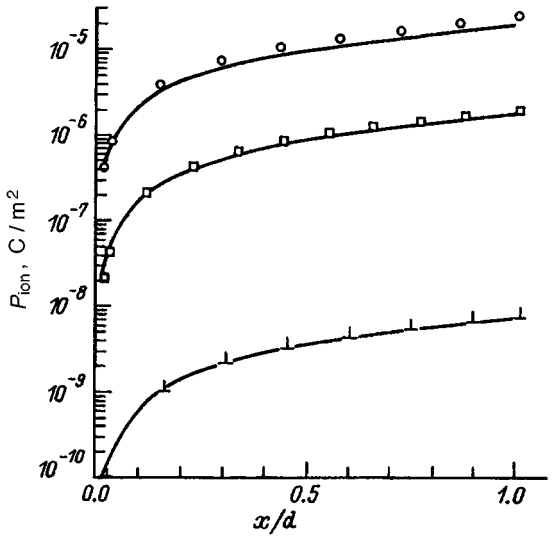

FIG. 3. Coordinate dependence of the ionic polarization. LC layer thickness $d=0.35$ (○), 0.9 (□), 13 μm (⊥).

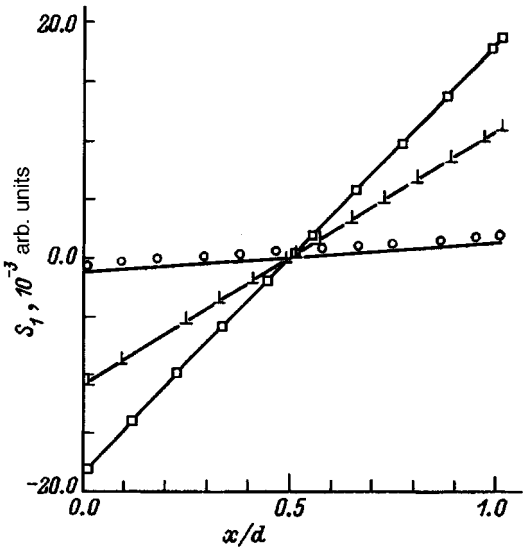

FIG. 4. Coordinate dependence of the order parameter. The LC layer thicknesses are the same as in Fig. 3.

sity of molecules is N. To take into account the degree of ordering of the dipole moments, we introduce, following Ref. 24, an order parameter $S_1=\langle P_1(\sin\theta)\rangle$, where P_1 is the first-order Legendre polynomial. Then we obtain the following expression for the polarization vector:

$$P_{\text{dip}}(x)=pNS_1(x). \tag{20}$$

The order parameter S_1 is a function of the coordinate x because of the presence of a nonuniform ordering electric field $E_{LC}(x)$, and the order parameter takes on its maximum (in modulus) values at the boundaries of the layer ($x=0$ and $x=d$), while $S_1=0$ at the center of the layer ($x=d/2$). Let us assume that

$$S_1(x)=E_xB, \tag{21}$$

where B is a certain quantity that depends on the thickness of the LC layer. Using this, we obtain for the dipole component of the polarization

$$P_{\text{dip}}(x)=pNBE_{LC}(x). \tag{22}$$

The second polarization mechanism, ionic polarization, is due to the nonuniformity of the charge distribution over the coordinate in the LC layer. It is clear that

$$P_{\text{ion}}(x)=\int_0^x \rho(x)dx=\int_0^x dx/\cos^2\left((Cn_0/2)^{1/2}\left(x-\frac{d}{2}\right)\right). \tag{23}$$

Here we have made use of the fact that the charge density $\rho(x)=qn(x)$. Figure 3 shows a plot of our calculated values of $P_{\text{ion}}(x)$ for different thicknesses of the LC layer.

To determine the dipole component of the polarization we substitute expression (22) together with (3) into Eqs. (17)–(19) and, on the basic of the dependence $F_0(\theta_0)$ found earlier, we find the parameter B, which we use to calculate the order parameter S_1 and the dipole component of the polarization P_{dip}. Figure 4 shows $S_1(x)$ for various thicknesses of the LC layer, and Fig. 5 shows how the boundary value of the dipole polarization $P_{\text{bndry}}=P_{\text{dip}}|_{x=0}$ depends on the layer thickness d. Let us discuss this last dependence in greater detail. The trend of this curve can be explained on the basis of the following considerations.

Near the surface bounding the LC there is a competition between two mechanisms. The quadrupole (nematic) order tries to orient the molecules parallel to the surface. At the same time, the internal electric field orients the molecules perpendicular to the substrate. As a result, a completely determined orientation is established, which corresponds, as we have said, to the minimum of the free energy (Fig. 1). In addition, as we have also mentioned, the dipole polarization is determined by the order parameter S_1, which in turn is proportional to the orienting field. At sufficiently large layer thicknesses ($d>10$ μm) the ordering effect of the electric field is substantially weakened (Fig. 6), leading to a decrease in the dipole polarization. This hypothesis is confirmed by

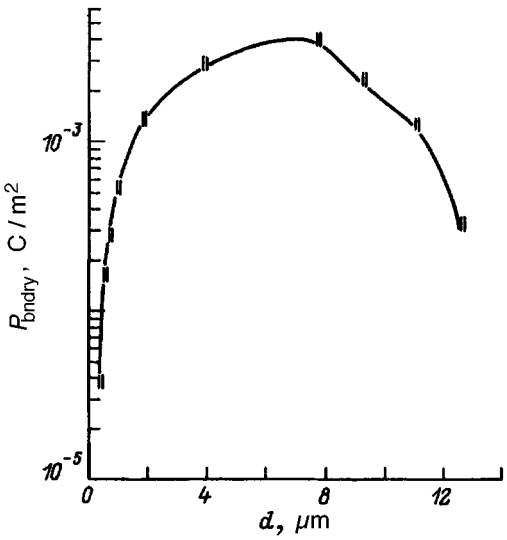

FIG. 5. Boundary value of the dipole polarization versus the thickness of the LC layer.

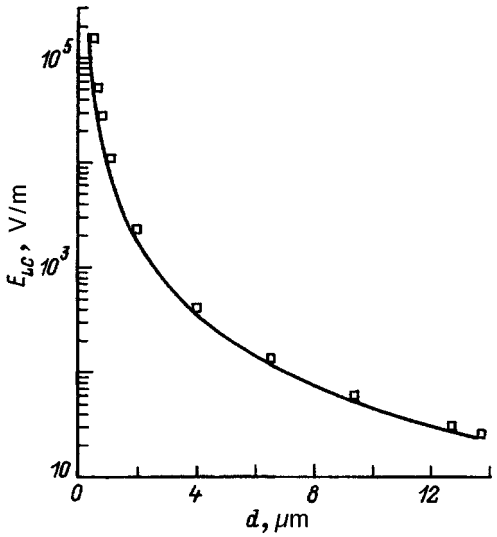

FIG. 6. Boundary value of the internal electric field versus the thickness of the LC layer.

the fact that the orientation correlation effects due to the influence of the substrate vanish in thick samples.

We note in conclusion that, although it was a metal–NLC–metal structure that was studied here, our results can also be used for analysis of other systems, in particular, for metal–LC–semiconductor systems, which have considerably broader applications.

The authors thank G. G. Plotnikov for his ongoing interest in this research.

This study was supported by the Russian Fund for Fundamental Research, Project No. 95–01-00531A.

[1] F. V. Losovskiĭ, E. G. Mansvetova, E. P. Nikolaeva, and A. V. Nikolaev, Zh. Éksp. Teor. Fiz. **103**, 213 (1993) [J. Exp. Theor. Phys. **76**, 116 (1993)].

[2] V. P. Voronov and F. Kholmurodov, Zh. Éksp. Teor. Fiz. **95**, 1324 (1989) [Sov. Phys. JETP **68**, 763 (1989)].

[3] Yu. S. Barash and O. I. Notych, Zh. Éksp. Teor. Fiz. **98**, 542 (1990) [Sov. Phys. JETP **71**, 301 (1990)].

[4] Yu. K. Kornienko and A. P. Fedchuk, Pis'ma Zh. Tekh. Fiz. **18**(12), 78 (1992) [Sov. Tech. Phys. Lett. **18**, 397 (1992)].

[5] A. E. Alekseev, Yu. K. Kornienko, L. D. Shevchenko, and A. P. Fedchuk, Pis'ma Zh. Tekh. Fiz. **16**(15), 77 (1990) [Sov. Tech. Phys. Lett. **16**, 593 (1990)].

[6] A. V. Sukhov, Zh. Éksp. Teor. Fiz. **98**, 919 (1990) [Sov. Phys. JETP **71**, 512 (1990)].

[7] A. N. Lachinov and A. N. Chuvyrov, Poverkhnost', No. 5, pp. 74–49 (1986).

[8] A. N. Lachinov and A. N. Chuvyrov, Kristallografiya **25**, 4 (1980) [*sic*.]

[9] L. G. Fel and E. B. Loginov, Zh. Éksp. Teor. Fiz. **103**, 1262 (1993) [J. Exp. Theor. Phys. **76**, 618 (1993)].

[10] L. M. Blinov, E. I. Kats, and A. A. Sonin, Usp. Fiz. Nauk **152**, 449 (1987) [Sov. Phys. Usp. **30**, 604 (1987)].

[11] Yu. K. Kornienko and A. P. Fedchuk, Zh. Tekh. Fiz. **64**(7), 1 (1994) [Tech. Phys. **39**, 633 (1994)].

[12] M. A. Osipov, Poverkhnost', No. 9, pp. 39–47 (1988).

[13] G. Barbero, N. V. Madhusudana, and G. Durand, Z. Naturforsch. Teil A. **39**, 1066. (1984).

[14] M. I. Barnik, L. M. Blinov, T. V. Korkishko *et al.*, Zh. Éksp. Teor. Fiz. **85**, 176 (1983) [Sov. Phys. JETP **58**, 102 (1983)].

[15] G. Barbero, N. V. Madhusudana, J. F. Palierne, and G. Durand, Phys. Lett. A. **103**, 385. (1984).

[16] G. Barbero and G. Durand, J. Phys. (Paris) **47**, 2129 (1986).

[17] Yu. K. Kornienko, S. F. Baranov, and I. V. Zaginaĭlo, Ukr. Fiz. Zh. **31**, 1528 (1986).

[18] Yu. K. Kornienko and A. P. Fedchuk, Zh. Tekh. Fiz. **62**(2), 140 (1992) [Sov. Phys. Tech. Phys. **37**, 191 (1992)].

[19] S. A. Pikin, *Structural Transformations in Liquid Crystals* [in Russian], Nauka, Moscow (1981).

[20] I. V. Zaginaĭlo, Candidate's Dissertation [in Russian], Odessa (1987).

[21] S. Valenti, M. Grillo, G. Barbero, and B. Taverna Valabrega, Europhys. Lett. **12**, 407 (1990).

[22] V. G. Podoprigora, L. A. Remizov, and V. F. Shabanov, Mater. Sci. **14**, 83 (1988).

[23] O. A. Skaldin, A. P. Krekhov, and A. N. Chuvyrov, Kristallografiya **34**, 951 (1989) [Sov. Phys. Crystallogr. **34**, 572 (1989)].

[24] T. V. Korkishko, V. G. Chigrinov, R. V. Galiulin *et al.*, Kristallografiya **32**, 452 (1987) [Sov. Phys. Crystallogr. **32**, 263 (1987)].

Translated by Steve Torstveit

Influence of heterogeneous chemical reactions on the slip velocity of an inhomogeneous multicomponent gaseous mixture

E. I. Alekhin, I. N. Golovkina, and Yu. I. Yalamov

Moscow Pedagogical University, 107005 Moscow, Russia
(Submitted December 4, 1995)
Zh. Tekh. Fiz. **67**, 29–33 (May 1997)

In constructing a theory of thermal diffusiophoresis of volatile aerosol particles it is necessary to have boundary conditions for the tangential velocity component which allow for the presence of heterogeneous chemical reactions. Conditions of this sort have been obtained by a number of authors [R. N. Gupta *et al.*, *Technical Papers, AIAA 22nd Aerospace Science Meeting, AIAA 19th Thermophysics Conference*, New York (1985), pp. 465–490; D. V. Kul'ginov, Tech. Phys. **63**, 940 (1993); A. V. Bogdanov *et al.*, Preprint No. 1051, Fiz-Tekh. Inst. Akad. Nauk. SSSR, Leningrad (1986)]. The results of Gupta *et al.* are in the form of analytical expressions, but their computations actually used Maxwell's method, which is of low accuracy. Kul'ginov *et al.* and Bogdanov *et al.* used the method of matched asymptotic expansions, which did not permit them to get simple analytical expressions. In the present paper the slip velocity is calculated by the Loyalka method. Analytical expressions are obtained for the slip coefficients, and the results of numerical calculations are presented. It is shown that in the presence of concentration gradients of the components of the gaseous mixture along the surface of an aerosol particle, the slip velocity can acquire new terms due to the change in the sticking coefficient along the catalytic surface. Expressions in final form are given for these terms. © *1997 American Institute of Physics.* [S1063-7842(97)00605-3]

STATEMENT OF THE PROBLEM

The derivation of macroscopic boundary conditions for the equations of gasdynamics must be based on the solution of the kinetic equations in the Knudsen layer.[4–8]

Let us consider an N-component gaseous mixture, inhomogeneous in temperature and concentration, moving near a plane catalytic surface. We choose the x axis normal to this surface and the y axis along the specified gradients of the temperature of the mixture (dT/dy) and of the relative concentrations of the components (dc_i/dy). The gas fills the half space $x>0$.

For describing the flow of the gaseous mixture in the Knudsen layer we use the Boltzmann equation[9]

$$\mathbf{v}_i\cdot\nabla f_i=\sum_j \int (f_i' f_j'-f_i f_j) g b\, db\, d\varepsilon\, d^3 v_i, \tag{1}$$

where f_i is the distribution function of the molecules of the ith component, $\mathbf{g}$ is the relative velocity of approach of two molecules, b is the impact parameter of the collision, ε is the azimuthal angle, and $\mathbf{v}_i$ is the velocity of a molecule.

In practice one ordinarily deals with mixtures in which the relative drops in temperature, concentrations of the components, and the mean-mass velocity over one mean free path are much less than 1. The heterogeneous reaction rate is limited by the slow processes of diffusion of the mixture components toward the surface. Under these conditions one can linearize the distribution function f_i about a local Maxwellian distribution function:

$$f_i=f_{i0}(1+\varphi_i), \tag{2}$$

where

$$f_{i0}=n_i\left(\frac{m_i}{2\pi kT}\right)^{3/2}\exp\left(-\frac{m_i V_i^2}{2kT}\right), \tag{3}$$

and φ_i is a small correction to the Maxwellian distribution function, proportional to the first powers of the gradients of the velocity, temperature, and concentration.

Substituting (2) into (1), we obtain the following system of linearized kinetic equations:

$$v_{xi}f_{i0}\frac{d\varphi_i}{dx}+v_{yi}f_{i0}\left(\frac{1}{c_i}\left(\frac{dc_i}{dy}\right)-S_{3/2}^1\left(\frac{d\ \ln\ T}{dy}\right)\right)$$

$$=-\sum_j n_i n_j I_{ij}(\varphi_i+\varphi_j), \tag{4}$$

where

$$I_{ij}(\varphi_i+\varphi_j)=\frac{1}{n_i n_j}\times\int f_{i0}f_{j0}(\varphi_i+\varphi_j-\varphi_i'-\varphi_j')g b\, db\, d\varepsilon\, d^3 v_i$$

is the linearized collision integral,[9] $S_{3/2}^1=5/2-w_i^2$ is the Sonin polynomial of order 3/2 (Ref. 9), and $w_i=\sqrt{m_i/2kT}v_i$ is the reduced velocity of the molecules.

CALCULATION OF THE SLIP COEFFICIENTS

To determine the slip velocity one must solve the system of equations (4) with the kinetic boundary conditions. At large distances from the surface the solution of system (4) has the form

$$\varphi_{i\infty}=\left(\frac{2m_i}{kT}\right)^{1/2}w_{yi}U(x)+\varphi_i^{CE}(w_i), \tag{5}$$

 1063-7842/97/050482-05$10.00

where $\varphi_i^{CE}(\omega_i)$ is the Chapman–Enskog function:[9]

$$\varphi_i^{CE}(w_i)=\frac{1}{n}w_{yi}A_i(w_i)\left(\frac{d\ \ln\ T}{dy}\right)+\frac{1}{n}w_{yi}\sum_j\ C_i^j(w_i)\times\left(\frac{dc_j}{dy}\right)+\frac{1}{n}w_{xi}w_{yi}B_i(w_i)\left(\frac{dU}{dx}\right). \tag{6}$$

Here $A_i(w_i)$, $C_i^j(w_i)$, and $B_i(w_i)$ are the thermal, diffusion, and viscosity functions;[9] $U(x=0)$ is slip velocity to be solved for. In the distribution function (6) we have kept only terms that give a nonzero contribution to the slip velocity. Solving the equation by the method of Loyalka,[10] we assume that the distribution of the incident molecules on the lower boundary of the Knudsen layer $(x=0)$ can be written in the following form:

$$\varphi_i^-=\left(\frac{2m_i}{kT}\right)^{1/2}w_{yi}\beta+\frac{1}{n}w_{yi}A_i(w_i)\left(\frac{d\ \ln\ T}{dy}\right)+\frac{1}{n}w_{yi}\sum_j\ C_i^j(w_j)\left(\frac{dc_j}{dy}\right)+\frac{1}{n}w_{xi}w_{yi}B_i(w_i)\left(\frac{dU}{dx}\right), \tag{7}$$

where β is some unknown quantity.

The distribution function of the molecules coming off the catalytic surface in the presence of heterogeneous chemical reactions is written as the sum[11]

$$f_i^+=f_i'+f_i'', \tag{8}$$

where the first term on the right-hand side of (8) describes molecules formed as a result of a chemical reaction, while the second term describes molecules reflected off the phase boundary.

We shall assume that the molecules formed as a result of chemical reactions leave the surface in a diffuse manner; then f_i' can be written as[11]

$$f_i'=\alpha_i^-\,n_i'\left(\frac{m_i}{2\pi kT}\right)^{3/2}\exp\left(-\frac{m_iV_i^2}{2kT}\right)=\alpha_i^-(1+\nu_i')f_{i0}, \tag{9}$$

where n_i' is the density of molecules formed as a result of the chemical reaction, $\nu_i'=(n_i'-n_{i0})/n_{i0}$, and α_i^- is the sticking coefficient, which characterizes the intensity of adsorption processes. It represents the probability that a molecule, having struck the catalytic surface, is not reflected back into the gas.

The distribution function of the reflected molecules is taken as

$$f_i''=(1-\alpha_i^-)q_in_{is}\left(\frac{m_i}{2\pi kT_{is}}\right)^{3/2}\exp\left(-\frac{m_iV_i^2}{2kT_{is}}\right)+(1-\alpha_i^-)(1-q_i)f_i^-. \tag{10}$$

Here the first term describes molecules reflected diffusely, while the second describes molecules reflected specularly; q_i is the accomodation coefficient of the tangential momentum, T_{is} and n_{is} are the temperature and density of the diffusely reflected molecules of the ith component. Linearizing the distribution function (8) about f_{i0} with allowance for (9) and (10), we get

$$f_i^+=f_{i0}(1+\varphi_i^+), \tag{11}$$

where

$$\varphi_i^+=\alpha_i^-\,\nu_i'+(1-\alpha_i')q_i(\nu_{is}-S_{1/2}^1\tau_{is})+(1-\alpha_i^-)(1-q_i)\hat{R}_x\varphi_i^-, \tag{12}$$

$S_{1/2}^1=3/2-w_i^2$ is a Sonin polynomial of order 1/2 (Ref. 9), $\hat{R}_x$ is an operator which changes the sign of the xth component of the velocity on reflection from the surface, $\nu_{is}=(n_{is}-n_{i0})/n_{i0}$, and $\tau_{is}=(T_{is}-T)/T$.

Multiplying Eqs. (4) by m_iw_i, $B_iw_{xi}w_{yi}$, integrating over all velocity space, and summing over i, we get

$$\frac{d}{dx}\sum_i\ (w_{xi}w_{yi},\varphi_i)_i=0,$$

$$\frac{d}{dx}\sum_i\ \sqrt{\frac{2kT}{m_i}}(w_{xi}^2w_{yi}B_i,\varphi_i)_i=\frac{2}{n}\sum_i\ m_i(w_{xi}w_{yi},\varphi_i)_i,$$

where the parentheses denote moments of the distribution function:

$$(P(w_{xi},w_{yi}),\varphi_i)_i=\int P(w_{xi},w_{yi})f_{i0}\varphi_i d^3v_i.$$

After integrating over x, we obtain on the interfacial surface between phases $(x=0)$

$$\sum_i\ (w_{xi}w_{yi},\varphi_i)_i=\text{const},$$

$$\sum_i\ \sqrt{\frac{2kT}{m_i}}(w_{xi}^2w_{yi}B_i,\varphi_i)_i=\text{const}. \tag{13}$$

Expression (13) is a system of linear algebraic equations for β and U. In evaluating the integrals that appear in Eq. (13), we keep (as usual) two terms in the expansions of the thermal and diffusion functions in Sonin polynomials, and one term in the expansion of the viscosity functions:

$$A_i\approx w_ia_{i0}+w_ia_{i1}S_{3/2}^1,$$

$$C_i^j\approx w_ic_{i0}^j+w_ic_{i1}^jS_{3/2}^1,\quad \hat{B}_i\approx w_i^0w_ib_{i0}.$$

Solving the system of equations (13), we find the slip velocity U. We write it in the following form:

$$U=\frac{\eta}{\rho}K_{Ts}\left(\frac{d\ \ln\ T}{dy}\right)+\sum_j\ D_{1j}K_{Ds}^j\left(\frac{dc_j}{dy}\right)+\lambda K_{Ms}\left(\frac{dU}{dx}\right), \tag{14}$$

where η/ρ is the viscosity of the gas, D_{1j} is the coefficient of multicomponent diffusion of the 1st and jth components; λ is the mean free path of the gas molecules, K_{Ts}, K_{Ds}^j, and K_{Ms} are the coefficients of thermal, diffusion, and isothermal slip:

$$\eta/\rho K_{Ts}=-\frac{1}{n}\sqrt{\frac{kT}{8}}\frac{1}{\Sigma_in_ib_{i0}}\times\left\{\Omega_{Ts}^1(\alpha_i^-)+\frac{\Omega_{Ts}^2(\alpha_i^2)\Omega_{Ts}^3(\alpha_i^-)}{\Omega_{Ts}^4(\alpha_i^-)}\right\}, \tag{15}$$

$$D_{1j}K_{Ds}^{j}=-\frac{1}{n}\sqrt{\frac{kT}{8}}\frac{1}{\Sigma_i n_i b_{i0}}\times\left\{\Omega_{Ds}^{1}(\alpha_i^-)+\frac{\Omega_{Ds}^{2}(\alpha_i^-)\Omega_{Ds}^{3}(\alpha_i^-)}{\Omega_{Ds}^{4}(\alpha_i^-)}\right\},\qquad(16)$$

$$\lambda K_{Ms}=-\frac{1}{n}\sqrt{\frac{kT}{2\pi}}\frac{1}{\Sigma_i n_i b_{i0}}\left\{\Omega_{Ms}^{1}(\alpha_i^-)+\frac{(\Omega_{Ms}^{2}(\alpha_i^-))^2}{\Omega_{Ms}^{3}(\alpha_i^-)}\right\},\qquad(17)$$

where

$$\Omega_{Ts}^{1}(\alpha_i^-)=\sum_i\frac{n_i}{\sqrt{m_i}}(a_{i0}-a_{i1})b_{i0}(1-(1-q_i)(1-\alpha_i^-)),$$

$$\Omega_{Ts}^{2}(\alpha_i^-)=\sum_i n_i b_{i0}(1+(1-q_i)(1-\alpha_i^-)),$$

$$\Omega_{Ts}^{3}(\alpha_i^-)=\sum_i n_i\left(a_{i0}-\frac{a_{i1}}{2}\right)(1-(1-q_i)(1-\alpha_i^-)),$$

$$\Omega_{Ts}^{4}(\alpha_i^-)=\Omega_{Ds}^{4}(\alpha_i^-)=\sum_i\sqrt{m_i}n_i(1-(1-q_i)(1-\alpha_i^-)),$$

$$\Omega_{Ds}^{1}(\alpha_i^-)=\sum_i\frac{n_i}{\sqrt{m_i}}(c_{i0}^{j}-c_{i1}^{j})b_{i0}(1-(1-q_i)(1-\alpha_i^-)),$$

$$\Omega_{Ds}^{2}(\alpha_i^-)=\sum_i n_i b_{i0}(1+(1-q_i)(1-\alpha_i^-)),$$

$$\Omega_{Ds}^{3}(\alpha_i^-)=\sum_i n_i\left(c_{i0}^{j}-\frac{c_{i1}^{j}}{2}\right)(1-(1-q_i)(1-\alpha_i^-)),$$

$$\Omega_{Ms}^{1}(\alpha_i^-)=\sum_i\frac{n_i b_{i0}}{\sqrt{m_i}}(1+(1-q_i)(1-\alpha_i^-)),$$

$$\Omega_{Ms}^{2}(\alpha_i^-)=\sum_i n_i b_{i0}(1+(1-q_i)(1-\alpha_i^-)),$$

$$\Omega_{Ms}^{3}(\alpha_i^-)=\frac{4}{\pi}\sum_i\sqrt{m_i}n_i(1-(1-q_i)(1-\alpha_i^-)).$$

In an inhomogeneous gaseous mixture there are concentration and temperature gradients along the surface of a particle. For this reason the sticking coefficient is not constant, and that means that there is an additional term in the slip velocity. Let us calculate this term. We assume that the relative drops in temperature and concentration along the surface are small. Taking into account that the sticking coefficient α_i^- depends on the concentrations of the mixture components and on the surface temperature, and keeping only the terms of zeroth and first orders, we obtain

$$\alpha_i^-=\alpha_{i0}^-(c_{i0},\ldots,c_{N0},T_0)+\frac{\partial\alpha_i^-}{\partial c_1}\Delta c_1+\ldots+\frac{\partial\alpha_i^-}{\partial c_N}\Delta c_N+\frac{\partial\alpha_i^-}{\partial T}\Delta T,\qquad(18)$$

where α_{i0}^- is the sticking coefficient at a certain point on the surface.

Substituting (18) into (15)–(17), we get

$$K_{Ts}=K_{Ts}^{0}+K_{Ts}^{1},$$

$$K_{Ds}^{1}=K_{Ds}^{j0}+K_{Ds}^{j1},$$

$$K_{Ms}=K_{Ms}^{0}+K_{Ms}^{1},$$

where K_{Ts}^{0}, K_{Ds}^{j0}, K_{Ms}^{0} are obtained from formulas (15)–(17) with α_{i0}^- in place of α_i^-.

The coefficients K_{Ts}^{1}, K_{Ds}^{j0}, and K_{Ms}^{1} describe the first-order correction to the slip velocity due to the nonuniformity of the coefficients α_i^- along the catalytic surface and have the following form:

$$K_{Ts}^{1}=-\frac{1}{n}\sqrt{\frac{kT}{8}}\frac{1}{\Sigma_i n_i b_{i0}}\times\left\{\Delta\Omega_{Ts}^{1}+\frac{\Omega_{Ts}^{2}(\alpha_{i0}^-)\Delta\Omega_{Ts}^{3}+\Omega_{Ts}^{3}(\alpha_{i0}^-)\Delta\Omega_{Ts}^{2}-\Omega_{Ts}^{2}(\alpha_{i0}^-)\Omega_{Ts}^{3}(\alpha_{i0}^-)\Delta\Omega_{Ts}^{4}/\Omega_{Ts}^{4}(\alpha_{i0}^-)}{\Omega_{Ts}^{4}(\alpha_{i0}^-)}\right\},$$

$$K_{Ds}^{j1}=-\frac{1}{n}\sqrt{\frac{kT}{8}}\frac{1}{\Sigma_i n_i b_{i0}}\times\left\{\Delta\Omega_{Ds}^{1}+\frac{\Omega_{Ds}^{2}(\alpha_{i0}^-)\Delta\Omega_{Ds}^{3}+\Omega_{Ds}^{3}(\alpha_{i0}^-)\Delta\Omega_{Ds}^{2}-\Omega_{Ds}^{2}(\alpha_{i0}^-)\Omega_{Ds}^{3}(\alpha_{i0}^-)\Delta\Omega_{Ds}^{4}/\Omega_{Ds}^{4}(\alpha_{i0}^-)}{\Omega_{Ds}^{4}(\alpha_{i0}^-)}\right\},$$

$$K_{Ms}^{1}=-\frac{1}{n}\sqrt{\frac{kT}{2\pi}}\frac{1}{\Sigma_i n_i b_{i0}}\left\{\Delta\Omega_{Ms}^{1}+\frac{2\Omega_{Ms}^{2}(\alpha_{i0}^-)\Delta\Omega_{Ms}^{2}-(\Omega_{Ms}^{2}(\alpha_{i0}^2))^2\Delta\Omega_{Ms}^{3}/\Omega_{Ms}^{3}(\alpha_{i0}^-)}{\Omega_{Ms}^{3}(\alpha_{i0}^-)}\right\},$$

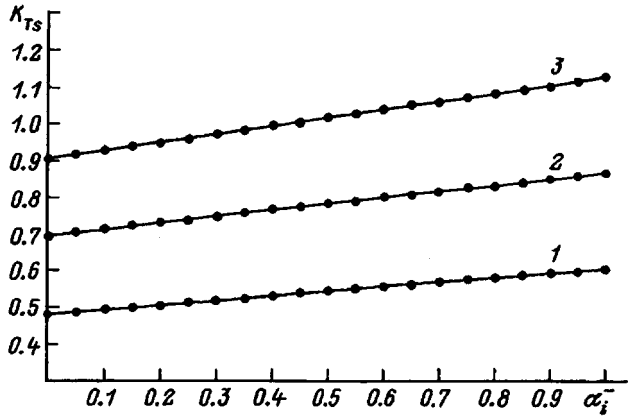

FIG. 1. Coefficients $K_{Ts}(\alpha_i^-)$, $q_{1,2}=0.5$, $R_1=0.9$. $M_i=0.1$ (*1*), 0.3 (*2*), 0.5 (*3*).

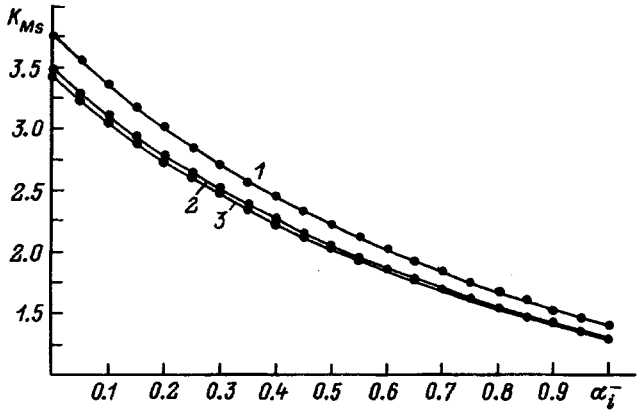

FIG. 2. Coefficients $K_{Ms}(\alpha_i^-)$, $q_{1,2}=0.5$, $R_1=0.9$. Curves *1–3* are the same as in Fig. 1.

where

$$\Delta\Omega_{Ts}^1=\sum_i \frac{n_i}{\sqrt{m_i}}(a_{i0}-a_{i1})b_{i0}(1-q_i)\Delta\alpha_i,$$

$$\Delta\Omega_{Ts}^2=\sum_i n_i b_{i0}(q_i-1)\Delta\alpha_i,$$

$$\Delta\Omega_{Ts}^3=\sum_i n_i\left(a_{i0}-\frac{ai1}{2}\right)(1-q_i)\Delta\alpha_i,$$

$$\Delta\Omega_{Ts}^4=\Delta\Omega_{Ds}^4=\sum_i \sqrt{m_i}n_i(1-q_i)\Delta\alpha_i,$$

$$\Delta\Omega_{Ds}^1=\sum_i \frac{n_i}{\sqrt{m_i}}(c_{i0}^j-c_{i1}^j)b_{i0}(1-q_i)\Delta\alpha_i,$$

$$\Delta\Omega_{Ds}^2=\sum_i n_i b_{i0}(q_i-1)\Delta\alpha_i,$$

$$\Delta\Omega_{Ds}^3=\sum_i n_i\left(c_{i0}^j-\frac{c_{i1}^j}{2}\right)(1-q_i)\Delta\alpha_i,$$

$$\Delta\Omega_{Ms}^1=\sum_i \frac{n_i b_{i0}^2}{\sqrt{m_i}}\Delta\alpha_i,\quad \Delta\Omega_{Ms}^2=\sum_i n_i b_{i0}(q_i-1)\Delta\alpha_i,$$

$$\Delta\Omega_{Ms}^3=\frac{4}{\pi}\sum_i \sqrt{m_i}n_i(1-q_i)\Delta\alpha_i,$$

$$\Delta\alpha_i=\frac{\partial\alpha_i^-}{\partial c_1}\Delta c_1+\ldots+\frac{\partial\alpha_i^-}{\partial c_N}\Delta c_N+\frac{\partial\alpha_i^-}{\partial T}\Delta T.$$

ANALYSIS OF THE RESULTS

The results of the numerical calculation of the slip coefficients for a binary gaseous mixture are plotted in Figs. 1, 2, and 3 [$M_1=m_1/(m_1+m_2)$, $R_1=r_1(r_1+r_2)$]. Analysis of these results shows that the coefficient of thermal slip K_{Ts} for fixed m_i and n_i increases monotonically with increasing α_i^-. This can be explained as follows: when α_i^- increases, so does the fraction of the molecules adsorbed by the surface; this entails an increase in the momentum transfer to the surface.

The isothermal slip coefficient K_{Ms} decreases monotonically with increasing α_i^-. This is explained by the fact that as α_i^- decreases, the fraction of the molecules reflected specularly from the surface increases, and in the presence of a velocity gradient near the surface the gas must move faster in the case of specular reflection of the molecules than for diffuse reflection, under otherwise equal conditions. The dependence of the coefficient of isothermal slip on the transverse momentum accomodation coefficient q_i is analogous to the dependence on α_i^-, as is particularly easy to notice for $\alpha_i^-=0$.

The dependence of K_{Ds} on α_i^- at fixed m_i and n_i is monotonic (functions (15) and (16) have the same form). As we see from the plots, the the coefficient of diffusion slip can change sign as α_i^- varies. Hence the rate of diffusiophoresis of aerosol particles having high thermal conductivity can change direction, since in the course of the chemical reaction the concentration of the components near the catalytic surface can vary, and α_i^- along with it.

In solving the problem of the slip of a multicomponent gaseous mixture in the presence of evaporation and conden-

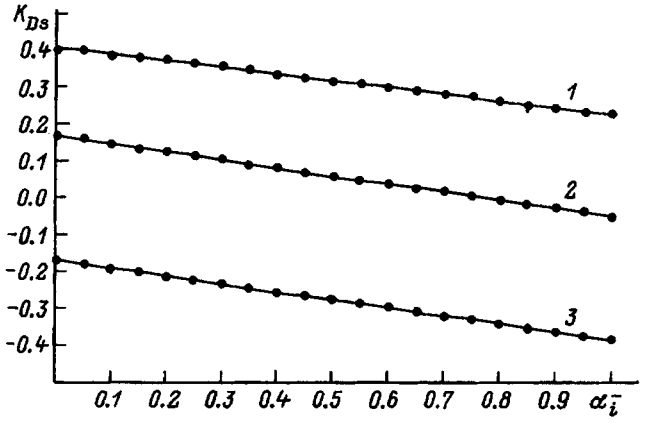

FIG. 3. Coefficients $K_{Ds}(\alpha_i^-)$, $q_{1,2}=0.5$, $R_1=0.9$. Curves *1–3* are the same as in Fig. 1.

sation processes, despite the difference in the form of the kinetic boundary conditions (the functions f_i'), the expressions for the slip coefficients come out the same as in the case of heterogeneous chemical reactions. Formally, the transition to the case of evaporation and condensation processes can be made by setting $K_{Ts}^1=0$, $K_{Ds}^{j1}0$, $K_{Ms}^1=0$ and equating the sticking coefficient α_i^- to the condensation coefficient.

In the absence of heterogeneous chemical reactions (sticking coefficient $\alpha_i^-=0$) formulas (14)–(17) agree with the expression for the slip velocity obtained in Ref. 12. For $\alpha_i^-=1$ the slip velocity is independent of the accomodation coefficient of the tangential momentum.

[1] R. N. Gupta, C. D. Scott, and J. N. Moss, *Technical Papers, AIAA 22nd Aerospace Science Meeting, AIAA 19th Thermophysics Conference*, New York (1985), pp. 465–490.

[2] D. V. Kul'ginov, Zh. Tekh. Fiz. **63**(11), 8 (1993) [Tech. Phys. **38**, 940 (1993)].

[3] A. V. Bogdanov, Yu. E. Gorbachev, D. V. Kul'ginov, and V. A. Pavlov, Preprint No. 1051 [in Russian], Fiz-Tekh. Inst. Akad. Nauk SSSR, Leningrad (1986), 36 pp.

[4] Yu. I. Yalamov, I. N. Ivchenko, and B. V. Deryagin, Dokl. Akad. Nauk SSSR **180**, 330 (1968) [Sov. Phys. Dokl. **13**, 446 (1968)].

[5] Yu. I. Yalamov, I. N. Inchenko, and V. B. Deriaguin, *Rarefied Gas Dynamics. VI International Symposium on Rarefied Gas Dynamics* (1969), Vol. 1, pp. 295–300.

[6] G. V. Deriaguin and Yu. I. Yalamov, *International Reviews in Aerosol Physics and Chemistry* (1972), Vol. 3, Pt. 2, pp. 1–200.

[7] I. I. Ivchenko and Yu. I. Yalamov, Izv. Akad. Nauk SSSR Mekh. Zhidk. Gaz., No. 4, pp. 22–29 (1971).

[8] Yu. I. Yalamov and A. A. Yushkanov, Dokl. Akad. Nauk SSSR **237**, 303 (1977) [Sov. Phys. Dokl. **22**, 655 (1977)].

[9] J. H. Ferziger and H. G. Kaper, *Mathematical Theory of Transport in Gases*, Elsevier–North-Holland, New York (1972) [Russ. transl., Mir, Moscow (1976)].

[10] S. K. Loyalka, Phys. Fluids **14**, 2291 (1971).

[11] E. I. Alekhin, Dep. VINITI No. 1698-B95 [Deposited ms, in Russian], Moscow (1995), 22 pp.

[12] E. I. Alekhin and Yu. I. Yalamov, *Mathematical Principles of Solving Boundary–Value Problems in the Kinetic Theory of Multicomponent Gases Near a Condensed Phase*, Instructional Materials for Advanced Course, N. K. Krupskaya Moscow District Pedagogical Institute (MOPI), 1991, 150 pp.

Translated by Steve Torstveit

CARS study of the vibrational kinetics of nitrogen molecules in the burning and afterglow stages of a pulsed discharge

K. A. Vereshchagin, V. V. Smirnov, and V. A. Shakhatov

Moscow Radio Engineering Institute, Russian Academy of Sciences, 113519 Moscow, Russia
(Submitted November 8, 1995)
Zh. Tekh. Fiz. **67**, 34–42 (May 1997)

The vibrational kinetics of the nitrogen molecule in the ground state $X^1\Sigma_g^+$ in the burning and afterglow stages of a pulsed discharge are investigated by coherent anti-Stokes Raman spectroscopy (CARS). The total cross section for vibrational excitation of the nitrogen molecule by electron impact to the first eight vibrational levels is determined. The rate constant for the associative ionization reaction involving nitrogen atoms in the metastable states ^2P and ^2D is estimated. It is found that the best agreement between the calculated and measured populations of the nitrogen molecules in the ground state $X^1\Sigma_g^+$ in the afterglow stage of a pulsed discharge is obtained when the rate constant for VV exchange K_{01}^{10} has the value predicted by the quantum–classical Billing–Fisher model. © *1997 American Institute of Physics.* [S1063-7842(97)00705-8]

INTRODUCTION

Pulsed high-current discharges in molecular nitrogen at medium pressures are widely used in technological processes. In optimizing operation of the devices employed one faces the problem of investigating the vibrational kinetics of the nitrogen molecules under conditions characterized by high densities of excited and charged particles at low gas temperatures.

Inelastic and superelastic collisions (collisions of the first and second kinds, respectively) of electrons with vibrationally excited molecules in the burning and afterglow regimes of the discharge play an important role in the formation of the electron energy distribution function (EEDF) and the distribution of the nitrogen molecules over vibrational levels ν in the ground state $X^1\Sigma_g^+$ (VDF).[1] The state of the electron component determines the integral characteristics of the gas discharge, such as the ionization and dissociation coefficients of the molecules, the drift velocity $v_{\rm dr}$, and the characteristic electron temperature D/μ.

The description of the integral characteristics in a highly nonequilibrium medium is made difficult by the lack of reliable data on the rate constants and cross sections of processes occurring in discharges in molecular nitrogen. It has been noted[2] that there is a large scatter in absolute values of the level cross sections σ_ν for vibrational excitation by electron impact for nitrogen molecules according to different measurements and calculations. For example, there is at least a fourfold disparity in the total cross section for vibrational excitation to the first eight vibrational levels v, i.e., $\sigma_\Sigma=\Sigma_{\nu=1}^{8}\sigma_\nu$ (at an electron energy $\varepsilon=2.3$ eV). The integral characteristics $v_{\rm dr}$ and D/μ of the electrons are rather insensitive to variations in the value of σ_Σ. Therefore, its determination in Ref. 2 from the measured data on $v_{\rm dr}$ and D/μ is only approximate. In our opinion a more systematic approach is to determine the value of σ_Σ for which the calculated and measured data for $v_{\rm dr}$, D/μ, and the VDF formed as a result of direct electron impact are in best agreement. The stage of excitation of the discharge consists of a set of kinetic processes that can influence the populations of the vibrational levels of the molecules and the EEDF. Thus a numerical modeling of the kinetics of vibrational excitation of the molecules for the purpose of determining σ_Σ must be included in a joint solution of the balance equations for the densities of the plasma components and the EEDF.

The rate constant K_{01}^{10} for VV exchange between vibrational levels $v=1$ and $v=0$ is an important quantity in the nonequilibrium vibrational kinetics of nitrogen molecules. The values of the rate constant measured in the afterglow of a discharge[3–5] differ from the results of Refs. 6 and 7 by as much as an order of magnitude. The accumulated experimental and theoretical data[1] on the kinetics of the processes occurring in the afterglow of a discharge permit one to carry out a more correct processing of the measurements of the VDF and its evolution in time.

In this paper we give the results of measurements made by coherent anti-Stokes Raman spectroscopy (CARS) and by numerical modeling of the VDF in pure molecular nitrogen in a pulsed high-current discharge and its afterglow. From a comparison of the results of the calculation and experiment we determine the values of the total cross section σ_Σ for vibrational excitation to the first eight vibrational levels and the rate constant K_{01}^{10} for VV exchange for the nitrogen molecule and we estimate the rate constant for the associative ionization reaction involving atoms in the metastable states N(^2P) and N(^2D).

DESCRIPTION OF THE EXPERIMENT

A diagram of the experiment is shown in Fig. 1. We investigated the nonequilibrium excitation of pure molecular nitrogen in a specially made cell at a pressure of 115 torr. The electrodes had a surface area of 45 mm^2 and were placed 1 cm apart. The experiment was done in a weak flow of nitrogen gas that had been purified in nitrogen traps. The parameter E/N (where E is the electric field and N is the density of gas molecules) in discharges with a duration of 40 ns (at half maximum of the current pulse) varied in the range

 1063-7842/97/050487-08$10.00

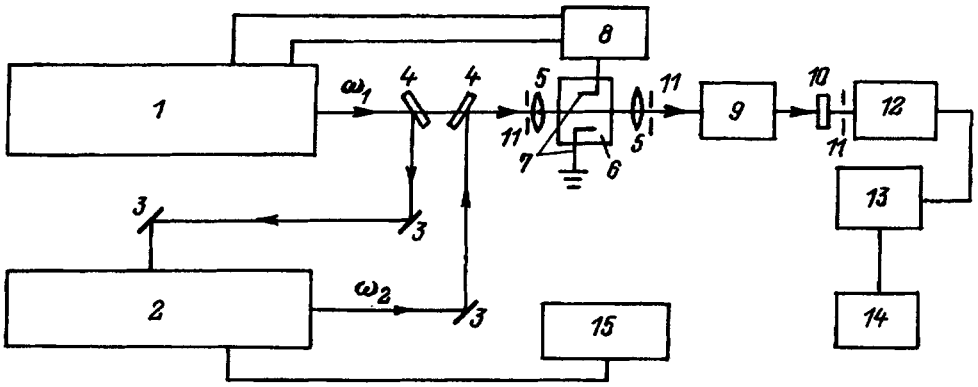

FIG. 1. Experimental apparatus for CARS spectroscopy of a pulsed high-current discharge. *1* — Nd:YAG laser, *2* — narrow-band dye laser, *3* — mirror, *4* — splitter, *5* — lens, *6* — discharge cell, *7* — electrode, *8* — synchronization unit, *9* — four-prism filter, *10* — wide-band filter, *11* — diaphragm, *12* — photomultiplier, *13* — integrator, *14* — chart recorder, *15* — control unit of dye laser.

190–200 Td. In the afterglow after the end of the current pulse, at $t>50$ ns, it had decreased to $E/N=5$–10 Td. The maximum value of the discharge current reached 1.4–1.5 kA. Figure 2 shows the time dependence of the electron density N_e inferred from a processing of the oscilloscope traces of the current in the discharge.

To obtain the CARS spectra the radiation of the second harmonic of a Nd:YAG laser at a frequency $\omega_1=18979$ cm^{-1} (with a power of up to 200 kW and a radiation linewidth of 0.05 cm^{-1}) together with the radiation of a narrow-band tunable dye laser (with a pulse power of up to 1 kW at a frequency $\omega_2=16475$ cm^{-1}) was focused to approximately the center of the discharge cell. The measurements of the gas temperature and the VDF were made in a collinear scheme of convergence of the laser beams, giving a spatial resolution of $50\times50\times1000$ μm. The useful CARS signal was discriminated from the background radiation of the pump lasers by means of a four-prism filter. The spectra were recorded in the integration mode of the photomultiplier, the signal from which was averaged by a boxcar integrator and sent to a chart recorder. The accuracy of determination of the delay time τ_d of the probe pulses relative to the current pulse was ±5 ns for $\tau_d<500$ ns and ±20 ns for $\tau_d>500$ ns.

We recorded experimentally the vibrational–rotational Raman spectrum of the Q bands of transitions 0–1, 1–2, 2–3, 3–4, 4–5, and 5–6 at delay times ranging from 30 ns to 20 μs from the start of the current pulse. The recorded CARS spectra were interpreted with the aid of numerical modeling by the technique of Ref. 8. Figure 3 shows the calculated and measured spectra of the Q branch for the transition 0–1 with a resolved rotational structure. From the experimental spectrum we recovered the distribution function of the nitrogen molecules over rotational levels in the ground state $X^1\Sigma_g^+$ (RDF), which is shown in Fig. 3a and, as we see, has a Boltzmann form. The RDFs obtained across the delay range all turned out to be Boltzmann distributions and, within the error limits of the measurements, were characterized by a temperature of 320 ± 10 K, which can be considered to be equal to the gas temperature in the investigated range of gas pressures.

COMPUTATIONAL MODEL; THE RATE CONSTANTS AND CROSS SECTIONS OF THE PROCESSES

For considering the plasma-chemical processes occurring in the burning and afterglow stages of a pulsed discharge in pure molecular nitrogen we integrated numerically a system of equations for the densities of the following species of particles: nitrogen molecules in the ground state $X^1\Sigma_g^+$ (47 vibrational levels, the vibrational level $v=46$ being taken as the level of dissociation of the nitrogen molecule via vibrational excitation) and in the electronically excited states $A^3\Sigma_u^+$, $B^3\Pi_g$, $C^3\Pi_u$, $B'^3\Sigma_u^-$, $a'^1\Sigma_u^-$, $W^3\Delta_u$, $a^2\Pi_g$, and $w^1\Delta_u$; nitrogen atoms in the ground state ^4S and excited states ^2P and ^2D; ions N^+, N_2^+, N_3^+, N_4^+; and electrons. Table I lists the reactions, involving neutral and charged particles, which were taken into account in the solution of the system of equations. We note that, depending on the numbers of the vibrational levels of the colliding molecules in Refs. 1, 7, and 9–12, different models were used in the calculations of the rate constants for VV processes (2).

To determine the rate constants for electron–molecule

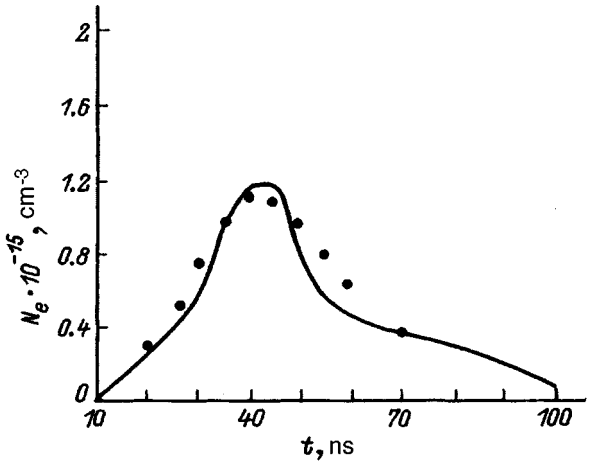

FIG. 2. Electron density versus time. ● — Experiment; solid curve — calculation.

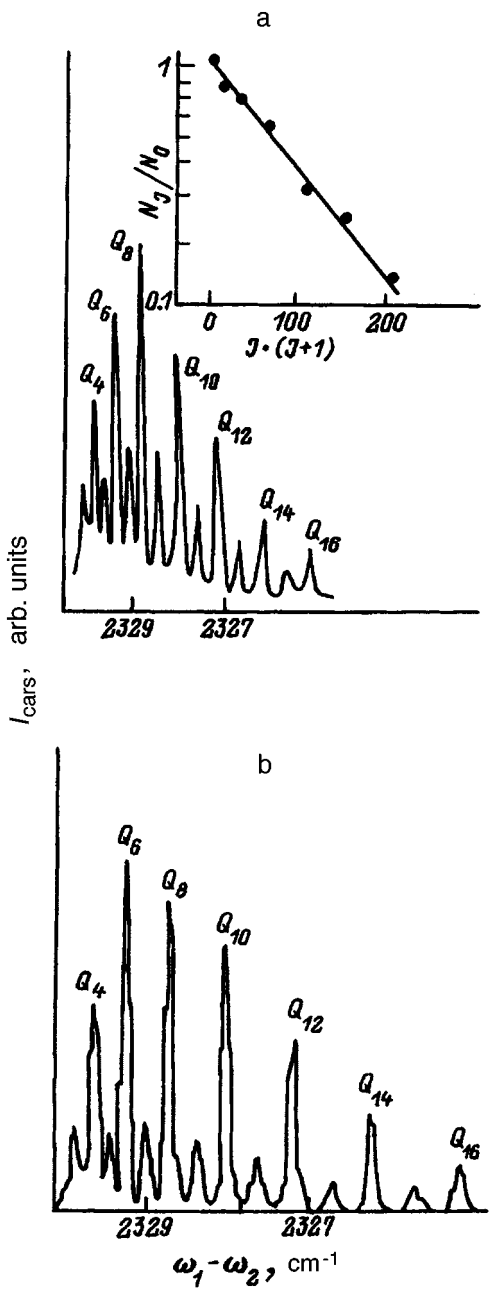

FIG. 3. Distribution of the radiation intensity in the CARS spectrum of the Q branch of the vibrational transition from $v=0$ to $v=1$ of the nitrogen molecule at a pressure of 115 torr and a current strength of 1.4 kA at $t=50$ ns. a: Experiment (N_J is the density of molecules on vibrational level J). b: Calculation for a temperature of 320 K.

collisions, we solved the Boltzmann transport equation numerically for the EEDF with allowance for elastic collisions of electrons with molecules and atoms, the excitation of rotational, vibrational, and electronic states of the molecules by electron impact, the dissociation and ionization of molecules from the ground state in collisions with electrons, and superelastic collisions of electrons with vibrationally excited molecules in the ground and electronically excited states.

The rate constants for excitation of electronic states of molecules (without resolution of the vibrational levels) from high vibrational levels of the ground state $X^1\Sigma_g^+$ were calculated on the basis of the relations from Refs. 42 and 43. And values of v_{dr}, T_e, and D/μ necessary for determining the electron density and the rate constants for dissociative electron–ion recombination were calculated according to the expressions from Ref. 44. Here $T_e=(2/3)\overline{\varepsilon}$, where $\overline{\varepsilon}$ is the average electron energy evaluated from the EEDF.

The total cross section σ_Σ for vibrational excitation was determined from a comparison of the calculated and experimental values of v_{dr}, D/μ, and the VDF under the assumption that in the burning stage of the discharge the change in populations at the lower vibrational levels $v=0-5$ of the molecules is due to direct electron impact. In the solution of the Boltzmann equation for the EEDF the values of the first eight level cross sections of vibrational excitation from Refs. 15 and 16 were suitably renormalized as the value of σ_Σ was varied. The resulting EEDF was used to calculate the rate constants for processes involving a discharge component and electrons. Then the system of equations for the densities were integrated numerically over the time interval $t=50$ ns. At the initial time the VDF corresponded to the Boltzmann distribution for $T=320$ K, and the densities of atoms, electronically excited molecules, and ions were assumed equal to zero. The procedure of varying the value of σ_Σ and performing the calculations was repeated until numerical agreement was reached between the measured and calculated values of the molecular populations on the vibrational levels $v=0-5$ and also for v_{dr} and D/μ. We might add that the calculations were done for both the measured and calculated time dependences of the electron density.

The value found for σ_Σ was used to determine the EEDF and the rate constants for electron–molecule collisions in a recombining atomic–molecular plasma. A numerical modeling of the decay of the plasma was carried out all the way to 20 μs. As the initial conditions in the equations for the densities we used the calculated composition for the time $t=50$ ns. The rate constants of electron–molecule collisions were recalculated in the course of solving the equations as the parameters of the recombining plasma were changed. In the calculations the rate constants of the little-studied associative ionization reaction (33) involving atoms in the metastable states N(^2P) and N(^2D) and also the rate constant K_{01}^{10} for VV exchange between the levels $v=1$ and $v=0$ were varied until the best agreement was achieved between the calculated and experimental values of the populations of the states with $v=0-5$.

DISCUSSION OF THE RESULTS

1) Burning regime. Figure 4 shows the results of calculations and measurements of the dependence of $\ln(N_v/N_0)$ on the vibrational level number v at a time $t=50$ ns in the burning stage of the pulsed discharge. A comparison of the experimental and calculated results for the VDF in the discharge shows that the best agreement between them is attained when the total cross section σ_Σ for vibrational excitation to the first eight vibrational levels is $(10.5\pm1.5)\times10^{-16}$ cm^2. This value of σ_Σ is in good agreement with the results of Refs. 45 and 46. As can be seen from Fig. 5, the results of the calculation for the drift velocity v_{dr} and the characteristic temperature D/μ of the EEDF,

TABLE I.

Reaction No.	Reaction and rate constants		Ref.
1	$N_2(X^1\Sigma_g^+, v+1) + N_2(X^1\Sigma_g^+) \longleftrightarrow N_2(X^1\Sigma_g^+, v) + N_2(X^1\Sigma_g^+)$	$k = K^m_{v+1,v}$	[1,7]
2	$N_2(X^1\Sigma_g^+, v+1) + N_2(X^1\Sigma_g^+, s-1) \longleftrightarrow N_2(X^1\Sigma_g^+, v) + N_2(X^1\Sigma_g^+, s)$	$k = K^{v,s}_{v+1,s-1}$	[1,7] [9–12]
3	$N_2(X^1\Sigma_g^+, v+1) + N(^4S) \longleftrightarrow N_2(X^1\Sigma_g^+, v) + N(^4S)$	$k = K^A_{v+1,v}$	[1,13]
4	$N_2(X^1\Sigma_g^+, v=46) + N_2(X^1\Sigma_g^+, v) \rightarrow N_2(X^1\Sigma_g^+, v-1) + N(^4S) + N(^4S)$	$k = K_d^{VV}$	[14]
5	$N_2(X^1\Sigma_g^+, v=46) + N_2(X^1\Sigma_g^+) \rightarrow N_2(X^1\Sigma_g^+,) + N(^4S) + N(^4S)$	$k = K_d^{VTm}$	[14]
6	$N_2(X^1\Sigma_g^+, v=46) + N(^4S) \rightarrow N(^4S) + N(^4S) + N(^4S)$	$k = K_d^{VTa}$	[14]
7	$e + N_2(X^1\Sigma_g^+, v) \longleftrightarrow e + N_2(X^1\Sigma_g^+, s)$; for k calculated from the cross section σ_v and the EEDF		[15–17]
8	$e + N_2(X^1\Sigma_g^+, v) \longleftrightarrow e + N_2(Y)$, $Y = A^3\Sigma_u^+, B^3\Pi_g, C^3\Pi_u, B'^3\Sigma_u^-, a'^1\Sigma_u^-, W^3\Delta_u, a''^1\Sigma_g^+, a^1\Pi_g, w^1\Delta_u$; for k calculated from the cross section σ_v and the EEDF		[18]
9	$N_2(A^3\Sigma_u^+) + N_2(A^3\Sigma_u^+) \rightarrow N_2(C^3\Pi_u) + N_2(X^1\Sigma_g^+)$	$k = 1.3 \cdot 10^{-10}$	[19]
10	$N_2(A^3\Sigma_u^+) + N_2(A^3\Sigma_u^+) \rightarrow N_2(B^3\Pi_g) + N_2(X^1\Sigma_g^+)$	$k = 7.7 \cdot 10^{-11}$	[20]
11	$N_2(A^3\Sigma_u^+) + N_2(X^1\Sigma_g^+, v>3) \rightarrow N_2(B^3\Pi_g) + N_2(X^1\Sigma_g^+)$	$k = 3.0 \cdot 10^{-11}$	[20]
12	$N_2(B^3\Pi_g) + N_2(X^1\Sigma_g^+) \rightarrow N_2(A^3\Sigma_u^+) + N_2(X^1\Sigma_g^+)$	$k \approx 2.0 \cdot 10^{-11}$	[21,22]
13	$N_2(B^3\Pi_g) \rightarrow N_2(A^3\Sigma_u^+) + h\omega$	$A = 1.2 \cdot 10^5$	[21]
14	$N_2(C^3\Pi_u) \rightarrow N_2(B^3\Pi_g) + h\omega$	$A = 2.7 \cdot 10^7$	[21]
15	$N_2(C^3\Pi_u) + N_2(X^1\Sigma_g^+) \rightarrow N_2(B^3\Pi_g) + N_2(X^1\Sigma_g^+)$	$k = 1.0 \cdot 10^{-11}$	[21]
16	$N_2(C^3\Pi_u) + N_2(X^1\Sigma_g^+) \rightarrow N_2(A^3\Sigma_u^+) + N_2(X^1\Sigma_g^+)$	$k = 1.0 \cdot 10^{-11}$	[21]
17	$N_2(a'^1\Sigma_u^-) + N_2(X^1\Sigma_g^+) \rightarrow N_2(B^3\Pi_g) + N_2(X^1\Sigma_g^+)$	$k = 2.0 \cdot 10^{-13}$	[22]
18	$N_2(C^3\Pi_u) + N_2(X^1\Sigma_g^+) \rightarrow N_2(a'^1\Sigma_u^-) + N_2(X^1\Sigma_g^+)$	$k = 1.0 \cdot 10^{-11}$	[22]
19	$N_2(X^1\Sigma_g^+) + e \rightarrow N(^4S) + N(^4S) + e$; for k calculated from the cross section σ_{diss} and the EEDF		[23,24]
20	$N(^4S) + e \rightarrow N(Y) + e$ $Y = {}^2P, {}^2D$; for k calculated from the cross section σ_Y and the EEDF		[25,26]
21	$N(^2D) + e \rightarrow N(^2P) + e$; for k calculated from the cross section σ_Y and the EEDF		[27]
22	$N(^4S) + N(^4S) + N_2(X^1\Sigma_g^+) \rightarrow N_2(Y) + N_2(X^1\Sigma_g^+)$		
	$Y = X^1\Sigma_g^+, A^3\Sigma_u^+,$	$k_r = 8.27 \cdot 10^{-34} \exp(-500/T)$	[28]
	$Y = C^3\Pi_u,$	$k_r = 1.0 \cdot 10^{-34}$	[29]
23	$N(^4S) + N(^2D) + N_2(X^1\Sigma_g^+) \rightarrow N_2(B^3\Pi_g) + N_2(X^1\Sigma_g^+)$	$k_r = 1.0 \cdot 10^{-34}$	[29]
24	$N(^4S) + N(^2P) \rightarrow N(^4S) + N(Y)$ $Y = {}^4S,$	$k = 1.8 \cdot 10^{-12}$	[29]
	$Y = {}^2D,$	$k = 6.0 \cdot 10^{-13}$	[22]
25	$N(^2P) + N_2(X^1\Sigma_g^+, 8 \le v \le 46) \rightarrow N(^4S) + N_2(A^3\Sigma_u^+)$	$k = 1.0 \cdot 10^{-10} \exp(-1300/T)$	[21]
26	$N(Y) + N_2(X^1\Sigma_g^+) \rightarrow N(^4S) + (X^1\Sigma_g^+)$ $Y = {}^2P,$	$k = 6.0 \cdot 10^{-14}$	[22]
	$Y = {}^2D,$	$k = 2.3 \cdot 10^{-14}$	
27	$N_2(A^3\Sigma_u^+) + N(^4S) \rightarrow N_2(X^1\Sigma_g^+) + N(Y)$ $Y = {}^2P,$	$k = 1.9 \cdot 10^{-10}$	[30]
	$Y = {}^2D,$	$k = 5.0 \cdot 10^{-11}$	[31]
28	$N_2(Y) + N(^4S) \rightarrow N_2(B^3\Pi_g) + N(^4S)$ $Y = B'^3\Sigma_u^-, a'^1\Sigma_u^-,$	$k = 10^{-10}$	[21]
	$Y = C^3\Pi_u, a^1\Pi_g, w^1\Delta_u,$	$k = 3 \cdot 10^{-11}$	
29	$N_2(Y) + N(^4S) \rightarrow N_2(X^1\Sigma_g^+) + N(^2P)$ $Y = a'^1\Sigma_u^-, B^3\Pi_g,$	$k = 10^{-10}$	[29]
	$Y = C^3\Pi_u,$	$k = 3 \cdot 10^{-10}$	
30	$N_2(Y) + N_2(X^1\Sigma_g^+) \rightarrow N_2(X^1\Sigma_g^+) + N_2(X^1\Sigma_g^+)$		
	$Y = A^3\Sigma_u^+,$	$k = 3.7 \cdot 10^{-16}$	[32]
	$Y = w^1\Delta_u, W^3\Delta_u, B^3\Pi_g,$	$k = 1 \cdot 10^{-11}$	[33,20]
	$Y = C^3\Pi_u,$	$k = 1 \cdot 10^{-11}$	[21]
	$Y = B'^3\Sigma_u^-, a'^1\Sigma_u^-,$	$k = 1.9 \cdot 10^{-10}$	[33,34]
	$Y = a^1\Pi_g,$	$k = 9.1 \cdot 10^{-12}$	[35]
31	$N_2(X^1\Sigma_g^+) + e \rightarrow N_2^+ + e + e$; for k calculated from the cross section σ_{ion} and the EEDF		[36]
32	$N_2(a'^1\Sigma_u^-) + N_2(a'^1\Sigma_u^-) \rightarrow e + Y$		[37]
	$N_2(a'^1\Sigma_u^-) + N_2(A^3\Sigma_u^+) \rightarrow e + Y$ $Y = N_4^+,$	$k = 5.0 \cdot 10^{-11}$	
33	$N(Y) + N(^2P) \rightarrow N_2^+ + e$ $Y = {}^2P, {}^2D,$	$k < 10^{-12}$	[21]
34	$N_2(A^3\Sigma_u^+) + e \rightarrow N_2^+ + e + e$; for k calculated from the cross section σ_{ion} and the EEDF		[38]
35	$N(^4S) + e \rightarrow N^+ + e + e$; for k calculated from the cross section σ_{ion} and the EEDF		[39]
36	$N_2^+ + e \rightarrow N(Y) + N(^4S)$ $Y = {}^4S,$	$k = 3.5 \cdot 10^{-7} \cdot (300/T_e)^{0.5}$	[29,40]
	$Y = {}^2D,$	$k = 2.0 \cdot 10^{-7} \cdot (300/T_e)^{0.5}$	
37	$N_3^+ + e \rightarrow N_2(Y) + N(^4S)$ $Y = X^1\Sigma_g^+,$	$k = 2 \cdot 10^{-7} \cdot (300/T_e)^{0.5}$	[29]
	$Y = A^3\Sigma_u^+, B^3\Pi_g,$	$k = 4.3 \cdot 10^{-7} (300/T_e)^{0.5}$	[41]
38	$N_4^+ + e \rightarrow N_2(X^1\Sigma_g^+) + N_2(X^1\Sigma_g^+)$	$k = 2.0 \cdot 10^{-6} \cdot (300/T_e)^{0.5}$	[29]

TABLE I. (*Continued.*)

Reaction No.	Reaction and rate constants		Ref.
39	$N_2^+ + e + e \rightarrow N_2(X^1\Sigma_g^+) + e$	$k_r = 1.0 \cdot 10^{-19} \cdot (300/T_e)^{4.5}$	[29]
40	$N^+ + e + e \rightarrow N(^4S) + e$	$k_r = 1.0 \cdot 10^{-19} \cdot (300/T_e)^{4.5}$	[29]
41	$N_2^+ + e + N_2(X^1\Sigma_g^+) \rightarrow N_2(X^1\Sigma_g^+) + N_2(X^1\Sigma_g^+)$	$k_r = 6 \cdot 10^{-27} \cdot (300/T_e)^{2.5}$	[29]
42	$N^+ + e + N_2(X^1\Sigma_g^+) \rightarrow N_2(X^1\Sigma_g^+) + N(^4S)$	$k_r = 6 \cdot 10^{-27} \cdot (300/T_e)^{2.5}$	[29]
43	$N_2^+ + N_2(X^1\Sigma_g^+) + N_2(X^1\Sigma_d^+) \rightarrow N_2(X^1\Sigma_g^+) + N_4^+$	$k_r = 10^{-28}$	[21]
44	$N(^4S) + N_2^+ \rightarrow N_2(X^1\Sigma_g^+) + N^+$	$k = 2.4 \cdot 10^{-15} \cdot T$	[21]
45	$N_2(X^1\Sigma_g^+) + N_4^+ \rightarrow N_2^+ + N_2(X^1\Sigma_g^+) + N_2(X^1\Sigma_g^+)$	$\log(k) = -14.6 + 0.0036 \cdot (T - 300),\ 300 < T < 900$	[28]
46	$N^+ + N_2(X^1\Sigma_g^+) + N_2(X^1\Sigma_g^+) \rightarrow N_2(X^1\Sigma_g^+) + N_3^+$	$k_r = 9 \cdot 10^{-30} \exp(-400/T)$	[21]
47	$N(^4S) + N_2^+ + N_2(X^1\Sigma_g^+) \rightarrow N_2(X^1\Sigma_g^+) + N_3^+$	$k_r = 9 \cdot 10^{-30} \exp(-400/T)$	[21]
48	$N_2^+ + N_2(A^3\Sigma_u^+) \rightarrow N(^4S) + N_3^+$	$k = 3.0 \cdot 10^{-10}$	[21]
49	$N_4^+ + N(^4S) \rightarrow N^+ + N_2(X^1\Sigma_g^+) + N_2(X^1\Sigma_g^+)$	$k = 10^{-11}$	[21]
50	$N_2^+ + N_2(X^1 + \Sigma_g^+, v > 40) \rightarrow N_2(X^1\Sigma_g^+) + N(^4S) + N^+$	$k = 1.2 \cdot 10^{-11}$	[21]
51	$N(^4S) + N^+ + N_2(X^1\Sigma_g^+) \rightarrow N_2^+ + N_2(X^1\Sigma_g^+)$	$k_r = 0.9 \cdot 10^{-29} \exp(-400/T)$	[28]
52	$N^+ + N_2(X^1\Sigma_g^+, 3 \le v \le 46) \rightarrow N(^4S) + N_2^+$	$k = 1.2 \cdot 10^{-11}$	[21]
53	$N_2^+ + N_2(X^1\Sigma_g^+, 21 \le v \le 46) \rightarrow N(^4S) + N_3^+$	$k = 5.5 \cdot 10^{-11}$	[21]
54	$N(^4S) + N_3^+ \rightarrow N_2^+ + N_2(X^1\Sigma_g^+)$	$k = 6.6 \cdot 10^{-11}$	[21]

Note. The rate constants k are given in units of cm^3/s; the radiative decay probabilities A are in units of s^{-1}; the rate constants k_r for processes involving three particles are in units of cm^6/s; K_d^i is the rate of dissociation of molecules by a purely vibrational mechanism in collisions with molecules, via *VV*-exchange processes (if $i=VV$) and *VT* relaxation (if $i=VTm$), and in collisions with atoms, via *VT* relaxation (if $i=VTa$); $K_{v+1,v}^m$, $K_{v+1,s-1}^{v,s}$, and $K_{v+1,v}^a$ denote the rate constants for *VV* exchange and *VT* relaxation of molecules on molecules and atoms, calculated using the model dependence on v and s.

recovered from the solution of the Boltzmann equation at a value of the total cross section $\sigma_\Sigma = 10.5 \times 10^{-16}$ cm^2, is also in good agreement with the measured values from Ref. 47 over a wide range of variation of the reduced electric field $E/N = 20$–80 Td. The discrepancy between the calculated and measured values of v_{dr} and D/μ is not over 5%.

Numerical modeling of the kinetics of vibrational excitation of the molecules in the discharge, both with the measured and with the calculated time dependence of the electron density (Fig. 2), confirms the presence of a characteristic feature in the shape of the VDF observed in experiment. By 50 ns after the start of the discharge current pulse the populations of the first five vibrational levels $v = 1$–5 have a

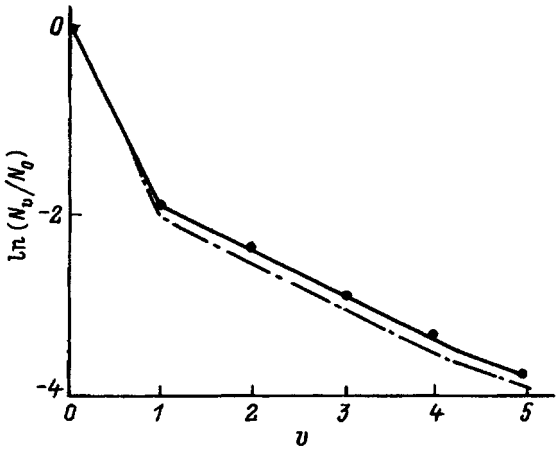

FIG. 4. Distribution $\ln(N_v/N_0)$ of the nitrogen molecules in the electronic ground state $X^1\Sigma_g^+$ over vibrational levels $v = 0$–5. ● — Experiment; solid curve — calculation with the experimental time dependence of N_e; dot-and-dash curve — calculation with the calculated time dependence of N_e.

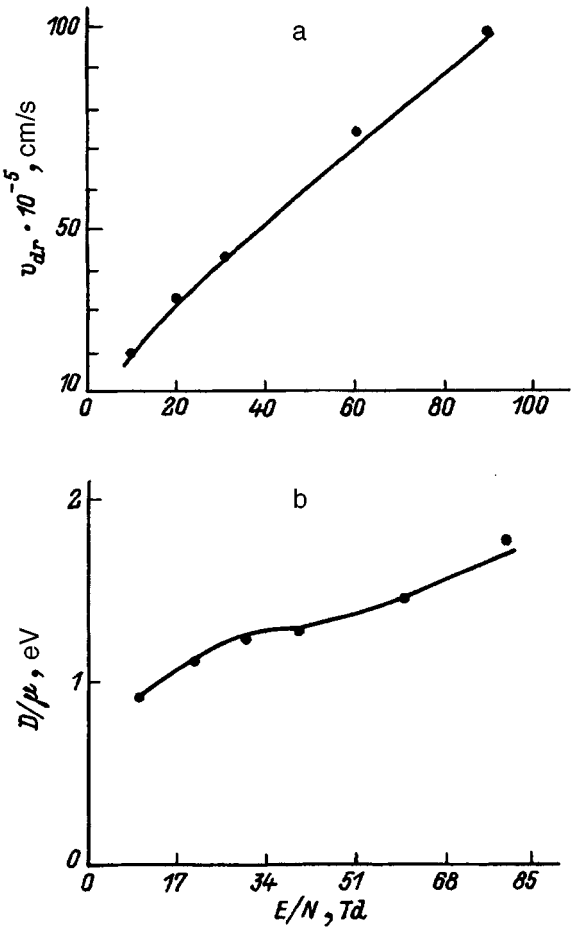

FIG. 5. Plots of electron parameters v_{dr} (a) and D/μ (b) versus E/N. ● — Experiment of Ref. 47; solid curve — calculation.

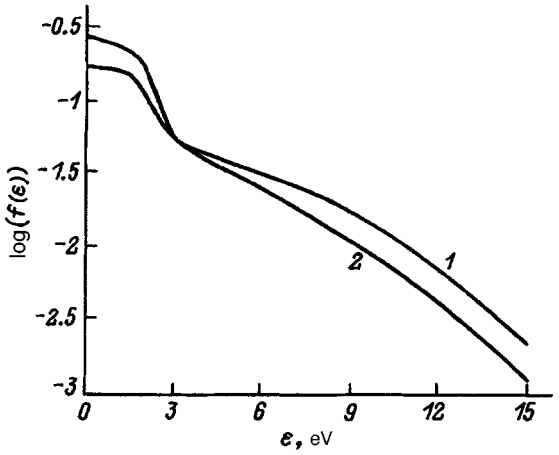

FIG. 6. Electron energy distribution function $f(\varepsilon)$ in the burning stage of the discharge for the VDF at time $t=50$ ns. $E/N=20$ Td, $T=320$ K, and $\sigma_\Sigma=10.5\times10^{-16}$ cm^2; *1* — $\alpha=0$, *2* — $\alpha=10^{-2}$.

Boltzmann distribution with a vibrational temperature $T_{v15}=7315$ K, which is in good agreement with its experimental value $T_{v15}=7290\pm350$ K. It is markedly different from the vibrational temperature of the first level $T_{v01}=1640$ K, as is also found in experiment: $T_{v01}=1740\pm200$ K. The presence of a ''bend'' in the VDF indicates that the initial stage of its evolution is governed mainly by the excitation of vibrational states of the molecules by direct electron impact. Indeed, calculations confirm that in calculating the populations of states with $v=1-5$ in the model under discussion one can neglect the influence of processes of excitation of electronic states of the molecules and the dissociation and ionization of molecules from the ground state as a result of collisions with electrons. Taking these processes into account in reactions describing the population dynamics of the lower levels leads to changes of not more than 10–20% in the absolute values, and that is within the experimental error.

The use of different recursion relations for the rate constants describing inelastic and superelastic collisions of electrons with vibrationally excited molecules will likewise not lead to substantial changes in the calculated populations of the lower levels. For example, the solution of the balance equations for the populations of the lower levels calculated using the relations from Ref. 48 differs by 10% from that calculated using the simplified relations from Refs. 37 and 49.

An analysis of the solutions of the Boltzmann transport equation for the EEDF implies that a key question from the standpoint of finding the rate constants for electron–molecule collisions is whether nitrogen atoms are formed in the ground state ^4S. According to calculations, under the actual experimental conditions the degree of dissociation $\alpha=N_a/N$ of the molecules in the discharge reaches values $\alpha\approx10^{-2}$. Figure 6 illustrates the fact that the EEDF differs strongly, over a wide range of energies, from the function calculated without taking into account the formation of atoms in the discharge. The change in the composition of the discharge is accordingly reflected in the rate constants for electron–molecule collisions, and, as a result, the populations of the lower levels are perturbed by 30–40%. We note that superelastic collisions of electrons with electronically excited molecules have a negligible influence on the EEDF.

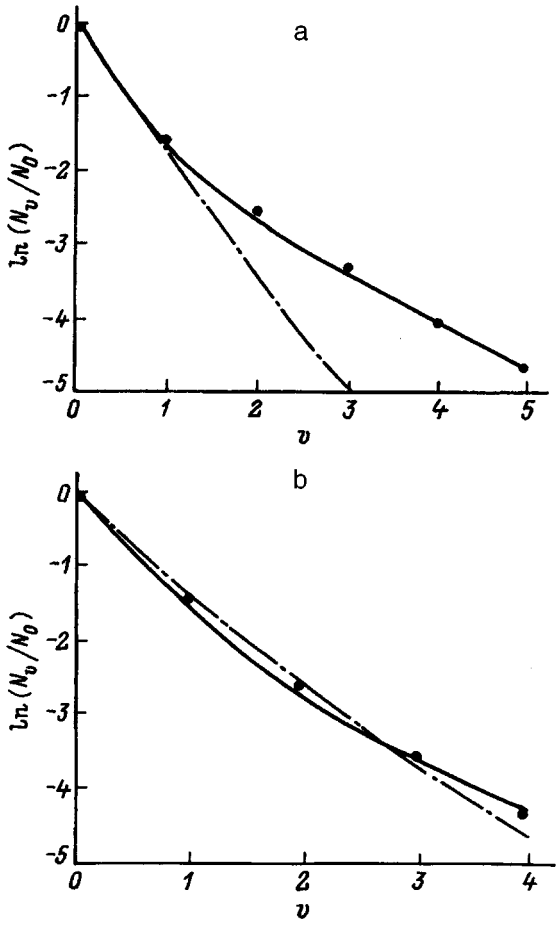

FIG. 7. Distribution of $\ln(N_v/N_0)$ of nitrogen molecules in the electronic ground state $X^1\Sigma_g^+$ over vibrational levels $v=0-5$ in the afterglow stage of the discharge. ● — Experiment; solid curve — calculation; dot-and-dash curve — Treanor distribution. $T=6$ μs (a), 20 μs (b); $T_v=1900$ K (a), 2500 K (b); $T=320$ K.

Thus the formation of atoms in the discharge indirectly affects the population kinetics of the lower vibrational levels. This makes it necessary to solve jointly the balance equations for the densities of molecules, atoms, and charged particles with the Boltzmann transport equation for the EEDF in order to evaluate σ_Σ.

2) Afterglow. According to the data obtained in the calculations and in experiment, the formation of the VDF in the afterglow can be divided temporally into two stages: an early stage, corresponding to the time interval from 50 ns to 15 μs, and a late stage, the period from 15 to 20 μs. In the early stage the populations in states with $v=0-5$ change mainly as a result of inelastic and superelastic collisions of electrons with vibrationally excited molecules. In the late stage the redistribution of the molecules over vibrational levels is dominated by *VV* exchange processes.

Figure 7a shows the calculated and experimental values of the populations for the lower vibrational levels from $v=0$ to $v=5$ at the time $t=6$ μs. The distribution of the molecules over levels differs from the Treanor distribution.

The vibrational temperatures of the first level, obtained both experimentally ($T_{\nu 01}=2130\pm 200$ K) and theoretically ($T_{\nu 01}=2000$ K), are considerably detached from the vibrational temperatures corresponding to the first five vibrational levels, $T_{\nu 15}=4420\pm 350$ and 4100 K, respectively. The quantitative agreement with experiment argues in favor of the interpretation that the change in the number of molecules on the lower levels in the early stage of the decay of the plasma occurs more efficiently as a result of inelastic and superelastic collisions of electrons with vibrationally excited molecules than as a result of their redistribution owing to *VV* exchange processes. Here processes of *VT* relaxation of the molecules on molecules and atoms at the measured gas temperature $T=320$ K do not affect the results of the calculations.

The rate of change of the populations depends on the densities of electrons and ions. Therefore, it is interrelated with the processes that govern the kinetics of the electron component and the ions in a recombining atomic–molecular plasma.

The main channels for neutralization of the charge are dissociative electron–ion recombination (36–38). An important role in the kinetics of the electron component is played by the conversion reactions (43 and 46) and the destruction of complex ions; these regulate the densities of N_3^+, N_4^+, and N_2^+ ions and thus raise or lower the significance of the dissociative recombination channels in which these ions participate. It must be noted that in view of the insignificant population of vibrational levels with $v>21$ the change in the species of nitrogen ion in reactions (50 and 53) is small and has only a weak effect on the value of the electron density.

Associative ionization processes (32 and 33) involving molecules and atoms in the metastable states $A^3\Sigma_u^+$, $a'^1\Sigma_u^-$, and ^2P, ^2D have an appreciable influence on the characteristic decay times of the plasma that is created. A decrease in the rate of decay on account of reactions (32) is manifested at times $t\approx 60$–800 ns. As the time increases to $t>1$ μs there occurs a depletion of the populations of electronically excited states mainly in collisions with molecules $N_2(X^1\Sigma_g^+)$ and atoms N(^{4}S). By the time $t=6$ μs the quenching of the molecules on atoms has significantly increased the density of atoms in the states ^2P and ^2D, and therefore the associative ionization channel (33) becomes dominant. On this basis we estimated the rate constant for reaction (33). Good quantitative agreement of the calculated and experimental results for the EEDF at the time $t=6$ μs is obtained with a value $k=10^{-13}$ cm^3/s.

The subsequent transformation of the VDF, corresponding to the late stage of the afterglow, is shown in Fig. 7b. The Treanor form of the distribution at the time $t=20$ μs and the results of the calculations confirm the hypothesis that *VV* exchange processes play the dominant role in the redistribution of molecules over the lower levels. This makes it possible to determine the rate constant K_{01}^{10} for *VV* exchange: the best agreement between the calculation and experiment is obtained at a value $K_{01}^{10}=(9\pm 1)\times 10^{-15}$ cm^3/s, which agrees with the value predicted in the quantum–classical Billing–Fisher model.[7] It should be added that the results of the calculations differed only slightly when different models[7,9–12] were used for the dependence of the rate constants for *VV* exchange, $K_{m+1,m}^{n,n+1}$, on the vibrational numbers *m* and *n* of the colliding molecules.

CONCLUSION

In summary, an experimental study and a numerical modeling of the temporal evolution of the populations of the lower vibrational levels $v=0$–5 of the ground state $X^1\Sigma_g^+$ in a high-current pulsed discharge (at $E/N=200$ Td, current strength 1.4 kA, gas temperature $T=320$ K) and in its afterglow at a pressure of 115 torr in pure molecular nitrogen have shown that in the burning stage of the discharge the change in populations of the vibrational levels of the molecules is governed predominantly by direct electron impact, that the formation of nitrogen atoms in the ground state ^4S has an appreciable influence on the EEDF and, accordingly, indirectly perturbs the populations of the lower vibrational levels, and that in the early stage of the afterglow of the discharge the change in populations occurs mainly as a result of inelastic and superelastic collisions of electrons with vibrationally excited molecules, while in the later stage of the afterglow the dominant process in the redistribution of the molecules over the lower levels is *VV* exchange.

The numerical model proposed for calculating the densities of the different components (molecules, atoms, and charged particles) and the EEDF gives good quantitative agreement with the experimental time dependence of the densities of electrons and vibrationally excited molecules and confirms the CARS measurements.

A comparison of the experimental and calculated results reveals that the total cross section σ_Σ for vibrational excitation to the first eight vibrational levels reaches 10.5×10^{-16} cm^2; a proposed refinement of the rate constant $k=10^{-13}$ cm^3/s for the associative ionization reaction involving the participation of atoms in the metastable states N(^2P) and N(^2D) improves the agreement of the calculated VDF with the measured functions at the time $t=6$ μs; the best agreement between the calculated and measured populations of the molecules at the time $t=20$ μs is found at a rate constant for *VV* exchange equal to $K_{01}^{10}=9\times 10^{-15}$ cm^3/s.

The information we have obtained about the values of σ_Σ, k, and K_{01}^{10}, we believe, can be recommended for numerical modeling of the vibrational and charge kinetics of molecular nitrogen in a pulsed discharge and its afterglow.

The authors are grateful to O. A. Gordeev for substantial assistance in providing data on the cross sections for electron-impact excitation of nitrogen molecules, for showing interest in this study, and for taking part in a discussion of it.

[1] M. Capitelli (ed.), *Nonequilibrium Vibrational Kinetics* (Vol. 39 of Topics in Current Physics series), Springer-Verlag, Berlin (1986) [Russ. transl., Mir, Moscow (1989)].
[2] O. A. Gordeev and D. V. Khmara, Teplofiz. Vys. Temp. **32**(1), 133 (1994).
[3] Yu. S. Akishev, A. V. Dem'yanov, I. V. Kochetov *et al.*, Teplofiz. Vys. Temp. **20**, 818 (1982).
[4] A. F. Suchkov and Yu. N. Shebeko, Khim. Vys. Energ. **15**, 279 (1981).
[5] A. A. Devyatov, S. A. Dolenko, A. T. Rakhimov *et al.*, Zh. Éksp. Teor. Fiz. **90**, 429 (1986) [Sov. Phys. JETP **63**, 246 (1986)].
[6] O. A. Gordeev and V. A. Shakhatov, *Materials of the Conference on*

Plasma Physics and Techniques [in Russian], Minsk, BGU (1994), p. 87.
[7] G. D. Billing and E. R. Fisher, Chem. Phys. **43**, 395 (1979).
[8] J. C. Luthe, E. J. Beiting, and F. Y. Yueh, Computer Phys. Commun. **42**, 73 (1986).
[9] B. F. Gordiets, A. I. Osipov, E. V. Stupochenko, and L. A. Shelepin, Usp. Fiz. Nauk **108**, 655 (1972) [Sov. Phys. Usp. **15**, 759 (1973)].
[10] S. D. Rockwood, J. E. Brau, W. A. Proctor, and G. H. Canavan, IEEE J. **QE-9**, 120 (1973).
[11] Yu. B. Konev, I. V. Kochetov, V. N. Pevgov, and V. F. Sharkov, Preprint No. 2821 [in Russian], I. V. Kurchatov Institute of Atomic Energy, Academy of Sciences of the USSR, Moscow (1977).
[12] A. V. Bogdanov, Yu. N. Gorbachev, and V. A. Pavlov, Preprint No. 833 [in Russian], A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, Leningrad (1983).
[13] P. V. Kozlov, S. A. Losev, V. N. Makarov *et al.*, Khim. Fiz. **14**(4), 57 (1995).
[14] J. Loureiro, C. M. Ferreira, and M. Capitelli, J. Phys. D **23**, 1371 (1990).
[15] G. J. Schulz, Phys. Rev. A **135**, 938 (1964).
[16] D. Spence, J. L. Mauer, and G. J. Schulz, J. Chem. Phys. **57**, 5516 (1972).
[17] M. J. W. Boness and G. J. Schulz, J. Chem. Phys. **57**, 2883 (1972).
[18] D. C. Cartwright, S. Trajmar, A. Chutjan, and W. Williams, Phys. Rev. A **16**, 1041 (1977).
[19] L. G. Piper, J. Chem. Phys. **88**, 231 (1988).
[20] L. G. Piper, J. Chem. Phys. **88**, 6911 (1988).
[21] D. I. Slovetskiĭ, *Mechanisms of Chemical Reactions in Nonequilibrium Plasma* [in Russian], Nauka, Moscow (1980).
[22] B. M. Smirnov, *Excited Atoms* [in Russian], Énergoatomizdat, Moscow (1982).
[23] H. F. Winters, J. Chem. Phys. **44**, 1472 (1966).
[24] E. C. Zipf and R. W. McLauglin, J. Planet Space Sci. **26**, 449 (1978).
[25] K. A. Berrington, P. G. Burke, and W. D. Roob, J. Phys. B **8**, 2500 (1975).
[26] S. Ormonde, K. Smith, B. W. Torres, and A. R. Davies, Phys. Rev. A **8**, 262 (1973).
[27] R. J. W. Henry, Phys. Rev. A **178**, 218 (1969).
[28] A. Y. Kostinsky *et al.*, Preprint No. 7, Institute of General Physics, Academy of Sciences of the USSR, Moscow (1990), 7 pp.
[29] L. S. Polak *et al.*, in *Plasma Chemistry* [in Russian], ed. by B. M. Smirnov, No. 5, Atomizdat, Moscow (1978), p. 328.
[30] L. G. Piper, J. Chem. Phys. **90**, 7087 (1989).
[31] J. Loureiro, Chem. Phys. **157**, 157 (1991).
[32] J. W. Dreyer and D. Perner, J. Chem. Phys. **58**, 1195 (1973).
[33] Rajesh Nagpal and P. K. Ghosh, J. Phys. D **23**, 1663 (1990).
[34] L. G. Piper, J. Chem. Phys. **87**, 1625 (1987).
[35] L. Magne, G. Cernogora, and P. Veis, J. Phys. D **25**, 472 (1992).
[36] D. Rapp and P. Englander-Golden, J. Chem. Phys. **43**, 1464 (1965).
[37] H. Brunet and J. Rocca-Serra, J. Phys. D **54**, 4957 (1983).
[38] V. L. Bychkov and O. A. Gordeev, Khim. Fiz. **22**, 1064 (1992).
[39] I. I. Sobel'man, A. A. Vaĭnshteĭn, and E. A. Yukov, *Cross Sections for Excitation of Atoms and Ions by Electrons* [in Russian], Nauka, Moscow (1973), p. 143.
[40] F. Brouillard and D. W. McGowan (eds.), *Physics of Ion–Ion and Electron–Ion Collisions* (Vol. 83 of NATO ASI Series B, Physics), Plenum Press, New York (1983) [Russ. transl., Mir, Moscow (1986)].
[41] L. S. Polak, D. I. Slovetskiĭ, and R. D. Todesaĭte, Khim. Vys. Energ. **10**, 64 (1976).
[42] A. V. Berdyshev, I. V. Kochetov, and A. P. Napartovich, Fiz. Plazmy **14**, 741 (1988) [Sov. J. Plasma Phys. **14**, 438 (1988)].
[43] M. Cacciotore, M. Capitelli, and G. Gorse, Chem. Phys. **66**, 141 (1982).
[44] L. G. H. Huxley and R. W. Crompton, *The Diffusion and Drift of Electrons in Gases*, Wiley, New York (1974) [Russ. transl., Mir, Moscow (1977), p. 191].
[45] K. Tachibana and A. V. Phelps, J. Chem. Phys. **71**, 3544 (1979).
[46] D. Levron and A. V. Phelps, Bull. Am. Phys. Soc. **24**, 129 (1979).
[47] J. Dutton, J. Phys. Chem. Ref. Data **4**, 577 (1975).
[48] M. Capitelli and M. Dilonardo, Chem. Phys. **20**, 417 (1977).
[49] Yu. S. Akishev, K. V. Baiadze, V. M. Vetsko *et al.*, Fiz. Plazmy **11**, 999 (1985) [Sov. J. Plasma Phys. **11**, 582 (1985)].

Translated by Steve Torstveit

The effect of the anode dimensions on the characteristics of a hollow-cathode glow discharge

S. P. Nikulin

Institute of Electrophysics, Ural Division of the Russian Academy of Sciences, 620219 Ekaterinburg, Russia
(Submitted December 25, 1995)
Zh. Tekh. Fiz. **67**, 43–47 (May 1997)

An analysis is made of the effect of the anode dimensions and the gas pressure on the possibility of achieving various conditions of burning of a hollow-cathode glow discharge (with negative or positive anode fall), and also with an anode plasma and an electrostatic double layer. Analytical relations that agree with experimental data are found for the voltage across the cathode sheath and the double layer as a function of the anode area and the pressure. Simple expressions are derived for the critical pressure below which no discharge of a particular kind is possible. © *1997 American Institute of Physics.* [S1063-7842(97)00805-2]

INTRODUCTION

It has been shown[1,2] that by taking into account ionization processes in the cathode sheath of a hollow-cathode glow discharge one can calculate the current-voltage characteristics of the discharge in satisfactory agreement with experiment. As the discharge current I_d increases and accordingly the thickness d of the cathode sheath decreases, the cathode fall U_c, which in this situation is practically the same as the discharge voltage U_d, attains a constant value

$$U_0=\frac{W}{e\gamma}+\frac{E_i}{e}\approx\frac{W}{e\gamma}, \tag{1}$$

where W is the average energy used in a single ionization event, γ is the ion-electron emission coefficient, e is the electron charge, and E_i is the threshold ionization energy.

The approximate equality in (1) can be taken over with little loss of accuracy, since $\gamma\ll 1$, and thus the discharge voltage is 1–2 orders of magnitude higher than the ionization potential. The reason U_c attains a constant value is that for small d the ionization is caused almost entirely by γ-electrons, which take on an energy eU_c in the cathode fall, while the contribution to the ionization from slow secondary electrons is insignificant. The average number of ions n generated in the cavity by a single electron is given by the relation

$$n=\frac{(eU_c-E_i)}{W}\approx\frac{eU_c}{W}, \tag{2}$$

and it is easy to see that the condition for a self-consistent discharge is

$$\gamma n=1 \tag{3}$$

which is satisfied for U_c given by Eq. (1)

The expression for n is valid only when the electron emerges from the gas discharge gap having lost all of its energy. The results of Refs. 3 and 5 show that as the pressure is reduced, the voltage U_d increases because of the increased loss of fast electrons through the aperture of the cavity[3] or at the anode.[4,5] A decrease of the aperture area or the anode area S_a facilitated burning of the discharge; i. e., it resulted in a lower U_d and critical pressure p_{cr} at which the discharge cannot burn. However, this occurs only until S_a reaches the value $\sim S_c\sqrt{m/M}$, where S_c is the cathode area and m and M are the electron and ion mass, respectively. A further decrease in S_a caused an increase in U_d. Moreover, in Refs. 4 and 5 it was observed that if an anode was used with an area several times smaller than $S_c\sqrt{m/M}$, the discharge voltage again began to decrease. The purpose of the work reported in the present article is to analyze the reasons for the ambiguous effect of the anode size on the characteristics of a hollow-cathode glow discharge.

DISCHARGE BURNING WITH A NEGATIVE ANODE FALL

In Ref. 3 it was shown that for values of S_a higher than $\sim S_c\sqrt{m/M}$ a discharge is obtained with a negative anode fall U_a. The negative fall U_a is small as a rule, $\sim kT_e/e$, where k is Boltzmann's constant and T_e is the temperature of the plasma electrons. About the same small potential fall occurs in the plasma, and on the whole it can be said that U_d is essentially equal to U_c. Let us consider how U_c, and accordingly U_d would change if the pressure were reduced. We shall neglect the ionization processes in the cathode sheath and assume that the ionization in the cavity occurs only by γ electrons generated after acceleration in the cathode sheath to an energy eU_c.

As the electron energy decreases from $E+dE$ to E, the electron will have formed a number $n=dE/W$ ions. However, these ions are formed only if the electron with an energy E remains in the gas discharge gap and does not escape to the anode. Therefore if we allow for the possibility that the electron can escape, then dn must be written as

$$dn=\frac{P(E,\,eU_c)dE}{W}, \tag{4}$$

where $P(E,eU_c)$ is the probability that an electron with an initial energy eU_c is still located in the cavity when its energy has been reduced to E.

In Ref. 3 an expression was derived for the average path length L traversed by an oscillating electron in the cavity before it exits

 1063-7842/97/050495-04$10.00

$$L=\frac{4V}{S_a}, \tag{5}$$

where V is the volume of the cavity.

Accordingly we can write the probability that an electron is still in the cavity after traversing a distance l as

$$P(l)=\exp\left(-\frac{l}{L}\right)=\exp\left(-\frac{lS_a}{4V}\right). \tag{6}$$

Using the approximation of continuous slowing, we can write down the expression for the energy loss dE in a path length dx:

$$dE=-\frac{Wdx}{l_i}=-W\sigma_i\,Ndx, \tag{7}$$

where l_i is the mean range before ionization, σ_i is the ionization cross section, and N is the density of the neutral gas.

To simplify subsequent calculations, we take $\sigma_i\sim$ const. Then integrating Eq. (7), we obtain a simple expression that connects the path traveled by an electron to its energy

$$l=\frac{eU_c-E}{W\sigma_i N}. \tag{8}$$

Substituting this expression into Eq. (6), we obtain

$$P(E,eU_c)=\exp\left(-\frac{eU_c-E}{W\sigma_i\,NL}\right), \tag{9}$$

and integrating Eq. (4) between the limits from 0 to eU_c, we arrive at an expression for the average number of ions formed by an electron in the cavity

$$n=\sigma_i NL\left(1-\exp\left(-\frac{eU_c}{W\sigma_i\,NL}\right)\right). \tag{10}$$

It would perhaps be more correct physically to take E_i rather than 0 as the lower limit of integration, but since $E_i\ll eU_c$, as discussed above, replacing E_i by 0 is entirely admissible. Substituting this expression into the self-consistency condition for a discharge, Eq. (3), and using Eq. (5) and the relation $N=p/kT$, where p and T are pressure and temperature of the gas, we obtain an expression that relates the voltage to the pressure

$$\frac{4\gamma pV\sigma_i}{S_akT}\left(1-\exp\left(-\frac{ekTU_cS_a}{4pVW\sigma_i}\right)\right)=1. \tag{11}$$

Introducing the notation

$$p_0=\frac{kTS_a}{4\gamma V\sigma_i} \tag{12}$$

and using Eq. (1), we can transform Eq. (11) to a more compact form

$$\frac{U_c}{U_0}=-\frac{p}{p_0}\ln\left(1-\frac{p_0}{p}\right). \tag{13}$$

With the expression written this way, the physical meaning of p_0 becomes obvious. At $p=p_0$ the logarithm in Eq. (13), and hence U_c, goes to infinity. That is, p_0 is none other than the critical pressure at which the discharge cannot burn. The value of U_0 corresponds to the discharge voltage at fairly high pressures, where the electrons deposit nearly all their energy before exiting at the anode. It is easy to see that this formula in the appropriate limit describes how at high enough p the voltage comes to a constant value, as has been experimentally observed. Indeed, taking $p\gg p_0$ and expanding the logarithm in a series, we obtain $(U_c/U_0)\approx-(p/p_0)(-p_0/p)=1$.

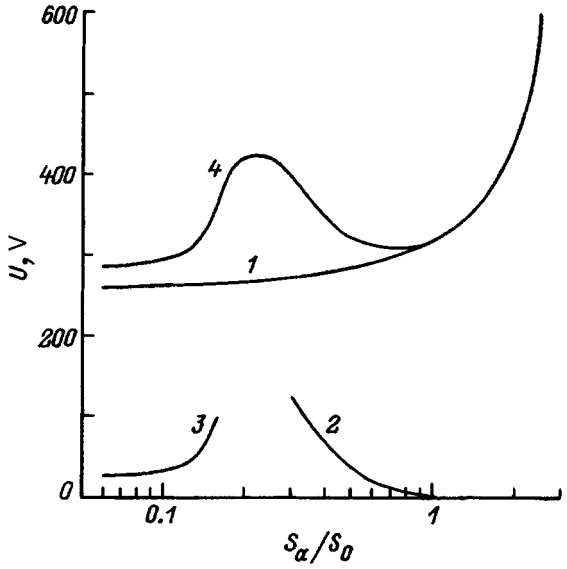

FIG. 1. Cathode fall (*1*), positive anode fall (*2*), voltage across the double layer (*3*), and discharge voltage (*4*) plotted against the anode area. Argon, $p=$ 0.1 Pa, $I_d=30$ mA, $S_c=271$ cm^2, $a=b=1,\gamma=0.1$.

The qualitative shape of the function $U_c(p)$ agrees with experiment. The main result of this investigation has been to obtain a relation that explains the experimentally observed right-ward shift of the characteristic with increasing S_a, which results in a proportional increase in the critical pressure. However, it should be noted that in expression (12), $p_0\propto S_a/V$, whereas in Ref. 3 the determining role of the parameter S_a/S_c was emphasized. This is likely due to the fact that under the conditions of the experiments reported in Ref. 3 the value of S_c essentially varied, as did V, in proportion to the length of the cylindrical cavity. In Refs. 4 and 5, where the cavity was of smaller volume for the same parameters S_a/S_c, the discharge burned at higher pressures. Therefore our relation $p_{cr}\propto S_a/V$ reflects more accurately the behavior characteristic of a hollow-cathode glow discharge.

If we assume that p is specified, we can regard expression (13) as a relation between U_c and S_a. In Fig. 1, curve *1* shows the characteristic form of $U_c(S_a)$. In the calculations the working gas was assumed to be argon, for which $\sqrt{M/m}=271$, and $W=26$ eV (Ref. 1). The other parameters used in the calculations are shown in the caption to the figure.

An estimate of p_{cr} from Eq. (12) is difficult because the coefficient γ is not known exactly, Nevertheless, we can make this estimate by using the experimental value of U_0. Then, from Eq. (1) and the known value of W we can determine γ and estimate p_{cr}, using some characteristic ionization cross section for σ_i. These estimates agree in order of

magnitude with the experimental values of p_{cr} under various conditions.

Thus expression (13), which we have derived for U_c with the parameter p_0 determined from Eq. (12), describes the main qualitative behavior of a hollow-cathode glow discharge under conditions of a negative anode fall.

DISCHARGE WITH A POSITIVE ANODE FALL IN THE ANODE PLASMA

When the anode area is smaller than $\sim S_c\sqrt{m/M}$ the anode fall becomes positive. Denoting by S_0 the anode area for which the anode fall is zero, it is clear that

$$S_0 = aS_c\sqrt{m/M}, \tag{14}$$

where a is a coefficient of the order of unity.

As the anode size decreases below S_0, the surface area of the plasma from which the electrons reach the anode does not change, but remains equal to S_0, which is necessary to maintain equality between the ion current to the cathode and the electron current to the anode, while the positive anode fall U_a and the length of the anode sheath l_a increase. If we ignore ionization processes in the anode sheath, we can estimate U_a using the formula for the ''3/2'' law in various geometries.[6] In cylindrical geometry, if we neglect the electron current to the end of the anode rod, we can estimate the radius R_0 of the plasma surface from the formula

$$R_0 = \frac{S_0}{2\pi L_a}, \tag{15}$$

where L_a is the length of the anode.

Curve *2* in Fig. 1 shows the characteristic shape of $U_a(S_a)$ in cylindrical geometry. In the calculations the radius R_a of the anode was taken to be 1 mm, and S_a varied because of variations in the length L_a. As one can see from the figure, even for a small discharge current I_d (in the calculation and in the experiment of Ref. 5, I_d was taken to be 30 mA), U_a can make a substantial contribution to the total voltage drop across the discharge gap, which in the present case can be taken to be $U_d=(U_c+U_a)$. This is the reason why, as observed in Ref. 5, U_d increases with decreasing diameter of the planar anode or decreasing length of the anode rod below some optimum value. With increasing I_d the increase in U_a, and hence in U_d will be all the greater.

As l_a and U_a increase, so does the probability of ionization processes, which can result in breakdown of the electrode sheath and the formation of a secondary plasma near the electrode, separated from the main discharge plasma by an electrical double layer.[7] In the double layer the ratio of the ion flux to the electron flux is $\sqrt{m/M}$ (Ref. 6). Therefore, in order for the plasma–double-layer system near the anode to exist the electrons entering the double layer from the main plasma must create ionization in the region of the anode plasma with a probability

$$P_i = b\sqrt{m/M}, \tag{16}$$

where b is a coefficient of the order of unity, which allows for the fact that some of the ions from the anode plasma can escape into the anode, and not enter the double layer.

Let us consider what the conditions are for satisfying relation (16) in the case of cylindrical geometry. Neglecting the thickness of the double layer, we can write the following relation for the ionization probability in the anode plasma

$$P_i = N\sigma_i(eU_l)(R_0-R_a), \tag{17}$$

where U_l is the voltage across the double layer and $\sigma(eU_l)$ is the ionization cross section for electrons that have been collected in the double layer with an energy eU_l.

Equating the right-hand sides of Eq. (16) and (17), and using Eqs. (14) and (15) and the relation $p=NkT$, we obtain a relation connecting U_l and p in the following implicit form

$$\sigma_i(eU_l) = \frac{2\pi bkTL_a\sqrt{m/M}}{p(S_0-S_a)}. \tag{18}$$

Since the form of $\sigma_i(E)$ is quite well known, the calculation of $U_l(p)$ or $U_l(S_a)$ does not present any particular difficulties. Curve *3* in Fig. 1 show the characteristic form of $U_l(S_a)$.

We note that since the ionization cross section has a maximum, for a given $S_a<S_0$ the plasma–double-layer system near the anode cannot exist if the pressure is lower than

$$p_l = \frac{2\pi bkTL_a\sqrt{m/M}}{\sigma_m(S_0-S_a)}, \tag{19}$$

where σ_m is the maximum ionization cross section.

Correspondingly, for a given p an increase in S_a above the value

$$S_l = S_0 - \frac{2\pi bkTL_a\sqrt{m/M}}{\sigma_m p} \tag{20}$$

results in decay of the plasma–double-layer system near the anode.

In Fig. 1 we have also plotted curve *4*, which shows the characteristic form of $U_d(S_a)$, which is given by the relations $U_d=U_c+U_l$ for $S_a<S_l$, $U_d=U_c+U_a$ for $S_l<S_a<S_0$, and $U_d=U_c$ for $S_a>S_0$. If S_a is slightly larger than S_l, we cannot calculate the function $U_a(S_a)$ by the 3/2-power law, since ionization processes proceed rapidly in this part of the anode sheath. In this section curve *4* has been approximated by a spline. The shape of curve *4* agrees qualitatively with the experimental results obtained in Ref. 5. Using the coefficients a, b, and γ as fitting parameters, we can obtain fair quantitative agreement.

It formally follows from Eq. (19) that with decreasing L_a the critical pressure p_l can be reduced to an arbitrarily small value. However, as L_a is reduced we can no longer neglect the electrons that arrive at the ends of the anode rod. The results of Ref. 7 have shown that the surface of the plasma region that forms near a small electrode has a nearly spherical shape. If we assume that the surface of the main discharge plasma is nearly spherical and has an area S_0, we can derive the following expression for the critical pressure p_m below which the plasma–double-layer system cannot exist for any arbitrarily small anode area:

$$p_m = \frac{bkT}{2\sigma_m}\sqrt{\frac{\pi m}{MS_0}}. \tag{21}$$

For $p<p_m$ the function $U_d(S_a)$ has only two characteristic parts, corresponding to negative and positive anode falls. Estimates calculated with Eq. (21) agree with experimental results.[4,5] The point at which the curve $U_d(S_a)$ is a minimum must of necessity coincide with S_0, because if S_a falls below S_0 then U_c continues to decrease, and this decrease must at some point outweigh the beginning of the increase in U_a. When the discharge burns through an aperture to the anode, which is located some distance l_{aa} from it, a decrease in its area below S_0 and the transition to discharge conditions with a positive anode fall result in an abrupt increase in the voltage ($U \propto l_{aa}^{4/3}$). The discharge can then occur with a small U_d only when there is an electric double layer and an anode plasma. The conditions for the existence of a double layer can be determined by assuming that the ions entering the double layer are generated in a region having a characteristic size of about the radius r_a of the aperture. Then the relation connecting U_l and p takes the form

$$\sigma_i(eU_l)=\frac{kT\sqrt{m/M}}{r_a p}. \tag{22}$$

The qualitative form of $U_l(p)$ calculated from Eq. (22) agrees with the experimental results given in Ref. 3. A fair quantitative agreement with experiment can be obtained by varying T, which in Ref. 3 varied with increased discharge current. However, in the region of large p the calculated and experimental curves are quite different because the calculations neglected the initial energy the electrons have when they enter the double layer. Because the electrons have an initial energy $\sim kT_e$ it is possible for the double layer to exist for U_l less than the ionization potential. The value of U_l increases with decreasing p, and for p lower than

$$p_2=\frac{kT\sqrt{m/M}}{r_a\sigma_m} \tag{23}$$

no double layer can exist in the aperture of the cavity. As discussed above, and as has been confirmed experimentally,[3] the discharge voltage increases abruptly.

Since the function $\sigma_i(E)$ is nonmonotonic, two solutions to Eqs.(22) and (18) are generally possible: one on the ascending part and the other on the descending part. The states of the plasma– double-layer system described by the second solution are unstable. Indeed, we assume that the rate of the ionization processes has decreased randomly, causing a decrease in the electron space-charge compensation in the double layer by ions entering from the anode plasma. As a result, the voltage across the double layer increases, which on the descending part of $\sigma_i(E)$ causes an additional decrease in the ionization so that the system goes out of the state of equilibrium, while on the ascending part an increase in U_l enhances the ionization and the random deviation is compensated. Of course, this argument cannot be said to have exhausted the possibilities. To solve the problem of the stability of any particular state, and also to determine the parameters a and b requires a more detailed analysis of the processes occurring in the anode region of the hollow-cathode glow discharge. This may be the subject of further work.

CONCLUSIONS

This investigation has analyzed the conditions for the existence of various regimes of burning of a hollow-cathode glow discharge — with a negative and a positive anode fall, and also the regime with an anode plasma and an electric double layer. The relations obtained, Eqs. (13), (18) and (22), which relate the voltage across the cathode and across the double layer to the dimensions of the anode or the aperture and to the pressure, are in qualitative agreement with experimental results. Simple expressions, (12), (21), and (23), have been obtained for the critical pressure below which the various kinds of discharge cannot occur. These expressions have provided estimates that agree in order of magnitude with experimental results.

[1] A. S. Metel', Zh. Tekh. Fiz **55**, 1928 (1985) [Sov. Phys. Tech. Phys. **30**, 1133 (1985)].
[2] S. P. Nikulin, Zh. Tekh. Fiz **62**(12), 21 (1992) [Sov. Phys. Tech. Phys. **37**, 1142 (1992)].
[3] A. S. Metel', Zh. Tekh. Fiz **54**, 241 (1984) [Sov. Phys. Tech. Phys. **27**, 141 (1984)].
[4] V. V. Bersenev, N. V. Gavrilov, and S. P. Nikitin, *Proceedings of the 22nd International Conference on Phenomena in Ionized Gases*, Hoboken, NJ (1995), Vol. 2, pp. 103–104.
[5] V. V. Bersenev, N. V. Gavrilov, and S. P. Nikulin, *Proceedings of the Conference ''Physics of Low-Temperature Plasmas''* [in Russian], Petrozavodsk (1995), Part 2, pp. 251–253.
[6] V. L. Granovskiĭ, *Electric Current in Gases. Steady-State Current* [in Russian] (Nauka, Moscow, 1971).
[7] B. N. Klyarfel'd, A. A. Timofeev, and N. A. Neretina, Zh. Tekh. Fiz **25**, 1581 (1955).

Translated by J. R. Anderson

Optical breakdown threshold in the electron-thermal model of defect generation

V. L. Komolov

S. I. Vavilov All-Union Scientific Center, State Optical Institute, 199034 St. Petersburg, Russia
(Submitted December 27, 1995)
Zh. Tekh. Fiz. **67**, 48–53 (May 1997)

An investigation is made of the conditions for the nucleation of optical breakdown in transparent material when recombination-stimulated defect-formation reactions occur in it. It is shown that a positive feedback between the conduction electron concentration and point defects activates defect formation even if the medium is not heated. Under real conditions, where heating of the medium by the light is important, lowering of the activation barrier by thermal defect generation aided by conduction electrons results in optical breakdown of the medium at light intensities much lower than predicted in the classical ''semiconductor'' or ''thermochemical'' models of thermal breakdown. The analysis confirms that optical breakdown of transparent condensed media is due to electron-aided defect formation reactions over a broad range of illumination conditions. © *1997 American Institute of Physics.* [S1063-7842(97)00905-7]

INTRODUCTION

The causes of nucleation of optical breakdown in weakly absorbing nonconducting media illuminated by intense optical radiation in the visible and infrared regions and the conditions under which this nucleation takes place are still matters of controversy. Despite considerable efforts extending over more than two decades, the physical models proposed to describe optical breakdown and light-induced damage in transparent materials have not yet provided correct predictions of the behavior of materials subjected to intense irradiation by light waves, and frequently they contain internal contradictions. These difficulties are overcome by introducing into the theoretical model fitting parameters, the choice of whose values is not always well grounded.

The most notable progress in the theoretical description of optical destruction has come with the notion of the thermal nature of optical breakdown.[1–5] The model of optical breakdown is based on the idea that light-induced destruction of transparent materials is due to heating of the medium to a high temperature at which the structure and properties of the crystal lattice of the material change (for example it may melt, evaporate, lose elasticity, etc.). The difficulty in accounting for the heating of a material that initially absorbs light only slightly is overcome by introducing into the theoretical model positive feedback, which provides further absorption as the medium is heated.

Taking into account feedback in describing the nucleation of thermal breakdown solves another problem in data analysis — the problem of selecting the breakdown criterion. If the optical parameters of the medium depend strongly enough on the temperature (in most model this dependence reflects an activation law, $\propto \exp(-U/kT)$), its heating has a threshold. When the light intensity q is above some value q^* appropriate to the given experimental conditions, the temperature of the medium begins to rise exponentially with the time, and it can be assumed that under conditions of progressive thermal instability in the medium any criterion of beam damage will be satisfied almost immediately after the heating enters the exponential regime.

Some models of thermal breakdown taking into account various feedback mechanisms exist at present.

In the so-called ''semiconductor'' model, which was first put forth in Ref. 1 and experimentally verified for germanium heated by infrared laser radiation,[2] the positive feedback is due to the valence electrons being promoted into the conduction band by thermal fluctuatins, which is a process whose probability increases exponentially with the temperature of the medium. The height W of the activation barrier in this model is related to the width E_g of the band gap of the semiconductor: $W=E_g/2$. The absorption of light in free–free transitions, which is proportional to the concentration n_e of free carriers, causes the temperature T to rise, with a consequent further increase in n_e and in the absorption coefficient $\varkappa$, during which process the higher the light intensity the larger the feedback coefficient.

In the thermochemical model[3] the onset of optical breakdown in a solid is related to thermal generation of point defects, a process whose probability also is activated. A new aspect that comes into this model as compared to the semiconductor model is that along with the generation of point defects, electronic states appear in the band gap of the material, bringing in fundamental changes in the mechanisms for the increase in the light absorption with temperature. These changes involve a large increase in the concentration n_e of free carriers above the equilibrium concentration at the given temperature through multistep excitation of carriers by the light into the conduction band via intermediate energy levels. As a result, the concentration of conduction electrons and the light absorption induced by these electrons take on a complicated temperature dependence, which is determined not only by the band gap as in the semiconductor model, but also by the activation energy W_0 of a point defect. It has been shown[3] that in this model thermal instability sets in at temperatures 2–3 times lower than that predicted by the semiconductor model.

Electron-thermal generation of defects

The models mentioned above for the thermal instability of weakly absorbing materials have helped to advance

greatly our understanding of the mechanisms for the nucleation of optical breakdown. In particular, modifications of the models so they will describe local absorption have allowed investigators to interpret a number of experimentally observed features in the behavior of optical damage.[4,5] Moreover, it has become quite apparent that these models contain internal limitations in that the feedback mechanisms mentioned above operate only when the media are heated to high temperatures, and this heating becomes harder to attain as the material becomes purer and the band gap wider (or the defect activation energy higher). This circumstance compels us to revisit the problem of thermal optical breakdown and search for processes that are activated at lower temperatures than the processes investigated heretofore. This article presents the results of an analysis of one of the mechanisms for lowering the activation barrier to the generation of point defects in a weakly absorbing crystal and to see how it influences the onset of optical breakdown.

The concepts used in the thermal-fluctuation mechanism for the formation and annihilation of point defects in a crystal lattice have been widely used to describe the results of laser action on materials with a relatively low binding energy. It is clear, however, that for media in which the activation energy for defect formation is 2–3 eV the probability of thermal generation of a defect is essentially zero when the material is heated to only a few hundred degrees.

For a medium with a high concentration of free carriers (including those created by the radiation), another known mechanism in addition to the thermal fluctuation mechanism for defect formation, which is related to the generation of defects in the decay of electronic excitations, is the mechanism of recombination-stimulated defect reactions.[6] The participation in the process by an excited carrier trapped in a local level of a defect as the latter is created effectively lowers the height of the activation barrier for the thermal-fluctuation process of defect formation. The cause of lowering the barrier is that the electron by its field deforms the lattice in its vicinity, and this deformation in turn facilitates localization of the electron at the place where the ideal periodic structure is distorted. The concepts associated with recombination-stimulated defect reactions have been of practical use in analyzing the causes for the degradation of the parameters of semiconductor heterolasers.[7] The microscopic theory of nonradiative recombination at a defect and the use of this theory for analyzing a number of recombination-stimulated reactions are presented in the review, Ref. 8.

In most articles (e.g., Refs. 6–8) the light-induced defect formation has been analyzed for conditions where the photon energy is sufficiently high to produce electron–hole pairs by interband absorption of the light in the crystal. The analysis devolves into two consecutive steps: 1) calculation of the concentration of light-generated free carriers and 2) calculation of the probability of defect formation (thermally generated or electron-induced) for a known concentration of free carriers. A qualitatively different picture is seen when the semiconductor is illuminated in the region of its transparency.[3] Then electron–hole pairs are not generated by interband excitation, and the only supplier of electrons to the conduction band are lattice defects, whose energy levels create a ''ladder'' of permitted states in the band gap of the crystal. Such a ladder permits multistep excitation of electrons into the conduction band when the crystal is illuminated with photons of energy less than the band gap. When the electrons appear in the conduction band the light absorption of the crystal increases, and the crystal heats up, which stimulates thermal generation of additional defects, while in turn the generation of defects produces new ladders in the band gap, which increase the rate of excitation of electrons into the conduction band, and so on.

In the recombination-stimulated generation of defects in a light-irradiated transparent material an important factor is that the presence of free carriers in the conduction band aids in the generation of defects even if the material is not heated up.[9] Because the electron-aided rate of defect generation is proportional to the electron concentration n_e, and this concentration is in turn determined by the concentration of intermediate levels in the band gap, that is, the concentration of defects, N_v, the concentrations n_e and N_v are interdependent. This positive feedback between the electron and defect concentrations causes a fast increase in the number of defects or their steady accumulation with repeated exposure to light. The first case corresponds to optical breakdown characterized by a sharp threshold and results in breakdown of the crystal, and the second corresponds to a gradual degradation of its optical characteristics with defect formation (the so-called accumulation effect).

Absorption of light in the model of defect generation.

Following the analysis method used in Ref. 9, we consider a material with an initial concentration of point defects, N_0 (to be specific, we take these to be vacancies), whose deep acceptor levels E_v are filled with electrons n_v (their initial concentration is $n_{v0}=N_0$). If light is incident on the material with an intensity J ($\mathrm{cm^{-2}s^{-1}}$) and a photon energy $\hbar\omega$ sufficient for stepwise excitation of electrons from the valence band p to the conduction band c through an intermediate level v (the absorption cross sections for the p–v and v–c transitions are σ_1 and σ_2, respectively) we obtain a system of balance equations for the nonequilibrium vacancy concentration N_v and electron concentration in the conduction band, n_e and in the acceptor level, n_v

$$\frac{dN_v}{dt}=\lambda+\eta n_e-\lambda' N_I(N_v-n_v)-\eta' N_I n_v,$$

$$\frac{dn_v}{dt}=J\sigma_1(N_v-n_v)-J\sigma_2 n_v+\eta n_e+\gamma_v(N_v-n_v)n_e - \eta' N_I n_v,$$

$$\frac{dn_e}{dt}=J\sigma_2 n_v-\eta n_e-\gamma_v(N_v-n_v)n_e+\eta' N_I n_v-n_e/\tau_r,$$

$$N_I=N_v-N_0, \quad N_v(0)=n_v(0)=N_0, \quad n_e(0)=0. \qquad (1)$$

Here $\lambda+\eta n_e$ is the total rate of defect generation (the second term describes the electron-aided process), $\lambda' N_l(N_v-n_v)$, and $\eta' N_l n_v$ are the rates of annihilation of empty vacancies and vacancies containing an electron at the level E_v (in the latter case the electron is promoted to the

conduction band); $\gamma_v(N_v-n_v)n_e$ is the rate of trapping conduction electrons at the level E_v, and $1/\tau_r$ is the rate of recombination of conduction electrons.

The coefficients that enter into Eqs. (1) depend on the temperature through an activation relation and have the form[10]

$$\eta=a^3\nu_0N^2/Q\exp\left(-\frac{W+E-E_v}{T}\right),$$

$$\lambda'=\eta'=a^3\nu_0\exp\left(-\frac{E}{T}\right),$$

$$\lambda=a^3\nu_0N^2/Q\exp\left(-\frac{W+E}{T}\right), \tag{2}$$

where a is the lattice constant, ν_0 is the frequency of atomic vibrations, N is the concentration of lattice sites, Q is the density of states in the band, and E is the activation energy for the diffusion of a defect.

For the subsequent analysis of optical breakdown of the medium we must have the explicit form of $\varkappa(T,q)$ in the framework of our electron-heating model. Under steady-state conditions

$$\frac{dN_v}{dt}=\frac{dn_v}{dt}=\frac{dn_e}{dt}=0$$

the system of Eqs. (1) can be written in a somewhat simplified form as

$$\eta n_e-\eta' N_v n_v=0, \tag{3a}$$

$$J\sigma_1(N_v-n_v)-J\sigma_2 n_v+\eta n_e-\eta' N_v n_v=0, \tag{3b}$$

$$J\sigma_2 n_v-\eta n_e+\eta' N_v n_v-n_e/\tau_r=0. \tag{3c}$$

In writing Eqs. (3) we have made the following simplification: we neglect the thermal-fluctuation-induced generation of defects without the participation of an electron, since, as shown in Ref. 9, $\lambda\ll\eta n_e$ for any realistic values of n_e, and have neglected the terms $\lambda' N_I(N_v-n_v)$ and $\gamma_v(N_v-n_v)n_e$, which have little effect on the results of the analysis for $(N_v-n_v)/N_v<1$, and further we have assumed that under steady-state conditions a fairly large number of defects have been generated $(N_v\gg N_0)$ $N_I=N_v-N_0\approx N_v$.

The correctness of these assumptions and their physical meaning will be discussed later. With these assumptions, a comparison of Eqs. (3a) and (3b) gives the relation

$$n_v=\frac{\sigma_1}{\sigma_1+\sigma_2}N_v, \tag{4}$$

and after substituting Eq. (4) into (3a), we obtain

$$n_e=\frac{\eta'}{\eta}N_v n_v=\frac{\eta'}{\eta}\frac{\sigma_1}{\sigma_1+\sigma_2}N_v^2. \tag{5}$$

On the other hand, we find from Eqs. (3a) and (3b) that

$$n_e=J\sigma_2\tau_r n_v=J\sigma_2\tau_r\frac{\sigma_1}{\sigma_1+\sigma_2}N_v. \tag{6}$$

A comparison of Eqs. (5) and (6) gives an expression for the defect concentration N_v

$$N_v=J\sigma_2\tau_r\frac{\eta'}{\eta}, \tag{7}$$

while substituting Eq. (7) into (4) and (5) gives the expressions for n_v and n_e, respectively

$$n_v=J\sigma_2\tau_r\frac{\sigma_1}{\sigma_1+\sigma_2}\frac{\eta}{\eta'}, \tag{8}$$

$$n_e=(J\sigma_2\tau_r)^2\frac{\sigma_1}{\sigma_1+\sigma_2}\frac{\eta}{\eta'}. \tag{9}$$

Using the fact that the crystal is heated mainly through the absorption of the light by free carriers, we obtain the form of the dependence of the absorption $\varkappa$ on the intensity q of the light flux and the temperature T

$$\varkappa(T,q)=\sigma_e n_e=(J\sigma_2\tau_2)^2\frac{\sigma_e\sigma_1}{\sigma_1+\sigma_2}\frac{\eta(T)}{\eta'(T)}, \tag{10}$$

or after substituting in the explicit forms of the temperature dependences (2) for $\eta(T)$ and $\eta'(T)$

$$\varkappa(T,q)=\frac{N^2}{Q}\frac{1}{(\hbar\omega)^2}(q\sigma_2\tau_r)^2\frac{\sigma_e\sigma_1}{\sigma_1+\sigma_2}\exp\left(-\frac{W-E_v}{T}\right)$$

$$=q^2\beta\exp\left(-\frac{W-E_v}{T}\right)=\varkappa_0(q)\exp\left(-\frac{U}{T(r)}\right). \tag{11}$$

Critical temperatures and breakdown threshold

In order to analyze the heating and optical breakdown of the medium, the system of equations (3) must be supplemented with the heat conduction equation with a source whose power is determined by the intensity of the light flux, $q=\hbar\omega J$, and the absorption coefficient $\varkappa$, which itself is a function of the temperature T $(\varkappa=\varkappa(T))$

$$\frac{\partial T}{\partial T}=a\nabla^2T+\frac{\varkappa(T)q}{\rho c} \tag{12}$$

with the appropriate initial and boundary conditions.

Putting aside a complete analysis of the solutions of the system of equations (3) and (12), we estimate the critical temperature and the breakdown threshold by employing some simple but universal relations derived in the context of the semiconductor model of thermal instability[3–5]

We recall that the instability of the temperature field during heating of the medium by light, where the absorption coefficient is a rapidly rising function of the temperature, results in conditions where the light-induced light absorption in the medium exceeds the initial absorption.[5] The method of analyzing the loss of stability by a temperature field was introduced into high-power optics from the theory of thermal explosions in exothermal chemical reactions[11] and is carried out according to the scheme described below for the very simple case of heating of a medium in the vicinity of an absorbing inclusion. In a transparent medium a steady-state temperature profile of the form

$$T(r)=\frac{qAR}{K}\frac{R}{r}=T_0\frac{R}{r}\quad(\text{for } r>R) \tag{13}$$

is formed around an inclusion or radius R with an absorption coefficient A, where q is the intensity of the incident light flux, K is the thermal conductivity of the medium, and $T_0=(qAR)/K$ is the temperature inside the inclusion.

If an induced absorbing ''halo,'' $\varkappa[T(r)]$, is present in the vicinity of the inclusion, the temperature field is distorted because of an additional volume heat source. A simple estimate of the extra heating of the inclusion, taking into account the effective absorbance of the halo, gives the relation

$$T(q)=\frac{qAR}{K}+\frac{q}{K}\int_R^\infty \varkappa[T(r)]rdr \tag{14}$$

for the temperature $T(q)$ of the inclusion.

If $\varkappa$ depends strongly enough on T (for example if it is activated, and proportional to $\exp(-U/T)$ as is characteristic all three models mentioned above), then the solution (14) exists only in a limited range of light intensities $q<q_*$. Here the issue is a physically meaningful low-temperature solution, since the high-temperature steady-state solution $T\approx U$ for a source of limited power always exists. The stability threshold of the low-temperature solution is determined by the condition[4,5]

$$R\varkappa[T_0(q,R)]\gtrsim A, \tag{15}$$

in which T_0 is related to q and R by relation (13).

Let us return to expression (14) for the temperature. Recalling that according to Eq. (11),

$$\varkappa(T,q)=\varkappa_0(q)\exp\left(-\frac{U}{T(r)}\right),$$

by substituting $T(r)=T_0(R/r)$ and carrying out the integration, we find with accuracy to terms in T_0/U

$$T(q)=T_0\left[1+\frac{T_0}{U}\frac{\varkappa_0 R}{A}\exp\left(-\frac{U}{T_0}\right)\right]. \tag{16}$$

It immediately follows from Eq. (16) that the additional heating associated with absorption by the halo (the second term in (16)) becomes large if a criterion equvalent to relation (15) is satisfied:

$$\frac{\varkappa_o R}{A}\exp\left(-\frac{U}{T_0}\right)\geqslant 1, \tag{17}$$

i.e., starting from the temperatures

$$T_*\gtrsim\frac{U}{\ln\left(\dfrac{\varkappa_0 R}{A}\right)}. \tag{18}$$

Using the relation (13) between the temperature and the light flux, we obtain an expression for the critical light intensity q_* above which a thermal instability develops in the medium

$$q_*=\frac{KT_*}{AR}=\frac{KU}{AR}\frac{1}{\ln\left(\dfrac{\varkappa_o R}{A}\right)}. \tag{19}$$

Discussion and conclusions

Before presenting the final results of this analysis of the conditions for optical breakdown, let us discuss briefly on the admissibility of the assumptions made in writing down the system of equations (3).

1) The condition $\lambda\ll\eta n_e$ is satisfied beginning with

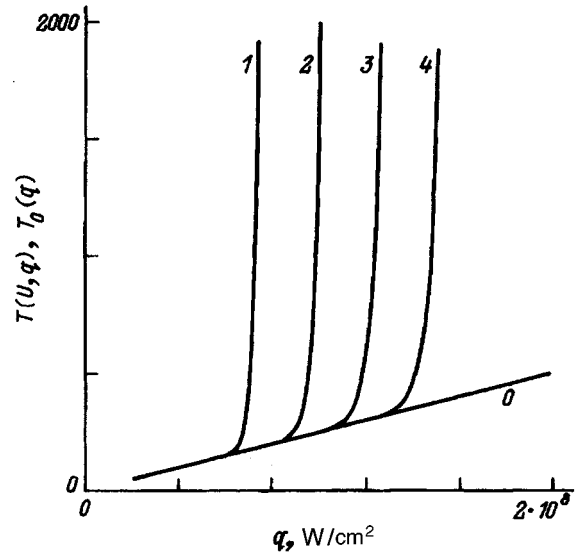

FIG. 1. Temperature of the medium versus the intensity of the light flux and the activation barrier height. U, K: *1* — 5000, *2* — 7000, *3* — 9000, *4* — 11 000. The straight line O is the heating without the induced absorption.

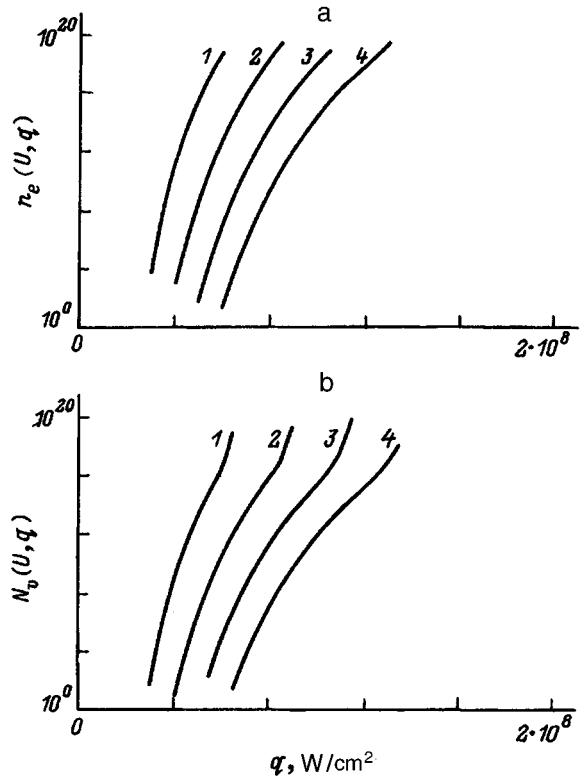

FIG. 2. Electron (*a*) and defect (*b*) concentrations versus the light flux intensity and activation barrier height. U, K: *1* — 5000, *2* — 7000, *3* — 9000, *4* — 11 000; the straight line has the same meaning as in Fig. 1.

$$n_e > n_* = \lambda/\eta = Q\exp\left(-\frac{E_v}{T}\right). \quad (20)$$

At $T=300$ K and $Q=5\times10^{18}$ cm^{-3}, the value of n_* varies from $n_*\sim10^2$ cm^{-3} for $E_v=1$ eV to $n_*\sim10^{12}$ cm^{-3} for $E_v=0.4$ eV; i.e., the thermal generation of defects without the participation of an electron can be neglected for any reasonable value of E_v.

2) According to Eq. (4), the condition $(N_v-n_v)/N_v<1$ imposes a constraint on the ratio of the transition cross sections σ_1 and σ_2 and is satisfied for $\sigma_1/\sigma_2<1$. This constraint, which was made to simplify the analysis, does not affect the overall result, but must be borne in mind in specific calculations using the scheme proposed here, or else one must make a more rigorous analysis of the populations on the basis of the complete set of equations (1).

3) The condition $N_V \gg N_0$ is satisfied better the higher the intensity q of the radiation and the more perfect the initial structure of the crystal. Generally the initial defects (N_0) act only as nuclei for the light-induced generation of vacancies, and in intrinsic materials their concentrations are negligible compared to the concentration of light-induced defects.

Therefore the assumptions made to simplify the estimates do little to limit the generality of the analysis, and have little effect on the main conclusions. The main conclusions can be summed up as follows.

1) Recombination-stimulated defect generation with the participation of conduction electrons plays the decisive role in laser modification of the structure of transparent crystalline media, and produces a large number of point defects even when the medium is not heated.

2) The proposed electron-thermal model of optical breakdown of transparent media is characterized by an activation barrier that is considerably lower (by 0.5–1 eV) than its typical value in the ''semiconductor'' or the ''thermochemical'' models.

3) The comparatively low activation energy U stimulates the fast growth in the defect concentration in a transparent material and produces thermal instability and optical breakdown of weakly absorbing media at reasonable temperatures T_* of the medium and moderate light fluxes q_*. This conclusion is illustrated in Figs. 1 and 2 and by the estimates listed in Table I for a typical set of parameters of a wide-gap semiconductor.

The calculations were carried out with Eqs. (18) and (19) for the following parameter values: $N=3\times10^{23}$ cm^{-3}, $Q=5\times10^{18}$ cm^{-3}, $\hbar\omega=1.17$ eV, $k=0.4$ W/cm·deg, $A=10^{-2}$, $R=10^{-4}$ cm, $\sigma_1=4\times10^{-18}$ cm^2, $\sigma_2=10^{-18}$ cm^2, $\sigma_e=10^{-18}$ cm^2, and $\tau_r=10^{-6}$ s.

TABLE I. Critical temperature T_* and breakdown threshold q_* for various activation barrier heights U.

U, K	$T_*(U)$, K	$q_*(U)$, MW/cm^2
5000	166	67.2
7000	235	94.2
9000	302	121.0
11000	369	148.0
13000	437	174.8
15000	504	201.7

The final conclusion of this analysis is that optical breakdown of a broad class of transparent condensed media under typical conditions of optical irradiation is due to electron-aided defect formation with the subsequent development of a thermal instability, whose onset has a distinct threshold. Below-threshold generation of defects is rapid, takes place even at low temperatures, and results in optical ''fatigue'' of the medium, i.e., gradual degradation of its optical characteristics with repeated illumination.

This work was carried out with the support of the International Science Foundation and the Russian Fund for Fundamental Research.

[1] É. M. Epshteĭn, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **15**, 33 (1972).
[2] P. A. Young, Appl. Opt. **10**, 638 (1971).
[3] A. M. Bonch-Bruevich, V. L. Komolov, M. N. Libenson, and A. G. Rumyantsev, Zh. Tekh. Fiz. **55**, 107 (1985) [Sov. Phys. Tech. Phys. **30**, 61 (1985)].
[4] S. I. Anisimov and B. I. Makshantsev, Fiz. Tverd. Tela (Leningrad) **15**, 1090 (1973) [Sov. Phys. Solid State **15**, 743 (1973)].
[5] I. V. Aleshin, S. I. Anisimov, and A. M. Bonch-Bruevich, Zh. Éksp. Teor. Fiz. **70**, 1214 (1976) [Sov. Phys. JETP **43**, 631 (1976)].
[6] V. N. Abakumov, A. A. Pakhomov, and I. N. Yassievich, Fiz. Tekh. Poluprovodn. **25**, 1489 (1991) [Sov. Phys. Semicond. **25**, 899 (1991)].
[7] M. I. Klinger, Ch B. Lushchik, T. V. Mashovets *et al.*, Usp. Fiz. Nauk **147**, 523 (1985) [Sov. Phys. Usp. **28**, 994 (1985)].
[8] I. N. Yassievich, Semicond. Sci. Technol. **9**, 1433 (1994).
[9] V. L. Komolov, Zh. Tekh. Fiz **64**(7), 64 (1994) [Tech. Phys. **39**, 667 (1994)].
[10] V. L. Vinetskiĭ and G. A. Kholodar', *The Statistical Interaction of Electrons and Defects in Semiconductors* [in Russian] (Naukova Dumka, Kiev, 1969).
[11] S. I. Anisimov and E. F. Nogotov, Teplofiz. Vys. Temp. **1**, 276 (1963).

Translated by J. R. Anderson

Limiting pulse repetition rate in copper vapor lasers

P. A. Bokhan and D. É. Zakrevskiĭ

Institute of Semiconductor Physics, Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia

(Submitted December 4, 1995)
Zh. Tekh. Fiz. **67**, 54–60 (May 1997)

Two groups of processes which limit the rate-power characteristics of copper vapor lasers are examined. The first of these is related to the inadequate relaxation rate of the metastable states and accordingly their high prepulse concentration, which degrades the lasing pulse parameters. The second group of processes is initiated by the high prepulse electron concentration. It is shown that if the effect of the prepulse electrons is neutralized in pulse-periodic copper vapor lasers, then lasing starts up again in the next pulse within 1 μs after the preceding pulse stops. It is concluded that the rate of deexcitation of the metastable states is so high that it has no effect whatever on the lasing parameters in real gas-discharge lasers based on copper vapor or copper bromide vapor, and the prospects are opened up for attaining a lasing power of $\sim$10 kW/m with electron-beam pumping. © *1997 American Institute of Physics.* [S1063-7842(97)01005-2]

Introduction and Statement of the Problem

In the first experiments on copper vapor lasers and other lasers based on self-limiting transitions of metal vapors it was postulated that the repetition rate of the laser pulses was determined by the removal of atoms from the lower metastable states, such as Cu(^2D_j), to the wall. However, direct experiments carried out with copper vapor lasers at repetition rates above 10 kHz (Ref. 2) showed that there must be another faster quenching process that permits such high repetition rates. It has been shown qualitatively[3,4] that this process is electron-induced quenching in the after-pulse period. Direct measurements of the population of the metastable states, N_{ms}, in the afterglow of a pulse-periodic discharge[5] revealed that under optimum conditions the residual population of the metastable states after 10–15 μs has no effect whatsoever on the parameters of the next laser pulse. This same conclusion could have been arrived at (but was not) from an analysis of the data of earlier work.[6]

The results of measuring the electron density N_e (Ref. 7) and the plasma conductivity (Ref. 4) were invoked to make the assumption that the main limitation on the rate-power characteristics of the copper vapor laser is the high pre-pulse electron concentration N_{e0}, which initiates a large number of phenomena in the pump and laser pulses. As a final result, the action of N_{e0} redistributes the rates of pumping the upper state and the lower state in the direction of the latter as N_{e0} increases, which also limits the rate-power characteristics.

A direct measurement of the electron quenching constants for the Cu(^2D_j) states, $k_e = 1.4\times10^{-8}$ cm$^3\cdot$s^{-1} and the metastable states of other atoms,[8] as well as high-rate pulse-periodic electron-beam pumping of lasers,[8] has reinforced our conviction that the mechanism that limits the rate-power characteristics operates through the effect of N_{e0}. In recent times the model of fast relaxation of the metastable states of the copper atoms, specifically the close relation between their population and the electron temperature T_e in the afterglow, has also been used in some theoretical articles,[9] in contrast to the earlier work.[10,11]

The principal advances in attaining high average power in lasers based on the vapor of copper and its salts, beginning in Refs. 5, 12, and 13, were achieved by introducing hydrogen into the active medium,[14–16] which increases the plasma recombination rate in the afterglow in large-diameter tubes. In this case, limitations on the rate-power characteristics due to prepulse population N_{ms0} of the metastable states, if such occur (this has not been demonstrated by direct experiment), are of a thermal nature, i. e., they are due to overheating of the active medium, particularly at the center of the tube. However, the methods for overcoming this overheating, as in CO_2 lasers, are more of an engineering exercise, and differ from methods that might be initiated by the slow relaxation of the metastable states.

However, according to Tabata *et al.*,[14] for a proper arrangement of the laser pumping and optimum temperature and composition of the active medium (relative to the output power) for a tube with a diameter up to 80 mm and an output power of 100 W/m, most of the increase in the lasing power comes from the increase in the plasma recombination rate in between pulses, a result of adding hydrogen to the active mixture. The introduction of baffles to cool the mixture in the central region of the tube[17] increased the lasing power to 110 W/m. Those investigators believe that the positive effect comes both from the decrease in the gas temperature in the center of the tube and the faster recombination, also in the center of the tube . The former factor greatly reduces N_{ms0}, while the latter factor increases the rate of penetration of the pumping electric field to the center of the tube, thereby improving the uniformity of the radial profile of the pumping and lasing. However the total improvement (27% in Ref. 17 and 10% compared to the results of Ref. 14) is not so great as to call it a dramatic effect of N_{ms0} on the lasing parameters, even in tubes with such a large diameter, a conclusion that in general agrees with the data of Ref. 14.

Another promising means of increasing the rate-power

characteristics is to pump the lasers by beams of low-energy (5–10 keV) runaway electrons formed directly in the active medium of the laser.[8] According to our calculations and model experiments, the lasing power per unit length for the copper vapor lasers with electron-beam pumping can exceed 10 kW/m without using any flow-through system.[18] This is two orders of magnitude higher than that attained at the present time.[14–17] This power level can be realized only at high repetition rates of pulsed pumping and a high rate of electron quenching of the metastable states in the afterglow, so that the latter will have only a slight effect on the lasing parameters.

From the results of some theoretical[10,11] and experimental works[19–21] there are still some lingering doubts in this respect. Therefore before going on to a further step in creating high- power lasers based on copper vapor and the vapors of other metals, it would be desirable to determine how reliable is the foundation for the assumptions underlying these predictions in Ref. 8. This topic is the subject of the present paper.

Experimental method and experimental apparatus

An elucidation of the limiting characteristics of lasers based on self-limiting transitions is carried out by the methods of doubled, regular, or trains of pulses. By these methods it is very difficult to separate out the effects of N_{e0} and N_{ms0} on the lasing parameters, either theoretically or experimentally. The results of Refs. 22 and 23, in which such a separation was made, are not always taken into account.[11,19–21,24] Therefore in the present investigation we have used a new method for measuring the effect of metastable states on the lasing characteristics, which we call the method of spatially separated regions of absorption and lasing.

The essence of the method is the following. Let us assume that we have an active medium with two zones of pumping and lasing, *1* and *2*, of different length and diameter (Fig. 1). Each part can be pumped independently and with a controllable delay between the pump pulses. The aone *3* is used to reduce any effect the discharge in one part of the laser may have on the other part. The tube itself can be placed in a cavity with a low-transmission mirror *4* and an exit mirror *5*. We shall show that by shifting the pulses of the pumping zones we can measure the absolute effect of N_{ms0} on the lasing parameters and separate this effect from that of N_{e0}.

The equation for the number of photons in the cavity can be written in the following way

$$\frac{\partial N_p}{\partial t} \pm c\frac{\partial N_p}{\partial x} = B\left(N_r - \frac{g_r}{g_{ms}}N_{ms}\right)N_p^{\pm} + AN_r(d\Omega/4\pi), \tag{1}$$

where N_p, N_r, and N_{ms} are, respectively, the photon density and the densities of resonant and metastable states, g_r and g_{ms} are their statistical weights, B is the Einstein coefficient for stimulated emission, A is the probability of spontaneous decay through the operating transition, $d\Omega/4\pi$ is the fractional solid angle the cavity subtends, and the sign $\pm$ corresponds to opposite directions of propagation of light.

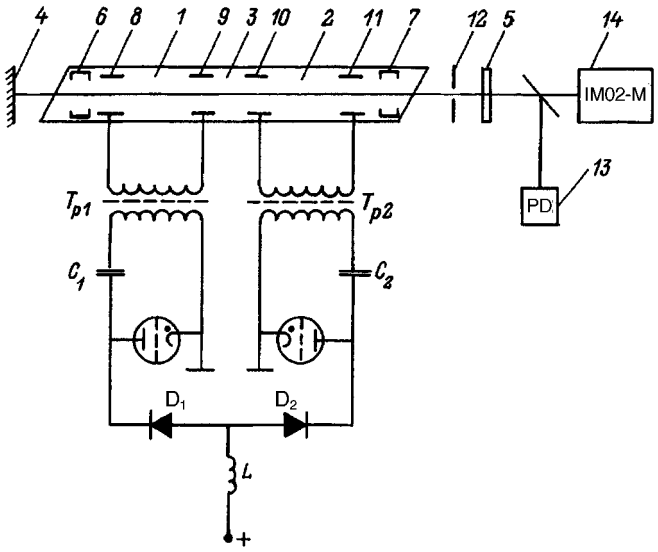

FIG. 1. Diagram of the experiment to determine the rate-power characteristics by the method of spatially separated regions of absorption and lasing.

Let us consider first the operation of the tube as an amplifier. We assume that the first zone contains a medium with zero prepulse population N_{ms}^0 and some population N_r, while the second zone contains a medium with zero population N_r and some population N_{ms}. It is clear that the result of the passage of light through this amplifier will be exactly what it would be if the light passed through any single zone where the active medium has a population N_r or N_{ms}, and it is not important from which direction the light approaches from the master oscillator. Specifically, Eq. (1) implies a small-signal gain of $K=\exp(kl)$, where k is the gain, which is proportional to the quantity

$$k \propto B\left(N_r - \frac{g_r}{g_{ms}}N_{ms}\right), \tag{2}$$

and l is the length of one part of the tube.

For a strong signal the increase in the emitted energy, ΔW, with allowance for the self-limiting nature of the lasing, is

$$\Delta W = h\nu V\left(N_r^m - \frac{g_r}{g_{ms}}N_{ms}^m\right)\frac{g_{ms}}{(g_{ms}+g_r)}, \tag{3}$$

where $h\nu$ is the radiated photon energy, V is the volume of the active medium (for simplicity in the analysis we assume that the population inversion is uniform over the volume); N_r^m and N_{ms}^m are the populations of the levels at the time of maximum inversion.

If $K \gg 1$, then under lasing conditions a result analogous to Eq. (3) is obtained if we integrate Eq. (1) over time and over the coordinate x for any boundary and initial conditions identical to the amplification conditions, where here too, it is not important in which zone of the laser, the first or second, the population N_{ms} is zero.

Hence when parts *1* and *2* of the laser are identical and are pumped identically, we can measure the absolute effect of the population of the metastable states on the lasing parameters by shifting the pump pulses relative to one another.

This effect is reflected in the lasing energy of the delayed pulse. We can clarify this in the following way. If the parts of the laser are pumped simultaneously with regular pulses with a repetition rate F (a period T), then each light pulse is formed under the same initial conditions of N_{msT} and N_{eT} and has an emission energy W_0 ($W_0/2$ from each part of the laser). We begin by delaying the pumping in the first zone of the laser. Then the lasing in it develops, according to Eqs. (1)–(3), under the combined initial conditions of N_{ms} in the first and second zones of the tube. The concentration of metastable states in the first zone of the tube is N_{msT}, i. e., the prepulse concentration of metastable states in the regular-pulse mode with a time between pulses of $t=T$. In the second zone of the tube, the concentration of metastable states is $N_{ms\tau}$, i.e., at the time $t=\tau$ after the pulse, where τ is the delay time of the pump pulse in the first zone relative to the pulse in the second zone. If $\tau \ll T$ then $N_{ms\tau} \gg N_{msT}$ (see e. g., Ref. 25), i. e., the development of lasing in the first zone of the tube actually occurs under the initial conditions on N_{ms0} from the second zone. According to Eqs. (1)–(3), the way in which N_{ms} in the second zone of the tube influences the lasing in the first zone is identical to the action due to the given concentration N_{ms} if it occurs in the first zone. At the same time, the initial conditions on N_e for the first zone of the tube remain unaltered.

In this way we separate the effect of N_{ms0} from that of N_{e0} on the parameters of the delayed lasing pulse in real conditions of operation of a pulsed-periodic laser. As pointed out before, the same result was obtained in Refs. 22 and 23. However, in those investigations the independently controlled action of N_{ms0} was produced in spatially coincident absorption and lasing zones and mostly at low repetition rates, i.e., when thermal effects can be neglected. Moreover, the auxiliary discharge, which produces the metastable states can, under the conditions of Refs. 22 and 23, influence the rate of recovery of the lasing, increasing or decreasing it.

The experimental studies were carried out with a copper bromide laser with a quartz tube having a discharge zone 40 cm long and 2.7 cm in diameter (Fig. 1). To protect the exit window from the being covered with particles, we installed the diaphragms *6* and *7* with an inner diameter of 22 mm, which at the same time limited the size of the laser spot. The copper electrodes *8–11* were connected with a molybdenum foil which was soldered into the quartz and acted as an electrical feedthrough. The length of the buffer zone *3* between electrodes *9* and *10* was 10 cm. This part of the tube was held at the operating temperature by means of an external heater. The laser tube had two branch pieces holding the copper bromide. The branch pieces were displaced within the tube by 5 cm from the external electrodes. The pressure of the CuBr vapor was regulated by an external heater, while the working temperature in zones *1* and *2* was maintained by the energy dissipated in the discharge.

The power supply consisted of two oscillators based on TGI2-500/20 thyratrons with a controlled delay between their ignition times. In addition, they could also tuned to operate simultaneously or nearly so, with the simultaneity of operation monitored by observing the current and voltage on an oscilloscope, and also from the maximum lasing power.

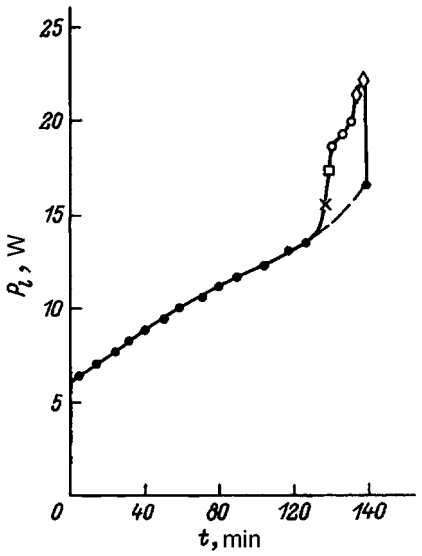

FIG. 2. Evolution of the average lasing power P_l with the introduction of hydrogen.

The pump pulses were formed by the discharge of the capacitors C_1 and C_2 through a thyratron and the transformers T_{p1} and T_{p2}. The secondary windings of these transformers were not connected directly to ground, which simplified the conditions of their operation; in particular, the discharge could be operated under conditions where the internal electrodes acted as anodes or cathodes. The laser was installed so as to satisfy the requirements of minimum inductance of the supply circuit. A typical current pulse width at half maximum was $\sim$50 ns for capacitances $C=C_1=C_2=680$ pF.

RESULTS AND DISCUSSION

After the outgassed tube was loaded with copper bromide in an inert-gas atmosphere, the lasing power with the thyratrons operating simultaneously was 10–15 W. As the copper bromide continued to outgas, the lasing power decreased. As in the case with pure copper vapors, this decrease was due to the steady loss of hydrogen from the active medium, where it accumulates after being released from the insufficiently outgassed components of the laser, primarily the copper bromide. Therefore all the later experiments were carried with mixtures containing hydrogen.

Unlike the work of Refs. 16, 26, and 27, the present experiments did not use a hydrogen-containing mixture prepared ahead of time; rather the hydrogen was admitted from a special generator. By way of example, we show in Fig. 2 the evolution of the lasing power in a carefully outgassed tube for a 17 kHz pulse repetition rate and simultaneous operation of the thyratrons as the hydrogen is let into the tube, which is filled with neon at a pressure $p_{\mathrm{Ne}} \approx 2.9$ kPa. The first stage (0–2 h, 6 min) corresponds to a low rate of hydrogen treatment. With a constant discharge voltage $U=6.25$ kV, the hydrogen treatment results in a steady reduction in the power consumption P from the rectifier from 2.56 to 2.49 kW. In Fig. 2 the mark $\times$ denotes the transition to the new optimum CuBr pressure, which increases with increasing H_2 pressure, while the symbols $\square$ and $\bigcirc$ corre-

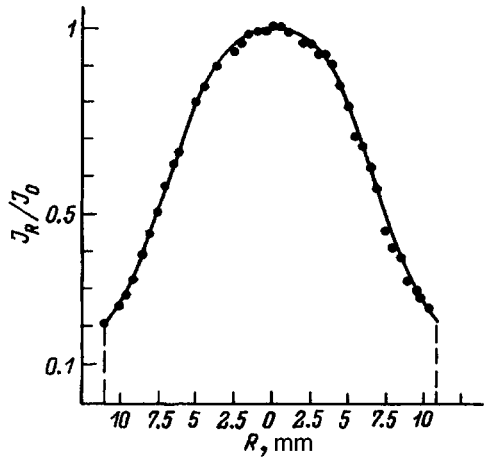

FIG. 3. Radial distribution of the intensity of coherent laser radiation, p_{Ne} = 2.9 kPa, $U=7$ kV

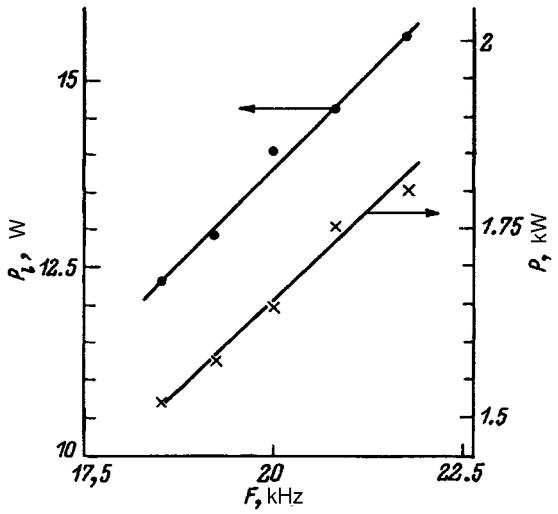

FIG. 4. Rate-power characteristics of the laser (for one operating tube): $p_{Ne}\approx 2.9$ kPa, $U=7$ kV.

spond to higher levels of pumping, which the mixtures with hydrogen can tolerate (respectively, □ refers to $U=6.75$ kV, $P=2.97$ kW, and ○ to $U=7$ kV, $P=3.18$ kW). Then the next two symbols ○ correspond to a stage of more intense hydrogen treatment. The first ◇ corresponds to a new pumping level, $U=7.5$ kV, $P=3.6$ kW, while the second corresponds to a further increase in the hydrogen treatment, which is accompanied by a reduction in the power consumption, $U=7.5$ kV, $P=3.49$ kW. As an example, after the second symbol ◇ the figure shows a transition to the pump voltage corresponding to the initial part of the curve: $U=6.25$ kV and $P=2.44$ kW. The dashed part of the curve corresponds to the stage of increase in the output power for more intense hydrogen treatment with $U=6.25$ kW.

As one can see from Fig. 2, introduction of hydrogen can cause large changes in the lasing power and efficiency. As the H_2 pressure increases, the diameter of the laser spot increases substantially. Under optimum conditions it has a sharp boundary (Fig. 3). Since the region of lasing is larger than the diameter of the diaphragms *6* and *7*, it is clear that this reduces the power and the practical efficiency of the laser. We note that under our conditions, because the hydrogen pressure can be regulated precisely, the effect of that gas on the lasing parameters is greater than on lasing in the pure copper vapor[5,12,13,28] or the vapors of its salts,[26] obtained earlier. Another property of this laser was that placing a peaking capacitor in the second winding of the transformer always lowered the output power of the laser.

In Fig. 4 we show the rate-power characteristics of the laser for $F>17$ kHz, taken for one operating zone of the tube (length $l=40$ cm). One can see from this figure that near $F=17$ kHz, where we performed most of the experiments to measure the effect of N_{ms0} on the lasing parameters, the increase of the output power is proportional to the repetition rate. From this result we can conclude that at these repetition rates and pump power, neither N_{ms0} nor N_{e0} particularly affect the lasing parameters.

To investigate the influence of metastable states on the lasing power, a diaphragm *12* was used to select out a central region 5 mm in diameter. The lasing parameters were measured with a germanium photodiode PD *13* and the power was measured calorimetrically with an IMO-2M calorimeter, *14*. As a rule, the measurements were made at a pump pulse repetition rate of 17 kHz with the optimum composition of the active medium.

In Fig. 5 we show the typical pattern of the recovery of lasing energy in the delayed pulse with varying delay time τ between the pump pulse of zone *1* and that of zone *2*. Curves *1–3* were measured as the operating voltage on the rectifier was varied from $U=6.25$ to 7.5 kV or the energy of the pump pulse was varied from $E=0.145$ to 0.18 J (by the pump energy we mean the quantity $E=P/F$; this is a somewhat higher value than $E=CU_0^2/2$, where U_0 is the initial voltage across the capacitor, since in the latter case the stray capacitance of the supply circuit is not taken into account). Curves *1* and *3* were measured when the electrodes *8* and *11* were operated as cathodes, and curve *2* was measured when they were the anodes. The form of the curves is also essentially independent of which of the zones of the laser is pumped first, the first or the second.

As is shown in Fig. 5, when the prepulse population of the metastable states and the initial concentration of electrons act separately (in other words, when the effect of N_{e0} is neutralized), N_{ms0} loses its influence over the lasing parameters for several microseconds, and lasing recurs in a time

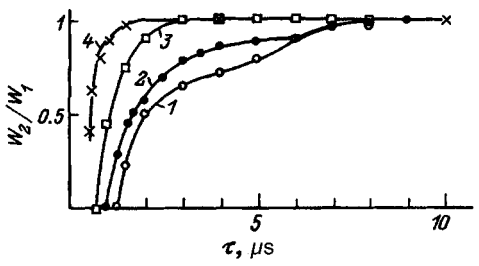

FIG. 5. Recovery of lasing energy in a copper bromide vapor laser (*1*)–(*3*) and lead vapor laser (*4*) for various conditions. $p_{Ne}\approx 2.9$ kPa, $F=17$ kHz; U, kV: *1* — 6.25, *2* — 6.8, *3* — 7.5, *4* — results of Ref. 22.

$\tau \sim 1$ μs and less after the previous pump pulse. A more characteristic and interesting property is that the rate of recovery increases as the pump pulse energy increases.

If we follow the arguments of Ref. 10, then we conclude that 38% of the energy stored in the capacitor goes into ionization, under conditions where most of the ions are singly ionized copper ions. Similar results were also obtained in an analysis of other theoretical and experimental studies. Here the electron concentration in the near afterglow is $N_e = 1.25 \times 10^{14}$ cm^{-3} (curve *1*) and 1.55×10^{14} cm^{-3} (curve *3*). The actual difference may be even greater, since under conditions of the curves *1* and *2* the current pulse was drawn out into a tail on account of the impedance mismatch due to the inadequate conductance of the plasma, which caused a bend in the curve of the recovery of the lasing energy in the interval $\tau = 5-7$ μs. The weaker the pump pulse the more prominent the bend.

If we consider the case of curve *3*, with $N_e = 1.55 \times 10^{14}$ cm^{-3}, then using $k_e = 1.4 \times 10^{-8}$ cm^3/s the characteristic quenching rate is $k_e N_e = 2.2 \times 10^6$ s^{-1}, which corresponds to a lifetime of ~ 0.46 μs. In practice, under these conditions lasing starts in $\tau = 0.7$ μs, or 1.5 decay lifetimes, which corresponds to a decay in the population of the metastable states by a factor of 4.5. Lasing is restored to the 0.9 level with respect to energy in $\tau = 2$ μs, i. e., at a metastable state population $N_{ms0} = 1.3 \times 10^{-2} N_{ms}^{\max}$. As in Ref. 5, the effect of N_{ms} was not observed within $\tau > 10$ μs even when the pump pulse was protracted, that is, the conditions under which curves *1* and *2* were obtained. We recall that the repetition period in these experiments was $T = 58$ μs.

The specific lasing energy in the first pulse, allowing for distribution over the diameter (Fig. 3) is $w_0 = 5.7 \times 10^{-6}$ J/cm^3, which corresponds to $N_p = 1.57 \times 10^{13}$ cm^{-3} photons in a lasing pulse. Consequently, using the condition $k = 0$, we estimate $N_{ms0} \sim (g_{ms}/g_r) N_p = 2.36 \times 10^{13}$ cm^{-3} under threshold conditions for $\tau = 0.7$ μs. Accordingly, $N_{ms0} = 1.06 \times 10^{14}$ cm^{-3} for the maximum value and 1.38×10^{12} cm^{-3} for $\tau = 2$ μs, or 9% of the number of photons in a pulse.

In principle, the situation that arises here has been studied many times, and the results of this investigation can be considered well established. The issue concerns lasers with a phototropic shutter, using absorption on the resonance transitions. This model appears as follows. A phototropic shutter is placed in a copper vapor laser with a potential specific number of photons emitted per pulse, e. g., $N_p = 1.57 \times 10^{13}$ cm^{-3}, and then absorption in the shutter is caused by an equivalent density of nonrelaxing atoms $N_{ms0} = 1.38 \times 10^{12}$ cm^{-3} ($\tau = 2$ μs). The question is, how much does this attenuate the lasing energy? According to Ref. 29, this attenuation amounts to the energy loss ε that goes into bleaching the shutter, i. e.,

$$\varepsilon = \frac{g_r}{(g_r + g_{ms})} N_{ms} h\nu, \tag{4}$$

and corresponds to a time 1.6 μs, for which

$$\varepsilon / w_0 = 0.1, \tag{5}$$

which is close to the observed value. Another parameter by which we can determine the effect of N_{ms} on the recovery of lasing is the ratio of the prepulse absorption coefficient k_0 to the maximum gain $k^{\max}$ obtained during the pump pulse for this repetition rate, when the effect of N_{ms0} is negligible. From Eqs. (2)–(5) it can be shown that

$$\varepsilon / w_0 = k_0 / k^{\max}, \tag{6}$$

and the criterion that the effect of N_{ms0} on the lasing parameters be small is that the condition

$$k_0 \ll k^{\max} \tag{7}$$

be satisfied.

The quantities entering into the right-hand side of Eq. (6) are readily and accurately measurable experimentally. Consequently we can always use this method in real experiments to estimate the actual effect of N_{ms0} on the lasing parameters and separate it from the effect of N_{e0}.

These estimates do not take into account a number of difficulties, in particular, structure in the lasing spectrum, subpopulations of metastable states during the time of the lasing pulse, and in the after-pulse period occurring because of the impedance mismatch, etc. Nevertheless these estimates give a correct picture of the mechanism and of the degree to which N_{ms0} influences the lasing pulse parameters, and they reflect the true order of magnitude of the relaxation rate of metastable states and the recovery of lasing capability by the active medium, which is governed by the prepulse population of metastable states. In support of this picture, Fig. 5 shows a curve (curve *4*) of the recovery of lasing in a lead vapor laser with electron-beam pumping, where the population in the afterglow was not determined from the lasing behavior, but was measured independently.[22] The figure shows that the lasing recovers according to relation (4).

However, unlike the investigation with the copper bromide laser, the recovery rate in the lead vapor beam-pumped laser does not depend on the electron concentration in the afterglow (in other words, on the pump energy). This is accounted for in the following way. Even at the minimum specific energy in the experiments of Ref. 22, the specific energy of a beam-pumped Pb laser, $w_0 = 10^{-6}$ J/cm^3, and a physical efficiency of 2%, the concentration of electrons at the termination of the lasing pulse is $N_e \sim 4 \times 10^{13}$ cm^{-3}, which for $k_e = 2.1 \times 10^{-7}$ cm^3/s for the lead state Pb($^1D_2$) (Ref. 8) gives a quenching rate $k_e N_e = 0.8 \times 10^7$ s^{-1}. This is higher than the rate of electron cooling. For this reason, the rate of recovery in Ref. 22 did not depend on N_e, while in the present case of the CuBr laser it does so depend, since in this case the cooling rate is higher than the quenching rate.

In lasers based on pure copper vapor the electron concentration as a rule is higher than in lasers based on the halides.[7,30] If we take this fact into account, the expected rate of lasing recovery in the Cu laser in the near afterglow can be even higher than shown in Fig. 5 for a copper bromide laser. Indirect support for this statement comes from the fact that lasing starts in the second pulse 300 ns after the termination of the first pulse.[4]

It is also interesting to note that the rate constant for quenching the metastable states in Ba and Pb is higher than

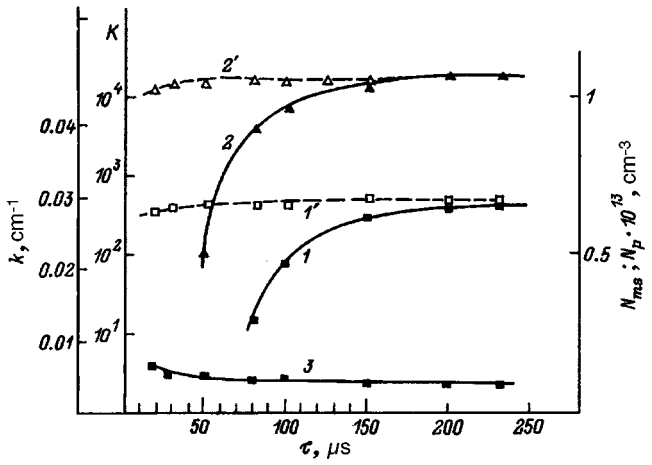

FIG. 6. Gain and the amplification coefficient for the number of photons in a test pulse.

for Cu.[8] However, In the present investigation the conditions for lasing in terms of N_{ms0} in the copper bromide laser are recovered more rapidly than in gas-discharge lasers based on Pb and Ba vapor.[22,23] This result may be due to the following two circumstances: a) the supply circuit is better constructed in our experiments (shorter pulse, no reflections in the current pulse) and b) a higher efficiency in exciting the resonance states of the copper atom as compared to the resonance states of Ba or Pb.

It is further of interest to compare the data we have obtained with the results of Ref. 14, shown in Fig. 6, where the curves *1* and *2* show experimental measurements of the small-signal gain K on the axis and a distance $R=3.5$ cm from the axis, respectively, as functions of the displacement of the test pulse relative to the regular repetition period T. When the length of the tube ($l=210$ cm) and the energy output for $T=230$ μs ($w_0=4$ μJ/cm^3) are taken into account, these curves, plotted in the appropriate coordinates, show a change in the gain k and the number of photons taken from a unit volume. Curve *3* shows the variation in the concentration of metastable states on the axis of the tube. At a distance $R=3.5$ cm, N_{ms} is much smaller than on the axis,[14] and in this figure it is not shown.

A comparison of curves *1* and *3* shows that $N_{ms\tau} \ll N_{pT}$ over the entire range of measured delays, up to $\tau=20$ μs. Using formulas (3), (4), and (6), one can readily see that if the attenuation of the lasing in the delayed pulse were determined by the effect of the prepulse metastable states, then we would find the curve shown as curve *1′* (for $R=3.5$ cm, the appropriate curve is *2′*), instead of the actual curves *1* and *2*, and the corresponding lasing power more than 1 kW/m instead of the actual $\sim$100 W/m. Curves *1′* and *2′* are similar to the curves for Pb vapor lasers[22] and Ba vapor lasers[23], obtained by direct measurement of the effect of the prepulse concentration of metastable states on the lasing parameters. In the light of this discussion it is easy to understand that even in a laser based on pure copper vapor in a tube 8 cm in diameter, the prepulse population of metastable states for a lasing power of $\sim$100 W/m has little influence on the lasing energy, even on the axis of the tube, at any rate, over the 20 μs after termination of the previous pulse. This results is in accord with the conclusion drawn by Tabata et al.[14] on the basis of another method for measuring the effect of N_{ms0}, and also with the results of the present investigation and much earlier papers.[5] Only the thermal nonuniformity together with the nonuniformity due to the skin effect can produce an attenuation of the lasing energy by a factor of $\sim$1.6 at the center of the tube, with an overall attenuation over the aperture, as discussed above, not more than 30%; i. e., neither the relaxation rate of the metastable states nor the Boltzmann thermal population can greatly influence the energy characteristics of the Cu laser. As long ago as 1979, the data of Ref. 6 together with the criteria (6) and (7) could have been used to conclude unambiguously that N_{ms0} has only a slight influence on the parameters of the pulse- periodic copper vapor laser.

Conclusions

The investigations reported in this article have shown that in real gas-discharge lasers the metastable states of the copper atoms are quenched in collisions with electrons in less than 1 μs. Consequently the energy parameters of the Cu laser can in no way be limited by an inadequate rate of quenching the metastable states. Therefore, under conditions where it is weakened or neutralized, the effect of the prepulse electrons, for example in electron-beam pumped lasers or in spatially separated active media, lasing in copper vapor lasers, as well as in lasers based on Ba, Mn, or Pb vapor, is restored to its initial level in a few microseconds after termination of the previous pump pulse. This situation opens up clear prospects for further substantial improvements in the power characteristics of copper vapor lasers, up to the values predicted in Ref. 18, i.e., up to 10 kW/m and above.

[1] W. T. Walter, M. Piltch, N. Solimene *et al.*, IEEE J. Quantum Electron. **QE-2**, 474 (1966).
[2] A. A. Isaev, M. A. Kazaryan, and G. G. Petrash, JETP Lett. **16**, 27 (1972).
[3] P. A. Bokhan and V. I. Solomonov, Kvantovaya Elektron. (Moscow) No. 6(18), 53 (1973) [Sov. J. Quantum Electron. **3**, 481 (1974)].
[4] P. A. Bokhan, V. A. Gerasimov, V. I. Solomonov *et al.*, Kvantovaya Elektron. (Moscow) **5**, 2162 (1978) [Sov. J. Quantum Electron. **8**, 1220 (1978)].
[5] P. A. Bokhan, V. I. Silant'ev, and V. I. Solomonov, Kvantovaya Elektron. (Moscow) **7**, 1264 (1980) [Sov. J. Quantum Electron. **10**, 724 (1980)].
[6] J. L. Miller and T. Kan, J. Appl. Phys. **50**, 3849 (1979).
[7] V. M. Batenin, V. A. Burmakin, P. A. Vokhmin *et al.*, Kvantovaya Elektron. 4, 1572 (1977) [Sov. J. Quantum Electron. **7**, 891 (1977)].
[8] P. A. Bokhan, Kvantovaya Elektron. (Moscow) **11**, 2081 (1984) [*sic*].
[9] R. J. Carman, D. J. W. Brown, and J. A. Piper, IEEE J. Quantum Electron. **QE-30**, 1876 (1994).
[10] B. L. Borovich and N. I. Yurchenko, Kvantovaya Elektron. (Moscow) **11**, 2081 (1984) [Sov. J. Quantum Electron. **14**, 1391 (1984)].
[11] B. L. Borovich, É I. Molodykh, L. A. Ryazanaskaya *et al.*, Kvantovaya Elektron. (Moscow) **17**, 1265 (1990) [Sov. J. Quantum Electron. **20**, 1173 (1990)].
[12] P.A. Bokhan, A. N. Maltsev, and V. I. Silantiev, *Abstracts of the Ninth Conference on Quantum Electronics of Nonlinear Optics*, Poznan, 1980, Sect. A, p. 333.
[13] P. A. Bokhan, A. N Mal'tsev, and V. I. Silant'ev, *Abstracts of the Tenth All-Union Conference on Coherent and Nonlinear Optics* [in Russian] Kiev, 1980, p. 201.
[14] Y. Tabata, K. Hara, and S. Ueguri, Proc. SPIE **1628**, 32 (1992).
[15] J. Chang, Appl. Opt. **32**, 5230 (1993).

[16] D. R. Jones, A. Maitland, and C. E. Little, IEEE J. Quantum Electron. **QE-30**, 2385 (1994).
[17] J. J. Chang, C. D. Boley, M. W. Martinez *et al.*, Proc. SPIE **2118**, 2 (1994).
[18] P. A. Bokhan and E. I. Molodykh, ''Pulsed metal vapour lasers'' in *Proceedings of NATO Advanced Research Workshop* (to be published).
[19] A. A. Isaev, V. V. Kazakov, M. A. Lesnoĭ *et al.*, Kvantovaya Elektron. (Moscow) **13**, 2302 (1986) [Sov. J. Quantum Electron. **16**, 1517 (1986)].
[20] A. A. Isaev, V. T. Mikhkel'soo, G. G. Petrash *et al.*, Kvantovaya Elektron. (Moscow) **15**, 2510 (1988) [Sov. J. Quantum Electron. **18**, 1577 (1986)].
[21] S. Sakato, K. Oohori, M. Higuchi *et al.*, IEEE J. Quantum Electron. **QE-30**, 2166 (1994).
[22] P. A. Bokhan, Kvantovaya Elektron. (Moscow) **12**, 945 (1985) [Sov. J. Quantum Electron. **15**, 622 (1985)].
[23] P. A. Bokhan, Kvantovaya Elektron. (Moscow) **13**, 1595 (1986) [Sov. J. Quantum Electron. **16**, 1041 (1986)].
[24] H. M. Pask and J. A. Piper, IEEE J. Quantum Electron. **QE-30**, 2376 (1994).
[25] D. N Astadzhov, N. K. Vuchkov, A. A. Isaev *et al.*, Kvantovaya Elektron. (Moscow) **14**, 396 (1987) [Sov. J. Quantum Electron. **17**, 245 (1987)].
[26] D. Astadjov, N. Sabotinov, and N. Vuchkov, Opt. Commun. **56**, 279 (1985).
[27] V. F. Elaev, G. D. Lyakh, and V. P. Pelenkov, Optika Atmosphery **2**, 1228 (1989).
[28] Z. C. Huang, K. Namba, and F. Shimizu, Jpn. J. Appl. Phys. **25**, 1677 (1986).
[29] B. R. Belostotskiĭ, Yu. V. Lyubavskiĭ, and V. M. Ovchinnikov, *The Fundamentals of Laser Technology* [in Russian] (Sov. Radio, 1972).
[30] A. A. Zayakin and I. I. Klimovskiĭ, Kvantovaya Elektron. (Moscow) **10**, 1092 (1983) [Sov. J. Quantum Electron. **13**, 698 (1983)].

Translated by J. R. Anderson

Nonstationary two-flux model of radiation transport for optical tomography of scattering media

S. V. Selishchev and S. A. Tereshchenko

Moscow Institute of Electronic Technology, Department of Theoretical and Experimental Physics, 103498 Moscow, Russia

(Submitted December 18, 1995)
Zh. Tekh. Fiz. **67**, 61–65 (May 1997)

A nonstationary two-flux model is formulated for the transport of radiation in an inhomogeneous scattering medium and is applied to the situation where such a medium is irradiated by the narrow beam of a pulsed laser. It is shown that when the time distribution of the transmitted photons is measured it is possible simultaneously to reconstruct the two spatial functions (the coefficients of absorption coefficient and of scattering of the radiation by the medium) by means of an inverse Radon transformation and the solution of a system of nonlinear differential equations on the projection lines. An analytic solution is obtained in quadratures for these differential equations. The results constitute a method of solving problems in optical tomography in an inhomogeneous scattering medium
[S1063-7842(97)01105-7]

INTRODUCTION

A large number of papers have recently appeared dealing with the possibility of tomographic approaches to reconstructing the internal structures in strongly scattering (turbid) media.[1,2] However, examples of reducing these problems to the well-known Radon transformation have been given only for weakly scattering (dilute) media.[3] As in the case of *x*-ray tomography, the idea is to irradiate with a continuous optical beam. Since not much progress has been made in the steady-state case, it is natural to turn to nonstationary models of the interaction of the radiation with the material. Moreover, the results of a large number of experiments[1,2] on the transmission of ultrashort optical pulses through a turbid medium have suggested that the use of pulsed radiation may be one of the most promising directions of research in this area.

The diffusion approximation to the transport equation is generally used to describe theoretically the transmission of optical radiation through a turbid medium, both in the steady-state case and in nonstationary cases.[4,5] We believe that the diffusion approximation does not make a good fit to the tomographic idea of passing narrow beams through the object, since a great deal of information is lost regarding the direction of the probe beam. An analysis of the propagation of a short optical pulse in a scattering medium, based on the time-dependent transport equation in the approximation of single and double scattering[6] is inappropriate for strongly scattering media.

The Kubelka and Munk approximation,[4] which is based on a model of two light beams, propagating, respectively, in the forward and reverse directions, is frequently used in the case of continuous optical radiation in a turbid medium. It would appear that this approximation, like the diffusion approximation, could be generalized to transient processes. Nonetheless, the two-flux (four-flux) approximation has so far been used only for analyzing steady-state processes.

The most general means of describing the interaction of optical radiation with matter (after Maxwell's equations) is the radiation transport equation.[4] The most important characteristics of the scattering medium that enter into the transport equation are the radiation absorption coefficient $\mu_a(\mathbf{r})$ and the differential angular scattering coefficient $\mu_s(\mathbf{r},\mathbf{\Omega}'\rightarrow\mathbf{\Omega}')$ where $\mathbf{r}$ denotes the point of the medium, and $\mathbf{\Omega}'$ and $\mathbf{\Omega}$ are the directions of the photon before and after scattering, respectively. Thus in the most general formulation of the problem, we can consider the reconstruction of two independent functions, which depend on three $(\mu_a(\mathbf{r}))$ and seven $(\mu_s(\mathbf{r},\ \mathbf{\Omega}'\rightarrow\mathbf{\Omega}))$ variables. Since the problem in this formulation is still too difficult to solve, investigators frequently have recourse to a simpler model and the corresponding simpler functions to reconstruct. We emphasize that the function that is to be reconstructed is determined by a theoretical model chosen to describe the interaction of radiation with the medium.

In accordance with the above discussion, the purpose of the work reported here is to analyze the nonstationary two-flux model of radiation transport in an inhomogeneous strongly scattering medium and to demonstrate the possibility of tomographic reconstruction of the characteristics of the medium defined by this model.

NONSTATIONARY TWO-FLUX MODEL OF RADIATION TRANSPORT IN AN INHOMOGENEOUS STRONGLY SCATTERING MEDIUM

Because of the complexity of the analytic solution of the complete transport equation, some approximation is generally used. We shall use the concept of two photon fluxes propagating in opposite directions. In the steady-state case, the model we shall use will be the well-known two-flux model of Kubelka and Munk.[4]

Let us consider the generalization of the two-flux approximation to the time-dependent process for describing the propagation of a point unidirectional pulse in a semi-infinite inhomogeneous scattering medium. An important point is that our approach differs from the model of Kubelka and Munk not only in that it takes into account the time depen-

 1063-7842/97/050511-05$10.00

dence in an inhomogeneous medium but it is also applied to a narrow (collimated) beam, whereas a plane wave source is ordinarily used for reducing the problem to a single dimension. This circumstance introduces definite changes in the physical significance of the characteristics that are to be reconstructed for the inhomogeneous strongly scattering medium.

As is usual in computational topography,[7] we introduce a fixed coordinate system (x,y,z) and a coordinate system (ξ,ζ,z) rotating about the z axis. The transport equation for the radiation in the one-velocity approximation in the rotating coordinate system can be described in the following way

$$\frac{1}{\nu}\frac{\partial}{\partial t}\Phi(\mathbf{r},\mathbf{\Omega},t)+\mathbf{\Omega}\,\mathrm{grad}\Phi(\mathbf{r},\mathbf{\Omega},t)+\mu(\mathbf{r})\Phi(\mathbf{r},\mathbf{\Omega},t)$$
$$-\int\Phi(\mathbf{r},\mathbf{\Omega},t)\mu_s(\mathbf{r},\mathbf{\Omega}'\to\mathbf{\Omega})d\mathbf{\Omega}'=S(\mathbf{r},\mathbf{\Omega},t), \qquad (1)$$

where ν is the velocity of light in the medium, $\Phi(\mathbf{r},\mathbf{\Omega},t)$ is the photon flux density at the point $\mathbf{r}=(\xi,\zeta,z)$ at time t, moving in the direction $\mathbf{\Omega}$; $\mu_s(\mathbf{r},\mathbf{\Omega}'\to\mathbf{\Omega})$ is the differential angular scattering coefficient of the radiation, $\mu_a(\mathbf{r})$ is the absorption coefficient of the radiation, $\mu(\mathbf{r})=\mu_a(\mathbf{r})+\int\mu_s(\mathbf{r},\mathbf{\Omega}'\to\mathbf{\Omega})d\mathbf{\Omega}'$; and $S(\mathbf{r},\mathbf{\Omega},t)$ is the radiation source distribution function.

We shall consider only those photons that travel along the axis of the initial beam. Since it is highly unlikely that the photons scattered away from the axis will return to travel again along the axis, we can assume that the scattering of photons reduces to backscattering along the axis. Consequently the (back) scattering coefficient can be written as $\mu_s(\mathbf{r},\mathbf{\Omega}'\to\mathbf{\Omega})=m_s(\mathbf{r})\,\delta(\mathbf{\Omega}'+\mathbf{\Omega})$, where $\delta(\cdot)$ is the Dirac delta function, $m_s(\mathbf{r})$ is a function of the coordinates, which we shall call the one-dimensional scattering coefficient, having in mind the one-dimensional nature of the two-flux model, and also the fact that this coefficient is a function of only one variable on the axis of the laser beam. Since the photons scattered in the direction of the x axis must be considered as absorbed energy in this model in order to preserve energy balance, the absorption coefficient $\mu_a(\mathbf{r})$ must be replaced by a function $m_a(\mathbf{r})$, which by analogy will be be called the one-dimensional absorption coefficient. We shall denote $m(\mathbf{r})=m_a(\mathbf{r})+m_s(\mathbf{r})$ and rewrite Eq. (1) as

$$\frac{1}{\nu}\frac{\partial}{\partial t}\Phi(\mathbf{r},\mathbf{\Omega},t)+\mathbf{\Omega}\,\mathrm{grad}\Phi(\mathbf{r},\mathbf{\Omega},t)+m(\mathbf{r})\Phi(\mathbf{r},\mathbf{\Omega},t)$$
$$-\mu_s(\mathbf{r})\Phi(\mathbf{r},-\mathbf{\Omega},t)=S(\mathbf{r},\mathbf{\Omega},t). \qquad (2)$$

For a short pulse of laser radiation emitted at time $t=0$ from the point $\mathbf{r}_0=(\xi,\zeta_0,z_0)$, we can write $S(\mathbf{r},\mathbf{\Omega},t)=U(\mathbf{\Omega})\,\delta(\mathbf{r}-\mathbf{r}_0)\,\delta(t)$. We shall assume that the laser pulse is directed along the ζ axis. We denote by $\mathbf{\Omega}_0$ the direction along the beam axis. Since we will consider only those photons that travel along the $\mathbf{\Omega}_0$ axis, while the photons scattered out of this direction can be considered as absorbed, we can write $\Phi(\mathbf{r},\mathbf{\Omega},t)=\Phi_0(\zeta,\mathbf{\Omega},t)\,\delta(\xi-\xi_0)\,\delta(z-z_0)$. We consider two directions, $\mathbf{\Omega}=\mathbf{\Omega}_0$ and $\mathbf{\Omega}=-\mathbf{\Omega}_0$. Then we obtain

$$\frac{1}{\nu}\frac{\partial}{\partial t}\Phi_0(\zeta,\mathbf{\Omega}_0,t)+\frac{\partial}{\partial\zeta}\Phi_0(\zeta,\mathbf{\Omega}_0,t)+m(\mathbf{r})\Phi_0(\zeta,\mathbf{\Omega}_0,t)$$
$$-\mu_s(\mathbf{r})\Phi_0(\zeta,-\mathbf{\Omega}_0,t)=U(\mathbf{\Omega}_0)\,\delta(\zeta-\zeta_0)\,\delta(t),$$
$$\frac{1}{\nu}\frac{\partial}{\partial t}\Phi_0(\zeta,-\mathbf{\Omega}_0,t)-\frac{\partial}{\partial\zeta}\Phi_0(\zeta,-\mathbf{\Omega}_0,t)$$
$$+m(\mathbf{r})\Phi_0(\zeta,-\mathbf{\Omega}_0,t)-\mu_s(\mathbf{r})\Phi_0(\zeta,\mathbf{\Omega}_0,t)$$
$$=U(-\mathbf{\Omega}_0)\,\delta(\zeta-\zeta_0)\,\delta(t). \qquad (3)$$

Using the notation, $F_+(\zeta,t)=\Phi_0(\zeta,\mathbf{\Omega}_0,t)$, $F_-(\zeta,t)=\Phi_0(\zeta,-\mathbf{\Omega}_0,t)$, $U(\mathbf{\Omega}_0)=U_0$, assuming that $U(-\Omega)=0$, and taking the source into account in the initial and boundary conditions, we obtain the main system of equations for the time-dependent two-flux model of radiation transport in an inhomogeneous strongly scattering medium

$$\frac{1}{\nu}\frac{\partial}{\partial t}F_+(\zeta,t)+\frac{\partial}{\partial\zeta}F_+(\zeta,t)+m(\mathbf{r})F_+(\zeta,t)-\mu_s(\mathbf{r})F_-(\zeta,t)$$
$$=0,$$
$$\frac{1}{\nu}\frac{\partial}{\partial t}F_-(\zeta,t)-\frac{\partial}{\partial\zeta}F_-(\zeta,t)+m(\mathbf{r})F_-(\zeta,t)-m_s(\mathbf{r})F_+(\zeta,t)$$
$$=0, \qquad (4)$$
$$F_+(\zeta_0,t)=U_0\delta(t),\quad F_-(\zeta\to+\infty,t)=0,$$
$$F_+(\zeta,0)=0,\quad F_-(\zeta,0)=0. \qquad (5)$$

Expression (5) determines the boundary and initial conditions.

Let us examine the solution of the system of equations (4) for a homogeneous medium, $m(\mathbf{r})=m=\mathrm{const}$ and $m_s(\mathbf{r})=m_s=\mathrm{const}$. Eliminating $F_-(\zeta,t)$ from Eqs. (4), we obtain the equation

$$\frac{1}{\nu^2}\frac{\partial^2}{\partial t^2}F_+-\frac{\partial^2}{\partial\zeta^2}F_++\frac{2m}{\nu}\frac{\partial}{\partial t}F_++(m^2-m_s^2)F_+=0, \qquad (6)$$

whose solution is

$$F_+(\zeta,t)=U_0 m\nu\,\delta(m\nu t-m\zeta')-\exp(-m\zeta')$$
$$+U_0\,\eta(m\nu t-m\zeta')\frac{m_s\nu\zeta'}{\sqrt{(\nu t)^2-\zeta'^2}}$$
$$\times I_1(m_s\sqrt{(\nu t)^2-\zeta'^2})\exp(-m\nu t), \qquad (7)$$

where $\eta(\cdot)$ is the Heaviside step function, $I_1(\cdot)$ are the modified Bessel functions of the first kind and first order, and $\zeta'=\zeta-\zeta_0$.

Expression (7) can be regarded as an approximate description of the radiation transmitted through a layer of thickness ζ', where the first term describes the unscattered (ballistic) component of the radiation transmitted though the layer of thickness ζ'. As one might expect, this component obeys the well-known exponential law of attenuation. The second term describes the scattered component of the radiation. The typical form of the time distribution Eq. (7) is shown in Fig. 1. The vertical line *1* corresponds to the un-

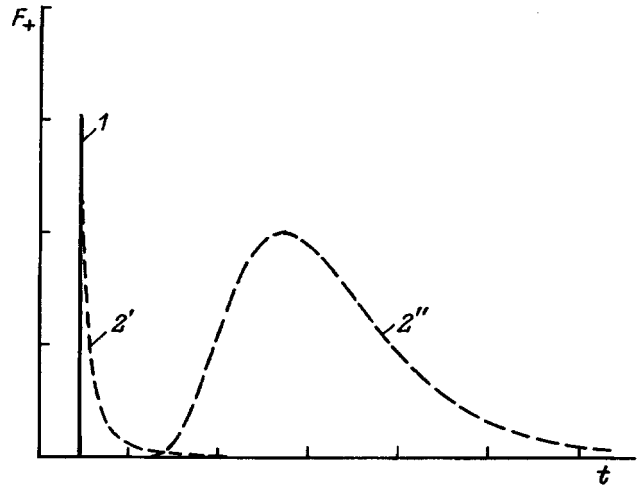

FIG. 1. Typical time distribution of a short optical pulse after passing through a scattering layer. *1* — Ballistic component; *2* — scattered component.

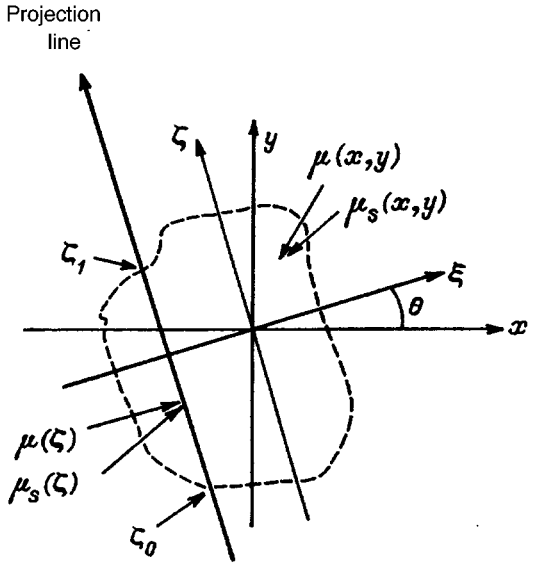

FIG. 2. Geometric arrangement of measurements with parallel projections in tomography.

scattered component. The scattered component has the shape of line *2′* for thin layers and *2″* for thicker layers.

The relative amounts of the unscattered ($U_+^{(ns)}$) and scattered ($U_+^{(s)}$) components can be found by integrating separately each term in Eq. (7) over the time

$$U_+^{(ns)}(\zeta)=\int_0^\infty U_0 m\nu\delta(m\nu t-m\zeta')\exp(-m\zeta')dt=U_0\exp(-m\zeta')=U_0\exp[-m(\zeta-\zeta_0)], \tag{8}$$

$$\begin{aligned}U_+^{(s)}(\zeta)&=\int_{\zeta'/\nu}^\infty \frac{U_0 m_s\nu\zeta'}{\sqrt{(\nu t)^2-\zeta'^2}}I_1\\&\times(m_s\sqrt{(\nu t)^2-\zeta'^2})\exp(-m\nu t)dt\\&=U_0[\exp(-\zeta'\sqrt{m^2-m_s^2})-\exp(-\zeta' m)]\\&=U_0[\exp(-(\zeta-\zeta_0)\sqrt{m^2-m_s^2})-\exp(-(\zeta-\zeta_0)m).\end{aligned} \tag{9}$$

OPTICAL TOMOGRAPHY OF A SCATTERING MEDIUM IN THE TWO-FLUX MODEL OF RADIATION TRANSPORT

To go over to optical tomography of a scattering medium it is necessary to analyze the propagation of an optical pulse in an inhomogeneous medium. Let us consider (Fig. 2) the usual geometrical scheme of measuring with parallel projections.[8] We shall examine the absorption and scattering coefficients in a particular plane $z=z_0$ in the stationary coordinate system (x,y,z): $m(x,y,z_0)=m(x,y)$; $m_s(x,y,z_0)=m_s(x,y)$ and in the rotating coordinate system (ξ,ζ,z) for each fixed value of ξ: $m(x(\xi,\zeta),y(\xi,\zeta),z_0)=m(\zeta)$; $m_s(x(\xi,\zeta),y(\xi,\zeta),z_0)=m_s(\zeta)$. Then instead of Eq. (6), we obtain the following equation

$$\frac{1}{\nu^2}\frac{\partial^2}{\partial t^2}F_+-\frac{\partial^2}{\partial\zeta^2}F_++\frac{1}{\nu}\left[2m(\zeta)+\frac{m_s'(\zeta)}{m_s(\zeta)}\right]\frac{\partial}{\partial t}F_+ +\frac{m_s'(\zeta)}{m_s(\zeta)}\frac{\partial}{\partial\zeta}F_++\left[m^2(\zeta)-m_s^2(\zeta)+m(\zeta)\frac{m_s'(\zeta)}{m_s(\zeta)}-m'(\zeta)\right]F_+=0. \tag{10}$$

Integrating Eq. (10) over the time, we find an equation for the total energy of a pulse transmitted through the scattering medium

$$U_+''-\frac{m_s'(\zeta)}{m_s(\zeta)}U_+'-\left[m^2(\zeta)-m_s^2(\zeta)+m(\zeta)\frac{m_s'(\zeta)}{m_s(\zeta)}-m'(\zeta)\right]U_+=0, \tag{11}$$

where

$$U_+(\zeta)=\int_0^\infty F_+(\zeta,t)dt \quad\text{and}\quad U_+(\zeta_0)=U_0,\ U_+(\zeta\to\infty)=0.$$

The equations for the unscattered ($U_+^{(ns)}$) and scattered ($U_+^{(s)}$) components have the form

$$[U_+^{(ns)}]''-[m^2(\zeta)-m'(\zeta)]U_+^{(ns)}=0, \tag{12}$$

$$[U_+^{(s)}]''-\frac{m_s'(\zeta)}{m_s(\zeta)}[U_+^{(s)}]'-\left[m^2(\zeta)-m_s^2(\zeta)+m(\zeta)\frac{m_s'(\zeta)}{m_s(\zeta)}-m'(\zeta)\right]U_+^{(s)}=-m_s^2(\zeta)U_0\exp\left(-\int_{\zeta_0}^{\zeta}m(\chi)d\chi\right),$$

$$U_+^{(ns)}(\zeta_0)=U_0,\quad U_+^{(ns)}(\zeta\to\infty)=0,$$

$$U_+^{(s)}(\zeta_0)=0,\quad U_+^{(s)}(\zeta\to\infty)=0. \tag{13}$$

The solution[9] of Eq. (12), as expected, expresses the law of exponential attenuation

$$U_+^{(ns)} = U_0 \exp\left[- \int_{\zeta_0}^{\zeta} m(\chi) d\chi \right]. \quad (14)$$

It is well known[7] that the relation (14), or equivalently, Eq. (12) in the problem of tomographic reconstruction of the two-dimensional function $m(x,y)$, leads to the inverse Radon transformation. By applying the inverse Radon transformation $\mathscr{R}^{-1}\{\cdot\}$ (Ref. 7) to the projections

$$p_+^{(ns)}(\xi,\theta) = -\ln \frac{U_+^{(ns)}(\zeta_1)}{U_0},$$

where θ is the angle of rotation of the coordinate system (ξ,ζ) about the stationary system (x,y) and $\zeta_1 = \zeta_1(\xi,\theta)$ is the point of intersection of the beam with the boundary of the medium (Fig. 2) we obtain the reconstructed image of the function being sought, $m(x,y)$. However, in the case of a scattering medium it is necessary to reconstruct simultaneously another two-dimensional function $m_s(x,y)$. One approach[10] is to postulate that the absorption coefficient $m(x,y)$ and the scattering coefficient $m_s(x,y)$ are proportional to the density $n(x,y)$ of the medium, which reduces two unknown functions to one. We can propose a more general approach. We shall assume that each measurement records the time distribution of the transmitted photons. Then from that distribution one can extract the part that corresponds to the ballistic (unscattered) photons that is described by Eq. (12). Then, using the inverse Radon transformation, we obtain one of the unknown functions $m(x,y)$

$$m(x,y) = \mathscr{R}^{-1}\{p_+^{(ns)}(\xi,\theta)\}. \quad (15)$$

To find the second unknown function $m_s(x,y)$ we proceed in the following way. Assuming that the solution $U_+(\zeta)$ of Eq. (11) is a nonnegative function, it can be written in exponential form in some function $M(\zeta)$

$$U_+(\zeta) = U_0 \exp\left[- \int_{\zeta_0}^{\zeta} M(\chi) d\chi \right]. \quad (16)$$

Applying the inverse Radon transformation to the projections

$$p_+(\xi,\theta) = -\ln \frac{U_+(\zeta_1)}{U_0}$$

we obtain the function $M(x,y)$

$$M(x,y) = \mathscr{R}^{-1}\{p_+(\xi,\theta)\}. \quad (17)$$

To find the relation between the functions $m(x,y)$, $m_s(x,y)$, and $M(x,y)$, we substitute expression (16) into Eq. (11)

$$m_s'(\zeta) = \frac{m_s^3(\zeta)}{m(\zeta) - M(\zeta)} + \frac{M^2(\zeta) - M'(\zeta) - m^2(\zeta) + m'(\zeta)}{m(\zeta) - M(\zeta)} m_s(\zeta). \quad (18)$$

Since the function $m(x,y)$ is already known, we must find $m_s(x,y)$ for each projection line by solving the nonlinear differential equation (18) for $m(\zeta)$. Equation (18) is the Bernoulli equation, which can be solved in quadratures,[11]

$$m_s(\zeta) = \frac{\exp\left(\int_{\zeta_0}^{\zeta} f_1(\chi) d\chi \right)}{\sqrt{C - 2\int_{\zeta_0}^{\zeta} f_3(\chi) \exp\left(2\int_{\zeta_0}^{\chi} f_1(\chi_1) d\chi_1 \right) d\chi}}, \quad (19)$$

where

$$f_1(\zeta) = \frac{M^2(\zeta) - M'(\zeta) - m^2(\zeta) + m'(\zeta)}{m(\zeta) - M(\zeta)}, \quad (20)$$

$$f_3(\zeta) = \frac{1}{m(\zeta) - M(\zeta)}, \quad (21)$$

$$C = \frac{1}{m^2(\zeta_0) - M^2(\zeta_0)}. \quad (22)$$

To reconstruct completely the function $m_s(x,y)$ it is sufficient to calculate the quantity $m_s(\zeta)$ from formula (19) over the entire length of the projection for any one value of the rotation angle θ. In the particular case where $m_s(x,y) = \beta m(x,y)$ we obtain the equation

$$m'(\zeta) = \frac{(1-\beta^2)}{M(\zeta)} m_s^3(\zeta) - \frac{M^2(\zeta) - M'(\zeta)}{M(\zeta)} m(\zeta), \quad (23)$$

whose solution is $m(\zeta) = (\sqrt{1-\beta^2})^{-1} M(\zeta)$, which agrees with the previously obtained result.[10]

CONCLUSIONS

To simplify the problem of reconstructing two independent functions depending on three variables (the absorption coefficient) and seven variables (the differential scattering coefficient of the radiation) we have formulated a nonstationary two-flux model of radiation transport in an inhomogeneous scattering medium applied to situations where the medium is irradiated with a narrow laser beam. We have introduced the underlying tomographical reconstruction and the model-dependent characteristics of the scattering medium, called nominally the one-dimensional absorption coefficient and scattering coefficient, but which depend nonetheless on three spatial variables. It is shown that when the time distribution of the transmitted beam photons is recorded it is possible simultaneously to reconstruct these two spatial functions with the aid of the inverse Radon transformation and an auxiliary solution of the set of nonlinear differential equations on the projection lines. An analytical solution in quadratures is obtained for these differential equations. In this way, by staying away from the equation of radiation transport, we have proposed a mathematical method for carrying out optical tomography of scattering media by illuminating them with narrow beams of a pulsed laser.

[1] *Proceedings SPIE. Medical Optical Tomography: Functional Imaging and Monitoring*, edited by G. I. Miller *et al.*, Vol. IS11 (1993).
[2] *Proceedings SPIE. Theoretical Study, Mathematical, Experimental Model for Photon Transport in Scattering Media and Tissue*, Vol. **2326** (1994).

[3] G. G. Levin and G. N. Vishnyakov, *Optical Tomography* [in Russian] (Radio i Svyaz', Moscow, 1989).
[4] A. Ishimaru, *Wave Propagation and Scattering in Random Media* (Academic Press, New York, 1978) [Russ. transl., Mir, Moscow, 1981].
[5] V. V. Lyubimov, Opt. Spektrosk. **76**, 814 (1994).
[6] A. M. Belyantsev, L. S. Dolin, and V. A. Savel'ev, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **10**, 489 (1967).
[7] F. Natterer, *The Mathematics of Computerized Tomography* (Teubner, Stuttgart, 1986) [Russian transl., Mir, Moscow, 1990].
[8] G. A. Fedorov and S. A. Tershchenko, *Calculational Emssion Tomography* [in Russian] (Energoatomizdat, Moscow, 1990).
[9] M. I. El'shin, Dokl. Akad. Nauk SSSR **18**, 141 (1938).
[10] S. V. Selishchev and S. A. Tereshchenko, Pis'ma Zh. Tekh. Fiz. **21**(12), 24 (1995) [Tech. Phys. Lett. **21**, 451 (1995)].
[11] E. Kamke, *Gewöhnliche Differentialgleichungen* (Akad. Verlag, Leipzig, 1959) [Russ. transl., 5th ed. Nauka, Moscow, 1976].

Translated by J. R. Anderson

Ultrawide-bandwidth gain in a coaxial dielectric Cherenkov maser

A. S. Shlapakovskiĭ

Scientific-Research Institute of Nuclear Physics at S. M. Kirov Tomsk Polytechnical Institute, 634061 Tomsk, Russia

(Submitted November 9, 1996)
Zh. Tekh. Fiz. **67**, 66–72 (May 1997)

A coaxial waveguide partially filled with a dielectric as the slow-wave structure of a dielectric Cherenkov maser is investigated. The dispersion of the fundamental mode of this waveguide is very weak at phase velocities close to the velocity of light, and for this reason a very wide gain bandwidth is possible under conditions of an interaction with a relativistic electron beam. The dispersion equation for an infinitely thin tubular beam in a coaxial waveguide with a dielectric liner adjoining either the outer or inner conducting surface is derived. The gain bandwidth as a function of the parameters of the electron beam and the slow-wave system are investigated on the basis of numerical solution of the dispersion equation, and a comparison with similar dependences for the conventional configuration of a dielectric Cherenkov maser is made. The structural features of the coaxial configuration which enable novel approaches to the problems of matching the microwave signal at the entrance and exit of the system are discussed. [S1063-7842(97)01504-3]

INTRODUCTION

A dielectric Cherenkov maser (DCM) is distinguished among the diverse ultrahigh-power devices in relativistic microwave electronics by the possibility of achieving very wide gain band. A dielectric waveguide as a smooth slow-wave system has no stop bands, and for high permittivities it can give a quite weak dispersion of electromagnetic waves in a wide frequency range at relativistic phase velocities. For this reason, a DCM can be regarded as a relativistic analog of the largest-bandwidth device in classical vacuum microwave electronics—a spiral traveling wave tube. In Ref. 1 it was shown that the wide gain band of a DCM is obtained only for definite values of the parameters of the electron beam and slow-wave system and the relative bandwidth at the -3 dB level can reach 40–50% with a maximum gain of 40 dB in the 3-cm wavelength range for electron energies and beam currents (300–600 keV, 2–10 kA) characteristic for high-current direct-action accelerators. These figures were obtained by solving numerically the dispersion equation for a DCM in the standard geometry, which was used in most experimental investigations of DCMs based on a high-current beam[2–6]—a thin tubular beam in a circular metal waveguide with a dielectric liner. However, it is also of interest to investigate a different possible configuration of a DCM—coaxial.

A coaxial geometry for a DCM was studied in Ref. 7 where the dispersion equation was analyzed approximately for the case when both conducting surfaces are coaxial and the outer and inner surfaces are loaded with a dielectric. Such a geometry was also recently realized experimentally in Ref. 8. The characteristic features of the coaxial configuration of a DCM, as compared with the conventional configuration, can be divided into three types. One type are the general physical features which were noted in Ref. 7: the possibility of increasing substantially the kinetic power of the electron beam by increasing the maximum current and also a lower electric field strength on the surface of the dielectric for a given microwave power. Another type are structural features: the presence of an internal rod complicates the structure, but in return it simplifies the problem of dumping the beam; in this case, if the dielectric is present only on the inner surface of the coaxial structure, then the beam can be dumped onto the outer surface at any location irrespective of the presence of the dielectric in a given section. Furthermore, novel designs of devices for extracting and introducing microwave energy are possible in the coaxial geometry. Finally, the third type of characteristic features are those associated with the character of the dispersion of the principal electromagnetic mode of a partially-filled coaxial waveguide—''quasi-TEM'' wave.

The quasi-TEM wave, similarly to a TEM wave, in a uniformly filled coaxial structure has no critical frequency, but in contrast to the TEM wave, it does have a field component E_z and it exhibits dispersion.[9] In the high-frequency limit, its phase velocity approaches $c/\sqrt{\varepsilon}$, where ε is the permittivity, and assumes its maximum value $v_{\max} < c$ at a frequency approaching zero. The quantity $v_{\max}$ is determined by the degree of filling of the waveguide. The less space occupied by the dielectric, the closer to c it is. On account of this dispersion, the Cherenkov interaction of the quasi-TEM wave with the electron beam depends on the ratio of the velocity $v_{\max}$ and the beam velocity u. In the case when u is much less than $v_{\max}$, the coaxial and conventional configurations are essentially similar. If $u > v_{\max}$, then synchronism between the beam and the quasi-TEM is impossible, which produces conditions for exciting the highest mode, so that the coaxial geometry of the DCM can be used to generate short-wavelength radiation with high-energy beams.[10] If the beam velocity is close to $v_{\max}$, then, conversely, the condition of synchronism with the quasi-TEM wave can be satisfied in a very wide range of frequencies. This feature is of interest from the standpoint of the possibility of obtaining an ultrawide gain band in a DCM.

This paper examines a coaxial configuration of a DCM

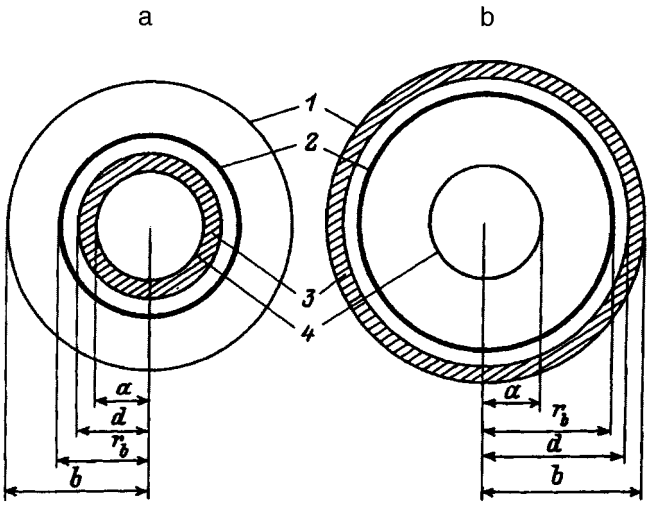

FIG. 1. Transverse section of the region of interaction of a coaxial DCM. a—Liner on the inner conductor, b—liner on the outer conductor; *1*—outer conductor, *2*—beam, *3*—dielectric, *4*—inner conductor.

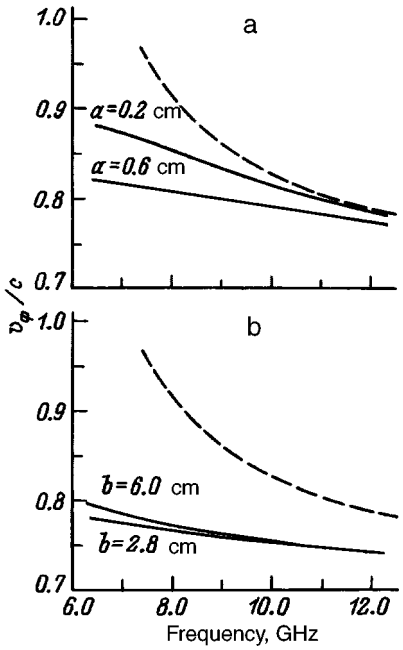

FIG. 2. Frequency dependences of the phase velocity. Dashed curve—E_{01} mode of a waveguide with a dielectric liner, solid curves: a—quasi-TEM mode of a coaxial structure with a liner on the outer surface and an inner rod with different radii, b—quasi-TEM mode of a coaxial structure with a liner on the inner rod and different radii of the outer surface. The liner parameters are fixed: outer radius 2 cm, inner radius 1 cm, $\varepsilon=2$.

with a single dielectric liner adjoining either the inner (a) or the outer (b) conducting surface. The geometry of the system is shown in Fig. 1. In contrast to the case when both surfaces are loaded with a dielectric,[7,8] the E_z component of the field of the quasi-TEM mode does not vanish anywhere inside the beam transport channel, so that the conditions of Cherenkov interaction of a beam and the quasi-TEM mode are most favorable in this system. In the sections below, the dispersion of the quasi-TEM mode is analyzed in comparison with the fundamental mode E_{01} of a circular waveguide with a dielectric liner, the dispersion equation for a coaxial system with an infinitely thin tubular beam, shown in Fig. 1, is derived, and the dependences of the width of the gain band on the parameters of the coaxial DCM are investigated in comparison with the same dependences for a DCM in the conventional geometry by solving the dispersion equation numerically, and the results of these investigations are discussed. The structural features and a possible design of the devices for introducing and extracting the microwave signal in a wide-band coaxial DCM amplifier are also discussed.

1. DISPERSION OF A QUASI-TEM WAVE

The dispersion characteristics of a quasi-TEM wave of a coaxial waveguide with a dielectric have been investigated in detail.[9] Here we present the computational results for the frequency dependences of the phase velocity in a form that is convenient for comparing with the similar characteristic of the working mode of a DCM with the conventional configuration—the E_{01} wave of a circular waveguide with a dielectric liner. The dispersion equation for E-type axisymmetric modes in a coaxial structure with a dielectric on the inner surface (the designations of the geometric dimensions are shown in Fig. 1) has the form

$$\frac{\varepsilon}{p}\frac{F_1(pd,pa)}{F_0(pd,pa)}=\frac{1}{q}\frac{\mathscr{F}_1(qd,qb)}{\mathscr{F}_0(qd,qb)}, \tag{1}$$

where

$$F_0(x,y)\equiv J_0(x)Y_0(y)-J_0(y)Y_0(x),$$
$$F_1(x,y)\equiv J_1(x)Y_0(y)-J_0(y)Y_1(x),$$
$$\mathscr{F}_0(x,y)\equiv I_0(x)K_0(y)-I_0(y)K_0(x),$$
$$\mathscr{F}_1(x,y)\equiv I_1(x)K_1(y)+I_0(y)K_1(x),$$
$$p^2=\varepsilon\frac{\omega^2}{c^2}-k^2,\quad q^2=k^2-\frac{\omega^2}{c^2}, \tag{2}$$

ω is the frequency, k is the longitudinal wave number, and $J_{0,1}$, $Y_{0,1}$, $I_{0,1}$, and $K_{0,1}$ are Bessel functions.

In the case when a dielectric is present on the outer surface of a coaxial line, the dispersion equation is obtained from Eq. (1) by interchanging a and b. If after doing so a is made to approach zero, an equation is obtained for a waveguide with a dielectric liner.

Figure 2 displays the computational results for the phase velocity for the cases when the same dielectric liner loads the inner or outer surface of the coaxial structure and also the surface of a circular waveguide. The liner parameters were chosen to be the same for which in Ref. 1 a wide gain band was demonstrated for a DCM with the conventional geometry in the 3-cm wavelength range. A comparison with the coaxial geometry is especially demonstrative when the dielectric adjoins the outer surface of the coaxial line (Fig. 2a); in this case, the conventional geometry is the particular case $a\to 0$. One can see that for a small radius of the inner conductor the dispersion of the E_{01} and quasi-TEM modes is virtually identical at frequencies above 10 GHz. At frequencies below 10 GHz the dispersion of the quasi-TEM mode is much weaker than that of E_{01}, but it is itself quite substan-

tial. If the radius of the inner rod is relatively large, then the dispersion becomes very weak in a wide frequency range.

The dispersion of the quasi-TEM wave of a coaxial waveguide with a dielectric on the inner surface is even weaker. Comparing Figs. 2a and 2b, it is easy to see that for the same gap width between the surfaces of the conductor and the dielectric the phase velocity of the wave is somewhat less and the dispersion is somewhat greater than in the case of a dielectric on the outer surface. One can see that for the curve $b=2.8$ cm in Fig. 2b the phase velocity changes very little over an octave range. At the same time, the gap width (8 mm) is entirely adequate for beam transport. When the beam interacts with the E_{01} mode, whose dispersion is shown in Fig. 2 by the dashed curve, the width of the gain band, as follows from Ref. 1, can reach 40–50%. Therefore it is natural to conjecture that a DCM amplifier with a coaxial configuration and a working quasi-TEM mode can possess an even wider band. To check this conjecture, it is necessary to obtain and investigate the dispersion equation for the system with a beam.

2. DISPERSION EQUATION OF A SYSTEM WITH A BEAM

We shall study the system displayed in Fig. 1. The tubular, electronic beam with radius r_b and current I_b is assumed to be monoenergetic, completely magnetized, and infinitely thin. In this formulation, in contrast to Ref. 7, the finite distance between the beam and the dielectric surface is taken into account. We assume that the electromagnetic fields and the disturbances of the beam density and velocity are axisymmetric and proportional to $\exp\{i(kz-\omega t)\}$. The equations and the conditions which the longitudinal component $E_z(r)$ of the electric field satisfies on the beam surface[11] follow from Maxwell's equations and the equations of relativistic hydrodynamics in the linear approximation

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dE_z}{dr}\right)+p^2E_z=0 \quad \text{in dielectric},$$

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dE_z}{dr}\right)-q^2E_z=0 \quad \text{in vacuum},$$

$$\{E_z\}_{r=r_b}=0,$$

$$\left\{\frac{dE_z}{dr}\right\}_{r=r_b}=-\frac{q^2}{(\omega-ku)^2}\frac{2eI_b}{\gamma^3 mur_b}E_z(r_b). \tag{3}$$

Here $\gamma=(1-u^2/c^2)^{1/2}$ is the relativistic factor of the beam; e and m are the electron charge and rest mass; p and q are defined in Eq. (2); and, the braces $\{\ldots\}$ denote a jump in the enclosed quantity. The conditions on the beam surface have the form

$$\{E_z\}_{r=d}=0, \quad \frac{\varepsilon}{p^2}\frac{dE_z}{dr}\bigg|_{r=d\pm 0}=-\frac{1}{q^2}\frac{dE_z}{dr}\bigg|_{r=d\mp 0} \tag{4}$$

(the upper sign in the subscripts corresponds to the liner on the outer surface and the lower sign corresponds to the liner on the inner surface of the coaxial structure). Taking account of the condition $E_z=0$ on the conducting surfaces, we shall write the solution for $E_z(r)$ in different regions.

For the dielectric on the inner conductor

$$E_z=\begin{cases} AF_0(pr,pa), & a<r<d,\\ BI_0(qr)+CK_0(qr), & d<r<r_b,\\ D\mathscr{F}_0(qr,qb), & r_b<r<b.\end{cases} \tag{5}$$

If, however, the dielectric adjoins the outer surface, then the following solutions are obtained:

$$E_z=\begin{cases} AF_0(pr,pb), & d<r<b,\\ BI_0(qr)+CK_0(qr), & r_b<r<d,\\ D\mathscr{F}_0(qr,qa), & a<r<r_b.\end{cases} \tag{6}$$

Substituting the boundary conditions (3) and (4) into Eqs. (5) and (6), the corresponding dispersion equations can be derived after simple transformations. The final result for the case of a dielectric on the inner surface has the form

$$\left[\frac{1}{p}\frac{F_1(pd,pa)}{F_0(pd,pa)}-\frac{1}{\varepsilon q}\frac{\mathscr{F}_1(qd,qb)}{\mathscr{F}_0(qd,qb)}\right](\omega-ku)^2$$
$$=\frac{2eI_b}{\gamma^3 mu}\,q^2\mathscr{F}_0(qb,qr_b)\,\frac{\mathscr{F}_0(qd,qr_b)}{\mathscr{F}_0(qd,qb)}$$
$$\times\left[\frac{1}{p}\frac{F_1(pd,pa)}{F_0(pd,pa)}-\frac{1}{\varepsilon q}\frac{\mathscr{F}_1(qd,qr_b)}{\mathscr{F}_0(qd,qr_b)}\right]. \tag{7}$$

The dispersion equation for a coaxial DCM with a dielectric on the outer surface is obtained from Eq. (7) by interchanging a and b and also changing the sign on the right-hand side. Then the dispersion equation for a DCM with the conventional configuration is obtained in the limit $a\to 0$.[11,12] In the absence of a beam ($I_b=0$) the equation (7) passes into Eq. (1).

3. INVESTIGATION OF THE DISPERSION EQUATION

The dispersion equation (7) derived above was investigated numerically. The approximate analytical analysis given in Ref. 7 does not make it possible to determine the width of the gain band and, moreover, it is inapplicable for the cases of practical interest when there is no direct synchronism between the quasi-TEM wave and the beam ($\omega=ku$) or amplification occurs at frequencies which are very far from synchronism. As already noted in Sec. 1, it is convenient to compare the DCM with the conventional and coaxial configurations for the case of a dielectric on the outer surface, varying the radius of the inner conductor. For this reason, we consider this case first.

Figure 3 displays the result of solving Eq. (7) numerically for a dielectric liner with the same parameters as in Fig. 2 and an electron beam with $\gamma=2$ and current $I_b=10.2$ kA passing 2 mm from the surface of the dielectric. One can see that as the radius of the inner conductor increases, the spatial growth rates gradually decrease (for $a<0.4$ cm, in this case, there are virtually no differences between the coaxial and conventional geometry), and the gain band shifts in the direction of low frequencies, narrowing in both absolute and relative values. It is important to note that for a waveguide with a dielectric liner the frequency of Cherenkov synchronism is comparatively close to the corresponding maximum of the increment, while for a coaxial line with $a=0.4$ cm the synchronism frequency is 2.4 times lower than the frequency

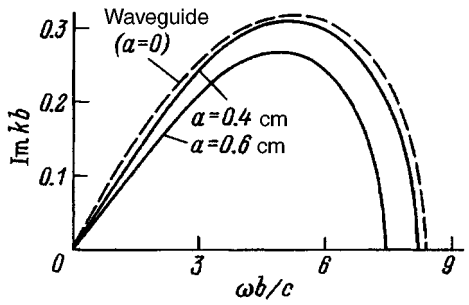

FIG. 3. Spatial growth rate versus frequency for a beam in a waveguide with a dielectric (dashed curve) and in a coaxial line with a dielectric on the outer surface (solid lines) for different radii of the inner conductor. The parameters of the dielectric liner and electron beam are fixed: $\varepsilon=2$, $b=2$ cm, $d=1$ cm, $r_b=0.8$ cm, $I_b=10.2$ kA, $\gamma=2$.

of the maximum increment, and for $a=0.6$ cm there is no synchronism at all, since in this case $u>v_{\max}$.

Given the maximum gain G, the widths of the gain band at the -3 dB level can be determined from the frequency dependences of the growth rate Im $k(\omega)$. Assuming that G is large, it is easy to obtain from the formulas for the gain of a traveling wave tube (TWT)[13]

$$\frac{\Delta(\mathrm{Im}\,k)}{(\mathrm{Im}\,k)_m}\approx\frac{3}{G[dB]+\alpha}. \tag{8}$$

Here $\Delta(\mathrm{Im}\,k)$ is the difference between the maximum spatial growth rate $(\mathrm{Im}\,k)_m$ and the growth rate corresponding to the -3 dB limit of the band; the coefficient α takes on values between 6.0 and 9.5 depending on the beam current (9.5 corresponds to the limit of low space charge and 6.0 corresponds to the opposite limit). Formula (8) makes it possible to perform an even clearer comparison of the conventional and coaxial configurations of a DCM.

Figure 4 displays the dependences of the gain band width on the beam current with a maximum gain of 40 dB. The set of parameters corresponding to the three curves presented are the same here as for Fig. 3. The absolute magnitude of the error, associated with the uncertainty of the coefficient α in Eq. (8), in determining the band equals 1–2% everywhere. It is obvious that the largest gain band width can be obtained here in the conventional configuration. When the

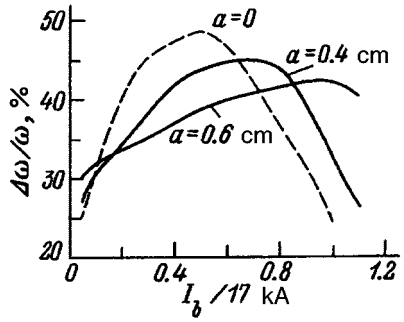

FIG. 4. Gain band width at the -3 dB level versus beam current. The other parameters of the beam and slow-wave systems are the same as in Fig. 3. The maximum gain equals 40 dB. Dashed curve—waveguide, solid curve—coaxial structure with a dielectric on the outer surface.

inner conductor is introduced, the dependence of the bandwidth on the current becomes flatter and the maximum value of this dependence decreases. The gain band of a coaxial DCM is even wider than for the conventional DCM for both high and low beam currents. It is important to keep in mind that a change in the beam current not only results in a change in the band width but it also results in a large change in the center frequency and the growth rates themselves (i.e. it leads to a substantial change in the distance over which a given gain is achieved). For low currents, the growth rates in a coaxial DCM are very small, especially for the case $a=0.6$ cm, when there is no synchronism. The maximum growth rate as a function of the current approximately coincides with the maximum band width.

Figure 4 is the main illustration of the band properties of a coaxial DCM in comparison with a conventional DCM. The weaker dispersion of the working mode in a coaxial configuration plays a role only for low beam currents. In this region of currents (<2 kA for the parameters in Fig. 4) the gain band in the coaxial case is wider than in the conventional case. But the band width itself is not very large here. This is due to the fact that in this case the gain band of a coaxial DCM lies in the low-frequency range, and for low frequencies the spatial growth rates are very small. It is well known that at low frequencies the growth rate of the Cherenkov instability (just as the spectral intensity of the spontaneous Cherenkov radiation) decreases directly as the decrease in the frequency. For this reason, large band widths are not reached despite the very weak dispersion.

As the current increases, the gain band shifts into the region of higher frequencies. In the process, the difference in the character of the dispersion of the quasi-TEM mode and the E_{01} mode of a waveguide with a dielectric liner is no longer as great (Fig. 2a), and since the center frequency for the coaxial case remains appreciably lower than for the conventional case, the maximum gain band is reached in the conventional configuration. Finally, for very high currents (>10 kA for the parameters of Fig. 4) the coaxial configuration has a large advantage with respect to the band width. But this is now due not to the character of the dispersion but rather to the fact that the frequency shift reaches a value for which the exponential dropoff of the field in the beam propagation channel becomes substantial. At high frequencies the beam–wave coupling starts to decrease rapidly because of the finite distance between the beam and the dielectric, and this results in a narrowing of the band. In the coaxial case the field within the propagation channel is redistributed, so that this effect comes into play at frequencies which are all the higher, the smaller the transverse size of the channel is. Therefore at high currents the gain band in a coaxial structure with a large radius of the inner rod is much wider than in the conventional geometry.

We shall now investigate the case when the inner conducting surface of the coaxial structure is loaded with a dielectric. Figure 5 displays the frequency dependences of the growth rate which make it possible to compare this case with the case of a dielectric on the outer surface. For this, geometries with identical dielectric liners, the same gap width between the dielectric and the unloaded conductor, the same

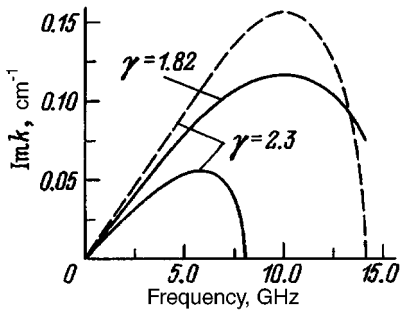

FIG. 5. Spatial growth rate versus frequency. $I_b=5.1$ kA, $\varepsilon=2$. Dashed line—coaxial structure with a liner on the outer surface ($b=2$ cm, $d=1.2$ cm, $a=0.4$ cm, $r_b=1$ cm, $\gamma=2.3$ cm); solid curves—coaxial structure with the same liner on the inner surface and the same beam–dielectric and beam–conductor distances ($b=2.8$ cm, $d=2$ cm, $a=1.2$ cm, $r_b=2.2$ cm, $\gamma=2.3$ and 1.82).

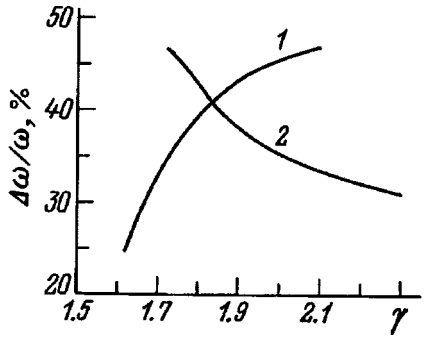

FIG. 6. Gain band width at the -3 dB level (maximum gain 40 dB) of a coaxial DCM with an liner on the inner surface versus electron energy. $\varepsilon=2$, $b=2.8$ cm, $d=2$ cm, $r_b=2.2$ cm. The center frequency is held constant (10 GHz) by varying the current with $a=1.2$ cm (*1*) and by varying a with $I_b=5.1$ kA (*2*).

distance between the beam and the dielectric, and the same beam current and energy were chosen. Comparing the solid and dashed curves with $\gamma=2.3$ in Fig. 5, it can be concluded that although the center frequency and the instability growth rate are much higher for the case of a dielectric on the outer surface, the relative band width for these cases is approximately the same. A calculation according to Eq. (8) gives an even somewhat larger value for the dielectric on the inner surface. Since for the geometries being compared different phase velocities correspond to the same frequency of the quasi-TEM mode, it is of interest to change the energy of the electron beam in the case of a dielectric on the inner surface in a manner so that the maximum growth rate would occur at the same frequency as for the dashed curve in Fig. 5 (10 GHz). This matching is achieved for $\gamma=1.82$. The corresponding curve shows that in this case the gain band in the configuration with a dielectric on the inner surface becomes much wider than for a dielectric on the outer surface, and the growth rate remains somewhat lower.

Of course, the range of amplified frequencies of the DCM can be changed not only by changing the electron energy. A change in the parameters of the dielectric liner as well as a change of the beam current strongly influences the frequency. The same center frequency can be obtained for different beam energies, if these parameters are also changed together with the energy. It is of interest to determine how the gain band changes in this case. Figure 6 displays the curves of the band width versus the electron energy when the center frequency is held constant. The point of intersection corresponds to the curve $\gamma=1.82$ in Fig. 5.

One can see that the two curves have opposite characters. Curve *1* illustrates well the situation in which the band width is small for low currents. The range of variation of the current within the curve 1 is from 170 A (for $\gamma=1.62$) to 25 kA (for $\gamma=2.1$). Increasing the energy simultaneously with the current results in broadening of the band, but for large values of γ very high currents are required to maintain the center frequency at 10 GHz and then the band changes very little. If, however, the center frequency is maintained by changing the thickness of the dielectric holding the current fixed, then as the energy increases, the band, conversely, becomes narrower on account of the fact that the degree of filling of the waveguide decreases.[1] Within the curve *2* the inner radius of the liner changes from 1.0 to 1.5 cm.

If at high energies the curve *2* in Fig. 6 can be continued without restrictions (a will approach d), then on the opposite side it reaches the natural limit of a situation of a waveguide with a dielectric rod of radius d. It is interesting to determine how the gain band width will change if the inner radius of the liner changes simultaneously with the outer radius, while the thickness of the liner remains constant. The result of such a calculation is displayed in Fig. 7. The radius of the beam was also changed so that the distance between the beam and the dielectric remained equal to 2 mm. In this case the frequency corresponding to maximum gain changes very little, and for the parameters of Fig. 7 it equals ≈ 10 GHz. Here the gain band becomes much wider as the liner decreases in size, despite the fact that the degree of filling of the waveguide decreases in this case.

4. STRUCTURAL FEATURES

Therefore the geometry with a dielectric loading the inner surface of the coaxial structure makes it possible to obtain a wider gain band with lower beam energies, though with somewhat lower values of the instability growth rate, than in the case of a liner on the outer surface. At the same time, such a configuration is extremely interesting because of

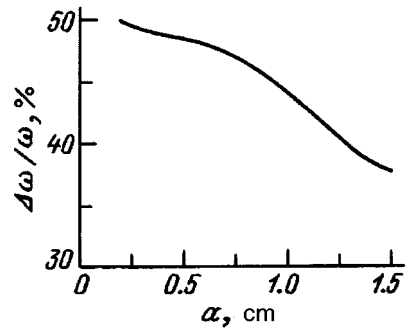

FIG. 7. Band width at the -3 dB level (maximum gain 40 dB) for a coaxial line with an liner on the inner conductor versus the inner radius of the liner with constant liner thickness ($d-a=8$ mm) and beam–dielectric gap width ($r_b-d=2$ mm). $\varepsilon=2$, $b=2.8$ cm, $\gamma=1.82$, $I_b=5.1$ kA.

its structural features. First, problems associated with the formation, transport, and doping of the electron beam are more easily solved in it. For a fixed current and thickness of the beam propagating outside the liner the current density is lower than in the case of propagation on inside liner, and hence it is easier to form and position the beam. But the main convenience is the possibility of dumping the beam in a decreasing magnetic field on the outer, unloaded metal surface. In a DCM, a section where the thickness of the dielectric liner decreases continuously is required in order to match the region of interaction and the output microwave channel. In the conventional configuration, because of the fact that the beam cannot be dumped onto the dielectric, the beam on this section continues to interact with the wave and this interaction can lead to a loss of microwave energy. Furthermore, the collector plasma that is formed also can absorb or reflect microwave energy, a substantial part of which in the conventional geometry is concentrated near the waveguide wall. In a geometry with a dielectric on the inner rod, however, the interaction can be cutt off at any location irrespective of the presence of a dielectric in a given section. Therefore the amplified signal can be transported noise-free farther along the dielectric and the matching section can be as long as desired. The collector plasma in this case also will be formed far from the dielectric.

Second, such a design admits novel possibilities for solving the microwave input–output problems. A metal rod coated with a dielectric is a ready-made transmission line. It can be assumed that in a coaxial DCM signal amplification occurs on some section of this line. Indeed, in this case it is possible to avoid the limits on the length of a such a dielectric waveguide which are due to the presence of the electron beam. For this, the beam generation scheme proposed for a triaxial relativistic klystron[14] can be used on the input side or the beam can be formed in a compact module of a linear induction accelerator. In the case of the output, a rod with a dielectric sheath (or simply a dielectric rod) can serve directly as an antenna.[15] Here it is important to note that an asymmetric mode, which produces sharply directed radiation, is employed in the widely used dielectric antennas. For this reason, it is of great interest to investigate a coaxial DCM with an asymmetric working mode.

CONCLUSIONS

A coaxial configuration of a DCM with one dielectric liner, which loads either the inner or outer conducting surface, was investigated. The gain band widths at the -3 dB level were determined as a function of the system parameters by solving numerically the dispersion relation obtained. It is found that from the standpoint of broadening of the band the coaxial configuration is no better than the conventional configuration, despite the much weaker dispersion of the working ''quasi-TEM'' mode. Nonetheless, as one can see from the curves presented, here the same band widths are achieved as in the conventional configuration—about 50% with a maximum gain of 40 dB for beam currents of 5–10 kA and beam energies of 400–600 keV, so that a coaxial waveguide with a dielectric liner can be used as an electrodynamic system for a wide-band microwave amplifier. The coaxial configuration admits a wider band than the conventional configuration, if the beam currents are very high (Fig. 4). A geometry with a dielectric liner on the inner conductor makes it possible to obtain a wider gain band with lower voltages than in the case of a dielectric on the outer conductor (for fixed center frequency). Furthermore, a coaxial structure with a dielectric on the inner surface is a structurally very attractive system, since it makes it easier to solve the problem of dumping the beam and admits novel possibilities for introducing and extracting the microwave signal.

This work was supported by the International Science Foundation (Grant No. NYD000) and a joint grant from the International Science Foundation and the government of the Russian Federation (No. NYD300).

[1] A. S. Shlapakovskii and K. A. Chirko, IEEE Trans. Plasma. Sci. **PS-22**, 544 (1994).
[2] J. E. Walsh, T. C. Marshall, and S. P. Schlesinger, Phys. Z. **20**, 709 (1977).
[3] A. N. Didenko, A. R. Borisov, G. P. Fomenko, and Yu. G. Shteĭn, Pis'ma Zh. Tekh. Fiz. **9**, 60 (1983) [Sov. Tech. Phys. Lett. **9**, 26 (1983)].
[4] W. Main, R. Cherry, and E. Garate, Appl. Phys. Lett. **55**, 1498 (1989).
[5] W. T. Main, E. Garate, J. Weatherall, and R. Cherry, IEEE Trans. Plasma. Sci. **PS-20**, 281 (1992).
[6] G. P. Fomenko, A. G. Pelyodov, A. S. Shlapakovskii, and Yu. G. Shtein, Nucl. Instrum. Methods Phys. Res. A **331**, 152 (1993).
[7] E. P. Garate, A. Fisher, and W. T. Main, IEEE Trans. Plasma Sci. **PS-18**, 831 (1990).
[8] T. J. Davis, L. Schächter, and J. A. Nation, IEEE Trans. Plasma Sci. **PS-22**, 504 (1994).
[9] Yu. A. Ilarionov, S. B. Raevskiĭ, and V. Ya. Smorgonskiĭ, *Calculations for Rippled and Partially Filled Waveguides* [in Russian], Sov. Radio, Moscow (1980).
[10] G. P. Fomenko and A. S. Slapakovskiĭ, Zh. Tekh. Fiz. **59**(8), 145 (1989) [Sov. Phys. Tech. Phys. **34**, 938 (1989)].
[11] N. I. Karbushev and A. S. Shlapakovskiĭ, Zh. Tekh. Fiz. **59**(3), 161 (1989) [Sov. Phys. Tech. Phys. **34**, 346 (1989)],
[12] D. S. Lemons and L. E. Thode, Phys. Rev. Lett. **56**, 2684 (1986).
[13] V. N. Shevchik and D. I. Trubetskov, *Analytical Computational Methods in Microwave Electronics* [in Russian], Sov. Radio, Moscow (1970).
[14] M. Friedman and V. Serlin, in *Proc. SPIE. The Inter. Soc. for Optical Eng.*, edited by Howard E. Brandt, Vol. 1872, 2 (1993).
[15] C. H. Walter, *Traveling Wave Antennas*, McGraw-Hill, New York, 1965 [Russian translation, Énergiya, Moscow (1970)].

Translated by M. E. Alferieff

Focusing a compensated proton beam with a high compression coefficient

H. Wuerz

Karlsruhe Research Center, 76021 Karlsruhe, Germany

G. A. Vyaz'menova, V. S. Kuznetsov, and V. I. Éngel'ko

D. V. Efremov Scientific Research Institute for Electrophysical Apparatus, 189631 St. Petersburg, Russia
(Submitted November 2, 1995)
Zh. Tekh. Fiz. **67**, 73–80 (May 1997)

An analysis is made of the possibility of focusing 10–30 keV charge-compensated proton beams with a current of 1 kA and an area compression coefficient $>1.5\times10^3$. To obtain a high degree of compression a combination of ballistic focusing and magnetic compression is used. The parameters of the focusing channel are determined and an estimate is made of the possible effect of the compensating electron background on the effectiveness of focusing of the proton beam. © *1997 American Institute of Physics.* [S1063-7842(97)01205-1]

Introduction

To investigate the erosion of the first wall and of the divertor material under conditions of thermal phase separation of the plasma current in the ITER installation a proton beam with a particle energy of 10–30 keV is required, with a pulse length greater than 10 μs and a power density higher than 10 MW/cm^2 over an area of at least 1 cm^2. In this investigation we examine the possibility of obtaining such a beam, based on the use of a vacuum diode with a multitip explosive emission cathode and a desorption anode for generating the beam, and a combination of ballistic focusing plus magnetic compression to attain the required power density at the target. Preliminary calculations and experiments[1] have shown that with a desorption anode as the source of protons the necessary beam parameters can be attained if the anode area is of the order of 10^4 cm^2, the electrode spacing in the diode is at least 3 cm, and the diode voltage is at least 20 kV. The effective pulse length is then about 10 μs, and the angular spread after passage of the multitip cathode is 25–30 mrad. The space charge of the beam in the drift region is neutralized by electrons from the cathode plasma.

To obtain the required power density at the target it is necessary to compress the beam from 10^4 cm^2 at the anode to 1 cm^2 at the target. The distance at which the focusing occurs must be not less than 1 m, since otherwise the longitudinal velocity of the particles at the boundary of the beam will be considerably lower than the total velocity. At this distance and with an angular spread of the particles of 20–30 mrad the radius of the beam at the crossover is 2.5–3 cm. Therefore to have a radius of 0.5 cm at the crossover requires additional compression of the beam. This compression can be obtained with the aid of a longitudinal magnetic field that increases in the direction of propagation of the beam. Injection of a proton beam in a longitudinally increasing magnetic field is required not only for focusing the beam, but also to carry out experiments under typical ITER conditions, where the magnetic induction reaches 5 T.

Magnetic compression of a charge-neutralized ion beam by a magnetic field has been studied in a number of papers. For example, under the assumption of a laminar quasineutral magnetized beam with a small transverse spread, an equation was found for the envelope of the beam.[2] The solution of this equation showed the possibility of adiabatic compression of the beam. However, the length of the magnetic compression channel for our conditions in this case is $\sim10^3$ m. Such dimensions are not feasible for constructing experimental apparatus.

We have assumed that in the present case the optimum method will be a combined method of focusing the beam, involving ballistic focusing in the first section and magnetic compression in the second section. The beam must be injected along the converging lines of force of the magnetic field, whose configuration must be as close as possible to the proton trajectories (Fig. 1). In the first section, which has a length L_1, the protons are not magnetized. In the second section, of length L_2, the protons are magnetized because of the increase in the magnetic field to the level at which the deviation of the particles from the lines of force becomes comparable with their cyclotron radius. The magnetic field in the second section increases adiabatically.

Let us make some estimates. First we consider a conical magnetic field with straight lines of force that converge to the axis. The deviation of an ion trajectory from the magnetic lines of force at the end of the ballistic section due to the initial angular spread $\Delta r_0'$ is

$$\Delta r_1=\frac{\Delta r_0' L_1}{\cos\alpha_1}, \tag{1}$$

where α_1 is the angle of inclination to the axis of the magnetic lines of force that pass through the edge of the emitter.

If we assume that the deviation Δr_1 is equal to the cyclotron radius at the end of this section, where the magnitude of the magnetic field is B_1, then we can determine the geometric parameters of the first sections

$$L_1=\frac{V_0 M\cos\alpha_1}{eB_1}, \tag{2}$$

$$\tan\alpha_1=R_0\left(1-\sqrt{\frac{B_0}{B_1}}\right)\frac{1}{L_1}, \tag{3}$$

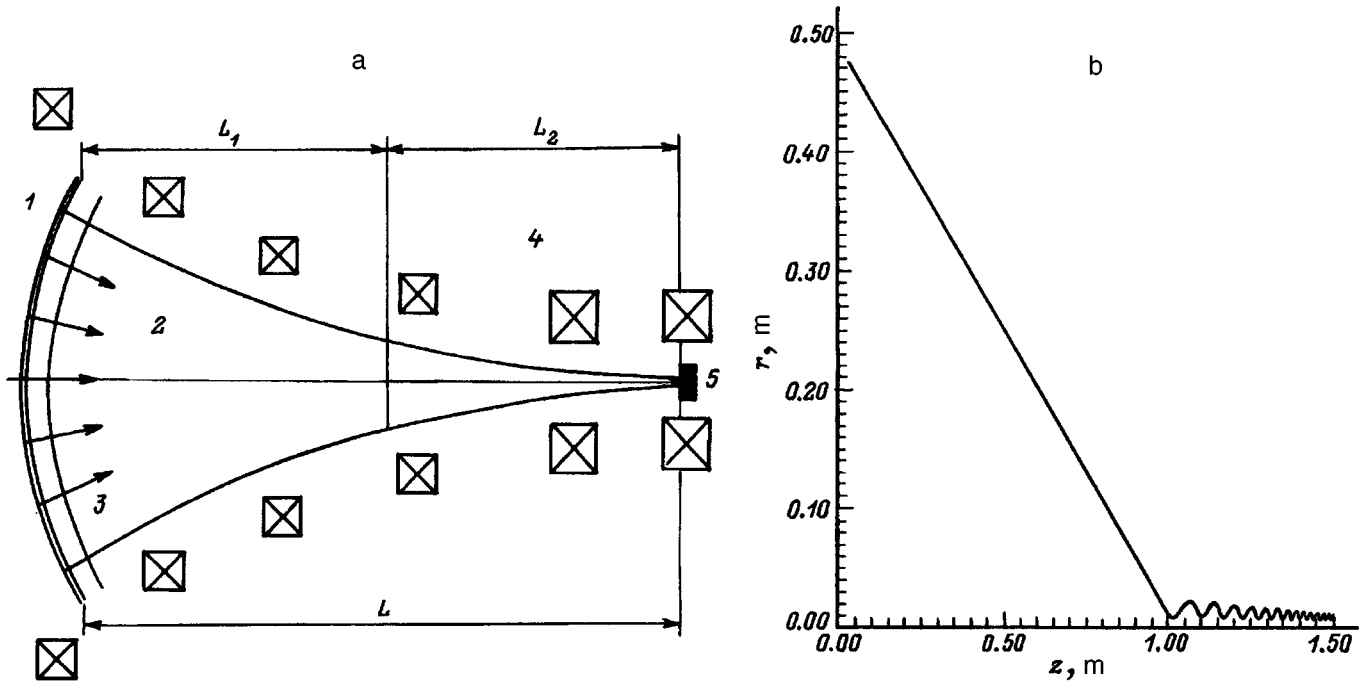

FIG. 1. (a) Diagram illustrating focusing of a proton beam. *1* — anode, *2* — protons, *3* — cathode, *4* — magnets, *5* — target; (b) — trajectory of outermost proton, of energy 15 keV in a double-conical magnetic field for $\Delta r_0'=0$. a — In integral form, b — in differential form.

where V_0, M, and e are the velocity, mass, and charge of the proton, and R_0 is the radial coordinate of the particle at the edge of the emitter.

The values of B_1 and α_1 also depend on the magnetic field B_m at the target through a relation that determines the limiting angle at the proton input to the second section for which the protons will reach the target

$$\sin^2\alpha_1 \leqslant \frac{B_1}{B_m}. \tag{4}$$

This same relation can be used to estimate the permissible initial angular spread

$$\Delta r_0' < \sqrt{\frac{B_1}{B_m}}. \tag{5}$$

The length L_2 of the second (magnetized) section is selected from the condition of adiabaticity

$$L_2 \geqslant \frac{2\pi M V_0 \cos\alpha_1}{eB_s}, \tag{6}$$

where $B_s=\sqrt{B_1 B_m}$ is the average magnetic field in the second section.

The initial ion energy corresponding to the velocity V_0 is selected to satisfy the condition that the beam radius r_m at the target be about two proton cyclotron radii at the target

$$V_0 \cong \frac{er_m\sqrt{B_m B_1}}{2\Delta r_0' M}. \tag{7}$$

The beam radius at the target is equal to the radius of the magnetic force lines plus the cyclotron radius, i.e.,

$$r_m = R_0\sqrt{\frac{B_0}{B_m}} + \frac{V_0 M}{eB_m}. \tag{8}$$

Now we can determine the average power density at the target as

$$P_s = \frac{MV_0^2}{2e} j = \frac{NV_0^2 I_0}{2\pi e r_m^2}, \tag{9}$$

where j is the average current density at the target and I_0 is the beam current.

Relations (2), (3), and (5)–(7) allow us to determine the parameters of the conical magnetic field focusing system. For $\Delta r_0'=0.03$, $B_m=5$ T, and $r_m=5$ mm, we have $V_0=1.2\times10^6$ m/s (this corresponds to a proton energy $E\sim10$ keV), $\alpha\geqslant7°$, $B_1\geqslant4.5\times10^{-3}$ T, $L_1=3$ m, and $L_2\geqslant0.5$ m. The total length of the focusing channel is $L=r_0/\tan\alpha\leqslant4$ m. These estimate show that in a conical magnetic field it is not possible to achieve adiabatic compression of a beam in the second section (since $L<L_1+L_2$).

We can improve the situation by using a more complicated magnetic field configuration. Let us consider, for example, a double-conical magnetic field. In the section L_1 the angle between the magnetic lines of force and the beam axis is α_1 and in section L_2 it is α_2. Let L_1 be equal to the ballistic focusing length, i.e., the distance between the diode and the beam crossover in the absence of the magnetic field. The motion of the protons in the first section is governed by the ballistics. Therefore the entrance angle of the outermost particles of the beam in the second section will be $\Delta\alpha=\alpha_1-\alpha_2$. We shall use for an estimate $\alpha_1\cong\Delta\alpha$. In or-

der for the proton to reach the target, which is situated in a magnetic field of 5 T, the following relation must hold

$$B_1 > B_m \sin^2 \alpha_1 . \tag{10}$$

We shall assume that the magnetic field in the second section increases adiabatically.

The transverse energy at the target is

$$W_m^{\perp} = W_0 \frac{B_m}{B_1} \sin^2 \alpha_1 , \tag{11}$$

where W_0 is the total proton energy, and so we determine B_1 as

$$B_1 = \frac{W_0}{W_m^{\perp}} B_m \sin^2 \alpha_1 . \tag{12}$$

Let us now consider the magnetic field configuration to be such that the cyclotron radius of the outermost particle at the entrance to the second section is equal to the radial coordinate of the outermost line of force, i.e.,

$$r_1 = r_c = \frac{V_0 M \sin \alpha_1}{e B_1} . \tag{13}$$

Assuming that in the first section the protons move along the magnetic lines of force, we can find the magnetic field at the emitter

$$B_0 = \frac{V_0^2 M^2 \sin^2 \alpha_1}{e^2 B_1 R_0^2} . \tag{14}$$

For the beam radius at the target we can write the expression

$$r_m \cong 2 r_1 \sqrt{\frac{B_1}{B_m}} . \tag{15}$$

Using Eqs. (12)–(15), we obtain

$$V_0 = \frac{e \sqrt{k} r_m B_m}{2M} , \tag{16}$$

where $k = W_0 / W_m^{\perp}$.

Now we can determine the average power density at the target

$$P_s = \frac{M V_0^2}{2e} j = \frac{I_0 B_m^2 k e}{8 \pi M} . \tag{17}$$

We choose by way of example $r_m = 5$ mm, $B_m = 5$ T, $L_1 = 1$ m, ($\alpha_1 = 26.6°$), and $k = 2$. From Eqs. (6), (12), (14), (16), and (17), we obtain $B_0 = 0.6 \times 10^{-4}$ T, $B_1 = 2$ T, $V_0 = 1.7 \times 10^6$ m/s, (15 keV), and $P_s = 2 \times 10^{11}$ W/m^2.

We note that these formulas give only an approximate estimate of the main parameters of the focusing channel. To determine more accurately the parameters of the magnetic field and the target, and also the effect of the initial beam divergence on the focusing, we have carried out the numerical calculations described below.

Calculational model

Let us consider the motion of the protons in an increasing magnetic field with the assumption of complete space-charge compensation of the proton beam. The proton trajectory in an axially symmetric field is determined by the equation

$$M \ddot{r} = \frac{M V_\theta^2}{r} + e V_\theta B_z , \tag{18}$$

$$M \ddot{z} = - e V_\theta B_r , \tag{19}$$

$$M \frac{d}{dt} (r V_\theta) = r (e V_z B_r - e V_r B_z), \tag{20}$$

where V_z, V_r, and V_θ are the axial, radial, and azimuthal velocities and B_z and B_r are the axial and radial components of the magnetic induction.

After making some uncomplicated manipulations, setting

$$B_r = - \frac{1}{2} r \frac{dB_z}{dz}, \tag{21}$$

and using the law of conservation of energy, we can write Eqs. (18)–(20) as

$$\frac{V_0^2 - V_\theta^2}{1 + r'^2} \ddot{r} = \frac{V_\theta^2}{r} + \frac{e}{M} V_\theta B_z - \frac{1}{2} \frac{e}{M} r \dot{r} \cdot V_\theta B_z' , \tag{22}$$

$$V_\theta = - \frac{1}{2} \frac{e}{M} r B_z \left(1 - \frac{r_0^2 B_0}{r^2 B_z} \right), \tag{23}$$

where r_0 and B_0 are the radius and axial component of the magnetic field at the cathode and V_0 is the total ion velocity.

Therefore the ion trajectory $r(z)$ can be obtained from Eqs. (22) and (23) if the distribution $B = B_z(z)$ is known.

In the numerical calculations the beam is divided into N_t current tubes. The current density is assumed to be the same in each tube and equal to the average current density in the ion source, $j_0 = I_0 / \pi R_0^2$. The current in the tube is calculated from the relation

$$I_{ti} = \frac{I_0}{R_0^2} (r_{0i}^2 - r_{0i-1}^2). \tag{24}$$

To allow for the angular distribution of the ions, each tube in turn is subdivided into N_α current tubes, each with a different initial angle of injection $\Delta r'_{0m}$, taking into account the angular distribution of the ions. The initial angle of the imth current tube is given by the relation

$$r'_{0im} = \frac{r_{0i}}{R_0} \tan \alpha_1 + \Delta r'_{0m} , \tag{25}$$

where r_{0i} is the radial coordinate of the ith particle at the emitter.

The angular distribution function was determined with the aid of the current density distribution in the vicinity of the crossover, which was calculated numerically.[1] The problem was solved for the case of zero magnetic field and a uniform current density at the diode. Then the current density at the crossover is related directly to the distribution of the particles over initial angles, since the radial coordinates of the particles at the crossover are $r_{cr} = \Delta r'_0 R_c$ (where R_c is the radius of curvature of the cathode). The fraction of particles with initial angles $\Delta r'_0 \leqslant \Delta r'_k$ is given by the expression

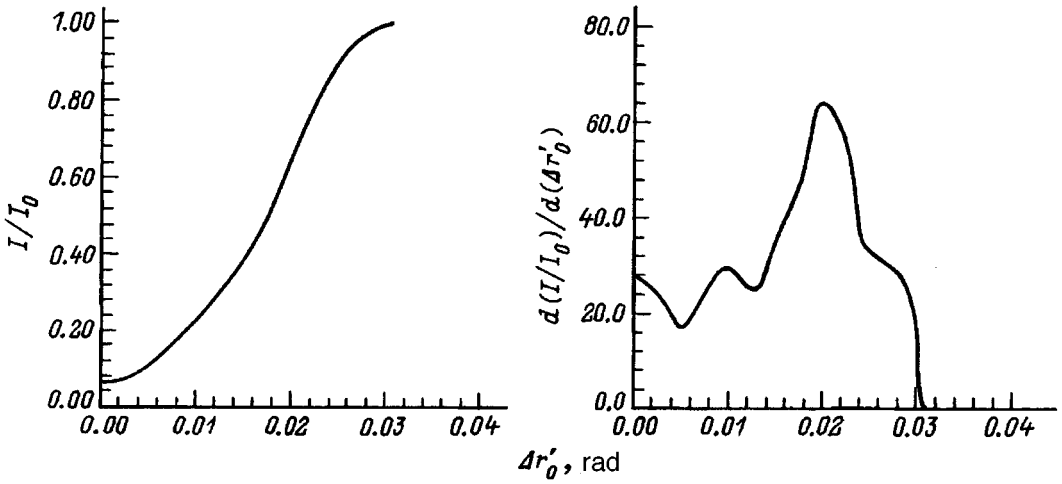

FIG. 2. Proton angular distribution function.

$$\frac{I(\Delta r_k')}{I_0} = \frac{2\pi}{I_0}\int_0^{r_k} j(r) r dr, \tag{26}$$

where $r_k = \Delta r_k' R_c$ and $j(r)$ is the current density distribution at the crossover.

Figure 2 shows the dependence of I/I_0 on $\Delta r_0'$, calculated by this method. In the calculation of I/I_0 it was assumed that $\Delta r'_{\max} = 0.03$, in accordance with the experimental results given in Ref. 1. In particular, one can see that 23% of the ions have $|\Delta r_0'| \leqslant 0.01$ and 66% have $|\Delta r_0'| \leqslant 0.02$.

The power density distribution at the target associated with the longitudinal component of the particle velocity is calculated by the formula

$$P_{zs}(r) = U_0 \sum_{i=1}^{N_t} \sum_{m=1}^{N_\alpha} k_m j_{im}(r) \left[\frac{V_{zim}}{V_0}\right]^2, \tag{27}$$

where U_0 is the accelerating potential, N_α is the number of intervals the angular distribution function is divided into, k_m is the weighting factor for the mth interval, and j_{im} is the density of the imth current tube at the target, i.e., $j_{im} = j_0(r_{0i}^2 - r_{0i-1}^2)/(r_{im}^2 - r_{i-1m}^2)$.

Results of the calculations

The numerical calculations confirmed the conclusions obtained from the estimates that it is not possible to focus the beam to the required power density in a conical magnetic field.

Let us examine the results of the calculations for the double-conical magnetic field. In Fig. 1b we have plotted the trajectory of the outermost particle in a double-conical field for the following parameters: $R_0 = 0.5$ m, $B_0 = 5 \times 10^{-4}$ T,

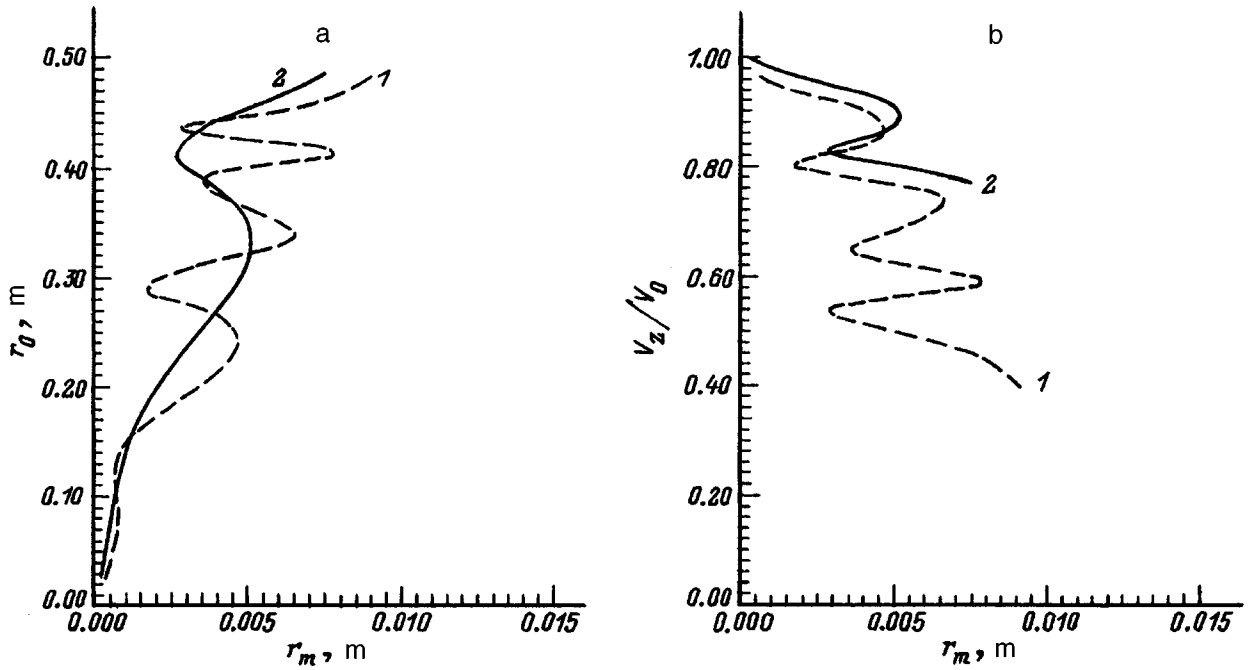

FIG. 3. Curves of $r_0(r_m)$, (*a*) and $V_z/V_0(r_m)$ (*b*) for a beam of 15 keV protons. $B_0 = 5 \times 10^{-4}$ T, $B_1 = 1$ T, $B_m = 5$ T, $\Delta r_0' = 0$, $L_2 = 0.5$ m, L, m: *1*—1.0, *2* — 1.5.

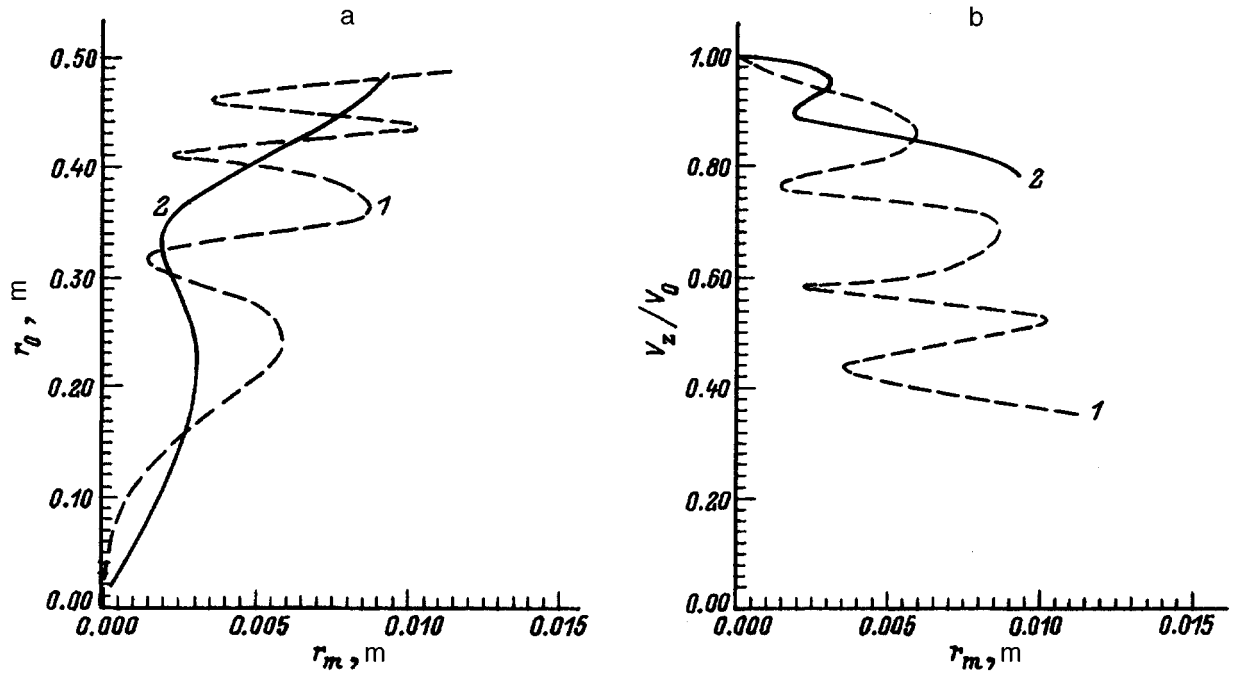

FIG. 4. Curves of $r_0(r_m)$ (a) and $V_z/V_0(r_m)$ (b) for a beam of 30 keV protons. The parameters are the same as in Fig. 3.

$B_1=1$ T, $B_m=5$ T, $L_1=1$ m, $L_2=0.5$ m, $\Delta r_0'=0$, and the ion energy is 15 keV. It can be seen that in the first section the ions move inertially and in the second section they are magnetized and move adiabatically. Figure 3 (curve *1*) shows $r_0(r_m)$ and $V_z/V_0(r_m)$ for this case. The beam radius at the target is 10 mm. The beam is not laminar, and the longitudinal velocity lies in the interval $V_z=(0.35–1)V_0$. By an increase of the length L_1 to 1.5 m (Fig. 3, curve *2*) it is possible to decrease the beam radius to $r_m=8$ mm, increase the longitudinal velocity to $(0.75–1)V_0$ and accordingly increase the average power density at the target.

An increase in the energy to 30 keV (Fig. 4) increases the average power density by 35–40%. The beam parameters at the target are very sensitive to the values of B_0, B_1, and L_1.

These calculations provided the optimum magnetic field configuration with the parameters $B_0=10^{-9}$ T, $B_1=0.25$ T, $L_1=1.5$ m, and $L_2=0.5$ m.

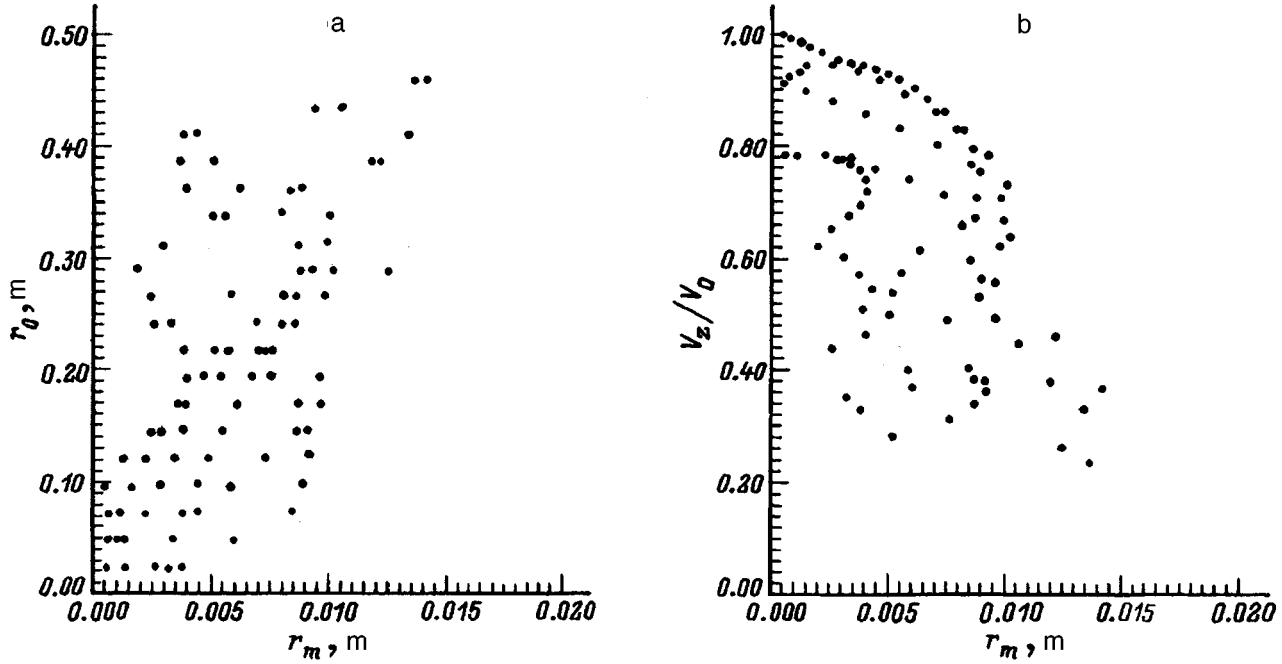

FIG. 5. Curves of $r_0(r_m)$ (a) and $V_z/V_0(r_m)$ (b) for a beam of 30 keV protons for the optimum magnetic field configuration. $B_0=10^{-9}$ T, $B_1=0.25$ T, $B_m=5$ T, $L_1=1.5$ m, $L_2=0.5$ m, $\Delta r_0'=0,\pm 0.01,\pm 0.02,\pm 0.03$.

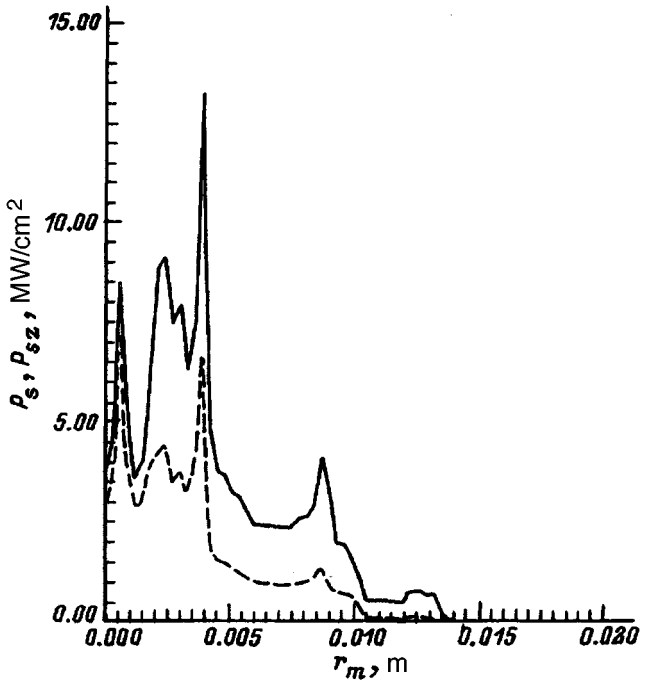

FIG. 6. Radial distribution of the total power density P_s (solid curve) and the longitudinal power density P_{sz} (dashed curve) for a beam of 30 keV protons.

In Fig. 5 we show the distributions $r_0(r_m)$ and $V_z/V_0(r_m)$ for various $|\Delta r_0'| \leqslant 0.03$. Figure 6 shows the distribution of the total power density P_s and the longitudinal power density P_{sz} (related to the longitudinal component of the velocity). From these results one can see that with such a magnetic field configuration it is possible to obtain close to the required power density in a region of radius less than 5 mm.

Effect of the electron background

In the previous sections the motion of the ions was analyzed without taking into account the behavior of the electrons that compensate the space charge of the beam. Actually, the compensating electron background can have a large effect on focusing the beam. Because of the different degree of magnetization of the electrons and ions, and because of the angular spread of the ions, a radial electric field can be set up that improves focusing the ion beam. In support of this statement, we present the following estimate. We assume that the electrons are immobile (rigidly fixed to the magnetic lines of force) and at each cross section of the beam the electron and ion charges are equal. Then the radial electric field can be written as

$$E_r = \frac{I(R)}{2\pi\varepsilon_0 V_z}\left(\frac{1}{r} - \frac{r}{R^2}\right), \tag{28}$$

where $R = R(z)$ is the radial coordinate of the magnetic lines of force, and $r(z)$ is the radial coordinate of the ion, which at $z = 0$ coincides with the radial coordinate of the magnetic line of force.

The equation that describes the radial motion of the ion can be written as

$$M\ddot{r} = \frac{MV_\theta^2}{r} + eV_\theta B_z + \frac{eI}{2\pi\varepsilon_0 V_0}\left(\frac{1}{r} - \frac{r}{R^2}\right), \tag{29}$$

where

$$V_\theta = \frac{eB_z}{2M} r\left(1 - \frac{r_0^2 B_{z0}}{r^2 B_z}\right). \tag{30}$$

We write r in the form

$$r = R + \Delta r, \tag{31}$$

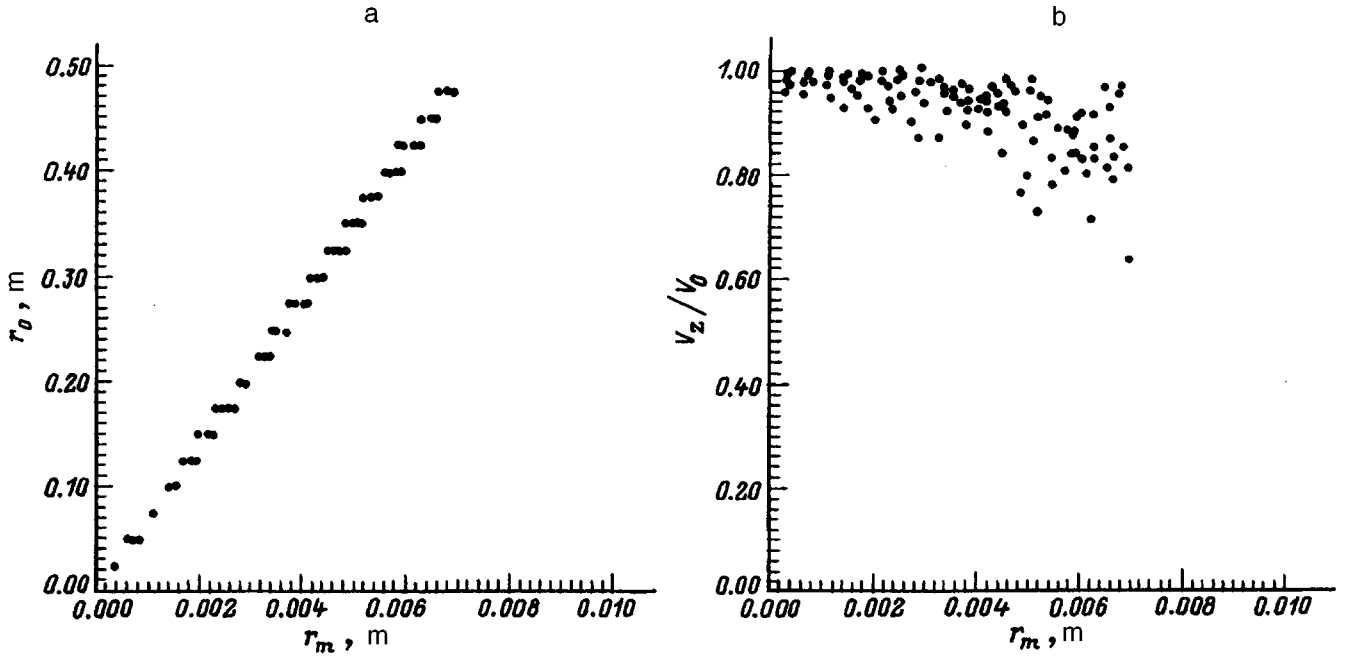

FIG. 7. Plots of $r_0(r_m)$ (a) and $V_z/V_0(r_m)$ (b) for a 30 keV proton beam, calculated with allowance for the radial electric field created by the electron background. $B_0 = 10^{-3}$ T, $B_1 = 0.1$ T, $B_m = 5$ T, $L_1 = 1.5$ m, $L_2 = 0.5$ m, $\Delta r_0' = 0, \pm 0.01, \pm 0.02, \pm 0.03$.

where $\Delta r \ll R$.

Substituting Eq. (31) into (30), we obtain

$$\Delta\ddot{r} + \omega^2 \Delta r = \ddot{R}, \tag{32}$$

where

$$\omega^2 = \frac{e^2 B_z^2}{M^2} + \frac{eI}{\pi\varepsilon_0 M V_z R^2} = \omega_c^2 + \omega_p^2, \tag{33}$$

and ω_p is the frequency of the plasma oscillations of the ions in the plasma with a density equal to $I/\pi R^2 V_z$.

The ratio $\omega_p/\omega_c = \sqrt{MI/\pi\varepsilon_0 e V_z^2 R^2 B_z}$ for $I = 1$ kA and a proton energy of 20 keV is larger than 17 over the entire region of transmission of the ion beam. Consequently it can be assumed that the electron background can hold the ions near the magnetic lines of force more effectively than the magnetic field can. Figure 7 shows curves of $r_0(r_m)$ and $V_z/V_0(r_m)$ obtained with allowance for the radial electric field created by the electron background for $B_0 = 10^{-3}$ T. The radius of the beam is $\sim$ 7 mm, and the power density (Fig. 8) is higher than 20 MW/cm^2. The permissible initial angular divergence (at which the power density at the target is not lower than 15 MW/cm^2) is $\sim$0.07. These estimates hence indicate that the electrons that neutralize the space charge of the proton beam can have a substantial effect on focusing the beam. For more reliable conclusions it is necessary to make calculations of greater accuracy, taking into account the mobility of the electrons along the magnetic lines of force, i.e., based on the theory of plasma optics[3] and the appropriate experimental investigations.

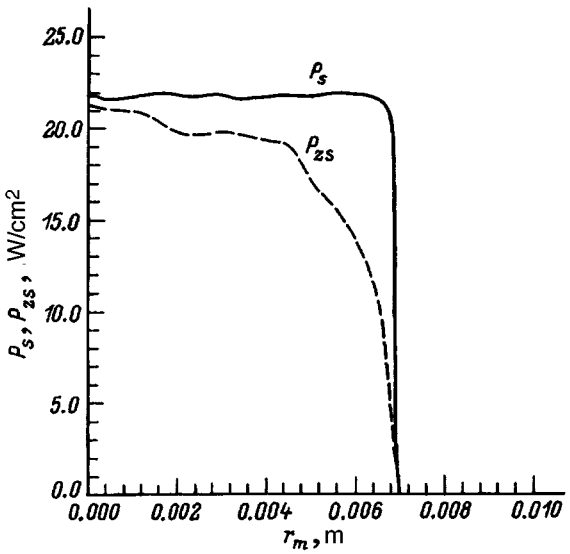

FIG. 8. Radial distribution of the total power density P_s and the longitudinal power density P_{sz} for a 30 keV proton beam with allowance for the radial electric field created by the electron background.

Conclusions

The results of these theoretical and calculational investigations have shown that the method proposed in Ref. 1 for focusing neutralized proton beams, which is based on a combination of ballistic focusing plus adiabatic magnetic compression, provides an area compression factor of $\sim 1.5\times 10^4$ and a 50% efficiency in beam transmission. For an initial proton beam diameter of $\sim$ 1 m it is possible to obtain a diameter of 1 cm at the target and a power density up to 10 MW/cm^2. The total length of the focusing channel is 2 m, the length of the ballistic section is 1.5 m, the magnetic field in the region of the diode is $\sim 10^{-3}$ T, at the end of the ballistic region it is $\sim 10^{-1}$ T, and at the target it is 5 T. the power density is sensitive to the initial angular distribution of the particles and to the magnetic field distribution, and depends only weakly on the initial kinetic energy of the particles. The spatial distribution of the electron background and its dynamics can have a large effect on the efficiency of focusing the beam.

[1] V. Engelko, Ch. Schultheiss, and H. Wuerz, Primärbericht. 31.02.03P. INR, Kernforschungszentrum Karlsruhe, GmbH. (1992).
[2] D. Mosher, Phys. Fluids **20**, 1148 (1977).
[3] A. I. Morozov and S. V. Lebedev, *Plasma Optics. Topics in Plasma Theory*, No. 8 [in Russian] (Atomizdat, Moscow, 1974).

Translated by J. R. Anderson

Relative contribution of various factors to the formation of the energy spectrum of fast, medium-energy, charged particles and ions transmitted through a thin target with fluctuations of the target thickness

N. N. Koborov, A. I. Kuzovlev, V. A. Kurnaev, and V. S. Remizovich

Engineering-Physics Institute, Moscow, Russia
(Submitted January 29, 1996)
Zh. Tekh. Fiz. **67**, 81–93 (May 1997)

The characteristics of the energy spectra of kiloelectron-volt protons transmitted through a free-standing foil are investigated theoretically and experimentally as functions of the angle of incidence of the beam on the target. Analytical expressions for the average characteristics of the transmitted-particle energy spectrum are determined for the case of small-angle scattering. The combined influence of various factors affecting the formation of the energy spectra is taken into account: systematic stopping of particles in the medium, fluctuations of the particle energy losses in inelastic collisions, bending of the particle trajectories due to multiple elastic scattering, and fluctuations of the target thickness. It is shown that the contributions of these factors to the width of the transmitted-particle energy spectrum depend differently on the angle of incidence of the beam on the target surface. On the basis of this differentiation it is inferred from the experimental dependence of the width of the energy spectra of kiloelectron-volt protons transmitted through a free-standing foil on the angle of incidence of the beam that fluctuations of the particle energy losses in inelastic collisions are the predominant factor in the formation of the proton energy spectra.
[S1063-7842(97)01305-6]

INTRODUCTION

One of the fundamental quantities in physics is the cross section for interaction of charged particles with an atom. In the case of fast charged particles ($v \gg v_{at}$, where v is the velocity of the incident particle, and v_{at} is a characteristic velocity of electrons in the atom) the scattering cross section has been thoroughly studied, both theoretically and experimentally. This study was possible because the Born approximation can be used in describing the scattering process for fast particles. Even in this case, however, it is necessary to know the wave functions of all states of the atom, which, as a rule, are not known. This consideration places special significance on the first and second moments of the scattering cross section, i.e., the specific energy losses $\overline{\varepsilon}_{at}(T)$ and the mean-square energy loss $\overline{\varepsilon_{at}^2}(T)$. For the scattering of fast charged particles the quantity $\overline{\varepsilon}_{at}(T)$ is given by the Bethe-Bloch formula,[1] and $\overline{\varepsilon_{at}^2}(T)$ is defined in the free-electron model using the Rutherford law.[2] In regard to the experimental feasibility of measuring the cross section for interaction of fast particles with an atom, this is accomplished without much difficulty. Since particles with such energies have a very long mean free path, a relatively thin target can be prepared, so that particles passing through it actually undergo single scattering by atoms[3] of the material.

The situation changes significantly when the particle energy is lowered. Even for particles of moderate energies ($v \sim v_{at}$) insurmountable mathematical difficulties are encountered in the theoretical determination of the scattering cross section, primarily in connection with the inadmissibility of the Born approximation for such energies. The difficulties of investigating the scattering cross section experimentally increase drastically. In this range the mean free path of the particles decreases to such an extent that it is technologically impossible to prepare targets thin enough to observe purely single scattering. In other words, multiple scattering always occurs in any experiment for particles of such energies. A hypothetically different situation is therefore encountered: The determination of the single-scattering characteristics [not only in regard to the cross section itself, but also to $\overline{\varepsilon}_{at}(T)$ and $\overline{\varepsilon_{at}^2}(T)$] requires the formulation of an appropriate procedure capable of utilizing the necessary information from the experimentally measured energy spectra of multiply scattered particles. This approach postulates the existence of an analytical theory that relates the energy spectra of multiply scattered particles to the single-scattering characteristics.

Of course, in no way are the preceding remarks meant to imply that attempts have not been made to determine the principal characteristics of the scattering cross section in the intermediate energy range. For example, the specific energy losses $\overline{\varepsilon}_{at}(T)$ have been calculated under various assumptions in Refs. 4 and 5. The functional relation obtained in this work, $\overline{\varepsilon}_{at}(T) \sim \sqrt{T}$ agrees quite well with experimental data on the range of intermediate-energy particles in the medium.[6] As for the quantity $\overline{\varepsilon_{at}^2}(T)$, scarcely any relevant theoretical or experimental data exist in the medium energy range [at any rate, in all the authors' many years devoted to transport theory they have never encountered data on $\overline{\varepsilon_{at}^2}(T)$ at energies in the range of several keV/nucleon]. This deficit can be attributed to the fact that the scattering of particles in a target in the given energy range, as mentioned above, is definitely of the multiple kind, and attempts to make thinner targets further accenuate the role of fluctuations in the target thick-

ness. A procedure capable of distinguishing the influence of these factors on the formation of the particle energy spectra has not yet been developed (e.g., this cannot be done for normal incidence of the particles on the target).

Here we give the results of an experimental and theoretical study of the energy spectra of particles in a beam obliquely incident on a free-standing thin target. We show that the contributions of target thickness fluctuations, curvature of the particle trajectories due to elastic scattering, and fluctuations of the inelastic energy losses in single scattering $[\sim\overline{\varepsilon_{\text{at}}^2}(T)]$ to the width of the energy spectrum of the transmitted particles depend differently on the angle of particle incidence on the target. This fact can be exploited to discern the quantity associated specifically with $\overline{\varepsilon_{\text{at}}^2}(T)$ in the energy spectrum of the transmitted particles and eventually to find the value of $\overline{\varepsilon_{\text{at}}^2}(T)$ in this energy range. In the article, therefore, we indicate a possible technique for determining experimentally the values of $\overline{\varepsilon_{\text{at}}^2}(T)$ in the medium energy range.

When a broad, initially monoenergetic beam of heavy charged particles ($m \gg m_e$, where m is the particle mass, and m_e is the electron mass) or ions is transmitted through a target, they lose energy, and the beam is no longer monoenergetic. The energy spectrum of the particle transmitted through the material is attributable to many factors, mainly the following: 1) systematic stopping of particles in the medium[7–9]; 2) the probabilistic character of the energy losses in a single inelastic scattering event[7–9]; 3) fluctuations of the particle ranges due to multiple elastic scattering[7–9]; 4) correlation between the scattering angle of a particle and the energy lost by it in a single scattering event[10–12]; 5) reflection of particles from the target[13]; 6) inhomogeneity and thickness fluctuations of the target.[14]

The relative role of these factors in the formation of the transmitted-particle energy spectrum depends on the initial energy of the particles, their mass and charge, the angle of incidence of the beam on the surface of the medium, and the material and thickness of the target. In real experiments it is often difficult to state beforehand which particles are predominant in the formation of the transmitted-particle energy spectrum. It is therefore necessary to investigate the combined influence of all factors theoretically.

The analytical treatment of the transmission of fast charged particles and ions of medium energy through a substance, taking into account all the above-stated factors, is rendered impossible at present by the extreme mathematical complexity of the problem. We shall therefore assume below that particle reflection from the target can be disregarded. For heavy charged particles or ions this implies exclusion of grazing incidence, when the fraction of reflected particles can be appreciable.[15–18] We shall also ignore any correlation between the scattering angle of a particle and the energy lost by it in a single collision. The influence of this effect on the energy spectrum of multiply scattered particles has not been adequately studied theoretically and required separate consideration. For example, in the description of multiple scattering of particles in Ref. 12 the correlation between the scattering angle of a particle and the energy lost by it in elastic scattering by nuclei of target atoms was considered only for low energy losses and without regard for fluctuations of the particle ranges. A subsequent attempt[11] to explain the previously observed[10] sharp dependence of the average transmitted-particle energy loss on the angle of observation yielded a glaring discrepancy with the experimental data. In the ensuing discussion we merely indicate the situations in which this effect can be decisive in the formation of the energy spectrum of particles transmitted through the material.

As for the remaining factors, only in the case of normal incidence of the beam on the target has their influence on the energy spectrum of heavy charged particles been studied to any degree of completion. For example, the transformation of the particle energy spectrum as a result of systematic stopping in the medium, the probabilistic nature of the particle energy losses in inelastic collisions, and curvature of the particle trajectory by multiple elastic scattering have been analyzed in detail.[7–9] It has been shown[14] that the width of the energy spectrum of light kiloelectron-volt ions transmitted through a thin graphite foil prepared by vacuum deposition is mainly attributable to fluctuations of the film thickness.

STATEMENT OF THE PROBLEM; THEORETICAL ANALYSIS

Let a broad beam of heavy charged particles with energy T_0 be incident at the angle θ_0 relative to the normal on a planar target having an average thickness L. The planes $z=0$ and $z=L$ coincide with the left and right boundaries of the target, respectively. We assume that the velocity vector of the incident particles is situated in the xz plane in the direction defined by the azimuth angle $\varphi_0=0$. We describe the direction of propagation of the particles by two angles: the polar angle θ (θ is the angle between the direction of motion of the particles and the z axis) and the azimuth angle φ.

The particle flux density at the depth z, $N(z,\theta,\varphi,T)$, obeys the transport equation[7–9]

$$\cos\,\theta\frac{\partial N}{\partial z}=\hat{I}_{\text{el}}+\hat{I}_{\text{in}}\,. \tag{1}$$

Here $\hat{I}_{\text{el}}$ and $\hat{I}_{\text{in}}$ are the elastic and inelastic collision integrals, respectively. We assume that the angle of incidence θ_0 is not too large ($\theta_0<60°$), so that reflected particles are essentially nonexistent,[15–18] and their influence can be disregarded. In this case the boundary condition to Eq. (1) has the form

$$N(z=0,\theta,\varphi,T)=N_0\delta(\cos\,\theta-\cos\,\theta_0)\,\delta(\varphi)\,\delta(T-T_0). \tag{2}$$

The elastic scattering of heavy charged particles takes place primarily at small angles ($\vartheta_{\text{eff}}\sim\lambda/r_{\text{at}}\simeq\sqrt{m_e/m}\sqrt{13.6\text{ eV}/T}Z^{1/3}\ll 1$, where ϑ_{eff} is an effective single-scattering angle, λ is the de Broglie wavelength of the particles, r_{at} is a characteristic atomic radius of the medium, m_e is the electron mass, m is the particle mass, and Z is the atomic number of the target). Consequently, the diffusion approximation (Fokker-Planck approximation) with respect to the angles can be used for the elastic collision integral:[7–9]

$$\hat{I}_{\rm el} \simeq \frac{1}{4}\langle \Theta_S^2(T)\rangle \left\{ \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right\} N(z,\theta,\varphi,T), \tag{3}$$

where $\langle \Theta_S^2(T)\rangle$ is the mean-square scattering angle of a particle with energy T per unit path.

For heavy charged particles the energy ε lost in a single inelastic collision is small in comparison with the energy T of the particle itself [$\varepsilon \sim (m_e/m)T \ll T$]. We therefore write the inelastic collision integral in the Fokker-Planck approximation:[7–9]

$$\hat{I}_{\rm in} \simeq \frac{\partial}{\partial T}\{\overline{\varepsilon}(T)N(z,\theta,\varphi,T)\} + \frac{1}{2}\frac{\partial}{\partial T}\left\{\overline{\varepsilon^2}(T)\frac{\partial N}{\partial T}\right\}, \tag{4}$$

where $\overline{\varepsilon}(T)$ is the stopping power of the medium which, in general, includes both inelastic and elastic energy losses associated with the scattering of particles by atoms of the medium,[9] $\overline{\varepsilon}(T) = \overline{\varepsilon}_{\rm ion}(T) + \overline{\varepsilon}_{\rm nuc}(T)$, and $\overline{\varepsilon^2}(T)$ is the mean-square energy lost by a particle with energy T per unit path.

We note the admissibility of using elastic and inelastic collision integrals of the form (3) and (4) in the range of intermediate energies, such that nuclear interactions associated with collisions of heavy particles with the nuclei of target atoms can be ignored. Otherwise, the scattering is not small-angle, and additional inelastic processes are possible, for example, the absorption of particles.[9]

In Eq. (1) we can transform from the energy T to the variable $s = \int_T^{T_0} dT'/\overline{\varepsilon}(T')$. In the continuous-slowing model[7–9] s represents the path over which the particle loses energy from T_0 to T. From now on we refer to the variable s simply as the path. Changing variables in Eq. (1), we obtain the following equation for the quantity $N(z,\theta,\varphi,s) = \overline{\varepsilon}(T)N(z,\theta,\varphi,T)$:

$$\cos\theta\frac{\partial N(z,\theta,\varphi,s)}{\partial z} = \frac{\langle \Theta_S^2(s)\rangle}{4}\left\{\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial N}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2 N}{\partial\varphi^2}\right\} - \frac{\partial N}{\partial s} + \frac{1}{2}\frac{\partial}{\partial s}\left\{\frac{\overline{\varepsilon^2}(s)}{\overline{\varepsilon}(s)}\frac{\partial}{\partial s}\frac{N}{\overline{\varepsilon}(s)}\right\}. \tag{5}$$

If the particles move along a straight line, the path traversed by them at the depth z would be equal to $z/\cos\theta_0$. Because of elastic scattering, the particle trajectories are not straight, and we therefore introduce the variable $u = s - z/\cos\theta_0$. Transforming from the variables z and s to the variables $z' = z/\cos\theta_0$ and u in Eq. (5), we obtain

$$\frac{\cos\theta}{\cos\theta_0}\frac{\partial N(z',\theta,\varphi,u)}{\partial z'} + \frac{\cos\theta_0 - \cos\theta}{\cos\theta_0}\frac{\partial N}{\partial u} = \frac{\langle \Theta_S^2(z'+u)\rangle}{4}\left\{\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial N}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2 N}{\partial\varphi^2}\right\} + \frac{1}{2}\frac{\partial}{\partial u}\left\{\frac{\overline{\varepsilon^2}(z'+u)}{\overline{\varepsilon}(z'+u)}\frac{\partial}{\partial u}\frac{N}{\overline{\varepsilon}(z'+u)}\right\}. \tag{6}$$

So far, we have not relied on any approximations other than the diffusion approximation with respect to the angles and energies. To further simplify the transport equation, we make use of the fact that elastic scattering is a small-angle process for heavy charged particles over the entire path, and the fluctuations of the energy losses are small.[7–9] This means that the path traversed by a particle in the medium differs very little from z', i.e., $u \ll z'$, and the directions of motion of the majority of the particles fall within a narrow cone around the initial direction (θ_0, φ_0). Since the functions $\langle \Theta_S^2\rangle$, $\overline{\varepsilon}$, and $\overline{\varepsilon^2}$ in Eq. (6) are sufficiently smooth functions of their arguments, we can ignore the quantity u in them in comparison with z' and write

$$\langle \Theta_S^2(z'+u)\rangle \simeq \langle \Theta_S^2(z')\rangle, \quad \overline{\varepsilon}(z'+u) \simeq \overline{\varepsilon}(z'), \quad \overline{\varepsilon^2}(z'+u) \simeq \overline{\varepsilon^2}(z').$$

We now transform to a coordinate frame, whose polar axis is directed along the initial particle direction. In this frame the angular distribution of the particles can be described by the angles α and β, which are related to the angles θ and φ by

$$\sin\alpha = \sin\theta\cos\varphi\cos\theta_0 - \cos\theta\sin\theta_0,$$
$$\sin\beta = \sin\theta\sin\varphi. \tag{7}$$

By this definition the angles α and β have the geometrical significance that α is the angle between the direction of propagation of the particles and the plane through the y axis and the initial velocity vector of the particles, and β is the angle between the direction of propagation of the particles and the xy plane. The small-angle character of the scattering of heavy charged particles implies that the angles α and β are small, so that we can write

$$\sin\alpha \simeq \alpha, \quad \sin\beta \simeq \beta$$

and regard the angles α and β as constants in the infinite limits $-\infty < \alpha$ and $\beta < +\infty$. In normal incidence ($\theta_0 = 0$) we infer from (7) that $\alpha \simeq \theta\cos\varphi$ and $\beta \simeq \theta\sin\varphi$, i.e., $\alpha = \theta_x$ and $\beta = \theta_y$, where θ_x and θ_y are the angles between the direction of particle propagation and the yz and xz planes, respectively.[7–9]

Carrying out the above-described procedures, we obtain the following equation for the particle flux density $N(z',\alpha,\beta,u)$:

$$\frac{\partial N(z',\alpha,\beta,u)}{\partial z'} + \left(\frac{1}{2}(\alpha^2+\beta^2) + \alpha\tan\theta_0\right)\frac{\partial N}{\partial u} = \frac{\langle \Theta_S^2(z')\rangle}{4}\left\{\frac{\partial^2 N}{\partial\alpha^2} + \frac{\partial^2 N}{\partial\beta^2}\right\} + \frac{1}{2}\frac{\overline{\varepsilon^2}(z')}{(\overline{\varepsilon}(z'))^2}\frac{\partial^2 N}{\partial^2 u} \tag{8}$$

subject to the boundary condition

$$N(z'=0,\alpha,\beta,u)=N_0\delta(\alpha)\delta(\beta)\delta(u). \tag{9}$$

The simplest way to get from (5) to (8) is to take into account the invariance of the elastic collision integral under rotation of the axes and, hence, the fact that it has the same form as in the case of normal incidence: $\Delta_{\theta,\varphi}\simeq\partial^2/\partial\theta_x^2+\partial^2/\partial\theta_y^2$ ($\theta_x\to\alpha$, $\theta_y\to\beta$). We also make use of the fact that

$$\cos\,\theta=\cos\,\theta_0\sqrt{1-\sin^2\,\alpha-\sin^2\,\beta}-\sin\,\alpha\,\sin\,\theta_0$$
$$\simeq\cos\,\theta_0\left(1-\alpha\,\tan\,\theta_0-\frac{1}{2}(\alpha^2+\beta^2)\right),$$

and retain the first nonvanishing terms in α and β. In Eq. (8) we have not disregarded the terms α^2 and β^2 in comparison with $\alpha\tan\theta_0$, because in normal incidence ($\theta_0=0°$) $\tan\theta_0=0$. Equations (8) and (9) describe the propagation of a broad beam of heavy charged particles in the medium in both normal and oblique incidence on the surface of the medium for small-angle scattering relative to the initial direction of the particles.

Equation (8) differs from the equation describing particle propagation for normal incidence of the beam on the medium[7–9] by the term $\alpha\tan\theta_0(\partial N/\partial u)$, which depends on the angle of incidence θ_0. We seek a solution of Eq. (8) in the form

$$N(z',\alpha,\beta,u)=\frac{N_0}{2\pi r}\int dp\,\frac{\exp(pu)}{\pi A(z',p)}\times\exp\left\{C(z',p)-\frac{\beta^2+(B(z',p)+\alpha)^2}{A(z',p)}\right\}. \tag{10}$$

Substituting Eq. (10) into (8) and equating coefficients of like powers of α and β, we find

$$B(z',p)=\left(1-\exp\left\{-\frac{p}{2}\int_0^{z'}dz''A(z'',p)\right\}\right)\tan\,\theta_0,$$

$$C(z',p)=\frac{p^2}{2}\int_0^{z'}dz''\frac{\overline{\varepsilon^2}(z'')}{(\overline{\varepsilon}(z''))^2}-\frac{p}{2}\int_0^{z'}dz''A(z'',p)+\frac{p}{2}\left(z'-\int_0^{z'}dz''\times\exp\left\{-p\int_0^{z''}dz'''A(z''',p)\right\}\right)\tan^2\,\theta_0. \tag{11}$$

The function $A(z',p)$ satisfies the first-order ordinary differential equation

$$\frac{dA(z',p)}{dz'}+\frac{1}{2}pA^2(z',p)=\langle\Theta_S^2(z')\rangle,\quad A(z'=0,p)=0. \tag{12}$$

Equation (12) and its solution for fast charged particles have been analyzed in detail.[8,9]

If the target had exactly the same thickness L over its entire area, we could solve the stated problem by setting $z'=L/\cos\,\theta_0$ in Eq. (10). In other words, the distribution (10), in conjunction with (7) and the equation

$$u=\int_T^{T_0}dT'/\overline{\varepsilon}(T')-L/\cos\,\theta_0,$$

determines the angular and energy spectra of particles transmitted through a layer of homogeneous material of fixed thickness L.

However, it is entirely obvious that the thickness of a thin foil is a random function of position on the foil surface, $L=L(x_0,y_0)$, and depends on the procedure used to prepare the foil. Consequently, in the case of a broad particle beam incident on a target, in general, it is first necessary to consider the propagation of a narrow particle beam and then to average over the transverse coordinates of the point where the particles enter the medium. For a wide-angle particle detector, i.e., a detector that records all particles independently of their point of exit, this procedure is equivalent to averaging over the target thickness z:

$$N=\int\frac{dx_0dy_0}{\Sigma}N(L(x_0,y_0))=\int dzN(z)\int\frac{dx_0dy_0}{\Sigma}\delta(z-L(x_0,y_0))=\int dzN(z)f(z),$$

where Σ is the area of the target.

The transition from averaging over the transverse coordinates of the beam entry point to averaging over the target thickness is illustrated in the Appendix in the example of a simpler model in which the propagation of particles in the medium is treated in the so-called forward-backward approximation.[7] We assume that the probability $f(z)$ of the target having a thickness z is described by the normal probability density function

$$f(z)=\frac{1}{\sqrt{2\pi}\Delta L}\exp\left\{-\frac{(L-z)^2}{2(\Delta L)^2}\right\}, \tag{13}$$

where L is the foil thickness averaged over the transverse directions, and ΔL is the variance of the thickness, which for a thin foil depends mainly on its preparation technique.

To permit the target to be treated as a planar foil, we make the natural assumption that $\Delta L\ll L$.

Multiplying Eq. (10) by (13), integrating with respect to z, and taking into account the smallness of the variance of the distribution along the thickness, we obtain the following expression for the flux density of particles transmitted through a plane layer of thickness L (from now on we omit the word ''average'' when referring to the average thickness):

$$N(L,\alpha,\beta,s)=\frac{N_0}{2\pi r}\int dp\,\frac{\exp\{p(s-L/\cos\,\theta_0)\}}{\pi A(L/\cos\,\theta_{0,p})}\times\exp\left\{\frac{(\Delta L)^2}{2\cos^2\,\theta_0}p^2+C(L/\cos\,\theta_0,p)-\frac{\beta^2+(B(L/\cos\,\theta_0,p)+\alpha)^2}{A(L/\cos\,\theta_0,p)}\right\}. \quad (14)$$

Integrating the distribution (14) over all scattering angles α and β, we obtain expressions for the energy spectrum of the transmitted particles, irrespective of the angles of observation (wide-angle detector):

$$N(L,s)=\frac{N_0}{2\pi r}\int dp\,\exp\left\{p(s-L/\cos\,\theta_0)+\frac{(\Delta L)^2}{2\cos^2\,\theta_0}p^2+C(L/\cos\,\theta_0,p)\right\}. \quad (15)$$

We see that even in this case it is necessary to know the explicit form of the function $A(z',p)$ over the entire range of p in order to find the energy spectrum. The equation for $A(z',p)$ (12) is a nonlinear differential equation of the Riccati type. A solution cannot be found for an arbitrary function $\langle\Theta_S^2(z')\rangle$. A sufficiently reliable expression has yet to be found for $\langle\Theta_S^2(T')\rangle$. These considerations make it impossible to calculate $A(z',p)$ for arbitrary values of p. We therefore confine the ensuing calculations to certain average characteristics of the energy spectrum of the transmitted particles, when it is not necessary to know the explicit expression for $A(z',p)$ for all p.

Integrating the distribution (15) over the traversed paths s, we obtain an expression for the total flux $N(L)$ of particles transmitted through a layer of thickness L:

$$N(L)=\int_0^{R_0}dsN(L,s),$$

where $R_0=\int_0^{T_0}dT'/\overline{\varepsilon}(T')$ is the mean free path of particles of energy T_0 in the material.[7–9]

If the target thickness L satisfies the inequality

$$R_0-L/\cos\,\theta_0\ll\langle\Theta_S^2(T_0)\rangle R_0^2, \quad (16)$$

the particles do not really stop in the target.[7–9] Hence, the upper limit of integration with respect to s in the equation for $N(L)$ can be replaced by infinity. We then have

$$N(L)=N_0. \quad (17)$$

Consequently, for targets whose thicknesses satisfy condition (16) the total flux of transmitted particles does not depend on their thickness and is equal to the flux of particles incident on the surface of the medium.

We now consider the mean path traversed by particles in the target and the variance of the traversed paths. We assume that the thickness L satisfies condition (16). We therefore assume from now on that the upper limit of integration with respect to s is infinite.

a) Mean path traversed by particles, and variance of the traversed paths, irrespective of the angles of observation (wide-angle detector)

When inequality (16) holds, the moments of the distribution (15) can be calculated, in general, without having to seek a solution of Eq. (12), and it suffices merely to determine the values of $A(z',p)$, $\partial A(z',p)/\partial p$, etc., for $p=0$. Indeed, taking into account (17) and the equations

$$\frac{1}{2\pi r}\int_0^\infty duu^n\int dp\,\exp\{pu\}F(p)=(-1)^n\frac{\partial^n}{\partial p^n}F(p)\bigg|_{p=0},$$

we obtain

$$\langle s\rangle_L=\frac{L}{\cos\,\theta_0}-\frac{\partial}{\partial p}C(L/\cos\,\theta_0,p)\bigg|_{p=0}=\frac{L}{\cos\,\theta_0}+\frac{1}{2}\int_0^{L/\cos\,\theta_0}dzA(z,p=0), \quad (18)$$

$$\langle(s-L/\cos\,\theta_0)^2\rangle_L=\frac{(\Delta L)^2}{\cos^2\,\theta_0}+\frac{\partial^2}{\partial p^2}C(L/\cos\,\theta_0,p)\bigg|_{p=0}+\left(\frac{\partial}{\partial p}C(L/\cos\,\theta_0,p)\right)^2\bigg|_{p=0}=\frac{(\Delta L)^2}{\cos^2\,\theta_0}+\int_0^{L/\cos\,\theta_0}dz\frac{\varepsilon^2(z)}{(\varepsilon(z))^2}+\int_0^{L/\cos\,\theta_0}dz\int_0^z dz'A(z',p)\tan^2\,\theta_0-\int_0^{L/\cos\,\theta_0}dz\frac{\partial}{\partial p}A(z,p)\bigg|_{p=0}+\frac{1}{4}\left(\int_0^{L/\cos\,\theta_0}dzA(z,p=0)\right)^2. \quad (19)$$

Setting $p=0$ in (12), we find that $A(z,p=0)$ is equal to the mean-square scattering angle of particles traversing the path $s=z$ in the medium:

$$A(z,p=0)=\langle\theta^2\rangle_z=\int_0^z ds\langle\Theta_S^2(s)\rangle. \quad (20)$$

Differentiating Eq. (12) with respect to p and then setting $p=0$, we find

$$\frac{\partial}{\partial p}A(z,p)\bigg|_{p=0}=-\frac{1}{2}\int_0^z dsA^2(s,p=0)=-\frac{1}{2}\int_0^z ds(\langle\theta^2\rangle_s)^2. \quad (21)$$

Substituting relations (20) and (21) into (18) and (19), we obtain the following expressions for the values of the mean path $\langle s\rangle_L$ and the variance of the paths $D_s(L)$, irrespective of the angles of observation:

$$\langle s\rangle_L=\frac{L}{\cos\ \theta_0}+\frac{1}{2}\int_0^{L/\cos\ \theta_0}ds\langle\theta^2\rangle_s, \tag{22}$$

$$D_s^2(L)=\langle(s-\langle s\rangle_L)^2\rangle_L=\frac{(\Delta L)^2}{\cos^2\ \theta_0}+\int_0^{L/\cos\ \theta_0}dz\,\frac{\overline{\varepsilon^2}(z)}{(\overline{\varepsilon}(z))^2}$$
$$+\int_0^{L/\cos\ \theta_0}dz\int_0^z ds\langle\theta^2\rangle_s\tan^2\ \theta_0$$
$$+\frac{1}{2}\int_0^{L/\cos\ \theta_0}dz\int_0^z ds(\langle\theta^2\rangle_s)^2. \tag{23}$$

For normal incidence of the particles on the target ($\theta_0=0$), as should be expected, expressions (22) and (23) coincide with the relations given in Ref. 9.

We now analyze the results (22) and (23) in greater detail. The first term in (22) represents the path that the particles would traverse in the medium without elastic scattering. As a result of scattering, the particle trajectories bend, and the mean path traversed by the particles in the medium increases, since the path is not bounded in the direction of longer ranges but is bounded in the direction of shorter ranges by the thickness of the scatterer L. This fact is also revealed by the second term in (22). The variance of the particle ranges $D_s^2(L)$ represents the sum of three variances associated with fluctuations of the film thickness (first term), fluctuations of the particle energy losses in a single inelastic scattering event (second term), and fluctuations of the path due to multiple elastic scattering (third and fourth terms). Since all the fluctuations have been assumed from the outset to be small, the variances associated with these three factors are additive in Eq. (23).

b) Mean path traversed by particles, and variance of the traversed paths for fixed angles of observation (narrow-angle detector)

By analogy with the preceding subsection, we can use the distribution (14) to find the mean path traversed by the particles $\langle s\rangle_{L,\theta\varphi}$ and the variance of the traversed paths $D_s(L,\theta,\varphi)$ for fixed angles of observation θ and φ. To simplify the final results, we consider only the case in which the detector records particles in the plane of their incidence, i.e., $\varphi=\varphi_0=0°$. In this case $\beta=0$, $\alpha\simeq\theta-\theta_0$, and we obtain

$$\langle s\rangle_{L,\theta,\varphi=0}=\langle s\rangle_L-\frac{1}{2\langle\theta^2\rangle_{L/\cos\ \theta_0}}\int_0^{L/\cos\ \theta_0}ds(\langle\theta^2\rangle_s)^2$$
$$+\frac{(\theta-\theta_0)^2}{2(\langle\theta^2\rangle_{L/\cos\ \theta_0})^2}\int_0^{L/\cos\ \theta_0}ds(\langle\theta^2\rangle_s)^2$$
$$+\frac{(\theta-\theta_0)\tan\ \theta_0}{\langle\theta^2\rangle_{L/\cos\ \theta_0}}\int_0^{L/\cos\ \theta_0}ds\langle\theta^2\rangle_s, \tag{24}$$

$$D_s^2(L,\theta,\varphi=0)=D_s^2(L)$$
$$-\int_0^{L/\cos\ \theta_0}dz\,\frac{\langle\theta^2\rangle_s}{\langle\theta^2\rangle_{L/\cos\ \theta_0}}\int_0^z ds(\langle\theta^2\rangle_s)^2$$
$$+\frac{1}{4}\left(\int_0^{L/\cos\ \theta_0}ds\,\frac{(\langle\theta^2\rangle_s)^2}{\langle\theta^2\rangle_{L/\cos\ \theta_0}}\right)^2$$
$$-\frac{\tan^2\ \theta_0}{2\langle\theta^2\rangle_{L/\cos\ \theta_0}}\left(\int_0^{L/\cos\ \theta_0}ds\langle\theta^2\rangle_s\right)^2$$
$$+\frac{(\theta-\theta_0)^2}{2(\langle\theta^2\rangle_{L/\cos\ \theta_0})^3}\left\{2\langle\theta^2\rangle_{L/\cos\ \theta_0}\int_0^{L/\cos\ \theta_0}\right.$$
$$\times dz\langle\theta^2\rangle_z\int_0^z ds(\langle\theta^2\rangle_s)^2$$
$$\left.-\left(\int_0^{L/\cos\ \theta_0}ds(\langle\theta^2\rangle_s)^2\right)^2\right\}$$
$$+\frac{(\theta-\theta_0)\tan\ \theta_0}{(\langle\theta^2\rangle_{L/\cos\ \theta_0})^2}\left\{\langle\theta^2\rangle_{L/\cos\ \theta_0}\int_0^{L/\cos\ \theta_0}\right.$$
$$\times dz\int_0^z ds(\langle\theta^2\rangle_s)^2$$
$$+\frac{1}{2}\langle\theta^2\rangle_{L/\cos\ \theta_0}\left(\int_0^{L/\cos\ \theta_0}ds\langle\theta^2\rangle_s\right)^2$$
$$\left.-\int_0^{L/\cos\ \theta_0}dz\langle\theta^2\rangle_z\int_0^z ds(\langle\theta^2\rangle_s)^2\right\}. \tag{25}$$

Here $\langle s\rangle_L$ and $D_s^2(L)$ are defined in Eqs. (22) and (23), and the following relation is taken into account:

$$\frac{\partial^2}{\partial p^2}A(z',p)\bigg|_{p=0}=\int_0^{z'}dz\langle\theta^2\rangle_z\int_0^z ds(\langle\theta^2\rangle_s)^2.$$

It is evident from expressions (24) and (25) that in the case of normal ($\theta_0=0$) incidence of particles on the target scattering always increases the mean range of particles in the target and increases the variance of the ranges. Only squared terms $\sim\theta^2$ are left in Eqs. (24) and (25). The situation changes significantly for the oblique incidence ($\theta_0\neq0$) of particles on the target. If the angle of observation θ is larger than the angle of incidence θ_0 ($\theta>\theta_0$), these quantities increase in the presence of scattering, but in the opposite case

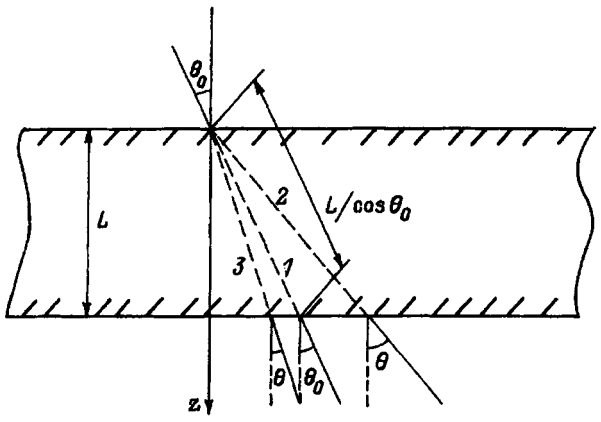

FIG. 1. Conceptual representation of the possible ''trajectories'' of a particle entering the medium at an angle θ_0. Trajectories: *1* — $\theta=\theta_0$; *2* — $\theta>\theta_0$; *3* — $\theta<\theta_0$.

$\theta < \theta_0$ they decrease. This result is physically evident from the geometry of the problem. Thus, if the detector records particles at an angle $\theta > \theta_0$, it can only capture those which are deflected from the initial direction toward the upper boundary of the target, i.e., move along concave trajectories (Fig. 1). Consequently, the path traversed by them in the medium is longer than $L/\cos\theta_0$. In the opposite case, when particles are recorded at an angle $\theta < \theta_0$, their trajectories bend toward the lower boundary, i.e., are convex, and their path is shorter than $L/\cos\theta_0$. However, despite the apparent simplicity of the effect, the situation is not all that trivial; for example, if the angle of observation $\theta = \theta_0$, the mean range of the particles in the medium is certainly not equal to $L/\cos\theta_0$.

CHARACTERISTICS OF THE ENERGY DISTRIBUTIONS OF THE TRANSMITTED PARTICLES

In the preceding section we have determined average characteristics of the path traversed by particles in a target. In practice, however, we work with the energy spectra of the particles. It is therefore necessary to make a transition from the average characteristics of the distribution along the paths to the characteristics of the energy distribution.

We characterize the particle energy distribution by the most probable energy loss $\Delta T_{mp} = T_0 - T_{mp}$, where T_{mp} is the energy at which a maximum is observed in the energy spectrum, and the width of the energy spectrum at the half-maximum points is $\Delta T_{1/2}$.

The particle energy distribution $N(T)$ is related to the path distribution $N(s)$ by the equation $N(T) = N[s(T)]/\overline{\varepsilon}(T)$. For heavy particles the variance of the path distribution D_s is small in comparison with the mean path $\langle s\rangle$ ($D_s \ll \langle s\rangle$); i.e., the spectrum $N(s)$ is narrow. The quantity $\overline{\varepsilon}(T)$, on the other hand, is a comparatively smooth function of the particle energy T. Consequently, the presence of the coefficient $[\overline{\varepsilon}(T)]^{-1}$ does not really have any appreciable influence on the position of the maximum or on the width of the energy spectrum, i.e.,

$$\Delta T_{mp} = T_0 - T(s_{mp}), \quad \Delta T_{1/2} T(s_1) - T(s_2), \tag{26}$$

where s_{mp} is the most probable path traversed by the particles in the medium, and the quantities s_1 and s_2 are the roots of the equation $N(s) = N(s_{mp})/2$.

By virtue of the narrowness of the distribution $N(s)$ ($s_2 - s_1 \ll s_{mp}$) we can disregard the asymmetry of $N(s)$ in the vicinity of its maximum $s_1 < s < s_2$ and assume a Gaussian distribution

$$N(s) \sim \exp\left\{ -\frac{(s-\langle s\rangle)^2}{2D_s^2} \right\}.$$

In this case

$$s_{mp} = \langle s\rangle, \quad s_{1,2} = \langle s\rangle \mp \sqrt{2\ln 2} D_s. \tag{27}$$

The relation between the energy T and the range s has a different form $T(s)$ for different intervals of the energy T_0 of the incident particles. We therefore consider certain limiting cases.

If the thickness of the scatterer is small, $L/\cos\theta_0 \ll R_0$, we can ignore the energy dependence of the quantities $\overline{\varepsilon}(T)$, $\overline{\varepsilon^2}(T)$, and $\langle \Theta_S^2(T)\rangle$. In this case $T(s) = T_0 - \overline{\varepsilon}(T_0)s$, and we obtain

$$\Delta T_{mp} = \overline{\varepsilon}(T_0)\langle s\rangle, \quad \Delta T_{1/2} = 2\sqrt{2\ln 2}\,\overline{\varepsilon}(T_0) D_s, \tag{28}$$

where

$$\begin{aligned}
\langle s\rangle = \langle s\rangle_L &\simeq \frac{L}{\cos\theta_0} + \frac{\langle \Theta_S^2(T_0)\rangle}{4}\frac{L^2}{\cos^2\theta_0}, \\
D_s^2 = D_s^2(L) &\simeq \frac{\overline{\varepsilon^2}(T_0)}{(\overline{\varepsilon}(T_0))^2}\frac{L}{\cos\theta_0} + \frac{(\Delta L)^2}{\cos^2\theta_0} \\
&+ \frac{\langle \Theta_S^2(T_0)\rangle}{6}\frac{L^3}{\cos^3\theta_0}\tan^2\theta_0 \\
&+ \frac{(\langle \Theta_S^2(T_0)\rangle)^2}{24}\frac{L^4}{\cos^4\theta_0}
\end{aligned} \tag{29}$$

for a detector that records particles, irrespective of their exit angle.

For fixed angles of observation we have

$$\begin{aligned}
\langle s\rangle = \langle s\rangle_{L,\theta,\varphi=0} &\simeq \frac{L}{\cos\theta_0} + \frac{\langle \Theta_S^2(T_0)\rangle}{12}\frac{L^2}{\cos^2\theta_0} \\
&+ \frac{(\theta-\theta_0)^2}{6}\frac{L}{\cos\theta_0} + \frac{\theta-\theta_0}{2}\frac{L}{\cos\theta_0}\tan\theta_0, \\
D_s^2 = D_s^2(L,\theta,\varphi=0) &\simeq \frac{\overline{\varepsilon^2}(T_0)}{(\overline{\varepsilon}(T_0))^2}\frac{L}{\cos\theta_0} + \frac{(\Delta L)^2}{\cos^2\theta_0} \\
&+ \frac{\langle \Theta_S^2(T_0)\rangle}{24}\frac{L^3}{\cos^3\theta_0}\tan^2\theta_0 \\
&+ \frac{(\langle \Theta_S^2(T_0)\rangle)^2}{360}\frac{L^4}{\cos^4\theta_0} + (\theta-\theta_0)^2 \\
&\times \frac{\langle \Theta_S^2(T_0)\rangle}{90}\frac{L^3}{\cos^3\theta_0} + (\theta-\theta_0) \\
&\times \frac{\langle \Theta_S^2(T_0)\rangle}{24}\frac{L^3}{\cos^3\theta_0}\tan\theta_0.
\end{aligned} \tag{30}$$

In the case of thicker scatterers, on the other hand, it is required to know the energy dependences of the quantities $\overline{\varepsilon}(T)$, $\overline{\varepsilon^2}(T)$, and $\langle \Theta_S^2(T)\rangle$. For example, in the case of fast ($v \gg v_{at}$, v is the particle velocity, v_{at} is a characteristic velocity of electrons in the target atoms, $T \gg 2.5\cdot 10^{-5} Z^{4/3} mc^2$, and c is the speed of light), heavy, charged particles in the nonrelativistic energy range $T \ll mc^2$ these dependences have the form[9]

$$\overline{\varepsilon}(T) = 0.3\rho\frac{Z}{A}z^2\frac{m_e mc^4}{T}L_{ion}(T),$$

$$\overline{\varepsilon^2}(T) = 0.6\rho\frac{Z}{A}z^2(m_e c^2)^2 K(T),$$

$$\langle \Theta_S^2(T)\rangle = 0.6\rho \frac{Z(Z+1)}{A} z^2 \frac{(m_e c^2)^2}{T^2} L_Q(T), \tag{31}$$

where L_{ion} is the ionization logarithm, K is a coefficient characterizing the binding of electrons in an atom, L_Q is the Coulomb logarithm, ρ is the material density of the target, g/cm^3, z is the charge of the incident particles in units of the proton charge, and A is the atomic weight of the target atoms.

Using relations (31) and disregarding the energy dependence of L_{ion} and K, we find that $T = T_0\sqrt{1-s/R_0}$ for fast nonrelativistic particles. Consequently, using Eqs. (26) and (27), we obtain

$$\Delta T_{mp} = T_0(1-\sqrt{1-\langle s\rangle/R_0}),$$

$$\Delta T_{1/2} \simeq 2\sqrt{2\ln 2}\,\overline{\varepsilon}(T_0) D_s/\sqrt{1-\langle s\rangle/R_0}, \tag{32}$$

where $R_0 = 0.5T_0/\overline{\varepsilon}(T_0)$ is the total range of the particles in the medium.

For a detector that records particles, irrespective of their exit angle, the quantities $\langle s\rangle$ and D_s have the form

$$\langle s\rangle = \langle s\rangle_L \simeq \frac{L}{\cos\theta_0} + \frac{1}{4}\gamma R_0 f_1\left(\frac{L}{R_0\cos\theta_0}\right),$$

$$D_s^2 = D_s^2(L) \simeq 2vR_0\frac{L}{\cos\theta_0}\left(1-\frac{1}{2}\frac{L}{R_0\cos\theta_0}\right) + \frac{(\Delta L)^2}{\cos^2\theta_0} + \frac{1}{8}\gamma R_0^2 f_2\left(\frac{L}{R_0\cos\theta_0}\right)\tan^2\theta_0 + \frac{1}{32}\gamma^2 R_0^2 f_3\left(\frac{L}{R_0\cos\theta_0}\right), \tag{33}$$

where

$$\gamma = \frac{\langle \Theta_S^2(T_0)\rangle T_0}{\overline{\varepsilon}(T_0)} \simeq 2(Z+1)\frac{m_e}{m}\frac{L_Q}{L_{\text{ion}}},$$

$$v = \frac{\overline{\varepsilon^2}(T_0)}{T_0\overline{\varepsilon}(T_0)} \simeq 2\frac{m_e}{m}\frac{K}{L_{\text{ion}}},$$

$$f_1(\xi) = \xi + (1-\xi)\ln(1-\xi),$$

$$f_2(\xi) = 3\xi^2 - 2\xi - 2(1-\xi)^2\ln(1-\xi),$$

$$f_3(\xi) = 7\xi^2 - 6\xi - 6(1-\xi)^2\ln(1-\xi) + 2(1-\xi)^2\ln^2(1-\xi).$$

In the case of fixed angles of observation we obtain the following expressions for $\langle s\rangle$ and D_s:

$$\langle s\rangle = \langle s\rangle_{L,\theta,\varphi=0} \simeq \frac{L}{\cos\theta_0} + \frac{1}{4}\gamma R_0 f_4\left(\frac{L}{R_0\cos\theta_0}\right) + \frac{1}{2}(\theta-\theta_0)^2 R_0 f_5\left(\frac{L}{R_0\cos\theta_0}\right) + (\theta-\theta_0)R_0 f_6 \times\left(\frac{L}{R_0\cos\theta_0}\right)\tan\theta_0,$$

$$D_s^2 = D_s^2(L,\theta,\varphi=0) \simeq 2vR_0\frac{L}{\cos\theta_0}\left(1-\frac{1}{2}\frac{L}{R_0\cos\theta_0}\right) + \frac{(\Delta L)^2}{\cos^2\theta_0} + \frac{1}{8}\gamma R_0^2 f_7\left(\frac{L}{\cos\theta_0}\right)\times\tan^2\theta_0 + \frac{1}{32}\gamma^2 R_0^2 f_8\left(\frac{L}{R_0\cos\theta_0}\right) + \frac{(\theta-\theta_0)^2}{16}\gamma R_0^2 f_9\left(\frac{L}{R_0\cos\theta_0}\right) + \frac{(\theta-\theta_0)}{8}\gamma R_0^2 f_{10}\left(\frac{L}{R_0\cos\theta_0}\right)\tan\theta_0, \tag{34}$$

where

$$f_4(\xi) = 2-\xi+\frac{2\xi}{\ln(1-\xi)},$$

$$f_5(\xi) = \xi - 1 + 2\frac{1-\xi}{\ln(1-\xi)} + \frac{2\xi}{\ln^2(1-\xi)},$$

$$f_6(\xi) = \xi - 1 - \frac{\xi}{\ln(1-\xi)},$$

$$f_7(\xi) = 2\xi - \xi^2 + \frac{2\xi^2}{\ln(1-\xi)},$$

$$f_8(\xi) = \xi^2 - 2\xi + 2 + \frac{5\xi(2-\xi)}{\ln(1-\xi)} + \frac{8\xi^2}{\ln^2(1-\xi)},$$

$$f_9(\xi) = -2(1-\xi)^2 + \frac{10(1-\xi)^2}{\ln(1-\xi)} - \frac{\xi(21\xi-26)}{\ln^2(1-\xi)} + \frac{16\xi^2}{\ln^3(1-\xi)},$$

$$f_{10}(\xi) = -2(1-\xi)^2 - \frac{\xi(10-7\xi)}{\ln(1-\xi)} - \frac{8\xi^2}{\ln^2(1-\xi)}.$$

The parameters γ and v introduced in (33) characterize the fluctuations of the ranges traversed by particles in the medium due to multiple elastic scattering ($\gamma \sim \langle\Theta_S^2\rangle$) and the fluctuations of the energy losses due to the probabilistic nature of the inelastic collision ($v \sim \overline{\varepsilon^2}$), respectively. The smallness of these parameters ($\gamma, v \sim m_e/m$) for heavy particles accounts for the small fluctuations.

It is clear from Eqs. (33) and (34) that the contributions to the half-width of the energy spectrum from fluctuations associated with different factors depend differently on the angle of incidence of the particles on the target θ_0. This fact can be utilized to determine the degree of influence of various factors on the formation of the transmitted-particle energy spectrum from the dependence of $\Delta T_{1/2}(T_0,\theta_0)$ on the angle of incidence θ_0. We note that this method can take on special importance in cases where such differentiation cannot be made from the energy dependence of $\Delta T_{1/2}(T_0,\theta_0)$.

EXPERIMENTAL DATA AND THEIR ANALYSIS

Experiments on the transmission of protons with energies of 10–25 keV through a free-standing Formvar (polyvi-

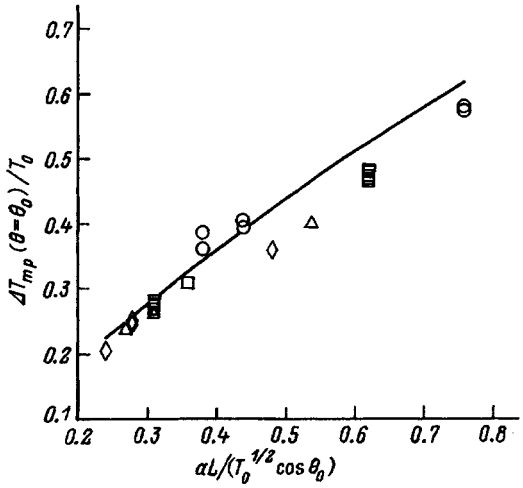

FIG. 2. Most probable energy losses of a broad proton beam in transmission through a free-standing Formvar film. Angles of incidence of the beam on the target $\theta_0=0°$, 30°, 60°. Energy: ○ — 10 keV; □ — 15 keV; △ — 20 keV; ◇ — 25 keV.

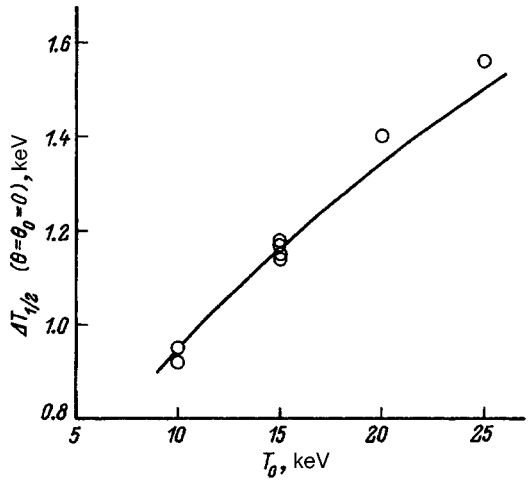

FIG. 4. Width of the energy spectrum of a broad beam of protons transmitted through a free-standing Formvar film versus initial energy T_0.

nyl formaldehyde plastic) film have been carried out on the Moscow Engineering-Physics Institute (MIFI) mass monochromator.[19]

A beam of protons with less than 0.4° angular spread was incident on the film, which was mounted on a goniometer to permit rotation of the film in two mutually perpendicular planes. The particles transmitted through the film were recorded by means of an automated energy analyzer for ions and neutral atoms. This device could be rotated relative to the target and incorporated an energy-integrating detector with an adjustable sensitivity. The input aperture of the energy analyzer was $6.8\cdot10^{-4}$ sr, and it had a resolution of 0.006 or better.

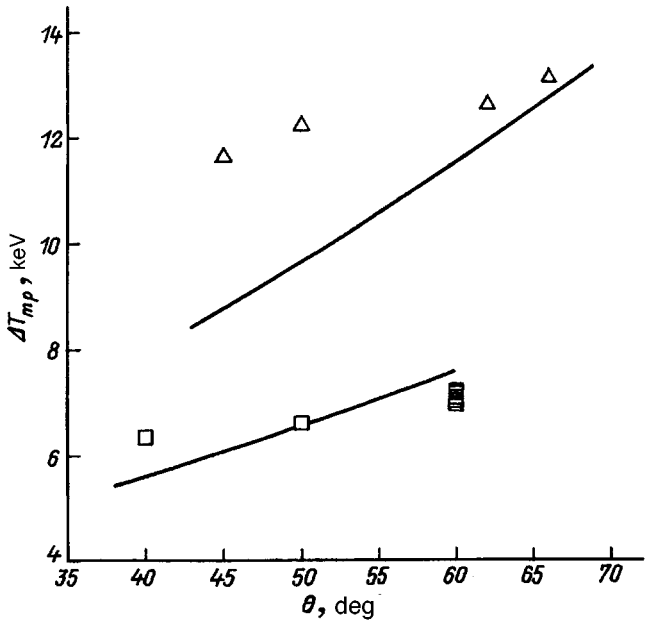

FIG. 3. Most probable energy loss of a broad beam of protons transmitted through a free-standing Formvar film versus angle of observation. Experimental values: □ — $T_0=15$ keV, $\theta_0=60°$; △ — $T_0=20$ keV, $\theta_0=70°$. The solid curves are calculated from Eq. (37).

The charging of the surface of the insulating Formvar film was monitored by an electron gun with its beam directed parallel to the surface and onto a luminophor. The film surface potential could be monitored from the shifting of the spot on the luminophor. The input currents of the ion gun were set lower than 10^{-9} A; this measure, in conjunction with beam scanning, made it possible to prevent charging of the film surface. The feedback potential of the film was additionally monitored from the position of the energy spectra of the positive and negative charge fraction of hydrogen ions transmitted through the foil.

The experiments were carried out for angles of incidence of the beam on the target $\theta_0=0°$, 30°, 60°, 70°. The energy spectra of the transmitted particles were measured in the plane of incidence ($\varphi=\varphi_0$) for various polar angles θ. The experimental results are shown in Figs. 2–5.

It is evident from the experimental data that the width of the energy spectra is small in comparison with the most probable energy loss, $\Delta T_{1/2}\ll\Delta T_{mp}$. This results indicates that fluctuations are small in the observed situation. Consequently, the results of our previous theoretical analysis can be used to analyze the experimental data.

The stopping power of the medium for protons with energies of the order of 10 keV can be expressed in the form $\overline{\varepsilon(T)}=aT^{\alpha}$. It follows from the theoretical treatment that $\alpha=1/2$ (Refs. 4, 5, and 20), and it follows from the approximation of the experimental data that $\alpha\simeq0.45$ (Ref. 6). Since these values are close, we shall assume for simplicity that the stopping power of the medium for 10-keV protons is

$$\overline{\varepsilon(T)}=a\sqrt{T}. \tag{35}$$

Making use of the smallness of the fluctuations, we can discard the term $\sim\langle\Theta_S^2\rangle$ in Eqs. (22) and (24). From Eqs. (26) and (27) we then obtain the following equation for the most probable energy loss when the axis of the detector coincides with the direction of particle incidence ($\theta=\theta_0$, $\varphi=\varphi_0$):

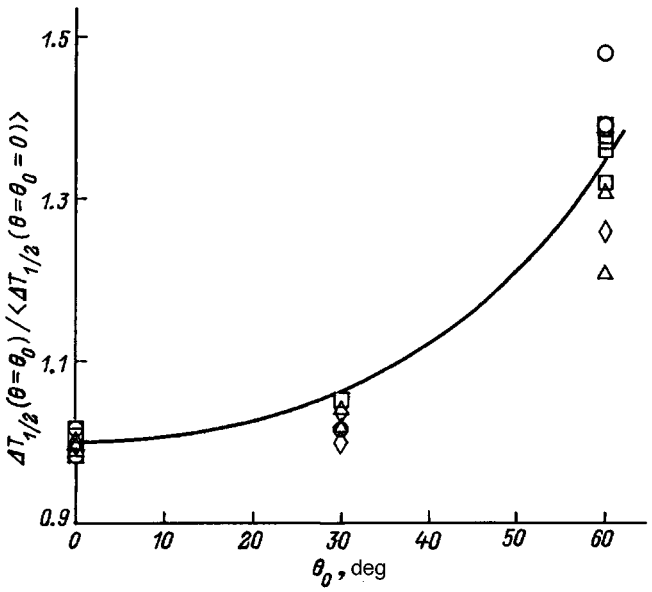

FIG. 5. Width of the energy spectrum of a broad proton beam versus angle of incidence of the beam on a free-standing Formvar film. The solid curve approximates the experimental data by the function $\cos^{-\beta}\theta_0$ with $\beta=0.43$. Energy: ○ — 10 keV; □ — 15 keV; △ — 20 keV; ◇ — 25 keV.

$$\frac{\Delta T_{mp}(\theta=\theta_0,\varphi=0)}{T_0}\simeq\frac{aL}{\sqrt{T_0}\cos\theta_0}\left(1-\frac{aL}{4\sqrt{T_0}\cos\theta_0}\right). \tag{36}$$

We note that the dependence of $\Delta T_{mp}/T_0$ on the energy of the incident particles T_0 and the angle of incidence θ_0 is characterized by their combination $\sqrt{T_0}\cos\theta_0$.

In Fig. 2 the dependence (36) is compared with the experimental data. The value of $aL=1.2$ was obtained by least-squares approximation of the experimental data for normal incidence of particles on the medium ($\theta_0=0°$) by the dependence (36). Clearly, the theoretical curve agrees quite well with the experimental data. The deviation of the curve (36) from the experimental data for a beam angle of incidence on the target $\theta_0=60°$ can be attributed both to the imprecision of the dependence (35) and to the fact that the reflection of particles from the target can no longer be ignored for protons of energy $T_0<15$ keV at such angles of incidence of the beam on the target.[18]

Unfortunately, reliable theoretical or experimental data on the dependence of $\overline{\varepsilon^2}(T)$ and $\langle\Theta_S^2(T)\rangle$ on the energy T for kiloelectron-volt protons are not found in the literature. We therefore use relations (28) and (30) for further analysis of the experimental data. This is especially justifiable in that the quantity $L/R_0\cos\theta_0$ is small in the given range of initial energies T_0 and angles of incidence θ_0 [$L/(R_0\cos\theta_0)\simeq aL/(2\sqrt{T_0}\cos\theta_0)<0.4$]; i.e., the case of a relatively thin scatterer is established.

In this approximation the dependence of the most probable energy loss on the angle of observation θ for the angle of incidence θ_0 assumes the form

$$\Delta T_{mp}(\theta,\theta_0)\simeq\frac{aL\sqrt{T_0}}{\cos\theta_0}\left[1+\frac{1}{6}(\theta-\theta_0)^2+\frac{1}{2}(\theta-\theta_0)\tan\theta_0\right]. \tag{37}$$

Note that for $\theta=\theta_0$ Eq. (37) transforms to Eq. (36) if the second term in the brackets, which takes into account the dependence of $\overline{\varepsilon}(T)$ on the energy T (35), is disregarded in (36).

The structure of Eq. (37) is very similar to that obtained when the quantity $\Delta T_{mp}\simeq\overline{\varepsilon}(T_0)\langle s\rangle$ is estimated on the rather crude assumption that particles exiting from the target at the angle θ have traversed an average range $L/\cos\theta$ in it ($\langle s\rangle\simeq L/\cos\theta$). Here, expanding $\cos\theta$ in the vicinity of θ_0, we obtain $\Delta T_{mp}\simeq\overline{\varepsilon}(T_0)L[1+(-\theta_0)^2/2+(\theta-\theta_0)\tan\theta_0]/\cos\theta_0$. Comparing this equation with (37), we see that despite the identical structure of these expressions, they differ considerably in the quantitative sense because of the total difference in their numerical coefficients. We emphasize that the numerical coefficients in (37) are obtained from the general equation (24) on the assumption that the scatterer is thin ($L/\cos\theta_0\ll R_0$). In general, however, these coefficients are rather complex functions of the quantity $\xi=L/(R_0\cos\theta_0)$. In the case of fast nonrelativistic particles, for example, it follows from (34) that the coefficient of $(\theta-\theta_0)^2$ is determined by the function $f_5(\xi)$, and the coefficient of $(\theta-\theta_0)$ is determined by the function $f_6(\xi)$.

The dependence (37) is compared with the experimental data in Fig. 3. We see that the experimental values in the range of observation angles $\theta_0-\theta>20°$ greatly exceed the theoretical results. From all appearances this discrepancy is attributable to our use of the diffusion approximation with respect to the angles (3) for the elastic collision integral. The approximation well describes the propagation of particles in the range of relatively small scattering angles, in which by far most of the particles are concentrated, but it lowers the results drastically in the range of relatively large scattering angles.[21] It has been shown[21] that scattering at relatively large angles is achieved most efficiently through single scattering. In this case the correlation between the scattering angle of a particle and the energy lost by it in a single elastic scattering event must be taken into account. Consequently, the cause of the departure of (37) from the experimental data can be twofold: the application of the diffusion approximation with respect to the angles in the interval of elastic collisions (beyond the scope of this approximation, at present there is no theory to correctly account for bending of the particle trajectories by multiple elastic scattering); the disregard of correlation between the scattering angle of a particle and the energy lost by it in a single elastic collision event.

The following expression for the width of the energy spectrum of the transmitted particles is obtained from Eqs. (28) and (30) with allowance for (35) in the case where the detector axis coincides with the direction of particle incidence ($\theta=\theta_0$, $\varphi=\varphi_0$):

$$\Delta T_{1/2}(\theta=\theta_0,\varphi=0)\simeq 2\sqrt{2\ln 2}a\sqrt{T_0}$$

$$\times\left\{\frac{(\Delta L)^2}{\cos^2\theta_0}+\frac{\overline{\varepsilon^2}(T_0)}{(\overline{\varepsilon}(T_0))^2}\frac{L}{\cos\theta_0}\right.$$

$$+\frac{\langle\Theta_S^2(T_0)\rangle}{24}\frac{L^3}{\cos^3\theta_0}\tan^2\theta_0$$

$$\left.+\frac{(\langle\Theta_S^2(T_0)\rangle)^2}{360}\frac{L^4}{\cos^4\theta_0}\right\}^{1/2}. \qquad (38)$$

Figure 4 shows experimental values of $\Delta T_{1/2}$ for the case of particles in normal incidence on a target ($\theta_0=0$). The dependence $\Delta T_{1/2}\sim\sqrt{T_0}$ is also plotted in this figure (solid curve). We see that the curve $\Delta T_{1/2}\sim\sqrt{T_0}$ agrees quite well with the experimental data. At first glance, their agreement suggests that the principal factor in establishing the width of the particle energy spectrum should be fluctuations of the film thickness [the first term under the square root sign in Eq. (38), which does not depend on the energy T_0]. However, since we do not know how $\overline{\varepsilon^2}(T_0)$ and $\langle\Theta_S^2(_0)\rangle$ depend on the energy, we are not prone to rush into any conclusions. For example, it could turn out that the ratio $\overline{\varepsilon^2}(T_0)/[\overline{\varepsilon}(T_0)]^2$ is also independent of T_0. We therefore analyze the dependence of $\Delta T_{1/2}$ on the angle of incidence θ_0.

Figure 5 shows experimental values of $\Delta T_{1/2}(\theta=\theta_0,\varphi=0)/\langle\Delta T_{1/2}(\theta=\theta_0=0,\varphi=0)\rangle$; the angle brackets $\langle\cdots\rangle$ indicate the average experimental value at a fixed initial energy T_0. Also shown in this figure is an approximation of the experimental data by the curve $1/\cos^\beta\theta_0$, where the exponent $\beta\simeq 0.43$ is evaluated by the least-squares method. The value of β is close to 1/2. We conclude from this result on the basis of (38) that the main contribution to the formation of $\Delta T_{1/2}$ is from fluctuations of the energy losses in a single inelastic scattering event [the second term under the square root sign in (38)]. The slight deviation of β from 1/2 can be attributed to neglect of the energy dependence of $\overline{\varepsilon}(T)$ and $\overline{\varepsilon^2}(T)$ in (38).

The final conclusion of the foregoing analysis is that the influence of fluctuations of the ranges of the transmitted particles due to multiple elastic scattering and the influence of fluctuations of the film thickness on the width of the transmitted-particle spectrum are slight in comparison with the influence of energy loss fluctuations in inelastic collisions. Whereas this statement could have been made earlier about the influence of fluctuations of the particle ranges,[8,9] since the effective value of Z is small for a Formvar film ($Z_{\text{eff}}\sim 7$), the same could not be said about the influence of fluctuations of the film thickness without first analyzing the dependence of $\Delta T_{1/2}$ on the angle of incidence θ_0. On the other hand, an analysis of the energy dependence of $\Delta T_{1/2}$ for particles in normal incidence ($\theta_0=0$) forced us to conclude that the fluctuations of the film thickness are large. Now, however, we see that the probabilistic nature of inelastic collisions is the dominant factor in establishing the width of the transmitted-particle energy spectrum under the given experimental conditions.

CONCLUSIONS

We have carried out a theoretical and experimental study of the characteristics of the energy spectra of kiloelectron-volt protons transmitted through a free-standing foil and their dependence on the angle of incidence of the beam on the target.

We have obtained analytical expressions for the average characteristics of the energy spectrum of transmitted particles or ions in the case where the angle of observation θ deviates relatively little from the angle of incidence θ_0, i.e., when the scattering angle is small. We have taken into account the combined influence of various factors affecting the formation of the energy spectra: systematic stopping of the particles in the medium, fluctuations of the particle energy losses in inelastic collisions, bending of the particle trajectories due to multiple elastic scattering, and fluctuations of the target thickness. We have shown that the contributions of these factors to the width of the transmitted-particle energy spectrum exhibit different dependences on the angle of incidence of the beam on the target surface. This fact can be exploited to determine from the experimental dependence of the width of the energy spectra on the angle of incidence θ_0 which of these factors are predominant in forming the energy spectra of the particles. This capability is particularly important in cases where the energy dependence of the scattering and stopping properties of the medium (for example, in application to medium-energy ions) is not known.

The difference in the dependences of the terms in the expression for the width of the energy spectrum on the angle of incidence θ_0 in fact provides a means for solving the inverse problem of transport theory: to extract information about elementary elastic and inelastic collision processes from experimental data on the transmission of particles through the medium.

The authors are grateful to D. B. Rogozkin for interest in the study and for numerous discussions of related problems.

APPENDIX. TRANSMISSION OF A BROAD PARTICLE BEAM THROUGH A THIN, HOMOGENEOUS TARGET WHOSE THICKNESS IS A RANDOM FUNCTION OF POSITION ON ITS SURFACE; THE FORWARD-BACKWARD APPROXIMATION

Here we consider the transmission of a broad particle beam through a homogeneous target, whose surfaces are not plane but have a random relief.

The surfaces of the target are parallel, on the average (Fig. 6). The target can be called planar in this sense. We direct the z axis downward and perpendicular to the target (i.e., perpendicular to the average surfaces of the target); the top and bottom surfaces of the target can then be described by random functions $f_1(x,y)$ and $f_2(x,y)$, respectively. We describe the direction of motion of the incident particles by the polar angle θ_0 and the azimuthal angle φ_0. We choose the x axis in such a way that the velocity vector of the incident particles is situated in the zx plane, making the azimuth angle of incidence equal to zero ($\varphi_0=0$).

A broad particle beam can be visualized as a superposition of narrow beams. We therefore begin with the transmis-

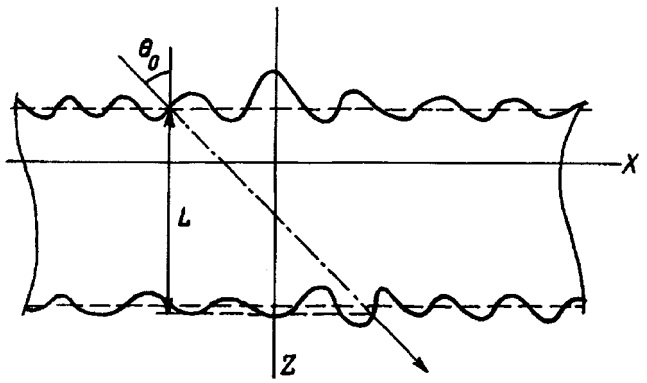

FIG. 6. Conceptual diagram of the transmission of a narrow particle beam through a target whose surfaces have a random relief.

sion of a narrow particle beam through such a target. In the so-called forward-backward approximation the directions of motion of the particles do not change, and the transport equation for the particle flux density in a narrow beam $N(z,x,y,T|x_0,y_0,T_0)$ has the form[7]

$$\cos\ \theta_0 \frac{\partial N(z,x,y,T|x_0,y_0,T_0)}{\partial z} + \sin\ \theta_0 \frac{\partial N}{\partial x} = \hat{I}_{\text{in}}[N]\,\eta(f_2(x,y)-z). \tag{A1}$$

Here $\hat{I}_{\text{in}}[N]$ is the inelastic collision integral. Equation (A1) is written for zero azimuth angle ($\varphi_0=0$), and the function

$$\eta(\xi)=\begin{cases}1 & \text{for } \xi>0,\\ 0 & \text{for } \xi<0\end{cases}$$

takes into account the absence of scattering outside the target. The boundary condition for Eq. (A1) has the form

$$N(z=f_1(x_0,y_0),x,y,T|x_0,y_0,T_0) = N_0\delta(T_0-T)\,\delta(x_0-x)\,\delta(y_0-y). \tag{A2}$$

It characterizes the fact that a narrow particle beam with an initial energy T_0 is incident on the top surface of the target at the point (x_0,y_0).

Since the particles move along a straight line in the forward-backward approximation, their coordinates at any depth z are described by the equations $x=x_0+[z-f_1(x_0,y_0)]\tan\ \theta_0$, $y=y_0$. By the same token, if we take into account the fact that scattering does not occur at $z>f_2(x_0,y_0)$, we can write the particle flux density $N(z,x,y,T|x_0,y_0,T_0)$ in the form

$$N(z,x,y,T|x_0,y_0,T_0) = \tilde{N}\left(\int_{f_1(x_0,y_0)}^{z} dz'\,\eta\{f_2(x_0+(z'-f_1(x_0,y_0))\tan\ \theta_0,y_0) - z'\};T|T_0\right)\delta(x-x_0-(z-f_1(x_0,y_0))\tan\ \theta_0)\,\delta(y_0-y), \tag{A3}$$

where $\tilde{N}(z,T|T_0)$ is the particle flux density in a broad beam propagating into a target with plane boundaries.

Now in fact, substituting Eq. (A3) into (A1) and (A2), we obtain

$$\cos\ \theta_0 \frac{\partial \tilde{N}}{\partial z} = \hat{I}_{\text{in}}[\tilde{N}], \quad 0<z<L,$$

$$\tilde{N}(z=0,T|T_0)=\delta(T-T_0). \tag{A4}$$

Here L is the target thickness, which satisfies the equation

$$L=f_2(x_0+(L-f_1(x_0,y_0))\tan\ \theta_0,y_0)-f_1(x_0,y_0). \tag{A5}$$

We see that the target thickness $L(x_0,y_0)$ is a random function of position of the point of particle entry (x_0,y_0). Consequently, the flux density of particles transmitted through the target in the case of a narrow incident beam has the form

$$\begin{aligned} N_{\text{tr}}(x,y,T|x_0,y_0,T_0) &= N(z>f_2(x,y);x,y,T|x_0,y_0,T_0)\\ &= \tilde{N}(L(x_0,y_0),\ T|T_0)\,\delta(x-x_0)\\ &\quad -(z-f_1(x_0,y_0))\tan\ \theta_0)\,\delta(y_0-y). \end{aligned} \tag{A6}$$

If the detector records all particles transmitted through the target, i.e., is wide-angle, we can average the distribution (A6) over the coordinates x, y and obtain

$$N_{\text{tr}}(T|x_0,y_0,T_0)=\int \frac{dxdy}{\Sigma} N_{\text{tr}}(x,y,T|x_0,y_0,T_0) = \frac{1}{\Sigma}\tilde{N}(L(x_0,y_0),T|T_0). \tag{A7}$$

Here Σ is the area of the average target surface. Making the transition from a narrow particle beam to a broad beam, we now have

$$\begin{aligned} N_{\text{tr}}(T|T_0) &= \int dx_0dy_0 N_{\text{tr}}(T|x_0,y_0,T_0)\\ &= \int \frac{dx_0dy_0}{\Sigma}\tilde{N}(L(x_0,y_0),\ T|T_0)\\ &= \int dz\tilde{N}(z,T|T_0)\int \frac{dx_0dy_0}{\Sigma}\,\delta(z-L(x_0,y_0)). \end{aligned} \tag{A8}$$

If we introduce the probability density function $f(z)$ of a target having the thickness z, we can use the relation

$$f(z)=\int \frac{dx_0dy_0}{\Sigma}\,\delta(z-L(x_0,y_0)) \tag{A9}$$

to establish the fact that the flux density of particles transmitted through a target with a random relief, instead of planar boundaries, can be determined from the equation

$$N_{\text{tr}}(T|T_0)=\int dz f(z)\tilde{N}(z,T|T_0). \tag{A10}$$

We now show that the function $f(z)$ defined by relation (A9) is the probability density function of the target having

the thickness z. Integrating Eq. (A9) with respect to z, we find that the function $f(z)$ is normalized to unity:

$$\int dz f(z) = \int dz \int \frac{dx_0 dy_0}{\Sigma} \delta(z - L(x_0, y_0))$$

$$= \int \frac{dx_0 dy_0}{\Sigma} = 1.$$

Multiplying Eq. (A9) by z and integrating with respect to z, we obtain

$$\int z dz f(z) = \int \frac{dx_0 dy_0}{\Sigma} L(x_0, y_0)$$

$$= \int \frac{dx_0 dy_0}{\Sigma} (f_2(x_0, y_0) - f_1(x_0, y_0))$$

$$= \langle f_2 \rangle - \langle f_1 \rangle.$$

We see that $f(z)$ is indeed the probability density function.

[1] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, Pergamon Press, New York–Oxford (1974) [Russ. transl., Nauka, Moscow (1974).
[2] B. B. Rossi, *High-Energy Particles*, Prentice-Hall, New York (1952) [Russ. transl., Gostekhizdat, Moscow (1955)].
[3] S. V. Maleev and B. P. Taperverg, Zh. Éksp. Teor. Fiz. **78**, 315 (1980) [Sov. Phys. JETP **51**, 158 (1980)].
[4] O. B. Firsov, Zh. Éksp. Teor. Fiz. **36**, 1517 (1959) [Sov. Phys. JETP **9**, 1076 (1959)].
[5] J. Lindhard and M. Scharff, Phys. Rev. **124**, 128 (1961).
[6] H. H. Andersen and J. F. Ziegler, *Hydrogen Stopping Powers and Ranges in All Elements*, Academic Press, New York (1977).
[7] N. P. Kalashnikov, V. S. Remizovich, and M. I. Ryazanov, *Collisions of Fast Heavy Particles in Solids* [in Russian], Atomizdat, Moscow (1980).
[8] V. S. Remizovich, D. B. Rogozkin, and M. I. Ryazano, Élem. Chastitsy At. Yadra (Dubna) **17**, 929 (1986).
[9] V. S. Remizovich, D. B. Rogozkin, and M. I. Ryazanov, *Fluctuations of the Ranges of Charged Particles* [in Russian], Énergoatomizdat, Moscow (1988).
[10] R. Iohiwari, N. Shiomi, and N. Sakamoto, Phys. Rev. A **25**, 2524 (1982).
[11] A. Gras-Marti, Nucl. Instrum. Methods. Phys. Res. B **9**, 1 (1985).
[12] V. S. Remizovich, D. B. Rogozkin, and V. V. Frolov, At. Énerg. **61**, 256 (1986).
[13] A. I. Kuzovlev and V. S. Remizovich, Phys. Rev. A **48**, 465 (1993).
[14] V. A. Kurnaev, Zh. Tekh. Fiz. **46**, 627 (1976) [Sov. Phys. Tech. Phys. **21**, 363 (1976)].
[15] A. I. Kuzovlev and V. S. Remizovich, Dokl. Akad. Nauk SSSR **266**, 1118 (1982) [Sov. Phys. Dokl. **27**, 856 (1982)].
[16] A. I. Kuzovlev and V. S. Remizovich, Pis'ma Zh. Tekh. Fiz. **9**, 710 (1983) [Sov. Tech. Phys. Lett. **9**, 305 (1986)].
[17] V. A. Kurnaev, E S. Mashkova, and V. A. Malchanov, *Reflection of Light Ions from the Surface of a Solid* [in Russian], Énergoatomizdat, Moscow (1985).
[18] N. N. Koborov, A. I. Kuzovlev, V. A. Kurnaev, and V. S. Remizovich, Pis'ma Zh. Tekh. Fiz. **17**(20), 6 (1991) [Sov. Tech. Phys. Lett. **17**, 718 (1991)].
[19] N. N. Koborov, V. A. Kurnaev, and V. A. Urusov, *Interaction of Ions and Plasma with the Surface of a Solid* [in Russian], Énergoatomizdat, Moscow (1985).
[20] Yu. V. Gott, *Interaction of Particles with Matter in Plasma Studies* [in Russian], Atomizdat, Moscow (1978).
[21] A. S. Kompaneets, Zh. Éksp. Teor. Fiz. **15**, 236 (1948).

Translated by James S. Wood

Beam-plasma discharge in the propagation of a long-pulse relativistic electron beam in a medium-pressure rarefied gas

M. V. Gladyshev and M. G. Nikulin

Radio Engineering Institute, Russian Academy of Sciences, 113519 Moscow, Russia
(Submitted December 27, 1995)
Zh. Tekh. Fiz. **67**, 94–98 (May 1997)

A theoretical model is given, along with a numerical analysis of the evolution of beam-plasma discharge in the propagation of a long-pulse relativistic electron beam in a rarefied gas at medium pressure. It is shown that the self-stabilization of beam-plasma discharge as a result of longitudinal inhomogeneity of the density of the discharge plasma makes it possible for the beam to traverse the beam chamber with relatively low total energy losses, including ionization losses and energy losses in the generation of oscillations. During the dissociative recombination of electrons and ions of the discharge-driven plasma, heat is released and spent in raising the temperature of the gas. The investigated collective-discharge mechanism underlying heating of the gas for a relativistic beam can be more efficient than the classical heating mechanism due to ionization losses of the beam in pair collisions of its electrons with gas particles. [S1063-7842(97)01405-0]

INTRODUCTION

Numerous experiments (see, e.g., Refs. 1 and 2) on the transport of relativistic electron beams with typical parameters in various gases in the self-focusing regime at a distance L of the order of a few betatron lengths, i.e., $L\sim 1$ m, have shown that the most efficient transport can be achieved in a medium pressure range $p\sim 1$ Torr. The propagation of relativistic electron beams is impeded at lower pressures by beam-plasma instability, which leads to energy losses and beam scattering, and at higher pressures by resistive hose instability, which deflects the beam from the axis and at the worst extreme can force the beam to strike the walls of the drift chamber.

A numerical experiment[3] on a model developed for a short-pulse beam (with a duration $\tau\sim 100$ ns) has shown that an important characteristic of the propagation of relativistic electron beams under the above-stated conditions is the self-stabilization of beam-plasma instability at a level such that it does not interrupt transport but still maintains beam-plasma discharge.[4] This consideration offers certain added possibilities. In particular, at medium pressures the beam-plasma discharge can be used effectively to achieve plasma-chemical reactions,[5] and in the case of long-pulse and repetitive beams it also provides a potential gas-heating source with a higher heat-transfer rate than that attained through the direct ionization losses from the relativistic electron beam.

Here we report the results of an experimental study of collective and discharge processes associated with the interaction of a long-pulse relativistic electron beam with a slightly ionized, medium-pressure gas. In particular, we examine the evolution of beam-plasma discharge in microsecond time intervals and the rise in temperature in this process as a result of heat transfer due to the dissociative recombination of plasma electrons and ions. It has been shown[3] that by the time the discharge reaches the leading edge of the short pulse it is in a quasisteady state, which gradually evolves with time; as a prototype we use the model proposed in Ref. 3, augmenting it with the heat-conduction equation for the gas and eliminating the equations describing the inductive excitation of a reverse plasma current at the leading edge of the beam and its dissipation.

STATEMENT OF THE PROBLEM AND METHOD OF SOLUTION

We consider a relativistic electron beam with current $I\gtrsim 50$ A, energy $W\sim 1$ MeV, radius $a\sim 1$ cm, and duration $\tau\sim 100$ μs and investigate its interaction with a slightly ionized gas (air) at an initial pressure $p_0\sim 1$ Torr in a drift chamber of length $L\simeq 1$ m without an external magnetic field, using a model in which the beam is described by large particles, the rf electric field of the plasma modes is determined from the one-dimensional Poisson equation by the method of slowly varying amplitudes, and the discharge kinetics is described by balance equations for the density and average energy of plasma electrons. We disregard the reverse plasma current, since the current of long-pulse beams rises smoothly.

In using the approach outlined above, we assume that the beam has a low density in comparison with the plasma density, is relativistic and monoenergetic, and is held in equilibrium by its own magnetic field. We assume that thermal motion in the plasma and the self-induced magnetic field do not influence the dispersion of the beam-generated waves and that the radial boundedness of the beam and the plasma do not affect the growth rate of instability. Furthermore, we assume that collisions of plasma electrons with heavy particles suppress all transverse disturbances without hindering the development of longitudinal oscillations. According to Ref. 6, the following inequalities must be satisfied for this condition to be achieved:

$$\frac{6T_e}{mc^2}\ll\left(\frac{\nu_e}{\omega_e}\right)^{3/2}\frac{\omega_e a}{c}\left(\frac{I_0\gamma^3}{I}\right)^{1/2}, \tag{1}$$

 1063-7842/97/050542-05$10.00

$$\frac{I}{I_0}, \quad \left(\frac{I}{I_0}\right)^2 \ll \left(\frac{\omega_2 a}{c}\right)^2 \ll \nu_e \tau, \tag{2}$$

$$\left(\frac{c}{\omega_e}\right)^2 \ll \frac{\nu_e}{\omega_e} \ll 1, \tag{3}$$

where ν_e is the frequency of collision of plasma electrons with heavy particles, $\omega_e = (4\pi n_e e^2/m)^{1/2}$ is the Langmuir frequency of plasma electrons, $I_0 = mc^3/e$, $-e$ and m are the rest electron charge and mass, and c is the speed of light in vacuum.

We assume that the variations of all characteristics of the process averaged on these scales are insignificant within the period and wavelength of the plasma oscillations.

To describe the ionization of the gas by plasma electrons, we use the elementary theory of discharge in relatively weak microwave electric fields, assuming that ionization is induced by the heating of electrons in the fields of the oscillations as they maintain a Maxwellian velocity distribution. The amplitude E_0 of the microwave electric field in this case must be much lower than the level at which the average energy of the plasma electron oscillations becomes commensurate with their thermal energy,[7]

$$E_0(\text{kV/cm}) \ll E_c = 3.7\frac{\omega_e}{\nu_e} p(\text{Torr}). \tag{4}$$

Under the stated assumptions the system of equations of the beam-plasma discharge model has the form

$$\frac{\partial}{\partial z}\left[\frac{\partial^2}{\partial t^2} + \nu_e \frac{\partial}{\partial t} + \omega_e^2(z)\right]\frac{\partial \tilde{\psi}}{\partial z} = 4\pi e\left(\frac{\partial^2}{\partial t^2} + \nu_e \frac{\partial}{\partial t}\right)\tilde{n}_b, \tag{5}$$

$$\tilde{\psi} = \sum_{n=1}^{N} [a_n(z)\cos(k_n z - \omega_n t) + b_n(z)\sin(k_n z - \omega_n t)], \tag{6}$$

$$\tilde{n}_b = n_{b0} \sum_{i=1}^{M} \int_z^{z+\Delta z} \delta(z - \tilde{z}_i)\, dz, \tag{7}$$

$$\frac{d\tilde{z}_i}{dt} = \tilde{v}_{bi}, \tag{8}$$

$$(1 - \tilde{v}_{bi}^2/c^2)^{-3/2}\frac{d\tilde{v}_{bi}}{at} = \frac{e\partial\tilde{\psi}}{m\partial z}, \tag{9}$$

$$\frac{\partial n_e}{\partial t} = \nu_{ib} n_b + \nu_{ie} n_e - \alpha_{dr} n_e^2 + \frac{\partial}{\partial z}\left(D_a \frac{\partial n_e}{\partial z}\right) - D_a \frac{n_e}{\Lambda_e^2}, \tag{10}$$

$$\frac{\partial \varepsilon_e}{\partial t} = \left(\frac{E_0^2}{8\pi n_e} - \delta\varepsilon_e\right)\nu_e - (\varepsilon_e + \varepsilon_i)\nu_{ie} + (\varepsilon_\delta - \varepsilon_e)\frac{n_b}{n_e}\nu_{ib} + \frac{1}{n_e}\frac{\partial}{\partial z}\left(\varkappa_e \frac{\partial \varepsilon_e}{\partial z}\right), \tag{11}$$

$$E_0^2 = \sum_{n=1}^{N}\left[\frac{1}{2}k_n^2(a_n^2 + b_n^2) + k_n\left(\frac{\partial a_n}{\partial z}b_n - a_n\frac{\partial b_n}{\partial z}\right)\right], \tag{12}$$

$$\delta(\varepsilon_e) = \begin{cases} 10^{-3}, & \varepsilon_e(\text{eV}) < 1.2, \\ 7\times 10^{-4}\varepsilon_e^2(\text{eV}), & 1.2 < \varepsilon_e 9(\text{eV}) < 23.4, \\ 0.4, & 23.4 < \varepsilon_e(\text{eV}), \end{cases} \tag{13}$$

$$\nu_e(\text{s}^{-1}) = 2.9\times 10^{15}\rho(\text{g/cm}^3), \tag{14}$$

$$\nu_{ib}(\text{s}^{-1}) = 1.0\times 10^{15}(\nu_b/c)\rho(\text{g/cm}^3), \tag{15}$$

$$\nu_{ie}(\text{s}^{-1}) = 1.5\times 10^{13}\rho(\text{g/cm}^3) \times \frac{\varepsilon_e^{1/2}(\text{eV})[5.6 + 0.6\varepsilon_e(\text{eV})]}{\exp[18.75/\varepsilon_e(\text{eV})]}, \tag{16}$$

$$D_a(\text{cm}^2/\text{s}) = 3.5\times 10^{-2}\varepsilon_e(\text{eV})/T_g^{1/2}(\text{K})\rho(\text{g/cm}^3), \tag{17}$$

$$\varkappa_e(\text{eV/cm}\cdot\text{s}) = 5.9\times 10^{15} n_e(\text{cm}^{-3})\varepsilon_e(\text{eV})/\nu_e(\text{s}^{-1}). \tag{18}$$

Here the tilde identifies quantities oscillating at the wave frequency ω_n, distinguishing them from all the other quantities averaged over fast oscillations, $k_n = \omega_n/\nu_b$, M is the number of large beam particles in the system, N is the number of modes, z_i is the coordinate of the ith particle, $n_{b0} = n_b(0,t)$ is the density of the beam at the input to the system, a_n and b_n are the amplitudes, slowly varying with respect to z, of the potential $\tilde{\psi}$, ν_{ib} is the frequency of impact ionization of air by the beam, ν_{ie} is the frequency of ionization of air by the beam, averaged over the Maxwellian distribution,[8] $\varepsilon_e = (3/2)T_e$ is the average energy of the plasma electrons, δ is the fraction of energy lost by a plasma electron in collision with a heavy gas particles,[10] $\varepsilon_\delta = 21.4$ eV is the average energy of δ-electrons generated in the ionizing collision of a relativistic beam electron with an air molecule, $\varepsilon_i = 12.5$ eV is the average energy of ionization of air, D_a is the ambipolar diffusion coefficient of the plasma, $\Lambda_e = a/2.4$ is the diffusion length for the plasma density n_e (Ref. 9, p. 159), and $\varkappa_e$ is the electron thermal conductivity (Ref. 9, p. 218). We see that the coefficients (14)–(18) depend on the parameters of the gas, and the model must therefore take into account the possibility that the parameters of the gas in discharge can change in the long-pulse regime.

To describe the heating of the gas, we use a heat-conduction equation in which the source is heat released in the dissociative recombination of plasma electrons and ions. We consider two limiting cases to simplify this part of the problem. In the first case we assume that the beam occupies a small part of the drift chamber ($a/b \ll 1$, where b is the radius of the chamber). The heating of the gas in the zone occupied by the beam can be regarded as isobaric in this case, i.e., because of the low thermal conductivity of the gas, the volume of the heated zone is approximately equal to the volume occupied by the beam, and since this volume is small in comparison with that of the chamber (see the estimates below), the gas pressure remains constant. In the second extreme case we assume that the beam fills up the entire drift chamber ($a/b = 1$). Now the gas is heated isochorically, i.e., its density remains constant. Consequently, the equations describing the rise in temperature of the gas and closing the beam-plasma discharge model in the long-pulse regime have the form

$$\rho c_g \frac{\partial T_g}{\partial t} = 0.75 \times \alpha_{dr} n_e^2 (\varepsilon_i - \varepsilon_d) + \frac{\partial}{\partial z}\left(\lambda \frac{\partial T_g}{\partial z}\right) - \lambda \frac{T_g}{\Lambda_T^2}, \tag{19}$$

$$\rho(g/\mathrm{cm}^3) = 1.54 \times 10^{-6} p_0(\mathrm{Torr}) \begin{cases} 300/T_g(\mathrm{K}), & a/b \ll 1, \\ 1, & a/b = 1, \end{cases} \tag{20}$$

$$\alpha_{dr}(\mathrm{cm}^3/\mathrm{s}) = 1.0 \times 10^{-4}/\varepsilon_e^{1/3}(\mathrm{eV})[9 + \varepsilon_e(\mathrm{eV})] T_g(\mathrm{K}), \tag{21}$$

$$\lambda\left(\frac{\mathrm{W}}{\mathrm{K}\cdot\mathrm{cm}}\right) = 1.14 \times 10^{-5} T_g^{1/2}(\mathrm{K}) \times \begin{cases} [1.25 + 7.5 \times 10^{-4} T_g(\mathrm{K})], \\ 300 < T_g(\mathrm{K}) < 10^3, \\ [2 + 2.5(10^{-3} T_g(\mathrm{K}) - 1)^2], \\ 10^3 < T_g(\mathrm{K}) < 2.5 \times 10^3, \\ [10.77 - 35(10^{-3} T_g(\mathrm{K}) - 2.8)^2], \\ 2.5 \times 10^3 < T_g(\mathrm{K}) < 3.2 \times 10^3. \end{cases} \tag{22}$$

Here the specific heat of the gas $c_g = 1.0$ (J/g·K)$=1.0$ for $a/b \ll 1$ and 0.71 J/g·K for $a/b = 1$, α_{dr} is the dissociative recombination coefficient, which is expressed by an approximation of the experimental curve in Ref. 9 (p. 139), $\varepsilon_d = 9.76$ eV is the energy of dissociation, λ is the thermal conductivity of the gas,[11], and $\Delta_T = a/2.4$ is the thermal diffusion length for the gas, which is analogous to the diffusion length Λ_e. The numerical factor in the first term on the right side of Eq. (19) characterizes the efficiency of heating of the gas by dissociative recombination and is calculated from the results of Ref. 12.

We note that for typical parameters of the problem [$\varepsilon_e = (3/2)T_e \simeq 3$ eV, $T_g \simeq 1000$ K, $p_0 \simeq 1$ Torr] the characteristic times of ambipolar diffusion of plasma electrons $\tau_e \sim \Delta_e^2/D_a$ and gas cooling $\tau_T \sim \rho c_g \Lambda_T^2/\lambda$ in the isobaric regime ($a/b \ll 1$) for $a = 1$ cm, according to the above expressions, are approximately equal to 40 μs and 120 μs, respectively, which are comparable with the beam duration $\tau = 50$ μs. Consequently, as assumed above, the volume occupied by the plasma and the heated gas is approximately equal to the volume occupied by the beam. Transport processes accelerate as the temperature of the gas increases and, of course, this fact limits the applicability of the model in the given regime. Nonetheless, our extreme simplification of the gasdynamic part of the problem does not detract from the possibility of finding what we want to know: the mechanism of energy transfer from the electron beam to the gas through collective and discharge processes.

In the proposed long-pulse, beam-plasma discharge model Eqs. (5)–(9) describe beam-plasma interaction in the one-dimensional approximation under the conditions of three-dimensional development of a one-dimensional dissipative instability. Equation (10) models the replenishment of electrons through impact ionization of the gas by the beam and thermal ionization by plasma electrons, along with the depletion of electrons by dissociative recombination and ambipolar diffusion. Equation (11) describes the heating of plasma electrons by Joule heat dissipation of rf Langmuir fields, the loss of energy in collisions of electrons with heavy gas particles, the change of energy in ionization of the gas by the beam and by plasma electrons, and energy transfer during electronic heat conduction. Equation (19) describes the heating of the gas in dissociative recombination of plasma electrons and ions and its cooling by heat conduction.

A comparison of the first two terms on the right side of Eq. (10) shows that the replenishment of the plasma in beam-plasma discharge in the investigated situation for $T_e \gtrsim 2$ eV and $n_e \gtrsim 10^3 n_b$ takes place more rapidly than direct ionization of the gas by the relativistic electron beam. Also, since the main source of heat transfer for both ionization channels, according to (19), is dissociative recombination, which is proportional to n_e^2, the collective-discharge mechanism of gas heating under the stated conditions becomes more efficient than the classical mechanism of heating due to ionization losses of the relativistic electron beam in pair collisions of its electrons with gas particles.

The given system of equations has been reduced to dimensionless form and solved numerically on a personal computer. The first step was to solve Eqs. (5)–(9) for the specified plasma parameters. The amplitudes a_n and b_n were solved by a first-order implicit scheme. The equations of motion of the beam electrons were calculated by a first-order Eulerian technique with refinement. The resulting distribution of the microwave field was then used to solve Eqs. (10)–(12) and (19) with the coefficients (13)–(18) and (20)–(22).

The time interval for recomputing the amplitudes of the microwave field was varied from 0.6 ns (at the start of the leading edge of the beam) to 0.3 μs (when the beam-plasma discharge attains the quasisteady state). Since the startup transient processes, with a duration roughly equal to the transit time L/v_b, cannot be described in the proposed model, the system was assumed to be filled at the initial time with a slightly ionized gas having a density $n_{e0} = 10^{10}$ cm^{-3} and an electron temperature $T_{e0} = (2/3)\varepsilon_e = 0.15$ eV and by a beam whose density, beginning at time $t = 0$, increases from $0.01 n_b$ to the maximum value n_b within the time span of the leading edge of the beam $\tau_f \gg L/v_b$.

Plasma oscillations were created at the input to the system ($z = 0$) by modulating the beam electron density at the local plasma frequency $\omega_e(0,t)$. The modulation factor was chosen to avoid any appreciable influence on the beam-plasma discharge characteristics.

RESULTS OF THE CALCULATIONS

The calculations were carried out for the following parameters of the beam, the gas, and the plasma: beam density $n_b = 10^{10}$ cm^{-3}; beam radius $a = 1$ cm; particle energy $W = 0.3$ MeV; duration of the leading edge $\tau_f = 5$ μs; pulse duration $\tau = 50$ μs; gas (air) pressure $p = 1$ Torr; initial gas temperature $T_{g0} = 300$ K; initial plasma density $n_{e0} = 10^{11}$ cm^{-3}; initial plasma temperature $T_{e0} = 0.15$ eV; length of the drift chamber $L = 1$ m; radius of the chamber $b \gg a$ for the first regime and $b = a$ for the second regime.

The results of the calculations are shown in the figures in the form of distributions with respect to the coordinate z at

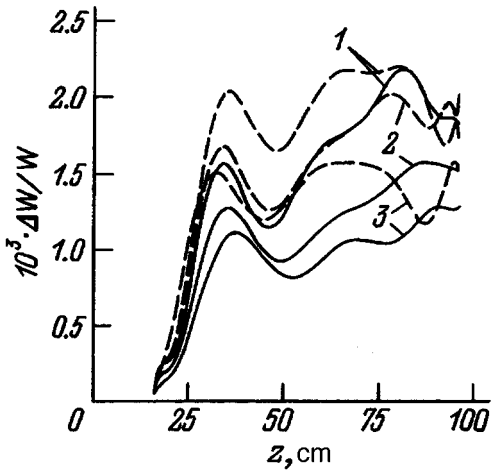

FIG. 1. Distribution of relative beam energy losses $\Delta W/W$ with respect to the coordinate z at three different times in the regimes $b \gg a$ (solid curves) and $b=a$ (dashed curves). *1* — $t=5$ μs; *2* — 20 μs; *3* — 50 μs.

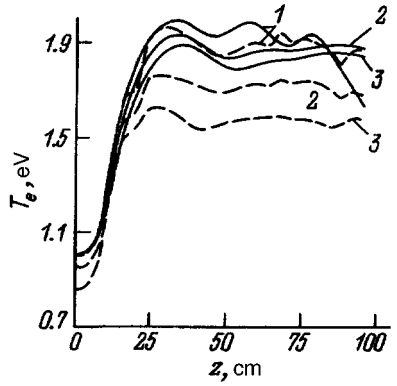

FIG. 3. Distribution of the plasma electron temperature T_e with respect to the coordinate z. The parameters are the same as in Fig. 1.

three different times t for the relative beam energy losses $\Delta W/W$ (Fig. 1), the amplitude of the electric field E_0 (Fig. 2), the temperature T_e (Fig. 3) and density n_e (Fig. 4) of the electronic plasma component, and the temperature of the gas T_g (Fig. 5).

According to the calculations, a dissipative, three-dimensional, two-stream, electron−electron instability with the plasma frequency develops at the leading edge of the beam current pulse ($t<\tau_f$), with beam−plasma interaction entering a nonlinear stage over the length of the system. Rapid heating of the plasma electrons and ionization of the gas take place in the interval of maximum amplitudes of the unstable oscillations near the end of the system. In accordance with the results of Ref. 13, the longitudinal inhomogeneity of the density of the generated plasma takes the instability into a nonresonant state, in which the beam loses less than 1% of its energy over the length of the system. The beam-plasma discharge enters a quasisteady state, which slowly evolves with time and is characterized by an abrupt increase in the beam energy losses, the energy of the electric field, and the temperature and density of the electronic component of the plasma over the first third of the length of the system, where the instability is driven up to a nonlinear level, and by a slightly nonuniform distribution of these quantities over the remaining path of the beam in the drift chamber (Figs. 1–4, curves *1*). The temperature of the gas toward the end of the leading edge of the beam exhibits an appreciable rise only in the second regime ($b=a$) (Fig. 5, curve *1*), in which the specific heat of the gas is lower than in the first regime.

With the passage of time ($t>\tau_f$) the gas temperature T_g rises considerably as a result of heat release in the dissociative recombination of plasma electrons and ions (Fig. 5, curves *2* and *3*). Here, in the first regime ($b \gg a$) under the conditions of constant pressure p and a decreasing gas density ρ, the collision frequency ν_e (20) of plasma electrons with neutrals decreases, as do the Joule losses of the microwave field in the plasma. As a result, the electron temperature T_e drops (Fig. 3, solid curves *2* and *3*), and the cumulative ionization frequency of the plasma ν_{ie} [Eq. (16)] diminishes. The result is an appreciable decrease in the plasma density n_e (Fig. 4, curves *2* and *3*).

In the second regime, the rise in the gas temperature T_g with passing time under the conditions of a constant gas density ρ primarily causes the dissociative recombination coefficient α_{dr} [Eq. (21)] to decrease, so that the plasma density n_e increases substantially (Fig. 4, dashed curves *2* and *3*), and its temperature T_e drops significantly (Fig. 3, dashed curves *2* and *3*). Since the heating rate of the gas, according

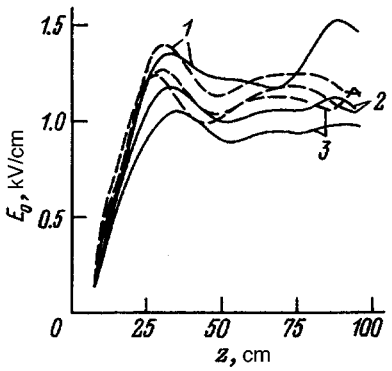

FIG. 2. Distribution of the amplitude of the electric field E_0 with respect to the coordinate z. The parameters are the same as in Fig. 1.

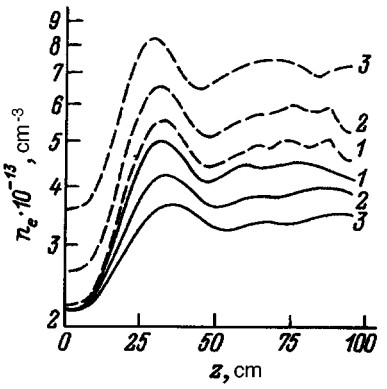

FIG. 4. Distribution of the plasma density n_e with respect to the coordinate z. The parameters are the same as in Fig. 1.

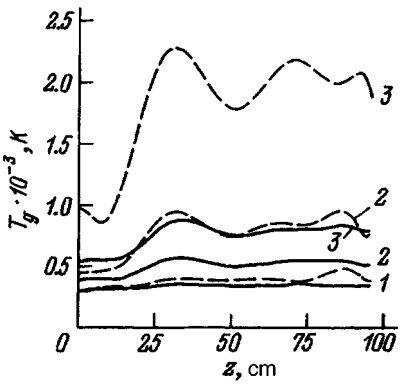

FIG. 5. Distribution of the gas temperature T_g with respect to the coordinate z. The parameters are the same as in Fig. 1.

to (19), is proportional to n_e^2, the gas temperature toward the end of the beam pulse is found to be much higher than in the first regime, attaining a value $T_g = 2.3 \times 10^3$ K (Fig. 5, dashed curve *3*). The more pronounced heating of the gas in the second regime corresponds to the increase in the beam energy losses and the level of the field in the plasma on the average over the length of the beam−plasma interaction zone (dashed curves in Figs. 1 and 2).

The calculations in other regimes with parameters similar to those described above, where the beam-plasma instability is able to attain a nonlinear level over the interaction length, yield qualitatively similar results.

CONCLUSIONS

To summarize, the results of calculations using the model developed in the study for beam-plasma discharge in a slightly ionized, medium-pressure gas in the long-pulse regime thus suggest following scheme of the process. A one-dimensional, spatial, dissipative beam-plasma instability evolves at the leading edge of a relativistic electron beam, attaining a nonlinear stage over the length of the system, but without leading to scattering or defocusing of the relativistic electron beam. In the interval of maximum amplitudes of the beam-generated plasma oscillations, which is situated close to the output end of the system, the plasma electrons are subjected to rapid heating, which causes them to begin ionizing the gas. The density inhomogeneity of the revitalized plasma stabilizes the instability at a level low enough that the beam can traverse the drift chamber with relatively low energy losses, but still high enough to maintain discharge. Heat is released during the dissociative recombination of plasma electrons and ions, raising the temperature of the gas. Since the plasma replenishment rate in the beam-plasma discharge begins at a certain time (in the investigated situation when $T_e \gtrsim 2$ eV and $n_e \gtrsim 10^3 n_b$) to exceed the rate of ionization of the gas by relativistic beam electrons, and the heat released in dissociative recombination varies as the square of the plasma density n_e, the collective-discharge mechanism of heating of the gas could be far more efficient than the classical heating mechanism due to ionization losses of the relativistic electron beam.

The authors are grateful to A. T. Kunavin, P. M. Tokar′, and A. V. Danilov for a discussion and the opportunity to review the results of their experiments, and also to O. A. Gordeev for consultation on certain aspects of plasma kinetics.

[1] S. Jordan, A. Ben-Amar Baranga, G. Benford *et al.*, Phys. Fluids **28**, 366 (1985).
[2] V. M. Batenin, V. S. Jivopistsev, A. O. Ikonikov *et al.*, Fiz. Plazmy **17**, 434 (1991) [Sov. J. Plasma Phys. **17**, 254 (1991)].
[3] M. V. Gladyshev, M. G. Nikulin, and A. B. Sionov, Fiz. Plazmy **17**, 938 (1991) [Sov. J. Plasma Phys. **17**, 546 (1991)].
[4] I. F. Kharchenko, Ya. B. Faĭnberg, A. A. Kornilov *et al.*, Zh. Tekh. Fiz. **31**, 761 (1961) [Sov. Phys. Tech. Phys. **6**, 551 (1962)].
[5] A. A. Ivanov, Fiz. Plazmy **2**, 277 (1976) [Sov. J. Plasma Phys. **2**, 152 (1976)].
[6] A. V. Baĭtin and M. G. Nikulin, Fiz. Plazmy **16**, 358 (1990) [Sov. J. Plasma Phys. **16**, 203 (1990)].
[7] N. D. Borisov, A. V. Gurevich, and G. M. Milikh, *Artificially Ionized Region in the Atmosphere* [in Russian], IZMIR AN SSSR, Moscow (1986).
[8] Yu. A. Lupan, Zh. Tekh. Fiz. **46**, 2321 (1976) [Sov. Phys. Tech. Phys. **21**, 1367 (1976)].
[9] Yu. P. Raĭzer, *Physics of Gas Discharge* [in Russian], Nauka, Moscow (1987).
[10] J. H. Yee, R. A. Alvarez, D. J. Mayhall *et al.*, Phys. Fluids **29**, 1238 (1986).
[11] Pen Tszai-Chen and A. L. Pindroh, Am. Rocket Soc. Paper No. 1995, (1958) [Russ. transl., Vopr. Raket. Tekh. Kosmonavt., No. 12, 3 (1962)].
[12] V. P. Konovalov and É. E. Son, Khim. Prom., No. 14, 194 (1987).
[13] A. V. Baĭtin, M. G. Nikulin, and A. B. Sionov, Fiz. Plazmy **16**, 363 (1990) [Sov. J. Plasma Phys. **16**, 206 (1990)].

Translated by James S. Wood

Equal-path dynamics of pulsed ion streams

A. A. Matyshev

State Technical University, 195251 St. Petersburg, Russia
(Submitted December 30, 1995)
Zh. Tekh. Fiz. **67**, 99–102 (May 1997)

A class of alternating electromagnetic fields, in which the ion trajectories from a short-duration packet do not depend on the initial energy (velocity) of ions, is described. It is suggested that particle dynamics in such fields be called equal-path dynamics. As an example, the deflecting properties of a parallel-plate capacitor, with an equal-path dependence of the voltage on time, are studied.[S1063-7842(97)01505-5]

Methods for analyzing the energies, masses, and angular and charge spectra of fluxes of charged and neutral particles have emerged lately as an important part of physics and engineering. The development of new methods and more sophisticated instruments is pursued in response to the attendant problems and requirements, including the need to extract the maximum available information in a single instrument (e.g., the simultaneous determination of the energy and mass spectra). However, the design of energy-mass analyzers using static electromagnetic fields is thwarted by the fundamental impossibility of determining the specific charge or mass of a particle from the position of the detector in a pure electric field. In static magnetic fields, on the other hand, it is required either to implement the energy focusing of ions or to generate a monoenergetic flux, making it impossible to determine the initial ion energy. In this light, the problem of simultaneously analyzing the energies, masses, and exit angles of an ion flux can be effectively attacked by switching to time-dependent fields using the time-of-flight technique.

However, with the exception of instruments utilizing harmonic alternating fields (rf and quadrupole mass spectrometers), few papers have been published on nonmagnetic methods for the analysis of charged particles in alternating electric fields.[1] And those papers typically omit any mention of the principle underlying the choice of time dependence of the investigated alternating field. In other words, so far no attempts have been made to find classes of alternating electromagnetic fields endowed with any sort of prescribed corpuscular-optical properties for the analysis of charged particle fluxes.

Here we propose a physical principle that can be used to find a class of time-dependent electromagnetic fields, which is unique in its properties and of utmost interest from the standpoint of subsequent investigations of fields in the class. The new principle is the requirement that the spatial trajectory of ions moving in the electromagnetic field be independent of their initial kinetic energy (the equal-path principle).

The equal-path principle can be used to select fields for investigation with the property that energy dispersion is identically zero. This means that ions of different energies must move in such fields along ''stationary'' flight paths, one after the other in sequence, permitting their energy to be determined by the time-of-flight technique. On the other hand, the presence of mass dispersion and the capability of focusing the flux with respect to exit angles provide a technique for measuring ion mass.

It has been established[2,3] on the basis of a solution of the corresponding inverse problem that the equal-path principle is satisfied in the nonrelativistic approximation by a unique class of alternating electromagnetic fields whose electric component varies as t^{-2} and whose magnetic component varies as t^{-1}. The particle source must be pulsed and be situated in the field-free (drift) region.

We show that this is indeed the case. Let an alternating electromagnetic field exist in a certain region of space (as an example, see Fig. 1), where its effect on a charged particle is described by the Lorentz force

$$\mathbf{F}=q\left[\frac{\mathbf{E}(x,y,z)}{t^2}+\mathbf{V}\times\frac{\mathbf{B}(x,y,z)}{t}\right], \tag{1}$$

in which the time $t \gg t_{\min}>0$.

In addition, let charged particles of different masses and velocities v be emitted in all possible directions from a point source I at time $t=0$. We replace t by the new independent variable

$$\xi=\ln\,[vt/D(\vartheta,\varphi)], \tag{2}$$

where $D(\vartheta,\varphi)$ is the distance from the point source to the boundary of the field along the straight line defined by the initial exit angles ϑ and φ.

The transformation to the new variable ξ in Newton's equations of motion yields the equations of motion and initial conditions

$$m\left(\frac{d^2\mathbf{r}}{d\xi^2}-\frac{d\mathbf{r}}{d\xi}\right)=q\left[\mathbf{E}(\mathbf{r})+\frac{d\mathbf{r}}{d\xi}\times\mathbf{B}(\mathbf{r})\right], \tag{3}$$

$$\xi_0=0:\quad \mathbf{r}_0=\mathbf{r}_0(\vartheta,\varphi),\quad \left(\frac{d\mathbf{r}}{d\xi}\right)_0=D(\vartheta,\varphi)\mathbf{n}(\vartheta,\varphi), \tag{4}$$

where $\mathbf{n}(\vartheta,\varphi)$ is the unit vector in the direction of particle emission.

It follows from Eqs. (3) and (4) that the trajectories of particles emitted from the source at time $t=0$ in fields of the family (1) do not depend on their initial velocity (kinetic energy), but do depend on the specific charge. Particles of one species exiting from the source in a particular direction move along a single spatial trajectory in sequence, slower particles followed by faster ones. This phenomenon is easy to explain physically: For the class of fields (1) the balance

 1063-7842/97/050547-03$10.00

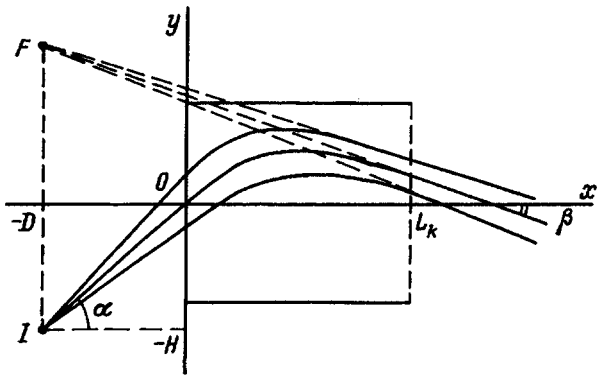

FIG. 1. Deflection of a pulsed ion beam in a plane-parallel capacitor with the electric field (5). *I* — Point source of particles; *F* — ideal virtual image.

between the kinetic energy and the potential energy of the particles remains constant, and in the final analysis this balance is what determines the configuration of the trajectory in any specific case. The same explanation accounts for the singularity of the fields (1) $t=0$, because an infinite field is required to keep a particle on trajectory in the limit $v\to+\infty$.

In practice, of course, the field must vary with time according to the law (1), but only beginning at a certain positive time $t_{\min}>0$ when the first particles from the source reach the boundary of the field.

Since the class of time-dependent fields (1) is based on the principle that the particle trajectories are independent of (invariant with respect to) their initial energy, we refer to the particle dynamics in such fields as equal-path dynamics.

In conservative systems all solutions of dynamical problems depend on the difference $(t-t_0)$, while in equal-path dynamics the solutions acquire an arbitrary constant in the form t/t_0. This property follows from Eqs. (3) and (4), which do not contain ξ in explicit form and, hence, give one of the constants in the combination $\xi-\xi_0$. Relation (2), in turn, transforms $\xi-\xi_0$ to t/t_0. From the mathematical standpoint, therefore, the equal-path property can be exploited to transform the time-dependent problem for the class of fields (1) to a time-invariant problem, which does not contain the independent variable in explicit form if the particles are launched at time $t=0$.

An analysis of the operation of very simple electrode configurations with an equal-path time dependence of the potential indicates significant and attractive characteristics of equal-path dynamics over the dynamics of conservative fields. For example, an electric quadrupole capacitor with a field that varies as t^{-2} implements focusing that does not depend on the particle exit angle (disregarding the influence of scattered fields), whereas a static field provides only first-order focusing with resect to exit angle.[4]

Another striking example is the operation of a plane-parallel capacitor as a deflecting system for pulsed ion fluxes. In the static regime a plane-parallel capacitor was first used to deflect charged particles in 1898.[5] In this case ideal deflection is possible only for a monoenergetic beam in the paraxial approximation in a model with a rectangular field boundary[6]; deflection is considered to be ideal when it is strictly proportional to the applied voltage, and the spot is undistorted. In general, however, the deflection depends on the particle energy and on the angle of entry of particles into the capacitor.

In the equal-path regime the deflection in the same model with a rectangular field boundary is ideal in the above-defined sense. We now prove this statement.

Let the electric field strength E have the form

$$E(t)=E_0\left(\frac{t_{\min}}{t}\right)^2, \tag{5}$$

and at time $t=0$ let a pulsed flux of particles be emitted from the point $(x_i,y_i)=(-D,-H)$ (Fig. 1). In the case of symmetrical feed to the plates the beam enters a zero-potential zone at the entry to the capacitor, the source itself is at zero potential. We know in this case that oblique incidence of the beam on the capacitor diminishes the influence of the edge field on the trajectories in the drift space.[6] Otherwise, this influence is totally disregarded in the model with a rectangular field boundary.

The simplicity of Newton's equations in this case obviates the need for the change of variable (2), and they can be integrated directly:

$$m\ddot{x}=0, \tag{6}$$

$$m\ddot{y}=qE_0\left(\frac{t_{\min}}{t}\right)^2. \tag{7}$$

We introduce the parameter L, which has the units of length:

$$L=\frac{|qE_0|}{m}t_{\min}^2, \tag{8}$$

and we take into account the initial data

$$t_0=\frac{D}{v\cos\alpha}:\begin{cases}x_0=0, & y_0=D\tan\alpha-H,\\ \dot{x}_0=v\cos\alpha, & \dot{y}_0=v\sin\alpha;\end{cases} \tag{9}$$

the solutions of the equations can then be written in the form

$$x=(t-t_0)v\cos\alpha, \tag{10}$$

$$y=L\ln(t/t_0)+(v\sin\alpha-L/t_0)(t-t_0)+D\tan\alpha-H. \tag{11}$$

Eliminating time from Eqs. (10) and (11), we obtain the geometrized trajectory equation

$$y=L[\ln(1+x/D)-x/D]+(x+D)\tan\alpha-H \tag{12}$$

and the slope of the trajectory

$$\frac{dy}{dx}=-\frac{L}{D}\frac{x}{(D+x)}+\tan\alpha. \tag{13}$$

We now determine the position of the virtual image F of the source I. We write the equation for the trajectory of a particle after it exits from the capacitor:

$$Y-y_k=\left(\frac{dy}{dx}\right)_k(X-L_k), \tag{14}$$

where

$$y_k = L[\ln(1+L_k/D) - L_k/D] + (L_k+D)\tan\alpha - H, \quad (15)$$

$$\left(\frac{dy}{dx}\right)_k = \tan\beta = \tan\gamma + \tan\alpha, \quad (16)$$

$$\tan\gamma = -\frac{L\cdot L_k}{D(D+L_k)}. \quad (17)$$

To determine the position of the virtual image, we differentiate Eq. (14) with respect to α:

$$-\frac{(D+L_k)}{\cos^2\alpha} = \frac{(X-L_k)}{\cos^2\alpha}. \quad (18)$$

From Eqs. (18) and (14) we finally obtain the coordinates of the image

$$X_f = -D, \quad Y_f = L\ln(1+L_k/D) - H. \quad (19)$$

We note that for large deflections in the static regime additional distortions occur as a result of beam refraction at the exit, since the particles leave the capacitor at points with nonzero potential. In the equal-path regime this problem is also ameliorated by the decay of the potential with time.

It is evident from Eq. (19) that the above-described dynamic capacitor has the effect of producing an ideal virtual image of a point source (disregarding scattered fields) after the beam is transmitted. The transformation of the beam in geometrized phase space can ultimately be written as a linear inhomogeneous transformation

$$\begin{pmatrix} y_k \\ \tan\beta \end{pmatrix} = \begin{pmatrix} y_c \\ \tan\gamma \end{pmatrix} + A\begin{pmatrix} y_0 \\ \tan\alpha \end{pmatrix}, \quad (20)$$

where A is the matrix of the drift space transformation,

$$A = \begin{pmatrix} 1 & (L_k+D) \\ 0 & 1 \end{pmatrix}, \quad (21)$$

$$y_0 = -H, \quad y_c = L[\ln(1+L_k/D) - L_k/D], \quad (22)$$

and $\tan\gamma$ is given by Eq. (17) and is the exit (turning) angle of particles leaving the capacitor along the x axis.

Simultaneously it follows from (16) that a parallel beam of ions is also transformed into a parallel beam, but the angle of deflection depends on the specific charge of the particle. Consequently, the above-described dynamic plane-parallel capacitor represents a prism, which turns a pulsed ion beam in different directions according to particle mass.

It should be recalled that the total particle energy is not conserved in time-dependent fields. The final velocity can be found from the equation

$$v\cos\alpha = v_k\cos\beta, \quad (23)$$

which is a consequence of the invariance of the longitudinal velocity component.

Another important question deserves our attention. The discussion so far is valid only for a pulse of zero duration. The real duration of the source pulse $\Delta T > 0$ can be taken into account as a certain spatial broadening of the source. Thus, if the particle source I is not a point source but has dimensions ΔD, ΔH, the dimensions of the virtual image can be determined from relations (19) and (24):

$$\Delta X_f = -\Delta D, \quad \Delta Y_f = -\frac{L\cdot L_k}{D(D+L_k)}\Delta D - \Delta H. \quad (24)$$

It follows from Eqs. (24) that the dimensions of the virtual image approximately correspond to the dimensions of the real source at $\Delta T = 0$. Now an upper bound of the pulse duration $\Delta T > 0$ can be estimated as follows: We associate $\Delta T > 0$ with a certain broadening of the source, roughly equal to $v_{\max}\Delta T$, where $v_{\max}$ is the maximum velocity with which particles are launched from the source. Equations (24) can then be used to determine the exact dimensions of the virtual image in each specific instance, depending on the angular distribution of the particles.

We note that an electrodynamic energy-mass analyzer utilizing a plane-parallel capacitor with the field configuration (5) has been around for more than twenty years.[8–13] The operation of this type of analyzer resembles that of a static plane-parallel capacitor in the specular reflection mode, but its operation as a dynamic particle deflector has never been described. However, despite persistent interest in the device, it has one very significant drawback: its fundamental incapability of focusing with respect to exit angles, and its basic time dependence has probably not been perceived as a universal quality inherent in any spatial distribution of electromagnetic fields.

We summarize briefly with the remark that from the theoretical standpoint equal-path dynamics is a very interesting object of investigation, and from the practical standpoint it offers a promising field for the design of a new type of dynamic analyzers.

[1] P. U. Arifov, R. Kh. Ayupov, and A. V. Shevchenko, in *New Possibilities in Corpuscular Spectrometry* [in Russian], Tashkent (1979), pp. 144–158.
[2] A. A. Matyshev, in *Proceedings of the Ninth All-Union Seminar on Computational Methods in Auger Electron Spectroscopy* [in Russian], Tashkent (1988), p. 76.
[3] A. A. Matyshev, Trudy Leningr. Politekh. Inst., No. 429, 67 (1989).
[4] A. A. Matyshev, Trudy Leningr. Gos. Tekh. Univ., No. 436, 61 (1991).
[5] E. Brüche and O. Scherzer, *Geometrical Electron Optics* [Russian translation], Leningr. Gazetno-Zhurn. Knizhn. Izd., Leningrad (1943).
[6] W. Glaser, *Grundlagen der Elektronenoptik*, Springer, Vienna (1952) [Russ. transl., Gostekhizdat, Moscow (1957).
[7] A. A. Matyshev, in *Proceedings of the Tenth All-Union Seminar on Computational Methods in Auger Electron Spectroscopy* [in Russian], Tashkent (1990), p. 59.
[8] M. Oron and Y. Paiss, Rev. Sci. Instrum., **44**, No. 1293 (1973).
[9] S. S. Chowdhury, R. M. Clement, and H. T. Miles, J. Phys. E **13**, 1099 (1980).
[10] R. M. Clement and H. T. Miles, J. Phys. E **16**, 377 (1983).
[11] J. Eicher, K. Rohr, and H. Weber, J. Phys. E **16**, 903 (1983).
[12] Yu. A. Bykovskiĭ, A. E. Gruzinov, and V. B. Lagoda, Pribory Tekh. Éksp., No. 4, 125 (1989).
[13] M. E. Belov, Yu. A. Bykovskiĭ, A. E. Gruzinov, and V. B. Lagoda, Pribory Tekh. Éksp., No. 2, 113 (1993).

Translated by James S. Wood

Backscattering of low-energy (0–8 eV) electrons by a silicon surface

O. B. Shpenik, N. M. Érdevdi, and T. Yu. Popik

Institute of Electron Physics, Ukrainian Academy of Sciences, 294016 Uzhgorod, Ukraine
(Submitted November 30, 1995)
Zh. Tekh. Fiz. **67**, 103–108 (May 1997)

An experimental apparatus and method for investigating elastic and inelastic backscattering (180°) of low-energy (0–8 eV) monoenergetic electrons by a solid surface are described and the first results are presented for the reflection of electrons by samples of pure single-crystalline silicon with a polished surface (Si), doped p-type single-crystalline silicon with a porous surface (Si-p) as well as H_2O and H_2O_2 passivated porous samples, Si-p + H_2O and Si-p + H_2O_2. A structure due to the excitation of surface plasmons has been observed for the first time in the loss spectra. Features corresponding to a resonance excited state of molecular nitrogen adsorbed on the surface of porous silicon have been observed in the constant residual energy spectra. [S1063-7842(97)01604-8]

INTRODUCTION

In the last ten years, substantial progress has been made in the study of the characteristic features of elastic and inelastic interaction of low-energy monoenergetic electrons with atoms and molecules. This is largely due to the development of new methods of investigation and new capabilities of the experimental technique. Together with the now conventional optical (spectroscopic) methods for investigating the excitation of atoms, the methods of electron and threshold spectroscopy, emitted-electron spectroscopy, and metastable spectroscopy have been extensively developed. It has been found that many of these methods can also be successfully used to investigate solid surfaces. For example, methods such as Auger electron spectroscopy, characteristic loss spectroscopy, and electron-photon emission have now been quite extensively developed, and high-resolution electronic spectrometers ($\Delta W = 6 - 20$ meV) in combination with ultrahigh vacuum technique ($P = 10^{-9}$ Pa) have made it possible to study the resonance features of electron scattering by both atomically pure surfaces and by surfaces containing adsorbed molecules.[1–3] This paper describes a new experimental apparatus and method for the studying angular and energy distributions of elastically and inelastically backscattered low-energy (0–10 eV) monoenergetic electrons under normal incidence on the sample surface and it reports the first results of these investigations. The role of the state of the surface of single-crystal and porous silicon samples in the scattering process is studied.

EXPERIMENTAL APPARATUS AND PROCEDURE

The experiments were performed on an apparatus consisting of the following basic units: a stainless steel vacuum chamber containing a trochoidal electron backscattering spectrometer and a goniometer holding the experimental samples, a power supply for the electron spectrometer, and a multichannel system for recording the primary and scattered electron currents.

To obtain a monoenergetic electron beam and to analyze the elastically and inelastically reflected electrons, in our experiments we used a trochoidal electron spectrometer which has been used for many years for studying the elastic scattering of electrons by atoms and the excitation and ionization of atoms and molecules.[4–6] It consists of two trochoidal electron energy analyzers positioned in tandem. The first one serves as an electron monochromator and the second one is used as an analyzer of elastically and inelastically reflected electrons (Fig. 1). This apparatus is unique because the longitudinal magnetic field allows it to operate at very low energies (starting virtually with zero) with a high degree of transmission (up to 95%).[4]

The property of a trochoidal electron monochromator, that in crossed electric fields electrons undergo, besides directed motion, drift in a direction perpendicular to **E** and **H** irrespective of the direction of the initial motion of the electrons,[7] was used in developing the backscattering (180°) electron spectrometer. As a result the elastically backscattered electrons move in the direction of the trochoidal electron monochromator and, having entered the crossed-field region, are displaced in the direction of drift of the primary electron beam. Therefore, having traversed the analyzer region, the elastically scattered electrons are displaced by twice the distance relative to the primary beam. For this reason, electrons elastically scattered by 180° can be detected by positioning an electron detector at the point of the indicated displacement at the analyzer exit.

The spectrometer units were made of molybdenum, and an indirectly-heated oxide cathode was used as the electron source. The spectrometer was placed inside a stainless steel vacuum chamber, which was evacuated by oil-free pumping (zeolite and gas-discharge pumps). The chamber was separated from the pump by a system of diaphragms to prevent the sputtered titanium from reaching the spectrometer electrodes and the experimental samples. The goniometer makes it possible to load up to five samples into the chamber simultaneously. Prior to the measurements, the electron spectrometer and samples were heated at $T = 500$ K in a 10^{-6} Pa vacuum for 50–60 h, after which the surface was cleaned for 1–2 h with high-energy electrons. The vacuum in the chamber during the measurements was 10^{-8} Pa. A uniform magnetic field (0.01–0.015 T), necessary for the operation of the spectrometer, was produced with Helmholtz coils with an

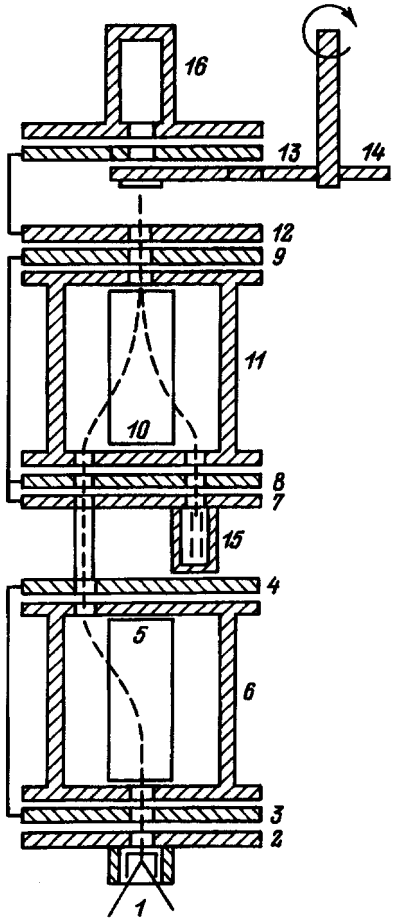

FIG. 1. Diagram of the electronic backscattering spectrometer. *1*—Indirectly heated oxide cathode; *2*—extracting electrode; *3*, *4*—entrance and exit electrodes of the selector; *5*, *6*—inner and outer electrodes of the cylindrical condensor of the selector; *7*–*9*—entrance and exit electrodes of the analyzer; *10*, *11*—inner and outer electrodes of the cylindrical condensor of the analyzer; *12*, *13*—accelerating electrodes of the spectrometer; *14*—sample; *15*—backscattered-electron collector; *16*—collector for monitoring the primary electron beam.

inner diameter of 210 mm. Power was supplied to the coils by a highly stabilized current source.

The power supply for the spectrometer includes a highly stabilized dc voltage source with noise not exceeding 1–2 mV. A step voltage generator with a step of 0.01–0.08 V was used to vary the accelerating (or decelerating) voltage smoothly.

The detection system consists of two V7-30 digital nanoammeters combined with multichannel information recording with the aid of a F-36 measurement-computing system. One channel of the system recorded the beam current at the sample and the second recorded the scattered-electron current.

Out instrumentation system enabled us to perform four types of experiments.

1. Measure the energy dependence of the elastically backscattered electrons. For this, optimal dc potentials are established on the monochromator and analyzer electrodes and the current on the collector *15* is measured with the electron energy varied by the potential difference on the electrodes *1* (cathode) and *14* (sample) (Fig. 1).

2. Study the energy-loss spectra as well as the angular distributions of both the elastically and inelastically reflected electrons (differential cross sections). In this case the energy E_0 of the incident electrons (the potential difference on the electrodes *1* and *14*) is maintained constant and the current on the collector *15* is measured while varying the potential difference between the sample *14* and the entrance electrode *9* of the analyzer. If backward scattering by 180° occurred, then in this regime we would obtain a ''pure'' energy-loss spectrum (as in the case of electrostatic analyzers). However, because in the present design the longitudinal component of the electron velocity is what is analyzed, electrons with the residual energy $E_r = E_0 - E_1$ give to the loss spectrum a reorientational continuum, extending in energy from E_0 (for scattering by 180°) to zero (for scattering by 90°).

3. Detect electrons with constant residual energy, including virtually zero energy (threshold spectrum). In this regime a constant potential difference is maintained between the analyzer *9* and the sample *14* with varying incident electron energy E. As a result, only the electrons which have the prescribed residual constant energy are detected.

4. Measure the energy dependence of the scattering cross section. In this case the analyzer potential V_9 relative to the cathode potential V_1 is maintained constant with varying incident electron energy E, and the current on the collector *15* is measured. However, it should be noted that as the incident electron energy is varied, the angular acceptance changes and therefore this factor must be taken into account when analyzing the curves.[5]

RESULTS AND DISCUSSION

We investigated four silicon samples. The first sample consisted of pure single-crystal silicon with a polished surface (Si), the second sample consisted of porous p-type doped silicon (Si-p), the third sample consisted of porous doped silicon passivated with water (Si-p + H_2O), and the fourth sample consisted of porous doped silicon passivated with hydrogen peroxide (Si-p + H_2O_2). Nanometer-size pores were obtained by electrochemical etching of the polished surface. The preparation technology and the properties of the samples are described in detail in Ref. 8. This choice of samples was dictated by the need to investigate in the course of a single experiment the effect of the state of the surface on the electron scattering processes.

The energy dependences for electrons elastically backscattered by the experimental samples in the energy range 0–4 eV are displayed in Fig. 2. The energy scale of the electron beam was calibrated according to the shift in the current-voltage characteristic of the electron current on the sample. The maximum of the electron energy distribution function was taken as the zero point of the scale. In this method, allowing for the high degree of monochromaticity of the electron beam, the calibration accuracy was not worse than $\pm 0.02 - 0.06$ eV and depended on the resistivity of the samples.

Analysis of Fig. 2 shows that the energy spectra of the elastically backscattered electrons are insensitive to the state of ordering of the atoms on the surface. There is a general tendency for the electron reflection intensity to decrease as

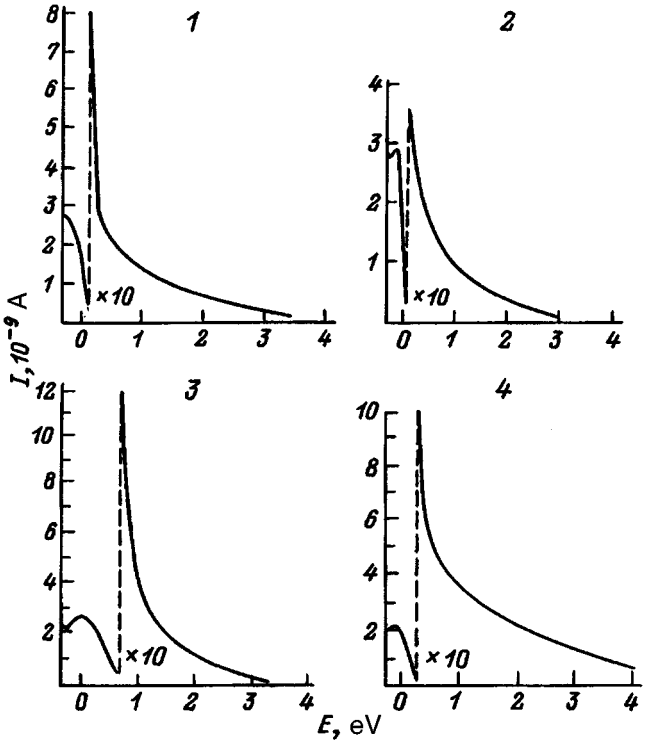

FIG. 2. Energy dependences of the elastic backscattering of electrons by surfaces. *1*—Si, *2*—Si-*p*, *3*—Si-*p* + H_2O, *4*—Si-*p* + H_2O_2.

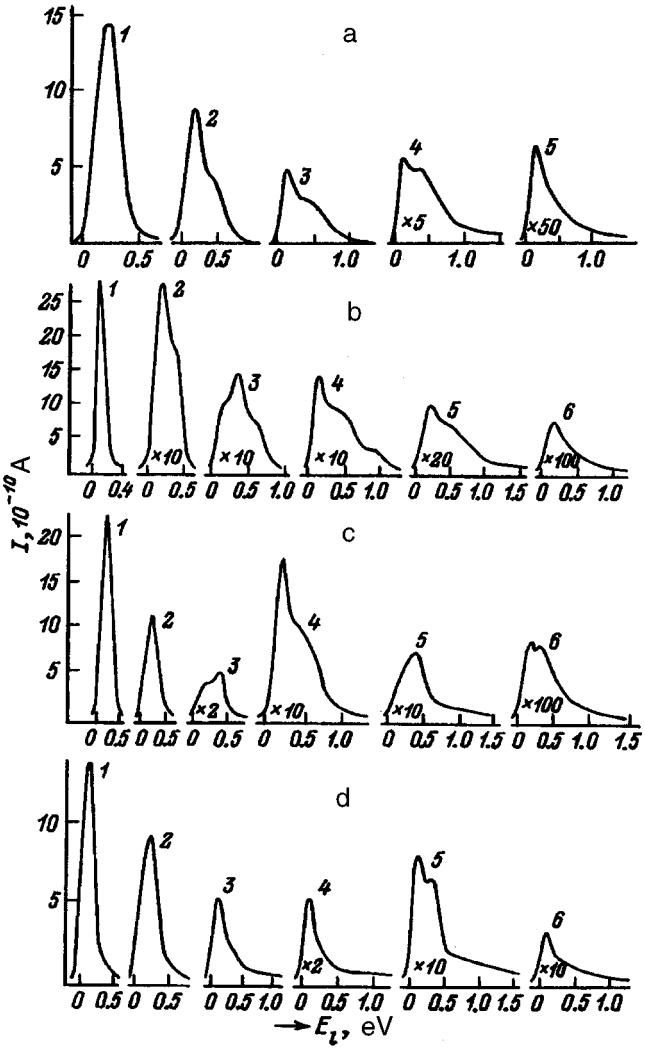

FIG. 3. Electron energy-loss spectra of surfaces (in parentheses E_0, eV). a—Si; (*1*—0.7, *2*—1.0, *3*—1.5, *4*—2.5, *5*—8.0); b—Si-*p* (*1*—0.4, *2*—0.7, *3*—1.0, *4*—1.5, *5*—2.5, *6*—8.0); c—Si-*p* + H_2O (*1*—1.0, *2*—1.5, *3*—2.0, *4*—2.5, *5*—3.5, *6*—8.0); d—Si-*p*+H_2O_2 (*1*—1.5, *2*—2.0, *3*—2.5, *4*—3.0, *5*—3.5, *6*—8.0).

the electron energy increases, and at energies above 4 eV it is negligibly small. Some characteristic features are nonetheless present: a) weak features and a more rapid dropoff of the intensity with increasing energy are observed in the energy dependence of the intensity of the elastically backscattered electrons for the Si and Si-*p* samples (Fig. 2, *1* and *2*); b) no anomalies are observed in the energy dependence of the intensity of the elastically backscattered electrons for the porous passivated samples, but in this case the intensity of the elastically reflected electrons decreases much less with increasing energy than for the Si and Si-*p* samples (Fig. 2, *3* and *4*).

The energy-loss spectra (quasielastic scattering) at some energies are displayed in Fig. 3. All curves are very similar at very low incident electron energies (Fig. 3, a–d, *1*). This is not surprising, since in this case processes other than elastic scattering are still not switched on, and because of the angular acceptance of electrons by the collector *15* is maximum (Fig. 1) we are measuring essentially the instrument function. As one can see from Fig. 3b, *1*, the total width of the electron energy distribution does not exceed 0.08 eV; in addition, characteristically, the instrumental function is practically symmetric with respect to the maximum. As the incident electron energy increases from 0.4 to 2.5–4 eV, the distribution function broadens and the curve becomes symmetric, and in some cases features in the form of steps (sometimes also very small peaks) appear, while at energies $E \sim 4-8$ eV all curves increase rapidly in the energy range corresponding to zero losses (elastic reflection) and drop off with increasing E_1; in addition, the width of the distribution function changes for different samples for large values of E.[1)] The gradual broadening of the energy spectrum is due to the angular distribution of the reflected electrons, and electrons scattered by 90° correspond to the ''cutoff'' towards zero energy, provided that their transverse motion is not limited by diaphragms in the path of the electrons from sample to collector *15*. Evidently, this factor is dominant at low energies $E \sim 0.4-4$ eV. At higher energies (8 eV) the curves for Si, Si-*p*, and Si-*p* + H_2O_2 become smooth, but their half-widths for all samples remain different. This attests to the fact that for Si, Si-*p*, and Si-*p* + H_2O_2 scattering by angles close to 180° is more likely, and for Si-*p* + H_2O the distribution curve is much wider relative to scattering by angles close to 180°. For all samples a maximum with definite energy $E_{l_1} \sim 0.1-0.2$ eV, corresponding to elastic scattering, is characteristically present in the loss spectrum. It is difficult to answer unequivocally the question as to the origin of the characteristic features in the curves. From the fact that for all samples anomalies in the energy dependence in the range 0–1 eV appear in a quite narrow range of incident electron

energies it can be concluded that the anomalies are not diffraction-related but rather of a resonance nature, i.e. they are due to the excitation of local centers (atoms, molecules). For pure single-crystal silicon, only one maximum with energy $E_{l_2} \sim 0.40 \pm 0.05$ eV for an incident electron energy 1.0–2.5 eV is observed in the loss spectrum. The intensity I_2 of the losses is always less than the intensity I_1 of the elastically scattered electrons.

For porous Si-p samples in this range of incident electron energies we observe two maxima with energies $E_{l_2} = 0.30 \pm 0.05$ eV and $E_{l_3} = 0.50 \pm 0.05$ eV (with $E = 1.0$ eV) or $E_{l_3} = 0.8 \pm 0.05$ eV (with $E = 1.5$ eV). We note that at a definite incident electron energy E_0 the intensity I_2 of the inelastically scattered electrons is greater than the intensity I_1 of the elastically scattered electrons (Fig. 3b, *3*); this was not observed in pure single-crystal Si (Fig. 3a). This factor and the decrease in E_{l_2} indicate that the local centers responsible for the appearance of E_{l_2} in porous Si-p and pure single-crystal Si are of a different nature. The large shift in E_{l_3} accompanying a very small increase in E can hardly be due to losses on the same local center. More likely, two different centers of different nature, whose transverse scattering cross sections have a sharp resonance dependence on E, are involved. We did not observe centers with such characteristics on the surface of pure single-crystal Si or H_2O or H_2O_2 passivated porous Si-p.

Only one maximum with $E_{l_2} = 0.35 \pm 0.05$ eV appears in the energy dependence of the losses in the energy range 0–1 eV for surfaces of porous p-type silicon with sorbed H_2O and H_2O_2 molecules. Depending on E, this maximum can be attributed to the same scattering centers which are characteristic of a porous Si-p surface. Sorption shifts E_{l_2} by only a small amount, but it destroys the centers with $E_{l_3} = (0.5 \pm 0.05)$ eV and (0.8 ± 0.05) eV appearing on the surface of porous Si-p. The fact that in all four types of experimental samples the maximum E_{l_2} in the energy-loss spectrum changes very little from $E_{l_2} = 0.4 \pm 0.05$ eV for pure single-crystal Si to $E_{l_2} = 0.30 \pm 0.05$ eV for porous Si-p and falls into an intermediate position for porous Si-$p + H_2O$ and Si-p + H_2O_2 can be attributed to the excitation of surface plasmons.

Let us now examine the constant residual energy spectra of scattered electrons near the elastic peak (Fig. 4). These curves have quite different shapes despite the fact that they were obtained under identical experimental conditions. For the case of scattering by Si-p + H_2O_2 the curves are very narrow at low residual energies (from 0 to 1.0 eV) (Fig. 4d), whereas for Si-p they are strongly broadened for the same residual energies (Fig. 4b) and pure Si and Si-p + H_2O exhibit a similar behavior (Fig. 4a and c). All energy dependences consist of asymmetric curves with a sharp peak at the initial energies.

We note that the energy positions of the anomalies observed in the energy dependences of the losses correlate well with the positions of the features in the energy dependences of the constant residual energy of the electrons. This confirms our previous conclusion that the anomalies are of a resonance nature.

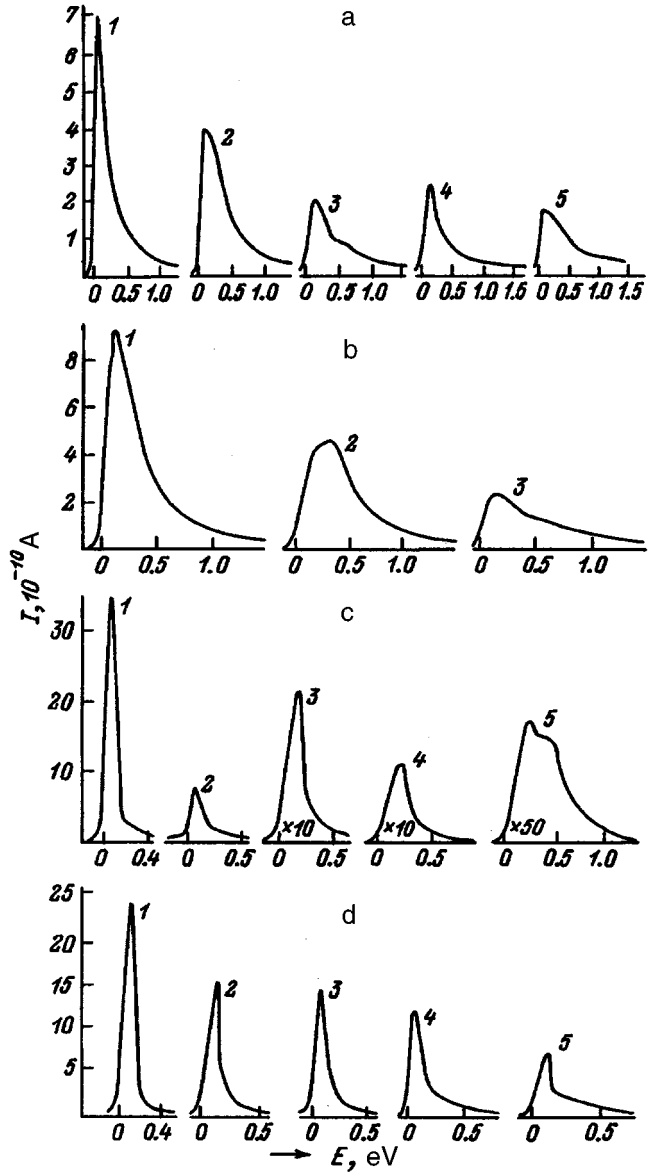

FIG. 4. Constant residual energy spectra with backscattering of electrons by surfaces. a—Si, b—Si-p, c—Si-p + H_2O, d—Si-p + H_2O_2; E_r, eV: *1*—0, *2*—0.2, *3*—0.4, *4*—0.7, *5*—1.0.

Our investigations show that one cannot investigate the structure of solid surfaces using 0–8 eV electrons scattered by 180°, since diffraction still does not occur at such low values of E. Such an electron energy is too low for exciting inner electrons in atoms and correspondingly for investigating the chemical composition of the surface. However, electrons with such energies are very convenient for investigating adsorbed molecules, for which the maxima in the backscattering cross sections are reached at an electron energy corresponding to a resonance state or near it.[9] N_2 molecules in residual gases are very convenient in this respect. The constant residual energy spectra displayed in Fig. 5 for backscattering of electrons by a Si-p surface attest to the presence of fine structure in the region 1.7–2.0 eV, which corresponds in form and position in energy to the excitation

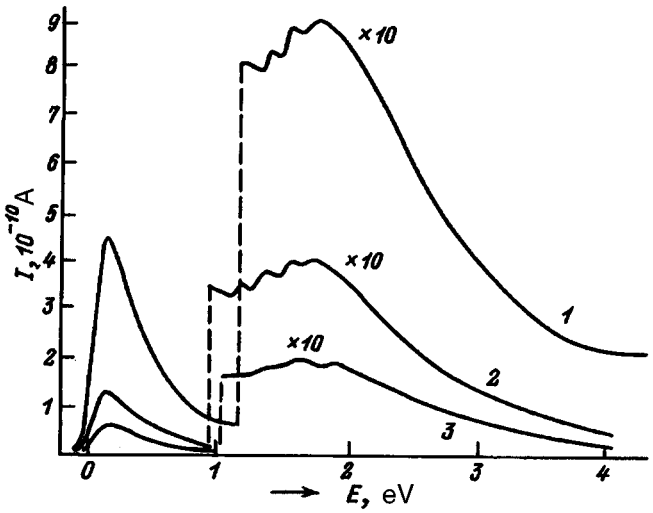

FIG. 5. Constant residual energy spectra with backscattering of electrons by a Si-*p* surface. E_r, eV: *1*—0, *2*—0.5, *3*—1.0.

of a $^2\Pi_g$ resonance state of the N_2 molecule.[1–3,9]

Completing the analysis of the results, it can be concluded that the electron backscattering spectrometer which we have developed makes it possible to investigate both elastic and inelastic electron scattering by different surfaces. To investigate the presence of adsorbed molecules with a large effective resonance backscattering cross section on the surface, it is desirable to investigate the electron energy loss spectra and residual electron energy spectra, which are more informative than the elastic backscattering spectra.

We thank É. B. Kaganovich for providing the samples and for fruitful discussions during the formulation of the problem of these investigations. We also thank A. I. Zavilopulo and I. O. Tsapfel for assisting in this work.

The investigations presented in this paper were conducted in collaboration with the George Soros International Program for the Support of Education in the Exact Sciences of the International ''Renaissance'' Fund (Grant No. ISSEP SPU042080) and were supported by the International Science Foundation and the Ukrainian government (Grant No. UDB200).

[1)]Only the initial region of the loss spectrum near the elastic peak is shown in the figures. This region was found to be more interesting, since in measuring the losses exceeding some value there appears a background due to the double reflection of electrons from the sample *14* and the diaphragm *9*.

[1] R. E. Palmer and P. J. Rous, Rev. Mod. Phys. **64**, 383 (1992).
[2] L. Sanche, J. Phys. B. **23**, 1597 (1990).
[3] L. Sanche in *XVIII ICPEAC, Abstracts of Contributed Papers,* AIP Press, N.Y., 1993, pp. 381–389.
[4] N. I. Romanyuk. O. B. Shpenik, and I. A. Mandi, Zh. Tekh. Fiz. **63,** 138 (1993) [Tech. Phys. **63**, 599 (1993)].
[5] N. I. Romanyuk, O. B. Shpenik, and F. F. Papp, Ukr. Fiz. Zh. **37**, 1639 (1992).
[6] O. B. Shpenik, A. N. Zavilopulo, and T. Yu. Popik in *XVIII ICPEAC, Abstracts of Contributed Papers,* AIP Press, N. Y., Vol. 11, p. 759 (1993).
[7] G. J. Shulz, Rev. Mod. Phys. **45**, 378 (1973).
[8] S. V. Savchnikov, A. V. Sachenko, and T. A. Sukach, Optoélektronika i poluprovodnikovaya tekhnika, No. 27, 3 (1994).
[9] O. B. Shpenik, N. I. Romanyuk, and I. V. Chernysheva, JETP Lett. **41**, 609 (1985).

Translated by M. E. Alferieff

Influence of temperature on the yield of ions desorbed as a result of electron transitions: relaxation model

S. Yu. Davydov

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted December 4, 1995; resubmitted September 10, 1996)
Zh. Tekh. Fiz. **66**, 109–115 (May 1997)

A relaxation model of electron-stimulated desorption is used to investigate the influence of temperature on the yield of desorbed ions and to calculate the corresponding temperature coefficient. Experimental data on the electron-stimulated desorption of lithium and sodium ions from a tungsten surface coated with a silicon monolayer are interpreted on the basis of the theoretical results. © *1997 American Institute of Physics.* [S1063-7842(97)01704-2]

INTRODUCTION

Electron-stimulated desorption (ESD) has lately gained a solid foothold in a variety of techniques for the investigation of surfaces.[1,2] There are several different approaches to the description of this phenomenon. The relaxation model rests on the following hypothesis.[3–5] Under the influence of an external force, i.e., when the adsorption system is bombarded by electrons or photons (Fig. 1), the bond between the constituent atoms (ions) 1 and 2 of an adsorption pair is broken by a redistribution of the electron density between them. Both atoms (ions) acquire a positive charge, whereupon the adsorption system decays and evolves under the influence of Coulomb repulsion of ions 1 and 2. This is the first stage of desorption.

The second stage sets in at some arbitrary time when a hole or two holes localized at the ion 1 disintegrate. If at this time the ion has acquired enough kinetic energy to surmount the barrier, it can be desorbed. A simple analytical theory of the phenomenon has been proposed earlier.[6,7]

As a rule, the influence of temperature on ESD has received little attention. The temperature effect has been investigated in detail in a study of the desorption of positive lithium and sodium ions from a tungsten surface coated with a monatomic silicon film.[8,9] It was found that the temperature coefficients of the yield of desorbed ions can differ substantially for the very same adsorption system, depending on the energy of the bombarding electrons. A qualitative explanation of this disparity has already been set forth in Refs. 8 and 9 on the assumption that the temperature coefficient is related to the direction of initial motion of the desorbed ion. This work has been the driving force behind the present author's interest in the problem.

ALLOWANCE FOR THE INFLUENCE OF TEMPERATURE IN THE RELAXATION MODEL

We consider the thermal vibrations of the atoms (ions) 1 and 2 forming an adsorption pair (Fig. 1). We take into account only vibrations involving a variation of the length of the bond l between ions 1 and 2. We can write

$$l = l_0 + \Delta l \cos(\omega t + \delta), \tag{1}$$

where l_0 is the equilibrium length of the bond, Δl is the amplitude of vibrations of frequency ω, and δ is the phase of the vibrations.

At the instant $t=0$ when the bond is broken under the influence of an external factor, ion 2 is removed to a distance $l = l_0 + \Delta l \cos\delta$. The vibration phase angle δ is a random variable at low densities of adatoms, as we shall assume to be the case here. Figure 2 illustrates the influence of such vibrations on desorption. Clearly, as the bond length decreases, the adatom acquires a high positive energy, which is lost as the bond length increases (this behavior is natural in view of the nonlinear dependence of the interaction on distance). On the average, therefore, the vibrating desorbed ion experiences stronger repulsion than in the frozen state. Consequently, its average (and maximum) kinetic energy must obviously increase, and the EDS ion yield should increase accordingly.

It will be shown below that the parameter governing the vibration effect is $(\Delta l/l_0)^2$. If we now interpret $(\Delta l)^2$ as the mean-square displacement, which is known to be proportional to the temperature T, we have a clear picture of how temperature effects can be taken into account in ESD.

As in Refs. 6 and 7, we consider a dipole desorption model, i.e., we regard the system formed by the charge z_1 and its image in the metal as a point dipole. Then, denoting the equilibrium distance of ion 2 from the surface of the metal by r_0 and denoting its displacement amplitude by Δr, we obtain the following equation in place of (1):

$$r = r_0 + \Delta r \cos\delta. \tag{2}$$

The first stage in the motion of the desorbed ion is described by the equation[6]

$$\dot{r} = \frac{4 z_1 z_2}{r^3} - \frac{z_2^2}{4 r^2} \tag{3}$$

subject to the initial conditions

$$r(0) = r_0, \quad \dot{r}(0) = 0. \tag{4}$$

We use dimensionless units everywhere: All distances are measured in units of the bond length d (Fig. 1), and time is measured in the units $\overline{t} = (M d^3/e^2)^{1/2}$, where M is the mass of the desorbed ion, and e is the positron charge. We can now use all the results of Ref. 6, setting $\theta_0 = 0$ in the equations derived therein, since our adsorption bond is normal to

 1063-7842/97/050555-06$10.00

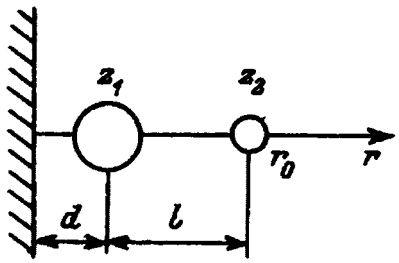

FIG. 1. Model of an adsorption pair.

the substrate surface, and replacing r_0 by r. In particular, for the energy distribution function of the desorbed ions (at low kinetic energies $E \to 0$) we have the following equation in place of Eq. (18) in Ref. 6:

$$f_E \sim \exp\left(-\frac{t_c^*}{\tau} \sqrt{1+(2E/V)} \right), \tag{5}$$

where the critical time t_c^*, i.e., the time required for ion 2 to acquire sufficient kinetic energy in the first desorption stage to surmount the barrier, is

$$t_c^* = r\sqrt{2V/W(2W-V)}, \tag{6}$$

$$W = 4z_1 z_2/r^2, \quad V = z_2^2/2r, \tag{7}$$

where τ is a characteristic lifetime of the two-hole state for ion 1 (it is of the order of 10^{-14} s), z_1 is the charge of ion 1 in the two-hole state, and z_2 is the charge of the desorbed ion.

CALCULATION OF THE YIELD OF DESORBED IONS

The yield of desorbed ions q is given by the expression

$$q = \int_{E_{\min}}^{E_{\max}} f_E dE, \tag{8}$$

where $E_{\min}=0$ at $\theta_0=0$, and $E_{\max}=(W-V)/2$ is dictated by the conservation of energy.[6]

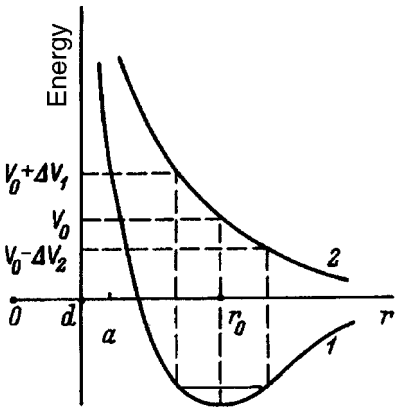

FIG. 2. Influence of vibrations of the adatom on the kinetic energy of electron-stimulated desorption of ions. *1* — Ground state potential; *2* — repulsive desorbing potential, $\delta V=\Delta V_1-\Delta V_2>0$. The origin coincides with the position of ion 1.

The function f_E described by Eq. (5) corresponds to the limit $E \to 0$, whereas in the limit $E \to E_{\max}$ we have a different functional relation [see Eq. (4) in Ref. 7]. However, since the main contribution to the integral (8) is from low-energy ions, we use Eq. (5) over the entire energy range in calculating q. We then readily obtain

$$q \sim V f(\theta,\eta)/\theta^2, \tag{9}$$

where $f(\theta,\eta)=f(\theta)-f(\eta)$, $f(A)=(1+A)\exp(-A)$, where $A=\theta,\eta$, $\theta=t_c^*/\tau$, and $\eta=\theta\sqrt{1+2E_{\max}/V}$.

Since the phase δ, which determines the instantaneous (at time $t=0$) value of r is a random variable, the average yield $\langle q \rangle$ must be defined by the equation

$$\langle q \rangle = \frac{1}{2\pi} \int_0^{2\pi} q\, d\delta. \tag{10}$$

Introducing the small parameter $\xi=\Delta r/r_0$ and evaluating the integral (11), we obtain the following expression in the second order with respect to ξ:

$$\langle q \rangle = q + \Delta q, \tag{11}$$

where q is given by Eq. (9) for $r=r_0$, and the increment Δq is described by the expression

$$\Delta q = \frac{V_0}{2\theta^2}\, \xi^2 [f(\theta)R_1 - f(\eta)R_2], \tag{12}$$

$$R_1 = 1 + 2(a_1+a_2) + 3a_1^2 + \left(a_1 + a_2 + a_1^2\, \frac{\theta+3}{2} \right) \frac{\theta^2}{1+\theta}, \tag{13}$$

$$R_2 = 1 + 2(a_1+a_2) + 3a_1^2 + \left[b_1(1+2a_1) + b_2 + \frac{\eta-1}{2}\, b_1^2 \right] \frac{\eta^2}{1+\eta}, \tag{14}$$

$$a_1 = (5W_0 - 2V_0)/(2W_0 - V_0),$$

$$b_1 = (8W_0 - 3V_0)/2(2W_0 - V_0),$$

$$a_2 = (5W_0^2 - 6W_0V_0 + V_0^2)/2(2W_0 - V_0)^2,$$

$$b_2 = (8W_0^2 - 4W_0V_0 + V_0^2)/8(2W_0 - V_0)^2. \tag{15}$$

All the quantities in Eqs. (11)–(15) are calculated for the equilibrium distance r_0. Defining the temperature coefficient of the EDS ion yield in the form

$$\gamma = \frac{\partial \ln\langle q \rangle}{\partial T}, \tag{16}$$

we obtain

$$\gamma = C\, \frac{\partial \xi^2}{\partial T}, \tag{17}$$

where

$$C = \frac{1}{2}\, \frac{f(\vartheta)R_1 - f(\eta)R_2}{f(\vartheta) - f(\eta)}. \tag{18}$$

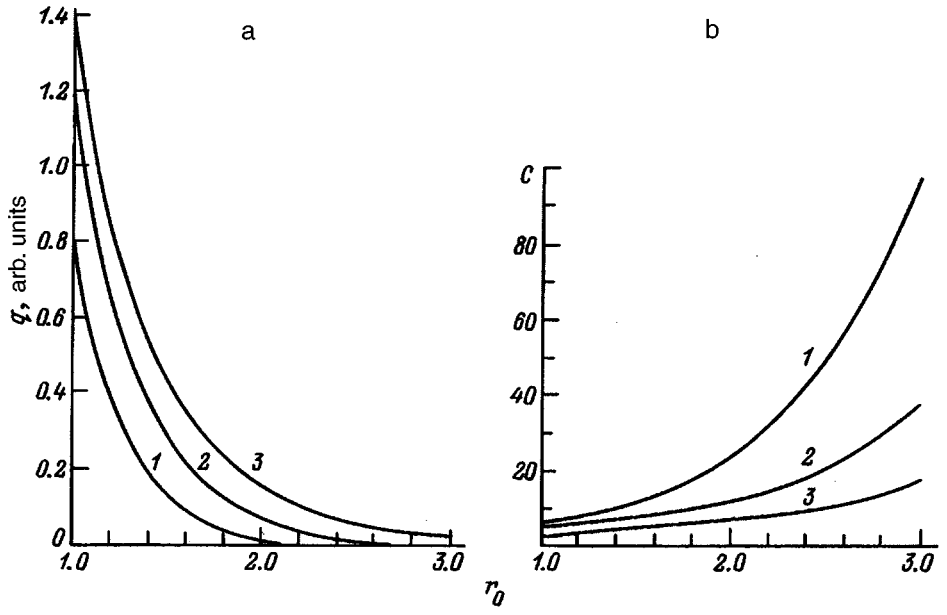

FIG. 3. ESD ion yield q (a) and its temperature coefficient C (b) vs initial position r_0 of the ion, $z_1=z_2=1$. *1* — $\tau=0.5$; *2* — 1; *3* — 2.

DEPENDENCE OF THE ION YIELD OF ELECTRON-STIMULATED DESORPTION AND ITS TEMPERATURE COEFFICIENT ON THE PARAMETERS OF THE MODEL

Figures 3 and 4 shows the ion yield q (in arbitrary units) and the coefficient C as functions of the equilibrium starting position r_0 of the desorbed ion 2 for various lifetimes of the two-hole state τ and as functions of the charge of ion 1. We see that as r_0 increases, the ion yield q decreases (Fig. 3a), and the coefficient C increases (Fig. 3b). The decrease in the yield is attributable to the fact that the repulsive desorbing potential diminishes as r_0 increases (Fig. 2), so that the kinetic energy of the ions decreases, and a smaller number of them can surmount the potential barrier. As τ increases, however, the residence time of ion 2 in the repulsive field of ion 1 increases, and so does the yield. Figure 3b shows that the coefficient γ, in accordance with its definition (16), is inversely proportional to q.

Figure 4 shows q and C as functions of the charge z_1. As z_1 increases, the ESD ion yield increases as a natural consequence of the increased repulsion. The temperature coefficient C decreases in this case. The critical charge of ion 1 has been determined previously:[6]

$$z_{1c}=z_2 r_0/8, \tag{19}$$

and it satisfies the condition that only for $z_1>z_{1c}$ does ion 2 escape to infinity, i.e., become desorbed. If we assume that

$$z_1=z_{1c}(1+\alpha),$$

where $\alpha\ll 1$, we can show that

$$q=\frac{1}{2}\,\alpha V_0 \exp(-R), \tag{20}$$

$$C=\frac{3}{2\alpha}\left(1+\frac{11}{12}\,R\right), \tag{21}$$

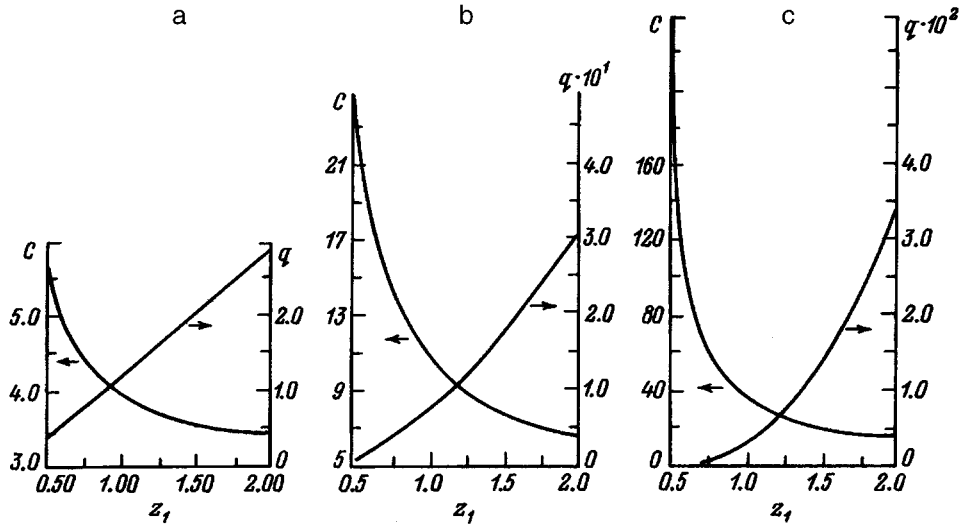

FIG. 4. Graphs of q and C vs charge z_1, $\tau=1$, $z_2=1$; a — $r_0=1$; b — 2; c — 3.

where

$$R=\frac{r_0}{\tau}\sqrt{2/V_0}.$$

Consequently, as $z_1\to z_{1c}$, the ion yield q fades to zero, and the coefficient C tends to infinity. If $z_2=1$, the critical charge for the cases represented in Figs. 4a–4c is equal to 0.125, 0.25, and 0.375, respectively.

The results of the calculations in this section also demonstrate that q and C both have wide ranges of variation (of the order of 10^2) for comparatively small variations of the parameters r_0, z_1, and τ.

ELECTRON-STIMULATED DESORPTION OF LITHIUM AND SODIUM IONS FROM A TUNGSTEN SURFACE COATED WITH A SILICON MONOLAYER

Having examined the general properties of the model, we now turn to the description of a specific system. We consider the desorption of alkali metal (Li^+ and Na^+) ions from a tungsten surface coated with a monolayer of silicon. A detailed investigation of this Si/W system is reported in Ref. 8.

For the calculations we need to determine the model parameters corresponding to the particular system. We begin with the Si/W structure. We interpret the distance d (Fig. 1) as the atomic radius of silicon: 1.17 Å (Ref. 10). We assume that the work function of the substrate is $\varphi=4.65$ eV, which corresponds to the (100) face. The ionization potential of the silicon atom is $I(\mathrm{Si})=8.15$ eV (Ref. 11).

We estimate the charge localized on a silicon atom situated on the tungsten surface. The position of the center of the quasilevel of the adatom relative to vacuum ε' can be calculated from the relation[12]

$$\varepsilon'=-I+3.6/d, \tag{22}$$

where d is measured in Å, and ε' in eV.

We then obtain $\varepsilon'=-5.07$ eV (having assumed that the image plane passes through the point at a distance d from the center of the silicon atom). The position of the adatom quasilevel relative to the Fermi level of tungsten is $\varepsilon=\varepsilon'+\varphi$, i.e., $\varepsilon=-0.42$ eV. According to Anderson's model,[13] the charge on the silicon atom is given by the expression

$$z(\mathrm{Si})=\frac{2}{\pi}\tan^{-1}\frac{\varepsilon}{\Gamma}, \tag{23}$$

where $\Gamma=\pi\rho_s V^2$ is the half-width of the quasilevel.

For the density of states of the substrate ρ_s we adopt Friedel's model[14] $\rho_s=10/W_s$, where W_s is the width of the substrate d band. To estimate the matrix element of interaction of the adatom with the metal V, following the convention of quantum chemistry and the theory of alloys, we assume that $V=\sqrt{t_s t_a}$, where t_s and t_a are the hopping integrals between nearest neighbors in the crystals formed by atoms of the substrate and atoms of the adsorption coating. Allowing for the fact that the width of the band in the strong-coupling approximation is $W=2nt$, where n is the number of nearest neighbors, we finally obtain

$$\Gamma=\frac{5\pi}{2n_s n_a}\,W_a. \tag{24}$$

Since $n_s=8$, $n_a=4$, and $W_a=12.5$ for our system,[15] we obtain $\Gamma=3.06$ eV, which gives $z(\mathrm{Si})=-0.09$.

Another way to estimate the silicon charge is based on the relation between the change of the work function $\Delta\varphi$ in the system and the dipole moment of the adatom p (Refs. 16 and 17):

$$\Delta\varphi=-2\pi eNp, \tag{25}$$

$$p=2edz, \tag{26}$$

where N is the number density of adatoms.

Inasmuch as the change of the work function in the Si/W system is 0.43 eV (Ref. 8) for a silicon monolayer on tungsten ($N=10^{15}$ atoms/cm^2; Ref. 18), we obtain $z(\mathrm{Si})=-0.02$.

We see that the estimates obtained by two altogether different methods give similar results, especially when one considers that adatoms are depolarized in the film,[17,19–23] so that Eq. (23) must give a higher absolute value of the charge than Eq. (24), and indeed it does. Consequently, silicon atoms on tungsten carry a very low negative charge.

We now estimate the charges of sodium and lithium ions adsorbed on Si/W. We stipulate at the outset that the estimates in this case are going to be very crude, first of all because we are looking at a ternary system and, secondly, because we do not know the structure of the silicon film. This consideration is no small matter, because experimental studies of the adsorption of alkali metals on thick silicon substrates[24] have shown that alkali metal atoms are almost completely ionized on the (111) face, where 7×7 structures are observed, whereas charge transfer is very slight for the 2×1 structure on the (100) surface. Moreover, the Li–Si bond parameters vary considerably for the Li/Si(111) system, even for very thin coatings, $\theta\cong0.05$ML (Ref. 25).

These considerations notwithstanding, we attempt to estimate the charges with the aid of expression (23). We regard the Si/W system as a ''monolithic'' substrate with a work function $\varphi(\mathrm{Si/W})=5.08$ eV (Ref. 8). Using the values of the ionization potentials $I(\mathrm{Li})=5.39$ eV and $I(\mathrm{Na})=5.14$ eV (Ref. 11) and interpreting d in Eq. (22) as the ionic radii $r_i(Li^+)$ and $r_i(Na^+)$ with values of 0.68 Å and 0.92 Å, respectively (Ref. 10), i.e., making the image plane coincide with the surface of the silicon film, for the positions of the levels we obtain $\varepsilon(\mathrm{Li})=4.98$ eV and $\varepsilon(\mathrm{Na})=3.85$ eV. Setting the transition integral t_a equal to the matrix element $V_{ss\sigma}=1.32(\hbar^2/mr_i^2)$ (Ref. 26), where m is the electron mass, and $\hbar$ is Planck's constant, we obtain $t_a(\mathrm{Li})=1.10$ eV and $t_a(\mathrm{Na})=0.75$ eV, which give $\Gamma=2.16$ eV and 1.47 eV for lithium and sodium, respectively. It then follows from Eq. (23) that $z(\mathrm{Li})=0.74$ and $z(\mathrm{Na})=0.77$. If we assume that a corresponding number of electrons transfer to the silicon (occupying, e.g., the outward-projecting sp^3 orbital), we can then assume that the charge on the silicon atoms in the alkali metal–Si/W system is negative and has an absolute value of 0.75–0.80. If we now create a two-hole state at the silicon atom (i.e., remove

two electrons from it), the charge of the silicon ions is of the order of 1.20–1.25. This is the value corresponding to the charge z_1 of our model by definition.

Another way to estimate the charges of Li and Na is to treat the silicon monolayer as an insulation spacer. Then in Eq. (22) d must be interpreted as the sum of the ionic radius of the alkali metal and the diameter of the silicon atom, i.e., the image plane must coincide with the tungsten surface. In this case we obtain $\varepsilon(\mathrm{Li})=0.88$ eV and $\varepsilon(\mathrm{Na})=1.04$ eV. Owing to the presence of the insulator film, the levels of the adatom become narrower, $\Gamma=\Gamma_0\exp(-2\zeta h)$, where h is the thickness of the film (diameter of the silicon atom). If we use the values previously obtained for Γ_0 and assume that ζ corresponds to the exponent of the radial wave function of the outer s-electron of the alkali atom, we can use data on ζ from Ref. 27 to show that Γ is more than two orders of magnitude smaller than Γ_0. From Eq. (23) we then obtain $z(\mathrm{Li,Na})\cong 1$ for the charges of lithium and sodium on the Si/W surface. If we assume once again that an electron transfers to the sp^3 orbital of silicon, i.e., assume that the silicon atom is transformed into a singly-charged negative ion, then after the two-hole state is created on the latter we immediately have an Si^+ ion. Consequently, both of these purely theoretical estimates of the silicon charge in the two-hole state give a value of z_1 of the order of 1–1.25.

However, there is also a third way to determine the charge z_1 directly: the ''experimental'' approach. We can use the expression for the maximum energy [see (8)] and experimental data on the delay curves.[8] It follows from Eq. (8) that

$$z_1=\frac{r_0^2}{2z_2}\left(E_{\max}+\frac{z_2^2}{4r_0}\right). \tag{27}$$

Having the atomic radius of silicon as the unit of length, we obtain $r_0(\mathrm{Li})=2.59$ and $r_0(\mathrm{Na})=2.79$; the unit of energy $E^*=e^2/d$ is equal to 12.31 eV. Choosing the experimental value of $E_{\max}$ is not so simple. We need to consider the fact that the high-energy tails of these curves, generally speaking, are not described by our theory, because the role of vibrations of the adsorption pair is taken into account only in the average, whereas the tails are associated with maximum contraction of the adsorption bond. Using the data from Ref. 8 (see Figs. 3 and 4 for the case of 300 eV bombarding electrons), we set $E_{\max}(\mathrm{Li})=3$ eV. From expression (27) we then find $z_1=1.13$. Adopting the energy $E_{\max}(\mathrm{Na})=4$ eV for sodium, we obtain $z_1=1.61$, which is considerably higher than our estimate. But if, as in Ref. 8, we assume that $E_{\max}(\mathrm{Na})=5$ eV, we obtain $z_1=2$. We can only assume, therefore, that we are dealing with a three-hole state. Ageev *et al.*[8] have arrived at the same conclusion on the basis of an analysis of the desorption thresholds and delay curves (the delay curve for Na has a kink and radically differs from the corresponding curve for Li).[1)] The theoretical estimates and estimates based on experimental data both lead to similar results.

All that remains is to estimate the parameter ξ and its temperature dependence. If we set $(\Delta r)^2$ equal to the mean-square displacement, in the high-temperature limit we have[28,29]

$$(\Delta r)^2=3(\hbar^2T/M\Theta^2k_B), \tag{28}$$

where Θ is the Debye temperature, and k_B is the Boltzmann constant.

However, for an adsorption system where such a parameter as Θ is not defined, it is more convenient to use the force constant of the adsorption bond

$$\kappa=(Mk_B^2\Theta^2/\hbar^2), \tag{29}$$

where M must be interpreted as the normalized mass of the Li, Na–Si/W molecule. To estimate the constant κ for vibrations corresponding to variation of the bond length, it is practical to use the approach developed in Refs. 30 and 31, setting

$$\kappa=E_{\mathrm{des}}/l^2, \tag{30}$$

where E_{des} is the desorption energy, and l is a characteristic length of variation of E_{des}, which we assume to be equal to r_0. Now

$$\xi^2=3(k_BT/E_{\mathrm{des}}). \tag{31}$$

According to experimental data, the desorption energy of alkali metals coated with a silicon monolayer on tungsten differs by no more than ten or twenty electron-volts from their energy of desorption from pure tungsten; we therefore set $E_{\mathrm{des}}(\mathrm{Li})=2.98$ and $E_{\mathrm{des}}(\mathrm{Na})=2.53$ (Refs. 32–34). Then $\partial\xi^2/\partial T=0.87\times10^{-4}$ K^{-1} and 0.98×10^{-4} K^{-1} for lithium and sodium ions, respectively.

For numerical calculations we also need to know the lifetime τ of the hole states. The simplest way to estimate it is from the experimental delay curves. It can be shown that the logarithm of the slope of the delay curve is proportional to the quantity $\vartheta=t_c^*/\tau$. If such an estimate is made by setting the charge of the silicon ion $z_1=1$ for the desorption of Li^+ ions and $z_1=2$ for Na^+, we obtain $\tau=4.84\times10^{-14}$ s and $\tau=1.43\times10^{-14}$ s, respectively.

The calculation of γ using these values of τ gives $\gamma=7.35\times10^{-4}$ K^{-1} and 7.63×10^{-4} K^{-1} for lithium and sodium, respectively. In the case of sodium this value essentially coincides with the experimental, whereas for lithium the experimental value is an order of magnitude lower. The reasons for such a large discrepancy in the latter case are attributable, first, to purely computational instability stemming from the fact that the numerator and denominator in Eq. (18) are small quantities, and the small numerator represents a difference between two large numbers. Second, the averaged description used here for the vibrations of the adsorption pair should doubtless be replaced by a more rigorous analysis for the lighter lithium.

We have thus shown that the temperature dependence of the yield of desorbed ions can be rationally described on the basis of a relaxation model. However, a certain reservation must be interjected. The results of our calculations are in good agreement with the experimental[8] for sodium ions desorbed in the irradiation of the Na–Si/W system with 300 eV electrons. But when the same system is irradiated with 100 eV electrons, the coefficient γ increases approximately threefold. An analysis of the desorption process in this case has brought Ageev *et al.*[8] to the conclusion that this behavior is a

consequence of the electron density becoming redistributed in the adsorption pair in such a way as to reverse the motion of the desorbed ion at the initial instant of desorption. Our model does not take this consideration into account. We intend in the future to generalize the approach developed here to account for the possibility of the reversal effect.

The author is indebted to V. N. Ageev, N. D. Potekhina, and B. V. Yakshinskiĭ for a discussion of the results and for valuable comments.

[1)]I am indebted to I. D. Potekhina for bringing this work to my attention.

[1] V. N. Ageev, O. P. Burmistrova, and Yu. Kuznetsov, Usp. Fiz. Nauk **158**, 385 (1989) [Sov. Phys. Usp. **32**, 588 (1989)].
[2] R. D. Ramsier and J. T. Yates, Surf. Sci. Rep. **12**, 244 (1991).
[3] V. N. Ageev and B. V. Yakshinskiĭ, Fiz. Tverd. Tela (Leningrad) **27**, 99 (1985) [Sov. Phys. Solid State **27**, 57 (1985)].
[4] V. N. Ageev, O. P. Burmistrova, A. I. Gubanov *et al.*, Poverkhnost', No. 12, 30 (1986).
[5] V. N. Ageev, O. P. Burmistrova, and B. V. Yakshinskii, Surf. Sci. **194**, 101 (1988).
[6] S. Yu. Davydov, Fiz. Tverd. Tela (St. Petersburg) **35**, 2524 (1993) [Phys. Solid State **35**, 1251 (1993)].
[7] S. Yu. Davydov, Fiz. Tverd. Tela (St. Petersburg) **37**, 1758 (1995) [Phys. Solid State **37**, 957 (1995)].
[8] V. N. Ageev, O. P. Burmistrova, A. M. Magomedov, and B. V. Yakshinskiĭ, Fiz. Tverd. Tela (Leningrad) **32**, 801 (1990) [Sov. Phys. Solid State **32**, 472 (1990)].
[9] V. N. Ageev, A. M. Magomedov, and B. V. Yakshinskiĭ, Fiz. Tverd. Tela (Leningrad) **33**, 158 (1991) [Sov. Phys. Solid State **33**, 91 (1991)].
[10] C. Kittel, *Introduction to Solid State Physics*, 6th ed., Wiley, New York (1986) [Russ. original, Nauka, Moscow (1978)].
[11] I. S. Grigor'ev and E. Z. Meĭlikhov (eds.), *Handbook of Physical Quantities* [in Russian], Énergoatomizdat, Moscow (1991).
[12] J. W. Gadzuk, Phys. Rev. B **1**, 2110 (1970).
[13] P. W. Anderson, Phys. Rev. **124**, 41 (1961).
[14] J. Friedel, in *The Physics of Metals*, J. M. Ziman (ed.), Cambridge Univ. Press, London (1969), Chap. 8.
[15] W. A. Harrison, *Electronic Structure and the Properties of Solids: The Physics of the Chemical Bond*, W. H. Freeman, San Francisco (1980) [Russ. original, Mir, Moscow (1983)].
[16] J. P. Muscast and D. M. Newns, J. Phys. C **7**, 2630 (1974).
[17] S. Yu. Davydov, Fiz. Tverd. Tela (Leningrad) **19**, 3376 (1977) [Sov. Phys. Solid State **19**, 1971 (1977)].
[18] V. N. Ageev and E. Yu. Afanas'eva, Poverkhnost', No. 7, 30 (1987).
[19] S. Yu. Davydov, Fiz. Tverd. Tela (Leningrad) **20**, 1752 (1978) [Sov. Phys. Solid State **20**, 1013 (1978)].
[20] S. Yu. Davydov, Fiz. Met. Metalloved. **47**, 481.
[21] S. Yu. Davydov, Fiz. Tverd. Tela (Leningrad) **20**, 1998 (1978) [Sov. Phys. Solid State **20**, 1153 (1978)].
[22] S. Yu. Davydov and V. I. Margolin, Fiz. Tverd. Tela (Leningrad) **24**, 1503 (1982) [Sov. Phys. Solid State **24**, 859 (1982)].
[23] S. Yu. Davydov, Poverkhnost', No. 8, 17 (1991).
[24] T. Hashizume, Y. Hasegawa, and T. Sakurai, Appl. Surf. Sci. **48–49**, 119 (1991).
[25] J. Chorst and D. Fisk, Surf. Sci. **251/252**, 18 (1991).
[26] W. A. Harrison, Phys. Rev. B **27**, 3592 (1983).
[27] A. A. Radtsig and B. M. Smirnov, *Handbook of Parameters of Atoms and Atomic Ions* [in Russian], Énergoatomizdat, Moscow (1986).
[28] M. Nishijima and F. M. Propst, Phys. Rev. B **2**, 2368 (1970).
[29] H. Böttger, *Principles of the Theory of Lattice Dynamics*, Physik-Verlag, Weinheim (1983) [Russ. Transl., Mir, Moscow (1986)].
[30] J. H. Rose, J. R. Smith, and J. Ferrante, Phys. Rev. B **28**, 1835 (1984).
[31] J. H. Rose, J. R. Smith, F. Guinea, and J. Ferrante, Phys. Rev. B **29**, 2963 (1984).
[32] É. G. Nazarov and U. Kh. Rasulev, *Transient Surface Ionization Processes* [in Russian], Fan, Tashkent (1991).
[33] V. N. Ageev and B. V. Yakshinskiĭ, Fiz. Tverd. Tela (Leningrad) **27**, 99 (1985) [Sov. Phys. Solid State **27**, 57 (1985)].
[34] P. R. Antoniewicz, Phys. Rev. B **21**, 3811 (1980).

Translated by James S. Wood

Self-excited oscillations in a parabolic system with nonlinear external feedback

A. S. Rudyĭ

State University, 150000 Yaroslavl, Russia
(Submitted March 14, 1995; resubmitted August 30, 1995)
Zh. Tekh. Fiz. **67**, 116–118 (May 1997)

[S1063-7842(97)01605-X]

Temperature is one of the most difficult thermodynamic parameters to control and at the same time is one of the most important. Patent difficulties of a general sort are encountered in regulating the temperature of a particular object, whether an industrial machine or an integrated microcircuit. One problem is that when the feedback ratio is increased, as is necessary in order to improve the precision of active temperature control, the temperature stabilization system can become unstable. We are therefore faced with the problem of determining the limit of stability of the temperature stabilizing and regulating system. In this article we investigate Hopf bifurcation in the class of proportional temperature control systems.

1. Regardless of the structure or function of such systems, they can always be conditionally partitioned into a nonlinear external feedback circuit with lumped parameters and a distributed controlled object (Fig. 1). A mathematical model of such a system can be constructed in the form of a thermal conduction boundary-value problem specified in dimensionless variables with a nonlocal boundary condition:[1]

$$\dot{T}(x,t)=T''(x,t), \quad T(0,t)=0, \tag{1}$$

$$T'(x,t)|_{x=1}=f[1-T(x_0,t)]\sigma[1-T(x_0,t)], \tag{2}$$

where $x_0\in[0,1]$, f is a certain smooth function of the temperature, and σ is the Heaviside unit step function, which cuts off the positive feedback branch from f.

The special case of a quadratic nonlinearity of the form $f=[1-T(0,5;t)]^2$ has already been investigated.[1] Since f can contain cubic terms, which influence the Hopf bifurcation regime (bifurcation of limit cycles), it is necessary to investigate the more general case (2) of nonlinear feedback.

2. We consider steady-state, periodic solutions of the boundary-value problem (1), (2). Its steady-state solution has the form $\overline{T(x)}=Cx$, where C is the root of the nonlinear equation

$$C=f(1-Cx_0). \tag{3}$$

The function f is positive concave over the entire domain of definition for nearly all proportional temperature control systems:

$$\left.\begin{aligned} f[1-T(x_0,t)]>0 \\ \partial f/\partial C\equiv A<0 \\ \partial^2 f/\partial T^2>0 \end{aligned}\right\} \forall T\in[0,1],$$

such that $f'(0)=0$. By virtue of these conditions Eq. (3) has a unique root, for which $A<0$ and the corresponding steady-state solution is stable.

3. Let us assume that the steady-state solution $\overline{T}=Cx$ acquires self-excited oscillations $T(x,t)=\overline{T}(x)+u(x,t)$ such that $|u(x_0,t)|\ll|\overline{T}(x_0)|$. The function f can then be expanded in a Taylor series in powers of the unsteady temperature $u(x_0,t)$:

$$f[1-\overline{T}(x_0)-u(x_0,t)] = C+\frac{A}{x_0}u(x_0,t)+f''\frac{u^2(x_0,t)}{2!}+f'''\frac{u^3(x_0,t)}{3!}\ldots+, \tag{4}$$

where $f^{(n)}$ is the nth derivative at $T=\overline{T}(x_0)$.

4. Next we consider the boundary-volume problem linearized in the steady state

$$\dot{u}(x,t)=u''(x,t),$$

$$u(0,t)=0,$$

$$u'(x,t)|_{x=1}=\frac{A}{x_0}u(x_0,t). \tag{5}$$

Its periodic solution is the function

$$u(x,t)=\xi\sinh(\sqrt{i\omega x})\exp(i\omega t) \tag{6}$$

and its complex conjugate. The condition imposed by feedback on the complex amplitude of the temperature oscillations at the points x_0 can be used to determine the critical value $A_c=A(\omega_c)$:

$$A_c=\frac{\mathrm{Re}\sqrt{i\omega_c}x_0\cosh\sqrt{i\omega_c}}{\sinh\sqrt{i\omega_c}x_0}, \tag{7}$$

where ω_c is that root of the equation

$$\frac{\cosh\sqrt{\omega/2}\cos\sqrt{\omega/2}-\sinh\sqrt{\omega/2}\sin\sqrt{\omega/2}}{\sinh(x_0\sqrt{\omega/2})\cos(x_0\sqrt{\omega/2})} = \frac{\cosh\sqrt{\omega/2}\cos\sqrt{\omega/2}+\sinh\sqrt{\omega/2}\sin\sqrt{\omega/2}}{\cosh(x_0\sqrt{\omega/2})\sin(x_0\sqrt{\omega/2})}, \tag{8}$$

for which $A(\omega_c)>A(\omega_n)$ for any n.

5. For $A>A_c$ the null equilibrium state of the boundary-value problems (1), (2) is locally exponentially stable, and for $A<A_c$ it loses stability by oscillating. Let

 1063-7842/97/050561-03$10.00

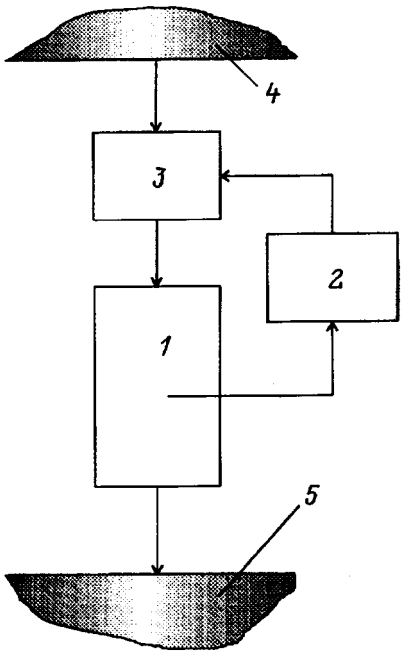

FIG. 1. Diagram of a temperature stabilization system. *1* — Distributed controlled object; *2* — feedback loop; *3* — controller; *4* — energy source; *5* — thermostat.

$$A=A_c(1+\varepsilon), \quad 0<\varepsilon\ll 1, \tag{9}$$

so that under the constraints (9) the Andronov–Hopf bifurcation theorem[2] is valid for proving the existence and stability of self-excited oscillations for the nonlinear boundary-value problem (1), (2).

We introduce the notation

$$\Theta(x,t)=\frac{f''}{2}u(x,t); \quad \gamma=\frac{f'''}{(f'')^2} \tag{10}$$

and transform the boundary-value problem as follows:

$$\dot{\Theta}(x,t)=\Theta''(x,t), \quad \Theta(0,t)=0,$$

$$\Theta'(x,t)|_{x=1}=\frac{A_c(1+\varepsilon)}{x_0}\Theta(x_0,t)+\Theta^2(x_0,t)+\frac{4}{3}\gamma\Theta^3(x_0,t). \tag{11}$$

Setting

$$t=(1+c)\tau; \quad c=c_2\xi^2+c_4\xi^4+\ldots;$$

$$\varepsilon=b_2\xi^2+b_4\xi^4+\ldots,$$

$$\Theta(x,\tau)=\xi\Theta_1(x,\tau)+\xi^2\Theta_2(x,\tau)+\ldots, \tag{12}$$

we obtain a recursive sequence of nonlinear, inhomogeneous, boundary-value problems for the determination of $\Theta_1,\Theta_2,\ldots$:

$$\dot{\Theta}_1(x,\tau)=\Theta_1''(x,\tau), \quad \Theta_1(0,\tau)=0, \tag{13}$$

$$\Theta_1'(x,\tau)|_{x=1}=\frac{A_c}{x_0}\Theta_1(x_0,\tau),$$

$$\dot{\Theta}_2(x,\tau)=\Theta_2''(x,\tau), \quad \Theta_2(0,\tau)=0,$$

$$\Theta_2'(x,\tau)|_{x=1}=\frac{A_c}{x_0}\Theta_2(x_0,\tau)+\Theta_1^2(x_0,\tau); \tag{14}$$

$$\dot{\Theta}_3(x,\tau)=\Theta_3''(x,\tau)+c_2\Theta_1''(x,\tau), \quad \Theta_3(0,\tau)=0,$$

$$\Theta_3'(x,\tau)|_{x=1}=\frac{A_c}{x_0}[\Theta_3(x_0,\tau)+b_2\Theta_1(x_0,\tau)]+2\Theta_1(x_0,\tau)\Theta_2(x_0,\tau)+\frac{4}{3}\gamma\Theta_1^3;\ldots, \tag{15}$$

the first of which coincides with problem (5). The solution of problem (14) has the form

$$\Theta_2=2\frac{|V_1(x_0)|^2}{1-A_c}x+V_2(x)e^{i2\omega\tau}+V_2^*(x)e^{-i2\omega\tau}, \tag{16}$$

where

$$V_2(x)=\frac{V_1^2(x_0)}{\sqrt{i2\omega}\cosh\sqrt{i2\omega}-(A_c/x_0)\sinh\sqrt{i2\omega}x_0}\times\sinh\sqrt{i2\omega}x, \tag{17}$$

and $V_1(x)$ is the solution of the Sturm-Liouville problem for the boundary-value problem (5).

From the solvability condition

$$\frac{c_2}{2}\left\{i\omega\frac{V_1(1)}{V_1(x_0)}-\frac{A_c}{x_0}\left[x_0\frac{V_1'(1)}{V_1(x_0)}-1\right]\right\}+\frac{A_c}{x_0}b_2+|V_1(x_0)|^2\left[\frac{2}{z}+4\left(\gamma+\frac{x_0}{1-A_c}\right)\right]=0 \tag{18}$$

for the third problem in the sequence we find the Lyapunov coefficients

$$c_2=4\frac{x_0}{A_c}|V_1(x_0)|^2\frac{z_2}{p_2|z|^2},$$

$$b_2=-\frac{x_0}{A_c}|V_1(x_0)|^2\left[\frac{2}{|z|^2}\left(z_1+\frac{p_1}{p_2}z_2\right)+4\frac{x_0}{1-A_c}\right]-4\frac{x_0}{A_c}|V_1(x_0)|^2\gamma, \tag{19}$$

where

$$z=z_1+iz_2=\frac{\sqrt{i2\omega}\cosh\sqrt{i2\omega}}{\sinh\sqrt{i2\omega}x_0}-\frac{A_c}{x_0},$$

$$p_1+ip_2=1+\frac{x_0}{|V_1(x_0)|^2}\left[\frac{i\omega}{A_c}V_1(1)V_1^*(x_0)-V_1'(1)V_1^*(x_0)\right]. \tag{20}$$

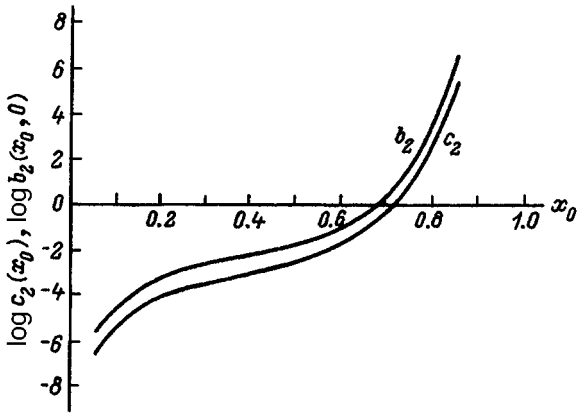

FIG. 2. Lyapunov coefficients $c_2(x_0)$ and $b_2(x_0,0)$ versus x_0 for a system with quadratic nonlinearity ($\gamma=0$).

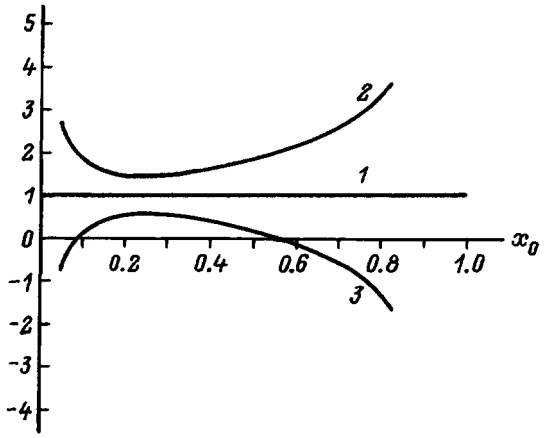

FIG. 3. Lyapunov coefficient $b_2(x_0,\gamma)$ normalized to $b_2(x_0,0)$ versus x_0. *1* — $\gamma=0$; *2* — $\gamma=0.01$; *3* — $\gamma=-0.01$.

Substituting Θ_1, Θ_2, and $\xi=\sqrt{\varepsilon/b_2}$ into $u(x,t)=(2/f'')[\xi\Theta_1(x,\tau)+\xi^2\Theta_2(x,\tau)]$, we obtain the periodic solution of problem (1), (2) in the final form

$$u(x,t)=\frac{2}{f''}\left\{\sqrt{\frac{\varepsilon}{b_2}}[V_1(x)e^{i\omega(\varepsilon)t}+V_1^*(x)e^{-i\omega(\varepsilon)t}]+\frac{\varepsilon}{b_2}\frac{|V_1|^2}{1-A_c}x+\frac{\varepsilon}{b_2}[V_2(x)e^{i2\omega(\varepsilon)t}+V_2^*(x)e^{-i2\omega(\varepsilon)t}]\right\},$$

$$\omega(\varepsilon)=\frac{\omega_c}{1+(c_2/b_2)\varepsilon}.^{1)} \qquad (21)$$

The latter relation is valid only when the coefficient b_2 is positive and soft Hopf bifurcation takes place, i.e., when the oscillation amplitude tends to zero in the limit $\varepsilon\to 0$. Graphs of the functions $c_2(x_0)$ and $b_2(x_0,0)$ are shown in Fig. 2. It is evident from Fig. 2 and the second relation (19) that the parameter b_2 is always positive for $f'''>0$, whereas for $f'''<0$ the sign of b_2 can be either positive or negative, depending on γ and x_0 (Fig. 3). This means that both soft and hard Hopf bifurcations are possible. Finally, for $f''=0$ the system is linear, and the oscillation amplitude in the supercritical region is infinite.

The results of our nonlinear analysis are in good agreement with general notions concerning nonlinearity as a factor limiting the amplitude of self-excited oscillations due to the redistribution of energy between the first and higher harmonics. From this point of view, the role of quadratic nonlinearity is purely ''dissipative,'' so that without cubic terms in the expansion (4) self-excited oscillations in a parabolic system always occur in the soft regime (Fig. 2). Cubic nonlinearity is more ''conservative'' since part of the energy returns in the form of oscillations in the fundamental mode. If the coefficient of the cubic nonlinearity is positive, i.e., if A_c and f''' have opposite signs, this additional energy is out-of-phase with the fundamental mode, diminishing the oscillation amplitude. Self-excited oscillations occur in the soft regime in this case as well [Fig. 3, curve *2*: $b_2(x_0,\gamma)>b_2(x_0,0)$ and $\xi(x_0,\gamma)<\xi(x_0,0)$]. But if $f'''<0$, the energy arrives in phase with the first harmonic, and its amplitude increases, $\xi(x_0,\gamma)>\xi(x_0,0)$ (Fig. 3, curve *3*). For a certain relation between the derivatives and the value of x_0 the recovery of energy by cubic nonlinearity exceeds its dissipation within one period as long as the oscillation amplitude remains small. Energy balance sets in when the amplitude of the self-excited oscillations attains a certain (not small) level. This self-excitation regime, which is said to be hard, corresponds to the range of negative values of the parameter b_2 on curve *3* in Fig. 3.

By far the majority of temperature control systems utilize Joule heat sources, for which f must be a quadratic function of the temperature. In reality, however, f contains several temperature-independent parameters, which introduce cubic terms in (4). For this reason, the behavior of systems with an identical temperature control scheme in the supercritical region can differ significantly. The results of the present study provide a straightforward explanation of this phenomenon.

[1)]The quantities $\omega(\varepsilon)$ and ω_c are mistakenly switched in Ref. 1.

[1] A. S. Rudy, Int. J. Thermophys. **14**, 159 (1993).
[2] Yu. S. Kolesov and D. I. Shvitra, *Self-Excited Oscillations in Systems with Delay* [in Russian], Mokslas, Vilnius (1979).

Translated by James S. Wood

Thermoconductometric detection of gases and gas flows by means of SAW delay lines

V. I. Anisimkin

Institute of Radio Engineering and Electronics, Russian Academy of Sciences, 103907 Moscow, Russia

S. A. Maksimov

Physicotechnical Institute, 141700 Dolgoprudnyi, Moscow Region, Russia

M. Penza

Centro Nazionale per la Ricerca e lo Sviluppo dei Materiali, Brindisi, Italy

L. Vasanelli

Istituto per lo Studio dei Nuovi Materiali per l'Elettronica del Consiglio Nazionale delle Ricerche and Università di Lecce, Italy

(Submitted September 21, 1995)
Zh. Tekh. Fiz. **67**, 119–123 (May 1997)

Surface acoustic wave (SAW) delay lines without gas-sensitive coatings are used as thermal sensors for the thermoconductometric detection of gases and gas flows. The forced convection of 13 gases is analyzed in the linear approximation without regard for their interaction with the environment. Quartz, $LiNbO_3$, $Bi_{12}GeO_{20}$, and $Bi_{12}SiO_{20}$ delay lines are used to detect H_2, He, Ar, CH_4, NH_3, N_2, and O_2 at frequencies $f=21$–263 MHz and temperatures $T=25$–165 °C. The SAW ''response'' is measured as a function of the gas concentration n, the flow rate U, the temperature coefficient of the SAW velocity (TCV), and the working temperature T_p. The feasibility of controlling the level of the gas ''response'' and imparting selectivity to the choice of TCV and T_p is demonstrated. The threshold gas concentrations are 0.35% for CH_4 and 0.1% for H_2 and NH_3 in nitrogen. A linear response is obtained in the interval $U=20$–200 ml/min.
[S1063-7842(97)01705-4]

INTRODUCTION

Because of their high-temperature sensitivity, surface acoustic wave (SAW) devices are well suited to temperature measurements of other physical parameters such as the velocities of gas and liquid flows, dew points, etc.[1,2]

In the present study SAW delay lines without gas-sensitive coatings are used to measure gas concentrations and velocities of gas flows. The characteristics of the detection technique are analyzed, and laboratory prototypes of sensors using inert (He, Ar, N_2) and other (H_2, CH_4, NH_3) gases are tested.

THEORETICAL ANALYSIS

A SAW device placed in a laminar gas flow and heated above ambient temperature can be regarded as a flat plate connected to a constant heat source. The change in temperature of the plate ΔT due to a change in the ambient thermal conductivity $\Delta\lambda$ or a change in the gas flow rate ΔU (or both at the same time) is detected at the output of the delay line as a change in phase $\Delta\Phi$ of the SAW signal.

An analysis of sources of heat losses[2–5] shows that the increment ΔT is proportional to the relative variation of the forced convection coefficient $-\Delta h_f/h_f$, which is given by the expression[4]

$$h_f=0.6795\lambda(\mu C_p/\lambda)^{1/3}(\rho U/I\mu)^{1/2}. \tag{1}$$

Here L is the length of the plate, and λ, μ, C_p, and ρ are the thermal conductivity, dynamic viscosity, specific heat at constant pressure, and density of the gas. Since the Prandtl number contained in the first parentheses of Eq. (1) is approximately the same for different gases,[5] the dependence of the quantity $-\Delta h_f/h_f$ on the gas constants can be written as follows in the linear approximation:

$$-\Delta h_f/h_f=-\Delta\lambda/\lambda-0.5(\Delta U/U)-0.5\times(\Delta\rho/\rho-\Delta\mu/\mu). \tag{2}$$

It is evident from Eq. (2) that the changes in the density and viscosity of the environment after the injection of a gas produce opposite effects. Moreover, the quantities $-\Delta h_f/h_f$ and, hence, ΔT are positive (heating of the substrate) for negative values of $\Delta\lambda$ and ΔU (e.g., when argon is injected into air or when the flow rate decreases), whereas both of these quantities become negative (cooling of the substrate) for positive increments $\Delta\lambda$ and ΔU (e.g., when helium is injected into air or when the flow rate increases).

An analysis of Eq. (2) for the gases listed in Table I relative to air shows that they can be classified into three groups: 1) H_2, He, Ar, and Ch_4, for which the modulus $|-\Delta h_f/h_f|$ is large, the term $0.5(\Delta\rho/\rho-\Delta\mu/\mu)$ is small, and the contribution of $\Delta\lambda/\lambda$ to $|-\Delta h_f/h_f|$ is dominant; gases of this kind can be identified only through the heat conduction effect, provided that the tested and calibrated gas (air) propagates at the same rate ($\Delta U=0$); on the other

TABLE I. Physical parameters of the gases (at 20°).

Gas	ρ, kg/m^3	C_p, 10^3 J/kg·K	λ, 10^{-2} W/m·K	μ, 10^{-7} kg/m·s
Dry air	1.2928	0.992	2.41	181.
N_2	1.2505	1.038	2.43	174.
NH_3	0.7714	2.244	2.18	97.
Ar	1.7839	0.523	1.62	222.
H_2	0.08988	14.273	16.84	88.
He	0.1785	5.238	14.15	194.
N_2O	1.9775	0.946	1.51	146.
O_2	1.42896	0.909	2.44	200.
CH_4	0.7168	2.483	3.02	109.
NO	1.3402	0.967	2.38	188.
CO	1.25	1.038	2.32	177.
CO_2	1.9768	0.846	1.45	144.8
Cl_2	3.22	0.519	0.72	132.

hand, the convective influence of gases in the first group differ from each other and also from air; 2) N_2, NH_3, N_2O, NO, CO, CO_2, H_2O, and O_2, for which the contributions of all terms except $\Delta U/U$ to $|-\Delta h_f/h_f|$ are small; these gasses cannot be identified through the $\Delta\lambda$ effect, but their response to the motion of the gas flow (convection) is approximately the same as for air; 3) Cl_2, for which all terms in (2) are mutually commensurate, and $|-\Delta h_f/h_f|$ is large; this gas cannot be identified through the $\Delta\lambda$ effect, and its convective response also differs from that of air.

These estimates have been made for 20°C (see Table I) and, of course, will differ in character at other temperatures.

In the analysis we have disregarded any kind of chemical interaction between the tested gas and the environment that could produce new gaseous substances. If that happens, Eq. (2) must be applied to each of these substances separately, and the tested gas can be detected through the influence of their heat conduction. In our experiments (see below), for example, the SAW response was very strong in the detection of 0.1% NH_3 in nitrogen, whereas Eq. (2) dictates that such a response is impossible even for 100% NH_3. The result can be explained in connection with the interaction of ammonia with environmental water vapor and the formation of gases, which greatly enhances the heat conduction effect.

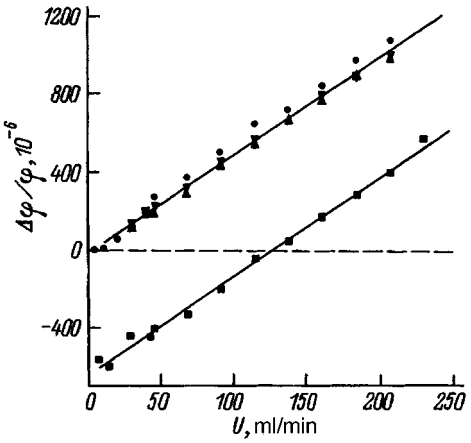

FIG. 2. SAW response $\Delta\phi/\phi$ versus gas flow rate U for a 128° YX-$LiNBO_3$ substrate, T_p=120 °C. ●—N_2; ▼—dry air; ▲—O_2; ■—Ar.

EXPERIMENTAL PROCEDURE

The measurements were performed at temperatures of 20–165°C, atmospheric pressure, and frequencies f=21–263 MHz. Ordinary SAW delay lines with AT- and X-SiO_2, YZ- and 128°Y, X-$LiNbO_3$, (001),⟨110⟩-$Bi_{12}GeO_{20}$, and (001),⟨110⟩-$Bi_{12}SiO_{20}$ substrates were used together with film heaters (Fig. 1). The devices had typical dimensions of 0.5×8×10 mm, the tested gases were H_2, He, Ar, CH_4, NH_3, N_2, O_2, N_2, and dry air, the gas concentrations were 0.1–100%, and the flow velocities were 20–200 ml/min.

The tested device was placed in a thermally insulated chamber capable of maintaining a stable temperature within ±0.5 °C limits for 5 h. The substrate was prevented from coming into direct contact with the chamber. A gas flowmeter (Lab-Box 1 Dimin 1204, Omicron Technologies) was placed between the gas source and the chamber so that the flow rate could be varied within 5-ml/min error limits. The substrate was oriented in the chamber perpendicular to the gas flow and was heated by a film heater (100 Ω/□) from a

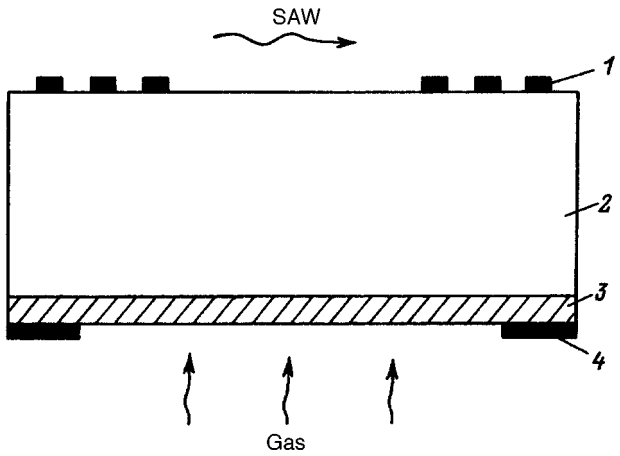

FIG. 1. Thermoconductometric SAW gas-sensitive element. *1*—Interdigital transducer; *2*—piezoelectric substrate; *3*—film heater; *4*—aluminum electrodes.

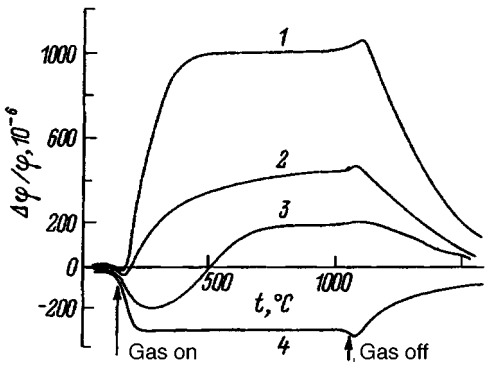

FIG. 3. Time profiles of the SAW response to flows of various gases for a 128° YX-$LiNbO_3$ substrate, T_p=120 °C. *1*—N_2, U=210 ml/min; *2*—N_2, U=115 ml/min; *3*—Ar, U=160 ml/min; *4*—Ar, U=70 ml/min.

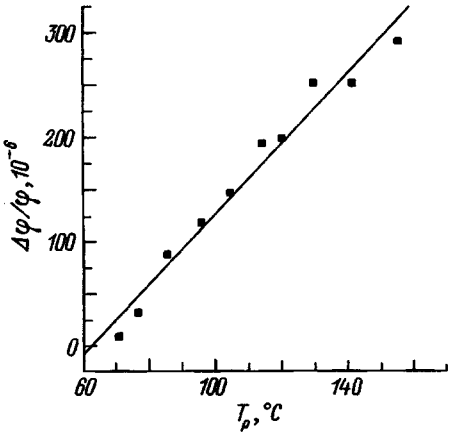

FIG. 4. SAW response to dry air flow versus temperature for a 128° YX-$LiNbO_3$ substrate, $U=50$ ml/min.

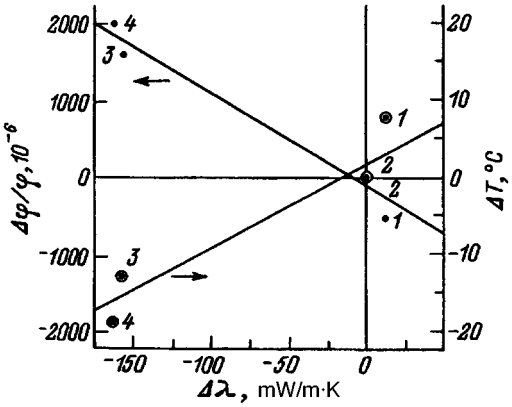

FIG. 6. SAW response $\Delta\phi/\phi$ and substrate temperature increment ΔT versus difference $\Delta\lambda$ between the thermal conductivities of the calibration gas (λ_{cg}) and the tested gas (λ_{tg} for a 128° YX-$LiNbO_3$ substrate, $T_p=120$ °C. Calibration gases: *1–3*—100% N_2; *4*—100% Ar. Tested gases: *1*—100% Ar; *2*—100% O_2; *3*—100% He; *4*—100% He. The solid lines are plotted by least-squares processing of the experimental data.

stabilized source (10 V, 0.1 A). The variation of the electrical resistance of the heater under the influence of the gases and temperature fluctuations was negligible.

For the gas flow rate measurements the tested gas was injected at a rate $U_{tg} \neq 0$ into the chamber with the air at rest ($U_a=0$).

For the gas concentration measurements the gas was injected into the chamber after the preliminary injection of a calibration gas (N_2) at the same rate ($U_{tg}=U_{cg}$). Before the measurements were begun, the substrate was heated to the working temperature and placed in a stream of N_2 in an effort to fully stabilize the phase of the output SAW signal.

The SAW phase angle ϕ was measured by means of an HP 8753 C instrument within $\pm 10\%$ error limits. The noise level was of the order of $\Delta\phi=2°$ ($\Delta\phi/\phi=2$ ppm). A thermocouple was used to monitor the working temperature T_p on the substrate surface.

GAS FLOW RATE MEASUREMENTS

The measurement results in Figs. 2 and 3 show the dependence of the SAW response on the flow rate U and the thermal conductivity λ of the tested gas. The calibration curves in Fig. 2 are close together for gases with values of λ approaching that of air (N_2, O_2, and dry air). The SAW response deviates considerably for Ar with its lower thermal conductivity; it is negative for low velocities and positive for high velocities U. It is evident from the parallelism of all the curves in Fig. 2 that the difference between the response level for Ar and the other gases does not depend on their injection rate U.

On the other hand, the profile of the SAW response for Ar does depend on U (Fig. 3). At $U<100$ ml/min the degree of convection cooling of the substrate [the second term in Eq. (2)] is lower than its degree of heating due to the decrease in the ambient thermal conductivity after the introduction of Ar [first term in (2)]. Consequently, at low rates U the temperature of the substrate increases, and the response becomes negative (Fig. 3). At $U>150$ ml/min the convection

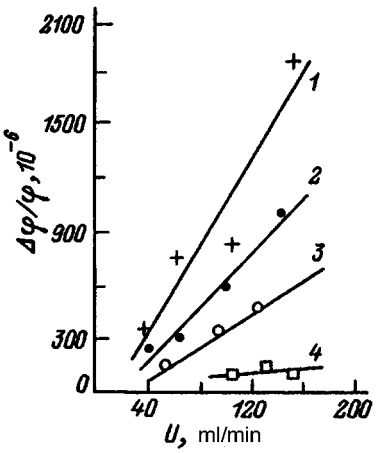

FIG. 5. SAW response versus flow rate of oxygen for various substrate materials ($T_p=120$ °C). *1*—(001),⟨110⟩-$Bi_{12}SiO_{20}$; *2*—YZ-$LiNbO_3$; *3*—128°YX-$LiNbO_3$; *4*—ST-SiO_2.

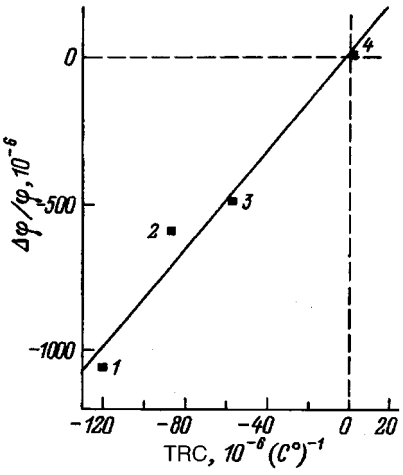

FIG. 7. Sensitivity of a calorimetric SAW sensor versus temperature coefficient of the flow rate (TCV), calibration gas N_2, $U=50$ ml/min, $T_p=120$ °C. *1*—(001),⟨110⟩-$Bi_{12}GeO_{20}$; *2*—YZ-$LiNbO_3$; *3*—128° YX-$LiNbO_3$; *4*—AT,X-SiO_2.

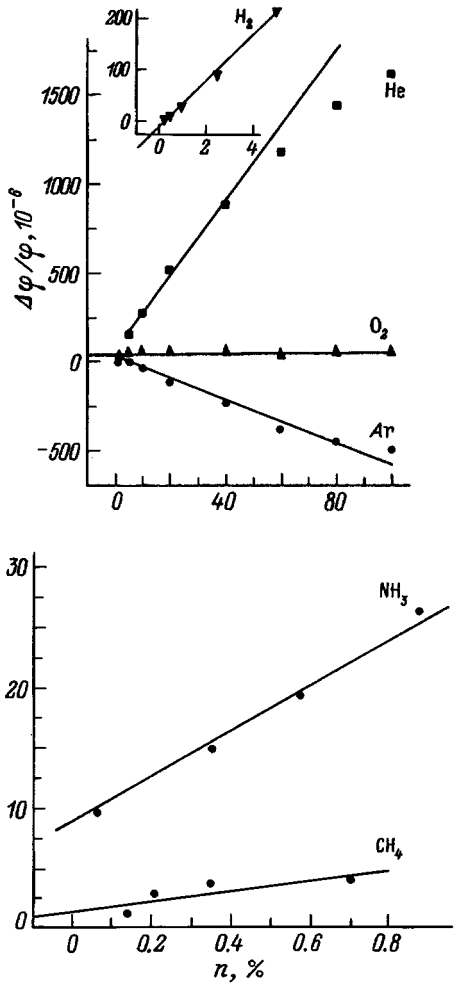

FIG. 8. Response of a thermoconductometric SAW sensor for a 128° XY-$LiNbO_3$ sensor, T_p=120 °C, calibration gas N_2, $U_{tg}=U_{cg}$=50 ml/min.

component becomes dominant, the substrates cools, and the SAW response changes to positive.

For all the investigated gases the SAW response is linear in the rate interval U=20–200 ml/min. Its level increases with the working temperature T_p (Fig. 4) and with the temperature coefficient of the velocity (TCV) of the surface acoustic waves (Fig. 5).

GAS CONCENTRATION MEASUREMENTS

Experiments have shown that the sensitivity of conductometric SAW gas detectors depends on the difference between the thermal conductivities of the calibration gas λ_{cg} and the tested gas λ_{tg} ($\Delta\lambda$) (Fig. 6), the material and the crystallographic cut of the substrate (through the TCV, Fig. 7), the working temperature T_p, and the concentration of the gas n (Fig. 8). For all the cases discussed here it does not depend on the gas injection rate, provided that the tested and calibration gases are injected into the chamber at the same rate.

In contrast with conventional thermoconductometric sensors,[6,7] SAW devices have the capability that their response can be increased, decreased, set to zero, or change sign, i.e., its level and sign can be controlled through an auxiliary parameter, the TCV. The same capability means that purely acoustical means can be used for certain gases to ensure that the selectivity of one of them will be small relative to another. This condition is illustrated in Fig. 9, in which the selectivity of He ($\Delta\lambda_{He}>0$, $\Delta T_{He}<0$) relative to Ar ($\Delta\lambda_{Ar}<0$, $\Delta T_{Ar}>0$) is more than 55 times greater. For the substrate used in our work the TCV is a linear function of the temperature and is equal to zero at T=174.1 °C. Consequently, the occurrence of Ar in air at a temperature T_p=174.1 °C−ΔT_{Ar} heats up the substrate by the amount ΔT_{Ar}, i.e., makes the working temperature equal to 174.1 °C, thereby reducing the TCV and the SAW response to zero. On the other hand, the introduction of He into the chamber at the same T_p leads to cooling of the substrate, imparting nonzero values of the TCV and the SAW response. The temperature dependence of the selectivity for this example is shown in Fig. 10.

CONCLUSIONS

The reported experimental results demonstrate the feasibility of detecting gases by their thermal conductivity using

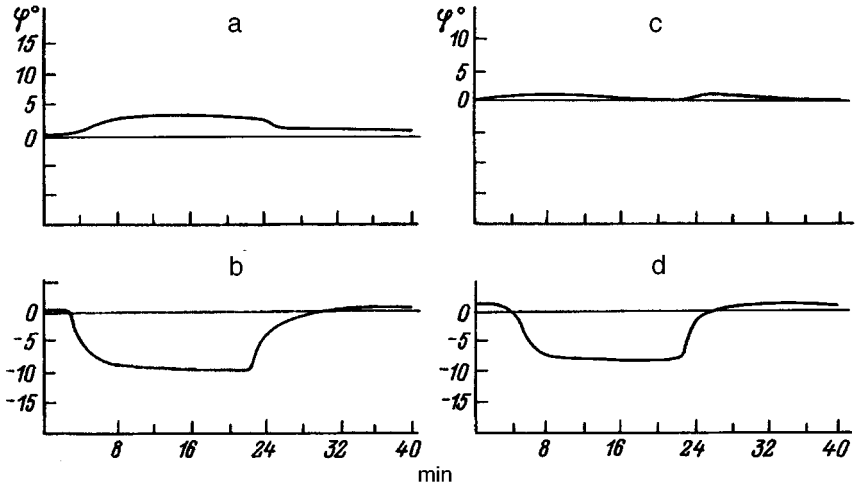

FIG. 9. Variation of the SAW responses with the working temperature T_p for an AT,X-SiO_2 substrate. a—Ar, T_p=102.9 °C; b—He, T_p=102.9 °C; c—Ar, T_p=163.4 °C; d—He, T_p=163.4 °C.

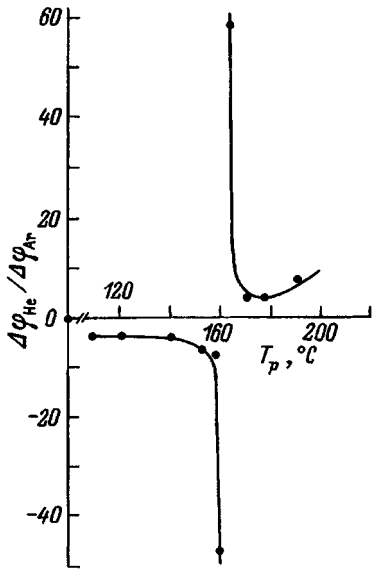

FIG. 10. Ratio of the SAW responses of He and Ar versus temperature for the example represented in Fig. 9.

SAW delay lines without gas-sensitive coatings. As in the case of conventional thermoconductometric sensors, the signal from the acoustic element is determined by the variation of the thermal conductivity of the gas mixture, thereby confining the use of the sensor to the analysis of binary mixtures of previously known gases. However, the response of the SAW element can be additionally controlled through the level and sign of the TCV, providing great flexibility in generating the output signal. In contrast with SAW gas sensors using gas-sensitive coatings, the latter are eliminated from thermoconductometric elements, and this simplifies the reproducibility and aging problems associated with such coatings. Moreover, direct contact between the tested gas and electromechanical transducers is eliminated in the thermoconductometric SAW element (an important asset in the detection of highly chemically active gases), and it is capable of recording different types of spectra of gases (including inert gases) over a wide range of concentrations (0.1–100%). The drawbacks of calorimetric SAW elements at the current stage are their comparatively high sensitivity threshold (0.1%), long activation times (of the order of a few minutes), and low selectivity. Promising solutions to these shortcomings in the future could be to utilize plate modes[4] and highly sensitive SAW temperature microsensors.[1] Also among the possible drawbacks of SAW elements are the need to ensure identical flow injection rates for the calibration gas and the tested gas, but this disadvantage has already been surmounted for conventional thermoconductometric sensors.[6]

[1] M. Hummady and D. Hauden, Sensors and Actuators **44**, 177 (1994).
[2] S. G. Joshi, Sensors and Actuators A **44**, 191 (1994).
[3] D. Rebiere, C. Dejous, J. Pistre *et al.*, Sensors and Actuators A **41-42**, 384 (1994).
[4] S. G. Joshi, Trans. IEEE Ultrason. Ferroelectr. Freq. Control **UFFC-38**, 148 (1991).
[5] V. K. Koshkin (ed.), *Fundamentals of Heat Transfer in Aircraft and Rocket Space Engineering* [in Russian], Mashinostroenie, Moscow (1975).
[6] P. T. Walsh and T. A. Jones, in *Sensors*, Vol. 2: *Chemical and Biochemical Sensors, Part 1*, W. Gopel, H. Hesse, and J. N. Zemel (eds.), New York–Basel–Cambridge (1991), pp. 529–570.
[7] E. A. Symons, in *Gas Sensors*, G. Sberveglieri (ed.), Kluwer Acad. Publ., Dordrecht–Boston–London (1992).

Translated by James S. Wood

Effectiveness of the repulsion of electromagnetic coils with strength and magnetic field limitations

V. V. Filatov

D. V. Efremov Scientific-Research Institute of Electrophysical Apparatus, 189631 St. Petersburg, Russia
(Submitted September 9, 1995; resubmitted December 18, 1995)
Zh. Tekh. Fiz. **67**, 124–127 (May 1997)

Mechanical power units utilizing electromagnetic interaction between coils are generally distinguished by compactness and ease of control. A coil-repulsion drive configuration is preferable for increasing the efficiency of such a device. One of the main limitations in choosing the parameters of stressed electromagnetic systems is mechanical strength. The application of superconducting coils has a number of additional limitations, for example, on the maximum magnetic field strength in the cross section of the coils. The problems of optimizing the dimensions of the electromagnetic system of a power unit consisting of two identical, coaxial, annular coils of rectangular cross section with opposing currents are investigated with allowance for strength and magnetic field limitations.
[S1063-7842(97)01805-9]

INTRODUCTION

Mechanical power units based on the electromagnetic interaction of coils are generally distinguished by their compactness and ease of control. An example of one such unit is the superconducting electromagnetic press,[1] which is designed for the long-term mechanical testing of samples. To enhance the efficiency and reliability of the power unit, it is preferable to use a coil-repulsion drive configuration.[2]

One of the main factors restricting the choice of parameters of stressed electromagnetic systems is their strength limitation. A number of additional limitations are encountered in the application of superconducting coils. For example, the minimum possible inside radius of a coil is dictated by the type of superconducting bus in terms of the maximum tolerable flexible strain of the superconductor, and the average current density in the coil cross section depends on the nominal current in the bus, which cannot exceed a critical value. Moreover, the maximum magnetic field in the coil must be below the limiting value at which the bus loses its superconducting properties.

The present article is devoted to optimizing the dimensions of the electromagnetic system of a power unit consisting of two identical, coaxial, annular coils of rectangular cross section with opposing currents.

EFFICIENCY OF INTERACTION OF ELECTROMAGNETIC COILS

Figure 1 shows the geometry of an electromagnetic system consisting of two identical, coaxial, annular coils of rectangular cross section with opposing currents. The current density is constant along the cross section in the case of multicoil systems.

In the theory of solenoids[3–6] it is customary to use dimensionless quantities referred to the inside radius of the coil r_{in} and the average current density j in the cross section of the coil:

$$\alpha = r_{\text{out}}/r_{\text{in}},$$

where r_{out} is the outside radius of the coil;

$$\upsilon = V/r_{\text{in}}^3,$$

where V is the volume of each coil;

$$\beta = b/r_{\text{in}} = \upsilon/[\pi(\alpha^2-1)],$$

where b is the axial width of each coil;

$$\delta = \Delta/(2r_{\text{in}}),$$

where Δ is the axial gap between the coils;

$$\sigma^* = \sigma/[\mu_0(jr_{\text{in}})^2],$$

where σ is the maximum stress in the coil, $\mu_0 = 4\pi\cdot 10^{-7}$ H/m is the absolute magnetic permeability;

$$B^* = B/(\mu_0 j r_{\text{in}}),$$

where B is the maximum magnetic induction on each coil;

$$F^* = F/[\mu_0(jr_{\text{in}})^2],$$

were F is the total repulsive force between the coils.

The force of electromagnetic interaction of the coils can be calculated analytically in series.[3,7] For certain system dimensions, however, the series converge very slowly. In this case numerical calculations are recommended, with the coils replaced by a system with many thin turns.

The closer together the coils are, the greater is the force of interaction of the currents, so that their repulsion for specified coil cross sections and currents (if the latter are opposing) increases when the coils are brought closer together, when their outside diameter is increased, or when the thickness of the coils is decreased.

To reduce the cost of the superconducting magnetic system, the volume of the coils must be made as small as possible for a given output force. The dimensions of the system are therefore optimal when the force F^* is a maximum for a given $\upsilon = \text{const}$. Figure 2 shows the dependence of $\mu_0 F^*$ on

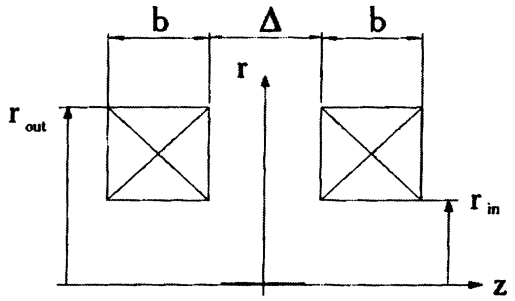

FIG. 1. Geometry of the electromagnetic system.

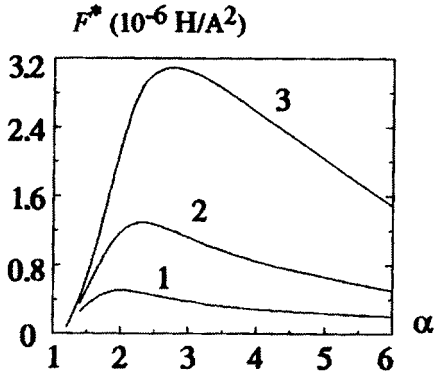

FIG. 3. Graph of $\mu_0 F^*$ versus α. *1* — $v=1$; *2* — $v=8$; *3* — $v=16$.

δ for coils of equal volume $v=4\pi$. The maximum force is attained for zero axial gap between the coils ($\delta=0$). This case is discussed below.

For $v=$const the function $F_v^*(\alpha)$ has a maximum (Fig. 3). For $j=$const this fact is attributable to a decrease in the total current for large values of α (thin pancake coil) and a rapid increase in the average distance between the coils for small values of α (long solenoids). For a given force F^* the optimum values of α_0 and $\beta_0=v/[\pi(\alpha_0^2-1)]$ correspond to minimum superconductor consumption, minimum stored energy, and minimum heating of the coils in the event of accidental power surges.

STRESSES IN THE COILS

For flat, thin, conducting pancake coils the stresses induced by electromagnetic loads are far greater than in long solenoids having the same cross sections and currents. Consequently, there are optimum system dimensions for achieving maximum repulsion of the coils for a given limiting stress level.

The strength problems of solenoids have already been studied in detail.[8,9] However, a system with opposing currents has special characteristics associated with the presence of a large radial field component and a complex configuration of magnetic lines of force between the coils. Consequently, the analytical relations customarily used in approximate strength calculations for solenoids are invalid in the given situation, and the combined problem of calculating the magnetic field and the stress of the strained state in the coil cross sections must be solved numerically.

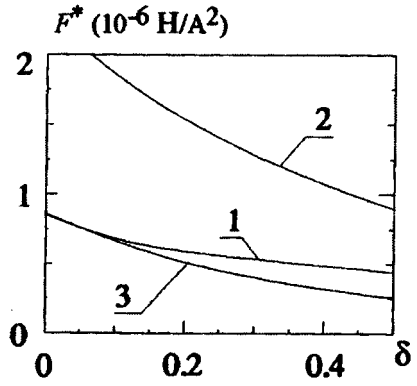

FIG. 2. Graph of $\mu_0 F^*$ versus δ for $v=4\pi$. *1* — $\beta=0.1$; *2* — $\beta=0.5$; *3* — $\beta=2.5$.

The author has developed a computer code called STRENG for the solution of such problems. The method of elliptical intervals[10] is used to calculate the field in the conductor; this method avoids the sizable error incurred by other methods in this application. The elasticity problem is solved on the basis of the resulting electromagnetic load distribution for boundary conditions corresponding to zero radial stresses at the inside and outside radii of a monolithic coil.[3]

The system of equations for the radial component r and the tangential component φ of the stresses σ in an element of volume of the coil has the form

$$\sigma_\varphi=\frac{E}{1-\mu^2}\left(\mu\frac{du}{dr}+\frac{u}{r}\right),$$

$$\sigma_r=\frac{E}{1-\mu^2}\left(\frac{du}{dr}+\mu\frac{u}{r}\right),$$

where u is the radial displacement of the element, E is Young's modulus, and μ is the Poisson ratio of the conductor material.

From the condition of equilibrium of the element we obtain

$$\frac{d^2u}{dr^2}+\frac{1}{r}\frac{du}{dr}-\frac{u}{r^2}+\frac{f_r}{E}(1-\mu^2)=0, \tag{1}$$

where $f_r=jB_z$ is the distributed volume radial electromagnetic force (per unit volume), and j is the current density in the coil element.

We solve the differential equation (1) numerically by the finite difference method with the boundary conditions

$$\sigma_r(r_{\text{in}})=0; \quad \sigma_r(r_{\text{out}})=0,$$

where r_{in} and r_{out} are the inside and outside radii of the coil, respectively.

The axial stresses in the repulsed coils

$$\sigma_z=\frac{F}{\pi r_{\text{in}}^2(\alpha^2-1)}$$

are at a maximum at the opposite ends of the coils. The equivalent stresses are

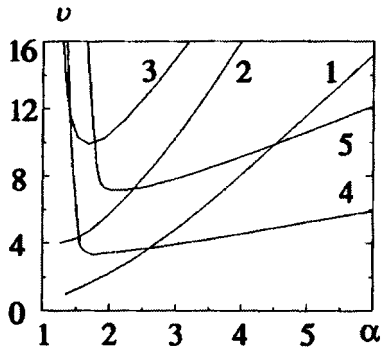

FIG. 4. Graphs of σ^* and F^* versus α and υ. *1* — $\sigma^*=0.032$; *2* — $\sigma^*=0.11$; *3* — $\sigma^*=0.21$; *4* — $F^*=0.28$; *5* — $F^*=0.80$.

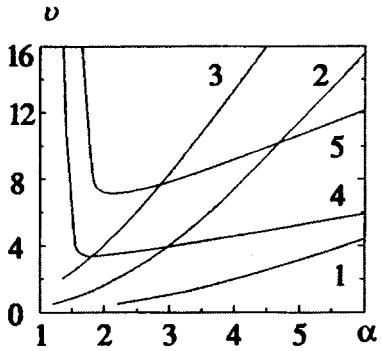

FIG. 5. Graphs of B^* and F^* versus α and υ. *1* — $B^*=0.040$; *2* — $B^*=0.14$; *3* — $B^*=0.28$; *4* — $F^*=0.28$; *5* — $F^*=0.80$.

$$\sigma=\sqrt{\sigma_\varphi^2+\sigma_z^2+\sigma_\varphi\sigma_z},$$

and their maximum is sought over the entire width of the coils.

Figure 4 shows $\sigma_\upsilon^*(\alpha)$ curves for $\upsilon=\text{const}$ and isolines for $F^*=\text{const}$). The operating point corresponds to the point of intersection of the equal-force line for a given value of F^* with the curve for the maximum allowed stress $\sigma^*=[\sigma]/(\mu_0 j^2 r_{\text{in}}^2)$ or, if they do not intersect, to the optimum value of α_0 in Fig. 3.

LEVEL OF THE MAGNETIC FIELD AT THE SUPERCONDUCTOR

A distinctive feature of superconducting coils is the added limitation on the strength of the magnetic field in the coil cross section at a given current density. This limitation must be taken into account in determining the most effective dimensions of the system.

Figure 5 shows $B_\upsilon(\alpha)$ curves and isolines for $F^*=\text{const}$. If the magnetic field at the operating point chosen from Fig. 4 exceeds the maximum allowed field $[B]$ for a given superconductor at a given current density $j=\text{const}$, the operating point must be chosen in Fig. 5 at the intersection of the F^* and $B^*=[B]/(\mu_0 j r_{\text{in}})$ isolines. Note that the dependence of the maximum field B on j is weaker than the same dependence of σ, so that for $B>[B]$ it is a good idea to lower the current density j by increasing the volume of the winding V in such a way that the operating points in Figs. 4 and 5 coincide.

CONCLUSIONS

The STRENG numerical code has been developed for the calculation of electromagnetic forces and mechanical stresses in coils. The calculated dependences can be used to choose the correct dimensions for a system of two identical, annular, coaxial coils of rectangular cross section with opposing currents, which are capable of generating a specified repulsive force with limitations on the strength and magnetic induction on the conductor.

The optimum coil dimensions and current distribution in terms of maximizing the force of interaction of the coils differ significantly from their optimum counterparts according to the condition of maximizing the field on the axis of a magnetic system. Possible approaches to further optimization of the electromagnetic system entail radial sectioning of the coil and redistribution of the current density toward the outer sections, for which the inductance and stored energy are higher.

[1] D. G. Akopyan, E. N. Bondarchuk, R. V. Chvartatsky *et al.*, IEEE Trans. Magn. **MAG-28**, 613 (1992).
[2] V. V. Filatov, NIIÉFA/TsNIIatominform Preprint No. 0929 [in Russian], Scientific-Research Institute of Electrophysical Apparatus and Central Scientific-Research Institute of Atomic Information, Moscow (1994).
[3] V. R. Karasik, *Physics and Engineering of High Magnetic Fields* [in Russian], Nauka, Moscow (1964).
[4] D. B. Montgomery, *Solenoid Magnet Design*, Wiley, New York (1969) [Russ. Transl., Mir, Moscow (1971)].
[5] D. H. Parkinson and B. E. Mulhall, *The Generation of High Magnetic Fields*, Plenum Press, New York; Heywood Books, London (1967) [Russ. transl., Atomizdat, Mosow (1971)].
[6] V. V. Filatov, in *Proceedings of the International Conference on Open Plasma Confinement Systems for Fusion*, World Scientific, Singapore (1994), pp. 561–577.
[7] H. B. Dwight, *Electrical Coils and Conductors*, London (1945).
[8] A. A. Kuznetsov, Zh. Tekh. Fiz. **30**, 592 (1960) [Sov. Phys. Tech. Phys. **5**, 555 (1960)].
[9] P. O. Carden, J. Sci. Instrum. **1**, 437 (1968).
[10] B. L. Alievskiĭ and V. L. Orlov, *Calculating the Magnetic Field Parameters of Axisymmetrical Coils* [in Russian], Énergoatomizdat, Moscow (1983).

Translated by James S. Wood

Kinetics of coalescence during phase transition in microcrystals

I. E. Yakovlev

State Technical University, 195251 St. Petersburg, Russia
(Submitted December 6, 1995)
Zh. Tekh. Fiz. **67**, 128–131 (May 1997)

[S1063-7842(97)01905-3]

Researchers working in the field of solid state physics have lately shown considerable interest in media containing microcrystalline inclusions. The characteristics of such materials are determined by the composition, shape, and size distribution of the inclusions. One method for the preparation of such media containing microcrystals is to create a supersaturated solid solution, for example, a semiconductor in a glass matrix, with subsequent phase melting to form microcrystallites.[1] The growth of such inclusions in isotropic matrices is discussed in Ref. 2, where an analytical expression is derived for the size distribution function of isotropic grains in the coalescence stage. In experimental work,[3] however, it has been observed that the distribution formed in this stage differs substantially in certain cases from what was expected.[2] In other experiments[4,5] a phase transition has been detected in the microcrystals during their growth. The objective of the present study is to investigate the nature of the possible influence of such phase transitions on the formation of the microcrystal size distributions in the coalescence stage.

The occurrence of phase transition in a microcrystal during its growth can be explained on the basis of qualitative considerations. The total energy of a microcrystal in a glass matrix is the sum of its internal energy and the energy of the boundary. For small grain sizes the total energy depends mainly on the boundary energy, and a more favorable (having the lowest energy) configuration of the microcrystal-matrix system from the energy standpoint is one having the least dissimilarity between the spatial structures of the microcrystal and the glass matrix. It follows from experimental results,[4] for example, that a cubic grain structure is ''closer'' to that of an isotropic matrix. For large microcrystals the total energy depends mainly on the internal energy of the inclusion, and a configuration in which the grains have a bulk crystal structure is more favorable. For example, the hexagonal modification of a number of II–VI microcrystals takes preference over the equally admissible cubic modification.[6] Consequently, structural transformation of the crystal lattice is possible when the microcrystal increases in size. The more rigid the glass matrix, i.e., the higher the energy density of the microcrystal–matrix boundary, the larger will be the dimensions of the microcrystal for a change to be observed in its structure, essentially in the form of a second-order transition. Such a phase transition has in fact been observed in the rigid silicate glasses.[4]

The instant of phase transition can be estimated from the following considerations. Let $\varepsilon_{\mathbf{a},\mathbf{b}}$ be the volume energy density of an inclusion having structure **a** or **b**, and let $\sigma_{\mathbf{a},\mathbf{b}}$ be the energy density of the boundary between the inclusion (with structure **a** or **b**) and the matrix; the total energy of the inclusion $E_{\mathbf{a},\mathbf{b}}$ is then

$$E_i=\frac{3}{4}\pi r^3\varepsilon_i+4\pi r^2\sigma_i, \quad i=\mathbf{a},\ \mathbf{b}, \tag{1}$$

where r is the radius of the inclusion.

Let the structure of the microcrystal change from **a** to **b** when it attains the radius r_0; then for $r<r_0$ the total energy of an inclusion with structure **a** is lower than the total energy of the same inclusion with structure **b** ($E_{\mathbf{a}}<E_{\mathbf{b}}$), and for $r>r_0$ the opposite relation holds, $E_{\mathbf{a}}>E_{\mathbf{b}}$, i.e., the quantity r_0 corresponds to the condition $E_{\mathbf{a}}=E_{\mathbf{b}}$. Consequently,

$$r_0=3\frac{\sigma_{\mathbf{b}}-\sigma_{\mathbf{a}}}{\varepsilon_{\mathbf{a}}-\varepsilon_{\mathbf{b}}}. \tag{2}$$

The necessary condition for the existence of phase transition is

$$(\varepsilon_{\mathbf{a}}-\varepsilon_{\mathbf{b}})(\sigma_{\mathbf{b}}-\sigma_{\mathbf{a}})>0, \tag{3}$$

which implies that the favorable grain structure from the standpoint of internal energy will be unfavorable from the standpoint of surface energy, and vice versa.

We investigate the coalescence process against the background of phase transition by numerical simulation. In constructing the model, we assume that the grains are spherical, the individual microcrystals are noninteracting by virtue of their smallness and large distances between them (this assumption holds true as a rule; in the example of glasses doped with II–VI semiconductor microcrystals the concentration of the semiconductor phase does not exceed 1–2%; see Ref. 3), and the external conditions are invariant.

For the growth rate of grains outside the size range for their structural transformation we use the standard relation[7]

$$\frac{\partial r}{\partial t}=\frac{D}{r}\left(\Delta-\frac{\alpha}{r}\right), \tag{4}$$

where r is the grain radius, D is the diffusion coefficient of the impurity in the glass, Δ is the excess of the concentration of the solid solution above the saturation level, and α is a coefficient proportional to the surface energy density of the microcrystal-matrix boundary.

In our model, however, α is not a constant, as it was in Ref. 2, but is a function of the radius of the microcrystal. As a result of phase transition with structural transformation, the surface energy increases (transition $\mathbf{a}\rightarrow\mathbf{b}$, $\sigma_{\mathbf{b}}>\sigma_{\mathbf{a}}$), causing the coefficient α to increase for microcrystals with structure **b** and, accordingly, the growth rate $\partial r/\partial t$ to decrease [see Eq. (4)].

 1063-7842/97/050572-04$10.00

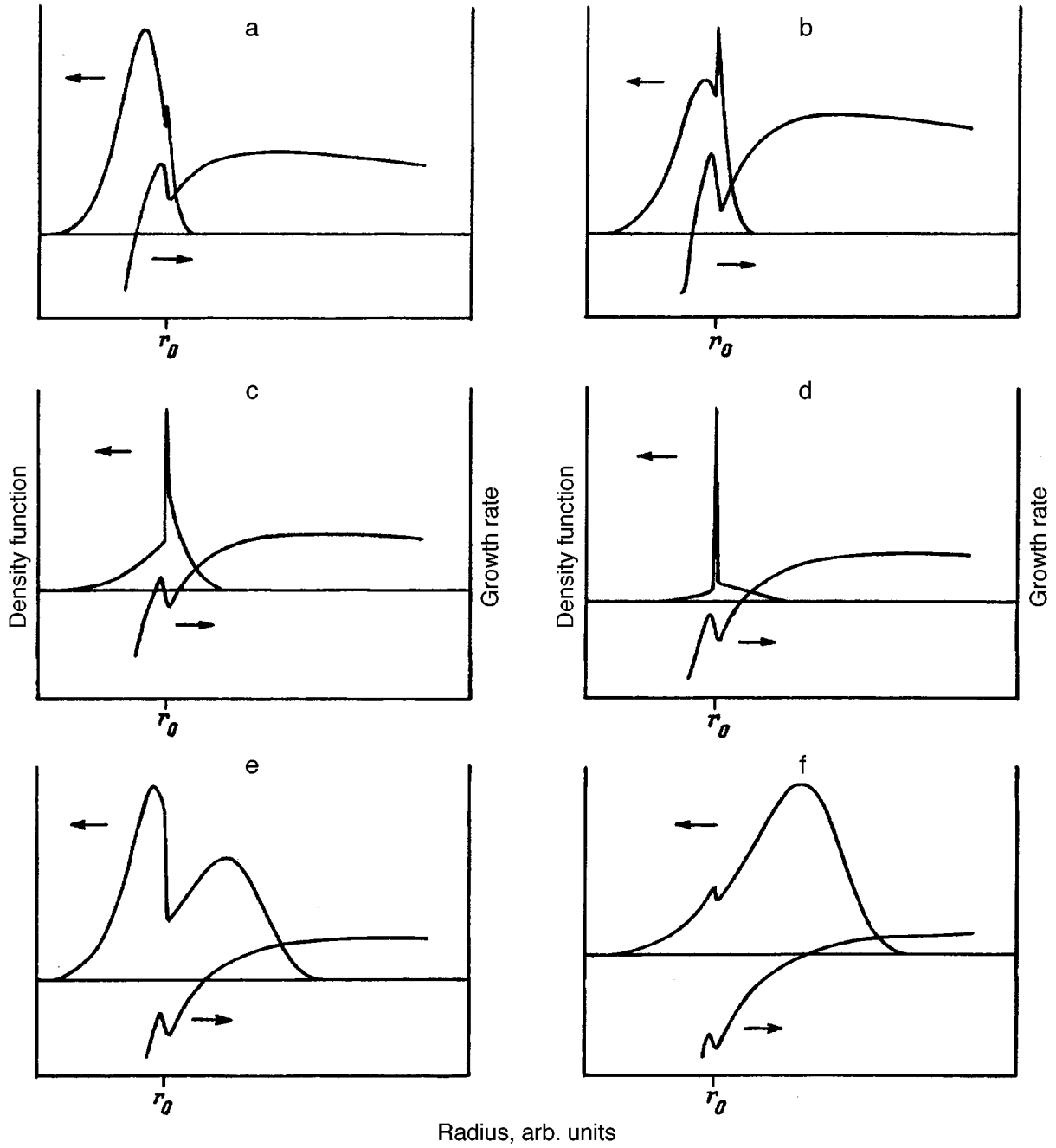

FIG. 1. Distribution function and growth rate of microcrystals during various stages of coalescence.

Within the framework of the proposed model we formulate and solve numerically a system of integrodifferential equations describing the evolution of the ensemble of microcrystals during the coalescence stage in the presence of the phase transition **a**→**b**. Microinhomogeneities of the medium, which create a difference in the values of r_0 for microcrystals distributed throughout the matrix, are taken into account by the introduction of a so-called width of the transition zone δr_0 about the phase-transition radius r_0, wherein the ensemble average of the coefficient α changes from $\alpha_{\mathbf{a}}$ for structure **a** to $\alpha_{\mathbf{b}}$ for structure **b** ($\delta r_0/r_0 \ll 1$).

The main results of the numerical calculations are summarized as follows.

1. As the average radius of the microcrystals $\langle r \rangle$ increases, the probability density function goes through the following characteristic stages (in Fig. 1 the probability density functions and the growth rates of the microcrystals $\partial r/\partial t$ are plotted as functions of the radius r at various stages of growth): 1) $\langle r \rangle < 2r_0/3$, all the microcrystals have structure **a**, and the density function coincides with the Lifshits–Slezov density function[2]; 2) $\langle r \rangle \approx 2(r_0+\delta r_0)/3$, the largest microcrystals undergo phase transition and have structure **b**; since the growth rate of type **b** microcrystals is lower than that of type **a** microcrystals, the density function acquires a peak on its right slope (Fig. 1a); 3) $\langle r \rangle \approx (r_0-\delta r_0)$, the peak on the right slope of the density function continues to rise,

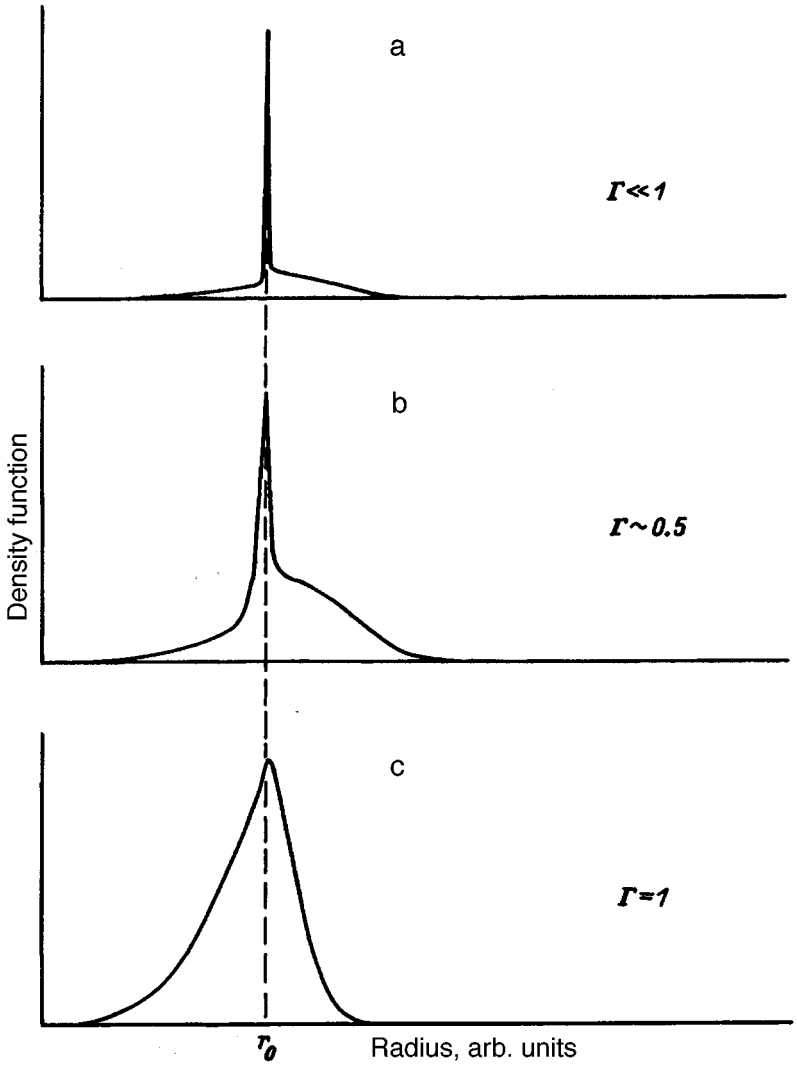

FIG. 2. Characteristic distribution functions for processes with different parameters Γ.

becoming equal to the principal maximum of the density function for type **a** microcrystals and still continuing to rise (Fig. 1b); 4) $\langle r\rangle \approx r_0$, in this stage the average radius of the microcrystals is almost constant, but there is an increase in the number of grains with structure **b**, i.e., of grains undergoing phase transition, accompanied by the following changes in the density function: The peak of the distribution continues to rise (Fig. 1c), the distribution acquires a quasimonodisperse form (Fig. 1d), and then a distribution of the Lifshits–Slezov type[2] begins to evolve from type **b** microcrystals having radii several times the phase-transition radius r_0, where the microcrystals responsible for the peak of the monodisperse distribution are now dissolved ($\partial r/\partial t<0$), reverting to structure **a**, and then the type **b** microcrystals form a distribution of the Lifshits–Slezov type[2] (Fig. 1e); 5) $\langle r\rangle \gg r_0$, almost all the microcrystals have structure **b**, and their density function differs very little from the Lifshits–Slezov form, while a minute fraction of type **a** microcrystals form a peak on the left slope of the probability density function (Fig. 1f).

2. The definition of these separate stages depends on the radius of the transition interval δr_0. In the course of simulation it has been observed empirically that the quantity

$$\Gamma = \frac{1}{2}\,\frac{\alpha_{\mathbf{b}}+\alpha_{\mathbf{a}}}{\alpha_{\mathbf{b}}-\alpha_{\mathbf{a}}}\,\frac{\delta r_0}{r_0} \tag{5}$$

well describes the various coalescence scenarios. The density function evolves similarly for processes with equal values of Γ: Essentially monodisperse distributions can emerge for $\Gamma \ll 1$ (Fig. 2a), in the case $\Gamma \approx 0.5$ (Fig. 2b) the results of the calculations are in fair agreement with experiment[3] (Fig. 3), and for $\Gamma > 1$ (Fig. 2c) the pattern is consistent with classical coalescence theory.[2]

3. In comparing the experimental results[3] and the results of numerical simulation of the given experimental situation, we have obtained the approximate value $r_0 \sim 3$ nm. Consequently, this value gives an idea of the range of microcrystal sizes where one should expect the distribution function to exhibit the indicated prominent features, including quasimonodisperse distributions, in the given situation.

In summary, on the basis of numerical analysis we have shown that phase transition in the stage of coalescence of microcrystals formed during the melting of a supersaturated solid solution can significantly alter the form of the size distribution function of these microcrystals. If the surface energy of the microcrystals changes appreciably as a result of

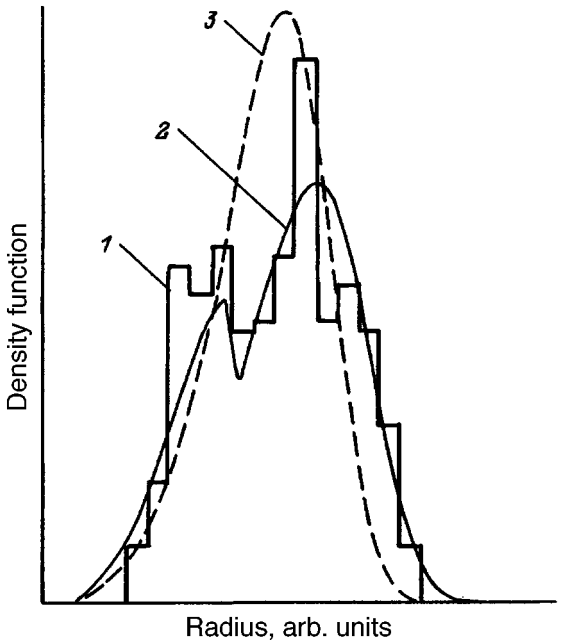

FIG. 3. *1* — Experimentally recorded size distribution function of microcrystals; *2* — calculated distribution function; *3* — Lifshits-Slezov distribution function.

the phase transition and if the medium is sufficiently homogeneous that the interval of microcrystal radii in which the phase transition takes place can be regarded as small, the microcrystals can acquire an almost-monodisperse size distribution.

The author is grateful to A. A. Lipovskiĭ for his undivided attention and a discussion of the results.

[1] A. I. Ekimov and A. A. Onushchenko, Fiz. Tekh. Poluprovodn. **16**, 1215 (1982) [Sov. Phys. Semicond. **16**, 775 (1982)].
[2] I. M. Lifshits and V. V. Slezov, Zh. Éksp. Teor. Fiz. **35**, 479 (1958) [Sov. Phys. JETP **8**, 331 (1958)].
[3] A. A. Lipovskii, N. V. Nikonorov, M. V. Kharchenko, and A. A. Sitnikova, Proc. SPIE **2291**, 327 (19994).
[4] S. A. Gurevich, A. I. Ekimov, and I. A. Kudryavtsev *et al.*, Fiz. Tekh. Poluprovodn. **28**, 830 (1994) [Semiconductors **28**, 486 (1994)].
[5] T. Gakoin, C. Train, F. Chaput *et al.*, Proc. SPIE **1758**, 565 (1992).
[6] *Handbook of Physicochemical Properties of Semiconductor Materials* [in Russian], Nauka, Moscow (1979).
[7] E. M. Lifshitz and L. P. Pitaevskii, *Physical Kinetics*, Pergamon Press, New York-Oxford (1981) [Russ. original, Nauka, Moscow (1979)].

Translated by James S. Wood

Interaction of ultrashort laser pulses with the surface of a semiconductor

E. V. Gorbunov

(Submitted January 10, 1995)
Zh. Tekh. Fiz. **67**, 132–133 (May 1997)

[S1063-7842(97)02005-9]

Ultrashort (of duration 10^{-9} s or less) pulses of various physical types (electromagnetic pulses, acoustic pulses, electron bunches, etc.) are currently a topic of growing interest.[1] In particular, one of the more interesting objects of investigation is the generation of ultrashort acoustic pulses when a semiconductor material is irradiated with ultrashort laser pulses.[2] On the other hand, the interaction of ultrashort laser pulses with a material has certain specific attributes associated with the need to treat the electronic and lattice (ionic) subsystems separately in writing out the system of governing equations.[3]

In this note we consider the interaction of ultrashort laser pulses with the surface of a semiconductor material and determine the density of nonequilibrium carriers and the temperatures of electrons and ions in it both during laser irradiation and after termination of the incident pulse.

We describe the interaction of an ultrashort laser pulse with the surface of a semiconductor by the system of equations

$$\frac{\partial n}{\partial t}=D\Delta n-\frac{n}{\tau_{\rm rec}},$$

$$C_e\frac{\partial T_e}{\partial t}=\lambda_e\Delta T_e-G(T_e-T_i)+(1-\beta)E_g\frac{n}{\tau_{\rm rec}},$$

$$C_i\frac{\partial T_i}{\partial t}=\lambda_i\Delta T_i+G(T_e-T_i)+\beta E_g\frac{n}{\tau_{\rm rec}}.$$

Here n is the density of nonequilibrium carriers, T_e and T_i are the temperatures in the electronic and ionic subsystems, C_e and C_i are the electronic and ionic heat capacities, λ_e and λ_i are the electronic and ionic thermal conductivities, D is the carrier diffusion coefficient, G is the energy transfer constant from electrons to the ionic lattice, E_g is the width of the semiconductor band gap, $\tau_{\rm rec}$ is the electron–hole recombination time, and β is the fraction of energy released by recombination that transfers to the ionic lattice. The system of governing equations is complemented by the following boundary conditions: The carrier densities, the electron and lattice temperatures, and their gradients are equal to zero outside the semiconductor and in the semiconductor at an infinite distance from the irradiated surface. The following conditions are also satisfied on the irradiated surface:

$$-\lambda_e\frac{\partial T_e}{\partial x}=(1-\eta)(1-R)I(x=0,y,z,t),$$

$$-D\frac{\partial n}{\partial x}=\frac{\eta(1-R)}{E_g}I(x=0,y,z,t),$$

and the initial conditions are written in the usual form

$$n(r,t)=0,\quad T_e(\mathbf{r},t)=T_i(\mathbf{r},t)=0,\quad t\leqslant 0.$$

The boundary conditions on the surface take into account the radiation reflection coefficient R, and account is taken of the fact that only a fraction η of the total absorbed intensity $(1-R)I$ is spent in generating electron-hole pairs, while the remainder $(1-\eta)$ is spent in heating the generated electrons.

Invoking the double Fourier transform with respect to the coordinates y, z and the Fourier cosine transform with respect to the coordinate x, we simultaneously reduce everything to dimensionless form. Introducing $t=\Delta t\cdot t$, $x,\ y,\ z=\Delta x(x,y,z)$, $T_{e,i}=\Delta T\cdot T_{e,i}$, $n=\Delta n\cdot n$, and $I=\Delta I\cdot I$ and taking into account the boundary conditions for the Fourier transforms of the unknown functions in the now-dimensionless variables $[\mathbf{k}=(k_x,k_y,k_z)]$, we obtain

$$\frac{\partial n^*}{\partial t}=-k^2Fd\cdot n^*-Nn\cdot n^*+\sqrt{\frac{2}{\pi}}\,In\ Fd\cdot I^*(k_y,k_z,t),$$

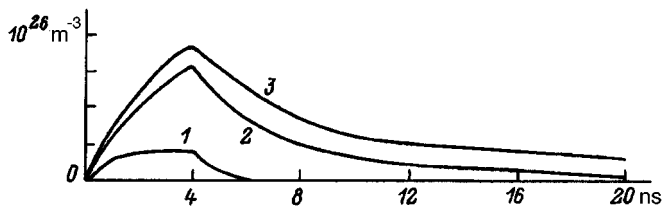

FIG. 1. Nonequilibrium carrier density versus time under the influence of an ultrashort laser pulse at a distance of 0.1 μm from the surface, $\tau=4$ ps, $I=8\times10^{12}$ W/m^2. *1* — $\tau_{\rm rec}=0.7$ ps; *2* — $\tau_{\rm rec}=7$ ps; *3* — $\tau_{\rm rec}=20$ ps.

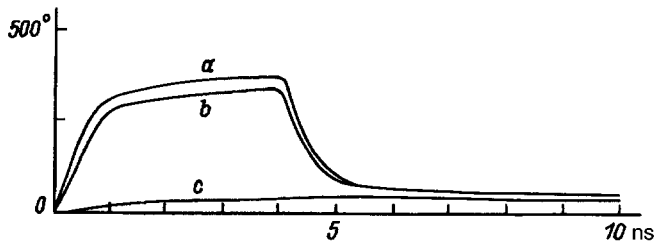

FIG. 2. Electron temperature (a, b) and lattice temperature (c) versus time under the influence of an ultrashort laser pulse at a distance of 0.1 μm from the surface, $\tau=4$ ps, $I=8\times10^{12}$ W/m^2. a — $\tau_{\rm rec}=0.7$ ps; b — $\tau_{\rm rec}=7$ ps.

 1063-7842/97/050576-02$10.00

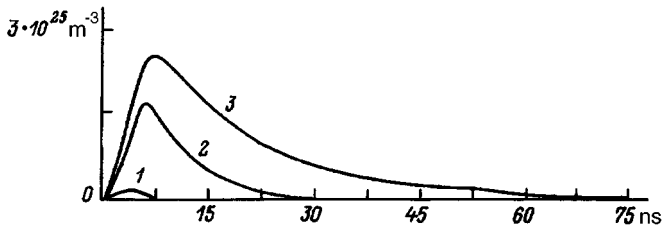

FIG. 3. Nonequilibrium carrier density versus time under the influence of an ultrashort laser pulse at a distance of 0.5 μm for the same values of τ and *I*. *1* — $\tau_{rec}=0.7$ ps; *2* — $\tau_{rec}=7$ ps; *3* — $\tau_{rec}=20$ ps.

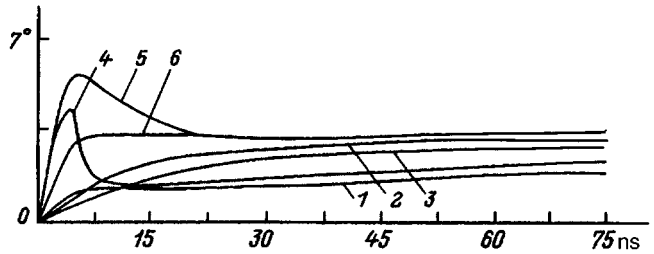

FIG. 4. Lattice temperature (*1–3*) and electron temperature (*4–6*) versus time under the influence of an ultrashort laser pulse at a distance of 0.5 μm for the same values of τ and *I*. *1*, *4* — $\tau_{rec}=0.7$ ps; *2*, *5* — $\tau_{rec}=7$ ps; *3*, *6* — $\tau_{rec}=20$ ps.

$$\frac{\partial T_e^*}{\partial t}=-k^2F_o^e-G_o(T_e^*-T_i^*)+Ne\cdot n^*+\sqrt{\frac{2}{\pi}}F_o^eIeI^*(k_y,k_z,t),$$

$$\frac{\partial T_i^*}{\partial t}=-k^2F_o^iT_i^*+Li(T_e^*-T_i^*)-Ni\cdot n^*.$$

Here we have introduced the dimensionless parameters

$$F_o^e=\frac{\lambda_e\Delta t}{C_e\Delta x^2},\quad F_o^i=\frac{\lambda_i\Delta t}{C_i\Delta x^2},\quad Fd=D\frac{\Delta t}{\Delta x^2},$$

$$Nn=\frac{\Delta t}{\tau_{rec}},\quad Go=\frac{G}{C_e\Delta t},\quad Lo=\frac{G}{C_i\Delta t},$$

$$Ne=\frac{(1-\beta)E_g\Delta n\Delta t}{\tau_{rec}C_e\Delta T},\quad Ni=\frac{\beta E_g\Delta n\Delta t}{\tau_{rec}C_i\Delta t},$$

$$In=\frac{\eta(1-R)\Delta I\Delta x}{D\Delta nE_g},\quad I_e=\frac{(1-\eta)(1-R)\Delta I\Delta x}{\lambda_e\Delta T}.$$

For the Fourier transform of the carrier density the solution has the form

$$n^*(\mathbf{k},t)=\sqrt{\frac{2}{\pi}}In\cdot Fn\int_0^\infty I^*(m)\times\exp((k^2Fd-Nn)(m-t))dm,$$

and the electron temperature is calculated from the equation

$$T_e^*(\mathbf{k},t)=\frac{1}{Lo}\frac{\partial T_i}{\partial t}+\frac{k^2F_o^i+Lo}{Lo}T_i^*-\frac{Ni}{Lo}n^*.$$

The Fourier transform of the lattice temperature is determined from the solution of the equation

$$\frac{\partial^2T_i^*}{\partial t^2}+A\frac{\partial T_i^*}{\partial t}+BT_i^*=Cn^*+\sqrt{\frac{2}{\pi}}EI^*(k_y,k_z,t),$$

where the factors *A*, *B*, *C*, and *E* represent combinations of the above-defined dimensionless parameters.

The final expression for the unknown temperatures and density is cumbersome. As an example we give the equations for the one-dimensional case, where the coordinate *x* is in the direction normal to the surface irradiated by a rectangular ultrashort laser pulse of duration τ and intensity *I*:

$$n(x,t)=I\int_0^\infty\cos(k_xx)A_1f_1(\Omega t)dk_x,$$

$$T_e(x,t)=I\int_0^\infty\cos(k_xx)[A_2f_2(s_1t)+A_3f_3(s_2t)+A_4f_1(\Omega t)]dk_x,$$

$$T_i(x,t)=I\int_0^\infty\cos(k_xx)[A_5f_2(s_1t)+A_6f_3(s_2t)+A_7f_1(\Omega t)]dk_x.$$

Here s_1, s_2, A_1–A_7, and f_1–f_2 are the expressions and functions determined by the dimensionless parameters and the parameter of integration k_x, and $\Omega=k_x^2Fd+Nn$. The results of calculations according to the proposed model are shown in Figs. 1–4 for an ''averaged'' semiconductor with its thermophysical parameters evaluated as the averages of the corresponding quantities. Figure 1 illustrates the influence of the recombination time τ_{rec} on the carrier density dynamics at a distance of 0.1 μm from the surface: Both the peak value and the lifetime of the elevated carrier density increase as τ_{rec} increases. It follows from Fig. 2, which shows the electron temperature (a, b) and the lattice temperature (c) for the same irradiation conditions, that the influence of τ_{rec} on the lattice temperature is insignificant, whereas the electron temperature is more sensitive to τ_{rec}.

Figures 3 and 4 show the carrier densities and temperatures at a distance of 0.5 μm from the irradiated surface. It is evident from the graphs that with increasing distance from the surface the position of the density and electron-temperature maxima shifts by an amount that increases as the carrier recombination time increases (up to 7.5 ps for $\tau_{rec}=20$ ps).

[1] S. A. Akhmanov, V. A. Vysloukh, and A. S. Chirkin, *Optics of Femtosecond Laser Pulses*, Nauka, Moscow (1988).
[2] S. M. Avanesyan and V. É. Gusev, Kvantovaya Élektron. **13**, 1241 (1986) [Sov. J. Quantum Electron. **16**, 812 (1986)].
[3] S. I. Anisimov, Ya. A. Imas, G. S. Romanov, and Yu. V. Kodyko, *Effects of High-Intensity Radiation on Metals* [in Russian], Nauka, Moscow (1970).

Translated by James S. Wood

Observation of the magnetization process of a thick ferromagnet by neutron radiography with refraction-produced contrast

K. M. Podurets, R. R. Chistyakov, and S. Sh. Shil'shteĭn

Russian Science Center ''Kurchatov Institute''
(Submitted December 18, 1995)
Zh. Tekh. Fiz. **67**, 134–136 (May 1997)

[S1063-7842(97)02105-3]

Magnetic powder patterns and the magnetooptic Kerr effect have been used to observe internal domain structure during the magnetization of a ferromagnet in the example of iron whiskers and silicon iron single-crystal plates.[1–3] In both cases the principal domains are layers running through the entire crystal; the domains of closure are wedge-shaped, and the directions of magnetization of the domains are easily determined from the pattern itself. Consequently, the transformation of the volume domain structure during magnetization can be assessed from the variation of the pattern. When the field is applied along the principal domains, those which are initially magnetized along the field grow at the expense of oppositely magnetized domains and domains of closure. In the vicinity of sharp angles (e.g., at the tips of whiskers) transversely magnetized domains of closure are preserved in fields up to 6 kG, despite the usual assumption that the magnetization processes have terminated much earlier.

It is impossible to investigate processes of magnetization of a thick ferromagnet by optical methods, which exhibit domains only on the surface. It has been shown[1,4] that neutron radiography with refraction-produced contrast can be used to observe individual domain boundaries in the interior of a thick ferromagnet and also in a region of nonuniform magnetic field. This consideration makes the problem of directly observing the magnetization process in a thick ferromagnet fully approachable. In previous papers, neutron-refraction techniques have been used to exhibit the pattern of the domain boundaries in the interior of cylindrical crystals of silicon iron.[4–6] The objective of the present study is to observe magnetization processes in such crystals by means of neutron radiography with refraction contrast.

In neutron radiography with refraction-produced contrast an image of the object is formed by neutrons transmitted through the object without refraction, so that the domain boundaries, at which neutrons are refracted,[7] show up dark on the radiograms.[3] The unrefracted beam is separated from the refracted beams by means of a double-crystal spectrometer with high angular resolution.[8] The sample is placed between the monochromator crystal and the analyzer crystal, and the cassette containing the converter (gadolinium foil) and the x-ray film is positioned behind the analyzer (Fig. 1). The spatial resolution in this arrangement depends on the beam spreading and the sample-film spacing; if the latter is equal to 6 cm, the resolution is 0.3 mm in the experiments described below. The dimensions of the beam were 20×30 mm, which were dictated by the dimensions of the germanium crystals used for the monochromator and analyzer and by the Bragg angle for the reflection (111) at a wavelength of 0.15 nm. The exposure time was 10 h.

The experimental samples were cylindrical Fe–Si (3% wt.) single crystals with [001] axis and a diameter of 13 mm, which were grown at the Physics Institute of the Czechoslovak Academy of Sciences in Prague by the floating-zone method. The cylinders were placed in a permanent magnet (Fig. 1) in such a way that the beam penetrated it perpendicular to its axis. Experiments were performed on sample 1 with a length of 22 mm in a field of 1.5 kG and on sample 2 with a length of 9 mm in fields of 3.5–8 kG as the width of the gap in the magnet was varied.

In the initial state the inner part of the cylinder is constructed of layers perpendicular to its axis and consisting of domains whose magnetization vectors [100] and [010] are situated in these layers (Fig. 2a); the outer part of the cylinder is constructed of domains magnetized along the cylinder axis (Fig. 2b).[6] Accordingly, in the given geometry of the experiment the boundaries between the layers and the boundary between the interior and surface zones with different directions of magnetization appear on the radiogram (Fig. 3a). In a field of 1.5 kG (Fig. 3b) both the diameter of the layered region and the thickness of the layers themselves decrease. In this case, therefore, the volume of the part of the crystal magnetized perpendicular to the field does in fact decrease. In all likelihood there is a simultaneous increase in the volume of the domains magnetized parallel to the field; this possibility is indicated by an increase in the thickness of the unlayered part of the cylinder near the surface. In a field of 3.5 kG we observe appreciable distortion of the layered structure inside the crystal (Figs. 4a and 4b), and in a field of 5.5 kG the initial layered domain structure of the crystal is completely destroyed (Fig. 4c). Neutron refraction is not observed in the middle part of the cylinder, suggesting that the domain boundaries have vanished, i.e., this part of the cylinder is completely magnetized. On the other hand, increased darkening appears on the images of the crystals near the ends of the cylinder, and these parts of the crystal become less ''transparent.'' In other words, the number of boundaries encountered along the path of the neutrons increases as a result of domain refinement. At the same time, images of the individual domain boundaries do not appear here, because the dimensions of the domains fall below the spatial resolution. Finally, in a field of 8 kG neutron refraction disappears throughout the entire volume of the cylinder (Fig. 4d), indicating the absence of domain boundaries in its interior, i.e., proximity to the state of total magnetization of the crystal.

The nature of the transformation of the domain structure

 1063-7842/97/050578-03$10.00

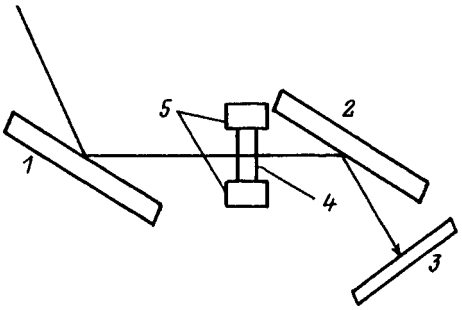

FIG. 1. Experimental arrangement. *1*, *2* — Monochromator and analyzer crystals; *3* — film cassette; *4* — sample; *5* — permanent magnet.

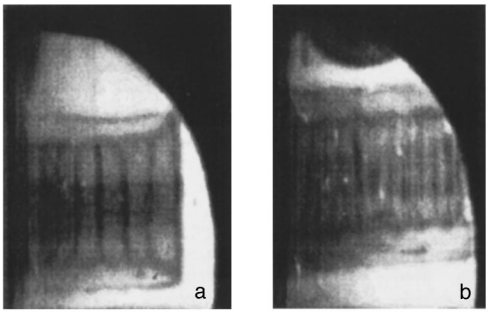

FIG. 3. Radiograms of an Fe–Si crystal (sample 1). a — Without a field; b — in a field of 1.5 kG.

in a field indicates the prevalence of displacement processes, where the volume of domains magnetized along the field increases, and the volume of transversely magnetized domains decreases. The nonuniformity of the magnetization in a field of 5.5 kG indicates a time lag of the magnetization near the ends of the crystal, which are perpendicular to the applied field, consistent with the data of Ref. 2. Here the middle part of the crystal is magnetized very close to saturation, and the density of the domain boundaries near the ends increases, i.e., the boundaries are displaced out of the middle part of the crystal. The general pattern of transformation of the volume domain structure of a thick crystal is similar to that observed previously in crystals with a through domain structure.

The observation of nonuniformity of the magnetization of thick samples could be the primary way to investigate magnetization processes in objects made from polycrystalline materials; this capability is important for discerning the details of processes governing the nature of losses in elements of machinery and equipment. In the observation of spatial nonuniformity of the magnetization it is not so important to have high spatial resolution as in the observation of individual boundaries. Consequently, bulkier objects can be used, and processes taking place in real devices, for example, in large casings, can be investigated. The specific characteristics of neutron penetration responsible for neutrons having a magnetic moment, combined with high penetrating power, are fully exploited in this case.

The authors are grateful to V. A. Somenkov for helpful discussions and to Yu. A. Bulanovskiĭ and A. I. Rogachev

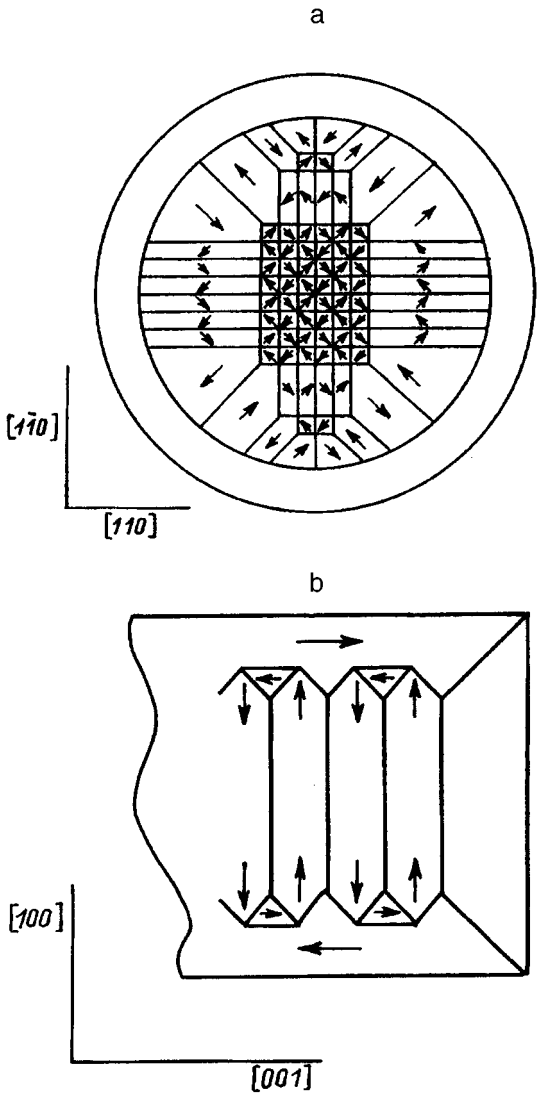

FIG. 2. Diagram of the domain structure of a cylindrical Fe–Si crystal. a — Magnetic flux closure in layers; b — between layers.

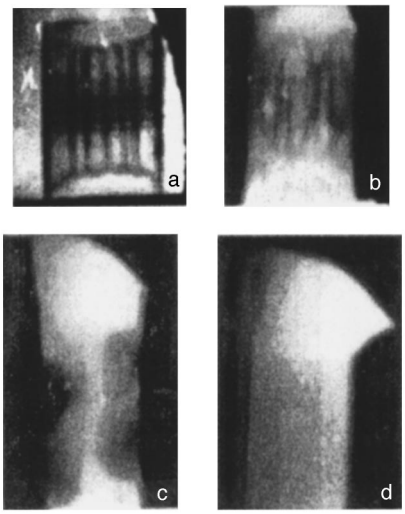

FIG. 4. Radiograms of an Fe–Si crystal (sample 2). a — Without a field; b — in a field of 3.5 kG; c — in a field of 8 kG.

for assisting with the experiment. This work has received support from the Russian Foundation for Fundamental Research, Project No. 95-02-04691-a.

[1] K. M. Podurets, V. A. Somenkov, and S. S. Shilstein, Physica B **156/157**, 691 (1989).
[2] H. J. Williams, R. M. Bozorth, and W. Schockley, Phys. Rev. **75**, 155 (1949) [Russ. transl. in *Physics of Ferromagnetic Domains*, IL, Moscow (1951), pp. 133–179].
[3] C. A. Fowler, E. M. Fryer, and D. Treves, J. Appl. Phys. **32**, 296S (1961).
[4] K. M. Podurets, V. A. Somenkov, R. R. Chistyakov, and S. Sh. Shil'shteĭn, Physica B **156/157**, 694 (1989).
[5] N. O. Elyutin, A. O. Bubleĭnik, V. A. Somenkov, and S. Sh. Shil'shteĭn, Pis'ma Zh. Éksp. Teor. Fiz. **18**, 318 (1973) [JETP Lett. **18**, 186 (1973)].
[6] K. M. Podurets, D. V. Sokol'skiĭ, P. P. Chistyakov, and S. Sh. Shil'shteĭn, Fiz. Tverd. Tela (Leningrad) **33**, 2954 (1991) [Sov. Phys. Solid State **33**, 1668 (1991)].
[7] D. J. Hughes, M. T. Burgy, R. B. Heller, and J. W. Wallace, Phys. Rev. **75**, 565 (1949) [Russ. transl. in *Physics of Ferromagnetic Domains*, IL, Moscow (1951), pp. 307–315].
[8] S. Sh. Shil'shteĭn, V. I. Marukhin, M. Kalanov, and V. A. Somenkov, Pribory Tekh. Éksp., No. 3, 70 (1971).

Translated by James S. Wood

Bouncing of the load and radiation of sound in dry friction

A. P. Kiselev and V. A. Lazarev

Institute of Problems in Machine Construction, Russian Academy of Sciences, 199178 St. Petersburg, Russia

(Submitted January 19, 1996)
Zh. Tekh. Fiz. **67**, 137–139 (May 1997)

[S1063-7842(97)02205-8]

Here we consider a simple model of sound generation in dry friction, taking into account ''bouncing'' of the load, i.e., vibrations of the load perpendicular to the sliding plane. We establish the fact that the bouncing effect is responsible for the radiation of sound waves with a complex, irregular velocity dependence at subsonic sliding velocities.

1. The only simple, general, and reliable result established in a millennium of research on dry friction is the Amontons-Coulomb law, which states that the force of friction is proportional to the static normal load.[1] It is natural, therefore, to be interested in the theoretical description of the dry friction process itself and its attendant phenomena on the basis of simple models. Of special note is a paper by Adirovich and Blokhintsev,[2] who have made a valiant attempt to form an elementary description of dry friction within the framework of the linear theory of ideal elasticity. The friction mechanism discussed in Ref. 2 was the radiation of sound from the contact zone.

2. We first retrace the train of thought in Ref. 2. We consider two semiinfinite, perfectly elastic bodies with periodic surfaces, one of which moves relative to the other with a subsonic sliding velocity v without coming into direct contact with it (Fig. 1). Interaction takes place through static (e.g., molecular) forces acting between the surfaces of the bodies. All that matters is the linear character of this interaction.

The first nontrivial step in Ref. 2 was to replace the effect of the moving body by a stationary equivalent load, periodic in the coordinate x and time, which is no longer applied to the rough body, but to the surface of the half-space $z>0$.

The second step was to calculate the power W of the waves radiated into the half-space $z>0$, and the third step was to set up the equality

$$W=F\cdot v, \tag{1}$$

where F is called the dry friction force.

The authors found that at low sliding velocities sound is radiated mainly by oscillatory modes of the load corresponding to normal and tangential vibrations of the surface $z=0$, and the radiated power is essentially independent of v. At least one consequence of this condition is the surprising result

$$F\sim\frac{1}{v}, \quad v\to 0. \tag{2}$$

The first of these steps cannot be implemented without taking considerable liberties, which can be avoided in the simplified contact model shown in Fig. 2. The simplicity here is excessive: Such a model totally fails to describe the generation of sound at subsonic sliding velocities.

However, the situation undergoes a qualitative change when ''bouncing'' of the load, i.e., its vibrations perpendicular to the slide plane, are taken into account. The assumption is (see Ref. 3 and the literature cited there) that bouncing is always accompanied dry friction.

3. Let us visualize one of the friction bodies as a homogeneous, perfectly elastic half-space $z>0$, over whose surface a perfectly rigid load moves in the direction of the x axis with a constant sliding velocity v (Fig. 2). The process does not depend on the coordinate y. The load, which has a periodic corrugated (rough) surface with period $2L$, acts on the body as before through surface forces.

The essence of the problem can be grasped if the processes at $z>0$ are described by the scalar equation

$$\mu\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial z^2}\right)-\rho\frac{\partial^2 u}{\partial t^2}=0, \tag{3}$$

where μ is the elastic modulus of the medium, and ρ is its density.

4. We first consider motion of the load without bouncing. Its action is modeled by the boundary condition

$$\mu\frac{\partial u}{\partial z}\bigg|_{z=0}=f(x-vt), \tag{4}$$

where the periodic function $f(x-vt)$ with period $2L$ is determined by the structure of the roughness and the form of the surface forces.

We interpret the load as one of the harmonics of the function f:

$$f(x-vt)=A_m\exp[i\Xi_m(x-vt)], \tag{5}$$

$$A_m=\text{const}; \quad \Xi_m=(m\pi)/L; \quad m=0, 1, \ldots .$$

The solution of problem (3)–(5) has the form

$$u=\frac{A_m}{i\mu\zeta_m}\exp i[\Xi_m(x-vt)+\zeta_m z], \tag{6}$$

where

$$\zeta_m=i\Xi_m\sqrt{1-\frac{v^2}{c^2}}, \tag{7}$$

and $c=\sqrt{\mu/\rho}$ is the sound velocity in the model.

In the most interesting case for the great majority of applications — a subsonic sliding velocity

 1063-7842/97/050581-03$10.00

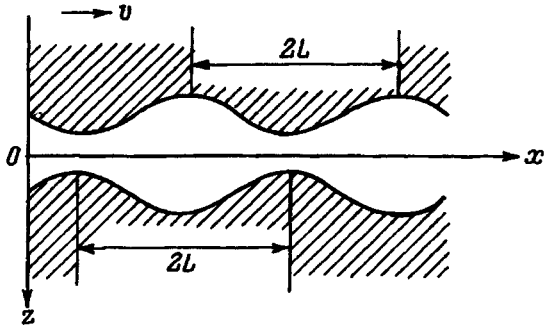

FIG. 1. Initial contact model in Adirovich and Blokhintsev's paper.

$$v<c, \tag{8}$$

ζ_m is purely imaginary. The wave process has a surface character and is not accompanied by sound radiation in the body. This is clearly also true of problem (4), (5) with any load periodic in $x-vt$.

5. We now consider the problem of a moving and simultaneously harmonically bouncing load of the form

$$\mu \frac{\partial u}{\partial z}\bigg|_{z=0} = \exp(-i\Omega t) f(x-vt), \tag{9}$$

where $f(x-vt)$, as before, is a periodic function of its argument, and Ω is the angular frequency of bouncing.

We do not assume any connection between Ω and the roughness period.

At first we confine the problem to one harmonic of the function f. The solution of problem (3), (4), and (9) has the form

$$u=\frac{A_m}{i\mu Z_m}\exp\, i[\Xi_m(x-vt)+Z_m z-\Omega t], \tag{10}$$

where

$$Z_m=\frac{i}{c}\sqrt{c^2\Xi_m^2-(\Omega+v\Xi_m)^2}. \tag{11}$$

Sound is radiated into the depth of the body by the mth bouncing mode of the load if Z_m is real, i.e., if $v>v_m$, where v_m is the corresponding critical velocity

$$v_m=c-\frac{\Omega}{\Xi_m}=c-\frac{\Omega L}{m\pi}, \tag{12}$$

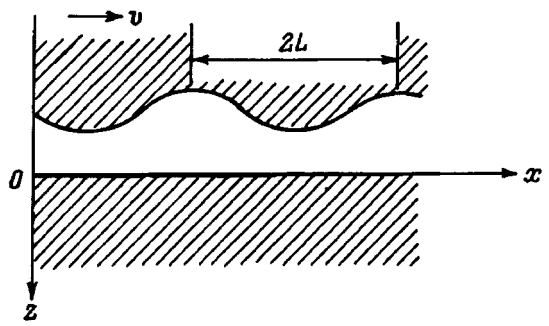

FIG. 2. Contact model used in the present paper.

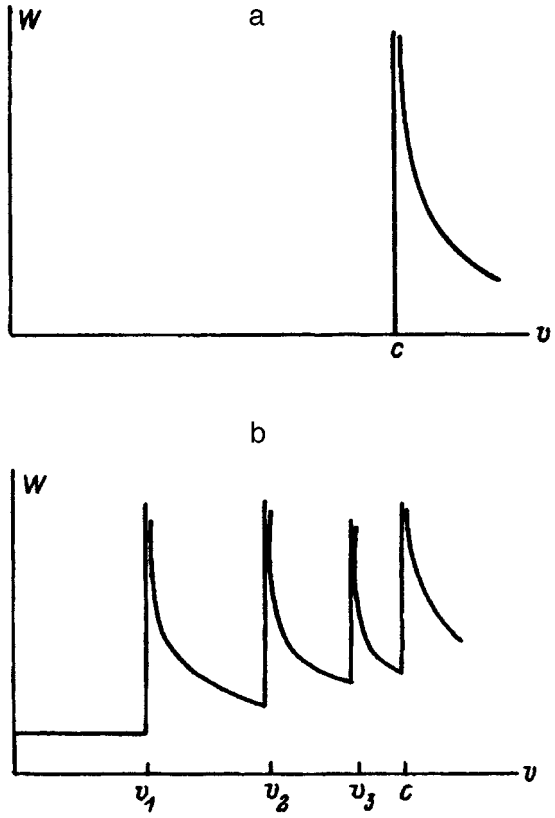

FIG. 3. Total specific power of radiated sound. a — Without bouncing; b — with bouncing.

which is lower than the sound velocity. For $m=0$ the given mode radiates sound for indefinitely small values of v.

6. It can be shown that the power density of the radiated sound, averaged over the roughness period and time, is expressed as follows for a load of the form (9) in the subsonic case (7):

$$w=\sum_{0\leqslant m\leqslant \frac{\Omega L}{\pi(c-v)}} \frac{\Xi_m v+\Omega}{\mu Z_m}|A_m|^2, \tag{13}$$

where A_m denotes the Fourier coefficients of the function f.

Figure 3 illustrates the irregular dependence of W on the sliding velocity.

The power extracted by the mode $m=0$, which corresponds to bouncing without sliding, does not depend on v. Waves corresponding to modes $m\neq 0$ remove energy from the contact zone for $v>v_m$; the contribution from each mode to the radiated power diminishes as v increases and tends to the asymptote $(c/\mu)|A_m|^2$.

The introduction of absorption (e.g., by treating complex-valued μ) makes the peaks at $v=v_m$ (Fig. 3) finite.

7. Consequently, the incorporation of bouncing in the model helps to explain several features observed in sound radiation associated with dry friction, specifically a complex and irregular dependence of the radiated power on the sliding velocity.

The analysis of more complex, linear models of an elastic body[4] alters several details, but leaves intact the above-stated conclusions as to the role of bouncing.

8. In the case of the elastic analog of the model in Fig. 2, without regard for bouncing of the load, sound is not generated at subsonic sliding velocities, and $W=0$. If bouncing is taken into account, on the other hand, the analog of the zeroth mode of the load (9) is radiated exclusively at low velocities v, where W does not depend on v (and does not vanish at $v=0$). Consequently, the quantity F introduced formally by Eq. (2) cannot in any way be interpreted as a tangential friction force.

Therefore, notwithstanding the interesting approach of Ref. 2 to the description of sound radiation, it does not provide a description of a conservative mechanism of dry friction.

The authors are grateful to É. L. Aéro, to whom we owe the idea of treating the bouncing phenomenon, and also to S. V. Krysov for a discussion of dry friction problems.

This work has received partial support from Russian Fund for Fundamental Research, Grant No. 95-02-05630.

[1] A. S. Akhmanov, *Molecular Physics of Boundary Friction* [in Russian], Fizmatgiz, Moscow (1963).
[2] E. Adirovich and D. Blokhintzev, J. Phys. USSR **7**, 29 (1943).
[3] M. E. Levinshteĭn and S. L. Rumyantsev, Pis'ma Zh. Tekh. Fiz. **18**(5), 42 (1992) [Sov. Tech. Phys. Lett. **18**, 144 (1992)].
[4] W. Nowacki, *Teoria Pelzania* (*Theory of Elasticity*), Arkady, Warsaw (1963) [Russ. transl., Mir, Moscow (1975)].

Translated by James S. Wood

Production of atomic hydrogen in an rf gas discharge and mass spectrometer diagnostics of the process

Yu. A. Akulov, B. A. Mamyrin, and P. M. Shikhaliev

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted March 7, 1996)
Zh. Tekh. Fiz. **67**, 140–142 (May 1997)

[S1063-7842(97)02305-2]

The production and diagnostics of atomic hydrogen are investigated in the present study as part of the preparation for a series of experiments on the influence of variations in the electron environment on the lifetime of the β-active tritium nucleus.[1] In the first experiment of the series it is proposed to investigate the chemical shift of the half-life for an atomic-molecular tritium pair. Since the energy of dissociation of molecular tritium is close to that of molecular protium, and since atomic tritium, because of its higher mass and lower mobility, recombines more slowly than atomic protium, the procedure used to obtain atomic tritium can be worked out in experiments with protium (called hydrogen from now on).

Atomic hydrogen is produced by thermal dissociation in a tungsten furnace,[2] by low-frequency and radio-frequency gas discharges,[3] or by means of a plasma arc.[4] In the present study we have used the rf gas discharge method to obtain atomic hydrogen. The high efficiency of this method helps to minimize the heating of the discharge cell and can be implemented without the intrusion of electrodes in the discharge volume, thereby ensuring high purity of the atomic hydrogen product.

The problems associated with the stabilization and detection of atomic hydrogen stem from its high chemical activity, by virtue of which atoms rapidly recombine at the walls of the discharge volume and in mutual collisions. For this reason, most diagnostic techniques applied to atomic hydrogen are based on measurements of secondary effects attributable to the presence of atomic hydrogen. For example, a chemical target of molybdenum oxide has been used as the atomic hydrogen detector in some studies.[3,5] When hydrogen atoms are incident on such a target, the oxide undergoes reduction, and the target changes color. However, the small range of measurement of the flux density (10^{13}–10^{15} atoms/s $\cdot$ cm^2) and low accuracy ($\pm 30\%$ error) limit the possibilities of this method. Bass and Broida[6] have measured the optical spectra of hydrogen atoms in the condensed and gaseous phases, but did not establish any relationship between the concentration of hydrogen atoms and the line intensities of the optical spectra. One approach that looks promising is the determination of hydrogen atoms by electron spin resonance (ESR). This method utilizes the presence of an unpaired electron due to the paramagnetism of the hydrogen atom, a property that can be exploited for detection by ESR. Only papers in which ESR has been used to measure atomic hydrogen in the condensed phase are known in the literature.[3]

Here we report mass spectrometer measurements of atomic hydrogen produced in an rf gas discharge. The experimental apparatus (Fig. 1) consists of an rf oscillator *1* (1 MHz, 300 W), a discharge tube *2*, an MI 1201 mass spectrometer *3*, and a system for the transport of atomic hydrogen into the mass spectrometer. The materials of the discharge tube and the transport system must have low recombination coefficients for atomic hydrogen on their surfaces. The materials with the lowest recombination coefficients are Teflon ($\sim 10^{-5}$) and molybdenum glass ($\sim 10^{-3}$).[7] The discharge cell comprises a molybdenum glass tube of length 50 cm with an inside diameter of 10 mm and a wall thickness of 1 mm. A Teflon tube of length 25 cm and inside diameter 3 mm is used to transport hydrogen atoms from the discharge tube into the mass spectrometer. The hydrogen atoms are sent from the discharge volume into the Teflon tube through a capillary orifice ($d = 0.1$ mm, $l = 2$ mm) in a Teflon stopper, which is hermetically sealed into the discharge tube. An initial H_2 pressure of 0.5 Torr ensures a molecular flow regime of hydrogen through the capillary orifice; the hydrogen flux through the ion source *7* of the mass spectrometer in this case is $\sim 5 \times 10^{15}$ atoms/s. Ionization is induced in the ion source by electron impact, and the currents in the beams of atomic and molecular hydrogen ions are measured by means of a multiplier tube and Faraday cylinders *8*.

The light emitted from the discharge tube is transmitted along an optical fiber to a silicon photodiode, which is sensitive to lines of the Balmer series of the atomic hydrogen spectrum. The leading edge of the photodiode signal, corresponding to the initiation of discharge, determines the zero point of the time scale for analyzing the dynamics of the transmission of hydrogen atoms from the discharge tube to the ion source of the mass spectrometer. The photodiode signal is also used to determine the statistical weight of the excited component of the hydrogen atoms for a high degree of dissociation of H_2 molecules.

Figure 2 shows the dependence of the ion current H^+ in the mass spectrometer on the power absorbed in the discharge tube. Discharge is triggered at a power ~ 5 W. The curve enters a plateau at power levels of 40–60 W, corresponding to the almost total dissociation of the molecular hydrogen. Since atomic hydrogen recombines mainly at the walls of the discharge tube at pressures of 0.1–0.5 Torr, a high degree of dissociation at lower absorbed power levels in the discharge can be achieved by decreasing the recombination coefficient of atoms at the walls. The probability of recombination of hydrogen atoms at the walls can be lowered by means of special coatings such as metaphosphoric acid, potassium tetraborate, or dimethylchlorosilane.[3] However, in view of the requirement for high purity of the hydrogen used in the experiment to measure the chemical shifts of the β-decay constant, we did not subject the walls of the dis-

 1063-7842/97/050584-02$10.00

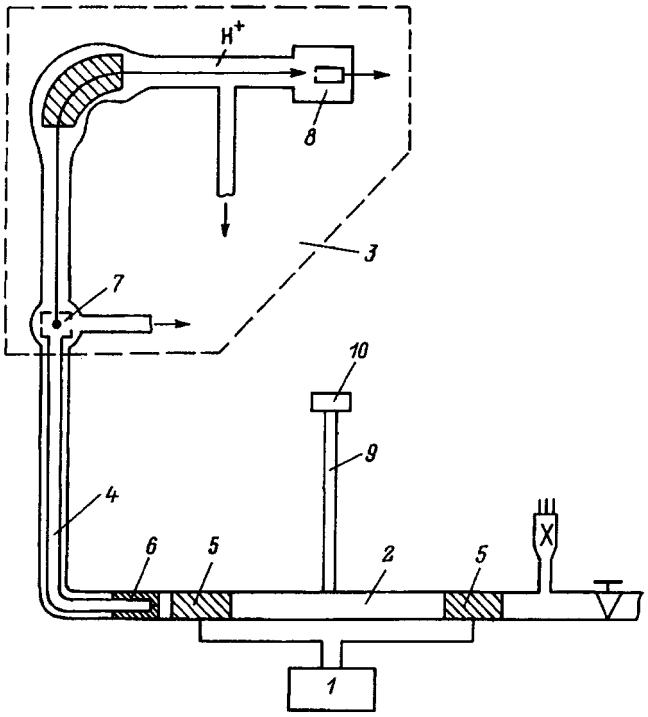

FIG. 1. Experimental apparatus. *1* — Radio-frequency oscillator; *2* — discharge tube; *3* — MI 1201 mass spectrometer; *4* — Teflon atom-transport tube; *5* — annular electrodes; *6* — Teflon stopper with a capillary orifice; *7* — ion source; *8* — multiplier tube and Faraday cylinder; *9* — optical fiber; *10* — photodiode.

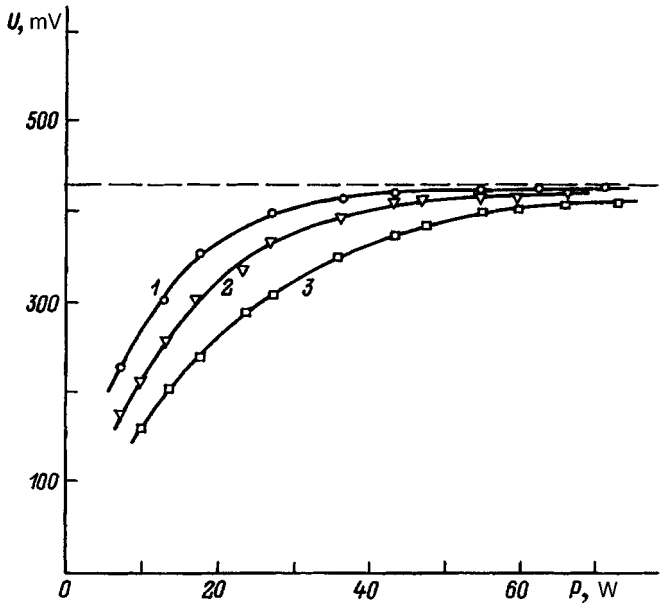

FIG. 2. Ion current H^+ in the mass spectrometer versus power absorbed in the discharge tube. *1* — $P=0.1$ Torr; *2* — 0.5 Torr; *3* — 1.0 Torr.

charge tube to any kind of chemical treatment.

The power absorbed in the discharge was determined by direct calorimetric measurements. For this purpose the discharge tube was placed in a thermostat with vacuum thermal insulation and a mirrored inner surface. The emission of atomic hydrogen in the near vacuum ultraviolet region of the spectrum was absorbed by the walls of the discharge tube, and its emission in the visible and infrared regions was reflected by the walls of the thermostat. The temperature variation in the thermostat was recorded by a special thermometer, whose scale was calibrated in power units by means of a standard heater placed in the thermostat.

We can use the curves of the degree of dissociation versus the discharge power to estimate the energy utilization efficiency, which is defined as the number of dissociation events per unit absorbed energy. In our case it is equal to 1.8×10^{-2} dissociation events/1 eV, which is consistent with a previous calculation[3] of the energy utilization efficiency for a molybdenum glass discharge tube, 2×10^{-2}, but is lower than the experimental value obtained in the same paper, 5.8×10^{-2}. The energy utilization efficiency was observed to increase in Ref. 3 after the walls of the discharge tube had been chemically treated to lower the recombination coefficient of the hydrogen atoms.

In our discharge tube geometry, stable triggering of discharge was achieved for an electrode separation of 5 cm to 50 cm; the absorbed power in this case was proportional to the indicated distance.

Since experiments involving a measurement of the chemical shifts of the decay half-life require absolutely positive identification of the state of the electron environment of the tritium nucleus, it is necessary to determine the statistical weights of the excited and ionized components of the hydrogen atoms in the discharge. Using the values of the absorbed power, the excitation energy, and the concentration and lifetime of atomic hydrogen, we can estimate the stationary fraction of excited states for 60–95% dissociation as $\sim10^{-5}$. We have also estimated the fraction of excited states from the amplitude of the photodiode signal, which is proportional to the line intensities of the Balmer series in the discharge emission and in this case has a value $\sim4\times10^{-6}$. The fraction of ionized states, estimated from the average current flowing through the discharge and the carrier mobilities is $\sim10^{-5}$. Consequently, the stationary statistical weights of excited and ionized states relative to the weight of the ground state of hydrogen atoms in a gas discharge for the indicated degrees of dissociation are negligible and cannot influence the accuracy of measurement of the chemical shift of the half-life of tritium as a result of the difference in the structure of the phase space accessible to a β-electron; according to preliminary theoretical estimates, this shift attains 0.1–0.5% level for the unexcited atomic and molecular systems.

This work has received financial support from the Russian Fund for Fundamental Research (Project No. 95-03-09329).

[1] Yu. A. Akulov, B. A. Matyrin, and P. M. Shikhaliev, Pis'ma Zh. Tekh. Fiz. **19**(18), 72 (1993) [Tech. Phys. Lett. **18**, 559 (1993)].
[2] J. Grosser and H. Haberland, Chem. Phys. Lett. **7**, 442 (1930).
[3] A. M. Bass and H. P. Broida, *Formation and Stabilization of Free Radicals* [Russian translation], IL, Moscow (1962).
[4] K. R. Way, S. Yang, and W. C. Stwalley, Rev. Sci. Instrum., **47**, No. 1049 (1936).
[5] V. B. Leonas, Usp. Fiz. Nauk **82**, 287 (1964) [Sov. Phys. Usp. **7**, 121 (1964)].
[6] A. M. Bass and H. P. Broida, Phys. Rev. **101**, 1740 (1956).
[7] H. Berg and D. Klepner, Rev. Sci. Instrum., **33**, 248 (1962).

Translated by James S. Wood