Macroscopic Einstein equations to second order in the interaction constant

A. V. Zakharov*)

Kazan State University, 420008 Kazan, Russia

(Submitted 7 February 1997)
Zh. Éksp. Teor. Fiz. **112**, 1153–1166 (October 1997)

The present paper is a direct continuation of an earlier paper [JETP **83**, 1 (1996)] devoted to the derivation of the macroscopic Einstein equations to within terms of second order in the interaction constant. Ensemble averaging of the microscopic Einstein equations and the Liouville equation for the random functions leads to a closed system of macroscopic Einstein equations and kinetic equations for one-particle distribution functions. The macroscopic Einstein equations differ from the classical equations in that their left-hand side contains additional terms due to particle interaction. The terms are traceless tensors with zero divergence. An explicit covariant expression for these terms is given in the form of momentum-space integrals of expressions dependent on one-particle distribution functions of the interacting particles of the medium. The given expressions are proportional to the cube of the Einstein constant and the square of the particle number density. The latter relationship implies that interaction effects manifest themselves in systems of very high density (the universe in the early stages of its evolution, dense objects close to gravitational collapse, etc.). *© 1997 American Institute of Physics.* [S1063-7761(97)00110-8]

1. INTRODUCTION

The present paper is a direct continuation of an earlier paper (see Ref. 1) devoted to the derivation of macroscopic Einstein equations correct to second order in the interaction constant.

The macroscopic Einstein equations reduce to

$$G_{ij}+\varphi^k_{ij;k}+\mu_{ij}=\chi T_{ij}\,,$$

where G_{ij} is the Einstein tensor, T_{ij} is the energy–momentum tensor, and χ is Einstein gravitational constant. The semicolon indicates a covariant derivative. The new terms on the left-hand side are responsible for particle interaction and are traceless tensors with zero divergence. Below we give a covariant expression for these terms in the form of momentum-space integrals of expressions dependent on one-particle distribution functions of the interacting particles of the medium. The given expressions are proportional to the cube of the Einstein constant and the square of the particle number density. The latter implies that interaction effects manifest themselves in systems of very high density (the universe in the early stages of its evolution, dense objects close to gravitational collapse, etc.) or in macrosystems consisting of highly massive objects (clusters of galaxies).

2. MACROSCOPIC EINSTEIN EQUATIONS

The method of deriving the macroscopic Einstein equations is discussed in Ref. 1. The notation used in the present paper coincides with that of Ref. 1.

Briefly, the method used to obtain the macroscopic Einstein equations consists in the following.

We start by writing the microscopic Einstein equations

$$\tilde{G}^{ij}=\chi\tilde{T}^{ij}. \tag{1}$$

Here $\tilde{G}^{ij}$ is the Einstein tensor in a Riemannian space with metric $\tilde{g}_{ij}$, $\chi=8\pi k/c^4$ is Einstein's constant (where k is the gravitational constant), $\tilde{T}^{ij}$ is the microscopic energy–momentum tensor,

$$\tilde{T}^{ij}=\sum_a c\int\frac{d^4\tilde{p}_a}{\sqrt{-\tilde{g}}}\,\tilde{p}^i_a\tilde{u}^j_a N_a(q^i,\tilde{p}_i), \tag{2}$$

where $\tilde{g}$ is the determinant of $\tilde{g}_{ij}$, $\tilde{p}^i_a$ is the momentum of particles of species a, $\tilde{u}^{\,i}_{\,a}=\tilde{p}^{\,i}_{\,a}/\sqrt{\tilde{g}_{kj}\tilde{p}^{\,k}_{\,a}\tilde{p}^{\,j}_{\,a}}$, and $N_a(q^i,\tilde{p}^{\,i}_{\,a})$ is the Klimontovich random function:[2]

$$\tilde{N}_a(q^i,\tilde{p}_j)=\sum_{i=1}^{n_a}\int d\tilde{s}\,\delta^4(q^i-q^i_{(l)})\,\delta^4(\tilde{p}_j-\tilde{p}^{(l)}_j(\tilde{s})), \tag{3}$$

where n_a is the number of particles belonging to species a, $\tilde{s}$ is the canonical parameter along the particle trajectories, $d\tilde{s}=\sqrt{\tilde{g}_{ij}dq^idq^j}$, and $q^i_{(l)}$ and $\tilde{p}^{(l)}_j$ are the coordinates and momentum of the lth particle of the a species, which are found by solving the equations of motion (the geodesic equations). The function $\tilde{N}_a$ satisfies the Liouville equation

$$\tilde{p}^{\,i}\frac{\partial\tilde{N}_a}{\partial q^i}+\tilde{\Gamma}_{j,ik}\tilde{p}^j\tilde{p}^k\frac{\partial\tilde{N}_a}{\partial\tilde{p}_i}=0. \tag{4}$$

Next we write the metric $\tilde{g}_{ij}$ as

$$\tilde{g}_{ij}=g_{ij}+h_{ij}\,, \tag{5}$$

where $g_{ij}=\langle\tilde{g}_{ij}\rangle$ is the ensemble average of the metric.[1,2] In the system of equations (1), (2), and (4) we go from quantities measured in the metric $\tilde{g}_{ij}$ to quantities measured in the metric g_{ij}. As a result Eqs. (1), (2), and (4) become

$$R_{ij}+\nabla_m\Omega^m_{ij}-\nabla_j\Omega^m_{im}+\Omega^m_{mn}\Omega^n_{ij}-\Omega^m_{jn}\Omega^n_{im}$$

$$=\chi\sum_a \int \frac{d^4p}{\sqrt{(-g)}}\alpha\sqrt{\frac{g}{\tilde{g}}}\Bigg[\tilde{g}_{ik}\tilde{g}_{jm} - \frac{1}{2}\tilde{g}_{ij}\tilde{g}_{km}\Bigg]p^kp^mN_a, \tag{6}$$

$$p^i\frac{\partial N_a}{\partial q^i}+\Gamma_{j,ik}p^kp^j\frac{\partial N_a}{\partial p_i}=\frac{\partial}{\partial p_i}(\Omega^m_{jk}\Delta_{mi}p^jp^kN_a). \tag{7}$$

Here p^i is the momentum measured in the metric g_{ij}, N_a is a random function in the space with metric g_{ij} (the expression for N_a can be obtained from (3) by dropping all tildes), $\Omega^m_{kj}=\tilde{\Gamma}^m_{kj}-\Gamma^m_{kj}$ is the difference between the Christoffel symbols of the second kind for the metrics $\tilde{g}_{ij}$ and g_{ij}, $\Delta_{ij}=g_{ij}-u_iu_j$, and $u_i=p_i/\sqrt{p_lp^l}$.

In the next step we must expand Eqs. (6) in power series in the h_{ij} to within second-order terms and find the ensemble averages of the resulting equations. If we limit ourselves to averaged equations correct to second order in the interaction constant, we arrive at a closed system of equations for the one-particle distribution function $f_a(q,p)=\langle N_a\rangle/n_a$ and the averaged metric g_{ij}.

The equation for f_a was derived in Refs. 1 and 3 and has the following form:

$$u^i\frac{\partial f_a}{\partial q^i}+\Gamma_{j,ik}\frac{\partial f_a}{\partial p_i}=\sum_b\int\frac{d^4p}{\sqrt{(-g)}}E_{ij}(p,p')\times\left(\frac{\partial f_a}{\partial p_j}f'_b-\frac{\partial f'_b}{\partial p'_j}f_a\right), \tag{8}$$

where

$$E_{ij}(p,p')=\frac{2\pi k^2Ln_b}{c^6[(u,u')^2-1]^{3/2}}[2(u,u')(p,p')-(u,p)\times(u',p')]^2\{-g_{ij}[(u,u')^2-1]-u_iu_j-u'_iu'_j+(u,u')(u_iu'_j+u'_iu_j)\}, \tag{9}$$

with $(u,u')=u'_iu^i$, $(u,u)=u_iu^i$, etc. Primed quantities refer to particles belonging to species b, nonprimed to particles belonging to species a, and L is an analog of the Coulomb logarithm:[3,4]

$$L=\int_{k_{\min}}^{k_\infty}\frac{dk}{k}. \tag{10}$$

The equations for the averaged metric g_{ij} reduce in Ref. 1 to the following form:

$$R_{ij}+\Lambda_{ij}=\chi(T_{ij}-\tfrac{1}{2}Tg_{ij}), \tag{11}$$

where T_{ij} is the macroscopic energy–momentum tensor, $T=g^{ij}T_{ij}$ is the trace of that tensor, R_{ij} is the Ricci tensor, and Λ_{ij} is an additional term on the right-hand side of the Einstein equations, which emerges after the terms quadratic in h_{ij} are averaged over the ensemble,

$$\Lambda_{ij}=(\delta^k_n\delta^s_j-\delta^k_j\delta^s_n)[-\tfrac{1}{2}\nabla_kP^n_{is}+Q^n_{kis}]+\lambda_{ij}. \tag{12}$$

Here

$$P^n_{is}=\langle h^n_l\Omega^l_{is}\rangle, \tag{13}$$

$$Q^n_{kis}=\langle\Omega^n_{kl}\Omega^l_{is}\rangle, \tag{14}$$

and

$$\lambda_{ij}=-\sum_a\frac{\chi}{m_a}\int\frac{d^4p}{\sqrt{(-g)}}\Bigg\{-\frac{1}{2}p_ip_ju^ku^m-\frac{1}{4}g_{ij}p^kp^m-\frac{1}{2}p_ip_jg^{km}+\frac{1}{4}m_a^2c^2g_{ij}g^{km}+p_ip^k\delta^m_j+p_jp^k\delta^m_i-\frac{1}{2}m_a^2c^2\delta^k_i\delta^m_j\Bigg\}\langle N_ah_{km}\rangle. \tag{15}$$

Indices are raised and lowered using the metric g_{ij}.

To simplify (13) and (14) still further, we only have to calculate h_{ij} and Ω^l_{ij} inside the region determined by the correlation radius and the corresponding correlation time. Only distant collisions provide the main contribution to the calculated macroscopic quantities. To allow for this contribution it is enough to find h_{ij} from the Einstein equations linearized with respect to the metric g_{ij}. Since we assume that the metric inside the correlation region is constant, we can introduce a locally Lorentzian reference frame near the point at which we wish to calculate the values of h_{ij}.

Thus, subsequent calculations do not have a covariant form, but they are all done for the purpose of determining the components of the tensor Λ_{ij} at some (arbitrary) point q in the selected reference frame, in which $g_{ij}=\eta_{ij}$ is the Minkowski tensor. In this reference frame the interval ds^2 has the form

$$ds^2=(d\eta)^2-(dq^1)^2-(dq^2)^2-(dq^3)^2.$$

The final result must be written in covariant form.

The expressions for h_{ij} and Ω^l_{ij}, derived from the Einstein equations linearized with respect to the Minkowski metric (which we still denote by g_{ij}) were found in Ref. 1:

$$h_{ij}(\eta,\mathbf{q})=\sum_b\int d^4p'\int d^3q'\int d^3k\int_{-\infty}^{\eta}d\eta'\times e^{-i\mathbf{k}\cdot(\mathbf{q}-\mathbf{q}')}h^{(b)}_{ij}(\eta,\eta',p',\mathbf{k})\Phi_b(\eta',\mathbf{q}',p'), \tag{16}$$

$$\Omega^i_{jk}(\eta,\mathbf{q})=\sum_b\int d^4p'\int d^3q'\int d^3k\int_{-\infty}^{\eta}d\eta'\times e^{-i\mathbf{k}\cdot(\mathbf{q}-\mathbf{q}')}\Omega^{i(b)}_{jk}(\eta,\eta',p',\mathbf{k})\times\Phi_b(\eta',\mathbf{q}',p'), \tag{17}$$

where $\mathbf{q}=(q^1,q^2,q^3)$ is a three-dimensional radius vector in the given reference frame, and $\mathbf{k}=(k_1,k_2,k_3)$. Here $\Phi_b=N_b-n_bf_b$,

$$h^{(b)}_{ij}(\eta,\eta',p',\mathbf{k})=-\frac{i\chi m_bc^2}{(2\pi)^3k}\left(u'_iu'_j-\frac{1}{2}g_{ij}\right)\{e^{ik(\eta'-\eta)}-e^{-ik(\eta'-\eta)}\}, \tag{18}$$

and

$$\Omega^{i(b)}_{jk}(\eta,\eta',p',\mathbf{k})=\frac{\chi m_bc^2}{2(2\pi)^3k}\Bigg\{\Bigg[\left(u'_ju'_k-\frac{1}{2}g_{jk}\right)k^i_+$$

$$
\begin{aligned}
&-\Bigl(u'_j u'^{i}-\frac{1}{2}\delta^i_j\Bigr)k^+_k-\Bigl(u'_k u'^{i}\\
&-\frac{1}{2}\delta^i_k\Bigr)k^+_j\Bigr]e^{ik(\eta'-\eta)}-\Bigl[\Bigl(u'_j u'_k\\
&-\frac{1}{2}g_{jk}\Bigr)k^i_- -\Bigl(u'_j u'^{i}-\frac{1}{2}\delta^i_j\Bigr)k^-_k\\
&-\Bigl(u'_k u'^{i}-\frac{1}{2}\delta^i_k\Bigr)k^-_j\Bigr]e^{ik(\eta'-\eta)}\Bigr\}.
\end{aligned}
\tag{19}
$$

In (18) and (19) we introduced the following vectors:

$$k^+_i=(k,\mathbf{k}),\quad k^-_i=(-k,\mathbf{k}),$$

where

$$k=\sqrt{[(k_1)^2+(k_2)^2+(k_3)^2]}=|\mathbf{k}|.$$

Obviously, $k^-_i(\mathbf{k})=-k^+_i(-\mathbf{k})$.

Substitution of (16)–(19) in (13)–(15) yields the following expressions for p^n_{is}, Q^n_{kis}, and λ_{ij}:

$$
\begin{aligned}
P^n_{is}=&\sum_{bc}\int d^4p'\int d^4p''\int d^3q'\int d^3q''\int_{-\infty}^{\eta}d\eta'\int_{-\infty}^{\eta}d\eta''\\
&\times\int d^3k'\int d^3k''e^{-i\mathbf{k}'\cdot(\mathbf{q}-\mathbf{q}')}e^{-i\mathbf{k}''\cdot(\mathbf{q}-\mathbf{q}'')}h^{n(b)}_l\\
&\times(\eta,\eta',p',\mathbf{k}')\Omega^{l(c)}_{is}(\eta,\eta'',p'',\mathbf{k}'')n_bn_cg_{bc}(x',x''),
\end{aligned}
\tag{20}
$$

$$
\begin{aligned}
Q^n_{kis}=&\sum_{bc}\int d^4p'\int d^4p''\int d^3q'\int d^3q''\int_{-\infty}^{\eta}d\eta'\\
&\times\int_{-\infty}^{\eta}d\eta''\int d^3k'\int d^3k''e^{-i\mathbf{k}'\cdot(\mathbf{q}-\mathbf{q}')}e^{-i\mathbf{k}''\cdot(\mathbf{q}-\mathbf{q}'')}\\
&\times\Omega^{n(b)}_{kl}(\eta,\eta',p',\mathbf{k}')\Omega^{l(c)}_{is}\\
&\times(\eta,\eta'',p'',\mathbf{k}'')n_bn_cg_{bc}(x',x''),
\end{aligned}
\tag{21}
$$

$$
\begin{aligned}
\lambda_{ij}=&-\sum_{ab}\frac{\chi}{m_a}\int d^4p\int d^4p'\int d^3q'\int_{-\infty}^{\eta}d\eta'\\
&\times\int d^3k'e^{-i\mathbf{k}'\cdot(\mathbf{q}-\mathbf{q}')}\Bigl\{-\frac{1}{2}p_ip_ju^ku^m-\frac{1}{4}g_{ij}p^kp^m\\
&-\frac{1}{2}p_ip_jg^{km}+\frac{1}{4}m_a^2c^2g_{ij}g^{km}+p_ip^k\delta^m_j+p_jp^k\delta^m_i\\
&-\frac{1}{2}m_a^2c^2\delta^k_i\delta^m_j\Bigr\}h^{(b)}_{km}(\eta,\eta',p',\mathbf{k}')n_an_bg_{ab}(x,x').
\end{aligned}
\tag{22}
$$

In these expressions unprimed, primed, and double-primed quantities refer to particles belonging to species a, species b, and species c, respectively.

The quantities $g_{ab}(x,x')$ in (20)–(22) are two-particle correlation functions, which inside the correlation region have the following form:[1,3]

$$
\begin{aligned}
g_{ab}(x,x')=&\int d^3k\int_{-\infty}^{\eta}\frac{d\tau}{p^0}\Bigl[\frac{\partial}{\partial p_i}(p^lp^m\Delta_{ij}f_a(x))\Bigr]_{\tau}\\
&\times\int_{-\infty}^{\tau}\frac{d\tau'}{u'^0}f_b(x')\Omega^{j(b)}_{lm}(\tau,\tau',p',\mathbf{k})\\
&\times\exp\Bigl[-i\mathbf{k}\cdot(\mathbf{q}-\mathbf{q}')+\frac{i}{c}(\mathbf{k}\cdot\mathbf{v})(\eta-\tau)\\
&+\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}')(\tau'-\eta')\Bigr]+\int d^3k\int_{-\infty}^{\eta'}\frac{d\tau'}{p'^0}\\
&\times\Bigl[\frac{\partial}{\partial p'_i}(p'^lp'^m\Delta'_{ij}f_b(x'))\Bigr]_{\tau'}\int_{-\infty}^{\tau'}\frac{d\tau}{u^0}f_a(x)\\
&\times\Omega^{j(a)}_{lm}(\tau',\tau,p,\mathbf{k})\exp\Bigl[-i\mathbf{k}\cdot(\mathbf{q}'-\mathbf{q})\\
&+\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}')(\eta'-\tau')+\frac{i}{c}(\mathbf{k}\cdot\mathbf{v})(\tau-\eta)\Bigr].
\end{aligned}
\tag{23}
$$

Here τ' indicates that into the corresponding expression dependent on the primed coordinates $\mathbf{q}'$ and momenta p'_i at time η' we must substitute τ' and $\mathbf{q}'+\mathbf{v}'(\tau'-\eta')/c$ for η' and $\mathbf{q}'$, respectively. The index τ has a similar meaning.

If we now substitute (23) in (20)–(22) and integrate with respect to $\mathbf{q}'$, $\mathbf{q}''$, $\mathbf{k}'$, and $\mathbf{k}''$, we arrive at the following expressions for P^n_{is}, Q^n_{kis}, and λ_{ij}:

$$
\begin{aligned}
P^n_{is}=&\sum_{bc}n_bn_c(2\pi)^6\int d^4p'\int d^4p''\int_{-\infty}^{\eta}d\eta'\int_{-\infty}^{\eta}d\eta''\\
&\times\int d^3k\Bigl\{\int_{-\infty}^{\eta'}\frac{d\tau'}{p'^0}\Bigl[\frac{\partial}{\partial p'_t}(p'^rp'^m\Delta'_{tj}f_b(x'))\Bigr]_{\tau'}\\
&\times\int_{-\infty}^{\tau'}\frac{d\tau''}{u''^0}f_c(x'')\Omega^{j(c)}_{rm}(\tau',\tau'',p'',-\mathbf{k})\\
&\times\exp\Bigl[\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}'')(\eta''-\tau'')+\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}')(\tau'-\eta')\Bigr]\\
&+\int_{-\infty}^{\eta''}\frac{d\tau''}{p^{0\prime\prime}}\Bigl[\frac{\partial}{\partial p''_t}(p''^rp''^m\Delta''_{tj}f_c(x''))\Bigr]_{\tau''}\\
&\times\int_{-\infty}^{\tau''}\frac{d\tau'}{u'^0}f_b(x')\Omega^{j(b)}_{rm}(\tau'',\tau',p',\mathbf{k})\\
&\times\exp\Bigl[\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}'')(\eta''-\tau'')+\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}')(\tau'-\eta')\Bigr]\Bigr\}\\
&\times h^{n(b)}_l(\eta,\eta',p',-\mathbf{k})\Omega^{l(c)}_{is}(\eta,\eta'',p'',\mathbf{k}),
\end{aligned}
\tag{24}
$$

$$
\begin{aligned}
Q^n_{kis}=&\sum_{bc}n_bn_c(2\pi)^6\int d^4p'\int d^4p''\int_{-\infty}^{\eta}d\eta'\int_{-\infty}^{\eta}d\eta''\\
&\times\int d^3k\Bigl\{\int_{-\infty}^{\eta'}\frac{d\tau'}{p'^0}\Bigl[\frac{\partial}{\partial p'_t}(p'^rp'^m\Delta'_{tj}f_b(x'))\Bigr]_{\tau'}\\
&\times\int_{-\infty}^{\tau'}\frac{d\tau''}{u''^0}f_c(x'')\Omega^{j(c)}_{rm}(\tau',\tau'',p'',-\mathbf{k})
\end{aligned}
$$

$$
\exp\left[\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}'')(\eta''-\tau'')+\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}')(\tau'-\eta')\right]
$$

$$
+\int_{-\infty}^{\eta''}\frac{d\tau''}{p^{0''}}\left[\frac{\partial}{\partial p''_t}(p''^r p''^m \Delta''_{tj} f_c(x''))\right]_{\tau''}\int_{-\infty}^{\tau''}\frac{d\tau'}{u'^0}
$$

$$
\times f_b(x')\Omega_{rm}^{j(b)}(\tau'',\tau',p',\mathbf{k})\exp\left[\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}'')(\eta''-\tau'')\right.
$$

$$
\left.\left.+\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}')(\tau'-\eta')\right]\right\}\Omega_{kl}^{n(b)}(\eta,\eta',p',-\mathbf{k})\Omega_{is}^{l(c)}
$$

$$
\times(\eta,\eta'',p'',\mathbf{k}), \tag{25}
$$

$$
\lambda_{ij}=-\sum_{bc}\frac{\chi(2\pi)^3 n_b n_c}{m_c}\int d^4p'\int d^4p''\left[-\frac{1}{2}p''_i p''_j u''^k u''^m\right.
$$

$$
-\frac{1}{4}g_{ij}p''^k p''^m-\frac{1}{2}p''_i p''_j g^{km}+\frac{1}{4}m_c^2c^2 g_{ij}g^{km}+p''_i p''^k\delta_j^m
$$

$$
\left.+p''_j p''^k\delta_i^m-\frac{1}{2}m_c^2c^2\delta_i^k\delta_j^m\right]\int_{-\infty}^{\eta}d\eta'\int d^3k h_{km}^{(b)}(\eta,\eta',p,
$$

$$
-\mathbf{k})\left\{\int_{-\infty}^{\eta'}\frac{d\tau'}{p'^0}\left[\frac{\partial}{\partial p'_r}(p'^l p'^s\Delta'_{nr}f_b(x'))\right]_{\tau'}\right.
$$

$$
\times\int_{-\infty}^{\tau'}\frac{d\tau''}{u''^0}f_c(x'')\Omega_{ls}^{n(c)}(\tau',\tau'',p'',-\mathbf{k})
$$

$$
\times\exp\left[\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}'')(\eta-\tau'')+\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}')(\tau'-\eta')\right]
$$

$$
+\int_{-\infty}^{\eta}\frac{d\tau''}{p''^0}\left[\frac{\partial}{\partial p''_r}(p''^l p''^s\Delta''_{nr}f_c(x''))\right]_{\tau'}
$$

$$
\times\int_{-\infty}^{\tau''}\frac{d\tau'}{u'^0}f_b(x')\Omega_{ls}^{n(b)}(\tau'',\tau',p',\mathbf{k})
$$

$$
\left.\times\exp\left[\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}'')(\eta-\tau'')+\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}')(\tau'-\eta')\right]\right\}. \tag{26}
$$

To simplify (24)–(26) still further, we proceed as follows. First, we assume that inside the correlation region the distribution function changes little, so that in calculating the integrals in (24)–(26) we can ignore, in the first approximation, the dependence of f on the temporal coordinate. We substitute the explicit expressions for $h_{ij}^{(b)}$ and $\Omega_{kj}^{i(b)}$ (Eqs. (18) and (19)) in (24)–(26) and evaluate the integrals with respect to τ', τ'', η', η'', and $\mathbf{k}$. Then the expression for P_{is}^n becomes

$$
P_{is}^n=\sum_{bc}\frac{\chi^3 m_b m_c n_b n_c c^6}{8(2\pi)^3}\int d^4p'
$$

$$
\times\int d^4p'' f_b(x')f_c(x'')[1-10(u'u'')^2]\left(u'^n u'_l-\frac{1}{2}\delta_l^n\right)
$$

$$
\times\left[\left(u''_i u''_s-\frac{1}{2}g_{is}\right)g^{lf}-\left(u''^l u''_i-\frac{1}{2}\delta_i^l\right)\right.
$$

$$
\left.\times\delta_s^f-\left(u''^l u''_s-\frac{1}{2}\delta_s^l\right)\delta_i^f\right](m_c u'^m K_{fm}^{(1)}(u',u'')
$$

$$
+m_b u''^m K_{fm}^{(2)}(u',u''))-\sum_{bc}\frac{\chi^3 m_b m_c n_b n_c c^7}{4(2\pi)^3}
$$

$$
\times\int d^4p'\int d^4p''\left(u'^n u'_l-\frac{1}{2}\delta_l^n\right)\left[\left(u''_i u''_s-\frac{1}{2}g_{is}\right)g^{lf}\right.
$$

$$
\left.-\left(u''^l u''_i-\frac{1}{2}\delta_i^l\right)\delta_s^f-\left(u''^l u''_s-\frac{1}{2}\delta_s^l\right)\delta_i^f\right]
$$

$$
\times\left\{\left[(u'u'')^2(\delta_j^m+u'_j u'^m)-\frac{1}{2}(\delta_j^m-u'_j u'^m)\right.\right.
$$

$$
\left.-2(u'u'')u'^m u''_j\right]f_c(x'')\frac{\partial f_b(x')}{\partial p'_j}K_{fm}^{(1)}(u',u'')
$$

$$
+\left[(u'u'')^2(\delta_j^m+u''_j u''^m)-\frac{1}{2}(\delta_j^m-u''_j u''^m)\right.
$$

$$
\left.\left.-2(u'u'')u''^m u'_j\right]f_b(x')\frac{\partial f_c(x'')}{\partial p''_j}K_{fm}^{(2)}(u',u'')\right\}. \tag{27}
$$

Here we have introduced the notation $K_{fm}^{(1)}(u',u'')$ and $K_{fm}^{(2)}\times(u',u'')$ for the tensors that in a locally Lorentzian reference frame, in which $g_{ij}=\eta_{ij}$ is the Minkowski tensor, have the following form:

$$
K_{fm}^{(1)}(u',u'')
$$

$$
=\frac{1}{u'^0u''^0}\int\frac{d^3k}{k^3}\int_{-\infty}^{\eta}d\eta'\int_{-\infty}^{\eta}d\eta''\int_{-\infty}^{\eta'}d\tau'\int_{-\infty}^{\tau'}d\tau''
$$

$$
\times(e^{ik(\eta'-\eta)}-e^{-ik(\eta'-\eta)})(k_f^+e^{ik(\eta''-\eta)}-k_f^-e^{-ik(\eta''-\eta)})
$$

$$
\times(k_m^+e^{-ik(\tau''-\tau')}-k_m^-e^{ik(\tau''-\tau')})\exp\left[\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}'')(\eta''-\tau'')\right.
$$

$$
\left.+\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}')(\tau'-\eta')\right],
$$

$$
K_{fm}^{(2)}(u',u'')
$$

$$
=\frac{1}{u'^0u''^0}\int\frac{d^3k}{k^3}\int_{-\infty}^{\eta}d\eta'\int_{-\infty}^{\eta}d\eta''\int_{-\infty}^{\eta''}d\tau''\int_{-\infty}^{\tau''}d\tau'
$$

$$
\times(e^{ik(\eta'-\eta)}-e^{-ik(\eta'-\eta)})(k_f^+e^{ik(\eta''-\eta)}-k_f^-e^{-ik(\eta''-\eta)})
$$

$$
\times(k_m^+e^{-ik(\tau''-\tau')}-k_m^-e^{ik(\tau''-\tau')})\exp\left[\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}'')(\eta''-\tau'')\right.
$$

$$
\left.+\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}')(\tau'-\eta')\right].
$$

After we carry out the integrals with respect to τ', τ'', η', and η'' these expressions assume the form

$$
K_{fm}^{(1)}(u',u'')
$$

$$
=\frac{2\pi c^5}{u'^0u''^0}\int\frac{d^3k}{k^2}\delta(\mathbf{k}\cdot\mathbf{v}''-\mathbf{k}\cdot\mathbf{v}')
$$

$$\times\frac{k_f^+k_m^+}{(kc-\mathbf{k}\cdot\mathbf{v}'')(kc+\mathbf{k}\cdot\mathbf{v}'')^3}+\frac{k_f^+k_m^-+k_f^-k_m^+}{(kc-\mathbf{k}\cdot\mathbf{v}'')^2(kc+\mathbf{k}\cdot\mathbf{v}'')^2}$$

$$\left.+\frac{k_f^-k_m^-}{(kc-\mathbf{k}\cdot\mathbf{v}'')^3(kc+\mathbf{k}\cdot\mathbf{v}'')}\right\}=K_{fm}(u',u''),\qquad(28)$$

$$K_{fm}^{(2)}(u',u'')=-K_{fm}^{(1)}(u',u'')=-K_{fm}(u',u'').\qquad(29)$$

The above equalities hold only in a locally Lorentzian reference frame. To obtain covariant expressions for the tensors $K_{fm}^{(1)}(u',u'')$ and $K_{fm}^{(2)}(u',u'')$, we take into account the following fact. The quantities $K_{fm}^{(1)}(u',u'')$ and $K_{fm}^{(2)}(u',u'')$ appeared in (27) after the correlation function $g_{ab}(x',x'')$ was substituted in (20) and the result was integrated with respect to $\mathbf{q}'$, $\mathbf{q}''$, $\mathbf{k}'$, and $\mathbf{k}''$. But the expression (23) for the two-particle correlation function is a sum of two terms, which differ in that primed quantities referring to particles of species a are replaced by double-primed quantities referring to particles of species b, and vice versa. It is after these terms were integrated with respect to $\mathbf{q}'$, $\mathbf{q}''$, $\mathbf{k}'$, and $\mathbf{k}''$ that $K_{fm}^{(1)}$ $\times(u',u'')$ and $K_{fm}^{(2)}(u',u'')$ appeared in (27). Obviously, the two must be calculated in the same reference frame, for which it is convenient to take the center-of-mass reference frame, in which

$$\mathbf{v}'=\mathbf{v},\quad \mathbf{v}''=-\mathbf{v},\quad u'^0=u''^0=\frac{1}{\sqrt{1-v^2/c^2}}=u^0.$$

In this reference frame,

$$K_{00}=K_{0\alpha}=0,\quad K_{\alpha\beta}=\frac{2\pi^2c}{vu_0^2k_{\min}^2}\left(\delta_{\alpha\beta}-\frac{v_\alpha v_\beta}{v^2}\right).\qquad(30)$$

Here $v=\sqrt{v_1^2+v_2^2+v_3^2}$, where $v_\alpha=v^\alpha=u^\alpha/u^0$ are the spatial components of the vector $\mathbf{v}$.

A covariant generalization of (30) has the form

$$K_{ij}(u'u'')=\frac{4\pi^2}{k_{\min}^2[(u'u'')^2-1]^{3/2}}\{-[(u'u'')^2-1]g_{ij}$$

$$-u_i'u_j'-u_i''u_j''+(u'u'')(u_i'u_j''+u_i''u_j')\}.$$

$$(31)$$

The expressions for $K_{fm}^{(1)}(u',u'')$ and $K_{fm}^{(2)}(u',u'')$ diverge as $k\rightarrow 0$, i.e., for large impact parameters. The reason is that we integrate over an infinite region, while actually we should integrate only over the correlation region, where the metric is assumed to vary only weakly. This difficulty is resolved, just as it is when the kinetic equation is derived, by introduce a cutoff procedure in the divergent integral $\int_0^\infty k^{-3}dk$.

We set the lower integration limit to $k_{\min}=1/r_{\max}$ rather than zero, where $r_{\max}$ is the size of the correlation region (the correlation radius). Then the above integral assumes the value $1/2k_{\min}^2=(1/2)r_{\max}^2$. As the experience of deriving the relativistic kinematic equation[4,5] shows, more thorough investigations suggest that the integrals become convergent as $r\rightarrow\infty$, with the contribution from the region where $r>r_{\max}$ being infinitesimal. In Refs. 4 and 5 there are estimates for $r_{\max}$ in the case where the average metric g_{ij} is the metric of the isotropic cosmological model.

The tensor (31) possesses the following properties:

$$K_{ij}(u',u'')=K_{ij}(u'',u'),$$

$$K_{ij}u'^i=K_{ij}u''^i=0,\quad K_{ij}=K_{ji}.\qquad(32)$$

Because of this the expression for P_{is}^n simplifies considerably. The macroscopic Einstein equations incorporate not P_{is}^n but the tensor $\varphi_{ij}^k=-\frac{1}{2}(\delta_n^k\delta_j^s-\delta_j^k\delta_n^s)$. The expression for this tensor can be written as follows:

$$\varphi_{ij}^k=-\sum_{bc}\frac{\chi^3m_b^2m_c^2n_bn_cc^7}{8(2\pi)^3}$$

$$\times\int\frac{d^4p'}{\sqrt{(-g)}}\int\frac{d^4p''}{\sqrt{(-g)}}\left[\frac{1}{2}g^{fk}u_i''u_j''+u'^k(u'u'')\right.$$

$$\left.\times(\delta_j^fu_i''+\delta_i^fu_j'')\right]\left((u'u'')^2-\frac{1}{2}\right)K_{fr}(u',u'')$$

$$\times\left(f_c(x'')\frac{\partial f_b(x')}{\partial p_r'}-f_b(x')\frac{\partial f_c(x'')}{\partial p_r''}\right).\qquad(33)$$

Note that

$$g^{ij}\varphi_{ij}^k=0,\quad \varphi_{ij}^i=0,\quad \varphi_{ij}^k=\varphi_{ji}^k.\qquad(34)$$

Reasoning along similar lines, we can simplify the expression for the tensor

$$\mu_{ij}=(\delta_n^k\delta_j^s-\delta_j^k\delta_n^s)Q_{kis}^n+\lambda_{ij},\qquad(35)$$

which assumes the following form:

$$\mu_{ij}=\sum_{bc}\frac{\chi^3m_bm_c\eta_b\eta_cc^6}{16(2\pi)^3}$$

$$\times\int\frac{d^4p'}{\sqrt{(-g)}}\int\frac{d^4p''}{\sqrt{(-g)}}f_b(x')f_c(x'')(1-10(u'u'')^2)$$

$$\times\left\{u'^qu_j'\delta_i^r-u''^ru_j''\delta_i^q+g^{qr}\left[\left((u'u'')^2\right.\right.\right.$$

$$\left.\left.+\frac{1}{2}\right)u_i''u_j''+\frac{1}{2}\left((u'u'')^2-\frac{1}{2}\right)g_{ij}-2(u'u'')u_i'u_j''\right]$$

$$\left.-\left((u'u'')^2-\frac{1}{2}\right)\delta_i^q\delta_j^r\right\}(m_cu'^mJ_{rqm}^{(1)}(u',u'')$$

$$+m_bu''^mJ_{rqm}^{(2)}(u',u''))$$

$$-\sum_{bc}\frac{\chi^3m_b^2m_c^2n_bn_cc^7}{8(2\pi)^3}\int\frac{d^4p'}{\sqrt{(-g)}}\int\frac{d^4p''}{\sqrt{(-g)}}$$

$$\times\left\{u'^qu_j'\delta_i^r-u''^ru_j''\delta_i^q+g^{qr}\left[\left((u'u'')^2+\frac{1}{2}\right)u_i''u_j''\right.\right.$$

$$\left.+\frac{1}{2}\left((u'u'')^2-\frac{1}{2}\right)g_{ij}-2(u'u'')u_i'u_j''\right]$$

$$\left.-\left((u'u'')^2-\frac{1}{2}\right)\delta_i^q\delta_j^r\right\}\left\{\left[\left((u'u'')^2-\frac{1}{2}\right)\delta_f^m\right.\right.$$

$$\left.+\left((u'u'')^2+\frac{1}{2}\right)u_f'u'^m-2(u'u'')u_f''u'^m\right]J_{rqm}^{(1)}$$

$$\times(u',u'')f_c(x'')\,\frac{\partial f_b(x')}{\partial p'_f}+\left[\left((u'u'')^2-\frac{1}{2}\right)\delta_f^m\right.$$

$$\left.+\left((u'u'')^2+\frac{1}{2}\right)u''_f u''^m-2(u'u'')u'_f u''^m\right]J^{(2)}_{rqm}$$

$$\left.\times(u',u'')f_b(x')\,\frac{\partial f_c(x'')}{\partial p''_f}\right\}. \tag{36}$$

Here we have introduced the notation $J^{(1)}_{rqm}(u',u'')$ and $J^{(2)}_{rqm}(u',u'')$ for tensors that in a locally Lorentzian reference frame have the form

$$J^{(1)}_{lmn}(u',u'')$$

$$=\frac{1}{u'^0u''^0}\int\frac{d^3k}{k^3}\int_{-\infty}^{\eta}d\eta'\int_{-\infty}^{\eta}d\eta''\int_{-\infty}^{\eta'}d\tau'\int_{-\infty}^{\tau'}d\tau''$$

$$\times(k_l^+e^{-ik(\eta'-\eta)}-k_l^-e^{ik(\eta'-\eta)})(k_m^+e^{ik(\eta''-\eta)}$$

$$-k_m^-e^{-ik(\eta''-\eta)})(k_n^-e^{ik(\tau''-\tau')}-k_n^+e^{-ik(\tau''-\tau')})$$

$$\times\exp\left[\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}'')(\eta''-\tau'')+\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}')(\tau'-\eta')\right],$$

$$J^{(2)}_{lmn}(u',u'')$$

$$=\frac{1}{u'^0u''^0}\int\frac{d^3k}{k^3}\int_{-\infty}^{\eta}d\eta'\int_{-\infty}^{\eta}d\eta''\int_{-\infty}^{\eta''}d\tau''\int_{-\infty}^{\tau''}d\tau'$$

$$\times(k_l^+e^{-ik(\eta'-\eta)}-k_l^-e^{ik(\eta'-\eta)})$$

$$\times(k_m^+e^{ik(\eta''-\eta)}-k_m^-e^{-ik(\eta''-\eta)})(k_n^-e^{ik(\tau''-\tau')}$$

$$-k_n^+e^{-ik(\tau''-\tau')})\exp\left[\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}'')(\eta''-\tau'')\right.$$

$$\left.+\frac{i}{c}(\mathbf{k}\cdot\mathbf{v}')(\tau'-\eta')\right].$$

After evaluating the integrals with respect to η', η'', τ', and τ'', we arrive at

$$J^{(1)}_{lmn}(u',u'')$$

$$=\frac{c^4}{u''^0u''^0}\fint\frac{d^3k}{k^3}\,\frac{1}{(\mathbf{k}\cdot\mathbf{v}''-\mathbf{k}\cdot\mathbf{v}')}$$

$$\times\left\{\frac{k_l^+k_m^+k_n^+}{(kc+\mathbf{k}\cdot\mathbf{v}'')^3}+\frac{k_l^+k_m^+k_n^-+k_l^+k_m^-k_n^++k_l^-k_m^+k_n^+}{(kc+\mathbf{k}\cdot\mathbf{v}'')^2(kc-\mathbf{k}\cdot\mathbf{v}'')}\right.$$

$$\left.+\frac{k_l^+k_m^-k_n^-+k_l^-k_m^+k_n^-+k_l^-k_m^-k_n^+}{(kc+\mathbf{k}\cdot\mathbf{v}'')(kc-\mathbf{k}\cdot\mathbf{v}'')^2}+\frac{k_l^-k_m^-k_n^-}{(kc-\mathbf{k}\cdot\mathbf{v}'')^3}\right\}. \tag{37}$$

The expression for $J^{(2)}_{lmn}(u',u'')$ can be derived from the expression for $J^{(1)}_{lmn}(u',u'')$ by replacing $\mathbf{v}'$ with $\mathbf{v}''$ and vice versa. The slash indicates that the integral is calculated as a principal value.

Just as we did in the previous case, we specify (37) in the center-of-mass reference frame, where $\mathbf{v}'=\mathbf{v}$, $\mathbf{v}''=-\mathbf{v}$, and $u'^0=u''^0=1/\sqrt{1-v^2/c^2}=u^0$. In this reference frame the components of $J^{(1)}_{lmn}(u',u'')$ have the following form (the spatial indices of the three-dimensional velocity vector v^α are lowered using the three-dimensional Kronecker symbol $\delta_{\alpha\beta}$)

$$J_{000}=-\alpha(v)\,\frac{v^2}{c^2},\quad J_{00\alpha}=-\alpha(v)\,\frac{v_\alpha}{c},$$

$$J_{0\alpha\beta}=-\alpha(v)\,\delta_{\alpha\beta}+\beta(v)\left(\delta_{\alpha\beta}-\frac{v_\alpha v_\beta}{v^2}\right), \tag{38}$$

$$J_{\alpha\beta\gamma}=-\frac{c^2}{v^2}\,\alpha(v)\left[\delta_{\alpha\beta}\frac{v_\gamma}{c}+\delta_{\alpha\gamma}\frac{v_\beta}{c}\right.$$

$$\left.+\delta_{\beta\gamma}\,\frac{v_\alpha}{c}-2\,\frac{v_\alpha v_\beta v_\gamma}{cv^2}\right]+\frac{c^2}{v^2}\beta(v)$$

$$\times\left[\left(\delta_{\alpha\beta}-\frac{v_\alpha v_\beta}{v^2}\right)\frac{v_\gamma}{c}+\left(\delta_{\alpha\gamma}-\frac{v_\alpha v_\gamma}{v^2}\right)\frac{v_\beta}{c}\right.$$

$$\left.+\left(\delta_{\beta\gamma}-\frac{v_\beta v_\gamma}{v^2}\right)\frac{v_\alpha}{c}\right]. \tag{39}$$

The functions α and β in (38) and (39) depend only on the velocity $v=\sqrt{v_1^2+v_2^2+v_3^2}$ and have the explicit form

$$\alpha=\frac{\pi c^3}{u_0^2v^3k_{\min}}\left[\frac{2(v/c)(1+v^2/c^2)}{(1-v^2/c^2)^2}+\ln\frac{1-v/c}{1+v/c}\right], \tag{40}$$

$$\beta=\frac{\pi c^3}{2u_0^2v^3k_{\min}}\left[\frac{2(v/c)(3-2(v^2/c^2)+3(v^4/c^4))}{(1-v^2/c^2)^2}\right.$$

$$\left.+3\left(1+\frac{v^2}{c^2}\right)\ln\frac{1-v/c}{1+v/c}\right]. \tag{41}$$

Here we have introduced the following notation for the integral:

$$\frac{1}{k_{\min}}=\int_{k_{\min}}^{\infty}\frac{dk}{k^2}.$$

Reasoning in the same way as we did earlier, we set the lower integration limit to $k_{\min}=1/r_{\max}$.

A covariant generalization of these results, which were obtained in the locally Lorentzian center-of-mass reference frame, to arbitrary reference frames has the form

$$J^{(1)}_{ijk}(u',u'')=J^{(2)}_{ijk}(u'',u')=J_{ijk}(u',u''), \tag{42}$$

with

$$J_{ijk}(u',u'')$$

$$=A[(g_{ij}u'_k+g_{ik}u'_j+g_{jk}u'_i)-z(g_{ij}u''_k+g_{ik}u''_j+g_{jk}u''_i)$$

$$-(u'_iu''_ju''_k+u''_iu'_ju''_k+u''_iu''_ju'_k)+3zu''_iu''_ju''_k]$$

$$+C[u'_iu'_ju'_k-z(u'_iu'_ju''_k+u'_iu''_ju'_k+u''_iu'_ju'_k)$$

$$+z^2(u'_iu''_ju''_k+u''_iu'_ju''_k+u''_iu''_ju'_k)-z^3u''_iu''_ju''_k], \tag{43}$$

where $z=(u'u'')=(u'^iu''_i)$,

$$A=-\frac{2\pi\sqrt{2}}{k_{\min}}\left[\frac{z-2}{(z-1)^2(z+1)^{1/2}}\right.$$

$$+\frac{2z-1}{(z+1)(z-1)^{5/2}}\ln(z+\sqrt{z^2-1})\Bigg], \tag{44}$$

$$C=-\frac{2\pi\sqrt{2}}{k_{\min}}\Bigg[\frac{z-6}{(z-1)^3(z+1)^{3/2}}$$
$$+\frac{6z-1}{(z+1)^2(z-1)^{7/2}}\ln(z+\sqrt{z^2-1})\Bigg]. \tag{45}$$

The tensor $J_{ijk}(u',u'')$ satisfies the identity

$$J_{ijk}(u',u'')u''^k=0. \tag{46}$$

By virtue of the properties (42) and (46), we can write the expression (36) for μ_{ij} in the following covariant form:

$$\mu_{ij}=-\sum_{bc}\frac{\chi^3 m_b^2 m_c^2 n_b n_c c^7}{8(2\pi)^3}$$
$$\times\int\frac{d^4p'}{\sqrt{(-g)}}\int\frac{d^4p''}{\sqrt{(-g)}}\Bigg\{\Bigg[\left(z^2+\frac{1}{2}\right)(u''_i u''_j+u'_i u'_j)$$
$$+\left(z^2-\frac{1}{2}\right)g_{ij}-2z(u'_i u''_j+u''_i u'_j)\Bigg]g^{qr}$$
$$-2\left(z^2-\frac{1}{2}\right)\delta_i^q\delta_j^r\Bigg\}f_c(x'')\frac{\partial}{\partial p'_f}\Bigg\{f_b(x')\Bigg[\left(z^2-\frac{1}{2}\right)\delta_f^m$$
$$+\left(z^2+\frac{1}{2}\right)u'_f u'^m-2zu''_f u'^m\Bigg]\Bigg\}J_{rqm}(u',u''). \tag{47}$$

where we used the identity

$$\frac{\partial}{\partial p'_f}\left[\left(z^2-\frac{1}{2}\right)\delta_f^m+\left(z^2+\frac{1}{2}\right)u'_f u'^m-2zu''_f u'^m\right]$$
$$=\left(5z^2-\frac{1}{2}\right)\frac{u'^m}{m_b c}.$$

Note that the tensor μ_{ij} is traceless:

$$g^{ij}\mu_{ij}=0. \tag{48}$$

From (34) and (48), the macroscopic Einstein equations for a system of self-gravitating particles can be written as

$$G_{ij}+\varphi^k_{ij;k}+\mu_{ij}=\chi T_{ij}, \tag{49}$$

where the semicolon stands for a covariant derivative in a space with the metric g_{ij}, G_{ij} is the Einstein tensor of this space, and T_{ij} is the energy-momentum tensor.

The tensors φ^k_{ij} and μ_{ij} are expressed via (33) and (47) in terms of one-particle distribution function f_b specified in the eight-dimensional phase space in which all four components of the four-dimensional momentum are independent. The transition to the seven-dimensional distribution function F_b is made according to the formula

$$n_b f_b(q^i,p_j)=F_b(q^i,p_\alpha)\delta(\sqrt{g^{lm}p_l p_m}-m_b c).$$

Here the functions F_b depend only on the spatial components of momentum (we use greek indices to denote spatial components).

By integrating (33) and (47) with respect to p'_0 and p''_0 we can reduce the tensors φ^k_{ij} and μ_{ij} to the form

$$\varphi^k_{ij}=-\sum_{bc}\frac{\chi^3 m_b^3 m_c^3 c^9}{8(2\pi)^3}$$
$$\times\int\frac{d^3p'}{p'^0\sqrt{(-g)}}\int\frac{d^3p''}{p''^0\sqrt{(-g)}}$$
$$\times\left[\frac{1}{2}g^{fk}u''_i u''_j+u'^k(u'u'')(\delta_j^f u''_i+\delta_i^f u''_j)\right]$$
$$\times\left((u'u'')^2-\frac{1}{2}\right)K_{f\alpha}(u',u'')$$
$$\times\left(F_c(x'')\frac{\partial F_b(x')}{\partial p'_\alpha}-F_b(x')\frac{\partial F_c(x'')}{\partial p''_\alpha}\right), \tag{50}$$

$$\mu_{ij}=-\sum_{bc}\frac{\chi^3 m_b^3 m_c^3 c^9}{8(2\pi)^3}$$
$$\times\int\frac{d^3p'}{p'^0\sqrt{(-g)}}\int\frac{d^3p''}{p''^0\sqrt{(-g)}}$$
$$\times\Bigg\{\Bigg[\left(z^2+\frac{1}{2}\right)(u''_i u''_j+u'_i u'_j)+\left(z^2-\frac{1}{2}\right)$$
$$\times g_{ij}-2z(u'_i u''_j+u''_i u'_j)\Bigg]g^{qr}-2\left(z^2-\frac{1}{2}\right)\delta_i^q\delta_j^r\Bigg\}$$
$$\times F_c(x'')\frac{\partial}{\partial p'_f}\Bigg\{F_b(x')\Bigg[\left(z^2-\frac{1}{2}\right)\delta_f^m+\left(z^2+\frac{1}{2}\right)$$
$$\times u'_f u'^m-2zu''_f u'^m\Bigg]\Bigg\}J_{rqm}(u',u''). \tag{51}$$

Here

$$\frac{d^3p'}{p'^0\sqrt{(-g)}}\quad\text{and}\quad\frac{d^3p''}{p''^0\sqrt{(-g)}}$$

are the invariant volume elements in the three-dimensional momentum space of particles of species b and c, respectively. The greek index α in (50) assumes only the values 1, 2, and 3 (the spatial index). The derivative with respect to p'_f in (51) should be calculated as if all four component of momentum are independent. The dependence of p'_0 on p'_α is taken into account only after differentiation with respect p'_f is completed.

The tensors φ^k_{ij} and μ_{ij} must obey the additional condition

$$g_{lj}(\varphi^k_{ij;k}+\mu_{ij})_{;l}=0, \tag{52}$$

since the divergences of G_{ij} and T_{ij} vanish.

Equation (52) imposes certain restrictions on the dependence of the parameter $k_{\min}$ on the coordinates and the relative velocity of particles (the latter can be expressed in terms of z).

The right-hand side of the macroscopic Einstein equations contains the macroscopic energy–momentum tensor, which can be expressed in terms of the one-particle distribution functions F_b as follows:

$$T_{ij}=\sum_b \int \frac{d^3p}{p^0\sqrt{(-g)}}\,p_i p_j F_b(p). \tag{53}$$

The system of equations (49)–(53) must be augmented by the kinetic equation for F_b, which can be obtained from (8) by integrating with respect to p_0 and has the form of Eq. (53) of Ref. 1

3. POSSIBLE APPLICATIONS OF THE THEORY

The equations of the gravitational field for continuum media obtained here differ from the classical Einstein equations by the presence of two additional terms, $\varphi^k_{ij;k}$ and μ_{ij}, on the left-hand side.

These terms are proportional to the cube of the Einstein constant and to the square of the particle number density. Hence these terms can play an important role only in continuous media of very high density. Such densities are possible in the early stages of the evolution of the universe and inside objects close to gravitational collapse. Therefore, it is natural to look for applications of the derived equations primarily in the theory of early stages of the evolution of the universe and in gravitational collapse theory.

Let us apply the derived macroscopic equations to a medium that is in a state of local thermodynamic equilibrium, where each particle species is described by a general relativistic distribution function of the type[6]

$$F_a(q^i,p_\alpha)=A_a\exp\left[-\frac{c(v_ip^i)}{k_BT}\right]. \tag{54}$$

Here v_i is the macroscopic 4-velocity of the medium, k_B is the Boltzmann constant, T is the temperature, and A_a is a normalization constant.

In this case the tensor φ^k_{ij} is identically equal to zero and the tensor μ_{ij} has the form

$$\mu_{ij}=-\chi\epsilon\left(\tfrac{4}{3}v_iv_j-\tfrac{1}{3}g_{ij}\right). \tag{55}$$

If we move μ_{ij} from the left-hand side of the macroscopic equations (49) to the right-hand side, the equations become the ordinary Einstein equations with an additional energy–momentum tensor that resembles the energy–momentum tensor of an ideal liquid with an equation of state $P=\epsilon/3$ but with a negative ‘‘energy density.’’

In the nonrelativistic limit $mc^2\gg k_BT$ the absolute value of this energy density is (we drop all intermediate calculations)

$$\sum_{ab}\frac{15k^2r_{\max}}{2k_BT}m_a^2m_b^2N_aN_b. \tag{56}$$

Here N_a and N_b are the particle number densities of species a and b, respectively. In the ultrarelativistic limit $mc^2\ll k_BT$ the expression for this negative energy density differs from (56). Here we will not write the corresponding formula. Note that at $mc^2=k_BT$ the result coincides in order of magnitude with (56).

As noted in the Introduction, due to the smallness of the gravitational constant, the contribution of a negative energy density to the total density of matter manifests itself only at high densities (in the early stages of the evolution of the universe) or in macrosystems consisting of extremely massive objects (clusters of galaxies).

It must also be noted that the additional terms on the left-hand side of the macroscopic equations for the gravitational field in continuous media have been obtained only with allowance for a system of self-gravitating particles of the medium. In media where other types of interaction play an important role these interactions must also be taken into account. For instance, electromagnetic interactions play the leading role in plasmas. Hence in deriving the equations for the gravitational field in a relativistic plasma (in particular, a plasma at the radiation-dominated stage of the evolution of the universe), we must primarily account for the electromagnetic interactions.

This work was sponsored by the Russian Fund for Fundamental Research (Grant No. 95-02-05734-a).

---

*)E-mail: Alexei.Zakharov@ksu.ru

[1] A. V. Zakharov, Zh. Éksp. Teor. Fiz. **110**, 3 (1996) [JETP **83**, 1 (1996)].
[2] Yu. L. Klimontovich, Zh. Éksp. Teor. Fiz. **37**, 735 (1959) [Sov. Phys. JETP **10**, 535 (1960)].
[3] A. V. Zakharov, Zh. Eksp. Teor. Fiz. **96**, 769 (1989) [Sov. Phys. JETP **72**, 437 (1989)].
[4] K. S. Bisnovatyĭ-Kogan and I. G. Shukhman, Zh. Éksp. Teor. Fiz. **82**, 3 (1982) [Sov. Phys. JETP **55**, 1 (1982)].
[5] A. V. Zakharov, Astron. Zh. **66**, 1208 (1989) [Sov. Astron. **33**, 624 (1989)].
[6] N. A. Chernikov, Nauchn. Dokl. Vyssh. Shkoly, Fiz. Mat., No. 1, 168 (1959).

Translated by Eugene Yankovsky

Bose–Einstein condensation in a static scalar field from the standpoint of the Goursat problem

A. A. Borgardt and D. Ya. Karpenko

Donetsk Physicotechnical Institute, Ukrainian National Academy of Sciences, 340114 Donetsk, Ukraine

(Submitted 31 December 1996)
Zh. Éksp. Teor. Fiz. **112**, 1167–1175 (October 1997)

The Goursat problem, developed by the present authors in previous papers [Ukr. Fiz. Zh. (Russ. Ed.) **27**, 1602 (1982); Differentsial'nye Uravneniya **20**, 302 (1984); J. Math. Phys. **33**, 233 (1996)], is used to study the energy spectrum of a scalar relativistic particle in a static axisymmetric external scalar field of an attractive nature. This is obviously a model. It is shown that the problem formulated in this way has no unstable solutions, i.e., solutions increasing with time, in contrast to the Cauchy problem, where such solutions appear when the square of the particle frequency (energy) vanishes (in other words, in a Bose–Einstein condensation) © *1997 American Institute of Physics.* [S1063-7761(97)00210-2]

1. INTRODUCTION

The Goursat problem can only be applied to hyperbolic, and hence relativistic, equations. In such a formulation one of the spatial coordinates (for the sake of definiteness we assume it to be z, but the procedure is relativistically covariant) is singled out and the initial value of the wave function is specified along the characteristics $\xi = ct - z = 0$ and $\eta = ct + z = 0$ (Ref. 4). A solution is then sought in the time-like region, the two-dimensional Minkowski subspace $\mathscr{M}^2_{(+)}$, where $\xi\eta = c^2t^2 - z^2 \geqslant 0$, or in the space-like region $\mathscr{M}^2_{(-)}$, where $\xi\eta = c^2t^2 - z^2 \leqslant 0$. What is important is that this is the only formulation of the problem that distributes the solutions of the wave equation among the regions of two-dimensional (four-dimensional) Minkowski space, which precludes violation of the uniqueness theorem, e.g., the emergence of two Green's functions in electrodynamics.[5]

The Goursat problem, which is one-dimensional in its classical formulation, is generalized to the three-dimensional wave equation in Refs. 1–3 and 6.

Here we apply this method to study the behavior of a relativistic particle in a static external scalar field, which is added to the square of the mass μ according to the rule

$$\mu^2 \to \mu^2 + \frac{\mu V}{c^2}, \tag{1}$$

where V is the external scalar field. Such fields are examined by Schiff *et al.*,[7] Bethe,[8] Kokkedee,[9] Migdal,[10] and many others. Equation (1) is probably as fundamental in nonelectromagnetic interactions as the minimality principle $\mathbf{p} \to \mathbf{p} + (e/c)\mathbf{A}$ is in electromagnetic interactions.

If the external scalar field V is static, spherically symmetric $(V = V(|\mathbf{r}|))$, and attractive, the Klein–Fock–Gordon (KFG) equation may, if formulated in terms of the Cauchy problem, have solutions that increase in time (unstable solutions). Below we briefly summarize results obtained by such an approach (a detailed description can be found in Ref. 10, Chap. 11, p. 103). The KFG equation for a static field $V(r)$ has the form

$$\left(\frac{\omega^2}{c^2} + \Delta - \left(\frac{\mu c}{\hbar}\right)^2 - \frac{\mu V(r)}{\hbar^2}\right)\phi(\mathbf{r}) = 0, \tag{2}$$

where the total wave function is $\psi(\mathbf{r},t) = \exp(-i\omega t)\phi(\mathbf{r})$.

Multiplying Eq. (2) by ϕ^* and integrating over three-dimensional space, we obtain

$$\frac{\omega^2}{c^2} = \frac{\overline{p^2}}{\hbar^2} + \left(\frac{\mu c}{\hbar}\right)^2 + \frac{\mu\overline{V}}{\hbar^2}, \tag{3}$$

where

$$\overline{p^2} = \int (d\mathbf{r})|\nabla\phi|^2,$$

and $\overline{V}$ is the mean value of the potential. Suppose that $V(r)$ is a well of depth $-V_0$. Then Eq. (3) becomes

$$\frac{\omega^2}{c^2} = \frac{\overline{p^2}}{\hbar^2} + \frac{\mu^2c^2}{\hbar^2} - \frac{\mu V_0}{\hbar^2}. \tag{4}$$

We see that for certain values of the well parameters, $\omega^2 = 0$ (in Ref. 10 this phenomenon is called Bose condensation), and further deepening of the well leads to negative values of ω^2. This means that there are solutions that increase in time (are unstable):

$$\psi(\mathbf{r},t) = \exp(|\omega|t)\phi(\mathbf{r}).$$

If we study what is known as helical motion (the particle moves freely along the z axis, while in the plane perpendicular to this axis there is an axisymmetric field $V = V(x_T)$ acting on the particle, with $x_T = (x_1^2 + x_2^2)$), the solution of the KFG equation formulated in terms of the Cauchy problem is

$$\psi(\mathbf{r},t) = \exp\left(-i\omega t + \frac{ip_z z}{\hbar}\right)\phi(x_1,x_2). \tag{5}$$

The same line of reasoning can be used to show that there are certain values of the parameters of a two-dimensional well (critical parameters) at which the square of the frequency, ω^2, vanishes (Bose condensation sets in); then ω^2 becomes negative and unstable solutions emerge.

We now show how helical motion shows up in our problem. In contrast to the solution (5), in the Goursat problem the wave function of the KFG equation

$$\left(\frac{\partial^2}{\partial z^2}-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}+\Delta_T-\left(\frac{\mu c}{\hbar}\right)^2-\frac{\mu V}{\hbar^2}\right)\psi(\mathbf{x}_T,z,t)=0, \tag{6}$$

with $\Delta_T=\partial^2/\partial x_1^2+\partial^2/\partial x_2^2$, is an eigenfunction of the one-dimensional wave operator

$$\hat{L}=\frac{\partial^2}{\partial z^2}-\frac{\partial^2}{c^2\partial t^2}=-4\frac{\partial^2}{\partial\xi\partial\eta},$$

so that

$$\hat{L}\psi=Q^2\psi,\quad \psi|_{\xi=0,\eta=0}=\phi(\mathbf{x}_T). \tag{7}$$

The fundamental solution (the Riemann function) of the one-dimensional wave equation

$$\left(4\frac{\partial^2}{\partial\xi\partial\eta}+Q^2\right)\psi(\xi,\eta)=0$$

is well-known:

$$\psi(\xi,\eta)=J_0(\sqrt{\xi\eta Q^2})=J_0(\sqrt{(c^2t^2-z^2)Q^2}), \tag{8}$$

where $J_0(\tau)$ is the zeroth-order Bessel function of the first kind, and $\xi\eta Q^2\geqslant 0$.

If we consider Q^2 an elliptic operator that is independent of the characteristic variables (ξ,η) or (z,ct), i.e.,

$$Q^2=\left(\frac{\mu c}{\hbar}\right)^2-\Delta_T+\frac{\mu V(x_T)}{\hbar^2},$$

we can write the KFG equation in operator form:[3]

$$\psi(\mathbf{x}_T,z,t)=J_0\left(\sqrt{(c^2t^2-z^2)\left(\frac{\mu^2c^2}{\hbar^2}-\Delta_T+\frac{\mu V}{\hbar^2}\right)}\right)|0\rangle, \tag{9}$$

where $|0\rangle=\psi|_{ct-z=0,ct+z=0}=\phi(x_T)$ is the initial value of the wave function along the characteristics $\xi=ct-z=0$ and $\eta=ct+z=0$.

In this paper we examine the axial analog of Kepler's problem[11] and a two-dimensional oscillator potential. Using such potentials, which allow for an analytic solution, as examples, we can study the features of the energy spectrum that emerge in the Goursat problem.

2. AXIAL ANALOG OF KEPLER'S PROBLEM

Suppose that a relativistic particle moves in an axisymmetric field $V(x_T)=-g^2/x_T$, where the constant g is a free parameter that determines the interaction of the particle and the external scalar field. The parameter Q^2 in this case has the form

$$Q^2=\left(\frac{\mu c}{\hbar}\right)^2-\Delta_T-\frac{\mu g^2}{x_T\hbar^2},$$

and the eigenfunctions of transverse motion satisfy the equation

$$\left(\Delta_T+\frac{\beta k_0}{x_T}\right)\phi_{n_r,m}(x_T,\varphi)-\left(\frac{\beta k_0}{N}\right)^2\phi_{n_r,m}(x_T,\varphi), \tag{10}$$

where $k_0=\mu c/\hbar$, $\beta=g^2/\hbar c$ is the coupling constant, n_r is the radial quantum number, m is the azimuthal quantum number, n is the principal quantum number, $N=2n+1$, $n=n_r+|m|$, $x_T=\sqrt{x_1^2+x_2^2}$, and φ is the polar angle.

Plugging (10) into (9), we arrive at a solution of the KFG equation in the form

$$\psi(x_T,z,t)=J_0\left(\sqrt{(c^2t^2-z^2)k_0^2\left(1-\frac{\beta^2}{N^2}\right)}\right)\phi_{n_r,m}(x_T,\varphi). \tag{11}$$

We write the coupling constant $\beta=g^2/\hbar c$ as $\beta=2n_C+1$, where we call n_C the condensate quantum number. Then the solution (11) assumes a form that is convenient for stability studies:

$$\psi(x_T,z,t)=J_0\left(\sqrt{(c^2t^2-z^2)\left(\frac{\mu_n c}{\hbar}\right)^2(n-n_C)}\right)\times\phi_{n_r,m}(x_T,\varphi), \tag{12}$$

which contains the positive constant

$$\mu_n^2=\mu^2\frac{n+n_C+1}{(n+1/2)^2}.$$

If $g^2/\hbar c\geqslant 1$ (strong coupling), a finite number of lower levels that satisfy the inequality $n-n_C<0$ emerge. To avoid instability, we must go to subspace $\mathscr{M}^2_{(-)}$ (where $z^2-c^2t^2\geqslant 0$), since stable solutions ''migrate'' to that subspace.

Thus, in the case of strong coupling, the solution (12) consists of two parts,

$$\psi=\sum_{n=0}^{[n_C]}\psi^{(-)}+\sum_{n=[n_C]+1}^{\infty}\psi^{(+)}, \tag{13}$$

where $[n_C]$ is the integer part of the condensate number, and

$$\psi^{(-)}=J_0(\sqrt{(z^2-c^2t^2)(\mu_n c/\hbar)^2(n_C-n)})\phi_{n_r,m}(x_T,\varphi), \tag{14}$$

$$\psi^{(+)}=J_0(\sqrt{(c^2t^2-z^2)(\mu_n c/\hbar)^2(n-n_C)})\phi_{n_r,m}(x_T,\varphi), \tag{15}$$

$$\phi_{n_r,m}=C_{nm}e^{im\varphi-k_nx_T}F(-n_r,2|m|+1,2k_nx_T)x_T^{|m|}, \tag{16}$$

$$k_n=k_0\frac{n_C+1/2}{n+1/2}.$$

The time variable satisfies the inequality $|t|\leqslant|z|/c$ in the solution (14) and the inequality $|t|\geqslant|z|/c$ in the solution (15).

When the coupling is weak $(g^2/\hbar c<1)$, the condensate quantum number is negative and the subspace $\mathscr{M}^2_{(-)}$ contains no states.

We now establish how the solutions (14) and (15) in the Goursat problem are related to the plane waves of the Cauchy problem. To this end we employ the Fourier transform of the Bessel functions J_0. In the $\mathscr{M}^2_{(+)}$ space we have

$$J_0(\sqrt{(c^2t^2-z^2)(\mu_n c/\hbar)^2(n-n_C)})$$

$$=\frac{c}{\pi}\int_{-\infty}^{\infty}dp_z\,\frac{\exp(ip_zz/\hbar)\sin(\mathscr{E}_n(p_z)|t|/\hbar)}{\mathscr{E}_n(p_z)},\qquad(17)$$

where $\mathscr{E}_n(p_z)$ is the particle (antiparticle) energy spectrum:

$$\mathscr{E}_n(p_z)=\sqrt{c^2p_z^2+\mu_n^2c^4(n-n_C)},\quad n\geqslant[n_C]+1.\qquad(18)$$

In the $\mathscr{M}^2_{(-)}$ space we have

$$J_0(\sqrt{(z^2-c^2t^2)(\mu_nc/\hbar)^2(n_C-n)})$$
$$=\frac{1}{\pi c}\int_{-\infty}^{\infty}d\mathscr{E}\,\frac{\exp(-i\mathscr{E}t/\hbar)\sin(p_z(\mathscr{E})|z|/\hbar)}{p_z(\mathscr{E})},\qquad(19)$$

where

$$p_z(\mathscr{E})=p_{z,n}(\mathscr{E})=\frac{1}{c}\sqrt{\mathscr{E}^2+\mu_n^2c^4(n_C-n)},\quad 0\leqslant n\leqslant[n_C].\qquad(20)$$

Equations (18) and (20) clearly show that the eigenvalues of the one-dimensional wave operator $\hat{L}$ change sign under a transition from $\mathscr{M}^2_{(+)}$ to $\mathscr{M}^2_{(-)}$ and back.

Indeed, Eq. (18) implies that $\mathscr{E}^2-c^2p_z^2\geqslant 0$ and Eq. (20), that $\mathscr{E}^2-c^2p_z^2\leqslant 0$. In the Goursat problem, the boundary case $\mathscr{E}^2=c^2p_z^2$ corresponds to the emergence of a condensate.

The Fourier transform (17) suggests that the state of a particle (antiparticle) in $\mathscr{M}^2_{(+)}$ is a linear combination of plane waves that travel along the z axis and have an energy defined by (18). The group velocity of this packet is

$$v_z=\frac{d\mathscr{E}_n}{dp_z}=\frac{c^2p_z}{\mathscr{E}_n}=\frac{c^2p_z^2}{\sqrt{c^2p_z^2+\mu_n^2c^4(n-n_C)}}<c.\qquad(21)$$

In the $\mathscr{M}^2_{(-)}$ space the state is a linear combination of standing waves with group velocity

$$v_z=\frac{d\mathscr{E}}{dp_z}=\frac{c^2p_z}{\mathscr{E}_n}=\frac{c^2p_z^2}{\sqrt{c^2p_z^2-\mu_n^2c^4(n_C-n)}}>c.\qquad(22)$$

What is most interesting here is the limit $n_C\rightarrow n\mp 0$. In this case the external scalar field ''consumes'' the particle mass, in Migdal's terminology.[10] The massless particle moves along the z axis with velocity $v_z=c\mp 0$. Note that a strong attractive scalar field can compensate for the particle mass: a constant magnetic field or the field of a plane electromagnetic wave increases the particle mass. For instance, the energy levels of a relativistic particle in a constant magnetic field are

$$\mathscr{E}_n=\sqrt{c^2p_z^2+\mu_n^2c^4},\quad \mu_n^2=\mu^2\left(1+\frac{H}{H_S}(2n+1)\right),$$

where $H_S=|e|\hbar/\mu^2c^3$ is a constant Schrödinger field. These levels are always dependent on the mass μ.

A constant electric field can perhaps also compensate for particle mass. This question is discussed in terms of the Cauchy problem in Refs. 12–14 and in many other papers. In terms of the Goursat problem this question merits separate study.[15]

The minimum value of the condensate quantum number n_C for the wave function in $\mathscr{M}^2_{(+)}$ at which particle mass is compensated is observed at the coupling constant value $\beta=g^2/\hbar c\approx 3$, i.e., at $n_C=1-0$. In this case the $\mathscr{M}^2_{(-)}$ space contains only the ground state with $n=0$, whose wave function is

$$\psi_0^{(-)}=C_0\,\exp(-3k_0x_T)J_0(\sqrt{8k_0^2(z^2-c^2t^2)}).\qquad(23)$$

In substituting (23) into the KFG equation, one must take into account a property of the Bessel functions J_0, namely

$$\hat{L}J_0(\sqrt{8k_0^2(z^2-c^2t^2)})=-8k_0^2J_0(\sqrt{8k_0^2(z^2-c^2t^2)}).\qquad(24)$$

States with $n\geqslant 1$ belong to the subspace $\mathscr{M}^2_{(+)}$, and in the harmonic with $n=1$ the mass is compensated. This is a condensate state. The wave functions of the condensate state are threefold degenerate (the multiplicity of degeneracy is $2n+1$):

$$\psi_1^{(+)}=C_{1,0}\,\exp(-k_0x_T)(1-2k_0x_T)$$
$$+C_{1,\pm 1}\exp(\pm i\varphi-k_0x_T)x_T.\qquad(25)$$

When dealing with the Goursat problem, it is convenient to classify states not by energy but by the eigenvalues of the characteristic operator $\hat{L}$, i.e., by the difference

$$\mathscr{E}^2-c^2p_z^2=\mathscr{E}_T^2.$$

For $n_C=1-0$ we have the spectrum

$$\mathscr{E}_{T,0}^2=-8\mu^2c^4,\quad \mathscr{E}_{T,1}^2=0,$$
$$\mathscr{E}_{T,2}^2=(16/25)\mu^2c^4,\ldots,\quad \mathscr{E}_{T,\infty}^2=\mu^2c^4.$$

A spontaneous transition to an excited condensate state is possible if the momentum of translational motion is lost.

If $n_C=n+0$ $(n\geqslant 0)$, the mass compensation effect is observed in $\mathscr{M}^2_{(-)}$ at a smaller value of the condensate quantum number n_C, at $g^2/\hbar c\approx 1$, i.e., at $n_C=+0$. The wave function of the condensate ground state is

$$\psi_0^{(-)}=C_{0,0}\,\exp(-k_0x_T).\qquad(26)$$

The state with $n=1$ belongs to $\mathscr{M}^2_{(+)}$, and the wave function of this state is

$$\psi_1^{(+)}=\exp\left(-\frac{k_0x_T}{3}\right)\left(C_{1,0}\left(1-\frac{2}{3}\,k_0x_T\right)\right.$$
$$\left.+C_{1,\pm 1}\,x_T\,\exp(\pm i\varphi)\right)J_0(\sqrt{(8k_0^2/9)(c^2t^2-z^2)}).\qquad(27)$$

When the solution (27) is inserted into the KFG equation, we again take into account the properties of the Bessel functions (cf. (24)):

$$\hat{L}J_0(\sqrt{(c^2t^2-z^2)\tfrac{8}{9}k_0^2})=\tfrac{8}{9}k_0^2J_0(\sqrt{(c^2t^2-z^2)\tfrac{8}{9}k_0^2}).\qquad(28)$$

The energy spectrum starts at zero:

$$\mathscr{E}_{T,0}^2=0,\quad \mathscr{E}_{T,1}^2=(8/9)\mu^2c^4,\ldots,\quad \mathscr{E}_{T,\infty}^2=\mu^2c^4.$$

The condensate state may be lost when the velocity along the z axis drops to

$$v_z=\frac{p_z c^2}{\sqrt{c^2 p_z^2+(8/9)\mu^2 c^4}}\approx c-\frac{4}{9}\frac{\mu^2 c^2}{p_z^2}<c.$$

Since the condensate and noncondensate states belong to different subspaces of $\mathscr{M}^2$, and hence of $\mathscr{M}^4$, one should expect that a condensate state (even an excited one) is stable. For instance, for a condensate state to set in at $n_C=1-0$, the particle must give up the momentum of translational motion

$$p_z=\frac{1}{c}\sqrt{\mathscr{E}^2+8\mu^2 c^4}\geqslant 2\sqrt{2}\mu c.$$

3. AXIAL OSCILLATOR POTENTIAL (TWO-DIMENSIONAL OSCILLATOR)

Let the static scalar field be a two-dimensional oscillator potential:

$$V(x_T)=-V_0+\frac{\mu\Omega^2 x_T^2}{4}, \tag{29}$$

where the depth V_0 and the oscillator frequency Ω are free parameters of the problem. The wave function of transverse motion, $\phi(\mathbf{x}_T)$, is an eigenfunction of the oscillator equation

$$\left(\Delta_T-\frac{\mu^2\Omega^2 x_T^2}{4\hbar^2}\right)\phi_{n_r,m}(x_T,\varphi)$$
$$=-\frac{\mu\Omega}{\hbar}(n+1)\Phi_{n_r,m}(x_T,\varphi), \tag{30}$$

where $n=2n_r+|m|$ is the principal quantum number.

According to (9), the solution of this equation is

$$\psi(x_T,z,t)=J_0\left(\sqrt{(c^2t^2-z^2)\left(\frac{\mu c}{\hbar}\right)^2\left(1+\frac{\hbar\Omega(n+1)-V_0}{\mu c^2}\right)}\right)$$
$$\times\phi_{n_r,m}(x_T,\varphi), \tag{31}$$

where

$$\phi_{n_r,m}=C_{n,m}\exp\left(im\varphi-\frac{\mu\Omega x_T^2}{4\hbar}\right)$$
$$\times F\left(-n_r,|m|+1,\frac{\mu\Omega x_T^2}{2\hbar}\right)x_T^{|m|}. \tag{32}$$

Again, it is convenient to introduce the condensate quantum number n_C, taking

$$V_0=\mu c^2+\hbar\Omega(n_C+1). \tag{33}$$

Then the solution (31) assumes a form convenient for stability studies:

$$\psi(x_T,z,t)=J_0\left(\sqrt{(c^2t^2-z^2)\left(\frac{\mu_{\mathrm{eff}}c}{\hbar}\right)^2(n-n_C)}\right)$$
$$\times\phi_{n_r,m}(x_T,\varphi), \tag{34}$$

where the effective mass $\mu_{\mathrm{eff}}^2=\mu\hbar\Omega/c^2$ depends on the frequency Ω of the external field.

For $n\geqslant[n_C]+1$ the energy spectrum in $\mathscr{M}_{(+)}^2$ is

$$\mathscr{E}_n(p_z)=\sqrt{c^2p_z^2+\mu_{\mathrm{eff}}^2c^4(n-n_C)},$$

while for $0\leqslant n\leqslant[n_C]$ the wave function belongs to the subspace $\mathscr{M}_{(-)}^2$, with momentum quantized along the z axis:

$$p_{z,n}=\frac{1}{c}\sqrt{\mathscr{E}^2+\mu_{\mathrm{eff}}^2c^4(n_C-n)}.$$

The subspace $\mathscr{M}_{(-)}^2$ contains no bound states if $V_0<\mu c^2+\hbar\Omega$.

Since the energy spectrum is uniformly spaced, the transition from $\mathscr{M}_{(+)}^2$ to $\mathscr{M}_{(-)}^2$ and back is related to acquisition or loss of momentum along the z axis ($n_C=n\pm 0$):

$$p_z=\frac{1}{c}\sqrt{\mathscr{E}_{n\mp 1}^2\pm\mu c^2\hbar\Omega}. \tag{35}$$

We can now formulate a somewhat general conclusion. For there to be a purely relativistic effect of particle mass compensation by an external attractive field, the coupling constant or, more precisely, the condensate quantum number n_C must take on values that are close to positive integers. There will then be a finite number of bound states in $\mathscr{M}_{(-)}^2$. It is these states, considered in terms of the Cauchy problem, that lead to unstable wave functions (i.e., increasing with time).

4. CONCLUSION

We have discussed the problem of condensate states in terms of the Goursat problem rather than in terms of the ordinary Cauchy problem. The drawback of the Cauchy problem is that the initial conditions are specified in three-dimensional space at $t=0$, i.e., on a hyperplane in the space of superluminal velocities, while a solution is sought inside the light cone.

We have studied two exactly solvable and simple models. The first is the axial analog of Kepler's problem with a variable coupling constant (one free parameter). The second, more complicated, problem, which involves two free parameters (the depth and width of the potential well), is the two-dimensional oscillator. We show that at a certain value of the coupling constant, a finite number of bound states appear in the $\mathscr{M}_{(-)}^2$ space, where the velocities are superluminal. These states correspond to the emergence of unstable states in the Cauchy problem when $\mathscr{E}^2<0$. This involves allowing for the boson–boson interaction and introducing a nonlinear equation, which can be solved only numerically.[10] The Goursat problem resolves this difficulty.

In the Goursat problem the energy spectrum in $\mathscr{M}_{(-)}^2$ (see (18) and (20)) does not allow for states with $\mathscr{E}^2<0$, although a state with $\mathscr{E}=0$ is possible. For this reason, a linear combination of plane waves (see (19)) is always stable. If we look at $\mathscr{M}_{(+)}^2$, the energy spectrum there starts at $[n_C]+1$; these states are stable in both approaches. A transition from $\mathscr{M}_{(-)}^2$ to $\mathscr{M}_{(+)}^2$ is possible if there is a massless condensate state, but this must involve a loss in momentum p_z sufficient for the transition between the hyperbolas belonging to different sectors in the $(\mathscr{E},p_z)$ space.

The migration of levels is related to an increase in the value of the coupling constant, and in this way it resembles

K-capture instability of electrons in atoms with $Z > 137$. Here, at $Z = 137$, the velocity of an electron in the lowest state becomes equal to the speed of light.

[1] A. A. Borgardt and D. Ya. Karpenko, Ukr. Fiz. Zh. **27**, 1602 (1982).
[2] A. A. Borgardt and D. Ya. Karpenko, Differentsial'nye Uravneniya **20**, 302 (1984).
[3] A. A. Borghardt and D. Ya. Karpenko, J. Math. Phys. **33**, 233 (1996).
[4] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. 2: R. Courant, *Partial Differential Equations*, Wiley-Interscience, New York (1962).
[5] J. D. Jackson, *Classical Electrodynamics*, 2nd ed., Wiley, New York (1975).
[6] P. Hillion, J. Math. Phys. **31**, 1939 (1990).
[7] L. Schiff, H. Snyder, and J. Weinberg, Phys. Rev. **57**, 315 (1940).
[8] H. A. Bethe and R. Jackiw, *Intermediate Quantum Mechanics*, Addison-Wesley, Reading, Mass. (1980).
[9] J. J. J. Kokkedee, *The Quark Model*, W. A. Benjamin, New York (1969).
[10] A. B. Migdal, *Fermions and Bosons in Strong Fields* [in Russian], Nauka. Moscow (1978).
[11] A. A. Borgardt and D. Ya. Karpenko, Ukr. Fiz. Zh. **30**, 1564 (1985).
[12] V. D. Mur and V. S. Popov, Teor. Mat. Fiz. **27**, 81 (1976).
[13] A. A. Grib, S. G. Mamaev, and V. M. Mostepanenko, *Quantum Effects in Intense External Fields* [in Russian], Atomizdat, Moscow (1980).
[14] V. P. Oleĭnik, Ukr. Fiz. Zh. **18**, 105 (1979).
[15] A. A. Borgardt and D. Ya. Karpenko, Ukr. Fiz. Zh. **27**, 1572 (1982).

Translated by Eugene Yankovsky

‘‘Inversionless’’ stimulated emission of radiation by nonequilibrium ensembles of classical oscillators

A. V. Gaponov-Grekhov and M. D. Tokman

Institute of Applied Physics, Russian Academy of Sciences, 603600 Nizhniĭ Novgorod, Russia

(Submitted 18 February 1997)
Zh. Éksp. Teor. Fiz. **112**, 1176–1196 (October 1997)

We find classical analogs of quantum systems capable of stimulated emission of radiation in the absence of inversion. We show that cyclotron parametric instability in low-frequency modulation of the distribution function of resonant particles can amplify a bichromatic high-frequency field when amplification of each spectral component separately is impossible. We point to similar modes for a Cherenkov resonance and a model system with lumped parameters. Finally, we suggest using this effect for converting microwave radiation to a higher frequency.

[S1063-7761(97)00310-7]

1. INTRODUCTION

The contemporary physics of radiative processes in plasmas and microwave electronics devices is based on a deep analogy between the emission mechanisms of ensembles of quantum and classical oscillators. ‘‘Maser’’ mechanisms of emission by aggregates of charged particles occupy center stage in the theory of instabilities in laboratory[1] and cosmic[2] plasmas, and in the physics of free-electron masers and lasers[1)] (see Refs. 3 and 4). Here the instability criteria of ‘‘classical’’ distribution functions of charged particles against coherent radiation various sorts of electromagnetic waves are a natural generalization of the population inversion condition in an elementary two-level quantum system, while the relaxation of unstable distributions to a ‘‘plateau’’ in momentum space (when spontaneous emission and nonradiative relaxation are unimportant) is a direct analog of the process in which the level populations in a resonant transition become equal as a result of the interaction of the quantum system and high-power radiation (see, e.g., Ref. 2).

The question of how universal the requirement is that there must be a population inversion is obviously one of the most important aspects of the theory of interaction of radiation and matter. It would seem that apart from instability of quasimonochromatic electromagnetic modes in an inverted system and ‘‘pure noise’’ emission by an equilibrium ensemble, there might in principle be ‘‘intermediate’’ possibilities. In other words, can a nonequilibrium but noninverted system relax to equilibrium[2)] as a result of the instability of a nonmonochromatic electromagnetic field[3)]? In recent years such regimes of emission have been found to exist in quantum systems (here and below in this regard we cite the review in Ref. 6). It is found, in particular, that when one of the levels in an elementary two-level system is split (Fig. 1), under certain conditions there can be parametric instability in the system, which leads to emission of bichromatic radiation in the absence of ordinary population inversion. Here, nevertheless, electrons in the upper energy level serve as an energy reservoir, as in the case of ordinary ‘‘maser’’ instability.

The present paper is devoted to a search for a classical analog of an unstable quantum ‘‘inversionless’’ system. The fundamental aspect of this problem is that it obviously broadens our ideas about the classes of aggregates of charged particles that are able to generate stimulated radiation. The interest from the application angle is related to a common problem of quantum and classical electronics: as lasing frequency increases, population inversion becomes more and more difficult.[4)]

The plan is as follows. In Sec. 2 we discuss, with the thoroughness required by the problem of building a classical analog, the features of ‘‘inversionless’’ emission by a three-level quantum system and give the simplest classical analog, a system consisting of three coupled oscillators with nonlinear friction (see Fig. 2). Having analyzed this model system, we conclude that the analog of the ‘‘inversionless’’ stimulated emission of radiation by the quantum system for a classical ensemble of electrons is a parametric instability at resonant frequencies common to two high-frequency modes and their beats. In Sec. 3 we develop a linear theory of cyclotron instability of this type, and in Sec. 4 we use a quasilinear theory to analyze the corresponding energy relations and the features of the relaxation of the distribution function when ‘‘inversionless’’ instability develops. In the Conclusion (Sec. 5) we discuss the results, possible applications, and future research directions in this field.

2. SIMPLE MODELS OF ‘‘INVERSIONLESS’’ STIMULATED EMISSION OF RADIATION

‘‘Inversionless’’ stimulated emission of radiation by various quantum system is discussed in a large number of papers on quantum electronics (see, e.g., the review in Ref. 6 and the literature cited therein). We discuss this problem in the form and with the thoroughness that in our opinion are necessary for a consistent systematic, persuasive transition to a classical model. In particular, we discuss in greater detail than in the cited papers the conditions for spatial synchronism in ‘‘inversionless’’ radiation by a distributed system.

The most graphic example of a quantum system capable of stimulated emission of radiation without ordinary inversion is what is known as the Lambda scheme.[6] The scheme represents a three-level system (see Fig. 1) with energy levels $\mathscr{E}_j$ $(j=1,2,3)$, two high-frequency transitions $(3\leftrightarrow 1$ and

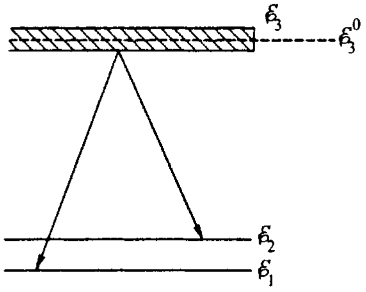

FIG. 1. Three-level Λ scheme.

$3\leftrightarrow 2$), and one low-frequency transition ($2\leftrightarrow 1$). Consider how this system interacts with the bichromatic field

$$\mathbf{E}=\mathbf{e}\ \mathrm{Re}[E_1\ \exp(-i\omega_1 t)+E_2\ \exp(-i\omega_2 t)], \tag{1}$$

where $\mathbf{e}$ is the unit polarization vector, and the frequencies $\omega_{1,2}$ are those of the high-frequency transitions, $\hbar^{-1}(\mathscr{E}_3-\mathscr{E}_{2,1})$, which means they must also satisfy the Raman synchronism condition

$$\omega_1-\omega_2=\omega_0=\frac{\mathscr{E}_2-\mathscr{E}_1}{\hbar}. \tag{2}$$

The transition probability per unit time (the transition rate) from level 3 to levels 1 and 2 is determined by the sum of the standard probabilities of the transitions $3\leftrightarrow 1$ and $3\leftrightarrow 2$ induced by resonant harmonics.[8,9] At the same time, for upward transitions, the standard transition probabilities in a system that is in a pure state are valid only if the initial condition corresponds to the definite energy $\mathscr{E}_1$ or $\mathscr{E}_2$.

The situation is different when the wave function obeys the ‘‘superpositional’’ initial condition

$$\psi_0=a_1\psi_1(q)\exp\left(-\frac{i\mathscr{E}_1 t}{\hbar}\right)+a_2\psi_2(q)\exp\left(-\frac{i\mathscr{E}_2 t}{\hbar}\right),$$

where $\psi_{1,2}$ are the energy-operator eigenfunctions corresponding to $\mathscr{E}_1$ and $\mathscr{E}_2$, q is the set of generalized coordinates of the quantum system, and the coefficients $a_{1,2}$ satisfy the normalization condition $|a_1|^2+|a_2|^2=1$. Then the transition probability calculated by the standard perturbation technique exhibits an ‘‘interference’’ pattern, since we sum the perturbations of the complex-valued wave function and not the partial transition probabilities.[10] Assuming, for the sake of definiteness, that level 3 is ‘‘smeared’’ (e.g., due to the finit lifetime of the upper level) over a small neighborhood of the energy $\mathscr{E}_3^0=\mathscr{E}_2+\hbar\omega_2=\mathscr{E}_1+\hbar\omega_1$, we get

$$W(\uparrow)=\frac{\pi}{\hbar^2}\int d\Delta\left\{\tau(\Delta)\,\delta(\Delta)\left[\sum_{p=1}^{2}|a_p d_{3p}E_p|^2+2\ \mathrm{Re}(a_1 a_2^* d_{31}d_{32}^* E_1 E_2^*)\right]\right\},$$

where d_{ij} is the matrix element of the projection of the electric-dipole moment in the direction specified by $\mathbf{e}$,

$$\Delta=\frac{\mathscr{E}_3-\mathscr{E}_1}{\hbar}-\omega_1=\frac{\mathscr{E}_3-\mathscr{E}_2}{\hbar}-\omega_2=\frac{\mathscr{E}_3-\mathscr{E}_3^0}{\hbar},$$

$\tau(\Delta)$ is the density of states that describes the smearing of the upper level (with dimensions of reciprocal frequency if the normalization condition is $\int\tau(\Delta)d\Delta=1$), and $\delta(\Delta)$ is a delta function. Obviously, if

$$a_{1,2}\neq 0,\quad \arg(a_1 a_2^*)=\pi-\arg(d_{31}d_{32}^* E_1 E_2^*), \tag{3}$$

the transition probability $W(\uparrow)$ is necessarily less than the sum of the ordinary probabilities of the $3\leftrightarrow 1$ and $3\leftrightarrow 2$ transitions with weights $|a_1|^2$ and $|a_2|^2$, respectively (see also Ref. 10). This illustrates that it is possible, at least in principle, to stimulate emission of radiation in a quantum oscillator in the absence of population inversion. If before the high-frequency field is turned on, the three-level atom is in a mixed state characterized by a density matrix ρ_{ij}, we can obtain the necessary condition (equivalent to (3)) for ‘‘inversionless’’ stimulated emission by employing a technique suggested, for example, in Ref. 9:

$$\rho_{21}\neq 0,\quad \arg(\rho_{21})=\pi-\omega_0 t+\arg(d_{31}d_{32}^* E_1 E_2^*). \tag{4}$$

Consider an ensemble of noninteracting three-level (for the sake of clarity) atoms of the above type that is distributed in space. One way to characterize this system is to introduce a set of density matrices ρ_{ij}^k, where the superscript k corresponds to the atom whose center is at the point with radius vector $\mathbf{r}=\mathbf{r}_k$. For wave fields ($E_{1,2}\rightarrow E_{1,2}\ \exp(i\mathbf{k}_{1,2}\cdot\mathbf{r})$) in the dipole approximation, the conditions (4) can be met by all atoms along the propagation path of the waves only if

$$\arg\ \rho_{21}^k=\varphi_0-\omega_0 t+(\mathbf{k}_1-\mathbf{k}_2)\cdot\mathbf{r}_k, \tag{5}$$

where φ_0 is a constant. The latter relationship shows that for ‘‘inversionless’’ amplification to occur in a distributed ensemble, a certain condition of spatial synchronism must be met, in addition to the temporal synchronism condition (2). It is convenient to formulate this new condition for the statistical matrix $\rho_{ij}(\mathbf{r},t)$, which depends on position $\mathbf{r}$ and time t. This matrix determines, among other things, the inhomogeneous dipole moment per unit volume $\mathbf{d}(\mathbf{r},t)$ (see also Ref. 9):

$$\mathbf{d}(\mathbf{r},t)d^3r=N(\mathbf{r})\mathbf{d}_{ji}\rho_{ij}(\mathbf{r},t)d^3r, \tag{6}$$

where

$$\rho_{ij}(\mathbf{r},t)=\frac{1}{N(\mathbf{r})}\sum_k \rho_{ij}^k\delta(\mathbf{r}-\mathbf{r}_k),$$

and $N(\mathbf{r})=\Sigma_k\delta(\mathbf{r}-\mathbf{r}_k)$ is the number density of the atoms.

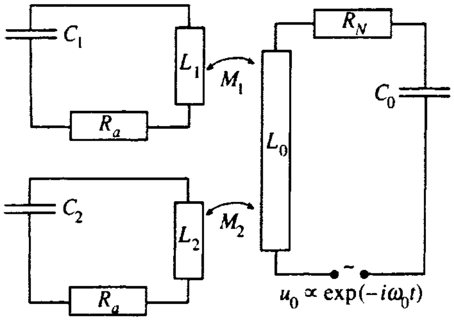

FIG. 2. Model system with lumped parameters.

If the distance between atoms is much less than the wavelengths $2\pi/k_{1,2}$, the definition (6) and the condition (5) imply that the matrix $\rho_{ij}(\mathbf{r},t)$ has a nonzero off-diagonal matrix element

$$\rho_{21}(\mathbf{r},t)=\sigma_{21}\exp(-i\omega_0 t+i\boldsymbol{\kappa}\cdot\mathbf{r})$$

synchronized with the bichromatic field by the parametric (Raman) resonance conditions (2):

$$\mathbf{k}_1-\mathbf{k}_2=\boldsymbol{\kappa}. \tag{7}$$

Such a state of electrons at the low-frequency transition is said to be coherent[5)] (see Refs. 6 and 11). Here the high-frequency dipole moment per unit volume of the medium acquires ''Raman'' components, which couple the wave harmonics $(\omega_1,\mathbf{k}_1)$ and $(\omega_2,\mathbf{k}_2)$:

$$\begin{aligned}(\mathbf{d}\cdot\mathbf{e})_{\text{comb}}={}&NE_2\exp[-i(\omega_1 t-\mathbf{k}_1\cdot\mathbf{r})]\\ &\times\frac{1}{\hbar}\int d\Delta\{\tau(\Delta)\xi(\Delta)d_{32}d_{31}^*\sigma_{21}\}\\ &+NE_1\exp[-i(\omega_2 t-\mathbf{k}_2\cdot\mathbf{r})]\\ &\times\frac{1}{\hbar}\int d\Delta\{\tau(\Delta)\xi(\Delta)d_{32}^*d_{31}\sigma_{21}^*\},\end{aligned}$$

where $\xi=\mathscr{P}/\Delta-i\pi\delta(\Delta)$, with $\mathscr{P}$ denoting a principal-value integral, and the atomic density is assumed to be uniform. In the absence of resonances, i.e., at $\tau(0)=0$, simultaneous amplification of all high-frequency waves is forbidden by the Manley–Rowe relations:[6)]

$$\sum_{1,2}\left(\frac{d}{dt}+2\gamma_{1,2}\right)Z_{1,2}\frac{|E_{1,2}|^2}{\omega_{1,2}}=0,$$

where

$$\frac{d}{dt}=\frac{\partial}{\partial t}+\mathbf{v}_{\text{gr }1,2}\cdot\boldsymbol{\nabla},$$

with $\mathbf{v}_{\text{gr }1,2}$ and $\gamma_{1,2}$ the group velocities and the standard linear damping constants of the modes, and

$$Z_{1,2}=\left[\frac{1}{\omega}\frac{\partial}{\partial\omega}(\omega^2\varepsilon)\right]_{\omega=\omega_{1,2}},$$

with ε the real-valued dielectric constant of the medium.

It turns out that the restriction imposed by the Manley–Rowe relations is lifted when we allow for radiative processes related to ''partial'' resonances for the high-frequency modes. To illustrate this we take the simple example in which the frequencies $\omega_{1,2}$ are the centers of the corresponding resonance lines (i.e., the integrals in the principal-value sense in the expression for the ''Raman'' dipole moment are zero). In this case the truncated equations for the two coupled wave harmonics are

$$\begin{aligned}&\frac{d}{dt}E_1+\gamma_1E_1=-G_1\sigma_{21}E_2,\\ &\frac{d}{dt}E_2+\gamma_2E_2=-G_2^*\sigma_{21}^*E_1,\end{aligned} \tag{8}$$

where

$$G_{1,2}=\frac{4\pi^2\omega_{1,2}\tau(0)d_{32}d_{31}^*N}{\hbar Z_{1,2}}.$$

The system of equations (8) is unstable (which can easily be verified) when

$$\gamma_1\gamma_2<G_1G_2^*|\sigma_{21}|^2.$$

If linear dissipation of the wave is due entirely to resonant absorption in the 3↔1 and 3↔2 transitions, the linear damping factors can be expressed in terms of the corresponding population differences in levels[7)] 1, 2, and 3 (see, e.g., Refs. 6 and 9):

$$\gamma_{1,2}=\frac{4\pi^2\omega_{1,2}\tau(0)|d_{31,2}|^2N(n_{1,2}-n_3)}{\hbar Z_{1,2}},$$

where $n_{1,2,3}$ are the level populations, with $n_1+n_2+n_3=1$. For our system with a ''smeared'' upper level, such an expression for the linear damping factors corresponds to the simplest possible probability distribution of the population in the upper level: $n_{e3}(\Delta)=n_3\tau(\Delta)$. The instability condition then reduces to the following inequality[6,11]:

$$(n_1-n_3)(n_2-n_3)<|\sigma_{21}|^2. \tag{9}$$

It is this inequality that demonstrates the possibility of stimulated emission of radiation in a medium without population inversion in the ordinary sense of the word,[8)] i.e., when $n_1,n_2>n_3$.

The medium we have just considered is of course nonequilibrium, if only because the off-diagonal elements of the density matrix are nonzero.[8,9] It is important to note, however, that the source of energy in ''inversionless'' emission of radiation is not the low-frequency polarization wave corresponding to this perturbation of the density matrix, but the electrons in the upper energy level[9)] (see Ref. 6). In this sense were are indeed dealing with the maser generation mechanism.

The simplest classical analog of the quantum lambda scheme is the following oscillatory system with lumped parameters (Fig. 2). Suppose that we have two high-frequency radio circuits with natural frequencies $\omega_{1,2}=1/\sqrt{L_{1,2}C_{1,2}}$ and negative resistance $R_a<0$. Individually, both circuits are unstable, but they are inductively coupled to a low-frequency circuit with $\omega_0=1/\sqrt{L_0C_0}$ and nonlinear resistance $R_N=R_0(1+\beta I_0)$, where I_0 is the current in the low-frequency circuit. Suppose that the system is stable in the linear approximation $(\beta=0)$. If to simplify matters we assume that coupling and dissipation are weak, or

$$\frac{M_{1,2}}{L_{1,2}},\ \frac{M_{1,2}}{L_0},\ -\frac{R_a}{2L_{1,2}\omega_{1,2}},\ \frac{R_0}{2L_0\omega_0}\ll 1$$

(here $M_{1,2}$ are the respective mutual inductance coefficients), the ''linear'' stability conditions are

$$\delta_{1,2}=\frac{M_{1,2}^2R_0}{2L_{1,2}L_0^2}+\frac{R_a}{2L_{1,2}}>0.$$

If, however, low-frequency oscillations of the current in the ''dissipative'' circuit ($I_0\approx\mathrm{Re}[I_L\exp(-i\omega_0t)]$) modulate the resistance,

$$R_N \approx R_0\{1+\beta\ \mathrm{Re}[I_L\ \exp(-i\omega_0 t)]\},$$

instability sets in under Raman resonance conditions:

$$\tilde{\omega}_1-\tilde{\omega}_2=\omega_0,$$

where

$$\tilde{\omega}_{1,2}=\omega_{1,2}\left(1-\frac{M_{1,2}^2}{2L_{1,2}L_0}\right)$$

are the frequencies of oscillation in the high-frequency circuits, slightly shifted due to inductive coupling. Here the truncated equations for the complex-valued amplitudes of the currents in the high-frequency circuits, $I_{1,2}$, correspond exactly to the system of equations (8) if we change the notation, i.e., if

$$E_{1,2}\rightarrow I_{1,2}, \quad \gamma_{1,2}\rightarrow\delta_{1,2}, \quad G_{1,2}\rightarrow\frac{1}{L_{1,2}},$$

$$\sigma_{21}\rightarrow\frac{M_2M_1R_0}{4L_0^2}\beta I_L,$$

and the instability condition becomes[10)]

$$\left(\frac{M_1^2R_0}{2L_1L_0^2}+\frac{R_a}{2L_1}\right)\left(\frac{M_2^2R_0}{2L_2L_0^2}+\frac{R_a}{2L_2}\right)<\frac{M_1^2M_2^2R_0^2\beta^2I_L^2}{16L_0^4L_1L_2}.$$

Clearly, in this instructive example, the high-frequency circuits model the $3\leftrightarrow1$ and $3\leftrightarrow2$ transitions, the negative resistance R_a models the nonzero population of the upper active level, and the low-frequency circuit models the $2\leftrightarrow1$ transition. The resistance in the low-frequency circuit corresponds to the population of the lower active levels of the quantum system. Finally, modulation of the resistance in the low-frequency circuit due to the oscillations generated in the circuit models the effect of excitation of a coherent state in the $2\leftrightarrow1$ transition on the absorptivity of the three-level lambda system.

The physical mechanism of instability in the proposed system with lumped parameters turns out to be extremely interesting. It can be shown that an unstable solution corresponds to an ''optimal'' relationship between the phase of beats of the high-frequency currents induced in the low-frequency circuit and the nonlinearly modulated resistance[11)]: the nodes of the beasts correspond to maximum resistance, and the crests to minimum resistance. Clearly, when the low-frequency oscillations are in antiphase with the current beats, absorption of the high-frequency oscillations decreases on the average, with the result that negative resistance can lead to a buildup of high-frequency oscillations, even if the system is stable when the low-frequency circuit has positive constant resistance.

The physical mechanism we have just discussed suggests an analog of ''inversionless'' emission of radiation for an ensemble of classical electron oscillators. From a macroscopic point of view, the system of equations (8) corresponds to the parametric interaction of modes in a medium with modulated conductivity, where instability emerges because of the aforementioned synchronization of the beats in the high-frequency field with the low-frequency conductivity oscillations. Such conductivity modulation in an ensemble of classical electrons may be due to low-frequency modulation of the distribution function of resonant particles under conditions of cyclotron or Cherenkov synchronism. An interesting case is that in which the distribution function is ''inversionless'' when averaging over generalized coordinates in the phase space and/or time. Just like an ensemble of quantum oscillators, the classical system under investigation is capable of amplifying monochromatic radiation, but under certain conditions it can impart energy to a pair of waves with an ''optimal'' phase of beats with respect to the phase of the low-frequency modulation.

Probably the main difference between ''inversionless'' generation in quantum and classical systems can be formulated as follows. The motion of a classical particle corresponds to a large number of transitions between elementary quantum states. Since a classical system is in principle ''multilevel,'' a pure lambda scheme cannot be realized in it. Instead, this scheme is combined with all conceivable variants and modifications (see Ref. 6). In particular, since low-frequency modulation inevitably affects not only the absorbing fraction of electrons but also the emitting particles, we must mention the so-called inverse lambda scheme, in which the upper active level is split rather than the lower one, i.e., in contrast to the ordinary lambda scheme.

The constant of motion of system (8),

$$\dot{W}+Q=0,$$

where

$$W=\frac{|E_1|^2}{2G_1}-\frac{|E_2|^2}{2G_2}, \quad Q=\gamma_1\frac{|E_1|^2}{2G_1}-\gamma_2\frac{|E_2|^2}{2G_2},$$

makes it possible to formally interpret the parametric instability of the system (8) as the interaction of two high-frequency modes with different signs of the energy.[11] We believe that such an interpretation is not very accurate, however, the point being that in the absence of parametric coupling, the electromagnetic waves involved are incapable of being amplified when additional linear absorption is introduced, i.e., not one of these waves possesses the necessary property of a mode with negative energy[12)] (see Refs. 1 and 13). In this regard we must mention the paper of Kiyashko *et al.*,[14] who pointed out the possibility of explosive parametric instability of waves with precisely positive energy in a medium with nonlinear conductivity. In the approximation of a given low-frequency pump, the system examined in Ref. 14 corresponds to the system (8) with $\gamma_{1,2}=0$.

Thus, for a classical analog of ''inversionless'' emission of radiation by an ensemble of quantum oscillators, we take the parametric instability of electromagnetic waves with periodic modulation of the distribution function of ''common'' resonant particles. Note that Sherman *et al.*[15] and Nikonov *et al.*[16] suggest a modified version of the optical klystron[17] as an analog of an ''inversionless'' quantum device. They propose using a two-section free-electron laser (or a Cherenkov device) and assume that in the klystron gap the electron paths are separated spatially in such a way that the time of flight through the gap is different for the ''emitting'' and ''absorbing'' fractions of the electron velocity distribution function. Here, for the ''absorbing'' electrons, the phase of the interaction with the high-frequency field changes during the time of flight by $\pi(2n+1)$, and for the ''emitting'' electrons it changes by, $2\pi n$. Thus the ''absorbing'' fraction of the distribution function should contribute nothing to the resonant interaction of the electron flux with the high-frequency field. Without discussing the implementation of such a scheme in practice, we note that the requirements on the velocity spread can be loosened by tightening the requirements on the spatial localization of the beam (otherwise it is difficult to spatially separate the paths of electrons with a small velocity difference). In any event, the above scheme has no place for emission of the parametrically coupled different-frequency modes typical of quantum systems.

3. CYCLOTRON PARAMETRIC INSTABILITY INVOLVING RESONANT PARTICLES (LINEAR THEORY)

Cyclotron radiation is a convenient classical analog of ''inversionless'' generation in an ensemble of quantum oscillators, because in the quantum limit a system of discrete Landau levels corresponds to the motion of particles in a magnetic field.[2,8] The obvious analog of the $3 \leftrightarrow 1$ and $3 \leftrightarrow 2$ transitions in this case is the resonance at the Nth and $(N-L)$th harmonics of the cyclotron frequency, while oscillations of the distribution function at the Lth harmonic of the cyclotron frequency correspond to the ''coherent'' state of the low-frequency $2 \leftrightarrow 1$ transition.

Consider two linearly polarized waves propagating perpendicular to a constant magnetic field $\mathbf{H}_0 = \mathbf{z}H_0$:

$$\mathbf{E} = \mathbf{y}_0 \ \mathrm{Re} \sum_{j=1}^{2} E_j \ \exp(ik_j x - i\omega_j t), \tag{10}$$

where $\omega_1 = \omega$, $\omega_2 = \omega - \Omega$, $k_1 = k$, $k_2 = k - \kappa$, and $\mathbf{z}_0$ and $\mathbf{y}_0$ are the corresponding unit vectors of the coordinate system (z,y,x). We assume that the frequency of the waves coincides with the harmonics of the relativistic cyclotron frequency of particles with energy $\mathscr{E}_0 = mc^2\gamma_0$ (here γ_0 is the Lorentz factor):

$$\omega_R = \frac{eH_0}{mc\gamma_0} = \frac{\omega}{N} = \frac{\omega - \Omega}{N-L} = \frac{\Omega}{L}. \tag{11}$$

If the perturbation of the energy of the resonant particles is weak, or $|\gamma - \gamma_0| \ll \gamma_0$, the motion of the particles in the vicinity of cyclotron resonances can be described by the equations for a nonlinear pendulum (see Refs. 7, 18, and 19). The motion induced by the resonant harmonics in such an approximation is described by the following system of truncated equations[18,20]:

$$\dot{w} = F(\theta,t), \quad \dot{\theta} = \omega_H(w),$$

$$\rho_\| = \mathrm{const}, \quad X = \mathrm{const}, \tag{12}$$

where

$$\omega_H = \omega_R\left(1 - \frac{w}{\gamma_0}\right),$$

$$F = \mathrm{Re} \sum_{j=1}^{2} g_j \alpha_j \ \exp(ik_j X + iN_j\theta - i\omega_j t). \tag{13}$$

In Eqs. (12) and (13), θ is the cyclotron rotation phase, $w = \gamma - \gamma_0$ is the perturbation of the relativistic gamma factor due to the perturbation of the transverse momentum $p_\perp$ (the longitudinal momentum $p_\| = \rho_\| mc$ does not change), $N_1 = N$, $N_2 = N - L$, and $g_j = (p_{\perp 0}/mc\gamma_0)J'_{N_j}(\chi_j)$, with $p_{\perp 0} = mc\sqrt{\gamma_0^2 - 1 - \rho_\|^2}$ and J'_{N_j} the derivative of the Bessel function with respect to the argument $\chi_j = k_j r_H$, where $r_H = p_{\perp 0}/m\gamma_0\omega_R$ is the gyroradius of the resonant particles. Finally, $\alpha_j = eE_j/mc$ are the normalized wave amplitudes, and X is the location of the center of the ''Larmor circle.''

Below we assume for simplicity that all electrons have the same longitudinal momentum $\rho_\|$ (the appropriate generalization is trivial). Let us examine the distribution function $f(\theta,w,X,t)$ specified by Liouville's equation in the space of truncated variables:[21]

$$\frac{\partial f}{\partial t} + \omega_H \frac{\partial f}{\partial \theta} + F\frac{\partial f}{\partial w} = 0. \tag{14}$$

For the wave amplitudes α_j we also use truncated equations. With a ''quasivacuum'' dispersion relation, we obtain

$$\dot{\alpha}_j = -\frac{2\pi e}{mc} I_j, \tag{15}$$

where the amplitudes I_j of the resonant harmonics of the current can be expressed in terms of the distribution function (see, e.g., Ref. 20):

$$I_j = ec\left\langle \int dw \ d\theta [f(\theta,w,X,t) \right.$$
$$\left. \times g_j \ \exp(-iN_j\theta - ik_j X + i\omega_j t)] \right\rangle. \tag{16}$$

Here the angle brackets indicate time averaging over the period $T = 2\pi/\Omega$, and averaging in the coordinate X over the length $l = 2\pi/\kappa$.

Suppose that initially (at time $t=0$) the distribution function is modulated in cyclotron rotation angle (the gyro-angle) θ and position X:

$$f_i(t=0) = f_0(w) + f_L(w)\cos(\varphi + L\theta + \kappa X). \tag{17}$$

In the absence of a high-frequency field acting on the system, the value of the distribution function f_i at subsequent times is determined by replacing θ with $\theta - \omega_H t$ in (17). Here the distribution function averaged over θ,

$$\frac{1}{2\pi}\int_0^{2\pi} f_i d\theta,$$

is ''inversionless'' at all times and at each point X if $\partial f_0/\partial w$ is negative in the energy range considered here. Since in the quantum limit the phase θ is not defined for a given energy $mc^2(\gamma_0 + w)$, we are indeed dealing with the classical analog of an ''inversionless'' state of the quantum system. The modulation of the distribution function in gyro-angle with a period equal to $2\pi/L$ is an analog of the coherent state of the low-frequency transition.

Now we examine times that are shorter that the ''ballistic'' relaxation time of the modulated component of the distribution function:[13)]

$$\omega_R t \ll \frac{\gamma_0}{\langle \Delta w \rangle L}, \tag{18}$$

where

$$\langle \Delta w \rangle \sim \left| \frac{f_L}{\partial f_L / \partial w} \right| \tag{19}$$

is the characteristic spread in the modulated component of the initial energy distribution. We solve the linearized kinetic equation

$$\left[\frac{\partial}{\partial t} + \omega_H(w) \right] \tilde{f} = -F(\theta,t) \left[\frac{\partial f_0}{\partial w} + \frac{\partial f_L}{\partial w} \times \cos(\varphi + L\theta + \kappa X - L\omega_H t) \right] \tag{20}$$

(condition (18) makes it possible to avoid differentiating the phase $\varphi + L\theta + \kappa X - L\omega_H(w)t$ with respect to w). In an approximation in which the damping factors are not too large, we can use the solution of Eq. (20) at fixed wave amplitudes $\alpha_{1,2}$ (see, e.g., Refs. 21 and 22). Allowing for the initial condition $\tilde{f}(t=0)=0$ and the synchronism conditions (11), we arrive at the following expression for the resonant harmonics of the distribution function:

$$\begin{aligned}\tilde{f} = &-\sum_{j=1}^{2} g_j \alpha_j \exp[i(N_j\theta + k_j X - \omega_j t)]\hat{\xi}(\Delta_j) \frac{\partial f_0}{\partial w} \\ &- \frac{1}{2} g_2 \alpha_2 \exp[i(N\theta + kX - \omega t)]\sigma\hat{\xi}(\Delta_2) \frac{\partial f_L}{\partial w} \exp(i\varphi) \\ &- \frac{1}{2} g_1 \alpha_1 \exp\{i[(N-L)\theta + (k-\kappa)X \\ &- (\omega - \Omega)t]\}\sigma^* \hat{\xi}(\Delta_1) \frac{\partial f_L}{\partial \omega} \exp(-i\varphi).\end{aligned} \tag{21}$$

Here

$$\hat{\xi}(x) = \frac{i}{x}[1 - \cos(xt)] + \frac{\sin(xt)}{x},$$

$$\Delta_1 = \omega - N\omega_H(w), \quad \Delta_2 = \omega - \Omega - (N-L)\omega_H(w),$$

$$\sigma = \exp[i(L\omega_H - \Omega)t] = \exp\left\{ i \frac{L\omega_R}{\gamma_0} wt \right\}.$$

In view of the inequality (18), in (21) we can put

$$\sigma = \sigma^* = 1. \tag{22}$$

If

$$\frac{\omega_{1,2}}{\gamma_0}\langle \Delta w \rangle t \gg 1, \quad \frac{\omega_{1,2}}{\gamma_0}\langle \Delta w_0 \rangle t \gg 1 \tag{23}$$

(here $\langle \Delta w_0 \rangle$ is the scale of the stationary distribution $f_0(w)$), the operator $\hat{\xi}$ tends to its standard form (see, e.g., Refs. 9 and 22):

$$\hat{\xi} \to \frac{i\mathscr{P}}{x} + \pi\delta(x). \tag{24}$$

Thus, we assume that the asymptotic solution corresponding to Landau's ''bypass rule'' is able to form in times much shorter than the ''ballistic'' relaxation time of the high-frequency modulation distribution function. Here the corresponding inequalities (18) and (23) can only be satisfied simultaneously if the problem contains an additional small parameter:

$$\frac{\Omega}{\omega} = \frac{L}{N} = \eta \ll 1. \tag{25}$$

Plugging (16), (21), (22), and (24) into the wave excitation equations (15) and allowing for (11), we obtain

$$\begin{aligned}&\dot{\alpha}_1 + (\gamma_l + i\delta_l)\alpha_1 = -a e^{i\varphi}(\gamma_n + i\delta_n)\alpha_2, \\ &\dot{\alpha}_2 + b(\gamma_l + i\delta_l)\alpha_2 = -e^{-i\varphi}(\gamma_n + i\delta_n)\alpha_1,\end{aligned} \tag{26}$$

where

$$\begin{aligned}&\gamma_l = -\frac{2\pi^2 e^2}{m} g_1^2 \int \frac{\partial f_0}{\partial w} \delta(\omega - N\omega_H(w))dw, \\ &\gamma_n = -\frac{\pi^2 e^2}{m} g_1 g_2 \int \frac{\partial f_L}{\partial w} \delta(\omega - N\omega_H(w))dw, \\ &\delta_l = -\frac{2\pi e^2}{m} g_1^2 \int \frac{\mathscr{P}\partial f_0/\partial w}{\omega - N\omega_H(w)} dw, \\ &\delta_n = -\frac{\pi e^2}{m} g_1 g_2 \int \frac{\mathscr{P}\partial f_L/\partial w}{\omega - N\omega_H(w)} dw,\end{aligned} \tag{27}$$

$$a = \frac{1}{1-\eta}, \quad b = \frac{g_2^2}{g_1^2} a. \tag{28}$$

The exponential solutions ($\alpha_{1,2} \sim e^{\mu t}$) of the system of equations (26) can be found from the characteristic equation

$$(\mu + \gamma_l + i\delta_l)[\mu + b(\gamma_l + i\delta_l)] = a(\gamma_n + i\delta_n)^2 \tag{29}$$

and a relationship that follows from (26):

$$\frac{\alpha_1}{\alpha_2} = K = \frac{-a e^{i\varphi}(\gamma_n + i\delta_n)}{\mu + \gamma_l + i\delta_l}. \tag{30}$$

In the simplest case, where $\delta_l = \delta_n = 0$[14], the system of equations (26) corresponds to Eqs. (8), which describe ''inversionless'' amplification of waves in the quantum system. When $\gamma_l < 0$, an unstable solution exists if

$$a\gamma_n^2 > b\gamma_l^2,$$

which reduces to a physically transparent analog of the ''quantum'' condition (9):

$$\frac{1}{2} \left| \frac{\partial f_L}{\partial w} \right|_{\omega_H = \omega/N} > -\left(\frac{\partial f_0}{\partial w} \right)_{\omega_H = \omega/N}. \tag{31}$$

This inequality guarantees the onset of strong ''inversion'' in the phase space that is periodic in the variables θ, X, and t.

The solution of the characteristic equation (29) in general form is cumbersome; nevertheless, there is no doubt that there are unstable solutions with Re $\mu > 0$ when $\partial f_L / \partial w$ is large. For instance, if

$$|\gamma_n + i\delta_n| \gg |(1-b)(\gamma_l + i\delta_l)|,$$

an unstable solution exists if an additional inequality holds, which differs from (31) only in the coefficient on the left-hand side:

$$\frac{1}{2} \rightarrow \frac{1+b}{4\sqrt{a}}\ .$$

Just as in the system with lumped parameters of Sec. 2, an unstable solution of Eqs. (26) corresponds to a certain ''optimum'' synchronization of the beats of the high-frequency oscillations and the low-frequency modulation of the system. Indeed, let us use the following solution of system (26) (a modified Manley–Rowe relation):

$$\left(\frac{\partial}{\partial t}+2\gamma_l\right)\frac{|\alpha_1|^2}{2a}+\left(\frac{\partial}{\partial t}+2b\gamma_l\right)\frac{|\alpha_2|^2}{2}$$
$$=-\mathrm{Re}(\gamma_n\alpha_2\alpha_1^* e^{i\varphi}).$$

For an unstable solution this yields

$$-\mathrm{Re}\ \gamma_n K^* e^{i\varphi}=\frac{\mathrm{Re}\ \mu+\gamma_l}{a}|K|^2+\mathrm{Re}\ \mu+b\gamma_l>0. \qquad (32)$$

If we allow for (27), Eq. (32) yields

$$g_1 g_2\ \mathrm{Re}\left[e^{i\varphi}K^*\left(\frac{\partial f_L}{\partial w}\right)_{\omega_H(w)=\omega/N}\right]>0. \qquad (33)$$

According to this condition, the nodes and the crests of the beats of the ''effective force''

$$F=\mathrm{Re}[\{\alpha_1 g_1+\alpha_2 g_2\ \exp[-i(L\theta+\kappa X-\Omega t)]\}$$
$$\times\exp[i(N\theta+kX-\omega t)]]$$

are at points with, respectively, a negative and positive energy derivative of the oscillating component of the distribution function,

$$\tilde{f}_L\approx f_L(w)\cos(\varphi+L\theta+\kappa X-\Omega t).$$

Thus far we have said nothing about the ways in which a modulated initial distribution (17) that satisfies (31) could be formed. Without going into detailed analysis of the problem, we will demonstrate that such a distribution can be formed.

Suppose that at time $t=-T$ there exists a distribution over the transverse energies of the ''smeared-ring'' type:

$$f_0(w)=\frac{n_b}{\sqrt{\pi}\langle\Delta w\rangle}\exp\frac{(w-w_b)^2}{\langle\Delta w\rangle^2}. \qquad (34)$$

The resonance energy $mc^2\gamma_0$ satisfying the synchronism condition (11) corresponds to the value $w\equiv\gamma-\gamma_0=0$. The particle energy $\mathscr{E}_b=mc^2(\gamma_0+w_b)$ corresponds to the maximum of the distribution function (at $w=w_b$). Let us see what happens when $w_b<0$, i.e., when the resonant particles are on the stable ''inversionless'' slope of the unperturbed distribution function.

Suppose that in the time interval from $t=-T$ to $t=0$ the system is subjected to an electromagnetic wave field with the ''low'' frequency $\Omega=L\omega_R$ and wave vector κ:

$$\mathbf{E}_0=\mathbf{y}_0\ \mathrm{Re}[E_0\ \exp(i\kappa x-i\Omega t)].$$

The corresponding linear perturbation of the distribution function at $t=0$ is

$$\tilde{f}_L=f_L\ \exp(iL\theta+i\kappa X), \qquad (35)$$

where

$$f_L=-g_0\alpha_0\frac{\partial f_0}{\partial w}\int_{-T}^{0}\exp\{i[L\omega_H(w)-\Omega]t\}dt,$$

$$\alpha_0=\frac{eE_0}{mc},\quad g_0=\frac{p_{\perp 0}}{\gamma_0 mc}J_L'(\chi_0),\quad \chi_0=\kappa r_H.$$

The simplest case is the one in which the pump period is short, or

$$T\omega_R\ll\frac{\gamma_0}{\langle\Delta w\rangle L}. \qquad (36)$$

In this approximation we arrive at the simple relationship

$$f_L=-g_0\alpha_0 T\frac{\partial f_0}{\partial w}.$$

Therefore, the instability criterion (31) is met if

$$\frac{T}{2}\left|g_0\alpha_0\frac{\partial^2 f_0}{\partial w^2}\right|_{w=0}>-\left(\frac{\partial f_0}{\partial w}\right)_{w=0}, \qquad (37)$$

which implies that the inequality

$$\langle\Delta w\rangle\gg\frac{T}{2}|\alpha_0 g_0|>|w_b|$$

guarantees that an ''inversionless'' instability appears even in the linear approximation in the pump amplitude. The case $w_b=0$ corresponds to a ''nonthreshold'' instability, since here the resonant particles are at the maximum of the distribution function (34) and the linear damping factors of the high-frequency waves vanish.

Above we examined the instability when the initial energy distribution is of a highly nonequilibrium nature. Nevertheless, in the region where resonant particles can exist the system is indeed ''inversionless'' from the standpoint of the standard linear theory. Here a resonance on the ''inversionless'' slope of the distribution function corresponds to the generation of waves with a frequency higher than in the case of a resonance with particles belonging to the inversion region. What is important here is that for initial distributions that monotonically decrease with increasing energy, an ''inversionless'' instability can probably exist only if the time during which the low-frequency field is applied to the system is comparable to the period of particle ''capture'' [15] in the field of a finite-amplitude wave,[16]

$$T_b=2\pi(L\omega_R g_0\alpha_0)^{-1/2}.$$

Concluding this section, we note that a similar parametric instability involving resonant particles can occur, at least in principle, in a Cherenkov resonance. This follows from the formal equivalence of Eqs. (12) and the equations of motion of charged particles in the field of longitudinal waves. However, the necessary conditions are met only if the longitudinal waves are retarded and exhibit no dispersion, or

$$\frac{\omega}{k}=\frac{\partial\omega}{\partial k}<c.$$

In this sense the cyclotron system considered here is more promising, since ‘‘vacuum’’ electromagnetic modes are sufficient for its realization.

4. QUASILINEAR THEORY OF CYCLOTRON PARAMETRIC INSTABILITY INVOLVING RESONANT PARTICLES

Analysis of the energy relations in the given system requires that we step outside the scope of linear theory. For a proper generalization, we use the machinery of the quasilinear theory of waves in plasmas. In that spirit (see, e.g., Refs. 1 and 24), we assume that in small neighborhoods of the cyclotron frequencies $N\omega_R$ and $(N-L)\omega_R$ the excited spectra are those with the set of frequencies $\omega_{1p}=\omega_p$ and $\omega_{2p}=\omega_p(N-L)/N$, where $p=1,2,\dots,\infty$. As in the standard approach of the quasilinear theory, we consider a set of harmonics with random phases. However, the complex amplitudes E_{1p} and E_{2p} of the pair of harmonics with frequencies ω_p and $\omega_p(N-L)/N$ are related by (30), since each such pair is a linear solution of the second-order system (26). Moreover, building a quasilinear theory in this case is more complicated because we must have a separate relationship for the low-frequency oscillations of the distribution function with frequency $\Omega\approx L\omega_R$.

So as not to complicate the discussion with cumbersome formulas, we introduce simplifying (but actually unimportant) assumptions. First, we assume that the ‘‘coupling constants’’ in the equations of motion (12) and (13) are approximately equal,

$$g\simeq g_{1,2}\gg|g_1-g_2|. \tag{38}$$

Allowing for the inequality (25), in the system of equations (26) we put

$$a\simeq 1,\quad b\simeq 1 \tag{39}$$

and use simplified relationships for the instability growth rate and the coupling constant that links the complex amplitudes of the amplified bichromatic radiation:

$$\mathrm{Re}\ \mu=\frac{2\pi^2e^2}{m}g^2\int\left[\frac{\partial f_0}{\partial w}+\frac{1}{2}\left|\frac{\partial f_L}{\partial w}\right|\right]\times\delta(\omega-N\omega_H(w))dw, \tag{40}$$

$$K=\frac{\alpha_1}{\alpha_2}=e^{i\varphi}\ \mathrm{sgn}\,\frac{\partial f_L}{\partial w}. \tag{41}$$

In addition, if we allow for the condition $|N\omega_R-\omega_p|\ll\omega_R$ and the inequalities (25) and (18), we can simplify the relationship that exists between the frequencies and wave vectors of the sets of resonant harmonics:

$$\omega_{1p}-\omega_{2p}\simeq\Omega,\quad k_{1p}-k_{2p}\simeq\kappa. \tag{42}$$

Because of these simplifications, the kinetic equation (14) reduces to the form

$$\frac{\partial f}{\partial t}+\omega_H(w)\frac{\partial f}{\partial\theta}=-g\sum_p\mathrm{Re}\{\alpha_{1p}\exp(iN\theta+ik_pX-i\omega_pt)\times[1+K\exp(i\Omega t-i\kappa X-iL\theta)]\}\frac{\partial f}{\partial w}. \tag{43}$$

We seek a solution averaged over the ensemble of random phases $\varphi_p=\arg\alpha_{1p}$ in the form

$$\langle f\rangle_{\varphi_p}=f_0(w,t)+f_L(w,t)\cos(\varphi+L\theta+\kappa X-\Omega t). \tag{44}$$

Note that the modulated component of the distribution function, $\langle f\rangle_{\varphi_p}$, can have such a form only if the condition (18) is met, i.e., when we can ignore the ‘‘ballistic term’’ 17) $i[L\omega_H(w)-\Omega]f_L$ in the kinetic equation (43). Next, in Eq. (43) we use the usual perturbation method to find the bilinear corrections $\sim\alpha^2$, and average over the ensemble of random phases φ_p. Allowing for (18), (39), and (41), and replacing the sum over the discrete spectrum by an integral over the continuous spectral intensity (see, e.g., Refs. 1 and 24),

$$\sum_p|\alpha_{1p}|^2(\cdots)\rightarrow\int_\omega|\alpha_1|^2_\omega(\cdots)d\omega,$$

we arrive at the following equations for the distribution function:

$$\frac{\partial f_0}{\partial t}=\frac{\partial}{\partial w}\left[D(w)\left(\frac{\partial f_0}{\partial w}+\frac{1}{2}\left|\frac{\partial f_L}{\partial w}\right|\right)\right], \tag{45}$$

$$\frac{\partial f_L}{\partial t}=\frac{\partial}{\partial w}\left[D(w)\left(\mathrm{sgn}\,\frac{\partial f_L}{\partial w}\right)\frac{\partial f_0}{\partial w}\right], \tag{46}$$

where

$$D(w)=g^2\pi\int_\omega|\alpha_1|^2_\omega\delta(\omega-N\omega_H(w))d\omega. \tag{47}$$

The equations for the spectral intensities $|\alpha_1|^2_\omega$ and $|\alpha_2|^2_\omega$ follow from linear theory. Using Eq. (40), we get

$$\frac{\partial}{\partial t}|\alpha_1|^2_\omega=\frac{4\pi^2e^2g^2}{m}|\alpha_1|^2_\omega\int\left(\frac{\partial f_0}{\partial w}+\frac{1}{2}\left|\frac{\partial f_L}{\partial w}\right|\right)\times\delta(\omega-N\omega_H(w))dw, \tag{48}$$

and Eq. (41) yields

$$|\alpha_1|^2_\omega\simeq|\alpha_2|^2_{\omega-\Omega}, \tag{49}$$

in view of which we can express the diffusion coefficient $D(w)$ in terms of the spectral intensity $|\alpha_1|^2_\omega$.

The system of equations (45)–(48) satisfies the conservation laws

$$\int_w f_0(w)dw=\mathrm{const},$$

$$mc^2\int_w(\gamma_0+w)f_0(w)dw+I_e=\mathrm{const}, \tag{50}$$

where

$$I_e=\frac{m^2c^2}{e^2}\int_\omega\left(\frac{|\alpha_1|^2_\omega}{8\pi}+\frac{|\alpha_2|^2_\omega}{8\pi}\right)d\omega=\frac{m^2c^2}{e^2}\int_\omega\frac{|\alpha_1|^2_\omega}{4\pi}d\omega \tag{51}$$

is the energy of the electromagnetic field. Moreover, Eq. (46) yields an extremely graphic relationship for the evolution of the amplitude of the modulated component of the ''slow'' distribution function:

$$\frac{\partial}{\partial t}\int f_L^2(w)\,dw = -\frac{1}{2}\int\left|\frac{\partial f_L}{\partial w}\right| D(w)\frac{\partial f_0}{\partial w}\,dw. \tag{52}$$

On the basis of the above relationships of the quasilinear theory we can draw some analogies with the corresponding quantum system:

(1) From Eqs. (45) and (48) it follows that in the approximation of fixed low-frequency modulation, the ''slow'' distribution function $f_0(w)$ relaxes not to the ''standard'' plateau but to a state with a negative derivative with respect to w. This state corresponds to the threshold in the parametric instability being discussed:

$$\frac{\partial f_0}{\partial w}\to -\frac{1}{2}\left|\frac{\partial f_L}{\partial w}\right| \tag{53}$$

in the region where $D(w)\neq 0$. Therefore, in the ''inversionless'' generation process, a decrease in the energy of the electron ensemble is accompanied by the formation of a distribution function with a sharper drop in energy in the resonant-particle region. This effect corresponds to depletion of the upper active level in a quantum ''inversionless'' system.[6]

(2) Equation (52) implies that in an ''inversionless'' system, i.e, a system in which $\partial f_0/\partial w<0$, the beats of the high-frequency field amplify the low-frequency modulation of the distribution function. This effect has an analog in quantum systems: the excitation and maintenance of the coherent state of a low-frequency transition by an external monochromatic field.[6,11]

Note that if we remain solely within the scope of Eqs. (45)–(49), the possibility of reaching some well-defined stationary regime is highly unlikely. The problem here is that the stationary state defined by (53) for the average component of the distribution function f_0 and the spectral intensities $|\alpha_{1,2}|^2_\omega$ applies only in an approximation in which the amplitude of low-frequency modulation of f_L is constant. On the other hand, the quantity $f_L(w)$, as Eqs. (46) and (52) imply, continues to increase even if condition (53) is met.[18)] This tendency, to a certain extent, probably reflects the ''explosive'' nature of parametric instability in media with nonlinear conductivity, when amplification of a pair of high-frequency modes is accompanied by the amplification of low-frequency oscillations.[14] The mechanisms by which stationary states set in are most likely to be found outside the scope of our simple model (''ballistic'' relaxation of low-frequency modulation, nonlinear shift of wave frequencies, particle ''capture,'' etc.).

The conservation laws (50) suggest that in ''inversionless'' generation the high-frequency field receives energy from the ''slow'' distribution function $f_0(w)$, as it does in ordinary maser instability. By themselves the equations of the quasilinear theory and the instability condition (31) do not generally restrict the class of initial distribution functions $f_0(w)$. Restrictions do appear, however, when one uses a specific method to modulate the distribution function of the resonant particles.

5. CONCLUSION

We have found that cyclotron parametric instability involving resonant particles is a direct classical analog of ''inversionless'' generation in a three-level quantum system. The way in which one can generalize this regime to the generation of a high-frequency signal formed by an arbitrary number of distinct harmonics is quite obvious. For the classical system, however, there is the question of whether the energy emitted into the high-frequency modes in the ''inversionless'' regime can exceed the energy supplied to the system in the excitation of low-frequency modulation. This possibility becomes obvious when for the instability in question to develop we need only a linear regime of low-frequency modulation of the initial distribution function (such an example is discussed in Sec. 3). In the general case, however, there is no single answer to this question. The answer depends on the type of initial distribution function, the specific pumping scheme, and the instability saturation mechanism. The main avenues of research in this field should probably be the following:

(1) Analyzing the efficiency of various schemes of preliminary low-energy pumping for various initial electron distribution functions.

(2) Studying the scheme with a constant source of low-frequency modulation.

(3) Building a self-consistent theory of ''inversionless'' instability that would allow for the evolution of the low-frequency modulation of the distribution function for the resonant particles as a result of ''ballistic'' relaxation and the effect of a bichromatic high-frequency field.

(4) Analyzing such nonlinear effects as the ''capture'' of resonant particles in a high-power field and the nonlinear shift of frequencies and/or wave vectors of the high-frequency modes.

Solving these problems will make it possible to establish the feasibility of practical applications of the effect. Our results suggest, at least, that cyclotron parametric instability involving resonant particles can be used as a method for converting microwave radiation to a higher frequency, for example, the emission of high-order cyclotron harmonics excited by pumping at low frequencies. In particular, to implement quasistationary ''inversionless'' generation, an electron beam can first be sent through a modulating section (where electromagnetic oscillations at a low harmonic of the cyclotron frequency have been excited) and then through an active cavity.

In comparison to standard cyclotron-frequency multiplication,[20] this regime would seem to have certain advantages, the point being that direct M-fold frequency multiplication of the low-frequency pump requires a nonlinearity of the corresponding order (or a nonlinearity that is not a power law). At the same time, for the parametric instability described above, a quadratic nonlinearity is quite sufficient even at extremely high harmonics of the cyclotron frequency, while a nonlinearity of lower order naturally corresponds to a lower-power low-frequency pump. In addition, the ''inversionless'' nature of this instability suggests that such a conversion method is not very sensitive to the quality of the electron beam.

In any case, we believe that studies of the phenomena associated with ''inversionless'' generation of stimulated radiation by aggregates of classical oscillators (electrons) comprise an interesting avenue of research in the physics of the interaction of radiation and matter.

The authors are grateful to A. A. Belyanin, V. V. Kocharovskiĭ, Vl. V. Kocharovskiĭ, O. A. Kocharovskaya, A. G. Litvak, M. I. Rabinovich, V. E. Semenov, I. D. Tokman, A. M. Feĭgin, and G. M. Fraĭman for discussing the various aspects of the problems studied in this paper. Partial support for this work was provided by the Russian Fund for Fundamental Research (Grant No. 96-15-96934).

1) Classical microwave electronics successfully borrows not only the ''ideology'' of quantum systems but also structural elements of quantum devices.[4,5]

2) Of course, there is the question of just what type of equilibrium this really is. It would also be interesting to know to what extent the tendency of the energy to relax to a plateau is universal.

3) Obviously, this possibility exists only in nonlinear regimes, since in a linear system a nonmonochromatic perturbation cannot lead to new effects due to the superposition principle.

4) In particular, in electron masers and lasers this tendency tightens the requirements on the quality of the electron beams.[7]

5) In this case, naturally, there is a low-frequency polarization wave in the system.

6) Actually here we are dealing with ordinary Raman scattering.

7) In the approach based on phenomenological equations for the density matrix, the expressions for the coefficients $\gamma_{1,2}$ and $G_{1,2}$ contain a damping constant instead of $\pi\tau(0)$, the normalized density of states.[6,11]

8) In Ref. 6 the inequality (9) is obtained with a model that uses phenomenological equations for the density matrix.

9) At $n_3=0$ the condition (9) cannot be met, since $|\sigma_{21}|^2\leqslant n_1 n_2$ in view of a property of the density matrix.[8] On the other hand, the instability condition (9) is entirely independent of the dipole-moment matrix element of the low-frequency transition.

10) If there is no negative constant resistance ($R_a=0$), the instability condition can be met only if there is extremely strong modulation ($\beta I_L>2$), whereupon the sign of the resistance in the circuit is manifestly alternating. This is a well-known property (see, e.g., Ref. 12) of oscillatory systems: modulation of an always-positive resistance (or, in a mechanical system, the friction coefficient) does not lead to instability, since under such modulation no energy is supplied to the system.

11) If the combination-resonance conditions are satisfies but only the reactive elements in the low-frequency circuit are modulated, the Manley–Rowe relations $\Sigma_{1,2}(d/dt+2\delta_{1,2})L_{1,2}\omega_{1,2}^{-1}I_{1,2}^2=0$ blocks instability development.

12) For a wave with negative energy (in the ordinary sense of the word) any dissipation is a ''feedback'' mechanism that makes it possible to induce an energy excess in the nonequilibrium medium. In this sense, parametric coupling to a wave with positive energy only acts as a kind of effective dissipation. As for the instability of positive-energy waves discussed here, the fact that the high-frequency field is nonmonochromatic is most important for its realization.

13) Obviously, the corresponding quantum system must meet a symmetric requirement: the time intervals in question must be short compared to the relaxation time of the coherent state of the low-frequency transition.

14) The frequency shifts δ_l and $b\delta_l$ can be ''removed,'' for example, by modifying the external electrodynamic system, and δ_n vanishes for a w-odd perturbation $f_L(w)$.

15) As applied to cyclotron resonance, the capture effect is described, for instance, in Refs. 19 and 21.

16) In a quantum system, a similar method of preliminary preparation of the necessary low-frequency coherence consists in applying a $\pi/2$-pulse of low-frequency radiation to the system.[23]

17) Another reason why this can be done is that for $N\gg L$ the relaxation time of the high-frequency field is much shorter than the time of ballistic ''smearing'' of the modulated component of the distribution function f_L.

18) The inverse is also true: the condition that the modulation amplitude must be stationary, $\partial f_0/\partial w=0$, which follows from (46) and (52), is not valid for the functions $f_0(w)$ and $|\alpha_{1,2}|_\omega^2$ if $\partial f_L/\partial w\neq 0$. As for the obvious stationary regime in which $\partial f_0/\partial w=\partial f_L/\partial w=0$ with $D(w)\neq 0$, the conservation laws (50) imply that with an unstable ''inversionless'' initial state the relaxation to such a stationary regime is forbidden by the condition $I_e>0$.

---

[1] B. B. Kadomtsev, *Collective Phenomena in Plasmas*, Pergamon Press, Oxford (1980).

[2] V. V. Zheleznyakov, *Electromagnetic Waves in Cosmic Plasma* [in Russian], Nauka, Moscow (1977).

[3] A. V. Gaponov, M. I. Petelin, and V. K. Ulpatov, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **10**, 1414 (1967).

[4] F. Hopf, T. Cooper, D. Moore, and M. Scully, in *Free-Electron Generators of Coherent Radiation* [in Russian], A. A. Rukhadze (ed.), Mir Publishers, Moscow (1983) [Russian translation of a collection of papers on free-electron lasers].

[5] A. V. Gaponov-Grekhov and M. I. Petelin, in *Science and Mankind* [in Russian], Znanie, Moscow (1980), p. 283.

[6] O. Kocharovskaya, Phys. Rep. **219**, 175 (1992).

[7] V. L. Bratman, N. S. Ginzburg, G. S. Nusinovich *et al.*, in *Relativistic High-Frequency Electronics*, A. V. Gaponov-Grekhov (ed.), Institute of Applied Physics of the USSR Academy of Sciences, Gorkiĭ (1979), p. 157.

[8] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-relativistic Theory*, 3rd ed., Pergamon Press, Oxford (1977).

[9] V. M. Faĭn, *Photons and Nonlinear Media* [in Russian], Sovetskoe Radio, Moscow (1972).

[10] E. Arimondo and G. Orriolis, Nuovo Cimento Lett. **17**, 333 (1976).

[11] O. Kocharovskaya and P. Mandel, Phys. Rev. A **42**, 523 (1990).

[12] L. I. Mandel'shtam, ''Notes for a monograph on parametric generation of alternating current,'' in *Complete Collected Works of L. M. Mandel'shtam* [in Russian], Vol. II, S. M. Rytov (ed.), USSR Academy of Sciences Press, Leningrad (1947), p. 342.

[13] A. B. Mikhaĭlovskiĭ, *Theory of Plasma Instabilities* [in Russian], Atomizdat, Moscow (1975), p. 32 [English transl.: 2 vols., Consultants Bureau (1974)].

[14] S. V. Kiyashko, M. I. Rabinovich, and V. P. Reutov, JETP Lett. **16**, 271 (1972).

[15] B. Sherman, G. Kurizki, D. E. Nikonov, and M. O. Scully, Phys. Rev. Lett. **75**, 4602 (1995).

[16] D. E. Nikonov, B. Sherman, G. Kurizki, and M. O. Scully, Opt. Commun. **123**, 363 (1996).

[17] N. A. Vinokurov and A. P. Skrinskiĭ, in *Relativistic High-Frequency Electronics*, A. V. Gaponov-Grekhov (ed.), Institute of Applied Physics of the USSR Academy of Sciences, Gorkiĭ (1981), p. 204.

[18] A. B. Kitsenko, I. M. Pankratov, and K. N. Stepanov, Zh. Tekh. Fiz. **45**, 912 (1975) [Sov. Phys. Tech. Phys. **20**, 575 (1975)].

[19] A. G. Litvak, A. M. Sergeev, E. V. Suvorov, M. D. Tokman, and I. V. Khazanov, Phys. Fluids B **5**, 4347 (1993).

[20] N. S. Ginzburg, Zh. Tekh. Fiz. **56**, 1433 (1986) [Sov. Phys. Tech. Phys. **31**, 571 (1986)].

[21] E. V. Suvorov and M. D. Tokman, Plasma Phys. **25**, 723 (1983).

[22] T. O'Neil, Phys. Fluids **8**, 2255 (1965).

[23] O. A. Kocharovskaya, Dr. Sci. (Phys.–Math.) thesis, Institute of Applied Physics of the Russian Academy of Sciences, Nizhniĭ Novgorod (1996).

[24] A. I. Akhiezer, I. A. Akhiezer, R. V. Polovin, A. G. Sitenko, and K. N. Stepanov, *Plasma Electrodynamics* [in Russian], Nauka, Moscow (1974), p. 443 [English transl.: 2 vols., Pergamon Press, Oxford (1975)].

Translated by Eugene Yankovsky

The generalized Bethe logarithm for tightly bound electrons

I. A. Goĭdenko and L. N. Labzovskiĭ

Physics Institute, St. Petersburg State University, 198904 St. Petersburg, Russia

(Submitted 24 March 1997)
Zh. Éksp. Teor. Fiz. **112**, 1197–1208 (October 1997)

We derive a closed relativistic expression that makes it possible to calculate the self-energy of multiply charged ions in an external Coulomb field without resorting to a series expansion in powers of αZ. The expression contains the generalized Bethe logarithm for tightly bound electrons. We do numerical calculations of the self-energy for the $1s_{1/2}$-electrons of multiply charged hydrogenlike ions. The proposed method allows for self-energy calculations for any values of the nuclear charge Z. *© 1997 American Institute of Physics.* [S1063-7761(97)00410-1]

1. INTRODUCTION

Experimental achievements in measuring the Lamb shift in multiply charged ions[1–3] require raising the accuracy of theoretical calculations and, in particular, allowing for quantum mechanical corrections of order α^2, where α is the fine-structure constant. Various groups of researchers have done such calculations,[4–6] whose main feature is the lack of a series expansion in powers of αZ, where Z is the atomic number, or nuclear charge.

What remains to be calculated is the correction to the electron self-energy in the second order. The general approach to calculating such corrections, including renormalization and removal of IR divergences, is discussed by Labzovsky and Mitrushenkov.[7] However, it has been established that not one of the methods used in self-energy calculations in the lowest order[8–10] is suitable here. In the present paper we propose another calculation method, which, we believe, can be generalized in a natural way to second order. The method is used to calculate the first-order electron self-energy and is compared with the existing methods.[11] In contrast to these, our method provides equally high accuracy for large and small values of Z. One other special feature of this method is that it actually produces a closed expression for the self-energy that contains the generalized Bethe logarithm.[12]

All theoretical calculations use the relativistic system of units, $m=\hbar=c=1$, where m is the electron mass.

2. THE NONREGULARIZED EXPRESSION FOR THE SELF-ENERGY

The Feynman diagram corresponding to the unrenormalized expression for the electron self-energy is depicted in Fig. 1. According to the classification scheme in Ref. 13, the diagram is irreducible, so that its contribution to the energy can be written

$$\Delta E_A=\langle\Phi_A|U^n|\Phi_A\rangle, \tag{1}$$

where U^n is the effective potential energy operator in nth-order perturbation theory, which in Ref. 13 is defined as

$$\langle\Phi_B|S^n|\Phi_A\rangle=-2\pi i\,\delta(E_A-E_B)\langle\Phi_B|U^n|\Phi_A\rangle. \tag{2}$$

Here S^n is the evolution operator in nth-order perturbation theory, and E_A and E_B are the electron's unperturbed energy levels in the external field of the nucleus.

According to the Feynman rules, for an electron in an external field (the Furry scheme), the contribution of diagram (a) can be written

$$\langle\Phi_A|S^2|\Phi_A\rangle=e^2\int d^4x_1d^4x_2(\overline{\Psi}_A(x_2)\times\gamma_\mu S(x_2,x_1)\gamma_\nu\Psi_A(x_1))D_{\mu\nu,\mathrm{ir}}(x_1,x_2), \tag{3}$$

where $S(x_1,x_2)$ is the electron propagator,

$$S(x_2,x_1)=\frac{1}{2\pi i}\int_{-\infty}^{\infty}d\omega\,\exp[i\omega(t_1-t_2)]\times\sum_n\frac{\Psi_n(\mathbf{r}_2)\overline{\Psi}_n^{-}(\mathbf{r}_1)}{E_n(1-i0)+\omega}, \tag{4}$$

and $D_{\mu\nu}(x_1,x_2)$ is the photon propagator,

$$D_{\mu\nu}(x_1,x_2)=\frac{4\pi}{i}\frac{\delta_{\mu\nu}}{(2\pi)^4}\int\exp[i\mathbf{k}\cdot(\mathbf{r}_1-\mathbf{r}_2)-\omega t]\times\frac{1}{\mathbf{k}^2-\omega^2-i0}d^3k\,d\omega, \tag{5}$$

with $\mathbf{k}$ and ω the wave vector and the frequency of the virtual photon.

Integrating in (3) with respect to time and frequency, we arrive at an expression for the electron's nonregularized self-energy in the lowest order:

$$\Delta E_A^{NR}=\frac{e^2}{2\pi i}\sum_n\left(\frac{1-\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2}{r_{12}}I_{nA}(r_{12})\right)_{AnnA}, \tag{6}$$

where $r_{12}=|\mathbf{r}_1-\mathbf{r}_2|$, and we have used the notation

$$(\hat{A})_{ABCD}=\int(\Psi_A^+(\mathbf{r}_1)\Psi_B^+(\mathbf{r}_2)\hat{A}\Psi_C(\mathbf{r}_1)\times\Psi_D(\mathbf{r}_2))d^3x_1d^3x_2. \tag{7}$$

The quantity $I_{nA,\mathrm{ir}}(r_{12})$ is defined as follows:

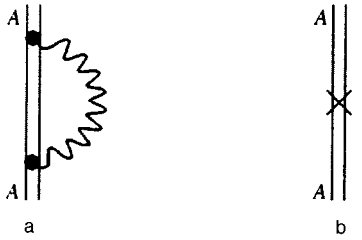

FIG. 1. Feynman diagram corresponding to the electron self-energy in lowest order. Diagram (a) corresponds to the unrenormalized self-energy. The double solid line represents the electron in the field of the nucleus and the wavy line, the photon. Diagram (b) corresponds to the counterterm subtracted from the contribution of diagram (a). The letter A stands for the state of the electron in the atom.

$$I_{nA}(r_{12})=\int_{-\infty}^{\infty}d\omega\int_{-\infty}^{\infty}d\kappa \times\frac{\kappa\exp(i|\omega|r_{12})}{(\omega-\kappa+i0)(\omega+\kappa-i0)[\beta_{nA}+\omega-i\,\mathrm{sgn}(E_n)\,0]}, \tag{8}$$

where $\beta_{nA}=E_n-E_A$. Integrating this expression with respect to ω and κ, we obtain

$$I_{nA}(r_{12})=2i\frac{|\beta_{nA}|}{\beta_{nA}}[\mathrm{ci}(|\beta_{nA}|r_{12})\sin(|\beta_{nA}|r_{12}) -\mathrm{si}(|\beta_{nA}|r_{12})\cos(|\beta_{nA}|r_{12})]+\frac{\pi i}{2}\left(1+\frac{E_n}{|E_n|}\right) \times\left(1-\frac{|\beta_{nA}|}{\beta_{nA}}\right)\exp(i|\beta_{nA}|r_{12}), \tag{9}$$

where si and ci are the sine and cosine integrals, respectively. For the positive part of the Dirac spectrum in the nonrelativistic limit we have $\beta_{nA}\sim(\alpha Z)^2$ and $r_{12}\sim(\alpha Z)^{-1}$. Hence a series expansion of (11) in powers of $\beta_{nA}r_{12}$ is equivalent to a series expansion in αZ, although the two do not coincide (see below). Allowing for the fact that in the limit $\alpha Z\ll 1$ the expression for the electron self-energy is of order $v\alpha(\alpha Z)^4$ (see Ref. 12), we should assume that this expansion begins with $(\beta_{nA}r_{12})^3$. However, the first term in the expansion contains $(\beta_{nA}r_{12})^0$. Furthermore, this term yields a divergent expression,

$$\sum_n\left(\frac{1}{r_{12}}\right)_{AnnA}=\int d^3r_1d^3r_2\Psi_A^+(\mathbf{r}_1)\,\delta(r_{12})\,\frac{1}{r_{12}}\Psi_A(\mathbf{r}_2)=\infty, \tag{10}$$

due to the completeness of the Dirac spectrum. This divergence and the discrepancy between the expansion in powers of αZ and the nonrelativistic limit disappear in the process of renormalization.

3. RENORMALIZATION

Following Ref. 8, we use the Pauli–Villars renormalization scheme. To this end we write the photon propagator in the form

$$D^{\Lambda}_{\mu\nu}(x_1,x_2)=\frac{4\pi}{i}\frac{\delta_{\mu\nu}}{(2\pi)^4}\int d^3k\,d\omega \times\exp\{i\mathbf{k}\cdot(\mathbf{r}_1-\mathbf{r}_2)-\omega t\} \times\left[\frac{1}{\mathbf{k}^2-\omega^2-i0}-\frac{1}{\mathbf{k}^2-\omega^2+\Lambda^2-i0}\right]. \tag{11}$$

Now we must subtract diagram (b) from diagram (a) (see Fig. 1), with the former expressed as follows in the Feynman gauge:

$$\delta m^{\Lambda}_{A}=\frac{e^2}{\pi}\left(\frac{3}{4}\ln\Lambda^2+\frac{3}{8}\right)(\beta)_{AA}, \tag{12}$$

where β is the Dirac matrix.

Plugging (11) into (3), we arrive at ΔE^{Λ}_{A} instead of (6), and to calculate the energy shift we must find the limit

$$\Delta E_A=\lim_{\Lambda\to\infty}(\Delta E^{\Lambda}_{A}-\delta m^{\Lambda}_{A}). \tag{13}$$

By analogy with (6), the expression for ΔE^{Λ}_{A} can be written

$$\Delta E^{\Lambda}_{A}=\frac{e^2}{2\pi i}\sum_n\left(\frac{1-\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2}{r_{12}}I^{\Lambda}_{nA}(r_{12})\right)_{AnnA}. \tag{14}$$

In calculating (14) we split I^{Λ}_{nA} into three terms:

$$I^{\Lambda1}_{nA}=-2i\frac{E_n}{|E_n|}\Lambda^2\int_0^{\infty}\frac{\kappa\sin(\kappa r_{12})d\kappa}{(\kappa^2-\beta^2_{nA})(\kappa^2-\beta^2_{nA}+\Lambda^2)}, \tag{15}$$

$$I^{\Lambda2}_{nA}=2i\beta_{nA}\int_0^{\infty}\frac{\sin(\kappa r_{12})d\kappa}{\beta^2_{nA}-\kappa^2} =2i\frac{\beta_{nA}}{|\beta_{nA}|}\left\{\sin(|\beta_{nA}|r_{12})\mathrm{ci}(|\beta_{nA}|r_{12}) -\cos(|\beta_{nA}|r_{12})\left[\mathrm{si}(|\beta_{nA}|r_{12})+\frac{\pi}{2}\right]\right\}, \tag{16}$$

$$I^{\Lambda3}_{nA}=-2i\beta_{nA}\int_0^{\infty}\frac{\kappa\sin(\kappa r_{12})d\kappa}{(\beta^2_{nA}-\kappa^2-\Lambda^2)\sqrt{\kappa^2+\Lambda^2}}. \tag{17}$$

The expression (16) is an analog of $I_{nA}(r_{12})$, which, however, lacks the term (12) containing the divergence. The term (15) contributes nothing to (14), so that we must calculate $I^{\Lambda3}_{nA}$.

4. CALCULATION OF $I^{\Lambda3}_{nA}$

Let us first assume that in (17)

$$\frac{\beta^2_{nA}}{\kappa^2+\Lambda^2}<1$$

for all β_{nA}. This means that for any state of the Dirac spectrum there exists a Λ such that $|\beta_{nA}|<\Lambda$. Then the integrand in (17) can be expanded in a series:

$$\frac{\kappa\,\sin(\kappa r_{12})}{(\beta_{nA}^2-\kappa^2-\Lambda^2)\sqrt{\kappa^2+\Lambda^2}}=\sum_{k=0}^{\infty}\frac{\kappa\,\sin(\kappa r_{12})}{\sqrt{\kappa^2+\Lambda^2}}\left(\frac{\beta_{nA}}{\kappa^2+\Lambda^2}\right)^{2k}. \tag{18}$$

Using the fact that

$$\int_0^{\infty}\frac{\kappa\,\sin(\kappa r_{12})d\kappa}{(\kappa^2+\Lambda^2)^{\gamma}}=\frac{(-\gamma+1)r_{12}}{2}\int_0^{\infty}\frac{\kappa\,\cos(\kappa r_{12})d\kappa}{(\kappa^2+\Lambda^2)^{\gamma-1}} \tag{19}$$

and comparing (17), (18), and (19), we find that

$$\begin{aligned}I_{nA}^{\Lambda 3}(r_{12})&=-2i\sum_{k=0}^{\infty}\beta_{nA}^{2k+1}\frac{r_{12}}{2k+1}\int_0^{\infty}\frac{\cos(\kappa r_{12})d\kappa}{(\kappa^2+\Lambda^2)^{k+1/2}}\\&=-2i\sum_{k=0}^{\infty}\beta_{nA}^{2k+1}\frac{r_{12}}{2k+1}\frac{\sqrt{\pi}}{(2\Lambda)^k}\frac{1}{\Gamma(k+1/2)}\\&\quad\times r_{12}^kK_k(\Lambda r_{12}),\end{aligned} \tag{20}$$

where $K_k(x)$ is the Bessel function of an imaginary argument. Expanding this function in a Taylor series, we finally obtain

$$\begin{aligned}I_{nA}^{\Lambda 3}(r_{12})&=-4i\sum_{k=0}^{\infty}\beta_{nA}^{2k+1}\frac{r_{12}}{2k+1}\frac{\sqrt{\pi}}{\Gamma(k+1/2)}\left(\frac{r_{12}}{2\Lambda}\right)^k\\&\quad\times\left\{\frac{1}{2}\sum_{m=0}^{k-1}(-1)^m\frac{(k-m-1)!}{m!(\Lambda r_{12}/2)^{k-2m}}\right.\\&\quad+(-1)^{k+1}\sum_{m=0}^{\infty}\frac{(\Lambda r_{12}/2)^{k+2m}}{m!(k+m)!}\left[\ln\frac{\Lambda r_{12}}{2}\right.\\&\quad\left.\left.-\frac{1}{2}\psi(m+1)+\frac{1}{2}\psi(m+k+1)\right]\right\},\end{aligned} \tag{21}$$

where

$$\psi(x)=\frac{d}{dx}\ln\,\Gamma(x).$$

Before we plug (21) into (14), let us briefly explain the further transformations.

5. CALCULATION OF MULTIPLE COMMUTATORS

The energy difference β_{nA} in (21) can be written in terms of commutators by plugging (21) into (14). Indeed,

$$(\beta_{nA}\hat{O})_{AnnA}=([\hat{h}_2,\hat{O}])_{AnnA}=-([\hat{h}_1,\hat{O}])_{AnnA}, \tag{22}$$

where $\hat{O}$ is an arbitrary operator, and $\hat{h}_i$ is the Dirac Hamiltonian in the i-coordinate space:

$$\hat{h}_i=\boldsymbol{\alpha}_i\hat{p}_i+\beta m-eU(r_i). \tag{23}$$

Here β is the Dirac matrix, $eU(r)$ is the potential energy, and $\mathbf{p}_i=-i\boldsymbol{\nabla}_i$.

By making use of formula (22) in expressions containing β_{nA} raised to any power it is possible to introduce multiple commutators. For instance, replacing β_{nA} by a commutator with $\hat{h}_2$, we obtain

$$(\beta_{nA}^k\hat{O})_{AnnA}=([\hat{h}_2,[\hat{h}_2,\ldots[\hat{h}_2,\hat{O}]\cdots]^k)_{AnnA}, \tag{24}$$

where $[\cdots]^k$ stands for k-fold commutator operation. Note that the same can be done in the $\mathbf{x}_1$-space, the only difference being that now we have $(-1)^k$ as a factor in the final expression.

Note that the momentum operator in the multicommutator expression (24) lowers the power of r_{12}. Let us examine the expression

$$F=\sum_n([\hat{h}_2[\hat{h}_2\cdots[\hat{h}_2,(1-\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2)r_{12}^{2k}]\cdots]^l)_{AnnA}. \tag{25}$$

Employing the fact that the Dirac spectrum is complete,

$$\sum_n\Psi_{n\alpha}^{+}(\mathbf{r}_1)\Psi_{n\gamma}(\mathbf{r}_2)=\delta(\mathbf{r}_1-\mathbf{r}_2)\delta_{\alpha\gamma} \tag{26}$$

(α and γ are spin indices), we can write (25) in the form

$$\begin{aligned}F=\int d^3r_1d^3r_2\Psi_A^{+}(\mathbf{r}_1)\delta(\mathbf{r}_1-\mathbf{r}_2)&\\\times\{[\hat{h}_2[\hat{h}_2[\cdots[\hat{h}_2,r_{12}^{2k}]\cdots]^l&\\-\delta_{12}\boldsymbol{\alpha}_1\cdot[\hat{h}_2[\hat{h}_2[\cdots[\hat{h}_2,\boldsymbol{\alpha}_2r_{12}^{2k}]\cdots]^l\}\Psi_A(\mathbf{r}_2).&\end{aligned} \tag{27}$$

In the latter expression the Kronecker delta δ_{12} refers to Dirac's α-matrixes. Only if $l\geqslant 2k$ is (27) not identically zero. Indeed, if after we write all the commutators in (27) explicitly there remains r_{12} raised to a power greater than zero, then $F=0$. This simple reasoning makes it possible to leave in (21) only terms in which $m=2k+1$.

6. EXPRESSION FOR THE SELF-ENERGY AND A SUM RULE

Simplifying (21) according to the rules formulated in Sec. 5 and plugging (21) and (16) into (14), we arrive at an expression similar to (6) which, however, is of the correct order in αZ in the nonrelativistic limit and has no divergences:

$$\begin{aligned}\Delta E_A=\frac{e^2}{\pi}\lim_{\Lambda\to\infty}\Biggl\{&\sum_n\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)!}[\psi(k+1)\\&-\psi(k)]\sum_n\beta_{nA}^{2k+1}([1-\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2]r_{12}^{2k})_{AnnA}\\&+2\left(\frac{1-\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2}{r_{12}}\left[\sin(\beta_{nA}r_{12})\mathrm{ci}(\beta_{nA}r_{12})\right.\right.\\&\left.\left.-\cos(\beta_{nA}r_{12})\left(\mathrm{si}(\beta_{nA}r_{12})+\frac{\pi}{2}\right)\right]\right)_{AnnA}\\&-2\sum_n\left(\frac{1-\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2}{r_{12}}\sin(\beta_{nA}r_{12})\ln\frac{\Lambda r_{12}}{2}\right)_{AnnA}\\&-\left(\frac{3}{4}\ln\,\Lambda^2+\frac{3}{8}\right)(\beta)_{AA}\Biggr\}.\end{aligned} \tag{28}$$

The finiteness of ΔE_A implies

$$\sum_n\left(\frac{1-\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2}{r_{12}}\sin(\beta_{nA}r_{12})\right)_{AnnA}=-\frac{3}{4}(\beta)_{AA}. \tag{29}$$

We call (29) a sum rule since it is similar to the nonrelativistic sum rules for transition probabilities. Note that the right-hand side of Eq. (29) is similar to the formula for the width of level A (see Ref. 13):

$$\mathrm{Im}(\Delta E_A) = -\frac{\Gamma_A}{2} = \frac{e^2}{\pi} \sum_n \left(\frac{1-\boldsymbol{\alpha}_1 \cdot \boldsymbol{\alpha}_2}{r_{12}} \sin(\beta_{nA} r_{12}) \right)_{AnnA}, \tag{30}$$

with $0 < E_n < E_A$.

We were unable to find an explicit analytic proof of the validity of the sum rule (29). In the Appendix we give a direct numerical proof of this rule for $Z=1$ and $A=1s_{1/2}$. In Sec. 5 we implicitly prove the sum rule (29) for $A=1s_{1/2}$ and an arbitrary value of Z.

Assuming that the sum rule (29) holds, we can derive an expression for the self-energy:

$$\begin{aligned}\Delta E_A = \frac{e^2}{\pi} \Bigg\{ & 2 \sum_n \Bigg(\frac{1-\boldsymbol{\alpha}_1 \cdot \boldsymbol{\alpha}_2}{r_{12}} \Bigg[\sin(\beta_{nA} r_{12}) \Bigg(\ln \frac{1}{2|\beta_{nA}|} \\ & + \sum_{k=1}^{\infty} \frac{(-1)^{k+1} (\beta_{nA} r_{12})^{2k}}{2k(2k)!} \Bigg) - \cos(\beta_{nA} r_{12}) \\ & \times \left(\mathrm{si}(\beta_{nA} r_{12}) + \frac{\pi}{2} \right) - \frac{1}{2} \\ & \times \sum_{k=1}^{\infty} \frac{(-1)^k (\beta_{nA} r_{12})^{2k+1}}{(2k+1)!} \sum_{p=1}^{k} \frac{1}{p} \Bigg] \Bigg)_{AnnA} \\ & - \frac{3}{8} (\beta)_{AA} \Bigg\}, \end{aligned} \tag{31}$$

where the ψ-functions are written explicitly.

The nonrelativistic expression for the contribution of the electron's self-energy to the Lamb shift at $A = ns_{1/2}$ has the form[12]

$$\begin{aligned}\Delta E_A^{NR} = & \frac{2\alpha}{3\pi} \sum_{m_{\mathrm{Sh}}} \left(r_{12}^2 \beta_{nA}^3 \ln \frac{1}{2|\beta_{nA}|} \right)_{AnnA} \\ & + \frac{e^3}{3\pi m^2} \frac{5}{6} \langle A | \Delta U | A \rangle \\ & + \frac{e}{4\pi m^2} \left\langle A \left| \boldsymbol{\sigma} \cdot \mathbf{l} \frac{1}{r} \frac{dU}{dr} \right| A \right\rangle, \end{aligned} \tag{32}$$

where summation is over the Schrödinger spectrum, $\boldsymbol{\sigma}$ are the Pauli matrices, and $\mathbf{l}$ is the angular momentum operator. The first term on the right-hand side of Eq. (32) is what is known as the Bethe logarithm.

The first term on the right-hand side of Eq. (31) in the nonrelativistic limit $\beta_{nA} r_{12} \ll 1$ becomes the Bethe logarithm. Note that actually the series expansion in powers of $\beta_{nA} r_{12}$ does not coincide with the expansion in powers in αZ, which means there is no way in which we can analytically obtain the terms in (32) by passing to the appropriate limit in (31). Such passage to the limit is done numerically in Sec. 7.

7. NUMERICAL CALCULATION OF THE SELF-ENERGY AND PROOF OF THE SUM RULE

Equation (31) can be written in the form

$$\Delta E_A = \sum_{k=0}^{\infty} (\alpha_k X_k + \gamma_k Y_k) - \frac{3e^2}{8\pi} (\beta)_{AA}, \tag{33}$$

where

$$X_k = \frac{2e^2}{\pi} \sum_n (\beta_{nA}^{2k+1} (1-\boldsymbol{\alpha}_1 \cdot \boldsymbol{\alpha}_2) r_{12}^{2k})_{AnnA}, \tag{34}$$

$$Y_k = \frac{2e^2}{\pi} \sum_n \left(\beta_{nA}^{2k+1} \ln \frac{1}{2|\beta_{nA}|} (1-\boldsymbol{\alpha}_1 \cdot \boldsymbol{\alpha}_2) r_{12}^{2k} \right)_{AnnA}, \tag{35}$$

and α_k and γ_k are numerical coefficients:

$$\begin{aligned}\alpha_k = & \sum_{p=0}^{k} \frac{(-1)^p}{(2p)!} \frac{(-1)^{k-p}}{(2k-2p+1)(2k-2p+1)!} \\ & + \sum_{p=0}^{k-1} \frac{(-1)^{p+1}}{(2p+1)!} \frac{(-1)^{k-p-1}}{(2k-2p)(2k-2p)!} \\ & - \frac{1}{2} \frac{(-1)^k}{(2k+1)!} \sum_{p=1}^{k} \frac{1}{p}, \end{aligned} \tag{36}$$

$$\gamma_k = \frac{(-1)^k}{(2k+1)!}. \tag{37}$$

The last term on the right-hand side of Eq. (33) can be written with allowance for the sum rule; after we expand the sine in a Taylor series, it becomes similar to the first term. The values of the coefficients γ_k remains unchanged, while the α_k change:

$$\Delta E_A = \sum_{k=0}^{\infty} (\tilde{\alpha}_k X_k + \tilde{\gamma}_k Y_k), \tag{38}$$

where

$$\tilde{\alpha}_k = \alpha_k + \frac{1}{2} \frac{(-1)^k}{(2k+1)!}, \tag{39}$$

$$\tilde{\gamma}_k = \gamma_k. \tag{40}$$

Table I lists the values of the coefficients $\tilde{\alpha}_k$ and $\tilde{\gamma}_k$, together with the values of X_k and Y_k calculated by formulas (34) and (35) for the state $A=1s_{1/2}$ for different values of Z via B-splines.[14,15] Splines were used for approximate calculations of sums over the Dirac spectrum. The last column contains the values of the self-energy taken from Ref. 11. Table I shows that the new method of calculating the self-energy is equally good for all values of Z.

In conclusion we note that the good agreement between our self-energy values and those obtained earlier implicitly proves the validity of the sum rule. Since at $Z=1$ the value of the last term on the right-hand side of Eq. (33) is five orders of magnitude larger than the self-energy and the results coincide to two decimal places, the sum rule is verified to seven significant figures. A similar result is achieved for

TABLE I. Numerical values of ΔE_A.

Z	k	$-X_k$, Ry	$-Y_k$, Ry	$\tilde{\alpha}_k$	$\tilde{\gamma}_k$	ΔE_A, Ry	
						Present work	Ref. 11
92	0	125.424	229.825	1.50000	1.00000	417.9614	
	1	303.501	514.988	-0.805556	-0.166667	87.64321	
	2	474.773	781.293	0.121111	8.33×10^{-3}	151.6543	
	3	14286.5	23335.9	-8.69×10^{-3}	-1.98×10^{-4}	22.74613	
	4	12379.0	23761.4	3.62×10^{-4}	2.76×10^{-6}	27.29876	
	5	18024.7	28277.2	-1.05×10^{-5}	-2.51×10^{-8}	27.10789	27.1072
80	0	86.5663	171.607	1.50000	1.00000	301.4567	
	1	218.195	381.149	-0.805556	-0.166667	62.16344	
	2	391.798	701.758	0.121111	8.33×10^{-3}	115.4626	
	3	11389.0	19947.5	-8.69×10^{-3}	-1.98×10^{-4}	12.43219	
	4	7844.25	15669.4	3.62×10^{-4}	2.76×10^{-6}	15.31876	
	5	8322.33	15750.7	-1.05×10^{-5}	-2.51×10^{-8}	15.23057	15.2315
70	0	46.3909	71.8354	1.50000	1.00000	141.4217	
	1	100.495	184.462	-0.805556	-0.166667	29.72343	
	2	140.618	241.298	0.121111	8.33×10^{-3}	48.76462	
	3	4542.50	7539.25	-8.69×10^{-3}	-1.98×10^{-4}	7.753638	
	4	4428.04	8238.59	3.62×10^{-4}	2.76×10^{-6}	9.381425	
	5	6842.79	10995.1	-1.05×10^{-5}	-2.51×10^{-8}	9.308960	9.3104
10	0	3.19×10^{-2}	5.363×10^{-2}	1.50000	1.00000	0.1014567	
	1	7.43×10^{-2}	0.122431	-0.805556	-0.166667	2.1163×10^{-2}	
	2	0.103819	0.207066	0.121111	8.33×10^{-3}	3.5462×10^{-2}	
	3	3.44458	5.37324	-8.69×10^{-3}	-1.98×10^{-4}	4.4321×10^{-3}	
	4	3.77247	6.94294	3.62×10^{-4}	2.76×10^{-6}	5.8187×10^{-3}	
	5	4.60871	6.92177	-1.05×10^{-5}	-2.51×10^{-8}	5.7699×10^{-3}	5.775×10^{-3}
1	0	8.54×10^{-6}	1.693×10^{-5}	1.50000	1.00000	2.9745×10^{-5}	
	1	2.17×10^{-5}	3.795×10^{-5}	-0.805556	-0.166667	5.9163×10^{-6}	
	2	2.81×10^{-5}	5.042×10^{-5}	0.121111	8.33×10^{-3}	9.7462×10^{-6}	
	3	9.76×10^{-4}	1.711×10^{-3}	-8.69×10^{-3}	-1.98×10^{-4}	9.1032×10^{-7}	
	4	1.03×10^{-3}	2.060×10^{-3}	3.62×10^{-4}	2.76×10^{-6}	1.2898×10^{-6}	
	5	1.37×10^{-3}	2.592×10^{-3}	-1.05×10^{-5}	-2.51×10^{-8}	1.2753×10^{-6}	1.297×10^{-6}

$Z=92$: $(\beta)_{AA}$ is two orders of magnitude larger, and from Table I we obtain four significant figures, which verifies the sum rule to six decimal places.

The authors are grateful to A. V. Nefedov, A. O. Mitrushchenkov, Yu. Yu. Dmitriev, T. A. Fedorova, G. Soff, H. Persson, I. Lindgren, P. Mohr, and W. Janchur for numerous discussions. This research was sponsored by the Russian Fund for Fundamental Research (Grant No. 96-02-17167). Part of the work was done during the visit of one of the authors (I.A.G.) to the Technical University of Dresden (Germany) under the REHE program.

APPENDIX: NUMERICAL PROOF OF THE SUM RULE

Here we prove Eq. (29) obtained for $Z=1$ and $A=1s_{1/2}$, which we do numerically by employing B-splines.[14,15]

We start by integrating over angles on the left-hand side of Eq. (29). Then we write the result as

$$S_A \equiv \sum_n \left(\frac{1-\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2}{r_{12}} \sin(\beta_{nA} r_{12}) \right)_{AnnA} = \sum_{l=0}^{\infty} \sum_{j=l\pm1/2} S_A^{(lj)}. \qquad \text{(A1)}$$

Next we take the state $A=1s_{1/2}$ of the hydrogen atom. In this case $(\beta)_{AA}\approx1.000$. The results of the calculations for $l=0,\ldots,4$ are listed in Table II.

Table II shows that the convergence of S_A in l is fairly poor. For this reason we did an asymptotic estimate of the series (A1) for $l>4$. To this end we first integrated the left-hand side of Eq. (29) over angular variables and to obtain

TABLE II. Numerical proof of the sum rule.

l	j	$S_A^{(lj)}$	$\Sigma^l_{l'} S_A^{(l'j)}$	$1/(\beta)_{AA}\,\Sigma^l_{l'}\Sigma_j S_A^{(l'j)}$
0	1/2	-0.607	-0.607	-0.607
1	1/2	0.045		
1	3/2	-0.135	-0.697	-0.697
2	3/2	0.017		
2	5/2	-0.048	-0.728	-0.728
3	5/2	0.007		
3	7/2	-0.015	-0.736	-0.736
4	7/2	0.001		
4	9/2	-0.001	-0.736	-0.736
$l\geqslant5$		-0.014	-0.750	-0.750

$$\begin{aligned}
\operatorname{Re}\Bigg[& \sum_{n,\kappa} F(\kappa,\kappa_A) \\
& \times \sum_{l=0}^{\infty} \int_0^{\infty} dr_1 \int_0^{\infty} dr_2 F_l \theta_l(r_1,r_2,\beta_{n\kappa,n_A\kappa_A}) \\
& \times [f_{n_A\kappa_A}(r_1) f_{n\kappa}(r_1) + g_{n_A\kappa_A}(r_1) g_{n\kappa}(r_1)] \\
& \times [f_{n_A\kappa_A}(r_2) f_{n\kappa}(r_2) + g_{n_A\kappa_A}(r_2) g_{n\kappa}(r_2)] \\
& + \theta_{l\pm 1}(r_1,r_2,\beta_{n\kappa,n_A\kappa_A})[F_{l_1} g_{n_A\kappa_A}(r_1) f_{n\kappa}(r_1) \\
& + F_{l_2} f_{n_A\kappa_A}(r_1) g_{n\kappa}(r_1)][F_{l_3} g_{n_A\kappa_A}(r_2) f_{n\kappa}(r_2) \\
& + F_{l_4} f_{n_A\kappa_A}(r_2) g_{n\kappa}(r_2)]\Bigg], \qquad \text{(A2)}
\end{aligned}$$

where $F, F_l, F_{l_1}, \dots F_{l_4}$ are numerical coefficients (some of which are given below), $f_{n\kappa}(r)$ and $g_{n\kappa}(r)$ are the large and small radial components of the relativistic wave functions $|n,\kappa\rangle$, $\kappa = \pm(j+1/2)$, and n is the principal quantum number. The functions θ_l have the form

$$\theta_l(r_1,r_2,\beta_{n\kappa,n_A\kappa_A}) = \begin{cases} j_l(\beta_{n\kappa,n_A\kappa_A} c r_1) h_l^{(1)}(\beta_{n\kappa,n_A\kappa_A} c r_2), \\ \qquad r_1 < r_2, \\ j_l(\beta_{n\kappa,n_A\kappa_A} c r_2) h_l^{(1)}(\beta_{n\kappa,n_A\kappa_A} c r_1), \\ \qquad r_1 > r_2, \end{cases} \qquad \text{(A3)}$$

where $j_l(z)$ and $h_l^{(1)}(z)$ are the spherical Bessel functions of the first and second kinds, respectively. All further calculations were done in the atomic system of units $(\hbar = e = m = 1)$, with c the speed of light. For $l \gg 1$ and $Z \sim 1$ the spherical Bessel functions have the asymptotic form

$$j_l \sim \left(\frac{eZ}{2l}\right)^l \left[1 + O\left(\frac{1}{l}\right)\right], \quad h_l^{(1)}(z) = j_{-l}(z).$$

At $Z=1$ the only component that is important is $f_{n_A\kappa_A}(r)$. Hence in (A2) we are interested only in the coefficient F, F_l, F_{l_1}, and F_{l_3}. For $l_n \gg 1$ we have

$$F \approx \sqrt{\frac{2}{2|\kappa|+1}}, \quad F_l \approx \frac{(-1)^{j_n}\beta_{nA}}{4j_n+2},$$

$$F_{l_1} \approx \frac{(-1)^{l_n}}{2|\kappa|+1}, \quad F_{l_3} \approx \frac{(-1)^{l_n}}{2|\kappa|+3}.$$

We consider summation over the discrete and continuous parts of the Dirac spectrum separately. For the discrete spectrum the radial components of the wave functions have the form[12] $(\alpha Z \ll 1)$

$$\begin{aligned}
\frac{1}{r} f_{n\kappa}(r) &\approx \frac{(2\lambda_\kappa)^{3/2+|\kappa|}}{(2|\kappa|)!} \sqrt{\frac{(n+|\kappa|)!(n+|\kappa|)}{2n(n-|\kappa|)!}}\, r^{|\kappa|} \\
&\quad \times \exp(-\lambda_\kappa r) \frac{|\kappa| + 2n_r - \kappa}{2|\kappa| + n_r},
\end{aligned}$$

$$\frac{1}{r} g_{n\kappa}(r) \approx 0,$$

where $\lambda_\kappa \approx Z/|\kappa| \ll 1$, and n_r is the radial quantum number. Equation (A2) can then be integrated with respect to r_1 and r_2. The result is a convergent series in n and l, which can be summed explicitly:

$$\sum_n \sum_{l=5}^{\infty} \left(\frac{1 - \boldsymbol{\alpha}_1 \cdot \boldsymbol{\alpha}_2}{r_{12}} I_{nA}^{\Lambda 1}(r_{12}) \right)_{AnnA} \approx 3.882 \times 10^{-2} \ \text{a.u.}$$

Both parts of the continuous spectrum were calculated in a similar manner. According to Ref. 12, for these parts we have $g_{E_-}(r) \simeq f_{E_+}(r)$, and for E_+ we have

$$\begin{aligned}
\frac{1}{r}[f_{E\kappa}(r)] &\approx 2^{3/2} \sqrt{\frac{mc^2+E}{E}} \exp\left(\pi \frac{\nu}{2}\right) \frac{|\Gamma(|\kappa| + i\nu)|}{\Gamma(2|\kappa|)} \\
&\quad \times (2pr)^{|\kappa|} \operatorname{Re}\{\exp[i(pr+\xi)] \\
&\quad \times F(|\kappa| - i\nu; 2|\kappa|; -2ipr)\},
\end{aligned}$$

where

$$p = \frac{1}{c}\sqrt{E^2 - (mc^2)^2}, \quad \nu = \frac{Z\alpha E}{p},$$

$$\begin{aligned}
\exp(-i\xi) &= \sqrt{\frac{|\kappa| - i\nu}{\kappa - imc^2\nu/E}} \\
&\approx \begin{cases} \kappa - (i\nu/2)(mc^2/E - 1), & \kappa > 0, \\ i\kappa + (\nu/2)(mc^2/E + 1), & \kappa < 0. \end{cases}
\end{aligned}$$

For the positive part of the continuous spectrum we are left with an expression with the coefficient F_{l_1} and for the negative part, with the coefficient F_{l_3}. After integrating with respect to r_1 and r_2 in (A2) we again arrive at a convergent series in n and l, which can also be summed explicitly:

$$\sum_n \sum_{l=5}^{\infty} \left(\frac{1 - \boldsymbol{\alpha}_1 \cdot \boldsymbol{\alpha}_2}{r_{12}} I_{nA}^{\Lambda 1}(r_{12}) \right)_{AnnA} \approx -5.317 \times 10^{-2} \ \text{a.u.}$$

[1] J. Schweppe, A. Belkacem, L. Blumenfeld, N. Claytor, B. Feinberg, H. Gould, V. E. Kostroun, L. Levy, S. Misawa, J. R. Mowat, and M. H. Prior, Phys. Rev. Lett. **66**, 1434 (1991).

[2] Th. Stöhlker, P. H. Mokler, K. Beckert, F. Bosch, H. Eickhoff, B. Franzke, M. Jung, Y. Kandler, O. Klepper, C. Kozhuharov, R. Moshammer, F. Nolden, H. Reich, P. Rymuza, P. Spüdtke, and M. Steck, Phys. Rev. Lett. **71**, 2184 (1993).

[3] H. F. Beyer, G. Menzel, D. Liesen *et al.*, Z. Phys. D **35**, 169 (1995).

[4] T. Beier and G. Soff, Z. Phys. D **8**, 129 (1988); S. M. Schneider, W. Greiner, and G. Soff, J. Phys. B **26**, L529 (1993).

[5] I. Lindgren, H. Persson, S. Salomonson *et al.*, J. Phys. B **26**, L503 (1993).

[6] A. Mitrushenkov, L. Labzowsky, I. Lindgren, H. Persson, and S. Salomonson, Phys. Lett. A **200**, 51 (1995).

[7] L. N. Labzowsky and A. O. Mitrushenkov, Phys. Rev. A **53**, 3029 (1996).

[8] P. J. Mohr, Ann. Phys. (N.Y.) **38**, 26 (1974).

[9] S. A. Blundell and N. L. Snyderman, Phys. Rev. A **44**, R1427 (1991).

[10] I. Lindgren, H. Persson, S. Salomonson, and A. Ynnerman, Phys. Rev. A **47**, R4555 (1993).

[11] P. J. Mohr, Phys. Rev. A **46**, 4421 (1992).

[12] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Systems*, Springer, Berlin (1958).

[13] L. N. Labzovskiĭ, in *Problems of the Quantum Theory of Atoms and Molecules* [in Russian], Issue 1, Leningrad State Univ. Press, Leningrad (1978), p. 52; L. Labzowsky, G. Klimchitskaja, and Yu. Dmitriev, *Relativistic Effects in the Spectra of Atomic Systems*, IOP Publ., Bristol (1993).

[14] W. R. Johnson, S. A. Blundell, and J. Sapirstein, Phys. Rev. A **37**, 307 (1988).

[15] C. deBoor, *Practical Handbook on Splines*, [Russian translation] Radio i Svyaz', Moscow (1984).

Translated by Eugene Yankovsky

Two-level electron dynamics in a strong variable external field

V. A. Burdov

N. I. Lobachevskiĭ Nizhniĭ Novgorod State University, 603600 Nizhniĭ Novgorod, Russia

(Submitted 24 February 1997)
Zh. Éksp. Teor. Fiz. **112**, 1209–1225 (October 1997)

This paper studies the quantum dynamics of an electron in a double-well potential subject to a strong time-periodic nonharmonic external field. The quasienergy spectrum of the system is calculated and an expression for the electron density distribution is derived. It is found that under certain conditions imposed on the shape of the excitation, the electron wave packet gets locked, into one potential well, as it were, and is unable to tunnel through the potential barrier.

1. INTRODUCTION

In recent years a large number of theoretical and experimental papers devoted to the study of various (including time-dependent) processes in quantum wells and heterostructures have been written. For instance, an extremely interesting effect was discovered theoretically[1–7]: when a periodic external field with a certain amplitude is applied to an electron in a symmetric double quantum well, the electron wave function is localized in one of the wells. This localization phenomenon is discussed in Refs. 1–3, 6 and 7, while the possibility of generating low-frequency dipole radiation in a system of quantum wells under conditions of localization is examined in Refs. 4 and 5.

Direct calculations of the wave function of the system and the system's dipole moment in an external variable field of arbitrary amplitude are extremely difficult, so that most results in Refs. 1–7 are obtained numerically, and only the approximations of a strong[2,6] or rapidly varying[4,5] field measurably simplify the problem. Furthermore, in all the aforementioned papers except Ref. 7, the solution is obtained only for a single initial condition, that amounts to electron localization in a well, which does not provide a complete and clear picture of the given phenomenon as a whole (in particular, the question of ''stability'' of such a solution remains unresolved).

Note that in all the cited papers only two types of time dependence of the external field are used: sinusoidal dependence,[1–6] and a symmetric sequence of rectangular pulses[6,7] (exceptions of a sort are Ref. 3, which employs a harmonic excitation modulated up to time $t=0$ by a slowly increasing exponential, and Ref. 5, where there is a constant external field in addition to a sinusoidal field).

The topic of the present paper is the quantum dynamics of an electron in a symmetric structure consisting of two quantum wells located in a time-periodic (but not necessarily harmonic) strong external field of frequency ω. Attention is focused on electron localization in one of the wells. In particular, we show that this effect is not only typical of a special (e.g., sinusoidal) time dependence of the external field, but can emerge over a much broader class of functions describing perturbations of a much more general form (at least at high field amplitudes). It is even unnecessary to specify the form of the potential energy of the unperturbed system, as in Refs. 1–3: we can limit ourselves to the two-level approximation, assuming that the distance between the two lowest levels, $E_1-E_0=\hbar\Delta$, is much smaller than the distance from this pair of levels to the next level E_2.

In our calculations we assume the external field to be so strong that the matrix element of the perturbation operator, V_{01}, is much larger than the characteristic energy scale of the problem, $\hbar\Delta$. At the same time, we must bound V_{01} from above, since we intend to use the two-level approximation. Combining these conditions yields

$$\hbar\Delta\ll V_{01}\ll E_2-E_1 .$$

Staying within these limits, we can obtain analytic results without resorting to numerical calculations.

It is well known that the most convenient way to describe the quantum evolution of a system in a periodic field is to invoke the formalism of quasienergies and quasienergy wave functions.[8–10] However, finding the quasienergy values is an extremely complicated problem, which cannot be solved in general form even for a two-level system. The strong-field approximation makes it possible to surmount these difficulties and find the quasienergies for any arbitrary time dependence of the perturbation.

2. HAMILTONIAN FORMALISM FOR A TWO-LEVEL SYSTEM

Since we are using the two-level approximation, it would seem natural to look for the wave function of a particle in a double-well potential and in the potential of a time-periodic force field $-F(t)$ in the form of an expansion in the basis consisting of the two stationary states $\chi_{0,1}(x)$ of the unperturbed system with the lowest possible energies $E_{0,1}$. It is clear, however, that the energy representation is unlikely to be the most convenient one, since the very idea of energy stationary states loses all meaning in variable fields that are so strong. Besides, below we are chiefly interested in the electron density distribution in the wells and the possible tunneling of charge through the separating barrier, while the level population dynamics is completely ignored.

For this reason, following Refs. 4–6, we introduce a new orthonormal basis $\Psi_{L,R}(x)$ constructed from the initial stationary state vectors

$$\Psi_L(x)=\frac{\chi_0(x)-\chi_1(x)}{\sqrt{2}},\quad \Psi_R(x)=\frac{\chi_0(x)+\chi_1(x)}{\sqrt{2}}.$$

Here we assume that the symmetric function $\chi_0(x)$ is always positive and the antisymmetric function $\chi_1(x)$ is positive when $x>0$. The basis functions $\Psi_L(x)$ and $\Psi_R(x)$ introduced in this manner are completely localized in the left and the right well, respectively. We can now write the wave function of the system as a linear combination of the vectors $\Psi_{L,R}(x)$ with time-dependent coefficients:

$$\Psi(x,\tau)=C_L(\tau)\exp\left(i\int \epsilon(\tau)d\tau\right)\Psi_L(x)$$
$$+C_R(\tau)\exp\left(-i\int \epsilon(\tau)d\tau\right)\Psi_R(x). \qquad (1)$$

Here we have introduced the dimensionless time variable $\tau=t\Delta$, and $\epsilon(\tau)$ is a periodic function with dimensionless period $T=2\pi\Delta/\omega$, and it depends on the ratio between the perturbation matrix element and the transition energy $\hbar\Delta$. The normalization condition imposed on the wave function implies that squaring the modulus of each of the coefficients $C_L(\tau)$ and $C_R(\tau)$ and adding the squares yields unity, with the square representing the probability of detecting an electron in the left or right well, respectively.

Choosing the origin of energy midway between the levels E_0 and E_1, we arrive at a system of equations describing the time evolution of the coefficient $C_{L,R}(\tau)$:

$$\begin{cases} i\dfrac{dC_L}{d\tau}=-\dfrac{C_R}{2}\exp\left(-2i\displaystyle\int \epsilon(\tau)d\tau\right), \\ i\dfrac{dC_R}{d\tau}=-\dfrac{C_L}{2}\exp\left(2i\displaystyle\int \epsilon(\tau)d\tau\right). \end{cases} \qquad (2)$$

Since the field is assumed strong, the characteristic values of $\epsilon(\tau)$ are much larger than unity.

We write Eqs. (2) in a more convenient form by introducing the notation

$$p=1-2|C_R(\tau)|^2,\quad q=\arg\{C_L(\tau)\}-\arg\{C_R(\tau)\},$$

where by ''arg'' we mean the phase of the complex number. This leads to a closed system of equations for p and q,

$$\begin{cases} \dfrac{dp}{d\tau}=\sqrt{1-p^2}\,\sin\left(q+2\displaystyle\int \epsilon(\tau)d\tau\right), \\ \dfrac{dq}{d\tau}=-\dfrac{p}{\sqrt{1-p^2}}\cos\left(q+2\displaystyle\int \epsilon(\tau)d\tau\right), \end{cases} \qquad (3)$$

which proves to be Hamiltonian, with the Hamiltonian function

$$H(p,q,\tau)=\sqrt{1-p^2}\,\cos\left(q+2\int \epsilon(\tau)d\tau\right).$$

Note that the Hamiltonian equations of motion (3) are not a full-fledged analog of the system (2): the solutions do not uniquely determine the coefficients $C_{L,R}(\tau)$ and hence the wave function $\Psi(x,\tau)$. However, the physically observable quantities, e.g., the various expectation values and probabilities, are fully determined by these solutions. In contrast to Eqs. (2), the Hamiltonian equations of motion are purely real, and the number of unknown variables in them is half the number of such variables in system (2), but the equations are nonlinear. This complicates matters somewhat, but at the same time the Hamiltonian form of these equations enables one to use the well-known and highly developed algorithms of the theory of Hamiltonian systems to solve them.

Obviously, the right-hand sides of Eq. (3) are rapidly oscillating functions of time, since the characteristic oscillation amplitude of $\epsilon(\tau)$ is much larger than unity. Hence by formally integrating Eqs. (3) with respect to τ we can easily see that the variations of the variables p and q within any finite time interval of order T are small.

Now let us solve the equations of motion (3). We must bear in mind that the function $\epsilon(\tau)$ consists of two parts: a constant part $\overline{\epsilon}$, which is the value of $\epsilon(\tau)$ averaged over one period, and a variable part $\widetilde{\epsilon}(\tau)$, whose average is already zero. Integration of the variable part $\widetilde{\epsilon}(\tau)$ yields a periodic function, which means we can expand the Hamiltonian $H(p,q,\tau)$ in Fourier series:

$$H(p,q,\tau)=\sqrt{1-p^2}\sum_{n=-\infty}^{\infty}\mu_n\,\cos(q+2\overline{\epsilon}\tau-n\Omega\tau+\psi_n), \qquad (4)$$

where $\Omega=2\pi/T$, and μ_n and ψ_n are, respectively, the coefficients and phases in the Fourier series, with

$$\mu_n\,\exp\{i\psi_n\}=\frac{1}{T}\int_0^T d\tau\,\exp\left\{i\left(n\Omega\tau+2\int d\tau\widetilde{\epsilon}(\tau)\right)\right\}. \qquad (5)$$

Below we derive explicit expressions for the coefficients μ_n and analyze their dependence on the perturbation amplitude.

The most interesting situation is the one in which the Hamiltonian (4) contains a slowly varying phase, which is possible if an integer l exists such that $\delta=2\overline{\epsilon}-l\Omega$ is much less than Ω. In this case the dynamics of the system proves to be resonant, and the external field is most effective.

Introducing the resonance phase $\theta=q+\tau\delta+\psi_l$, we try to find a canonical transformation that will destroy all nonresonant rapidly varying terms (with a frequency that is an integral multiple of Ω) in the Hamiltonian. The procedure, which formally amounts to averaging the Hamiltonian over rapid oscillations of frequency Ω, has come to be known as resonant perturbation theory.[11]

As a result of averaging, the Hamiltonian function of the system becomes

$$\mathscr{H}=p\delta+\mu_l\sqrt{1-p^2}\,\cos\theta,$$

and the equations of motion

$$\begin{cases} \dfrac{dp}{d\tau}=\mu_l\sqrt{1-p^2}\,\sin\theta, \\ \dfrac{d\theta}{d\tau}=\delta-\mu_l\dfrac{p}{\sqrt{1-p^2}}\cos\theta \end{cases} \qquad (6)$$

lead to the solution

$$p(\tau)=h\,\cos w+\sqrt{1-h^2}\,\sin(\sqrt{\delta^2+\mu_l^2}\tau-T_1)\sin w. \qquad (7)$$

Here we have introduced the notation

$$h=\frac{\mathscr{H}}{\sqrt{\delta^2+\mu_l^2}},\quad \cos w=\frac{\delta}{\sqrt{\delta^2+\mu_l^2}},$$

$$\sin w=\frac{\mu_l}{\sqrt{\delta^2+\mu_l^2}},$$

and T_1 is the constant of integration.

The solution (7) describes slow resonant oscillations of the momentum p in the vicinity of resonance. The slowness of the variation of p and θ is guaranteed by the smallness of the Fourier coefficients μ_l and the condition $\delta \ll \Omega$, whose validity has been assumed from the start. An increase in δ leads, on the one hand, to more rapid variations in p and θ; but on the other, these variations prove to be small, and in the limit $\sin w \to 0$ the momentum p becomes a constant of the motion.

3. QUASIENERGIES OF THE SYSTEM

We must now find the quasienergy spectrum of the system. According to Floquet's theorem, the solutions of the Schrödinger equation in a periodic external field with period T are functions with the property

$$U_\nu(x,\tau+T)=U_\nu(x,\tau)\exp\{-i\nu T\},$$

where the constant number ν is called the quasienergy. The Floquet functions form an orthonormal basis, and the wave function of the system can be expanded in the basis functions. As shown by Zel'dovich,[8] the expansion coefficients are time-independent. In our case of a two-level system, the number of quasienergies and quasienergy wave functions decreases to two, and it is easy to show that when the energy origin is midway between the levels, the quasienergies differ only in sign: $\nu_{1,2}=\pm\nu$. Indeed, if we assume that the system of equations (2) has a solution consisting of a pair of Bloch-type functions that correspond to the quasienergy ν,

$$C_L(\tau)\exp\left(i\int \epsilon(\tau)d\tau\right)=G(\tau)\exp(-i\nu\tau),$$

$$C_R(\tau)\exp\left(-i\int \epsilon(\tau)d\tau\right)=F(\tau)\exp(-i\nu\tau),$$

where $F(\tau)$ and $G(\tau)$ are periodic function with period T, then the other pair of linearly independent solutions of this system, as Eqs. (2) imply, consists of the functions

$$C_L(\tau)\exp\left(i\int \epsilon(\tau)d\tau\right)=F^*(\tau)\exp(i\nu\tau),$$

$$C_R(\tau)\exp\left(-i\int \epsilon(\tau)d\tau\right)=-G^*(\tau)\exp(i\nu\tau),$$

corresponding to the quasienergy $-\nu$ (the asterisk denotes complex conjugation).

Expanding the electron wave function $\Psi(x,\tau)$ in the quasienergy basis functions $U_\pm(x,\tau)$,

$$\Psi(x,\tau)=AU_+(x,\tau)+BU_-(x,\tau), \tag{8}$$

and comparing the expansion with (1), we find the general solution of the system of equations (2),

$$C_R(\tau)\exp\left(-i\int \epsilon(\tau)d\tau\right)=AF(\tau)\exp(-i\nu\tau)-BG^*(\tau)\exp(i\nu\tau),$$

$$C_L(\tau)\exp\left(i\int \epsilon(\tau)d\tau\right)=AG(\tau)\exp(-i\nu\tau)+BF^*(\tau)\exp(i\nu\tau),$$

and the expressions for the quasienergy functions:

$$U_+(x,\tau)=\exp(-i\nu\tau)[f(\tau)\exp\{i\varphi_F(\tau)\}\Psi_R(x)+g(\tau)\exp\{i\varphi_G(\tau)\}\Psi_L(x)],$$

$$U_-(x,\tau)=\exp(i\nu\tau)[f(\tau)\exp\{-i\varphi_F(\tau)\}\Psi_L(x)-g(\tau)\exp\{-i\varphi_G(\tau)\}\Psi_R(x)]. \tag{9}$$

The functions $f(\tau)$ and $g(\tau)$ introduced in (9) are the moduli and the $\varphi_{F,G}(\tau)$ are the phases of the complex functions $F(\tau)$ and $G(\tau)$, respectively. Now, according to the definition of $p(\tau)$, we look for

$$p(\tau)=1-2a^2f^2(\tau)-2b^2g^2(\tau)+4abf(\tau)g(\tau)\cos(2\nu\tau-\varphi(\tau)-\gamma), \tag{10}$$

where γ, a, and b denote the phase difference of the coefficients A and B and the moduli of these coefficients, and the periodic function $\varphi(\tau)$ consists of two terms, $\varphi(\tau)=\varphi_F(\tau)+\varphi_G(\tau)$.

Let us compare the expression (10) for the momentum p with the expression (7) obtained by solving the dynamical equations of motion. Obviously, pure quasienergy states correspond to the case in which either A or B vanishes (depending on the state that must remain) in the expression (8) for the wave function. In this case the squared moduli of the coefficients $C_{L,R}(\tau)$ must be strictly periodic in time with a period equal to that of the field. Naturally, the dynamical variable $p(\tau)$ defined earlier in terms of $C_R(\tau)$ also becomes a periodic function with the same period, a fact specified by (10).

On the other hand, the solution (7) obtained via resonant perturbation theory becomes a strictly periodic function (a constant in our case) if $h=\pm 1$. In the phase plane of the dynamical system (6), the points $h=\pm 1$ correspond to two stationary points, while any other value of h corresponds to slow periodic motion with a frequency that is the same for all phase trajectories and coincides, as Eq. (10) implies, with the quasienergy difference 2ν. Thus, we conclude that the values $h=\pm 1$ of the Hamiltonian correspond to pure quasienergy states of the quantum system.

Hence, by setting the coefficient B to zero in Eq. (10) and h to 1 in (7) (and then setting A to zero in (10) and h to -1 in (7)) we find $f(\tau)$ and $g(\tau)$, which determine the quasienergy functions $U_\pm(x,\tau)$:

$$f(\tau)=\sin\frac{w}{2},\quad g(\tau)=\cos\frac{w}{2}. \tag{11}$$

If we do not limit ourselves to pure quasienergy states and assume that A and B are arbitrary, then by comparing Eqs. (10) and (7) with allowance for the expressions we found for $f(\tau)$ and $g(t)$, we obtain

$$a^2=\frac{1+h}{2}, \quad b^2=\frac{1-h}{2}. \tag{12}$$

Obviously, according to the usual rules of quantum mechanics, the squared moduli of A and B can be interpreted as the probability of one of the two quasienergy values. Thus we find that the probability of a definite quasienergy value is specified by the dimensionless Hamiltonian h of the dynamical system (6), and since h is a constant of the motion, the phase trajectories of system (6) in the $\{p;\theta\}$ plane are the curves on which the quasienergy probabilities are constant.

We now determine the quasienergy values of the two-level system, to which end we go back to Eqs. (7) and (10). Direct comparison of the time dependence in the two expressions leads to

$$\nu_{\pm}=\pm\nu=\pm\frac{1}{2}\sqrt{\delta^2+\mu_l^2}. \tag{13}$$

As is known, quasienergies are defined to within the frequency of the external field Ω. Thus the resulting values (13) are quasienergies reduced to the ''first Brillouin zone.''

Let us discuss the dependence of the quasienergies on the parameters of the external perturbation. For this we will need the coefficients μ_n, which are defined in (5). First we note that only the variable part of the perturbation affects μ_n, while variations in the average value are only able to move the system away from resonance or closer to resonance, with the condition for resonance,

$$\delta=2\,\overline{\epsilon}-l\Omega=0, \tag{14}$$

having a clear physical meaning. The constant component of the perturbation moves the energy levels of stationary states apart by a distance of approximately $2\,\overline{\epsilon}$ (in strong fields). Hence to couple these levels in a resonant manner the system needs exactly the same energy as one quantum (or several quanta) of the external field.

In calculating the integral (5), to determine the Fourier coefficients μ_n we can use the method of stationary phase, provided that the equation

$$2\,\tilde{\epsilon}(\tau)+n\Omega=0 \tag{15}$$

has a solution. Moreover, the points of stationary phase must, as is known, be located far from one another (if there are several such points), so that integration near each point is done independently—in this case their contributions are additive.

For the sake of definiteness we assume that Eq. (15) has only two solutions for given n, $\tau=\tau_{1,2}(n)$, and consequently the derivatives $\dot{\epsilon}_{1,2}$ (of the function $\epsilon(\tau)$) at the points $\tau_{1,2}(n)$ have different signs. After doing the necessary calculations we arrive at

$$\mu_n^2=\frac{\Omega^2}{4\pi}\left(\frac{1}{\dot{\epsilon}_1}+\frac{1}{|\dot{\epsilon}_2|}+\frac{2}{\sqrt{|\dot{\epsilon}_1\dot{\epsilon}_2|}}\right.$$
$$\left.\times\sin\left\{\int_{\tau_1(n)}^{\tau_2(n)}d\tau(2\,\tilde{\epsilon}(\tau)+n\Omega)\right\}\right). \tag{16}$$

An estimate of the μ_n under the assumption that $\dot{\epsilon}_1\sim|\dot{\epsilon}_2|$ yields $\mu_n\sim\sqrt{\Omega/\epsilon_0}$, where ϵ_0 is the characteristic amplitude of variation of the function $\tilde{\epsilon}(\tau)$. If we assume that the perturbation frequency Ω is of order unity and $\epsilon_0\gg 1$, the coefficients μ_n prove to be small decreasing functions of ϵ_0. In addition we note that μ_n is an oscillating function of ϵ_0 because of the sine in (16).

If the number n is not large enough for Eq. (15) to become invalid, and yet its value is such that the points of stationary phase are too close to each other, the method of stationary phase used in calculating the coefficients μ_n must be slightly modified. Since here the two regions in which the phase changes only slowly begin to merge, the function $\tilde{\epsilon}(\tau)$ can be approximated in (5) by a parabola symmetric about the point τ_0 at which the two solutions of Eq. (15) merge, which results in the following expression for μ_n:

$$\mu_n=\frac{\Omega}{(\ddot{\epsilon}_0)^{1/3}}\,\mathrm{Ai}\left(\frac{2\,\tilde{\epsilon}(\tau_0)+n\Omega}{(\ddot{\epsilon}_0)^{1/3}}\right).$$

Here $\ddot{\epsilon}_0$ stands for the second derivative at $\tau=\tau_0$, and $\mathrm{Ai}(x)$ is the Airy function of argument x.

Note that this expression is also applicable to numbers n at which Eq. (15) has still no solution, but is close to becoming valid. In this case the argument of the Airy function is small and the value of the function is of order unity. Increasing the amplitude ϵ_0 of the external perturbation (i.e., actually the value of $|\tilde{\epsilon}(\tau_0)|$) for any preassigned value of n, we can ensure that while the argument of the Airy function has become negative it still continues to grow in absolute value. In this case μ_n proves to be an oscillating function of ϵ_0, with the size of oscillations slowly decreasing, and gradually the function is transformed into the one discussed above (see Eq. (16)).

If, on the other hand, we decrease the amplitude of variation of $\tilde{\epsilon}(\tau)$, the argument of the Airy function grows in the positive direction, which leads to rapid exponential decay of the coefficients μ_n as functions of ϵ_0. This situation corresponds to a complete absence of solutions of Eq. (15) and points of stationary phase in the integral (5). Note, however, that for all values of the argument of the Airy function, the coefficients μ_n prove to be small because of the small factor of the Airy function, which is equal in order of magnitude to $\epsilon_0^{-1/3}$.

Returning to the expression (13) for the quasienergies, it can be shown that the dependence of quasienergy on the amplitude ϵ_0 of the external field qualitatively follows the behavior of the coefficients μ_l and is an oscillating function if the integral (5) contains points of stationary phase. In particular, when the average value of the perturbation is identically zero or when the condition (14) is satisfied exactly, the quasienergies are completely determined by the value of μ_l.

In this case quasienergy resonance is possible, i.e., an integral number of external field quanta ''fit'' into the interval between two quasienergy levels, which corresponds to the crossing of the two quasienergy branches in the first Brillouin zone.

For instance, if the perturbation is purely harmonic, Eqs. (5) and (13) yield a value for the quasienergy difference equal to $J_0(2\epsilon_0/\Omega)$, where $J_0(x)$ is the zeroth-order Bessel function with argument x. Here quasienergy resonance occurs at values of ϵ_0 corresponding to the zeros of the Bessel function. Similar results are derived in Refs. 12 and 13, where the quasienergies are calculated with a low-frequency harmonic field in the WKB approximation.

When the integral (5) has no points of stationary phase, the quasienergies are essentially independent of ϵ_0, and only in an exponentially narrow range of values of $\overline{\epsilon}$ near $\overline{\epsilon}=l\Omega/2$ does the quasienergy also become a small exponentially decreasing function of ϵ_0, as are the coefficients μ_l, which, however, never vanish.

4. ELECTRON DENSITY DISTRIBUTION

Our analysis of the dynamical equations in Sec. 2 makes it possible to study the time evolution of the electron density in a system with two quantum wells. We start by finding the probability distribution density in the coordinate space at an arbitrary moment in time. Calculating the square of the modulus of the wave function (1) and going from the expansion coefficients $C_{L,R}(\tau)$ to the dynamical variables p and q, we arrive at the expression

$$\rho(x,\tau)=\frac{1+p(\tau)}{2}\Psi_L^2(x)+\frac{1-p(\tau)}{2}\Psi_R^2(x), \tag{17}$$

which implies that changes in the shape of the electron wave function are determined entirely by the dependence of p on τ. In deriving (17), we dropped terms containing the product $\Psi_L(x)\Psi_R(x)$, since as noted earlier, each of the two functions $\Psi_L(x)$ and $\Psi_R(x)$ is concentrated almost entirely in its own well, so that their overlap can be ignored. The estimate of $\Psi_L(x)\Psi_R(x)$ in Ref. 14 yields values of order $\sqrt{\hbar\Delta/E_2}\ll 1$, which supports the validity of the approximation used in (17).

In the general case, for arbitrary values of the Hamiltonian h and an angle w that varies from 0 to π, Eq. (17) describes smooth harmonic spatial oscillations of the electron density, which according to (7) take place between the two wells with a frequency equal to the difference 2ν of the quasienergy values.

We note here that there is a strong analogy between the behavior of a two-level system in a strong variable external field and the behavior in the absence of such a field, when there is also a flow of charge from well to well according to a harmonic law with unit frequency. This analogy is not accidental, and is due primarily to the fact that a periodic external field forces its own hierarchy of states on the quantum system, which leads to active replacement of stationary states by quasienergy states and of energy levels by quasienergy levels. As a result, the system's spectrum still consists of only two levels and the frequency of oscillations of the spatial dipole coincides with the problem's single characteristic frequency scale, the transition frequency between the quasienergy levels.

We now examine some interesting limiting cases in the evolution of the electron wave packet. In particular, we discuss the possibility of a static distribution of electron density in a double quantum well and of complete localization of the particle wave function in one of the wells.

First we assume that the parameters μ_l and δ do not vanish simultaneously, i.e., the quasienergy spectrum (13) is nondegenerate and the two quasienergy values $\nu_\pm$ correspond to the pair of quasienergy functions (9). To describe the function $p(\tau)$ we can then use the expression (7), which becomes a constant either when $\sin w\to 0$ or when $h\to\pm 1$. Let us examine in detail what happens in this case with the wave function (8) of the system and the quasienergy functions (9).

The parameters h and w are entirely independent, and while the value of the Hamiltonian h, which is uniquely determined by the initial condition, specifies according to (12) the weight of one of the two quasienergy states in the wave function $\Psi(x,\tau)$, the value of the angle w affects only the internal structure of the quasienergy functions, with no effect on the general structure of the expansion (8).

Let us assume that initially the wave function of the system coincides with one of the quasienergy functions, i.e., the value of h is either 1 or -1 and one of the two expansion coefficients in (8) vanishes. Since the coefficients in the expansion of the wave function in the quasienergy basis functions are time-independent, at all subsequent times the system will also be in a quasienergy state fixed by the initial condition, and described by the function $U_+(x,\tau)$ or $U_-(x,\tau)$.

Thus the square of the modulus of the wave function $\Psi(x,\tau)$, which determines the probability distribution density, coincides with the square of the modulus of one of the two functions $U_\pm(x,\tau)$ in (9). In accordance with (11), for the distribution density in a pure quasienergy state we have

$$\rho(x,\tau)=\Psi_{R,L}^2(x)\sin^2\frac{w}{2}+\Psi_{L,R}^2(x)\cos^2\frac{w}{2}, \tag{18}$$

where the first indices correspond to $U_+(x,\tau)$, and the second indices to $U_-(x,\tau)$.

From (18) it follows that the electron density distribution for a pure quasienergy state remains constant, while the functions $U_\pm(x,\tau)$ are time-dependent. Generally speaking, in the process of calculating the square of the modulus of the wave function (8), the expression for $\rho(x,\tau)$ can become time-dependent, but only in the mixed terms containing the product $\Psi_L(x)\Psi_R(x)$. As noted earlier, theses terms are extremely small, so that they can be dropped from (18).

Note that the filling factor of the left and right wells in a quantum state with definite quasienergy depends on $\sin^2(w/2)$ and $\cos^2(w/2)$, i.e., on the offset from resonance, δ, and the value of μ_l. Bearing in mind that the quasienergies (13) are specified by the same parameters, we can say that the filling factor of the wells is determined by one of these parameters (say δ) and the quasienergy ν, so that it differs for different quasienergy values within a single quasienergy zone.

We now discuss the other possibility of a static charge distribution forming in a double-well structure, i.e., we assume that the Fourier coefficient μ_l vanishes, and with it $\sin w$. If, in addition, the offset from resonance is positive, $\delta > 0$, then w is also zero. But if $\delta < 0$, then $w = \pi$. Generally speaking, μ_l can vanish only under certain conditions imposed on the shape of the function $\epsilon(\tau)$. We return to this question later. To ensure consistency of presentation, we now focus on the problem of finding the quasienergy functions.

Suppose that w is zero. Then, according to (11), $f(\tau) = 0$ and $g(\tau) = 1$, and the quasienergy functions $U_{\pm}(x,\tau)$ assume the form

$$U_{\pm}(x,\tau) = \Psi_{L,R}(x) \exp\left\{ \pm i \left(\varphi_G(\tau) - \frac{\delta}{2}\tau \right) \right\}, \tag{19}$$

where for $U_+(x,\tau)$ we must take the subscript ''L'' and for $U_-(x,\tau)$ the subscript ''R.'' Obviously, each of these two functions turns out to be localized entirely in its own well: $U_+(x,\tau)$ in the left well and $U_-(x,\tau)$ in the right. But if $w = \pi$, the function $f(\tau)$ becomes equal to unity while $g(\tau)$ vanishes. In this case, too, the quasienergy functions

$$U_{\pm}(x,\tau) = \Psi_{R,L}(x) \exp\left\{ \pm i \left(\varphi_F(\tau) + \frac{\delta}{2}\tau \right) \right\} \tag{20}$$

are completely localized, but now $U_+(x,\tau)$ is localized in the right well and $U_-(x,\tau)$ in the left.

Although in both cases the functions $U_{\pm}(x,\tau)$ also depend on time, the dependence has no effect on the density distribution $\rho(x,\tau)$. Squaring the modulus of the wave function (8), we obtain only the sum of the squares of the modulus of each term separately, since now, with $w = 0$ or π, the product of the quasienergy functions belonging to different quasienergies is again determined by the product $\Psi_L(x)\Psi_R(x)$, which is extremely small under conditions in which the two-level approximation is valid. As a result we again arrive at a static charge distribution,

$$\rho(x,\tau) = a^2 \Psi_{L,R}^2(x) + b^2 \Psi_{R,L}^2(x), \tag{21}$$

where the first subscripts correspond to $w = 0$ and the second to $w = \pi$. The coefficients a^2 and b^2 yield the filling factor of each of the two wells, and in turn are uniquely determined by the Hamiltonian h of the dynamical system (see (12)).

Either of the two cases of formation of a static charge distribution in a symmetric structure consisting of two quantum wells can serve as a starting point in solving the problem of existence of quantum states completely localized within one well. For this to occur we must be sure, as can easily be seen, that both conditions needed for the function $\rho(x,\tau)$ to be constant, i.e., $h = \pm 1$ and $\sin w = 0$, are met simultaneously, and that they can be applied in arbitrary order.

Indeed, let us set the Hamiltonian h either to 1 or to -1. The wave function then becomes one of the quasienergy functions $U_{\pm}(x,\tau)$. If we now require that μ_l vanish, which means that $\sin w$ also vanishes, we find, in accordance with (19) and (20), that the quasienergy functions are localized entirely in one of the wells. Thus, the wave function $\Psi(x,\tau)$ coincides to within a phase factor with one of the basis functions $\Psi_{L,R}(x)$, and the density distribution becomes a constant function entirely localized either in the left or right well:

$$\rho(x,\tau) = \Psi_{L,R}^2(x).$$

The subscript in this expression must be chosen in the following manner: the subscript ''L'' is chosen for $h = 1$ and $w = 0$ or $h = -1$ and $w = \pi$. Otherwise we must select ''R.''

Since, as noted earlier, each value of the Hamiltonian corresponds to a definite initial condition and the coefficients A and B in the expansion (8) do not change, we finally conclude that a static charge distribution localized entirely in one of the two wells can be obtained if we can initially build a wave packet corresponding to a pure quasienergy state. Here the coefficient μ_l must vanish, which corresponds to a certain form of the time dependence of the external field (as will be shown later) and furthermore, a definite value of the perturbation amplitude.

We now turn to the case in which the quasienergy spectrum (13) is degenerate, i.e., we assume that the two quasienergy values merge. This is possible only if the parameters δ and μ_l vanish simultaneously, which automatically nullifies the quasienergy values. Braun and Miroshnichenko[15] show that quasienergy branches can cross only if there is some sort of spatial symmetry in the problem, for example, what is known as generalized parity symmetry, for which the Schrödinger operator is invariant under the transformations $x \to -x$ and $\tau \to \tau + T/2$. Here the quasienergy branches corresponding to different parities cross if the system has no higher-order symmetries. Since we are discussing quantum dynamics under conditions of quasienergy resonance, we assume that the Hamiltonian possesses a symmetry that allows for quasienergy degeneracy.

In this case the quasienergy functions of the system are strictly periodic, with a period equal to that of the external field, T. Hence the dynamical variable $p(\tau)$ defined earlier in terms of $C_R(\tau)$ also becomes a periodic function with the same period. Since h, $\cos w$, and $\sin w$ have lost all meaning now, we should seek the solution $p(\tau)$ directly from the system of equations (6), where we must put $\mu_l = \delta = 0$. The solution, which proves to be a constant, $p(\tau) = p(0)$, must then be plugged into (17). As a result we find that when the quasienergy levels are degenerate, the density distribution also remains constant in time,

$$\rho(x,\tau) = \frac{1 + p(0)}{2} \Psi_L^2(x) + \frac{1 - p(0)}{2} \Psi_R^2(x), \tag{22}$$

and is determined entirely by the initial condition.

We would now like to establish the structure of the wave function in the case of quasienergy resonance. Zel'dovich[8] shows that when quasienergy branches cross, the choice of quasienergy functions becomes ambiguous, as illustrated by the following. Let us try to obtain the quasienergy functions by passing to the limits $\delta \to 0$ and $\mu_l \to 0$ in (9). It turns out that the final form of these functions then depends on the order in which these limits are taken, i.e., which of the two parameters δ and μ_l vanishes first.

If we require that the frequency of the external field land exactly in resonance (14), we find that $w=\pi/2$ and, in accordance with (11), that $f(\tau)=g(\tau)=1/\sqrt{2}$. We plug these values into Eqs. (9) for the quasienergy functions and set μ_l to zero. The result is

$$U_+(x,\tau)=\frac{1}{\sqrt{2}}[\Psi_R(x)\exp\{i\varphi_F(\tau)\}+\Psi_L(x)\exp\{i\varphi_G(\tau)\}],$$

$$U_-(x,\tau)=\frac{1}{\sqrt{2}}[\Psi_L(x)\exp\{-i\varphi_F(\tau)\}-\Psi_R(x)\exp\{-i\varphi_G(\tau)\}].$$

Each of the derived functions $U_\pm(x,\tau)$ yields a spatially symmetric distribution, with equal filling of the two wells.

Now let us change the order in which the limits are taken. Here $\mu_l=0$, and hence $\sin w=0$. The angle w is either 0 or π. If we set w to zero, which means that $\cos(w/2)=1$ and $\sin(w/2)=0$, we arrive at Eqs. (19) for $U_\pm(x,\tau)$ in which we must put $\delta=0$. If we set w to π, i.e., $\cos(w/2)=0$ and $\sin(w/2)=1$, we arrive at (20), where we must also put $\delta=0$. In this case the functions $U_\pm(x,\tau)$ prove to be localized entirely in only one well.

Note that as a result of the above manipulations with the passage to limits, we obtained three different pairs of quasienergy functions. These three sets of basis functions do not exhaust all possible sets, since there is an infinitude of ways in which, under conditions of degeneracy, we can obtain a pair of quasienergy functions that differs from all previous pairs. Thus, the choice of a pair of basis functions in the expansion (8) is really a question of taste. We note only that it is easy to go from one set of basis functions to another in (8) by redefining the expansion coefficients, and since any two basis vectors are quasienergy functions, the expansion coefficients always remain constant.

If for the quasienergy basis we take, for instance, the functions (19) with $\delta=0$, obtained by taking the limits $\mu_l=0$ and $w\to 0$, we can easily explain the localization of the electron wave function in one of the two wells, provided that for the initial condition we take a function that is localized entirely in a well, and thus essentially coincides with one of the quasienergy functions. This is the initial condition in Refs. 1–3, where the possibility of localization under conditions of quasienergy resonance is mentioned for the first time. For the sake of definiteness we assume that the initial condition has the form $\Psi(x,0)=\Psi_L(x)$. Obviously, the expansion coefficients in (8) are $A=\exp\{-i\varphi_G(0)\}$ and $B=0$, which yields the solution

$$\Psi(x,\tau)=\Psi_L(x)\exp\{i(\varphi_G(\tau)-\varphi_G(0))\},$$

since both A and B are constant. Naturally, such a wave function can be localized only in the left well, with the density distribution $\rho(x)$ equal to $\Psi_L^2(x)$ remaining unchanged.

In total contrast to all of the above cases, in which the function $\rho(x,\tau)$ is constant, is the situation corresponding to exact resonance (14) with equal populations of the quasienergy levels, i.e., $\delta=h=0$. Here we assume that the coefficient μ_l is nonzero, which automatically results in $\sin w$ being equal to unity. For simplicity we assume that both A and B in (8) are purely real and positive, an assumption that in no way affects the generality of the result.

In this case each of the two quasienergy functions (9) yields a symmetric coordinate distribution with a zero dipole moment, which is not true of the system's wave function $\Psi(x,\tau)$. Plugging $w=\pi/2$ and $A=B=1/\sqrt{2}$ into (11), (9), and (8) and grouping terms in the basis functions $\Psi_{L,R}(x)$, we obtain

$$\Psi(x,\tau)=\Psi_L(x)\exp\left[\frac{i}{2}\{\varphi_G(\tau)-\varphi_F(\tau)\}\right]\cos\frac{\mu_l\tau-\varphi(\tau)}{2}-i\Psi_R(x)\exp\left[\frac{i}{2}\{\varphi_F(\tau)-\varphi_G(\tau)\}\right]\times\sin\frac{\mu_l\tau-\varphi(\tau)}{2}. \tag{23}$$

If we were to assume that A and B are complex, we would have a common phase factor and constant phase shift, which can easily be removed by changing the point from which time is measured.

Obviously, the wave function (23) describes oscillations of the electron density with the maximum possible amplitude and with frequency $\mu_l/2$. Such a mode can be called the maximally time-dependent mode: the entire wave packet alternately shifts at an interval equal to half the period, from one well to the other, forming at these moments a distribution with a maximum dipole moment.

5. EFFECT OF THE TIME DEPENDENCE OF THE PERTURBATION ON THE EVOLUTION OF THE ELECTRON DENSITY

In this section we specify and discuss the conditions under which the Fourier coefficients μ_l can vanish. It turns out that the situation here depends largely on the relationship between the average perturbation $\overline{\epsilon}$ and the maximum value of the variable part $\tilde{\epsilon}(\tau)$, which earlier was denoted by ϵ_0. We start with the case in which the perturbation is a mostly positive function and $\overline{\epsilon}>\tilde{\epsilon}(\tau)$ for all values of τ.

Earlier we noted that the constant component of the perturbation, $\overline{\epsilon}$, moves the energy levels of the stationary states far apart, to a distance of approximately $2\overline{\epsilon}$. On the other hand, qualitatively we can think of the component $\tilde{\epsilon}(\tau)$ as a small perturbation, since it is much smaller than the level separation. It is well known, however, that even a small perturbation can be effective if its frequency (or the frequency of one of its harmonics) is close to the transition frequency between levels (see, e.g., the problem of resonant excitation in Ref. 16). In our case this corresponds to values of the parameter δ close to zero or at least comparable to μ_l, which yields values of $\sin w$ equal to unity or close to this value. The temporal variation of the wave function leads to a periodic variation of the level populations and filling of the potential wells, with the maximum variation attained at $h=0$ (see (23)).

Note that when the function $\epsilon(\tau)$ is positive at all values of τ, the Fourier coefficient μ_l is an exponentially small quantity (see Sec. 4). For this reason the size of the resonant region of δ, which coincides in order of magnitude with μ_l, also proves to be exponentially small. An increase in δ in comparison to μ_l takes us outside the resonant region, which drives $\sin w$ down, and in the limit $\sin w \ll 1$ leads to the onset of a charge distribution that is almost constant in time, since the lth perturbation harmonic is now unable to resonantly couple the two levels in the spectrum. Note that although the μ_l are extremely small, they never vanish, with the result that $\sin w$ is never exactly zero, and hence the expressions (19)–(21) for the quasienergy functions and the density distribution must be considered asymptotic expressions of some sort.

Now we go to the opposite case, in which the amplitude of the variable component of the external field exceeds the average value of the field: $\epsilon_0 > \overline{\epsilon}$. Here, naturally, even qualitatively we cannot assume, as we did earlier, that the variable part of the perturbation is small and that the action of the external field reduces to resonant excitation of the two-level system by a single lth harmonic. Under conditions in which the amplitude of the variable component is large, we must allow not only for single-quantum resonance processes, but also for multiquantum resonance processes, a situation automatically taken into account by the Fourier coefficient μ_l.

According to the inequality $\epsilon_0 > \overline{\epsilon}$, when the time is $\tau_1(l)$ or $\tau_2(l)$, the condition (15) with $n = l$ is met and the integral (5) contains no points of stationary phase. Now μ_l is specified by (16), which implies, in particular, that μ_l vanishes only if two conditions are met. First, the derivatives $\dot{\epsilon}_1$ and $\dot{\epsilon}_2$ must differ only in sign, remaining equal in absolute value, and second,

$$\int_{\tau_1(l)}^{\tau_2(l)} d\tau (2\tilde{\epsilon}(\tau) + l\Omega) = 2\pi\left(N - \frac{1}{4}\right), \tag{24}$$

where N is an arbitrary integer. In this case, the various resonant processes, both single-quantum and multiquantum, quench one another. Here the value of $\sin w$ is zero and the two-level system is not excited, which leads to the constant charge distribution (21) in the wells.

Obviously, the first condition determines the class of functions that describe the time dependence of the external field in which an electron can be localized in one of the wells. Assuming that $\dot{\epsilon}_1 = -\dot{\epsilon}_2$ at all values of the perturbation's amplitude and frequency, we conclude that the function $\epsilon(\tau)$ in this case is symmetric about $\tau = \tau_0$, where the two solutions of Eq. (15) merge, while the Schrödinger equation is invariant as $x \to -x$ and $\tau \to \tau + T/2$ (generalized parity).

The symmetry of $\epsilon(\tau)$ about τ_0 is a necessary but not sufficient condition for μ_l to vanish. For a static charge distribution to emerge, the perturbation amplitude ϵ_0 must have a certain value, which is determined by (24). Actually this second condition specifies the relationship between the external perturbation amplitude ϵ_0 and perturbation frequency Ω that makes μ_l vanish. For instance, when the perturbation is purely sinusoidal, ϵ_0 can be found as a function of Ω by solving the equation $J_0(2\epsilon_0/\Omega) = 0$ (see Ref. 2; this also follows from (24)), which leads to linear behavior of $\epsilon_0(\Omega)$ in the form $\epsilon_0 = x_i\Omega/2$, where the x_i are the roots of the zeroth-order Bessel function $J_0(x)$.

It is clear, however, that under certain conditions it is unimportant whether these requirements are rigorously met, since any small deviations from the conditions lead to equally small variations of the space and time dependence of the wave function and density distribution deviations that have no significant effect on the behavior of the system. For instance, when the system is far from quasienergy resonance, i.e., δ is nonzero, slight deviations of the perturbation from a given shape and amplitude lead to the emergence of a finite (but small in comparison to δ) coefficient μ_l, which has essentially no effect on the quasienergy functions (19) and (20) and the wave function (8), since $\sin w$ is much smaller than unity. In this sense, obviously, we can speak of the resulting solutions of the Schrödinger equation and the corresponding electron density configurations as being ''stable.''

But if the deviations from the specified conditions are significant, the value of $\sin w$ may approach unity. In this case the external field acts resonantly on the two-level system when δ is close to zero. Such action induces tunneling oscillations of the charge from one well to the other, with the maximum amplitude of such oscillations being reached when the populations of the quasienergy level become equal, in accordance with (23). This is precisely what happens in the case of quasienergy resonance, where no matter how small the emerging value of μ_l is, the value of $\sin w$ automatically becomes equal to unity, and the quasienergy functions uniformly fill the two wells. Obviously, the solution of the Schrödinger equation corresponding to quasienergy resonance is now ''unstable,'' in the sense that the smallest variation in the parameters of the system markedly alters the nature of the solution.

Thus, the ''stability'' of states localized in one of the wells depends on the extent to which the constant component of the external field distorts the double-well potential. If the offset from resonance δ is large compared to μ_l, localized states (built in this case from pure quasienergy states at values of the amplitude ϵ_0 of the external field corresponding to (24)) prove to be ''stable'' under small variations in the perturbation amplitude or initial condition. A decrease in δ leads to a narrowing of the stability region, and in the limit $\delta = 0$, which corresponds to quasienergy resonance, the localized states are completely destroyed by the slightest changes in ϵ_0.

In conclusion, we note once more that localization of an electron wave packet in one of the wells is possible if the time dependence of the periodic external field is described by a function symmetric with respect to any of its maxima or minima and if the external field is strong. Since there are no other restrictions on the type of the time dependence of the perturbation, we can say that the given phenomenon is of a fairly general nature.

[1] F. Grossmann, T. Dittrich, P. Jung, and P. Hänggi, Phys. Rev. Lett. **67**, 516 (1991); Z. Phys. B **84**, 315 (1991); F. Grossmann, T. Dittrich, and P. Hänggi, Physica B **175**, 293 (1991).

[2] F. Grossmann and P. Hänggi, Europhys. Lett. **18**, 571 (1992).

[3] R. Bavli and H. Metiu, Phys. Rev. Lett. **69**, 1986 (1992); Phys. Rev. A **47**, 3299 (1993).

[4] Y. Dakhnovskii and H. Metiu, Phys. Rev. A **48**, 2342 (1993); Y. Dakhnovskii and R. Bavli, Phys. Rev. B **48**, 11020 (1993).

[5] Y. Dakhnovskii and R. Bavli, Phys. Rev. B **48**, 11010 (1993).

[6] L. Wang and J. Shao, Phys. Rev. A **49**, R637 (1994).

[7] A. A. Gorbatsevich, V. V. Kapaev, and Yu. V. Kopaev, Zh. Éksp. Teor. Fiz **107**, 1320 (1995) [JETP **80**, 734 (1995)].

[8] Ya. B. Zel'dovich, Zh. Éksp. Teor. Fiz **51**, 1492 (1966) [Sov. Phys. JETP **24**, 1006 (1967)].

[9] V. I. Ritus, Zh. Éksp. Teor. Fiz **51**, 1544 (1966) [Sov. Phys. JETP **24**, 1041 (1967)].

[10] J. H. Shirley, Phys. Rev. B **138**, 979 (1965).

[11] A. J. Lichtenberg and M. A. Lieberman, *Regular and Stochastic Motion*, Springer, Berlin (1982).

[12] V. P. Kraĭnov and V. P. Yakovlev, Zh. Éksp. Teor. Fiz **78**, 2204 (1980) [Sov. Phys. JETP **51**, 1104 (1980)].

[13] I. Sh. Averbukh and N. F. Perel'man, Zh. Éksp. Teor. Fiz. **88**, 1131 (1985) [Sov. Phys. JETP **61**, 665 (1985)].

[14] V. A. Burdov, Teor. Mat. Fiz. **108**, 69 (1996).

[15] P. A. Braun and G. P. Miroshnichenko, Opt. Spektrosk. **49**, 1024 (1980).

[16] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-relativistic Theory*, 3rd ed., Pergamon Press, Oxford (1977).

Translated by Eugene Yankovsky

Extraction of information from the periodic system of the elements as an inverse problem

Yu. V. Tarbeev, N. N. Trunov, A. A. Lobashev, and V. V. Kukhar'

D. I. Mendeleev All-Russian Scientific-Research Institute of Metrology, 198005 St. Petersburg, Russia

(Submitted 12 March 1997)
Zh. Éksp. Teor. Fiz. **112**, 1226–1238 (October 1997)

The inverse problem of recovering an effective atomic potential that is universal for all elements of the periodic system is examined using a refined WKB method. The order of filling the shells is shown to be given by a linear combination of the principal and orbital quantum numbers with a coefficient that depends only on the asymptotic behavior of the potential. A relationship between the atomic numbers of the elements which open the shells and a certain functional (the strength) of the potential is found. Narrow ranges for the asymptotic index and strength of the potential are determined by comparison with data from the actual periodic system. The self consistency of this approach is demonstrated by good agreement between the calculated and actual atomic numbers of the elements. © *1997 American Institute of Physics.* [S1063-7761(97)00610-0]

1. INTRODUCTION

The periodic system of the elements discovered by D. I. Mendeleev in 1869 has been studied by physicists, chemists, and scientists in other specialties in its fundamental and applied aspects. Despite a variety of approaches, in reality two methods are used: either a phenomenological and/or semi-empirical systematization is carried out, for example, by choosing a combination of quantum numbers which describe the order of filling of the levels in most of the elements,[1] or the direct problem is solved, for example, by determining the electronic levels of the elements using a previously discovered form of the self consistent atomic potential. Two sources of difficulty are then inevitable. First, the potential itself is calculated on the basis of approximate quantum mechanical methods, e.g., by different versions of the Thomas–Fermi or Hartree–Fock methods. The resulting errors lead to ''exceptions to the rule,'' such as deviations from the empirical rule for filling of shells.[1] Second, studies of continuous variations, i.e., of the quantum numbers and level energies, etc., as functions of atomic number, will essentially require exact solutions of the Schrödinger equation for a given potential and this is a rare exception for realistic potentials.

Meanwhile, there is obvious interest in solving the inverse problem: beginning with the information contained in the modern periodic table, find the effective potential for many-electron atoms corresponding to the true interaction with the maximum possible accuracy and the other characteristics of the interactions. Evidently, here all the information in the periodic table must be regarded as ''regular'' and used to correct the approximations.

Like every somewhat complicated inverse problem, this problem becomes well posed and constructive only when additional limitations are imposed (on the allowable potential functions, etc.).

The purpose of this article is to extract information, such as limitations on the form of the effective one-electron potential, from the order of filling of the shells and the structure of the periodic table as a whole. Specifically, the following inverse problem is solved: we determine self-consistently a linear combination of the quantum numbers which controls the appearance and filling of new levels of outer electrons of all the elements with increasing atomic number and the effective potential itself, for a class of self-similar solutions such as the often used Tietz approximation,[2] but with an additional parameter.

This approach has a certain simplicity and constructability because of the following features: (1) exact solutions of the Schrödinger equation are not required and (2) the problem breaks up into two parts, i.e., the class of quantum numbers and the order of filling of the levels are uniquely determined by the asymptotic behavior of the potential, while the opening of new shells is controlled by a global parameter, the strength of the potential.

This paper is organized as follows: in Sec. 2 a new effective quantum number T, which depends on the asymptotic behavior of the potential and determines the order of filling of the levels with given quantum numbers, is introduced on the basis of a WKB analysis of the appearance of levels in a potential with a specified asymptotic behavior. In Sec. 3, limitations on the asymptotic behavior of the effective atomic potential are obtained by analyzing the behavior of the shells as the periodic system of the elements is filled in. In Sec. 4, parameters which completely characterize the potential in this problem are established, and a description is given of the classes of equivalent potentials with the same order of filling of the shells and the same atomic numbers for the elements at which an atomic shell with given quantum numbers n and l begins to fill up. Section 5 is devoted to determining the order of the elements for which electrons with a given orbital angular momentum first appear in a specific atomic model potential. This is a test of the self consistency of the approach developed here. The results are discussed and the directions for future studies are pointed out in the Conclusion.

Atomic units are used throughout this paper.

1063-7761/97/100666-07$10.00

2. WKB DETERMINATION OF THE QUANTUM NUMBER WHICH DETERMINES THE ORDER OF FILLING OF THE LEVELS

In order to solve the proposed inverse problem we use a WKB method for levels with zero binding energy in an effective (one-electron) atomic potential. In this section is sufficient just to know the asymptotic behavior of the effective potential $U(r,Q)$, where Q is the nuclear charge, which actually equals one of the integers Z that coincide with the order of an element in the periodic table. (In the intermediate calculations it is convenient to regard U as a continuous function of Q.)

We shall use a new, extremely effective quasiclassical quantization condition for levels with zero energy,[3,4]

$$\frac{1}{\pi}\int_0^\infty dr\sqrt{-2U(r,Q)}=N_l+\left(l+\frac{1}{2}\right)(\phi_0+\phi_\infty)-\frac{1}{2},\tag{1}$$

where the phases ϕ_0 and ϕ_∞ are determined by the asymptotic behavior of the potential $U(r,Q)$ at zero and at infinity,

$$r\to 0:\quad U\sim-\frac{1}{r^{1-\sigma}},\quad \sigma<1;\quad \phi_0=\frac{1}{1-\sigma};$$

$$r\to\infty:\quad U\sim-\frac{1}{r^{1+\nu}},\quad \nu>1;\quad \phi_\infty=\frac{1}{\nu-1},\tag{2}$$

with the Coulomb part explicitly separated from the asymptote. It is this value of ν which is meant for the asymptote throughout this paper.

We assume that the electrons move in a self consistent effective potential $U(r,Q)$ with a Coulomb asymptote for $r\to 0$ $(\sigma=0)$ and a power-law asymptote with the exponent ν (to be determined below) for $r\to\infty$.

As applied to our problem, the quasiclassical quantization condition (1) makes it possible to determine the times when new levels appear with increasing Q as a function of the power of the potential

$$\Phi(Q)=\frac{1}{\pi}\int_0^\infty\sqrt{-2U(r,Q)}dr\tag{3}$$

and its asymptotes (2). Note that $\Phi(Q)$ signifies the number of half waves in the WKB wave function, neglecting the centrifugal potential $W_l=l(l+1)/r^2$, while the additional phases ϕ_0 and ϕ_∞ account for both W_l and the exact form of the wave functions in the regions of very small and very large r (where the condition of quasiclassicality is not satisfied). Evidently, the N_lth level (in sequence) with orbital angular momentum l and binding energy $E=0$ appears for Z such that

$$\Phi(Z)=N_l+l\,\frac{\nu}{\nu-1}.\tag{4}$$

As Z and hence $\Phi(Z)$ increases, the binding energy of a level increases monotonically. (Throughout this paper we assume that the power of the potential $\Phi(Z)$ is a monotonically increasing function of the atomic number Z.)

Note that the first level with angular momentum l appears for the principal quantum number $n=l+1$; thus, $N_l=n-l$. Let us introduce n in Eq. (4):

$$\Phi(Z)=T(n,l,\nu)+\frac{t(\nu)}{2},\quad t(\nu)=\frac{1}{\nu-1}.\tag{5}$$

Here we have introduced a class of quantum numbers T which are linear in n and l and depend parametrically on ν:

$$T=n+lt(\nu).\tag{6}$$

The systematics of filling the levels in the periodic table is, therefore, completely determined by Eqs. (3), (5) and (6). One extremely convenient advantage of this approach is that the problem is separated into two steps. The actual order of filling the levels, which is determined by the increase in T, is uniquely fixed by the parameter ν or $t(\nu)$ and is independent of the specific form of $U(r,Z)$ provided that the asymptotes (2) do not change.

Let us consider (just in this paragraph) a potential with a noncoulomb asymptote at zero $(\sigma\neq 0)$ (2). Generalizing condition (1) to this case, we find that the order in which the levels are filled is determined only by the equivalency class of potentials with different specific forms and different pairs (σ,ν), but the same

$$t(\sigma,\nu)=\frac{1}{\nu-1}+\frac{\sigma}{1-\sigma}.$$

In our case, of course, $\sigma=0$ and $t(\nu)=t(0,\nu)$. It is useful to note, however, that the schemes for the appearance of the levels will be identical if $t(\nu)=t(\sigma_1,\nu_1)$. For example, they are the same for a potential which has a Coulomb form at the center and falls off exponentially at infinity $(\sigma=0,\ \nu=\infty)$ and for a potential which is bounded at zero $(\sigma=-1)$ and has an asymptote with $\nu=3$. This equivalency can be used to make a few estimates on choosing an exactly solvable potential; more complicated equivalency classes are examined in Sec. 4.

We now illustrate the accuracy of the proposed WKB method for a model potential which has been used to solve this same problem of the order of appearance of the levels. Exact solutions have been given[5,6] at zero energy for a specific potential of the form

$$U=-\frac{v}{r(1+r)^2}\tag{7}$$

(the simplest form of the solution occurs for $n-l=1$; see Ref. 4, p. 464). The appearance of the levels is controlled by the number $K=n+l$ (where n is the principal quantum number and l is the orbital quantum number), chosen empirically in Ref. 1, with

$$\sqrt{2v}=\sqrt{\left(K+\frac{1}{2}\right)^2-\frac{1}{4}}.\tag{8}$$

In the quasiclassical approximation for the potential (7), from our conditions (3) and (4) with $\nu=2$ it is easy to obtain

$$\sqrt{2v}=K+1/2.\tag{9}$$

It can be seen that even for $K \sim 1$ the relative difference is small and decreases rapidly with increasing K as $1/8(K+1/2)^2$. This is an estimate of the accuracy of the approach we have developed. In practice, it turns out to be higher.

A potential of the form (7) is the same, to within constant factors, as the Tietz approximation[2] for the self-consistent atomic potential in the Thomas-Fermi method, so that it was possible to treat,[5,6] with some reservations, the classification in terms of the number K as a consequence of the statistical model of the atom. It was proposed that the splitting of the levels in l be explained by a nonzero binding energy. Lacking exact solutions for arbitrary, slightly modified potentials, we have not been able to estimate the accuracy and stability of these results[5,6] before. The classification in terms of the number K has a number of disadvantages; in particular, the same value of K corresponds to the end of one period and the beginning of the next and for fixed K all the levels with different l appear to be degenerate, contrary to reality. In the classification with respect to the number T developed in the next section, these deficiencies are eliminated.

3. ANALYSIS OF THE FILLING OF SHELLS IN THE PERIODIC SYSTEM IN TERMS OF THE NUMBER *T* AND THE ASYMPTOTIC BEHAVIOR OF THE POTENTIAL

Let us use Eqs. (5) and (6) to analyze the information contained in the real periodic system. We begin with those consequences which require minimal information on the potential. As potential is specified in more detail, both the number of consequences and the accuracy of the estimates contained in them will increase.

The minimum requirements on the potential, which we assume to be satisfied everywhere, are the following:

(1) $\Phi(Z)$ increases strictly monotonically with increasing Z. Here shell or other effects can, in general, disrupt the monotonic rise in $U(r,Z)$ in isolated regions of an atom.

(2) The potential has power law asymptotes (2) with $\sigma=0$.

(3) With increasing Z the levels are filled in the same sequence in which they appear; in other words, a newly manifested level remains the shallowest until it is filled as Z increases. This assumption is natural in light of the smallness of E in real atoms and, thereby, the small departure in Z from the peak Z at which the level appears ($E=0$) and is filled. We are not interested in the possible competition between levels in filled shells as Z increases.

At this step, further specification of the potential is not required.

As noted above, the order of the appearance and, in view of condition 3, of the filling of the levels is completely determined by the parameter ν, i.e., by the number $T(n,l,\nu)$. Evidently, the values of Q (rounded off to the nearest integer Z) obtained by solving Eq. (5) for the given $T(n,l,\nu)$ determine when the (n,l) shells open up and are then filled as $T(n,l,\nu)$ increases, so that the correspondence of (n,l) and $T(n,l)$ one-to-one. By analyzing the real order of filling of the shells in the periodic system with increasing Z, we are able to find an interval of values of ν which will ensure the required ordering, i.e., a strictly monotonic rise in T as new shells are filled.

TABLE I. The change $T'-T$ in the quantum number $T(n,l,\nu)$ on going from the shell (n,l) to the next shell (n',l').

N	1		2				3			
(n,l)	1s		2s		2p		3s		3p	
$T'-T$		1		$1-\Delta$		Δ		$1-\Delta$		Δ

N	4						5					
(n,l)	4s		3d		4p		5s		4d		5p	
$T'-T$		$1-2\Delta$		Δ		Δ		$1-2\Delta$		Δ		Δ

N	6								7				
(n,l)	6s		4f		5d		6p		7s		5f		6d
$T'-T$		$1-3\Delta$		Δ		Δ		Δ		$1-3\Delta$		Δ	

The conditions for a rise in T on going from one shell to another (see Table I) require that the following inequalities be satisfied:

$$\Delta=\frac{\nu-2}{\nu-1}>0, \quad 1-k\Delta>0, \quad k=1,\,2,\,3. \tag{10}$$

This yields limits on Δ and ν,

$$0<\Delta<1/3, \quad 2<\nu<5/2. \tag{11}$$

The interval (11) can be narrowed by invoking specific additional information on the periodic system (see Sec. 5). It is natural to expect, as well, that the largest difference in the properties of the inert gases and alkali metals which begin a new period should be reflected in the magnitude of the gap between periods, in the component Δ. The requirement that Δ should exceed the other gaps from Eq. (10), $\Delta>1-k\Delta$, can be satisfied for $k=3$, which then yields

$$\frac{1}{4}<\Delta<\frac{1}{3}, \quad \frac{7}{3}<\nu<\frac{5}{2}. \tag{12}$$

The more stringent limitations on ν (12) are in full agreement with the result of Sec. 5 which follows from an analysis of the filling of the periodic system. (If we require the same inequality for $k=2$, we obtain the sole value $\Delta=1/3$, which probably exceeds the accuracy of the method; for $k=1$ this requirement is incompatible with $\Delta<1/3$.)

It is easy to see from Eqs. (6) and (11) and the real periodic system that the integral part of T specifies the number of the period in the periodic system, $N=[T]$. (See Table II as well. It shows values of $T(n,l,\nu)$ for the case of a specific model atomic potential.)

Besides their mutual ordering, the shells are characterized by a capacity $2(2l+1)$, i.e., the maximum number of electrons which can be placed in them, as well as by the total capacity

$$V(T)=2\sum_{n>l}\sum_{l}(2l+1), \quad n+t(\nu)l \leqslant T. \tag{13}$$

Here the sum is taken over all n and l which satisfy the inequality given here. (The values of T are not, in general,

TABLE II. The quantum number $T(n,l,\nu)$ and the element number $Z(n,l,\nu)$ as functions of $\nu_k=\{2.30,2.35,2.40,2.45\}$ calculated for the potential (22), together with the total capacity V of the shells and the corresponding elements of the periodic system with their atomic numbers.

N	l											
	s			f			d			p		
	T	Z	V	T	Z	V	T	Z	V	T	Z	V
1	1.00	0.5	0									
	1.00	0.5	2									
	1.00	0.5	H^{1}									
	1.00	0.5										
2	2.00	2.7	2							2.77	6.2	4
	2.00	2.7	4							2.74	6.2	10
	2.00	2.7	Li^{3}							2.71	6.1	B^{5}
	2.00	2.8								2.69	6.0	
3	3.00	7.4	10							3.77	14.3	12
	3.00	7.8	12							3.74	14.2	18
	3.00	7.9	Na^{11}							3.71	14.1	Al^{13}
	3.00	8.0								3.69	14.0	
4	4.00	16.8	18				4.54	23.8	20	4.77	27.3	30
	4.00	17.0	20				4.48	23.3	30	4.74	27.3	36
	4.00	17.3	K^{19}				4.43	22.9	Sc^{21}	4.71	27.2	Ga^{31}
	4.00	17.5					4.38	22.5	Cr^{24*}	4.69	27.2	
5	5.00	31.2	36				5.54	41.5	38	5.77	46.5	48
	5.00	31.6	38				5.48	40.9	48	5.74	46.6	54
	5.00	32.1	Rb^{37}				5.43	40.4	Y^{39}	5.71	46.7	In^{49}
	5.00	32.6					5.38	40.0	Nb^{41*}	5.69	46.8	
6	6.00	51.9	54	6.31	59.8	56	6.54	66.2	70	6.77	73.0	80
	6.00	52.8	56	6.22	58.5	70	6.48	65.7	80	6.74	73.4	86
	6.00	53.6	Cs^{55}	6.14	57.3	Ce^{58}*	6.43	65.2	La^{57}*	6.71	73.8	Tl^{81}
	6.00	54.4		6.07	56.2		6.38	64.8	Gd^{64}*	6.69	74.1	
7	7.00	80.3	86	7.31	90.8	88	7.54	99.2	102			
	7.00	81.8	88	7.22	89.4	102	7.48	98.8	112			
	7.00	83.1	Fr^{87}	7.14	88.0	Pa^{91}*	7.43	98.5	Cm^{96}*			
	7.00	84.4		7.07	86.8		7.38	98.2	Bk^{97}*			

integral.) Evidently, the atomic number of an element $Z(T)$ at which an electron with given (n,l) first appears, i.e., the (n,l) shell begins to be filled, and the capacity $V(T)$ must obey the relation

$$Z(T)\leqslant V(T), \tag{14}$$

which ensures an adequate number of sites for the electrons.

4. RELATIONSHIP BETWEEN THE POTENTIAL AND THE ATOMIC NUMBERS OF ELEMENTS FOR WHICH NEW SHELLS ARE OPENED

Extracting further information requires an examination of some general properties of the potential besides its asymptotic behavior. Here it is important that in Eq. (5) and the like, the potential only enters through a finite number of values of the power $\Phi(Q)$, i.e., through the WKB phase calculated for $Q=1, 2,\dots,Z_{\max}$ (where $Z_{\max}$ is the maximum number of an element in the periodic system). For brevity we write $Q\in\{Z\}$.

We anticipate the general analysis by a study of a particular class of self-similar potentials in which the dependence on position shows up only through the combination $x=Z^{\gamma}r$. The general form of these potentials is

$$-U(r,Z)=\frac{Z^{1+\gamma}}{x}\,\varphi(x), \tag{15}$$

where the screening function $\varphi(x)$ must satisfy the condition $\varphi(0)=1$.

The importance of this class of U originates, ultimately, in the following fact: for large T, on replacing the sum in Eq. (13) by an integral, we obtain the leading term of the total shell capacity as a function of T:

$$V(T) \sim T^3\left[1+O\left(\frac{1}{T}\right)\right]. \tag{16}$$

Since the energies of the next filled levels are small for all T, the deviation of Eq. (14) from equality should be small for all T, which requires the asymptotes

$$Z(T) \sim T^3\left[1+O\left(\frac{1}{T}\right)\right]$$

or, according to Eq. (5), for large Z,

$$\Phi \sim Z^{1/3}. \tag{17}$$

Substituting Eq. (15) in Eq. (3), we obtain

$$\Phi(Z)=Z^{(1-\gamma)/2}G[\varphi], \quad G[\varphi]=\frac{\sqrt{2}}{\pi}\int_0^\infty dx\sqrt{\frac{\varphi(x)}{x}}, \tag{18}$$

so that the self similar potential satisfies condition (17) for $\gamma=1/3$, and does this for all Z, which, however, imposes an excessively rigid limitation on the form of U at low Z. It is clear that the real potential is asymptotically close to the self similar one, but differs slightly from it at moderate and low Z. On setting $\gamma=1/3$, from Eqs. (5) and (18) we find, as functionals of φ, the atomic numbers $Z(n,l)$ of the elements at which electron levels with given (n,l) first appear:

$$Z(n,l)=\frac{1}{G[\varphi]^3}\left(T(n,l,\nu)+\frac{t(\nu)}{2}\right)^3. \tag{19}$$

It is clear from Eqs. (18) and (19) that the set of values $\{\Phi(Q)\}$, $Q\in\{Z\}$ for the self similar potentials is determined by a single parameter, the dimensionless power of the potential $G[\varphi]$ given by $G[\varphi]=\Phi(Z=1)$. Different $\varphi(x)$ can yield the same values of G. We shall refer to such potentials as equivalent.

The self-similar potential, therefore, is completely characterized by a pair of numbers (G,ν), where ν is the asymptotic index. All potentials can be divided into equivalence classes determined by points on the (G,ν) plane.

Potentials in general are completely characterized by a set of $2Z_{\max}$ numbers $\{\Phi(Q),\nu(Q)\}$, $Q\in\{Z\}$ and can be subdivided into equivalence classes within each of which fall all potentials with equal $\{\Phi(Q),\nu(Q)\}$, $Q\in\{Z\}$. Equation (19) generalizes to

$$Z(n,l)=\Phi^{-1}\left(T(n,l,\nu)+\frac{t(\nu)}{2}\right), \tag{20}$$

where the function Φ^{-1} inverse to Φ exists because of the strict monotonicity of Φ.

We now bring together the conditions imposed on the function Φ:

1. $\Phi(Z)$ is a monotonically increasing function, so that the natural relationship between the set $\{Z\}$ and the set $\{n,l\}$ of shells is preserved.

2. The following boundary conditions are satisfied: $Z=1$ maps onto the shell (1,0) and $Z=Z_{\max}$ maps onto the shell (6,2).

3. The function Φ is matched to the capacity of the shells and to the total capacity by Eq. (14).

The function $\Phi(Z)$ can be regarded as a mapping from the set $\{Z\}$ of elements into the set of shells $\{(n,l)\}$ consisting of 18 elements in the real periodic system. Since the solution of Eqs. (19) or (20) is rounded off to integers, the overall number of significantly different functions $\Phi(Z)$ is the same as the number of different sets of increasing integers $\{Z(1,0), Z(2,0), Z(2,1),\ldots,Z(6,2)\}$, with $Z(6,2) < Z_{\max}$. The actual number of allowed functions, however, is much smaller owing to the above mentioned conditions imposed on Φ by the real structure of the periodic system and by physical considerations. These questions will be investigated in a separate paper.

Note that the number $T(n,l)$ introduces an ordering relation in the set of shells: $(n,l)<(n',l')$ if and only if $T(n,l)<T(n',l')$.

5. SELF-CONSISTENT METHOD FOR DETERMINING THE PERIOD IN THE PERIODIC SYSTEM AND THE ATOMIC NUMBER OF AN ELEMENT WITH GIVEN *n*,*l*

We shall carry out a further analysis of the parameters of the potential and its consistency with the real periodic system by strengthening condition (3) of Sec. 3. Specifically, we identify the atomic numbers of the elements at which the new levels appear and are filled. First of all we find the limits on the possible values of G using a fact which follows from the periodic system: the element $Z_{\max}$ for a calculation using Eq. (19) with this value of G must fall into the (6,2) shell. Given the interval of possible values of ν (12) which follows from the order of filling of the shells, we obtain the interval of possible values of the right hand side of Eq. (5) for $n=6$, $l=2$. We set $Z_{\max}=104$ (the element kurchatovium, Ku^{104} with a $6d^27s^2$ electron configuration). Then Eqs. (5), (18), (19), and (12) yield the interval of possible values of G:

$$1.630<G<1.685.$$

Thus, the parameters G and ν of the effective self-consistent self-similar potential are

$$G=1.658\pm0.028, \quad \nu=2.42\pm0.08, \tag{21}$$

i.e., G is determined with a relative accuracy of $\pm1.7\%$ and ν with an accuracy of $\pm3.5\%$. (In the interval (11), $G=1.719\pm0.089$.)

We now move on to determining the $Z(n,l)$ according to Eq. (19) which correspond to the given $T(n,l,\nu)$ for the self-similar potentials (15) with $\gamma=1/3$. This problem, as we shall see, is a generalization and refinement of the old well known problem of determining the atomic number Z of an element for which an electron with given orbital angular momentum l first appears. (See Tietz,[2] as well as Abrahamson[8] and the references cited there.)

As a screening function we take

$$\varphi_\nu(x)=\frac{1}{(1+ax)^\nu} \tag{22}$$

with the two parameters a and ν. Recall that the standard Thomas-Fermi method[7,9] leads to a similarly self-similar (with $\gamma=1/3$) potential, while a function of the form (22) with $\nu=2$ (the Tietz approximation[7]) yields a fair approximation of the potential for small and moderate r. (For $r\to\infty$

the Thomas–Fermi potential behaves as a power law with $\nu = 3$. Our range of allowable values of ν (12) lies between the $\nu = 2$ and $\nu = 3$ characteristic of the Thomas–Fermi potential for different ranges of variation of the position.) Using Eq. (22) for all $0 \leqslant r < \infty$ with $\nu \neq 2$ implies an indirect accounting for the corrections to the standard Thomas–Fermi method.

Up to now we have not used any specific considerations from the statistical model of the atom, so that the discussion has been of an entirely general character. Now, however, we apply a normalization condition for φ to the total number of electrons which follows from the Poisson equation with an electron density calculated in the simple quasiclassical approximation.[7,9] In our variables (not including the conventional factor $b = 0.885$), this condition takes the form

$$\frac{8\sqrt{2}}{3\pi}\int_0^\infty dx\,\sqrt{x}\,\varphi^{3/2}(x) = 1. \tag{23}$$

The normalization condition (23) makes it possible to eliminate the parameter a, by expressing it explicitly in terms of ν. For functions $\varphi_\nu(x)$ of the form (22), the functional $G[\varphi]$ is easily calculated:

$$G[\varphi_\nu] = G(\nu) = \left(\frac{3}{4\pi^2}\right)^{1/3} \frac{B\left(\frac{1}{2}, \frac{1}{2}\nu - \frac{1}{2}\right)}{\left[B\left(\frac{3}{2}, \frac{3}{2}\nu - \frac{3}{2}\right)\right]^{1/3}}, \tag{24}$$

where $B(u,v)$ is the beta function.

We now calculate the atomic numbers $Z(n,l)$ of the elements for several values of ν in the case of the class of self-similar potentials (22) considered above. We set $\nu = \{2.30, 2.35, 2.40, 2.45\}$. The corresponding values of $G(\nu)$ given by Eq. (24) are $G = \{1.711, 1.698, 1.686, 1.674\}$. The pairs (G_k, ν_k) can be seen to lie in the (G, ν) plane inside a triangle bounded by the condition (11) and the corresponding condition for G: for $k = 3,4$ the pairs (G_k, ν_k) essentially satisfy the stronger condition (21). For each pair (ν_k, G_k) we use Eq. (19) to calculate the numbers $Z(n,l)$ of the elements for which an electron with the given n,l first appears for all the shells (n,l). Note that, according to the results of the preceding section, we are actually calculating $Z(n,l)$, not only for the model potential (22), but also for all potentials from the classes of equivalent self-similar potentials labelled by the given pairs (G_k, ν_k). The results of these calculations are shown in Table II (without rounding off to the nearest integer Z). The table was constructed as follows: the rows list the number N of the period in the periodic system and the columns, the quantum number l. Each cell lists the values of $T(n,l,\nu_k)$ and $Z(n,l,\nu_k)$ for the ν_k determined above (which ranges upward from $\nu_k = 2.3$ to $\nu_k = 2.45$), the total capacities V of the shells, and the symbols of the elements with their atomic numbers for which electrons with the given n, l first actually appear (in the periodic table). Also listed there are the elements which form an exception from the empirical rules[1] (denoted by asterisks).

We then discuss the calculated $Z(n,l)$. As can be seen from Table II, there is good agreement between the calculated and actual values of Z. The Z for which electrons with given l first appear are determined quite accurately. In addition, the atomic numbers of the elements in which the given l shows up later are also determined quite accurately.

Let us consider the atomic number of the elements in which d- and f-electrons appear. The calculated $Z(3,2)$ lie between 21 (Sc^{21}) and 24 (Cr^{24}*). The electronic configuration of Cr is $4s^1 3d^5$ and, therefore, it is an exception to the ideal empirical rule for the filling of shells. A similar pattern occurs in the 5th period, where there are a whole series of exceptions beginning with niobium Nb^{41} with $5s^1 3d^4$. For the f-elements, the calculated $Z(5,3)$ are close to 58, which is the atomic number of cerium Ce^{58} with $6s^2 5d^1 4f^1$, while $Z(6,3)$ is close to 91 (protactinium Pa^{91} with $5f^2 6d^1 7s^2$). In the 6- and 7th periods there are especially many misses (exceptions) in the order of filling the f- and d-shells. The calculated $Z(5,2)$ lies between 64 and 66. Gadolinium Gd^{64}, $4f^7 5d^1 6s^2$, with a d-electron is an exception in the lanthanide series; the elements before and after gadolinium do not have d-electrons. (Note that a d-electron first appears in the 6th period in La^{57}, which is also an exception.) A similar pattern occurs in the 7-th period: the $Z(6,2)$ are close to 98 or 99 and there are exceptions containing d-electrons in the actinide series, curium Cm^{96}, $5f^7 6d^1 7s^2$ and berkelium Bk^{97}, $5f^8 6d^1 7s^2$. This indicates that the potential is such that the number of d-electrons tends to increase in elements whose atomic numbers are close to the calculated values. Thus, we have also been able to explain the exceptions to the empirical filling rule.

The calculations of $Z(n,l,\nu)$ for $\nu < 2.3$ differ more from the empirical values of Z. These calculations agree completely with the conclusion of the previous section to the effect that the asymptotic index of the potential, ν, must be fairly close to 2.5 (12), while the power parameter of the self-similar potential lies within the interval (21).

The above analysis shows that our approach is self consistent and gives good results for the atomic numbers of elements with specified n and l. The exceptions to the empirical shell filling rule are also explained by it. This analysis places rather severe restrictions on the magnitude of the asymptotic index ν and the power parameter $G[\varphi]$ of the effective self-similar potential.

6. CONCLUSION

The basic results of the inverse problem examined above are the following: a quantum number $T(t)$, which is a linear combination of the quantum numbers of the centrally symmetric problem with a coefficient t lying within an extremely narrow interval, has been determined. It uniquely (without degeneracy) determines the appearance and filling of new levels for practically all the elements of the periodic system and ensures a substantial gap between periods. On leaving this interval, for example, on returning to the phenomenological number $K = n + l$, we find that the number of elements which fall out of the classification (the ''exceptions'') increases sharply. A unique relation has been established between the values of t and the asymptotic behavior of the effective potential for the electrons. These results depend only on the asymptotic behavior of the potential, but not on its specific form, and do not require the invocation of exact solutions.

The electron shells of atoms are represented as a set on which an ordering relation is introduced. By examining the relationships between the power of the effective potential $\Phi(Z)$ and the quantum number T, we have obtained strong limits on the parameters of the self-similar potentials.

Taking a simple generalization of the extremely good and widely applied Tietz approximation, which provides a necessary and already known asymptotic behavior, as an effective potential, we obtain an essentially exact correspondence between the calculated atomic numbers (elements) at which new levels appear and the numbers in the actual periodic system. It also provides a natural description of such subtle points as the exceptions to the empirical filling rules. Note that exact solutions of the Schrödinger equation are not required even at this stage. The above analysis shows that the real atomic potential is close to self-similar.

Here we point out some possible ways of extending this problem. By using a Thomas–Fermi–Dirac function including exchange, correlation, gradient, and other corrections, it is possible to search for an effective potential among the class of multiparameter functions by a variational technique. Then the given asymptotic behavior, which is closely related to the order of filling of the levels, i.e., to the real periodic system, and the power of the potential, which is related to the atomic numbers of the elements at which certain electron shells begin to fill, are preserved as additional conditions. These conditions, of course, affect the form of the solution. In a variational search in the class of self-similar potentials the imposition of limits on the power G of the potential leads to the corresponding isoparametric problem. For potentials of arbitrary form it is necessary to carry out an initial analysis of the restrictions on the function $\Phi(Z)$ similar to that carried out in this paper for the case of self-similar potentials. It seems probable that the allowable $\Phi(Z)$ will be close to the function (18) for self-similar potentials.

It should be kept in mind that all the corrections in the Thomas–Fermi–Dirac method[9–11] are admittedly approximations, while the requirements for the asymptotic behavior of the potential are essentially exact. We emphasize once again that the approach developed in this paper is valid to the same extent as the approximations used: the one-electron approximation for the effective atomic potential, the WKB approximation, and the approximation of zero energy for the last filled level.

The information contained in the self-consistent potential reduces to a finite set of parameters. This is indirect evidence of the high symmetry of the problem of constructing the real periodic system. Reducing all the characteristics of the periodic system to an algebraic structure makes it possible to begin a constructive search for the explicit operators and parameters of this symmetry. Group theoretical methods which currently seem extremely formal may then become suitable for application to the periodic system.[6,12,13]

It is appropriate to begin this kind of study with self-consistent potentials with extremely high symmetries; the number of parameters is just two instead of $2\cdot 10^2$ in the general case. Deviations from potentials of this type must be treated as a breakdown of symmetry.

The most symmetric problem on which the refinements could be based might be chosen differently: for example, an idealized version of the periodic system with zero energies for filled levels, in which $E\neq 0$ and splitting of the atomic numbers of the elements for which levels appear and the numbers at which they are filled are related as a consequence of a breakdown of symmetry.

In general, including levels with $E\neq 0$ in a WKB quantization scheme (5) makes it possible to invoke the energies of the ground and excited states as additional information.

[1] V. M. Klechkovskiĭ, *The Distribution of Atomic Electrons and the Filling Rule for the* $(n+j)$*-groups* [in Russian], Atomizdat, Moscow (1968).
[2] T. Tietz, J. Chem. Phys. **22**, 2094 (1954); Nuovo Cimento **1**, 955 (1955).
[3] M. S. Marinov and V. S. Popov, Zh. Éksp. Teor. Fiz. **67**, 1250 (1974) [Sov. Phys. JETP **40**, 621 (1974)].
[4] V. M. Galitskiĭ, B. M. Karnakov, and V. I. Kogan, *Problems in Quantum Mechanics*, Prentice-Hall, Englewood Cliffs, NJ (1963).
[5] Yu. N. Demkov and V. N. Ostrovskiĭ, Zh. Éksp. Teor. Fiz. **62**, 125 (1972) [Sov. Phys. JETP **35**, 66 (1972)].
[6] V. N. Ostrovsky, J. Phys. B **14**, 4425 (1981).
[7] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, Pergamon, Oxford (1977), Sections 70 and 73.
[8] A. A. Abrahamson, Phys. Rev. A **4**, 454 (1971).
[9] P. Gombás, *Die Statistische Theorie des Atoms und Ihre Anwendungen*, Springer-Verlag, Vienna (1949).
[10] D. A. Kirzhnits, Yu. E. Dozovik, and G. V. Shpatakovskaya, Usp. Fiz. Nauk **117**, 3 (1975) [Sov. Phys. Usp. **18**, 649 (1975)].
[11] D. A. Kirzhnits and G. V. Shpatakovskaya, Zh. Éksp. Teor. Fiz. **108**, 1238 (1995) [JETP **81**, 679 (1995)].
[12] O. Novaro and M. Berrondo, J. Phys. B **5**, 1104 (1972).
[13] T. Negadi and M. Kibler, J. Phys. A **25**, L157 (1992).

Translated by D. H. McNeill

Bifurcation in the rotational spectra of nonlinear symmetric triatomic molecules

I. N. Kozin and I. M. Pavlichenkov

Institute of Applied Physics, Russian Academy of Sciences, 603600 Nizhni Novgorod, Russia;
Kurchatov Institute Russian Scientific Center, 123182 Moscow, Russia

(Submitted 16 April 1997)
Zh. Éksp. Teor. Fiz. **112**, 1239–1256 (October 1997)

A microscopic theory is proposed for bifurcation in the rotational spectra of nonlinear AB_2-type molecules. The theory is based on a study of small-amplitude vibrational and precessional motion near the stationary states of a rotating molecule. Bifurcation leads to the formation of fourfold clusters of levels in the upper parts of the rotational multiplets, disruption of the symmetry of the molecule, and a transition from normal to local valence vibrations. The role of the centrifugal force of inertia in the development of these effects is clarified. Bifurcation and the accompanying phenomena are studied in the hydride molecules H_2O, H_2S, H_2Se, and H_2Te using empirical molecular potentials. *© 1997 American Institute of Physics.* [S1063-7761(97)00710-5]

1. INTRODUCTION

The cluster structure of molecular rotational spectra is an interesting phenomenon which has attracted the interest of researchers for over two decades.[1–4] In the past, however, attention has mainly been devoted to highly symmetric molecules of the spherical top type. One of the authors of the present paper, together with Zhilinskii,[5] predicted the possible formation of fourfold cluster states in the upper parts of the rotational multiplets of symmetric nonlinear triatomic molecules with the structure of an asymmetric top. The classical vibrational–rotational dynamics of these molecules has been studied using a model[6] of rigid valence bonds. It was shown that as the total angular momentum J of the molecule[1)] increases, local precession about the axis with the lowest moment of inertia evolves into delocalized precession about two equivalent axes lying in the plane of the molecule. The change in the rotational dynamics is caused by bifurcation at a critical point $J=J_c$. Bifurcation shows up as a softening of the precessional mode, i.e., the upper levels of the rotational multiplets come closer with increasing J. Prior to the prediction of this effect, experimental data showed a distinct convergence of levels in vibrational bands (ground state and ν_2) of the H_2O and H_2S molecules.[5,7] The observed effect was so unusual from the standpoint of existing concepts of rotation in asymmetric-top molecules, that immediately after the appearance of Ref. 5 numerical calculations of the rotational spectrum of water using a model of rigid valence bonds were published[8] which confirmed the level convergence effect.

The simplest theory of bifurcation proposed in Ref. 5 made it possible to estimate the critical angular momentum J_c.[9,7] It turned out that J_c decreases as the mass of the central core of the molecule increases. This result has been used in experimental studies of level clustering in the ground vibrational state of the H_2Se molecule.[10,11] It was found that, indeed, groups of four quasidegenerate levels (clusters) begin to form in the upper parts of the J-multiplets of this molecule for $J>J_c$. An analogous effect has recently been observed experimentally in H_2Te.[12] The MORBID method, which is based on numerical solutions of the Schrödinger equation, has been used successfully to describe cluster states in H_2Se,[13,14] H_2S,[15] and H_2Te.[16] The variational program MORBID[17] can be used to find the energy levels of the complete vibrational–rotational Hamiltonian of a molecule in an isolated electronic state. Calculations using realistic potentials, both *ab initio* and fitted to experimental data, reproduce well the observed energy levels of the rotational multiplets and can be used to follow the evolution of clusters in the ground vibrational state of the H_2Se, H_2S, and H_2Te molecules as J increases. These calculations agree with classical, semiclassical, and model quantum mechanical estimates of the magnitude of J_c. At the same time, numerical calculations of the rotational levels of the ground state band of the water molecule up to $J\sim 40$ using[18] the potential of Jensen *et al.*[19] and subsequently[20] with a more exact empirical potential have not revealed any distinct fourfold clusters. These results contradict estimates of the classical angular momentum for the H_2O molecule given in Refs. 5 ($J_c=27–28$) and 8 ($J_c=26$).

In this paper we eliminate the deficiency inherent in the model of absolutely rigid bonds. Bifurcation in the rotational spectrum of symmetric triatomic molecules is analyzed on the basis of the exact vibrational–rotational Hamiltonian obtained in Sec. 2. Classical equations are used for the vibrational–rotational motion of the molecules, since we are considering phenomena at large J and the de Broglie wavelengths of the nuclei are much shorter than the molecular dimensions. In Sec. 3 the stationary states of the system are found and in Secs. 4 and 5 the variations in the vibrational–rotational dynamics of the system are studied as functions of its total angular momentum. A method based on an examination of small harmonic oscillations (classical and quantum mechanical) near the stationary states of a rapidly rotating molecule is used.[21] The four-mode motion can be separated into a slow precession near the stationary rotation axis and fast vibrations near an equilibrium configuration that does not coincide with the configuration of the molecule in its ground state. This approach differs from the widely used Wilson–Decius–Cross theory[22] in that the vibrational and rotational degrees of freedom are separated near the stationary state of the rapidly rotating molecule. Because of centrifugal forces, this difference leads to some new phenomena: the precession of an AB_2 molecule about the axis with a minimum moment of inertia becomes unstable at the bifurcation point J_c. At high J, the molecule rotates uniformly about one of two equivalent (owing to the C_{2v} symmetry) axes located in the plane of the molecule between the principal axes with the minimum and intermediate moments of inertia. Quantum delocalization of the precession can lead to fourfold clusters in the rotational multiplets which are the spectroscopic manifestation of bifurcation. Another consequence is an asymmetric deformation of the molecule by the centrifugal forces of inertia. As a result, one of the A$-$B bonds becomes longer than the other, which in turn changes the vibrational dynamics of the molecule: the normal valence vibrations ν_1 and ν_3 change into local vibrations of the two A$-$B bonds. The transition discussed in Section 5 is in no way related to the anharmonicity of the vibrations, as happens in the generally accepted theory,[23] but is a consequence of the disruption of the symmetry of the molecule by the bifurcation. As applications of the theory we consider the hydrides H_2O, H_2S, H_2Se, and H_2Te, whose vibrational–rotational motion is still adiabatic compared to the electronic motion (despite the high J) and have energies lower than the dissociation energy of the molecules. In this series, the water molecule occupies a special place, demonstrating that level clustering is not, in general, a necessary consequence of bifurcation.

2. THE CLASSICAL VIBRATIONAL–ROTATIONAL HAMILTONIAN

Our derivation of the classical vibrational–rotational Hamiltonian is, on the whole, close to that of Wilson, Decius, and Cross.[22] The difference is that we do not use normal coordinates, since we are considering configurations which differ greatly from that of the molecular ground state. The kinetic energy of a molecule in the rotating coordinate system has the form

$$T=\frac{1}{2}\sum_{ij}\omega_i I_{ij}\omega_j+\sum_{i}\omega_i L_i+\frac{1}{2}\sum_{\lambda i} m_\lambda v_{\lambda i}^2, \tag{1}$$

where the subscripts i and j denote the x, y, and z axes of this system, ω_i are the projections of the angular velocity, and I_{ij} is the moment of inertia tensor of the molecule. The position of the nucleus λ with mass m_λ is given by the vector $\mathbf{r}_\lambda$ $(x_\lambda, y_\lambda, z_\lambda)$, its velocity is $\mathbf{v}_\lambda$, and $\mathbf{L}=\Sigma_\lambda m_\lambda \mathbf{r}_\lambda\times\mathbf{v}_\lambda$ is the angular momentum of the vibrational motion of the nuclei.

Let us consider a triatomic molecule $B_1A_2B_3$ with nuclear masses $m_1=m_3=m$ and $m_2=M$ and define a rotating coordinate system as follows:[24] place the molecule in the xz plane (i.e., $y_\lambda=0$) so that the bisector of the valence angle $\alpha=\widehat{B_1A_2B_3}$ is parallel to the x axis. The z axis is directed from nucleus B_3 to nucleus B_1, while the x axis is directed from the center of mass to nucleus A_2. In the following we also introduce the internal coordinates $q_1=r_1$ (the A_2-B_1 distance), $q_2=\alpha$ and $q_3=r_3$ (the A_2-B_3 distance). The internal coordinates are related to the projections of the vector $\mathbf{r}$ as follows:

$$\begin{aligned}
x_1&=-\frac{(M+m)r_1-mr_3}{M+2m}\cos\frac{\alpha}{2},\\
x_2&=\frac{m(r_1+r_3)}{M+2m}\cos\frac{\alpha}{2},\\
x_3&=\frac{mr_1-(M+m)r_3}{M+2m}\cos\frac{\alpha}{2},\\
z_1&=\frac{(M+m)r_1+mr_3}{M+2m}\sin\frac{\alpha}{2},\\
z_2&=-\frac{m(r_1-r_3)}{M+2m}\sin\frac{\alpha}{2},\\
z_3&=-\frac{mr_1+(M+m)r_3}{M+2m}\sin\frac{\alpha}{2}.
\end{aligned} \tag{2}$$

With this choice of a rotating coordinate system, the nonzero components of the inertia tensor are

$$\begin{aligned}
I_{xz}&=\frac{m(M+m)}{2(M+2m)}(r_1^2-r_3^2)\sin\alpha,\\
I_{xx}&=\frac{m}{M+2m}[M(r_1^2+r_3^2)+m(r_1+r_3)^2]\sin^2\frac{\alpha}{2},\\
I_{zz}&=\frac{m}{M+2m}[M(r_1^2+r_3^2)+m(r_1-r_3)^2]\cos^2\frac{\alpha}{2},\\
I_{yy}&=\frac{m(M+m)}{M+2m}(r_1^2+r_3^2)-\frac{2m^2}{M+2m}r_1r_3\cos\alpha.
\end{aligned} \tag{3}$$

We write the components of the vector $\mathbf{L}$ in the form $L_i=\Sigma_\nu G_{i\nu}\dot{q}_\nu$, where only three elements of the mixed matrix $\mathbf{G}$ are nonzero (the subscript ν refers to the internal coordinates):

$$\begin{aligned}
G_{y1}&=-\frac{m^2}{M+2m}r_3\sin\alpha,\\
G_{y2}&=\frac{m(M+m)}{2(M+2m)}(r_1^2-r_3^2),\\
G_{y3}&=\frac{m^2}{M+2m}r_1\sin\alpha.
\end{aligned} \tag{4}$$

The kinetic energy (1) in the new coordinates is

$$T=\frac{1}{2}\sum_{ij}\omega_i I_{ij}\omega_j+\sum_{i\nu}\omega_i G_{i\nu}\dot{q}_\nu+\frac{1}{2}\sum_{\nu\nu'}\dot{q}_\nu a_{\nu\nu'}\dot{q}_{\nu'}, \tag{5}$$

where the matrix $\mathbf{a}$ has the elements

$$\begin{aligned}
a_{11}&=a_{33}=\frac{m(M+m)}{M+2m},\\
a_{12}&=a_{21}=\frac{m^2}{2(M+2m)}r_3\sin\alpha,\\
a_{22}&=\frac{m}{4(M+2m)}[(M+m)(r_1^2+r_3^2)+2mr_1r_3\cos\alpha],
\end{aligned}$$

$$a_{23}=a_{32}=\frac{m^2}{2(M+2m)}\, r_1 \sin\alpha,$$

$$a_{13}=a_{31}=-\frac{m^2}{M+2m}\cos\alpha. \tag{6}$$

In order to obtain the equations of motion in Hamiltonian form, we introduce the total angular momentum **J** of the molecule, which is related to its angular velocity by the equation[22]

$$J_i=\sum_j I_{ij}\omega_j+L_i, \tag{7}$$

and the momenta

$$p_\nu=\frac{\partial T}{\partial \dot q_\nu}=\sum_\nu a_{\nu\nu'}\dot q_{\nu'}+\omega_y G_{y\nu}, \tag{8}$$

conjugate to the internal coordinates q_ν. It follows from the last equation that the Coriolis force in the rotating molecule (given by the second term in Eq. (5)) is analogous to a magnetic field, since $\mathbf{p}\neq 0$ holds when $\dot{\mathbf{q}}=0$. Solving Eq. (7) for ω_j, we find

$$\omega_j=\sum_i \mathring{\mu}_{ji}(J_i-L_i). \tag{9}$$

where $\mathring{\boldsymbol{\mu}}$ is the inverse matrix of **I** with elements

$$\begin{aligned}&\mathring{\mu}_{xy}=\mathring{\mu}_{yz}=0, \quad \mathring{\mu}_{yy}=1/I_{yy},\\ &\mathring{\mu}_{xx}=I_{zz}/(I_{zz}I_{xx}-I_{xz}^2), \qquad (10)\\ &\mathring{\mu}_{xz}=-I_{xz}/(I_{zz}I_{xx}-I_{xz}^2), \quad \mathring{\mu}_{zz}=I_{xx}/(I_{zz}I_{xx}-I_{xz}^2).\end{aligned}$$

Expressing $\dot{\mathbf{q}}$ in terms of **p** and **J** with the aid of Eqs. (8) and (9), we find the Hamiltonian of the rotating AB_2 molecule after some simple transformations:

$$H=\frac{1}{2}\sum_{ij} J_i\mu_{ij}J_j-J_y\sum_\nu u_{y\nu}p_\nu+\frac{1}{2}\sum_{\nu\nu'} p_\nu b_{\nu\nu'}p_{\nu'} +V(r_1,\alpha,r_3), \tag{11}$$

where $u_{y\nu}=\mathring{\mu}_{yy}\Sigma_{\nu'}G_{y\nu'}b_{\nu'\nu}$ and $\boldsymbol{\mu}$ is the matrix $\mathring{\boldsymbol{\mu}}$ modified by the Coriolis force. In our case, all the elements μ_{ij} are equal to $\mathring{\mu}_{ij}$ (10) except $\mu_{yy}=\mathring{\mu}_{yy}(1+\Sigma_\nu G_{y\nu}u_{y\nu})$. The matrix **b** is defined by the equation

$$\mathbf{b}=(\mathbf{a}-\mathbf{G}^T\mathring{\boldsymbol{\mu}}\mathbf{G})^{-1}. \tag{12}$$

where $\mathbf{G}^T$ is the transposed matrix. The elements of the matrix **b** are given by[22]

$$\begin{aligned}&b_{11}=b_{33}=\frac{1}{M}+\frac{1}{m}, \quad b_{12}=b_{21}=-\frac{\sin\alpha}{Mr_3},\\ &b_{22}=\left(\frac{1}{M}+\frac{1}{m}\right)\left(\frac{1}{r_1^2}+\frac{1}{r_3^2}\right)-\frac{2\cos\alpha}{Mr_1r_3}, \qquad (13)\\ &b_{23}=b_{32}=-\frac{\sin\alpha}{Mr_1}, \quad b_{13}=b_{31}=\frac{\cos\alpha}{M}.\end{aligned}$$

Finally, the last term in Eq. (11) is the potential V of the interaction of the nuclei in the molecule, which is a symmetric function with respect to interchange of identical nuclei.

The equations of motion for the Hamiltonian (11) have the form

$$\dot q_\nu=\sum_{\nu'} b_{\nu\nu'}p_{\nu'}-J_y u_{y\nu}, \tag{14}$$

$$\dot p_\nu=-\frac{1}{2}\sum_{ij} J_i\frac{\partial\mu_{ij}}{\partial q_\nu}J_j+J_y\sum_{\nu'}\frac{\partial u_{y\nu'}}{\partial q_\nu}p_{\nu'} -\frac{1}{2}\sum_{\nu'\nu''} p_{\nu'}\frac{\partial b_{\nu'\nu''}}{\partial q_\nu}p_{\nu''}-\frac{\partial V}{\partial q_\nu}. \tag{15}$$

and

$$\dot J_i=\sum_{jkl} e_{ijk}J_j\mu_{kl}J_l+\sum_{k\nu} e_{iyk}u_{y\nu}J_k p_\nu. \tag{16}$$

The last equation is obtained using the Poisson brackets $\{J_i,J_j\}=e_{ijk}J_k$, where e_{ijk} is the asymmetric unit tensor. It is easy to see that $\mathbf{J}^2$ is a constant of motion for Eqs. (14)–(16). In addition, the Hamiltonian and the equations of motion are invariant with respect to the $C_{2v}(M)$ group of the AB_2 molecule.[25,26]

3. STATIONARY STATES OF A ROTATING MOLECULE

First we shall find the stationary points of the equations of motion. Equating the time derivatives in Eqs. (14)–(16) to zero, we obtain a system of nonlinear algebraic equations which determine the configuration of the molecule and its axis of uniform rotation. In particular, Eq. (16) and the condition that $\mathbf{J}^2$ be conserved yield the stationary angular momentum $\mathbf{J}_s$ in the proper coordinate system. The three projections of this vector can satisfy the four equations if at least one of them is equal to zero. Thus, there are two types of stationary states, axial and planar. For small J only the axial states S_i exist: $J_{si}=\pm J$, $i=x, y, z$, in which the molecule rotates about one of the principal axes of inertia. Because the molecule is flat, the Coriolis force shows up only in the S_y state, which leads to a nonzero momentum with components

$$p_{\nu s}=J\sum_{\nu'} c_{\nu\nu'}u_{y\nu'}. \tag{17}$$

As a result, the additional term in the expression for μ_{yy} vanishes and the rotation of the molecule about the y axis is characterized by the moment of inertia I_{yy}. The equilibrium configuration of the molecule in state S_i is determined by the equations

$$\frac{1}{2}\frac{\partial\mathring{\mu}_{ii}}{\partial q_\nu}J^2+\frac{\partial V}{\partial q_\nu}=0, \quad \nu=1, 2, 3, \tag{18}$$

which are found from Eqs. (14) and (15). In the axial stationary state, the molecule has a symmetric configuration with bond lengths $r_{1s}=r_{3s}=r_s$, a valence angle α_s, and energies

$$E_i=\frac{J^2}{2I_{ii}(r_s,\alpha_s,r_s)}+V(r_s,\alpha_s,r_s). \tag{19}$$

TABLE I. The change $\Delta r_\nu = r_{\nu s} - r_{\nu e}$ in bond lengths owing to the centrifugal force in the S_{xz} stationary state for $J=20$ (it is assumed that the molecule rotates about an axis approximately perpendicular to the r_1 bond).

	Quantum mechanical calculation		Classical calculation	
Molecule	Δr_1,Å	Δr_3,Å	Δr_1,Å	Δr_3,Å
H_2Se	0.023 [14]	0.004 [14]	0.025	0.002
H_2S	0.024 [15]	0.009 [15]	0.022	0.010

This state is doubly degenerate in the direction of the total angular momentum. In the following we shall distinguish the equilibrium configurations of rotating and nonrotating molecules by the subscripts s and e.

We now consider the plane stationary state S_{xz} in which the molecule rotates around one of two equivalent axes, z' or z'', located in the xz plane symmetrically with respect to the x axis and forming angles β_s and $\pi - \beta_s$, respectively, with the z axis. The configuration of the molecule and the angle β_s are determined by the equations

$$\frac{1}{2}\left(\frac{\partial \mathring{\mu}_{xx}}{\partial q_\nu}\sin^2\beta_s + \frac{\partial \mathring{\mu}_{xz}}{\partial q_\nu}\sin 2\beta_s + \frac{\partial \mathring{\mu}_{zz}}{\partial q_\nu}\cos^2\beta_s\right)J^2 + \frac{\partial V}{\partial q_\nu} = 0, \quad \nu = 1, 2, 3. \tag{20}$$

and

$$\frac{1}{2}(\mathring{\mu}_{xx} - \mathring{\mu}_{zz})\sin 2\beta_s + \mathring{\mu}_{xz}\cos 2\beta_s = 0. \tag{21}$$

In the S_{xz} state the molecule has an asymmetric configuration with unequal bond lengths r_{1s} and r_{3s}. Of the two bonds, the one which forms the larger angle with the rotation axis is the longer (for example, r_{1s} for the z' axis). As J increases it is stretched by the centrifugal force and tends to be perpendicular to the axis of rotation. The symmetry disruption effect is illustrated in Table I, which lists the amounts of change in the bond lengths in the S_{xz} state relative to the equilibrium value for the nonrotating molecule. These results were obtained by classical and quantum mechanical calculations using empirical potentials for the H_2Se (Ref. 13) and H_2S (Ref. 15) molecules. The two methods appear to give approximately the same results.

Let us turn the coordinate system through an angle β around the y axis. Using Eq. (10) and noting that $I_{xx}I_{zz} - I_{xz}^2$ is invariant with respect to rotations, we find the relation between the matrix elements of the inverse moment of inertia in the initial (xyz) and rotated ($x'yz'$) systems to be

$$\begin{aligned}
\mathring{\mu}_{x'x'} &= \mathring{\mu}_{xx}\cos^2\beta + \mathring{\mu}_{zz}\sin^2\beta - \mathring{\mu}_{xz}\sin 2\beta,\\
\mathring{\mu}_{x'z'} &= -\frac{1}{2}(\mathring{\mu}_{xx} - \mathring{\mu}_{zz})\sin 2\beta - \mathring{\mu}_{xz}\cos 2\beta,\\
\mathring{\mu}_{z'z'} &= \mathring{\mu}_{xx}\sin^2\beta + \mathring{\mu}_{zz}\cos^2\beta + \mathring{\mu}_{xz}\sin 2\beta.
\end{aligned} \tag{22}$$

Comparing the expression for $\mathring{\mu}_{x'z'}$ with Eq. (21), we find that the z' axis (as well as the z'' axis) is the principal axis of inertia of an asymmetric molecule with a moment of inertia of

$$I_{z'z'} = \frac{1}{2}\left[I_{xx} + I_{zz} + \sqrt{(I_{xx} - I_{zz})^2 + 4I_{xz}^2}\right]. \tag{23}$$

Thus, the plane state S_{xz} with energy

$$E_{xz} = \frac{J^2}{2I_{z'z'}} + V(r_{1s}, \alpha_s, r_{3s}) \tag{24}$$

is fourfold degenerate. It has a lower symmetry than the axial state. This means that a transition from the axial sate S_z into the plane state with increasing J is accompanied by a C_{2v} type bifurcation.[7] Up to the bifurcation point J_c, a symmetric molecule rotates around the axis with the minimum moment of inertia I_{zz}. In the region $J > J_c$, an asymmetric molecule rotates around the principal axis with an intermediate moment of inertia $I_{z'z'}$. This ensures that the energy of a uniformly rotating molecule and its first derivative with respect to J are continuous at the critical point J_c. A change in the rotation regimes leads to an increase in the rotation energy in the S_{xz} state compared to the state S_z, i.e., to an increase in the energy of the upper level of the multiplet.

The equilibrium configuration of a uniformly rotating molecule is found according to Eqs. (18) and (19) from requiring the time-independence of the effective potential

$$V_{\text{eff}} = \frac{1}{2}\sum_{ij} J_i \mathring{\mu}_{ij} J_j + V. \tag{25}$$

On the other hand, the effective potential in the equilibrium configuration is equal to the energy of the corresponding stationary state, as this follows from Eqs. (19) and (24). The competition between the centrifugal and potential terms of Eq. (25) determines the change in the molecular configuration as J increases. Let us first examine the state S_z. Differentiating Eq. (18) with respect to J yields the following formulas:

$$\begin{aligned}
\frac{dr_1}{dJ} &= \frac{dr_3}{dJ} = \frac{J}{r_s I_{zz}}\left(\frac{\partial^2 V_{\text{eff}}}{\partial r_1^2}\right)_s^{-1},\\
\frac{d\alpha}{dJ} &= -\frac{J}{I_{zz}}\left(\frac{\partial^2 V_{\text{eff}}}{\partial \alpha^2}\right)_s^{-1}\tan\frac{\alpha_s}{2},
\end{aligned} \tag{26}$$

valid in the approximation

$$\left(\frac{\partial^2 V_{\text{eff}}}{\partial q_\nu^2}\right)_s \gg \left|\frac{\partial^2 V_{\text{eff}}}{\partial q_\nu \partial q_{\nu'}}\right|_s, \quad \nu \neq \nu', \tag{27}$$

which is satisfied for all the hydrides under consideration. Thus, in the stationary state S_z the bond lengths increase, while the valence angle decreases, with increasing J.

The approximation (27) is not applicable in the S_{xz} state and the change in the equilibrium configuration of the molecule as J increases depends on the features of its molecular potential. The rate of change of the coordinates of the equilibrium configuration is determined by the equation

$$\frac{dq_\nu}{dJ} = -\sum_{\nu'} g_{\nu\nu'}^{-1}\left(\frac{\partial \mathring{\mu}_{z'z'}}{\partial q_{\nu'}}\right)_s, \tag{28}$$

where **g** is the matrix of the second derivatives of V_{eff} at the stationary point. For rotation about the z' axis the bond length r_{1s} increases, while r_{3s} initially decreases and then

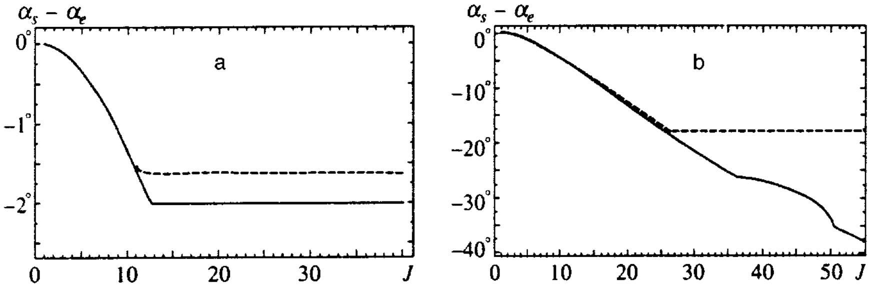

FIG. 1. The dependence on angular momentum J of the relative change in the valence angle in the stationary states S_z ($J<J_c$) and S_{xz} ($J>J_c$) corresponding to the upper level of the rotational multiplets of the ground state of the H_2Se (a) and H_2O (b) molecules. The smooth curves are calculations according to Eqs. (18), (20), and (21) ($J_c=13$ for H_2Se and 36 for H_2O); the dashed curves are from the absolutely rigid bond model ($J_c=12$ for H_2Se and 27 for H_2O).

slowly increases with rising J. For the heavy hydrides, however, the changes in the angle α_s and the bond r_{3s} are insignificant, i.e., these quantities are stabilized beyond the critical region. (See Fig. 1). In the rigid bond models,[5,9] the valence angle is independent of J for $J>J_c$ and equals its value at the critical point. It should be noted that the contribution of the potential energy to the overall energy of the molecule in the S_z and S_{xz} states is small. For the H_2S, H_2Se, and H_2Te molecules at the critical point, the rotational energy forms more than 95% of the total energy. With increasing J, the contribution of the rotational energy decreases. For the same values of J as for the heavy hydrides, the potential energy fraction for the water molecule and the lighter molecules is greater. This circumstance, together with the fact that the equilibrium angle α_e increases as the mass of the central nucleus is reduced, leads to higher values of the critical momentum J_c. For the water molecule the fraction of potential energy increases to 20% beyond the critical region. The stabilization is disrupted as a result: the angle α_s continues to decrease (see Fig. 1) and r_{3s} increases, while r_{1s} begins to decrease for $J>50$.

4. PRECESSIONAL MOTION OF THE MOLECULE NEAR STATIONARY STATES

The shift in the stationary rotation regimes of the AB_2 molecule with increasing J can be traced by studying the stability of the stationary states using the linearized equations (14)–(16) for small deviations in the internal coordinates $Q_\nu = q_\nu - q_{\nu s}$ (vibrations) and in the projections of the angular momentum $J_i - J_{is}$ (precession) from their stationary values. Let us begin with the stationary state S_y in which the vibrational motion is separated from the precession. Since $J_{xs}=J_{zs}=0$ holds in this state, while the momentum $p_{\nu s}$ is nonzero and given by Eq. (17), the precession equation has the form

$$\ddot{J}_x + J^2(\mathring{\mu}_{xx}-\mathring{\mu}_{yy})(\mathring{\mu}_{zz}-\mathring{\mu}_{yy})J_x = 0, \tag{29}$$

with the moments of inertia taken in the stationary state. Equation (29) coincides exactly with the precession equation for a solid asymmetric top.[27] The precession is stable, since its frequency

$$\Omega_y = J\sqrt{(\mathring{\mu}_{xx}-\mathring{\mu}_{yy})(\mathring{\mu}_{zz}-\mathring{\mu}_{yy})} \tag{30}$$

is real because $\mathring{\mu}_{xx}$ and $\mathring{\mu}_{zz}$ are always greater than $\mathring{\mu}_{yy}$.

The linearized equations of motion in the S_z state describe vibrational and precessional motion. Because of the symmetry of the molecule in this state, the equations break down into two independent subsystems. The coupled precession and asymmetric vibrations with coordinate $Q_a = (Q_1 - Q_3)/\sqrt{2}$ obey the equations

$$\begin{gathered}
\ddot{Q}_a + \left[(b_{11}-b_{13})f_{aa} + 2J^2 u_{y1}\frac{\partial\mathring{\mu}_{xz}}{\partial q_1}\right]Q_a + \sqrt{2}J \\
\times\left[(b_{11}-b_{13})\frac{\partial\mathring{\mu}_{xz}}{\partial q_1} - (\mathring{\mu}_{zz}-\mathring{\mu}_{xx})u_{y1}\right]J_x = 0, \\
\ddot{J}_x + J^2\left[(\mathring{\mu}_{zz}-\mathring{\mu}_{yy})(\mathring{\mu}_{zz}-\mathring{\mu}_{xx}) + 2u_{y1}\frac{\partial\mathring{\mu}_{xz}}{\partial q_1}\right]J_x \\
+\sqrt{2}J\left[u_{y1}f_{aa} - J^2(\mathring{\mu}_{zz}-\mathring{\mu}_{yy})\frac{\partial\mathring{\mu}_{xz}}{\partial q_1}\right]Q_a = 0.
\end{gathered} \tag{31}$$

We have introduced the force constant $f_{aa}=f_{11}-f_{13}$ for the asymmetric vibrations of the molecule using the matrix

$$f_{\nu\nu'} = \left(\frac{\partial^2 V_{\text{eff}}}{\partial q_\nu \partial q_{\nu'}}\right)_s . \tag{32}$$

The two remaining vibrational modes, which are independent of the rotation, will be examined below.

The precessional motion can be separated from the asymmetric vibrations in the adiabatic approximation $\omega_3 \gg \Omega_z$, which is valid up to the critical point J_c. In this approximation the precession frequency about the z axis equals

$$\Omega_z = J\sqrt{(\mathring{\mu}_{zz}-\mathring{\mu}_{yy})\left\{\mathring{\mu}_{zz}-\mathring{\mu}_{xx}+\frac{J^2}{f_{aa}}\left(\frac{\partial\mathring{\mu}_{xz}}{\partial r_a}\right)_s^2\right\}}, \tag{33}$$

where, according to Eqs. (3) and (10),

$$\mathring{\mu}_{zz}-\mathring{\mu}_{xx} = \frac{2(M+m)}{mMr_s^2\sin\alpha_s}\left(\frac{m}{M+m}-\cos\alpha_s\right). \tag{34}$$

For small J the last, nonadiabatic term in the curly brackets of Eq. (33) can be neglected. In this limit we obtain the precession frequency of a solid asymmetric top. It is real, since we are considering precession about the axis with the smallest moment of inertia and the angle α_e is greater than 90° for all the hydrides being considered. With increasing J, as we have seen, the stationary angle α_s decreases, $\mathring{\mu}_{zz}$ approaches $\mathring{\mu}_{xx}$, and the precession frequency Ω_z goes to zero when

$$\mathring{\mu}_{zz}-\mathring{\mu}_{xx}+\frac{J_c^2}{f_{aa}}\left(\frac{\partial\mathring{\mu}_{xz}}{\partial r_a}\right)_s^2 = 0. \tag{35}$$

Equations (35) and (18) determine the critical angular momentum J_c beyond which the stationary state S_z becomes unstable. The small nonadiabatic term in Eq. (35), which accounts for the deformability of the bonds, becomes important when the difference $\mathring{\mu}_{zz}-\mathring{\mu}_{xx}$ is small. Since it is positive, the critical angular momentum is greater, while the critical valence angle is less than the corresponding quantities in the rigid bond models. The critical angle α_c in the latter case is given by

$$\alpha_c=\cos^{-1}\left(\frac{m}{M+m}\right). \tag{36}$$

Note that the nonadiabatic term in Eq. (35) originates exclusively in the centrifugal inertia force. The Coriolis force, which changes the precession frequency by an amount of order $(\Omega_z/\omega_3)^2$, has no effect on the critical angular momentum J_c, which is explained by the planar configuration of the molecule.

The higher the asymmetric mode frequency ν_3 is, the closer J_c is to the value found[9] in the approximation of absolutely rigid bonds. Numerical calculations of the critical angular momentum for these hydrides using realistic potentials[13,15,16] are only slightly different from the previously determined values of J_c. We found $J_c=9.3$ for H_2Te, $J_c=12.5$ for H_2Se, and $J_c=18.9$ for H_2S. The corresponding values in the rigid bond models are 8.5, 11.4, and 16.9. It appears that the lighter the molecule, the greater the difference between these numbers. This difference is greatest for the water molecule. The potential of Ref. 19 gives $J_c=35.2$, which is much higher than the $J_c=26.5$ found in the approximation of rigid bonds.

For $J>J_c$ the stationary state S_z becomes a saddle point on the energy surface. It determines the maximum depth of the valley (a classically inaccessible region) separating the two symmetric maxima S_{xz} which develop as a result of the bifurcation. Near one of the maxima, the linearized equations of motion have a simple form in the coordinate system rotated about the y axis by an angle β_s (or $\pi-\beta_s$). The Hamiltonian and the equations of motion in this system are given by Eqs. (11) and (14)–(16) if the projection of the vector $\mathbf{J}$ and the elements of the matrix $\mathring{\mu}$ are related to the new primed axes. In this system, the total angular momentum vector of the state S_{xz} is directed along the z'. The linearized equations have the form

$$\begin{gathered}
\ddot{Q}_\nu+\sum_{\nu'}\left[\sum_{\nu''} b_{\nu\nu''}f_{\nu''\nu'}+J^2u_{y\nu}\left(\frac{\partial\mathring{\mu}_{x'z'}}{\partial q_{\nu'}}\right)_s\right]Q_{\nu'}\\
+J\left[\sum_{\nu'} b_{\nu\nu'}\left(\frac{\partial\mathring{\mu}_{x'z'}}{\partial q_{\nu'}}\right)_s+(\mathring{\mu}_{x'x'}-\mathring{\mu}_{z'z'})u_{y\nu}\right]J_{x'}=0,\\
\ddot{J}_{x'}+J^2\Bigg[(\mathring{\mu}_{z'z'}-\mathring{\mu}_{yy})(\mathring{\mu}_{z'z'}-\mathring{\mu}_{x'x'})\\
+\sum_{\nu} u_{y\nu}\left(\frac{\partial\mathring{\mu}_{x'z'}}{\partial q_\nu}\right)_s\Bigg]J_{x'}+J\sum_{\nu}\Bigg[\sum_{\nu'} u_{y\nu'}f_{\nu'\nu}\\
-J^2(\mathring{\mu}_{z'z'}-\mathring{\mu}_{yy})\left(\frac{\partial\mathring{\mu}_{x'z'}}{\partial q_\nu}\right)_s\Bigg]Q_\nu=0.
\end{gathered} \tag{37}$$

where the matrix elements $\mathring{\mu}_{i'j'}$ are given by Eqs. (22). This system of equations coincides in form with the linearized equations for the S_z with the sole difference that the three vibrational and precessional motions are not separated owing to the asymmetry of the molecule. Note that the force constant matrix $f_{\nu\nu'}$ differs from the matrix $g_{\nu\nu'}$ of Eq. (28). The second derivatives of the effective potential in the first case are calculated for a fixed angle β_s. These quantities are related by

$$\begin{gathered}
g_{\nu\nu'}=f_{\nu\nu'}+2[(\mathring{\mu}_{z'z'}-\mathring{\mu}_{x'x'})^2\\
+4\mathring{\mu}^2_{x'z'}]^{-1/2}\left(\frac{\partial\mathring{\mu}_{x'z'}}{\partial q_\nu}\right)_{\beta_s}\left(\frac{\partial\mathring{\mu}_{x'z'}}{\partial q_{\nu'}}\right)_{\beta_s},
\end{gathered} \tag{38}$$

where the subscript β_s indicates that the derivatives are taken for fixed values of this angle. Using the adiabatic approximation, we find the precession frequency for not too large J:

$$\Omega_{xz}=J\sqrt{(\mathring{\mu}_{z'z'}-\mathring{\mu}_{yy})\left[\mathring{\mu}_{z'z'}-\mathring{\mu}_{x'x'}+J^2\sum_{\nu\nu'}\left(\frac{\partial\mathring{\mu}_{x'z'}}{\partial q_\nu}\right)_s f^{-1}_{\nu\nu'}\left(\frac{\partial\mathring{\mu}_{x'z'}}{\partial q_{\nu'}}\right)_s\right]}. \tag{39}$$

It is easy to show that the frequency (39) goes to zero at the critical point J_c. Equation (39) is formally similar to Eq. (33) for the precession frequency in the S_z state, but the term in square brackets, which is proportional to J^2, is not small. It ensures that the frequency Ω_{xz} is real. At the critical point J_c this frequency goes to zero, since the expression in the square brackets of Eq. (39) transforms into the left hand side of Eq. (35).

The precessional–vibrational motion near the stationary state S_x obeys equations similar to Eq. (31) for the S_z state. Precession about this axis is unstable.

We now turn to the structure of the rotational multiplet levels. Equations (19) and (24) can be used to estimate with good accuracy the energies of the lower and upper levels in the multiplet for all J, including those in the $S_z\rightarrow S_{xz}$ transition region. The difference from experimental values or variational calculations for the H_2O, H_2S, H_2Se, and H_2Te molecules is less than 10%. The separatrix passing through the saddle point corresponding to the S_x state separates the region of precessional motion localized about the y or z axes and, therefore, determines the formation of K-doublets in the lower and upper parts of the multiplets.[3] Yet another separa-

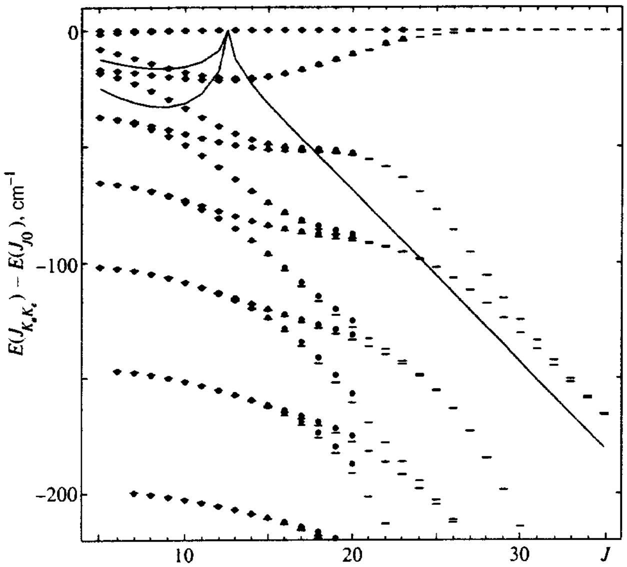

FIG. 2. Structure of the rotational levels of the upper parts of the J-multiplets of the ground vibrational state of the $H_2\,^{80}Se$ molecule. The multiplet levels are determined by the approximate quantum numbers for the projection of the angular momentum on the axes with the minimum (K_a) and maximum (K_c) moments of inertia. The energy of a level is taken relative to the energy of the upper level of the multiplet ($K_a = J$, $K_c = 0$) and equals the experimental (points) or theoretical (dashes) value from Ref. 14. The smooth curves are a harmonic approximation with frequencies (33) and (39).

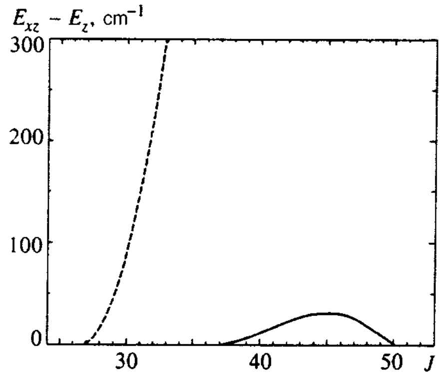

FIG. 3. The depth of the valley separating the two equivalent maxima of the S_{xz} state on the rotational energy surface of the water molecule calculated using Eqs. (19) and (24) (smooth curve) and in the approximation of absolutely rigid bonds (dashed curve).

trix of the S_z saddle arises as a result of the bifurcation. It separates the regions of precessional motion about the S_{xz} maxima and leads to clustering of the K-doublets. The depth of the valley between these maxima equals the energy difference between the stationary states S_{xz} and S_z. As J rises, this quantity increases in the H_2S, H_2Se, and H_2Te molecules and this shows up as a reduction in the splitting of the clustered levels. This behavior can be seen in Fig. 2, which shows the energy of the upper multiplet levels of the H_2Se molecule as a function of J. The energies given there have been calculated with the aid of the MORBID variational program[17] and the harmonic approximation of Eqs. (33) and (39) using the empirical potential from Ref. 13. The harmonic approximation clearly reproduces the qualitative behavior of the upper levels. It is not, however, applicable near the critical point and cannot describe level clustering.

The water molecule is a special case among the hydrides considered here. Calculations for this molecule with a new optimized potential from Ref. 19 show that the depth of the valley separating the S_{xz} maxima grows very slowly as the total angular momentum is increased. Figure 3 illustrates the qualitative difference between the models. While the valley rapidly becomes deeper with increasing J in the model of absolutely rigid bonds, in the calculations using Eqs. (20) and (21) its depth reaches a maximum at $J = 45$ and then begins to decrease. The valley vanishes at $J = 52$. This is a second critical point J_c' in the rotational spectrum where the two S_{xz} peaks merge into one S_z and for $J > J_c'$ the molecule is again rotating about the z axis. The two bifurcations are clearly visible in Fig. 1 from the characteristic breaks in the $\alpha_s(J)$ curve. The second critical point is too close to the first for a significant valley to form. Its maximum depth is comparable to the precession frequency. Thus, cluster levels similar to the levels in the heavy hydrides do not form in the water molecule. This provides a physical explanation for some numerical calculations[20] of the rotational spectrum of water. It is desirable to determine the molecular potential used there up to energies of order 10000 cm^{-1}, which corresponds to the potential energy at the point J_c'. Only then can we speak of the behavior of this molecule near the second critical point.

A type-C_{2v} bifurcation should also exist in the rotational spectra of the excited vibrational states, since the adiabatic approximation, and hence Eqs. (33) and (39), remain valid in this case. An analysis of experimental data[5,9] indicates that the upper levels of the rotational multiplets of the ν_2 vibrational state of the H_2O and H_2S molecules tend to cluster. Clusters have been found in rotational spectra of the ν_1/ν_3 vibrational states of the H_2Se (Refs. 13,14) H_2S, (Refs. 15) and H_2Te (Ref. 16) molecules using quantum mechanical variational calculations. Later, these clusters were observed experimentally in H_2Se, (Ref. 28) and H_2Te (Ref. 29).

5. VIBRATIONS OF A ROTATING MOLECULE

We now consider the change in the vibrational motion of a molecule during the transition at the critical point J_c. In the S_z stationary state ($J < J_c$) the adiabatic approximation gives the following equation of motion for the asymmetric mode ν_3:

$$\ddot{Q}_a + (b_{11} - b_{13}) f_{aa} Q_a = 0. \tag{40}$$

where, according to Eq. (13), $b_{13} \sim (m/M) b_{11}$. The symmetric ν_1 ($Q_{sm} = (Q_1 + Q_3)/\sqrt{2}$) and bending ν_2 (Q_2) modes do not depend on the precession and obey the coupled equations

$$\begin{gathered} \ddot{Q}_s + (b_{11} + b_{13}) f_{ss} Q_{sm} + (b_{11} + b_{13}) f_{s2} Q_2 = 0, \\ \ddot{Q}_2 + (b_{22} f_{22} + \sqrt{2} b_{12} f_{s2}) Q_2 + (b_{22} f_{s2} + \sqrt{2} b_{12} f_{ss}) Q_{sm} = 0. \end{gathered} \tag{41}$$

where $f_{ss} = f_{11} + f_{13}$ and $f_{s2} = \sqrt{2} f_{12}$, with $f_{11} = f_{33}$ and $f_{23} = f_{12}$ owing to the symmetry of the molecule in the S_z state. (See the definition of the force constants (32).) Clearly,

TABLE II. Normalized amplitudes of the internal coordinates and components of the total angular momentum of the four vibrational–precessional modes of the H_2Se molecule in the stationary S_z state for $J=10$.

Frequency, cm^{-1}	Normalized amplitudes			
	r_1-r_{1s}, Å	$\alpha-\alpha_s$, rad	r_3-r_{3s}, Å	J_x/J
$\omega_1=2426$	1*	−0.032	1.000	0
$\omega_2=1082$	−0.014	1*	−0.014	0
$\omega_3=2437$	−1.000	0	1*	0.004
$\Omega_z=16$	0.004	−0.004	0	1*

Note. *The normalization condition is $Q_n=1$.

the coupling takes place through the mixed derivatives of V_{eff} and the nondiagonal element b_{12}. The parameters

$$\frac{b_{11}f_{s2}}{b_{11}f_{ss}-b_{22}f_{22}}, \quad \frac{b_{12}f_{ss}}{b_{11}f_{ss}-b_{22}f_{22}}. \tag{42}$$

which determine the mode coupling, are small owing to the condition (27) and the inequality $m/M \ll 1$. Thus, for the hydrides the mode interaction is weaker and this allows us to classify the molecular vibrations in the S_z state according to the standard scheme:[22] the frequency ω_1 corresponds to symmetric valence vibrations ν_1, ω_3 to the asymmetric ν_3 mode, and ω_2 to the bending ν_2 mode. Unlike a nonrotating molecule, the centrifugal force of inertia changes the frequency of all the vibrations. This is illustrated in Table II, which lists the normalized coordinate amplitudes Q_1, Q_2, Q_3, and $Q_4=J_x/J$ for all four precessional–vibrational modes of the H_2Se molecule. These data were obtained by numerically solving the linearized equations for $J=10$. It is clear from the table that the interaction of the vibrational modes among themselves and with the precessional motion is indeed small.

The character of the vibrational motion changes radically after the critical point J_c. In the S_{xz} state the molecule becomes asymmetric, which leads to mixing of all three vibrational modes among themselves and with precession. For J that is not too large, the adiabatic approximation makes it possible to separate the vibrational motion, for which, according to Eq. (37), the equations have the form

$$\ddot{Q}_\nu + \sum_{\nu'\nu''} b_{\nu\nu'} f_{\nu'\nu''} Q_{\nu''} = 0. \tag{43}$$

They describe the coupled local vibrations Q_1, Q_2, and Q_3. And again, coupling takes place through the mixed derivatives and nondiagonal elements of the matrix $\mathbf{b}$. Since the inequality

$$\sum_{\nu''} \frac{b_{\nu\nu''} f_{\nu''\nu'}}{|b_{\nu\nu} f_{\nu\nu} - b_{\nu'\nu'} f_{\nu'\nu'}|} \ll 1, \quad \nu \neq \nu', \tag{44}$$

is satisfied for all the hydrides, the three local modes are quasi-independent vibrations. This effect is illustrated in Table III, which shows the normalized amplitudes of the vibrational–precessional motion of the H_2Se molecule given by Eq. (37) for $J=40$. The precession in the total angular momentum corresponds to the dimensionless variable $Q_4=J_{x'}/J$. It is clear that the precessional motion is mixed most strongly with bending vibrations.

TABLE III. Normalized amplitudes of the internal coordinates and components of the total angular momentum of the four vibrational–precessional modes of the H_2Se molecule in the stationary S_{xz} state for $J=40$ (the molecule rotates uniformly about an axis that forms an angle $\beta_s=42°$ with the z axis).

Frequency cm^{-1}	Normalized amplitudes			
	r_1-r_{1s}, Å	$\alpha-\alpha_s$, rad	r_3-r_{3s}, Å	$J_{x'}/J$
$\omega_1=2096$	1*	0.030	0.015	0.004
$\omega_2=1272$	−0.022	1*	−0.009	0.093
$\omega_3=2438$	−0.015	0.003	1*	−0.002
$\Omega_{xz}=212$	0	0.429	−0.009	1*

Note. *The normalization condition is $Q_n=1$.

Therefore, beyond the critical region the standard picture of the normal vibrations of symmetric triatomic molecules is not applicable. We note that the same result has been obtained in numerical calculations using the MORBID program.[14,15] Because of the asymmetry of the molecule, the elastic coupling constants of the central nucleus with the hydrogen nuclei become different. The difference becomes greater as J and the mass of the molecule increase, and this leads to a characteristic pattern in the vibrations for the local modes.[23] Usually the local modes arise as a result of the strong anharmonicity of the highly excited valence vibrations. In our case, the transition from normal to local vibrations takes place as a result of rotational excitation which disrupts the symmetry of the molecule owing to a bifurcation.

In the stationary S_y state, the vibrational motion independent of precession is described by the equation

$$\ddot{Q}_\nu + \sum_{\nu'\nu''} b_{\nu\nu'} (J h_{\nu'\nu''} \dot{Q}_{\nu''} + f_{\nu'\nu''} Q_{\nu''}) = 0. \tag{45}$$

The Coriolis force leads to an additional interaction of the vibrational modes which corresponds in Eq. (45) to the term with the first derivative and the antisymmetric matrix

$$h_{\nu\nu'} = \left[\frac{\partial(\mathring{\mu}_{yy} G_{y\nu})}{\partial q_{\nu'}} - \frac{\partial(\mathring{\mu}_{yy} G_{y\nu'})}{\partial q_\nu} \right]_s. \tag{46}$$

Since the molecule has a symmetric configuration in the S_y state, only one element of this matrix, $h_{12} \simeq -1/2r_s$, is nonzero. If in Eq. (46) we transform to the variables Q_{sm} and Q_a for the valence vibrations, use the smallness of the parameters (42), and neglect the nondiagonal terms b_{12} and b_{13} compared to the diagonal terms b_{11} and b_{22}, which are a factor of M/m greater than the former, then the symmetric valence vibrations are separated, while the antisymmetric and bending modes obey the coupled system of equations

$$\begin{aligned} &\ddot{Q}_a + b_{11}(\sqrt{2} J h_{12} \dot{Q}_2 + f_{aa} Q_a) = 0, \\ &\ddot{Q}_2 + b_{22}(-\sqrt{2} J h_{12} \dot{Q}_a + f_{22} Q_2) = 0. \end{aligned} \tag{47}$$

It is easy to see that the coupling parameter in these vibrations is of order $\omega/(\omega_3-\omega_2)$, where ω is the angular rotation frequency of the molecule. Since the difference in the frequencies of the two normal modes is $\omega_3-\omega_2 \sim 1000\ cm^{-1}$ for the heavy hydrides, the coupling

TABLE IV. Normalized vibration amplitudes of the H_2Se molecule in the stationary S_y state for $J=40$.

Frequency, cm^{-1}	Normalized amplitudes		
	r_1-r_{1s}, Å	$\alpha-\alpha_s$, rad	r_3-r_{3s}, Å
$\omega_1=2346$	1*	0.031	1.000
$\omega_2=1030$	-0.014	1*	-0.014
$\omega_3=2278$	-0.992	0.001	1*

Note. *The normalization condition is $Q_n=1$.

parameter is small if $J<100$. For these angular momenta the Coriolis force does not lead to significant distortion of the fundamental modes. This is illustrated by Table IV, which lists the normalized vibration amplitudes obtained by numerical solution of the exact equations (45). The table does not show the tiny phase shift in the coordinates of the normal vibrations owing to the Coriolis interaction.

The authors thank the Russian Fund for Fundamental Research for partial support of this work through grants Nos. 96-02-16115 and 97-02-16593. One of the authors (I.N.K.) thanks PECO and CIES for providing a stipend.

[1)] This quantity is measured in units of $\hbar$ in the following.

---

[1] A. J. Dorney and J. K. G. Watson, J. Mol. Spectrosc. **42**, 135 (1972).
[2] W. G. Harter and C. W. Patterson, J. Chem. Phys. **80**, 4241 (1984).
[3] W. G. Harter, Comput. Phys. Rep. **8**, 319 (1988).
[4] D. A. Sadovskii, B. I. Zhilinskii, J.-P. Champion, and G. Pierre, J. Chem. Phys. **92**, 1523 (1990).
[5] B. I. Zhilinskiĭ and I. M. Pavlichenkov, Opt. i Spektr. **64**, 688 (1988).
[6] J. T. Hougen, P. R. Bunker, and J. W. C. Johns, J. Mol. Spectrosc. **34**, 136 (1970).
[7] I. M. Pavlichenkov, Phys. Rep. **226**, 175 (1993).
[8] J. Makarewicz and J. Pyka, Mol. Phys. **68**, 107 (1989); J. Makarewicz, Mol. Phys. **69**, 903 (1990); J. Pyka, Mol. Phys. **70**, 547 (1990).
[9] I. M. Pavlichenkov, in *Physics of the Atomic Nucleus (Material from the XXIV Winter School at the Leningrad Inst. of Nuclear Physics, Leningrad (1989)* [in Russian], Leningrad Inst. of Nuclear Physics, Leningrad (1989), p. 69.
[10] I. N. Kozin, S. P. Belov, O. I. Polyansky, and M. Yu. Tretyakov, J. Mol. Spectrosc. **152**, 13 (1992).
[11] I. N. Kozin, S. Klee, P. Jensen, O. I. Polyansky, and I. M. Pavlichenkov, J. Mol. Spectrosc. **158**, 409 (1993).
[12] I. N. Kozin, P. Jensen, O. Polanz, S. Klee, L. Poteau, and J. Demaison, J. Mol. Spectrosc. **180**, 402 (1996).
[13] P. Jensen and I. N. Kozin, J. Mol. Spectrosc. **160**, 139 (1993).
[14] I. N. Kozin and P. Jensen, J. Mol. Spectrosc. **161**, 186 (1993).
[15] I. N. Kozin and P. Jensen, J. Mol. Spectrosc. **163**, 483 (1994).
[16] P. Jensen, Yan Li, G. Hirsch, R. J. Buenker, T. H. Lee, and I. N. Kozin, Chem. Phys. **180**, 179 (1995).
[17] P. Jensen, J. Mol. Spectrosc. **128**, 478 (1988); J. Chem. Soc. Faraday Trans. 2, **84**, 1315 (1988).
[18] P. Jensen and I. N. Kozin, unpublished results.
[19] P. Jensen, J. Mol. Spectrosc. **133**, 438 (1989); P. Jensen, S. Tashkun, and V. G. Tyuterev, J. Mol. Spectrosc. **168**, 271 (1994).
[20] O. L. Polyansky, P. Jensen, and J. Tennyson, J. Chem. Phys. **101**, 7651 (1994).
[21] I. N. Kozin and I. M. Pavlichenkov, J. Chem. Phys. **104**, 4105 (1996).
[22] E. Wilson, J. Decius, and P. Cross, *Molecular Vibrations: the Theory of Infrared and Raman Vibrational Spectra*, McGraw-Hill, N. Y. (1955).
[23] M. S. Child and I. Halonen, Adv. Chem. Phys. **57**, 1 (1984).
[24] B. T. Sutcliffe and J. Tennyson, Mol. Phys. **58**, 1053 (1986); Int. J. Quantum Chem. **39**, 183 (1991).
[25] P. R. Bunker *Molecular Symmetry and Spectroscopy*, Academic Press, London (1979)
[26] P. Jensen and P. R. Bunker, J. Mol. Spectrosc. **164**, 315 (1994).
[27] L. D. Landau and E. M. Lifshitz, *Mechanics* 3rd ed., Pergamon, Oxford (1976).
[28] J.-M. Fland, C. Camy-Peyret, H. Bürger, P. Jensen, and I. N. Kozin, J. Mol. Spectrosc. **172**, 194 (1995).
[29] J.-M. Flaud, M. Betrencourt, Ph. Arcas, H. Bürger, O. Polanz, and W. J. Lafferty, J. Mol. Spectrosc. **182**, 396 (1997).

Translated by D. H. McNeill

Generation of one-photon states by a quantum dot

A. V. Krasheninnikov and L. A. Openov

Moscow State Institute of Engineering Physics, 115409 Moscow, Russia

S. N. Molotkov and S. S. Nazin

Institute of Solid State Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

(Submitted 22 April 1997)
Zh. Éksp. Teor. Fiz. **112**, 1257–1272 (October 1997)

A technique for preparation of a one-photon wave packet through action of a classical electromagnetic field on a semiconducting quantum dot is proposed. We demonstrate that the Coulomb repulsion between charge carriers allows one to select the frequency, amplitude, and duration of an electromagnetic pulse so that one electron will transfer from an upper size-quantized level of the valence band to a lower size-quantized level of the conduction band with a probability close to unity. As a result of radiative recombination of the produced electron–hole pair, exactly one photon is emitted (a one-photon wave packet). This source of one-photon states can be used in quantum systems of data transmission and in quantum computers. © *1997 American Institute of Physics.* [S1063-7761(97)00810-X]

1. INTRODUCTION

''Classical'' states of the electromagnetic field containing a large (formally infinite) number of photons have by now been studied in detail. This includes coherent and compressed states,[1] in particular. In recent years, interest in ''constructing'' nonclassical photon states has risen. For example, nonlocal two-photon interferometry[2,3] and quantum cryptosystems[4–7] require so-called entangled photon states.[8] A line of research in modern quantum physics such as quantum cryptography is largely based on application of one-photon states, since the secrecy of quantum cryptographic systems depends on the impossibility of cloning (copying) *a priori* unknown quantum states.[9]

Presently the focus of studies in quantum cryptography has shifted from purely theoretical analysis of the problem to designing experimental facilities for quantum data transmission, for the time being over relatively short distances (30 km; see Ref. 5). This effort requires reliable sources of one-photon states. Here, reliability means that the sources generate only one-photon states (but not states with two or more photons).

What can be a source of such states? The simplest method of obtaining an ''almost one-photon'' state is to alternate a coherent laser output originally containing many Fock states[5] $|n\rangle$ (this technique was used in experiments in quantum cryptography based on phase coding[6,7]). The field at the output of an ideal laser above threshold can be described as a coherent state $|\alpha\rangle$ of the form

$$|\alpha\rangle=\exp\left(-\frac{|\alpha|^2}{2}\right)\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle. \qquad (1)$$

The coherent state attenuated downstream of an absorbing medium corresponds to the limit of small $|\alpha|$ (recall that the intensity, i.e., the average number of photons, of a monochromatic coherent field is proportional to $|\alpha|^2$). Expanding Eq. (1) in terms of $|\alpha|\ll 1$, we obtain

$$|\alpha\rangle\approx\left(1-\frac{1}{2}|\alpha|^2\right)|0\rangle+\alpha|1\rangle+\frac{|\alpha|^2}{\sqrt{2}}|2\rangle. \qquad (2)$$

This equation shows that the largest contribution to the coherent state of the photon field is due to the vacuum component $(1-|\alpha|^2/2\approx 1)$, whereas the contribution of the one-photon Fock component $|1\rangle$ is small insofar as $|\alpha|\ll 1$. The latter condition means that the probability of detecting a single photon is also small.

It would seem that the situation could be improved by increasing $|\alpha|$, i.e., the intensity. But then, although the contribution of the vacuum component to $|\alpha\rangle$ would decrease (see Eq. (2)), the probability of emitting two photons (Fock state $|2\rangle$) would markedly increase. In the quantum cryptography this would destroy the security of the cryptosystem, since a fraction of photons could be ''extracted'' from the line and used for eavesdropping.

Application of the attenuated state described by Eq. (2) with $|\alpha|\ll 1$ to quantum cryptography leads to another difficulty. The problem is that a cryptosystem based on the principle of phase coding[6] ensures secrecy only if the states with phase shifts 0°/180° and 90°/270° are orthogonal to one another (see Ref. 6 for details). The overlap of these states, however, is always nonzero, being proportional to $\exp(-|\alpha|^2/2)$, and the smaller the value of $|\alpha|$, the closer to unity this overlap.

Note that in quantum cryptography, the requirement that the one-photon state be monochromatic is not absolute. It suffices to prepare a one-photon packet that can be described by the formula[10–12]

$$|1\rangle_f=\sum_{\omega=0}^{\infty}f(\omega)\hat{a}^+(\omega)|0\rangle=\sum_{\omega=0}^{\infty}f(\omega)|1_\omega\rangle, \qquad (3)$$

where $\hat{a}^+(\omega)$ is the boson creation operator of monochromatic state $|1_\omega\rangle$. The one-photon nature of the wave packet $|1\rangle_f$ in Eq. (3) means that only states with population $n=1$

1063-7761/97/100682-08$10.00

are included in the expansion of $|1\rangle_f$ in Fock states $|1_\omega\rangle$, although they may have different frequencies. This state of the photon field is a ''one-particle'' quantum state, and it generates only a single activation of a light detector, or only one transition in a detecting system (it cannot be split so as to generate two detector outputs or two transitions). In this sense, the state $|1\rangle_f$ can be termed as a ''single nonmonochromatic photon.'' The function $f(\omega)$ defines the spectrum of the emitted wave packet, which formally contains an infinite number of modes $|1_\omega\rangle$, but with amplitudes $f(\omega)$ whose absolute values are smaller than unity. This function is normalized,

$$\sum_{\omega=0}^{\infty} |f(\omega)|^2 = 1,$$

and is assumed to be centered about a certain carrier frequency ω_0.

The one-photon wave packet (3) can be produced in the following manner. One electron, which is initially in a quasistationary excited state $|e\rangle$ with energy ε_e, is transferred to the ground state $|g\rangle$ with energy $\varepsilon_g < \varepsilon_e$ as a result of interaction with the photon field, which is initially in the vacuum state.

Thus, the problem of generating a one-photon state can be reduced to the problem of preparing a quasistationary excited state of a single electron. This problem has been discussed in the literature. For example, Imamoglu and Yamamoto[13] suggest a method for fabricating such a state in an Al–GaAs–GaAs $p-i_p-i-i_n-n$ heterostructure by applying an ac voltage with specially selected parameters. In this process, one electron and one hole resonantly tunnel from the contacts (across the potential barriers) into the structure and annihilate with emission of a photon. In their opinion, the effects of size quantization and Coulomb blockade should suppress quantum fluctuations associated with the injection of electrons and holes. This scheme, however, like other similar designs, imposes stringent requirements on the parameters of the nanostructure and applied electric field, so that experimental implementation is questionable.

One can suggest another method for exciting a single electron to a quasistationary level, namely, to apply an ac electromagnetic field to an isolated quantum dot with size-quantized levels in both the valence and conduction bands. At first sight, this suggestion seems trivial. Indeed, suppose that all levels of the valence band are occupied, and all conduction band levels are vacant (which is usually the case in a semiconductor). Let us denote by ε_v and ε_c the energies of the upper (occupied) size-quantized level $|v\rangle$ in the valence band and the lower (vacant) level $|c\rangle$ in the conduction band. It is well known that a periodic perturbation $\hat{V}\cos(\Omega t)$ (for example, a classical ac electric field) with frequency $\Omega = \varepsilon_c - \varepsilon_v$ (hereinafter we take $\hbar = 1$) leads to periodic oscillation of the probabilities p_v and p_c of detecting an electron in levels $|v\rangle$ and $|c\rangle$.[14–17] If $p_v(0)=1$ and $p_c(0)=0$ at the initial moment, then

$$p_c(t) = \sin^2(\omega_R t), \tag{4}$$

where $\omega_R = |\langle c|\hat{V}|v\rangle|/2$. Here $\langle c|\hat{V}|v\rangle$ is the matrix element of the interband transition. It follows from Eq. (4) that one can select the time T during which the perturbation is on (for example, $T = \pi/2\omega_R$) so that the condition $p_c(T)=1$ is satisfied (so-called π-pulse[14–17]). Then at time T, the electron is in the quasistationary state $|c\rangle$ with probability one.

This scheme, however, applies only when the levels $|v\rangle$ and $|c\rangle$ are nondegenerate. For example, suppose that they are doubly degenerate in the spin projection, and at $t=0$ level $|v\rangle$ contains not one but two electrons. One can easily calculate the probability that the periodic perturbation $\hat{V}\cos(\Omega t)$ has transferred only one electron (its spin projection is unimportant) to level $|c\rangle$, whereas the other has remained in level $|v\rangle$:

$$p_{vc}(t) = \frac{1}{2}\sin^2(2\omega_R t). \tag{5}$$

It follows from Eq. (5) that $p_{vc} \leqslant 1/2$, and one cannot be sure that exactly one electron has been transferred to the quasistationary state $|c\rangle$ if levels $|c\rangle$ and $|v\rangle$ are degenerate (the limitation $p_{vc} \leqslant 1/2$ is due to the ''transfer of probability'' to the state with two electrons on level $|c\rangle$, i.e., $p_{cc} \leqslant 1$).

Whereas the degeneracy in both angular momentum and its projection is usually lifted because a quantum dot is usually not spherically symmetric,[18] spin degeneracy persists in the absence of a magnetic field. To lift this degeneracy, one must apply a very strong magnetic field. Let us estimate its strength H. Our estimate is based on the assumption that the longest experimentally attainable electron lifetime in the excited level $|c\rangle$ of the quantum dot is $\tau \approx 10^{-10}$–10^{-9} s. The perturbation duration time T needed to transfer an electron from the valence to conduction band should be much shorter than τ, otherwise the probability of photon emission at $t<T$ is high, and the emitted photon will be ''lost'' in the perturbation field. Thus, we should have $T \approx 10^{-12}-10^{-11}$ s. But in this case the perturbating field contains harmonics in the frequency range $\delta\omega \approx 1/T \approx 10^{11}-10^{12}\ \mathrm{s}^{-1}$. The bandwidth $\delta\omega$ should be much smaller than the Zeeman splitting energy ($\approx \mu_B H$) of the size-quantized levels $|v\rangle$ and $|c\rangle$, since otherwise the external field will still mix all the states $|v\uparrow\rangle$, $|v\downarrow\rangle$, $|c\uparrow\rangle$, and as before $|c\downarrow\rangle$, i.e., although the spin degeneracy is formally lifted, its consequences persist. As a result, we have an estimate $H \gg 10$ T, which is very difficult to achieve experimentally.

In this paper we suggest an alternative approach to the problem. It is based on consideration of the Coulomb repulsion between electrons with different spin projections that occupy (in the absence of interaction between them) the same one-electron level (in a real quantum dot, Coulomb repulsion is always present, and the smaller the dot the stronger it is). It is clear from qualitative considerations that if the Coulomb repulsion (for example, in level $|c\rangle$) were infinite, the transition of a second electron to this level would be forbidden, and one could generate the one-electron state with probability $p_{vc}=1$ at some time $t=T$. But if the Coulomb repulsion energy is finite, this is not obvious.

Our aim is to check whether the transition of one electron from the valence to conduction band is possible under the action of a periodic perturbation on a quantum dot with finite Coulomb interaction. It will be shown that for realistic parameters of the system (Coulomb repulsion energy, frequency and amplitude of applied periodic electric field, and lifetime of the electron in the quasistationary state), one can select a pulse duration T such that the probability of driving one electron to the conduction band is close to unity. Thus, it is possible to design a source of one-photon wave packets on the basis of an isolated quantum dot without resorting to sophisticated techniques.

2. QUALITATIVE ESTIMATES. DESCRIPTION OF THE MODEL

Let us first approximately estimate the splitting $\Delta\varepsilon$ of the doubly degenerate (in the spin projection) size-quantized level $|v\rangle$ due to the Coulomb interaction between electrons in states $|v\uparrow\rangle$ and $|v\downarrow\rangle$:

$$\Delta\varepsilon\approx U=\left\langle v\uparrow\left|\frac{e^2}{\varepsilon|\mathbf{r}_1-\mathbf{r}_2|}\right|v\downarrow\right\rangle\approx\frac{e^2}{\varepsilon a} \tag{6}$$

is the matrix element of the Coulomb interaction, ε is the static dielectric constant, and a is the characteristic linear dimension of the quantum dot. It follows from Eq. (6) that at $\varepsilon\approx 10$ and $a\approx 10$ nm, we have $U\approx 10$ meV. Our estimate of U is in agreement with the experimental data to order of magnitude.[19] A similar estimate is, naturally, obtained for the splitting of level $|c\rangle$ in the conduction band due to the Coulomb repulsion between electrons in states $|c\uparrow\rangle$ and $|c\downarrow\rangle$.

As was demonstrated in Introduction, it is possible to excite exactly one electron to the conduction band with probability $p_{vc}=1$ only when $T\ll\tau$ and $1/T\ll\Delta\varepsilon$. Thus, we obtain from Eq. (6)

$$1/\tau\ll 1/T\ll U. \tag{7}$$

Since $1/\tau\approx 10^{-3}-10^{-2}$ meV, we have $1/T\approx 0.1-1$ meV, so condition (7) can be satisfied for $U\geqslant 1-10$ meV, i.e., as follows from Eq. (6), for $a\leqslant 10-100$ nm. The range of admissible T can, of course be broadened. To this end, one should increase τ and/or reduce a. For our purposes, however, this range of T is sufficient, since condition (7) is then satisfied at realistic values of τ and a in semiconducting quantum dots fabricated using contemporary technologies.

Note also that we would be satisfied, in principle, by the situation in which the maximum value of $p_{vc}(t)$ is not exactly unity, but very close to it. Therefore, we determine the time T of the external perturbation duration by the condition

$$p_{vc}(T)=\max[p_{vc}(t)]. \tag{8}$$

In addition, let us agree to select the minimal value of T among those satisfying condition (8) (otherwise the first of the inequalities in Eq. (7) can be violated).

Before proceeding to the description of our model, let us somewhat simplify the problem. Assume that both the valence and conduction bands contain only one size-quantized level that is doubly degenerate in the spin projection. Generally speaking, this is not so. Physically, this simplification means that we take into account only electronic transitions from the upper valence band level to the lower conduction band level, and ignore low-lying (fully occupied) levels of the valence band and upper (vacant) levels of the conduction band (the effect of these levels will be discussed below). In describing Coulomb repulsion, we use the Hubbard model,[20] which is widely used in the theory of systems with strong Coulomb correlations.[21]

Thus, taking into account the discussion above, we express the model Hamiltonian of a semiconducting quantum dot in a periodic electric field as

$$\hat{H}(t)=\sum_{\sigma}(\varepsilon_v\hat{a}^+_{v\sigma}\hat{a}_{v\sigma}+\varepsilon_c\hat{a}^+_{c\sigma}\hat{a}_{c\sigma})+U_v\hat{n}_{v\uparrow}\hat{n}_{v\downarrow}+U_c\hat{n}_{c\uparrow}\hat{n}_{c\downarrow}+E(t)\sum_{\sigma}(d_{cv}\hat{a}^+_{c\sigma}\hat{a}_{v\sigma}+d_{vc}\hat{a}^+_{v\sigma}\hat{a}_{c\sigma}), \tag{9}$$

where $\hat{a}^+_{q\sigma}(\hat{a}_{q\sigma})$ is the creation (annihilation) operator for an electron with spin $\sigma=\uparrow$ or $\downarrow$ in size-quantized levels $|q\rangle=|v\rangle$ and $|c\rangle$ (the subscript v refers everywhere to the valence band, and subscript c to the conduction band); $\hat{n}_{q\sigma}=\hat{a}^+_{q\sigma}\hat{a}_{q\sigma}$ is the particle number operator (population number); ε_q are energies of one-electron levels; U_q is the Hubbard repulsion energy between two electrons in the same level; d_{cv} is the matrix element of optical dipole transitions ($d_{vc}=d^*_{cv}$); and $E(t)$ is the external classical ac field.

Let us suppose that the external field is turned on at $t=0$ and turned off at $t=T$, and has a carrier frequency Ω, i.e.,

$$E(t)=E_0\cos(\Omega t)\theta(t)\theta(T-t),$$

where E_0 is the field amplitude, $\theta(t)$ is the Heaviside step function. The duration T of the applied field action and its frequency Ω are to be drived by maximizing the probability that only one electron is transferred from the valence band to the quasistationary level of the conduction band (Eq. (8)).

Since we consider two electrons with different spin projections in two different one-electron energy levels, the total number of two-electron states of the system is four. We denote these states by

$$|1\rangle=|v\uparrow,v\downarrow\rangle,\quad |2\rangle=|v\uparrow,c\downarrow\rangle,\quad |3\rangle=|v\downarrow,c\uparrow\rangle,\quad |4\rangle=|c\uparrow,c\downarrow\rangle, \tag{10}$$

where, for example, $|v\uparrow,c\downarrow\rangle$ means that the electron with $\sigma=\uparrow$ is in the valence band, and the electron with $\sigma=\downarrow$ is in the conduction band, etc. For the system under consideration, the set of states defined by Eq. (10) is complete, so the wave function $\Psi(t)$ can be expressed at any moment as

$$\Psi(t)=\sum_{i=1}^{4}A_i(t)\exp(-iE_it)|i\rangle, \tag{11}$$

where E_i are eigenvalues of the time-independent Schrödinger equation $\hat{H}|i\rangle=E_i|i\rangle$ in the absence of an applied field ($t\leqslant 0$). With due account of Eq. (10), we derive from Eq. (9)

$$E_1=2\varepsilon_v+U_v,\quad E_2=\varepsilon_v+\varepsilon_c,\quad E_3=\varepsilon_v+\varepsilon_c,\quad E_4=2\varepsilon_c+U_c. \tag{12}$$

Note once again that we use the complete set of two-electron states defined by Eq. (10) and schematically shown in Fig. 1.

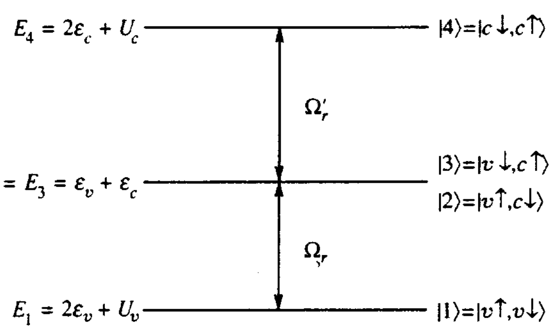

FIG. 1. Diagram of energy levels of two-particle states of a semiconducting quantum dot.

The values $A_i(0)$ define the system wave function at the initial moment (moreover, $A_i(0)=A_i(t<0)$ since the $|i\rangle$ are eigenstates of Hamiltonian (9) for $t\leqslant 0$). We assume that at $t\leqslant 0$ the semiconductor is in its usual state, namely, the valence band is fully occupied and the conduction band is empty, i.e., $A_1(0)=1$, $A_2(0)=A_3(0)=A_4(0)=0$. The probability $p_i(t)$ of detecting the system in state $|i\rangle$ at an arbitrary time t is $|A_i(t)|^2$. We are interested in the value of $p_{vc}(t)=p_2(t)+p_3(t)$, which is the probability (see Eq. (10)) that both the valence and conduction bands contain one electron each (it follows from the normalization condition (11) that, naturally, $p_1(t)+p_2(t)+p_3(t)+p_4(t)=1$ at any t).

The coefficients $A_i(t)$ in expansion (12) can be calculated by solving the time-dependent Schrödinger equation

$$i\,\frac{\partial\Psi(t)}{\partial t}=\hat H(t)\Psi(t), \tag{13}$$

where $\hat H(t)$ is given by Eq. (9), i.e., it explicitly depends on time when $0\leqslant t\leqslant T$.

3. RESONANT APPROXIMATION

In order to solve the problem as formulated, we use the resonant approximation,[14–17] which enables us to progress reasonably far analytically and obtain an accurate solution (in some limiting cases in analytic form). One can, certainly, go beyond the resonant approximation, but in this case the solution can be obtained only numerically and, as will be shown below, this will not lead to any significant changes in the results. In the resonant approximation, the Hamiltonian (9) has the form

$$\hat H(t)=\sum_\sigma (\varepsilon_v\hat a^+_{v\sigma}\hat a_{v\sigma}+\varepsilon_c\hat a^+_{c\sigma}\hat a_{c\sigma})+U_v\hat n_{v\uparrow}\hat n_{v\downarrow} + U_c\hat n_{c\uparrow}\hat n_{c\downarrow}+\sum_\sigma\left[\frac{\lambda}{2}\exp(-i\Omega t)\hat a^+_{c\sigma}\hat a_{v\sigma} + \frac{\lambda^*}{2}\exp(i\Omega t)\hat a^+_{v\sigma}\hat a_{c\sigma}\right], \tag{14}$$

where we have introduced the notation $\lambda=d_{cv}E_0$.

In what follows, we will obtain exact solutions. It is noteworthy that, unlike the elementary case of a resonant transition between two one-electron levels $|v\rangle$ and $|c\rangle$, which occurs at the frequency $\Omega=\varepsilon_c-\varepsilon_v$,[14] we consider all allowed transitions between four two-electron states of the system. Therefore the resonant frequency is different from $\varepsilon_c-\varepsilon_v$ in the presence of finite Coulomb repulsion.

To solve the Schrödinger equation (13) with Hamiltonian (14), it is convenient to go to a representation with a time-independent Hamiltonian. To this end, we use the unitary transformation

$$\hat U(t)=\exp\left[-\frac{i\Omega t}{2}\sum_\sigma(\hat a^+_{c\sigma}\hat a_{c\sigma}-\hat a^+_{v\sigma}\hat a_{v\sigma})\right], \tag{15}$$

which was used in Ref. 16 to describe the interaction between an intense electromagnetic field and a semiconductor.

We replace the wave function $\Psi(t)$ with $\tilde\Psi(t)=\hat U^+(t)\Psi(t)$. Substituting

$$\Psi(t)=\hat U(t)\tilde\Psi(t) \tag{16}$$

into the Schrödinger equation (13) for $\Psi(t)$, and taking into account Eq. (15), we obtain the Schrödinger equation for $\tilde\Psi(t)$:

$$i\,\frac{\partial\tilde\Psi(t)}{\partial t}=\hat{\tilde H}\tilde\Psi(t), \tag{17}$$

with the Hamiltonian

$$\begin{aligned}\hat{\tilde H}&=\hat U^+(t)\hat H(t)\hat U(t)-i\hat U^+(t)\frac{\partial\hat U(t)}{\partial t}\\ &=\sum_\sigma\left[\left(\varepsilon_v+\frac{\Omega}{2}\right)\hat a^+_{v\sigma}\hat a_{v\sigma}+\left(\varepsilon_c-\frac{\Omega}{2}\right)\hat a^+_{c\sigma}\hat a_{c\sigma}\right]\\ &\quad+U_v\hat n_{v\uparrow}\hat n_{v\downarrow}+U_c\hat n_{c\uparrow}\hat n_{c\downarrow}+\sum_\sigma\left(\frac{\lambda}{2}\hat a^+_{c\sigma}\hat a_{v\sigma}\right.\\ &\quad\left.+\frac{\lambda^*}{2}\hat a^+_{v\sigma}\hat a_{c\sigma}\right).\end{aligned} \tag{18}$$

This Hamiltonian $\hat{\tilde H}$ is independent of time; therefore the general solution of the time-dependent Schrödinger equation (17) for $0\leqslant t\leqslant T$ has the form

$$\tilde\Psi(t)=\sum_{i=1}^4 B_i\exp(-i\tilde E_i t)|\tilde i\rangle, \tag{19}$$

where $|\tilde i\rangle$ and $\tilde E_i$ are the eigenstates and eigenenergies of the time-independent Schrödinger equation

$$\hat{\tilde H}|\tilde i\rangle=\tilde E_i|\tilde i\rangle \tag{20}$$

with Hamiltonian $\tilde H$ in the form given by Eq. (18). The coefficients B_i in Eq. (19) should be derived from the initial conditions (predetermined values of $A_i(0)$ in Eq. (11)) and satisfy the normalization condition

$$\sum_{i=1}^4|B_i|^2=1. \tag{21}$$

We seek solutions of Eq. (20) in the form

$$|\tilde i\rangle=\sum_{k=1}^4 C_{ik}|k\rangle, \tag{22}$$

where the $|k\rangle$ are two-electron eigenstates defined by Eq. (10) for $t\leqslant 0$. Substituting Eq. (22) into Eq. (20), we obtain a set of equations that determines $\widetilde{E}_i$ and C_{ik}:

$$\sum_{k=1}^{4} C_{ik}(\langle i|\hat{\widetilde{H}}|k\rangle-\delta_{ik}\widetilde{E}_i)=0, \tag{23}$$

where $i=1,2,3,4$, $\langle i|\hat{\widetilde{H}}|k\rangle$ are the matrix elements of Hamiltonian (18) in terms of the basis states (10). The Hamiltonian matrix $\langle i|\hat{\widetilde{H}}|k\rangle$ has the form

$$\begin{pmatrix} 2\varepsilon_v+U_v+\Omega & \lambda^*/2 & \lambda^*/2 & 0 \\ \lambda/2 & \varepsilon_v+\varepsilon_c & 0 & \lambda^*/2 \\ \lambda/2 & 0 & \varepsilon_v+\varepsilon_c & \lambda^*/2 \\ 0 & \lambda/2 & \lambda/2 & 2\varepsilon_c+U_c-\Omega \end{pmatrix}.$$

From Eqs. (19) and (22) we obtain

$$\widetilde{\Psi}(t)=\sum_{i=1}^{4} D_i(t)|i\rangle, \tag{24}$$

where

$$D_i(t)=\sum_{k=1}^{4} B_k C_{ki}\exp(-i\widetilde{E}_k t). \tag{25}$$

Since $\widetilde{\Psi}(0)=\Psi(0)$ (see Eqs. (15) and (16)), we have $D_i(0)=A_i(0)$, where the coefficients $A_i(0)$ determine the state (11) for $t\leqslant 0$. Therefore, we obtain from Eq. (25) an equation to determine the coefficients B_i in terms of given $A_i(0)$:

$$A_i(0)=\sum_{k=1}^{4} B_k C_{ki},$$

whence

$$B_i=\sum_{k=1}^{4} A_k(0)C_{ki}^{-1}, \tag{26}$$

where C^{-1} is the matrix inverse of C. From Eqs. (25) and (26) we obtain

$$D_i(t)=\sum_{k=1}^{4}\sum_{l=1}^{4} A_l(0)C_{li}^{-1}C_{ki}\exp(-i\widetilde{E}_k t). \tag{27}$$

Given Eq. (16) relating the function $\widetilde{\Psi}(t)$ to $\Psi(t)$, and taking into account that the operator $\hat{U}(t)$ defined by Eq. (9) is unitary, we obtain an expression for the probability $p_i(t)$ for the transition to state $|i\rangle$:

$$p_i(t)=|D_i(t)|^2. \tag{28}$$

In the case under investigation, $A_1(0)=1$, $A_2(0)=A_3(0)=A_4(0)=0$ (at time $t=0$, all levels of the valence band are occupied and the levels of the conduction band are vacant; see Eq. (10)), expression (27) for $D_i(t)$ can be simplified somewhat. Nonetheless, the calculation of $D_i(t)$ still requires finding $\widetilde{E}_i$ and C_{ik} from Eq. (23). This eigenvalue problem reduces to the solution of a cubic algebraic equation. The resulting expressions for $\widetilde{E}_i$ and C_{ik} are cumbersome and difficult to analyze, so it is more convenient at this stage to solve Eq. (23) numerically. However, before discussing numerical calculations, we consider an important limiting case that admits of an analytic solution which is in excellent agreement with the numerical solution (obtained using the scheme described above) in the parameter range of interest.

Suppose that the Coulomb repulsion between electrons in the conduction band is infinite, $U_c=\infty$. The two-electron state $|4\rangle=|c\uparrow,c\downarrow\rangle$ (see Eq. (10)), whose energy is $E_4=2\varepsilon_c+U_c=\infty$, is then obviously always empty, i.e., the coefficient $A_4(t)$ in Eq. (11) is zero for all t. For the same reason, the matrix elements C_{i4} $(i=1-4)$ in Eq. (22) are also zero. Therefore, Eq. (23) can be solved by diagonalizing a 3×3 matrix, which can be done analytically. Thus, having calculated the eigenvalues $\widetilde{E}_i$ and the matrix of eigenvectors C_{ik}, we obtain from Eqs. (27) and (28) the following expressions for the transition probabilities $p_i(t)$ from the state $\Psi(0)=|1\rangle=|v\uparrow,v\downarrow\rangle$ to the state $|i\rangle$:

$$\begin{aligned} p_1(t)&=1-\left(1-\frac{\delta^2}{4\omega_R^2}\right)\sin^2(\omega_R t),\\ p_2(t)&=\frac{1}{2}\left(1-\frac{\delta^2}{4\omega_R^2}\right)\sin^2(\omega_R t),\\ p_3(t)&=\frac{1}{2}\left(1-\frac{\delta^2}{4\omega_R^2}\right)\sin^2(\omega_R t),\\ p_4(t)&=0, \end{aligned} \tag{29}$$

where

$$\delta=\Omega-(\varepsilon_c-\varepsilon_v-U_v),\quad \omega_R=\frac{\sqrt{\delta^2+2|\lambda|^2}}{2}. \tag{30}$$

From Eq. (29) we derive the desired probability $p_{vc}(t)=p_2(t)+p_3(t)$ for the transition of one electron to the conduction band:

$$p_{vc}(t)=\left(1-\frac{\delta^2}{4\omega_R^2}\right)\sin^2(\omega_R t). \tag{31}$$

It follows from Eqs. (30) and (31) that $p_{vc}(T)=1$ at $\delta=0$ and

$$T=\frac{\pi}{2\omega_R}. \tag{32}$$

The equality $\delta=0$ has a simple physical meaning: the frequency Ω is equal to $\Omega_r=E_2-E_1=E_3-E_1$ (see Eq. (12) and Fig. 1), i.e., the difference between the energy of the pair of degenerate levels $|2\rangle$ and $|3\rangle$, and that of level $|1\rangle$. Only these two-electron states are resonant at $\delta=0$. Since $\varepsilon_c-\varepsilon_v\approx 1$ eV and $U_v+U_c\ll\varepsilon_c-\varepsilon_v$, we have $\Omega_r\approx 10^{15}$ s^{-1}. When the frequency is offset from resonance $(\delta\neq 0)$, the value of $p_{vc}(T)$ derived from Eq. (8) deviates from the maximum value by a quantity of order $|\delta|^2/|\lambda|^2$, i.e., the following inequality should hold:

$$|\delta|\ll|\lambda|. \tag{33}$$

It is clear that at $\delta=0$ our results should also be valid when U_c is finite (the lower boundary will be determined below), since the system can then only be in states $|1\rangle$, $|2\rangle$, and $|3\rangle$ (if $U_c=\infty$, this statement is true at any δ). Moreover, if $\delta=0$, the sum U_v+U_c should be finite (i.e., U_c may be zero), which follows from the condition that the resonant frequencies Ω_r and $\Omega_r'=E_4-E_2=E_4-E_3$ (Fig. 1) for transitions between the levels $|1\rangle\leftrightarrow|2\rangle,|3\rangle$ and $|4\rangle\leftrightarrow|2\rangle,|3\rangle$, respectively, be different (at $U_v=U_c=0$ we have $p_{vc}\leqslant 1/2$; see Eq. (5)). At $\delta\neq 0$, the necessary condition for the validity of Eq. (31) for $p_{vc}(t)$ is obviously

$$|\delta|\ll U_v+U_c, \tag{34}$$

and the important parameter is the sum U_c+U_v, rather than U_v and U_c individually.

Inequalities (33) and (34) do not specify the ratio between $|\lambda|$ and U_v+U_c. Since for $|\delta|\ll|\lambda|$ we have $\omega_R\approx|\lambda|$, by generalizing condition (7), discussed in Sec. 2, we obtain $|\lambda|\ll U_v+U_c$, i.e., it follows from Eqs. (7) and (33) that

$$1/\tau,|\delta|\ll|\lambda|\ll U_v+U_c \tag{35}$$

(recall that τ is the radiative recombination time). Thus, if condition (35) is satisfied, the ''three-level approximation'' should apply to electronic transitions in the investigated four-level system under an applied an periodic perturbation.

We emphasize that conditions (35) can be satisfied in a real experiment. In fact, since $U_v+U_c\geqslant 1-10$ meV and $1/\tau\approx 10^{-3}-10^{-2}$ meV (see Sec. 2), we should have $|\lambda|\approx 0.1-1$ meV (the upper boundary of this interval can be increased by increasing U_v+U_c, i.e., decreasing the quantum dot size a, while the lower boundary is determined by the lifetime τ). Recalling that the parameter $|\lambda|$ is a product of the optical dipole matrix element $|d_{cv}|$ and the electric field amplitude E_0, and noting that $|d_{cv}|\approx ea$, where $a\approx 10-100$ nm, we obtain $E_0\approx 10^2$ V/cm, which can be easily realized in experiment. As concerns the condition $|\delta|\ll|\lambda|$, we have $|\delta|\leqslant 0.01$ meV, i.e., the frequency Ω of the external source should be accurate to within 10^{10} s^{-1}. Modern experimental methods make this possible.

We now compare the numerical solutions of Eqs. (23)–(28) to the analytic solution of Eqs. (30)–(32) in the ''three-level approximation.'' Let us take for definiteness $\varepsilon_c-\varepsilon_v\approx 1$ eV. Figure 2 shows p_{vc} as a function of t at $\delta=0$, $|\lambda|=0.01$ eV, and $U_v+U_c=0.1$ eV (the numerical calculations indicate that $p_{vc}(t)$ is only determined by the sum U_v+U_c, at least when $U_v+U_c\ll\varepsilon_c-\varepsilon_v$). Figure 2 demonstrates that the analytic solution is in good agreement with numerical calculations, as expected, since the selected values of the system parameters satisfy (34). The numerical calculations yield $p_{vc}^{\max}$ and T equal to 0.9935 and $1.4606\cdot 10^{-13}$ s, respectively, and the analytic solution yields 1 and $\pi/|\lambda|\sqrt{2}=1.4622\cdot 10^{-13}$ s, respectively.

Figures 3 and 4 illustrate the transition from the ''regime of noninteracting electrons'' ($U_v+U_c=0$, $p_{vc}^{\max}=1/2$, and $T=\pi/2|\lambda|=1.0339\cdot 10^{-13}$ s; see Eq. (5)) to the ''regime of strong Coulomb repulsion,'' for which $U_v+U_c\gg|\lambda|$ (Eq. (35)). This transition occurs when $U_v+U_c\approx\lambda$; when $U_v+U_c\geqslant 3\lambda$, the probability $p_{vc}^{\max}$ is larger than 0.95 and T is within 3.5% of its asymptotic value corresponding to $U_v+U_c=\infty$.

We now discuss the effect of an offset from resonance ($\delta\neq 0$) on $p_{vc}^{\max}$ and T. The corresponding curves are given in Figs. 5 and 6 for $U_v+U_c=0.1$ eV and $|\lambda|=0.01$ eV, which satisfy the second inequality in (35). One can see that the analytic solution is in excellent agreement with the numerical calculations even when the first inequality in (35) is not satisfied. Here, of course, $p_{vc}^{\max}\approx 1$ only when $|\delta|\ll|\lambda|$.

Figures 7 and 8 show $p_{vc}^{\max}$ and T calculated in the immediate neighborhood of resonance for $U_v+U_c=0.01$ eV

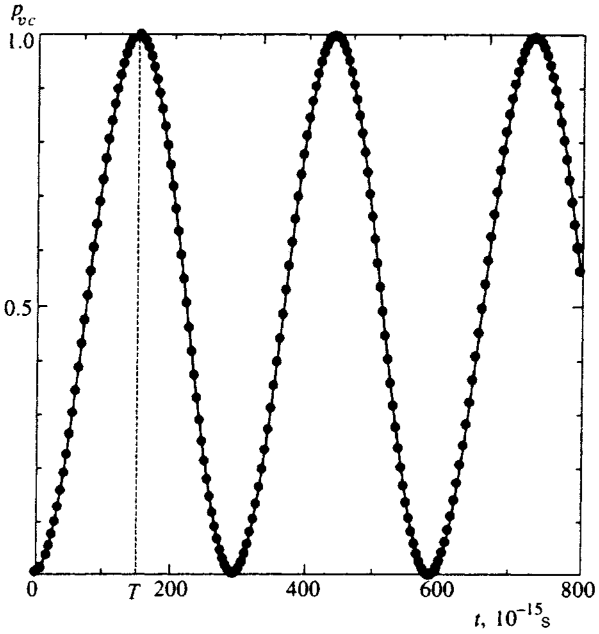

FIG. 2. Probability p_{vc} of one-electron transition from the valence band to the conduction band versus duration t of action of a periodic perturbation with frequency $\Omega=\varepsilon_c-\varepsilon_v-U_v$ ($\delta=0$); $\varepsilon_c-\varepsilon_v=1$ eV, $U_v+U_c=0.1$ eV, $|\lambda|=0.01$ eV. Numerical calculations with Eqs. (23)–(28) are plotted as dots, and the analytic solution defined by Eqs. (30)–(32) in the ''three-level approximation'' is shown by the solid line.

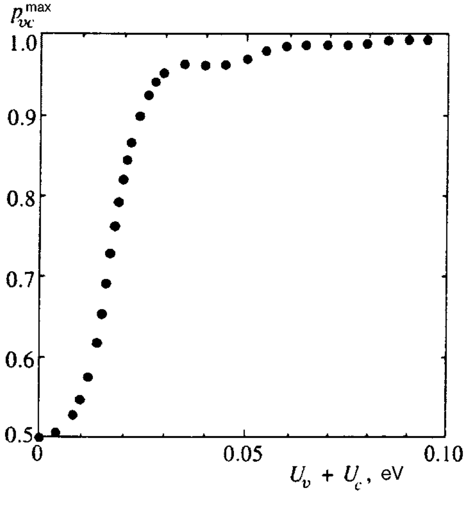

FIG. 3. Maximum probability $p_{vc}^{\max}$ of one-electron transition from the valence band to the conduction band due to a periodic perturbation with frequency $\Omega=\varepsilon_c-\varepsilon_v-U_v$ ($\delta=0$) as a function of the sum U_v+U_c of Coulomb repulsion energies in the valence band and conduction band; $\varepsilon_c-\varepsilon_v=1$ eV, $|\lambda|=0.01$ eV. The probability was calculated numerically using Eqs. (23)–(28).

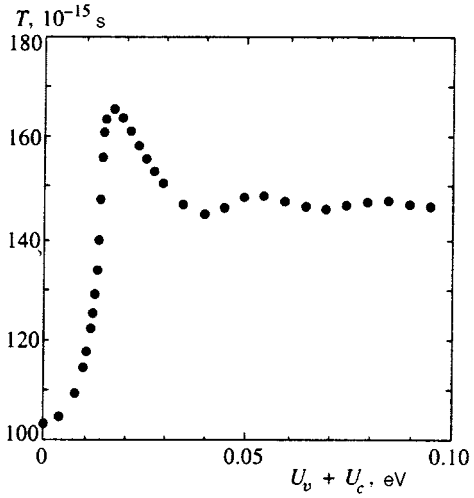

FIG. 4. Time T corresponding to the maximum probability $p_{vc}^{\max}$ of one-electron transition from the valence band to the conduction band due to a periodic perturbation with frequency $\Omega=\varepsilon_c-\varepsilon_v-U_v$ ($\delta=0$) as a function of the sum U_v+U_c of the Coulomb repulsion energies in the valence band and conduction band; $\varepsilon_c-\varepsilon_v=1$ eV, $|\lambda|=0.01$ eV. The time was calculated numerically using Eqs. (23)–(28).

and $|\lambda|=0.0001$ eV. For $|\delta|\leqslant 10^{-10}$ s^{-1}, we have $p_{vc}^{\max}\geqslant 0.998$, i.e., the transition probability is essentially indistinguishable from unity. In this region, T varies between $1.4607\cdot 10^{-11}$ and $1.4622\cdot 10^{-11}$ s, i.e., by 0.1%.

4. EFFECT OF NONRESONANT COMPONENTS

In the resonant approximation we neglected several ''nonresonant'' harmonics of the applied periodic field (see Sec. 3). To assess their effect on the final result, we numerically solved the time-dependent Schrödinger equation (13) with Hamiltonian (9) using the explicit Runge–Kutta method of order 7(8) based on the equations of Dormand and Prince with automatic selection of the step size.[22] We calculated the coefficients $A_i(t)$ in Eq. (11) with initial conditions $A_i(0)=\delta_{i1}$, and determined the value of $p_{vc}(t)=|A_2(t)|^2+|A_3(t)|^2$. As follows from our calculations, the function $p_{vc}(t)$ determined in the nonresonant approximation in the parameter range of interest (see Eq. (35)) coincides with that found analytically in the ''three-level resonant approximation'' based on Eqs. (30)–(32) to within a fraction of a percent. It is noteworthy, however, that when nonresonant terms are taken into account, the probability $p_{vc}^{\max}$ peaks at a very small but finite offset δ. The good agreement between the numerical and analytic results sug-

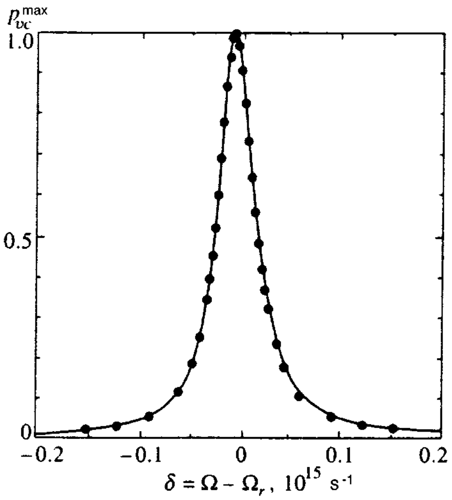

FIG. 5. Maximum probability $p_{vc}^{\max}$ of a one-electron transition from the valence band to the conduction band due to a periodic perturbation with frequency Ω as a function of $\delta=\Omega-\Omega_r$, where $\Omega_r=\varepsilon_c-\varepsilon_v-U_v$ is the resonant frequency; $\varepsilon_c-\varepsilon_v=1$ eV, $U_v+U_c=0.1$ eV, $|\lambda|=0.01$ eV. The dots show numerical calculations using Eqs. (23)–(28). The analytic solution of Eqs. (30)–(32) in the ''three-level approximation'' is plotted as a solid line.

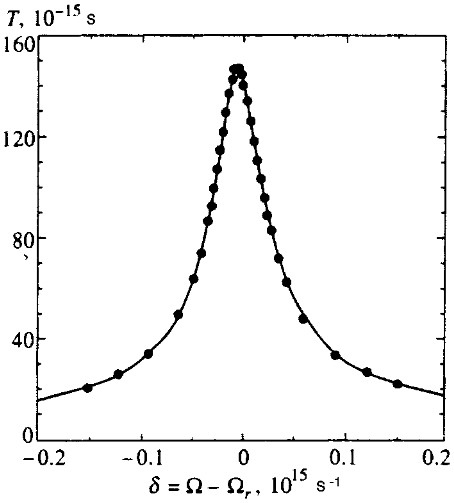

FIG. 6. Time T corresponding to the maximum probability $p_{vc}^{\max}$ of one-electron transition from the valence band to the conduction band due to a periodic perturbation with frequency Ω as a function of $\delta=\Omega-\Omega_r$, where $\Omega_r=\varepsilon_c-\varepsilon_v-U_v$ is the resonant frequency; $\varepsilon_c-\varepsilon_v=1$ eV, $U_v+U_c=0.1$ eV, $|\lambda|=0.01$ eV. The dots show numerical calculations using Eqs. (23)–(28). The analytic solution defined by Eqs. (30)–(32) in the ''three-level approximation'' is plotted as a solid line.

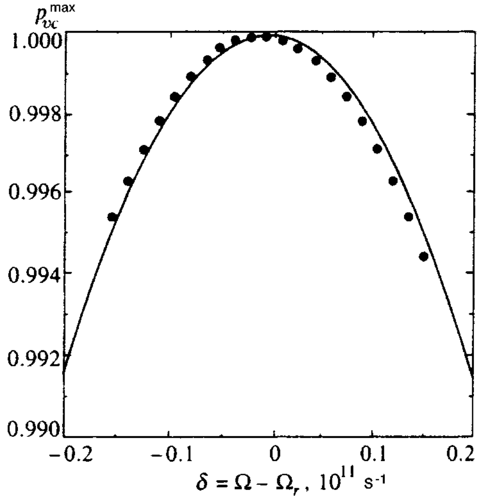

FIG. 7. Same as Fig. 5 with $U_v+U_c=0.01$ eV, $|\lambda|=0.0001$ eV.

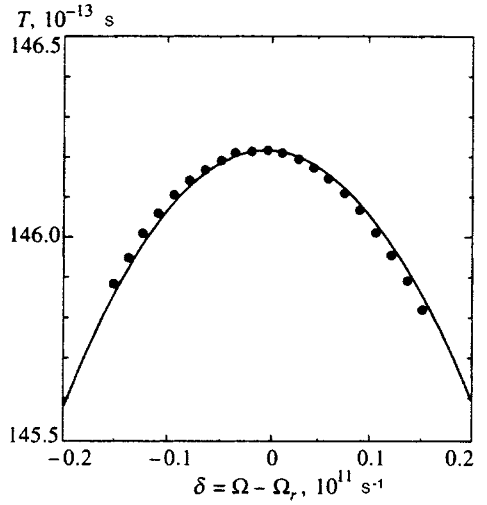

FIG. 8. Same as Fig. 6 with $U_v + U_c = 0.01$ eV, $|\lambda| = 0.0001$ eV.

gests that the nonresonant summands in the Hamiltonian (9) have little effect on system evolution. This result, however, was predictable for $|\lambda| \ll \Omega$.[14]

5. DISCUSSION OF RESULTS AND CONCLUSIONS

Thus, we have solved the problem of interaction between an hf electric field and a quantum dot with two size-quantized levels (one in the valence band and the other in the conduction band) and two electrons, taking into account Coulomb repulsion between the electrons. In the parameter range of interest, the problem has an easily interpretable analytic solution (30)–(32) in the so-called resonant three-level approximation (Sec. 3). We have shown that neither taking account of all two-electron levels of the quantum dot in the resonant approximation (Sec. 3) nor numerically solving the time-dependent Schrödinger equation with due account of ''nonresonant'' harmonics of the applied electric field (Sec. 4), have any perceptible effect on the time dependence of the energy-level populations.

With regard to the form of the model Hamiltonian (9), it does not include terms corresponding to the Coulomb interaction U_{vc} between two electrons in different levels $|v\rangle$ and $|c\rangle$ (if the valence band states are described in terms of holes, this interaction corresponds to excitonic effects). If the wave functions of the one-electron states $|v\rangle$ and $|c\rangle$ have different symmetry properties (as in GaAs/AlGaAs), the interaction energy U_{vc} should be significantly less than U_v and U_c. But even if it is not, it can be shown[18] that U_{vc} leads only to a shift in the resonant frequency, which becomes $\Omega_r = \varepsilon_c - \varepsilon_v - U_v + U_{vc}$, while the system dynamics remain unchanged. The frequency Ω_r of a specific quantum dot should be determined empirically.

Thus, we have shown that it is possible, in theory, to select the parameters of an external perturbation (Ω, $|\lambda|$, and T) and of a quantum dot ($U_v + U_c \approx e^2/\varepsilon a$) so that the probability of driving one electron from the valence band to the conduction band is extremely close to unity. Significantly, the time of this transition, $T = 10^{-13}$–10^{-11} s, is several orders of magnitude shorter than the recombination time $\tau = 10^{-10}$–10^{-9} s. Therefore, after the applied field is off (at $T \geqslant t$), the quantum dot is in a quasistationary excited state $|v\uparrow, c\downarrow\rangle$ or $|v\downarrow, c\uparrow\rangle$ with lifetime τ, after which the electron–hole pair will annihilate with generation of a one-photon wave packet $|1\rangle_f$ (Eq. (3)) with spectral distribution $f(\omega)$ centered at the frequency $\omega_0 = (2\varepsilon_c + U_v) - (\varepsilon_c + \varepsilon_v)$, i.e., near the resonant frequency $\Omega_r = \varepsilon_c - \varepsilon_v - U_v$ of the external field applied to the quantum dot. The process of photon generation is certainly probabilistic, so we cannot give the exact time when the photon is emitted. An important point is that if the recombination is radiative, only one photon will be emitted. This source of one-photon states can be used in quantum systems for data transmission and processing, including quantum cryptosystems.

In conclusion, note that the estimates of the required parameters of the quantum dot and external perturbation provide evidence in favor of the feasibility of such a source of one-photon states. Moreover, if the external field frequency cannot be varied over a broad band, the source resonance condition can probably be achieved by varying the energy difference between the size-quantized levels in the valence and conduction bands via application of static electric and magnetic fields to the quantum dot.

The authors are grateful to V. F. Elesin for discussion of the results of this work.

This work was supported by the Russian Fund for Fundamental Research (Project No. 96-02-18918).

[1] Trudy Fiz. Inst. Akad. Nauk **200** (1991).
[2] Z. Y. Ou and L. Mandel, Phys. Rev. Lett. **61**, 50 (1988).
[3] J. G. Rarity and P. B. Tapster, Phys. Rev. Lett. **64**, 2495 (1991).
[4] A. K. Ekert, J. G. Rarity, P. R. Tapster, and G. M. Palma, Phys. Rev. Lett. **69**, 1293 (1992).
[5] C. Marand and P. D. Townsend, Opt. Lett. **20**, 1695 (1995).
[6] S. J. D. Phoenix and P. D. Townsend, Contemp. Phys. **36**, 165 (1995).
[7] R. J. Hughes, D. M. Alde, P. Dyer, G. G. Luther, G. L. Morgan, and M. Schauer, Contemp. Phys. **36**, 149 (1995).
[8] C. H. Bennett, Phys. Today **48**, 24 (1995).
[9] W. K. Wotters and W. H. Zurek, Nature (London) **299**, 802 (1982).
[10] U. M. Titulaer and R. J. Glauber, Phys. Rev. **145**, 1041 (1966).
[11] H. Fearn and R. Loudon, Opt. Commun. **64**, 485 (1987); H. Fearn and R. Loudon, J. Opt. Soc. Am. B **6**, 917 (1989).
[12] R. A. Campos, B. E. Salech, and M. Teich, Phys. Rev. A **42**, 4127 (1990).
[13] A. Imamoglu and Y. Yamamoto, Phys. Rev. Lett. **72**, 210 (1994).
[14] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, 3rd ed., Pergamon Press, New York (1977).
[15] S. Flügge, *Practical Quantum Mechanics II*, Springer-Verlag, Berlin–Heidelberg–New York (1971), p. 180.
[16] V. M. Galitskii and V. F. Elesin, *Resonant Interaction between Electromagnetic Fields and Semiconductors*, Énergoatomizdat, Moscow (1986); V. M. Galitskii, S. P. Goreslavskii, and V. F. Elesin, Zh. Éksp. Teor. Fiz. **57**, 207 (1969) [Sov. Phys. JETP **30**, 117 (1970)].
[17] É. A. Manykin and A. I. Maimistov, *Interaction between Radiation and Matter: Coherent Processes*, MIEP Press, Moscow (1996).
[18] S. N. Molotkov, JETP Lett. **64**, 237 (1996).
[19] B. Meurer, D. Heitmann, and K. Ploog, Phys. Rev. Lett. **68**, 1371 (1992).
[20] J. Hubbard, Proc. R. Soc. London, Ser. A **276**, 238 (1963).
[21] Yu. A. Izyumov, Usp. Fiz. Nauk **165**, 403 (1995).
[22] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations. Nonstiff Problems*, Springer-Verlag, Berlin–Heidelberg (1987).

Translation was provided by the Russian Editorial office.

Laser photoelectron projection microscopy of insulating samples with sub-wavelength spatial resolution

S. K. Sekatskiĭ

Institute of Spectroscopy, Russian Academy of Sciences, 142092 Troitsk, Moscow Region, Russia

(Submitted 29 April 1997)
Zh. Éksp. Teor. Fiz. **112**, 1273–1284 (October 1997)

Operation of a laser-driven photoelectron projection microscope is analyzed in the case of low-conductivity samples. Under certain conditions, even with ''perfectly'' insulating samples, one can obtain photoelectron images with a high magnification factor and spatial resolution smaller than the laser wavelength owing to the effect that the diple field of a sample polarized by a strong external electric field has on the motion of emitted photoelectrons. The experimental investigation of lithium fluoride and glass samples reported in this paper can be considered a qualitative confirmation of the suggested model. © *1997 American Institute of Physics.* [S1063-7761(97)00910-4]

1. INTRODUCTION. PROBLEM STATEMENT

The study of various characteristics of samples with superhigh spatial resolution is a central problem of surface physics, microelectronics, biophysics, etc. Researchers' attention has been focused on experimental techniques that provide simultaneously high spatial resolution and spectral (chemical) selectivity, i.e., they allow one not only to visualize the surface topography, but also to identify various microstructures observed in a sample. The ultimate goal of these studies is to identify individual molecules and even parts of complex molecules.

One of the feasible techniques that can be applied to such studies is resonant photoelectron (photoion) microscopy, which was suggested in 1975.[1] Recently this technique was implemented for the first successful experiments in this field and allowed the researchers to detect, in particular, isolated F_2 color centers on a LiF surface.[2,3]

The underlying concept of the resonant laser photoelectron microscopy[2,3] (Fig. 1a) is recording laser-induced photoelectron images of samples shaped as sharp needles with small curvature radii r of their tips. The conditions of laser irradiation of tested samples (tips) are selected so that the external photoeffect in the sample material due to the laser light is present,[4] i.e., only selected centers in the sample absorbing the incident light are ionized. If photoelectrons are detected by a position-sensitive detector placed at distance L_1 from the sample, one can obtain a high magnification factor of the order of L_1/r and visualize individual absorbing centers.

Projection photoelectron laser microscopy has an additional advantage because, unlike the traditional techniques of field electron (field ion) microscopy, which is widely used in studies of surface topography and electrophysical properties of various metals, alloys, and some semiconductors with a high spatial resolution (see Refs. 5 and 6, and references therein), this method can be also applied to studies of dielectric materials. This can be done because the source of electrons in this case is external photoeffect, and the external electric field only conveys them to the detector, therefore there is no need for the superhigh electric fields necessary for efficient field emission of electrons or field ionization of imaging gas atoms in traditional field electron (field ion) microscopic techniques.[5,6]

If we feed to a metallic electrode supporting the needle (Fig. 1) a voltage V of several kilovolts, the ohmic voltage drop IR across the sample can take the most part of this voltage, whereas a voltage of 100–200 V at the tip (at that area with radius r from which electrons are emitted) is sufficient for fairly efficient detection of emitted photoelectrons. Given that the total resistance of a conic needle with top angle ϑ_0 is about $R \simeq \rho/\pi r \tan \vartheta_0$, where ρ is the needle material resistivity, and the voltage drop across the needle is about 3 kV, whereas the emitted current sufficient for recording photoelectron images with high resolution and a good signal-to-noise ratio is 0.1 pA, one can easily check that high-resolution projection photoelectron imaging of materials with resistivities of up to $\rho \sim 10^{12} \Omega \cdot \mathrm{cm}$ (and in some case even higher) are feasible. This conclusion is supported by recent experimental investigations of dielectric tips from lithium fluoride containing F_2 color centers.[2,3]

It obviously follows from this discussion that the possibility of imaging dielectric samples using the photoelectron projection microscopy, hence recording all basic parameters of photoelectron images, depends on many parameters, such as the sample conductivity, applied high voltage, collected photoemission current, etc., and investigation of this dependence is of considerable practical interest. One limiting case corresponding to a relatively small sample resistivity is very similar to the case of metallic samples studied in a field electron/ion microscope, which has been investigated in detail. This paper analyzes an opposite limiting case, namely, when samples are ''perfectly'' insulating and contain no free charges, which lead to concentration of electric lines of force around the tip.[5,6] The only field-induced effect is sample polarization (Fig. 1b). It will be shown that a considerable magnification factor and a spatial resolution smaller than the laser wavelength can be also achieved in this configuration. Experimental investigations of tips fabricated from lithium fluoride dielectric crystals and glass using the projection photoelectron microscopy technique can be considered as a confirmation of theoretical results.

1063-7761/97/100690-07$10.00

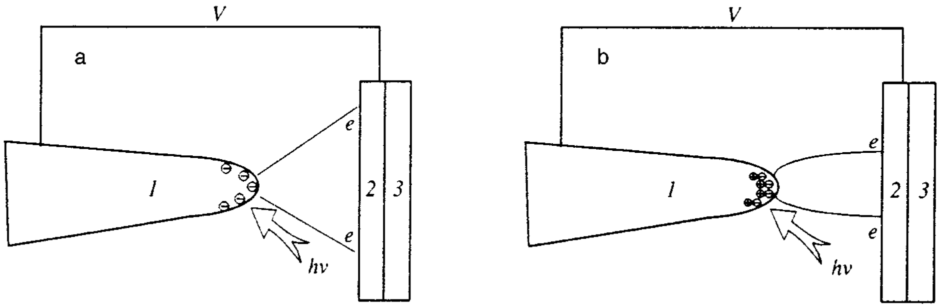

FIG. 1. Diagram of a laser photoelectron projection microscope: (a) conducting sample; (b) insulating sample ((*1*) dielectric sample; (*2*) microchannel plate; (*3*) luminescent screen).

2. MOTION OF PHOTOELECTRONS IN ELECTRIC FIELD DUE TO POLARIZED DIELECTRIC SAMPLES

Let us consider a simple model illustrated by Fig. 2. An insulating ball with radius r_0 and dielectric permittivity ϵ is placed between the plates of a plane capacitor with distance L between them and electric field $\mathbf{E}$. In this field, the ball obtains dipole moment $\mathbf{p}$ equal to[7]

$$\mathbf{p}=r_0^3\frac{\epsilon-1}{\epsilon+2}\mathbf{E}, \tag{1}$$

and the electric field outside the ball is calculated as $\mathbf{E}_{\text{ext}}=\mathbf{E}+\mathbf{E}_{\text{dip}}$, where $\mathbf{E}_{\text{dip}}$ is

$$\mathbf{E}_{\text{dip}}=\frac{3(\mathbf{pr})\mathbf{r}}{r^5}-\frac{\mathbf{p}}{r^3}. \tag{2}$$

By considering the meridional plane of the ball and introducing polar angle ϑ (Fig. 2), one can easily obtain for components E_x and E_y of the field E_{ext} outside the ball, i.e.,

$$E_x=\frac{3}{2}r_0^3\frac{\epsilon-1}{\epsilon+2}E\frac{\sin 2\vartheta}{r^3}, \tag{3}$$

$$E_y=E+r_0^3\frac{\epsilon-1}{\epsilon+2}E\frac{3\cos^2\vartheta-1}{r^3}. \tag{4}$$

Let us consider photoelectrons emitted with the initial kinetic energy equal to zero. At small angles ϑ we can assume that the distance $r(t)$ between a moving electron and the ball center as a function of time is fully determined by the field component E_y. Besides, if we assume that $E_y=E$, we can easily derive for the function $r(t)$ (e and m are the electron charge and mass)

$$t=\sqrt{\frac{2(r-r_0)m}{eE}}, \tag{5}$$

$$\frac{dr}{dt}=\sqrt{\frac{2(r-r_0)eE}{m}}. \tag{6}$$

Then, using the equations

$$\frac{dv_x}{dt}=\frac{eE_x}{m},\quad \frac{dv_x}{dt}=\frac{dv_x}{dr}\frac{dr}{dt},$$

substituting into them expressions (3) and (6), and integrating the resulting equation, we obtain the following expression for the velocity component v_x of an electron moving in the dipole field:

$$v_x=\beta r_0^3\vartheta\int_{r_0}^{\infty}\frac{dr}{\sqrt{r-r_0}\,r^3}, \tag{7}$$

where

$$\beta\equiv\frac{3(\epsilon-1)}{\epsilon+2}\sqrt{\frac{eE}{2m}},$$

and $\sin 2\vartheta$ is replaced with 2ϑ. The expression for the integral in Eq. (7) is given in reference books,[8] and after substituting its value $3\pi/8r_0^{5/2}$ we have

$$v_x=\frac{9\pi}{8}\frac{\epsilon-1}{\epsilon+2}r_0^{1/2}\vartheta\sqrt{\frac{eE}{2m}}. \tag{8}$$

Almost all the velocity v_x of an emitted electron is imparted in the close vicinity of the polarized ball, therefore the resulting electron displacement Δl along the x-axis can be expressed as

$$\Delta l=v_x t=\frac{9\pi}{8}\frac{\epsilon-1}{\epsilon+2}\vartheta\sqrt{r_0 L}, \tag{9}$$

where $t=\sqrt{2Lm/eE}$ is the total electron transit time across the capacitor. Thus the microscope magnification factor M is

$$M=\frac{\Delta l}{\vartheta r_0}=\frac{9\pi}{8}\frac{\epsilon-1}{\epsilon+2}\sqrt{\frac{L}{r_0}}. \tag{10}$$

After substituting into Eq. (9) $\epsilon=9$, which is the dielectric permittivity of lithium fluoride studied in the experiments,[9] $L=10$ cm, and $r_0=3$ μm, one can see that the magnification factor of the photoelectron microscope can be extremely high (about 470). Note also that the magnification factor M

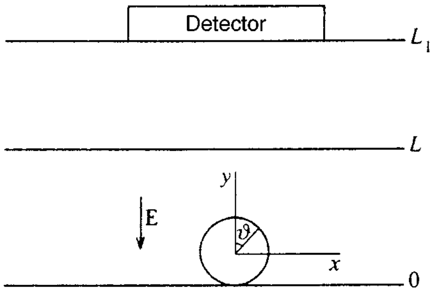

FIG. 2. Configuration for which motion of emitted electrons driven by the field generated by a polarized ball has been calculated (this diagram is not an accurately scaled drawing of the model).

is proportional to the square root of L/r_0, whereas in the case of conducting samples the corresponding function is linear:

$$M = L/\chi r_0, \tag{11}$$

where χ is a numerical factor depending on the system configuration and ranging between 1.5 and 2.[5]

Numerical calculations of the motion of photoelectrons with zero initial kinetic energy driven by the dipole field confirm that the magnification factor is proportional to the square root of L/r_0, whereas the absolute values of M proved to be a factor of about 1.5 smaller than those given by Eq. (10) (specifically, in the example discussed above $M = 310$), which is not surprising since the derivation of Eq. (10) contains a lot of simplifications. Moreover, the numerical calculations indicate that the magnification factor M depends on the initial photoelectron coordinate or angle ϑ, as shown in Fig. 3. This dependence leads to ambiguities in resulting photoelectron images of a polarized dielectric ball. Only the region of the sample surface where $0 \leqslant \vartheta \leqslant 0.57$ rad is mapped one-to-one, whereas in images of regions with $0.57 \leqslant \vartheta \leqslant \pi/2$ rad each point corresponds to two points on the sample surface. From the experimental viewpoint, this ambiguity can give rise to additional (noncentral) peaks in photoelectron images of polarizing dielectric samples, and this effect should be taken into account in interpretation of photoelectron images.

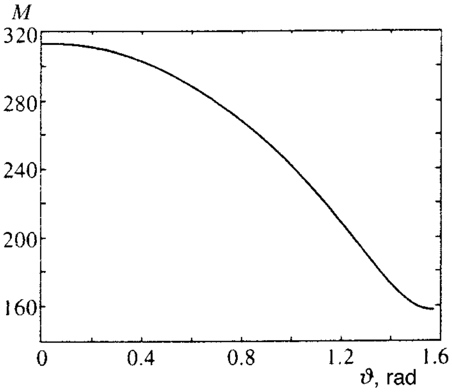

FIG. 3. Magnification factor of photoelectron images of a dielectric ball as a function of the angular coordinate of emitted electrons.

The magnification factor of the studied system can be increased considerably if after passing across a capacitor with distance L between its plates photoelectrons pass through a zero-field region with length $L_1 \gg L$, as shown in Fig. 2. For example, numerical calculations of the photoelectron motion in such a system in the case of $L = 1$ cm, $L_1 = 30$ cm, and $r_0 = 10$ μm yield at small angles ϑ the magnification factor $M = 870$. (Numerical estimates indicate that the distortions due to the grid separating the regions of L and L_1 are negligible if the mesh width is about ten micrometers or smaller.) The maximum value of the image coordinate is achieved at an angle $\vartheta_{\text{opt}} \simeq 1$ rad and equals 5.6 mm.

The microscope spatial resolution is determined, as in the case of a projection photoelectron microscope with con-

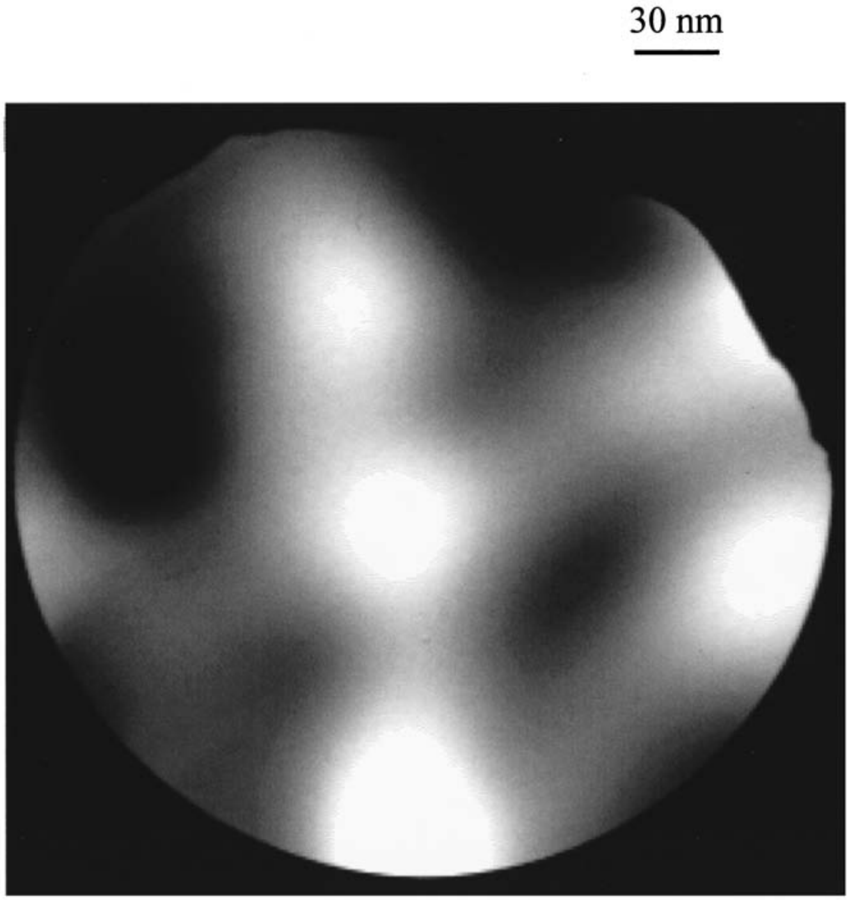

FIG. 4. Photoelectron image of a LiF:F_2 tip with a magnification factor of about 10^5. The tip curvature radius is 0.6 μm, the applied potential 2.5 kV. The bright white spots on the pattern correspond to isolated color centers on the tip surface.

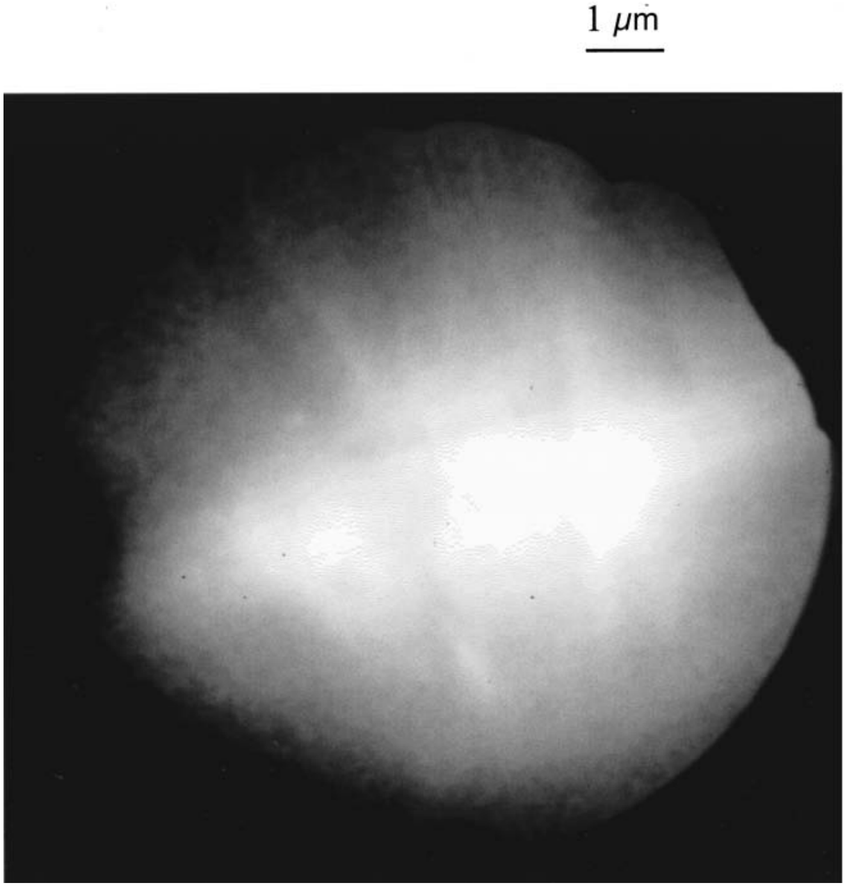

FIG. 5. Photoelectron image of a LiF:F_2 tip with a magnification factor of about 2500. The tip has an elongated shape with a large curvature radius of about 4 μm, its potential is 500 V.

ducting samples, by the initial spread of the photoelectron kinetic energy ΔE_0. The numerical calculation of motion of electrons with finite ΔE_0 has shown that a high resolution can be achieved only in the case of very small initial kinetic energies of electrons corresponding to liquid-helium temperatures. For example, under the conditions defined above at angle $\vartheta = 0.5$ rad, the electric field strength in the capacitor $E = 10$ kV/cm, and $\Delta E_0 = 10$ meV, the spatial resolution can be about 0.8 μm, whereas at $\Delta E_0 = 1$ meV the resolution is as small as 0.26 μm, etc.

3. EXPERIMENTAL OBSERVATION OF PHOTOELECTRON IMAGES OF DIELECTRIC TIPS

The analysis discussed above indicates that a considerable spatial magnification (and in some cases a spatial resolution smaller than the light wavelength) can be achieved in the technique of photoelectron microscopy of insulating samples. Naturally, the situation discussed above, when emitted photoelectrons are driven only by the dipole field generated by the polarized sample and the total free charge in the system is zero, is, in a sense, unrealistic. In the analysis of real photoelectron images of low-conductivity samples, one should take into account both the dipole field component and the component due to free electron charges, and also a number of other factors not discussed in the paper (for example, effects of the field penetration depth as a function of applied voltage and collected photocurrent, etc.), which make the problem extremely difficult. Moreover, the configuration of real experiments on photoelectron spectroscopy of dielectric samples (Fig. 1) is notably different from the model configuration given in Fig. 2.

Nonetheless, results of a set of experiments with lithium fluoride and glass samples can be considered, in our opinion, as a qualitative confirmation of the discussed calculations. In the first set of experiments (the results will be described in detail elsewhere), the photoelectron images of tips fabricated from lithium fluoride crystals with different concentrations of color centers were studied under radiation from an argon laser with a pumping power density 10^3-10^4 W/cm^2. The samples were exposed to all spectral lines generated by the argon laser or to separate lines with wavelengths of 488 and 514 nm. LiF crystals with color centers were manufactured at Institute of General Physics, Russian Academy of Sciences. Tips were fabricated by etching fragments of crystals in concentrated hydrochloric acid. Curvature radii of the tips ranged between 0.6 and 5 μm, the voltage fed to the samples was 0–4 kV. The detector was an assembly of a microchannel plate, a luminescent screen, and a fiber-optic plate, whose output was recorded by a CCD camera and fed to a computer, where the image was processed by a dedicated Argus-50 processor manufactured by Hamamatsu Photonics K. K. (Japan). The working area of the microchannel plate was 32 mm in diameter, thus, given the separation between the tip and plate of 10 cm, the angle from which photoelectrons were collected was about 20°.

Depending on the experimental parameters, such as the tip potential, photoemission current etc., substantially differ-

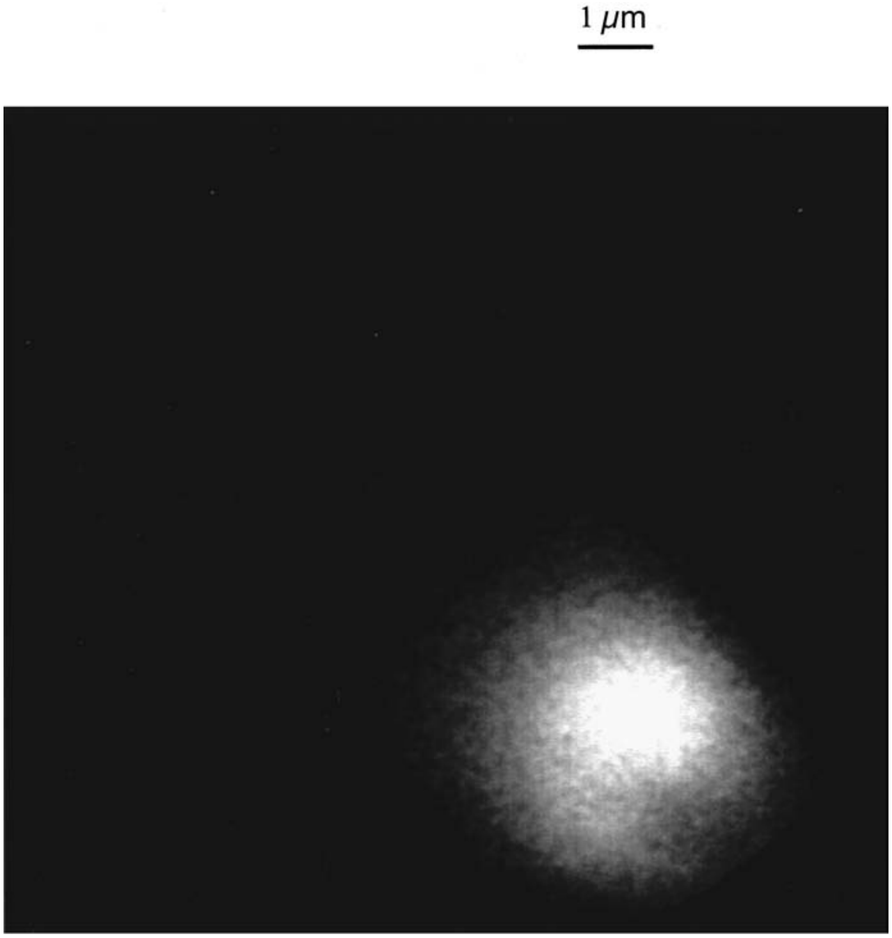

FIG. 6. Photoelectron image of a tip fabricated from KS-11 optical glass with a magnification factor of about 2000.

ent photoelectron images of LiF:F_2 tips were recorded. A typical photoelectron image of a high spatial resolution displaying isolated F_2 color centers is shown in Fig. 4. An interpretation of such photoelectron images of LiF:F_2 tips, as well as a more detailed description of experiments, is beyond the scope of this paper and can be found elsewhere.[2,3] The photocurrent as a function of the light intensity I fed to the tips is also not analyzed in this publication. Note only that this function is clearly nonlinear, but its accurate measurements were difficult because the dynamic range of the microchannel plate is limited, and the image shape depends on the total collected photocurrent. Another important point is that, in addition to high-resolution photoelectron images similar to that in Fig. 4 and, as follows from the analysis, adequately described by formulas applicable to experiments with conducting tips, the low-resolution photoelectron images similar to that shown in Fig. 5 were detected for some samples at small voltages applied to the tip (about 500 V).

Unlike Fig. 4, the magnification factor of the system in Fig. 5 is about an order of magnitude smaller, and the detector images the entire LiF tip, but not its small fraction. As follows from comparison with images of the same samples obtained by optical and electron microscopes, these low-resolution electron images, all in all, faithfully reproduce the tip shape (such as the tip shown in Fig. 5, which is really much longer in one direction). Moreover, these images often contain additional features (such as three photocurrent maxima in the case of Fig. 5), which could not be detected by studying these tips using other techniques.

In our opinion, such photoelectron low-resolution images should be considered as patterns which are largely due to the polarization component of the electric field generated in the system. This assumption is supported by the following arguments.

Above all, note that the lithium fluoride resistivity at room temperature is extremely high. As follows from our measurements and data given in Refs. 10 and 11, $\rho \sim 10^{14}-10^{15} \Omega \cdot \mathrm{cm}$. This value is much higher than the estimate of the resistivity of samples suitable for studies by traditional projection photoelectron spectromicroscopy, namely $\rho \sim 10^{12} \Omega \cdot \mathrm{cm}$, i.e., generation of ''polarization'' photoelectron images of LiF tips is quite feasible under certain conditions. (The lithium fluoride resistivity rapidly drops with temperature, and at $T = 250\,^\circ\mathrm{C}$ it is $\rho \sim 10^{9} \Omega \cdot \mathrm{cm}$.[11,12] This fact and expected effects of intense electric field and photoconductivity allow one to image LiF sample surfaces with a high resolution under different conditions.[3]) Furthermore, the magnification factor in this case is within one order of magnitude different from the factor calculated by the simplest model of a polarized ball (this model, as was noted above, cannot be considered as qualitatively accurate) and differs by more than one order of magnitude from the magnification typical of a conducting tip. Photoelectrons are emitted not into a solid angle close to $\pi/2$, as in the case of a conducting tip, but into a relatively narrow angle (less than 10°), which is in agreement with simple models of photo-

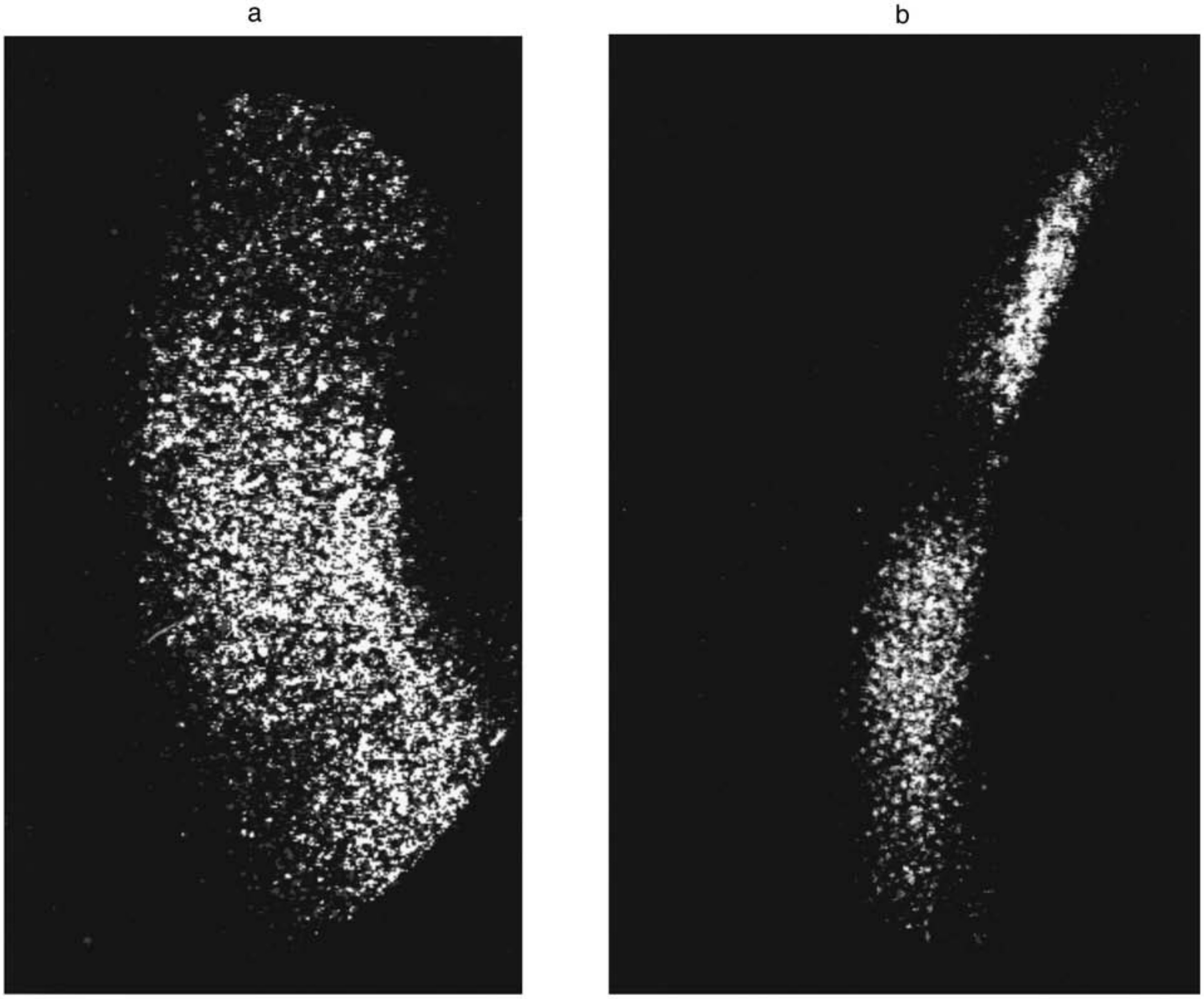

FIG. 7. Photoelectron images of sections of the surface of a conical needle fabricated from KS-11 glass obtained by focusing radiation from a copper-vapor laser on the sample surface at some distance from the tip: (a) 0.3 mm; (b) 1.0 mm.

electron motion in fields generated by polarized dielectrics.

An increase in the magnification factor with the decreasing tip curvature radius r and a growth in the total image dimension with increasing r (Eq. (9)) predicted by these models have been detected in experiments. Finally, an important point is that low-resolution images were recorded only when the tip potential was relatively low. A ''spread'' of images and transition to images similar to that in Fig. 4 occurred abruptly when the potential was increased. This behavior is predicted by simple models that take into account the voltage drop across the needle as a function of the total current across the sample and applied voltage.

Similar results were obtained in another set of experiments conducted using tips fabricated by etching needle-shaped fragments of red KS-11 and KS-15 optical filters in concentrated hydrofluoric acid and illuminating them with a copper vapor laser. The curvature radii of these tips were 1–3 μm, the duration of optical pulses generated by the copper-vapor laser was 18 ns, the repetition rate was 8.8 kHz, and the pumping power density was 10^7-10^8 W/cm^2. Both spectral lines with wavelengths of 511 and 572 nm generated by the laser were simultaneously used in experiments.

The resistivity of these glass samples at room temperature is extremely high (up to $\rho \sim 10^{15} \Omega \cdot \mathrm{cm}$) and rapidly drops with temperature,[13] which allows one, as with LiF tips, to obtain photoelectron images with either low or high resolution. Images of each sort have been really observed in experiments. A typical example of a low-resolution image is given in Fig. 6. As in the case of LiF samples, the image faithfully reproduces the overall tip shape, and its overall dimension also depends on the tip curvature radius.

In order to dispel all doubts concerning the feasibility of the photoelectron projection microscopy of dielectric samples and the fidelity of the recorded images, Fig. 7 shows images of glass tips recorded with a very low resolution. In this case, light generated by the copper-vapor laser was focused not on the tip apex, but on the conical surface of the sample at a small distance (about 0.3–1 mm) from the tip. It is beyond doubt that the graph shows the photoelectron images of rings on the conic sample surfaces exposed to the laser light, and the magnification factor is different for different curvature radii of the samples.

4. CONCLUSIONS

The main result of the reported work is the feasibility of photoelectron images of insulating tips with a fairly high magnification factor and spatial resolution, which is confirmed by experiments with lithium fluoride and glass tips. It is clear that, in the limit of ‘‘perfectly’’ insulating tips the magnification factor is much smaller than the spatial magnification factor characteristic of field-emission projectors with metallic tips and is described by a different function (square root) of the ratio between characteristic system dimensions R/r_0.

A detailed and consistent analysis of the transition between the regime of the photoelectron microscope using electric field generated by free charges to the ‘‘polarization’’ regime should take into account all the electric field components and the real system configuration, and demands additional experiments with samples of low conductivity. This research is of great interest, and remarkable progress in this field can be expected in the immediate future.

The work was supported by the Russian Fund for Fundaental Research and the U.S. Department of Defense. The author is grateful to V. S. Letokhov for support of this work and helpful discussions, to V. N. Konopskiĭ for help in experiments with LiF samples, T. T. Basiev for supplying lithium fluoride crystals doped with color centers, and Hamamatsu Photonics K. K. (Japan) for supplying experimental equipment.

[1] V. S. Letokhov, Kvantovaya Élektronika **2**, 930 (1975) [Sov. J. Quantum Electron. **5**, 506 (1975)].

[2] V. N. Konopskiĭ, S. K. Sekatskiĭ, and V. S. Letokhov, Izv. Ross. Akad. Nauk, Ser. Fiz. **59(12)**, 13 (1995).

[3] V. N. Konopsky, S. K. Sekatskii, and V. S. Letokhov, Appl. Surf. Sci. **94/95**, 148 (1996); Opt. Commun. **132**, 251 (1996).

[4] V. S. Letokhov and S. K. Sekatskiĭ, Optika i Spektroskopiya **76**, 303 (1994).

[5] T. T. Tsong, *Atom-Probe Field Ion Microscopy*, Cambridge Univ. Press, Cambridge (1990).

[6] A. Modinos, *Field-Emission, Thermal, and Secondary-Electron Spectroscopy* [in Russian], Nauka, Moscow (1990).

[7] D. V. Sivukhin, *Electricity* [in Russian], Nauka, Moscow (1983), p. 70.

[8] G. B. Dwight, *Tables of Integrals and Other Mathematical Formulas* [Russian translation], Nauka, Moscow (1977), p. 43.

[9] *Acoustic Crystals* [in Russian], ed. by M. P. Shaskol'skaya, Nauka, Moscow (1982), p. 88.

[10] W. E. Nelson, J. F. Merklin, and R. S. Lee, Nucl. Instrum. Methods **133**, 387 (1976).

[11] A. G. Lipson, V. A. Kuznetsov, D. M. Sakov, and Yu. P. Toporov, Poverkhnost' No. 10–11, 42 (1992).

[12] S. C. Jain and G. P. Sootha, Phys. Status Solidi **22**, 505 (1967).

[13] E. B. Shand, *Glass Engineering Handbook*, McGraw-Hill, New York (1958).

Translation was provided by the Russian Editorial office.

Ionization self-channeling of whistler waves in a collisional magnetized plasma

A. V. Kudrin, L. E. Kurina, and G. A. Markov

Nizhniĭ Novgorod State University, 603600 Nizhniĭ Novgorod, Russia

(Submitted 18 February 1997)
Zh. Éksp. Teor. Fiz. **112**, 1285–1298 (October 1997)

An investigation is made of the self-interaction of whistler waves (whistlers) involving the formation of waveguide channels in a collisional magnetoactive plasma as a result of its additional ionization by the field of the propagating wave. Simplified equations are derived to describe the behavior of the whistler field in a channel of enhanced plasma density in the presence of electron collisions. Self-consistent distributions of the field and the plasma corresponding to steady-state ionization self-channeling of whistlers are obtained by numerically solving the equations for the field together with balance equations for the electron density and energy. Our estimates indicate that this effect can be observed under laboratory conditions. © *1997 American Institute of Physics.* [S1063-7761(97)01010-X]

1. INTRODUCTION

Characteristic features of the formation of waveguide channels in a magnetized plasma as a result of various nonlinear effects accompanying the propagation of high-power wave beams have been investigated by many authors (see, for instance, Refs. 1–5 and the cited literature). Recently, these topics have attracted increased interest with the setting up of various new laboratory and ionospheric experiments to produce self-consistent plasma structures, with controllable properties, that extend along an external magnetic field.[6–10] Of particular interest as applied to an ionospheric plasma is the possible self-channeling of high-intensity electromagnetic waves in the whistler frequency range

$$(\Omega_H\omega_H)^{1/2} < \omega \ll \omega_H \ll \omega_p, \tag{1}$$

where ω_H and ω_p are the electron gyrofrequency and plasma frequency, respectively, and Ω_H is the ion gyrofrequency. In particular, this explains why whistler waves (whistlers) found under natural conditions in near-Earth space play a very important role in many fundamental and applied problems in cosmic plasma physics.[11] Studies of the propagation characteristics and self-interaction of high-power wave beams in the whistler range are of considerable interest for the development of new methods of investigating the ionosphere and the magnetosphere by actively controlling the near-Earth plasma, and also for various other similar applications (see Refs. 6 and 12). Problems involving the efficient excitation of whistlers and transferring their energy to a plasma are attracting increased attention because of the scope for using these waves in existing laboratory and technological plasma facilities, such as helicon plasma sources.[13–16]

Note that most theoretical studies dealing with self-interaction of whistler waves mainly consider broad collisionless or weakly collisional channels on the wavelength scale of the propagating wave, with a small transverse plasma density gradient. These channels (density ducts) are usually formed in a magnetized plasma where the influence of striction or thermal nonlinear effects predominates.[5,7,10] However, the situations changes substantially when the background plasma is additionally ionized by a wave field. The enhanced-density channels formed under conditions of ionization nonlinearity usually have widths comparable to the whistler wavelength, and frequently exhibit an appreciable radial density gradient.[6] In addition, even in the collisionless limit, the propagation constants of the waveguide modes guided by these plasma formations are complex in the range (1) because of their leakage into the surrounding medium as small-scale quasielectrostatic waves.[5,17] This unusual feature of whistler channeling means that any mathematical description of their ionization self-interaction resulting in the formation of enhanced-density ducts presents major difficulties. In particular, no theory of ionization self-channeling of whistlers has yet been constructed for a plasma with comparatively low effective electron collision frequency ν_e, where the dominant mechanism of energy loss from the channel is emission of quasielectrostatic waves into the background plasma. However, at high enough frequencies ν_e, corresponding to appreciable collisional damping of quasielectrostatic waves, this effect can be analyzed using a simplified model (provided that the condition $\nu_e \ll \omega_H$ is also satisfied). Here we consider the steady-state ionization self-channeling of whistler waves for this special case.

2. BASIC EQUATIONS

A description of the complete picture of ionization self-channeling of whistlers is fairly difficult to achieve. However, the key issue of whether steady-state waveguide structures sustained by additional ionization of the background medium by trapped whistler waves can exist in a collisional magnetized plasma, together with the most characteristic features of this effect, can be investigated if we neglect, to a first approximation, the dependence of the plasma channel parameters and the absolute value of the wave field amplitude in this channel on the longitudinal coordinate, i.e., we confine our analysis to homogeneous plasma formations in the direction of the external magnetic field $\mathbf{H}_0 = H_0\mathbf{z}_0$. As we shall confirm subsequently, this model is quite adequate for comparatively weak whistler absorption $(\nu_e \ll \omega_H)$. We shall subsequently only consider axisymmetric distributions of the field and the plasma. In this case, with these idealizations the transverse structure of the whistler field

$$\frac{1}{2}[\mathbf{E}(\rho)\exp(i\omega t - ihz) + \text{c.c.}],$$

$$\frac{1}{2}[\mathbf{H}(\rho)\exp(i\omega t - ihz) + \text{c.c.}]$$

propagating along $\mathbf{H}_0$ is described by the following equations:

$$\nabla_\perp^2 E_\varphi - \frac{E_\varphi}{\rho^2} + \left(\frac{k_0^4 g^2}{h^2 - k_0^2\varepsilon} - h^2 + k_0^2\varepsilon\right) E_\varphi = \frac{k_0^4 gh}{h^2 - k_0^2\varepsilon}\,\frac{dE_z}{d\rho},$$

$$\nabla_\perp^2 E_z + \frac{h^2}{h^2 - k_0^2\varepsilon}\,\frac{1}{\varepsilon}\,\frac{d\varepsilon}{d\rho}\,\frac{dE_z}{d\rho} - \frac{\eta}{\varepsilon}(h^2 - k_0^2\varepsilon)E_z$$

$$= \frac{h}{\varepsilon}(h^2 - k_0^2\varepsilon)\,\frac{1}{\rho}\,\frac{d}{d\rho}\,\frac{\rho g E_\varphi}{h^2 - k_0^2\varepsilon},$$

$$E_\rho = \frac{i}{h^2 - k_0^2\varepsilon}\left(h\frac{dE_z}{d\rho} - k_0^2 g E_\varphi\right),$$

$$\mathbf{H} = ik_0^{-1}\,\nabla\times\mathbf{E}, \quad \nabla_\perp^2 = \frac{1}{\rho}\,\frac{d}{d\rho}\left(\rho\frac{d}{d\rho}\right). \tag{2}$$

Here $k_0 = \omega/c$ is the wave number in vacuum, $h = k_0 p$ is the propagation constant of the whistler, ρ, φ, z are cylindrical coordinates, and ε, g, η are the components of the plasma permittivity tensor

$$\hat{\varepsilon} = \begin{pmatrix} \varepsilon & -ig & 0 \\ ig & \varepsilon & 0 \\ 0 & 0 & \eta \end{pmatrix}. \tag{3}$$

In the frequency range studied (1), provided that

$$(\Omega_H\omega_H)^{1/2} \ll |\omega - i\nu_e| \ll \omega_H, \tag{4}$$

the components of the tensor (3) have the form[18]

$$\varepsilon = \frac{\omega_p^2}{\omega_H^2}\left(1 - i\frac{\nu_e}{\omega}\right), \quad g = -\frac{\omega_p^2}{\omega\omega_H}\left(1 - i\frac{2\nu_e\omega}{\omega_H^2}\right),$$

$$\eta = -\frac{\omega_p^2}{\omega^2}\left(1 - i\frac{\nu_e}{\omega}\right)^{-1}. \tag{5}$$

If the effective collision frequency ν_e is not too low, the equations (2) can be simplified substantially. Before making the appropriate simplifications, we note that the field structure of the modes guided in the frequency range (1) by an enhanced plasma density channel, has two different transverse scales—a large scale corresponding to the whistler waves sustained by the channel and a small scale corresponding to the quasielectrostatic waves, which transfer some of the energy to the background plasma.[5,17] As a result of energy leakage, the propagation constants p become complex: $p = p' - ip''$. Note that in the collisionless limit, the small-scale component predominates in the radial and longitudinal components of the electric field and is fairly appreciable in the other components.[19] In the presence of collisions, an effect similar to the skin effect is observed for the quasielectrostatic waves—above a certain collision frequency ν_e, the small-scale component of the field is concentrated in a comparatively thin layer in the vicinity of the plasma density gradient $N(\rho)$. On the basis of some results from Ref. 19, generalized to a collisional plasma, it is easy to establish that for the case

$$\text{Im}(k_0 ap(-\eta/\varepsilon)^{1/2})$$

$$= k_0 ap'\,\frac{\omega_H}{\omega(1 - \nu_e^2/\omega^2)}\left(\frac{\nu_e}{\omega} - \frac{p''}{p'}\right) \gg 1, \tag{6}$$

where a is the characteristic scale of the plasma density distribution along the transverse coordinate, assuming also that

$$|p|^2 \gg 4|\varepsilon|, \tag{7}$$

the field structure in the central part of the channel is predominantly determined by the large-scale component. Note that if the density N has a small gradient over the channel radius, where $(N/N_0 - 1)^2 \ll \min\{1,\, k_0 ap'\nu_e/\omega\}$ and N_0 is the background plasma density, the attenuation p'' of the dominant mode is given by

$$p'' \approx \frac{\nu_e}{2\omega_H}p'. \tag{8}$$

Using this equation and condition (4), inequality (6) can be simplified:

$$k_0 ap'\,\frac{\omega_H\nu_e}{\omega^2 + \nu_e^2} \gg 1. \tag{9}$$

Condition (9) is obviously satisfied for large enough values of the parameter $(\omega_H/\omega)k_0 ap'$, even if the effective collision frequency ν_e is low compared with the angular frequency ω. Thus, in the frequency range (1), allowance for the influence of electron collisions on the characteristics of the modes guided by enhanced-density channels can be of fundamental importance, and in particular, can appreciably change the structure of the mode field, as compared with a collisionless plasma.

This behavior is clearly illustrated by the results of the field structure calculations plotted in Figs. 1 and 2, which were performed by solving the equations (2) for the simplest model density profile

$$N(\rho) = N_0 + (\tilde{N} - N_0)[1 - U(\rho - a)], \tag{10}$$

and for two special cases, $\nu_e = 0$ and $\nu_e = 0.25\omega$, for the parameters $k_0 a = 0.12$, $\omega_H/\omega = 8.8$, $\omega_{p0}/\omega = 56.5$, $\tilde{N}/N_0 = 1.5$, where U is the Heaviside step function, $\tilde{N}$ is the plasma density inside the channel, and ω_{p0} is the plasma frequency corresponding to the background density N_0. In both cases, only the lowest (dominant) mode can propagate in the channel, for which the complex propagation constant is $p = 21.10 - i\cdot 1.34\times 10^{-2}$ at $\nu_e = 0$ (see Fig. 1) and $p = 21.18 - i\cdot 0.79$ at $\nu_e = 0.25\omega$ (see Fig. 2). Note that these graphs do not show the dependence of $\text{Re}\,E_\rho$, $\text{Im}\,E_{\varphi,z}$, $\text{Im}\,H_\rho$, and $\text{Re}\,H_{\varphi,z}$ on the transverse coordinate ρ, which have substantially lower absolute values compared with the corresponding quantities $\text{Im}\,E_\rho$, $\text{Re}\,E_{\varphi,z}$, $\text{Re}\,H_\rho$, and $\text{Im}\,H_{\varphi,z}$.

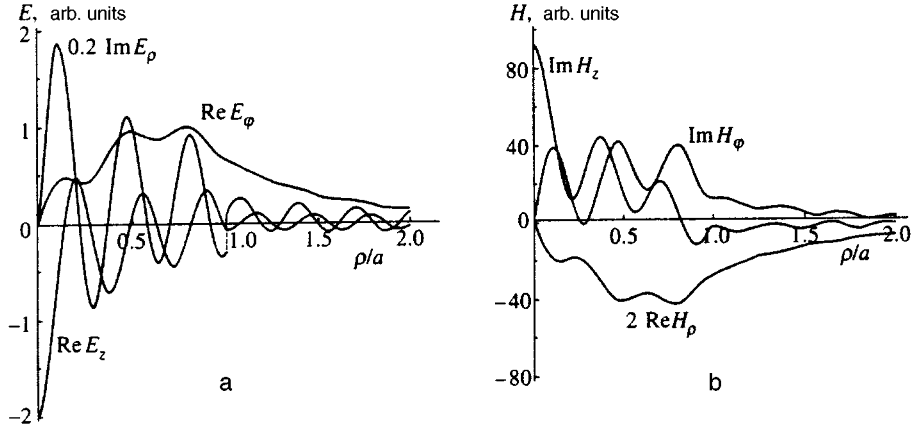

FIG. 1. Field structure of the dominant mode without collisions: $\nu_e=0$, $\widetilde{N}/N_0=1.5$, $\omega_{p0}/\omega=56.5$, $\omega_H/\omega=8.8$, and $k_0a=0.12$.

If conditions (6) and (7) are satisfied, the results of this analysis allow us to use comparatively simple approximate equations, easily derived from the rigorous equations (2), to describe the field in a collisional channel. For this purpose we rewrite the equations (2) in the form

$$\frac{d^2E_\varphi}{d\zeta^2}+\frac{1}{\zeta}\frac{dE_\varphi}{d\zeta}-\frac{E_\varphi}{\zeta^2}+(k_0a\hat{q}_1)^2E_\varphi=k_0ap\frac{g}{p^2-\varepsilon}\frac{dE_z}{d\zeta},$$

$$\chi^2\left(\frac{d^2E_z}{d\zeta^2}+\frac{1}{\zeta}\frac{dE_z}{d\zeta}+\frac{p^2}{p^2-\varepsilon}\frac{1}{\varepsilon}\frac{d\varepsilon}{d\zeta}\frac{dE_z}{d\zeta}\right)-\frac{\varepsilon}{p^2\eta}\hat{q}_2^2E_z$$

$$=-\chi\frac{p^2-\varepsilon}{g'}\frac{1}{\zeta}\frac{d}{d\zeta}\left(\frac{\zeta g'E_\varphi}{p^2-\varepsilon}\right), \qquad (11)$$

where

$$\hat{q}_1^2=\frac{g^2}{p^2-\varepsilon}-p^2+\varepsilon, \quad \hat{q}_2^2=-\frac{\eta}{\varepsilon}(p^2-\varepsilon),$$

$$\chi=\frac{\omega-i\nu_e}{\omega_H}\frac{1}{k_0ap}, \quad \zeta=\frac{\rho}{a},$$

$$g'=\mathrm{Re}\ g(|\mathrm{Re}\ g|\gg|\mathrm{Im}\ g|),$$

and allowance is made for the expressions (5). Recalling that $k_0ap'\gg1$ for azimuthally symmetric whistler modes in an enhanced-density channel,[19] it is readily established that in the special case (4) the following relations hold: $|\chi|\ll1$, $|\varepsilon\hat{q}_2^2/p^2\eta|\simeq1$ ($|\chi|\sim|\hat{q}_1|/|\hat{q}_2|$) so that a perturbation method can be used to analyze the system (11). Returning to the dimensional coordinate ρ, in the first order of perturbation theory ($\chi=0$) the first equation in this system gives

$$\frac{d^2E_\varphi}{d\rho^2}+\frac{1}{\rho}\frac{dE_\varphi}{d\rho}-\frac{E_\varphi}{\rho^2}+k_0^2\left(\frac{g'^2}{p^2}-p^2\right)E_\varphi=0. \qquad (12)$$

Here we have

$$E_\rho=-ig'p^{-2}E_\varphi \qquad (13)$$

and $E_z=0$. A nontrivial expression for the longitudinal component of the electric field is given by the next (first) order of perturbation theory:

$$E_z=-\frac{1}{k_0p\eta}\frac{1}{\rho}\frac{d}{d\rho}(\rho g'E_\varphi). \qquad (14)$$

In order to avoid misunderstanding, we note that in Eqs. (12)–(14) we have neglected some small terms of order ε/p^2, ν_e/ω_H. Significantly, when deriving these formulas it is not assumed that the variation in density N along the transverse coordinate is small. Equation (12) and the expressions (13) and (14) can be used to determine the structure of the large-scale components of the mode fields and also their real propagation constants $p=p'$ (when seeking p, small losses of energy to radiation and dissipation are simply neglected in this approximation)[1)] Since in case (6), the behavior of the field in channels of enhanced plasma density is determined by the large-scale components, these relations are extremely

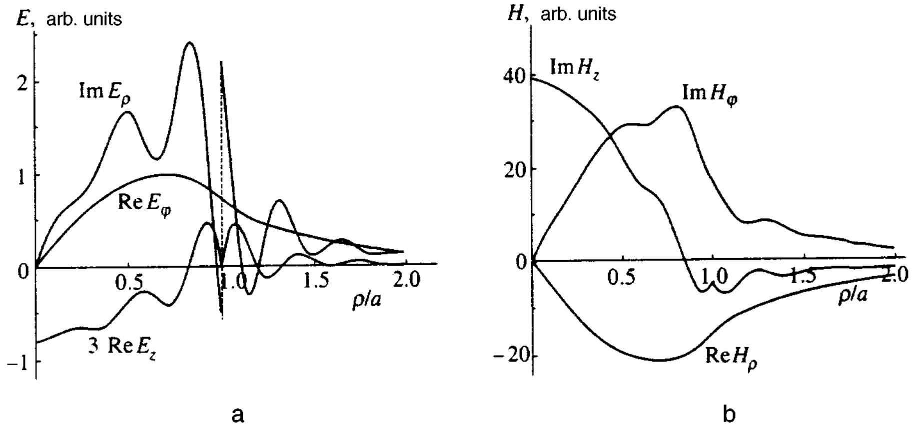

FIG. 2. Field structure of the dominant mode with collisions: $\nu_e=0.25\omega$; the other parameters are the same as in Fig. 1.

useful for analyzing the ionization self-interaction of whistler waves, resulting in the formation of these channels.

Equation (12) and expressions (13) and (14) should be analyzed together with the plasma energy and density balance equations.[21] We assume that in a plasma channel formed by breakdown and additional ionization of the background medium by a whistler field,

$$\nu_{ei} \ll \nu_{em}, \quad \delta_{em}\nu_{em} \ll \omega, \quad \delta_{ei}\nu_{ei} \ll \nu_{im}. \tag{15}$$

Here, ν_{im}, ν_{em}, and ν_{ei} are the frequencies of ion and electron collisions with neutral molecules and ions, respectively ($\nu_e = \nu_{em} + \nu_{ei}$), and δ_{em} and δ_{ei} are the average relative fractions of energy lost by the electrons in collisions with neutral molecules and ions, respectively. The last inequality in (15) allows the ion heating to be neglected.[21] Assuming that the characteristic inhomogeneity scales of the field amplitude are fairly large, we can neglect the contribution of heat conduction to the formation of the electron temperature profile T_e and we can express the steady-state distribution of T_e in the form

$$T_e = T_{e0} + \frac{e^2}{3m\delta_{em}\omega_H^2}\left[|E_\rho|^2 + |E_\varphi|^2 + 4\frac{\omega}{\omega_H}\,\mathrm{Im}(E_\rho E_\varphi^*) + \frac{\omega_H^2}{\omega^2+\nu_e^2}|E_z|^2\right], \tag{16}$$

where e and m are the electron charge and mass, and T_{e0} is the background electron temperature. The expressions (5) for the components of the plasma permittivity tensor were used to derive Eq. (16).

The steady-state distribution of the plasma density N over the transverse coordinate will be described using the density balance equation[21]

$$\frac{1}{\rho}\frac{d}{d\rho}\left(\rho D_\perp \frac{dN}{d\rho}\right) + (\nu_i - \nu_a)N - \alpha N^2 + q_{\mathrm{ext}} = 0, \tag{17}$$

where $D_\perp$ is the coefficient of plasma diffusion across the magnetic field, ν_i is the frequency of electron impact ionization of neutral molecules, ν_a is the electron attachment frequency, α is the electron–ion recombination frequency, q_{ext} is the intensity of the external ionization source sustaining the equilibrium density N_0:

$$q_{\mathrm{ext}} = (\alpha_0 N_0 + \nu_{a0} - \nu_{i0})N_0$$

(the zero subscript indicates background values of the appropriate quantities). Here it is assumed that, if the length of the plasma channel in the direction of the external magnetic field is sufficiently great as is the case for this particular range of parameters, as we shall see subsequently, longitudinal diffusion of the plasma can be neglected. In this case, the transverse diffusion is ambipolar. The coefficient $D_\perp$ under the conditions $T_e \gg T_i$ (T_i is the ion temperature) and $\omega_H\Omega_H \gg \nu_{em}\nu_{im}(1+\Omega_H^2/\nu_{im}^2)$, which are assumed to be satisfied here, is given by[21,22]

$$D_\perp = \frac{T_e\nu_e}{m\omega_H^2}. \tag{18}$$

Note that in Eq. (17) we also neglected the influence of thermal diffusion. This is valid if the characteristic transverse scale of electron temperature variation exceeds the corresponding scale of plasma density variation in the channel or if variations in the temperature T_e are fairly small. This condition is satisfied, for instance, when T_e is close to the ''breakdown'' value T_e^* which is determined by the relation

$$\nu_i(T_e^*) = \nu_a(T_e^*) + \alpha(T_e^*)N_0, \tag{19}$$

(in this case it is obvious that $q_{\mathrm{ext}} = 0$).

Equations (12) and (17) together with relations (13), (14), and (16) can be used to investigate steady-state self-consistent distributions of the field and the plasma formed in an unperturbed (background) plasma during the propagation of high-intensity whistler waves as a result of ionization nonlinear effects. The results of solving these equations are determined to a considerable extent by the dependences of the quantities contained in them on the temperature T_e and other factors. Thus, we need to specify these dependences to obtain some results.

3. SELF-CONSISTENT DISTRIBUTIONS OF THE FIELD AND THE PLASMA

Here we confine our analysis to ionization effects in air at typical discharge temperatures $T_e \leq 3$ eV. In this case, the temperature dependence of the quantities in expressions (16) and (17) can be described by the following model expressions:[21,23]

$$\nu_e = 1.23\times 10^{-7} N_m T_e^{\mu},$$

$$\nu_i = 2.7\times 10^{-8} N_m \left(\frac{T_e}{I}\right)^{1/2}\left(1 + 2\frac{T_e}{I}\right)\exp\left(-\frac{I}{T_e}\right),$$

$$\nu_a = \beta_a N_m, \quad \alpha = 5\times 10^{-7}\left(\frac{0.026}{T_e}\right)^{\lambda}, \tag{20}$$

where ν_e, ν_i, and ν_a are in reciprocal seconds, T_e is measured in eV, α is in $\mathrm{cm}^3\cdot\mathrm{s}^{-1}$, the concentration of neutral molecules N_m is in cm^{-3}, and I is the effective ionization potential of the neutral molecules, which is subsequently assumed to be $I = 14$ eV. The coefficient of attachment β_a and the values of μ, λ, and δ_{em} can be considered to be independent of T_e over this range of electron temperature: $\beta_a = 7.2\times 10^{-12}\ \mathrm{cm}^3\cdot\mathrm{s}^{-1}$, $\mu = 5/6$, $\lambda = 1.2$ (Ref. 21), and $\delta_{em} = 0.01$ (Ref. 24).

Despite the approximations and simplifications, Eqs. (12) and (17) can in general only be solved numerically. The results of the numerical calculations will be preceded by some analytic reasoning.

We postulate that the relative changes in the plasma density and the electron temperature are small quantities

$$n = \frac{N - N_0}{N_0} \ll 1, \quad \Theta = \frac{T_e - T_{e0}}{T_{e0}} \ll 1, \tag{21}$$

and the characteristic transverse dimension a of the channel is substantially greater than the diffusion length $(D_{\perp 0}/\alpha_0 N_0)^{1/2}$:

$$a^2 \gg D_{\perp 0}/\alpha_0 N_0. \tag{22}$$

Note that when (21) holds, the latter condition is known to be satisfied if

$$(k_0P)^2\frac{D_{\perp 0}}{\alpha_0N_0}\ll 1,$$

where P is the normalized propagation constant of a whistler wave strictly along the external magnetic field in a homogeneous plasma of density $N=N_0$ (in this approximation $P=\omega_{p0}/\sqrt{\omega\omega_H}$). Provided that these conditions are satisfied, the diffusion term in Eq. (17) can be neglected, along with the contribution of the longitudinal field component E_z to electron heating.[2)]

As a result, from Eqs. (16) and (17) we obtain

$$n=\frac{\Delta_0}{\alpha_0N_0}\left(1+\frac{q_{\mathrm{ext}}}{\alpha_0N_0^2}\right)^{-1}\Theta,\quad \Theta=2\frac{\omega^2}{\omega_H^2}\frac{|E_\varphi|^2}{E_p^2},\tag{23}$$

where

$$\Delta_0=T_{e0}\frac{\partial}{\partial T_e}(\nu_i-\nu_a-\alpha N_0)_{|T_e=T_{e0}},$$

$$E_p^2=3T_{e0}m\delta_{em}\omega^2/e^2.$$

Here it is assumed that for $n\ll 1$ we can set $E_\rho\approx iE_\varphi$ (see Eq. (13)). For the dependences used (20), the relation $\Delta_0>0$ is clearly satisfied, so that heating and additional ionization lead to an increase in the plasma density in regions of strong field. In this case, Eq. (12) has the form

$$\Delta_\perp E_\varphi-\frac{E_\varphi}{\rho^2}+2(k_0P)^2\left(n-2\frac{p-P}{P}\right)E_\varphi=0.\tag{24}$$

With the relations (23), Eq. (24) can be written as

$$\Delta_\perp E_\varphi-\frac{E_\varphi}{\rho^2}+2(k_0P)^2\left(\gamma^2|E_\varphi|^2-2\frac{p-P}{P}\right)E_\varphi=0,\tag{25}$$

where

$$\gamma^2=2\frac{\omega^2}{\omega_H^2}\frac{\Delta_0}{\alpha_0N_0}\left(1+\frac{q_{\mathrm{ext}}}{\alpha_0N_0^2}\right)^{-1}\frac{1}{E_p^2}.$$

It is known (see, for instance, Ref. 4) that an equation like (25) can have localized solutions satisfying the condition $E_\varphi\to 0$ at the origin ($\rho\to 0$) and at infinity ($\rho\to\infty$). Note that although a similar equation has been used in many studies of whistler self-interaction in a collisionless magnetoactive plasma[3)] in the frequency range (1) being considered, this equation is only valid for enhanced-density channels in case (9), i.e., when electron collisions are taken into account, as is clear from the analysis described above. It can be seen from Eq. (25) and relations (23) that when the dominant whistler mode undergoes self-channeling, the plasma radial distribution is nonmonotonic, with a density minimum on the axis $\rho=0$ and an annular layer of enhanced density surrounding the axial region. Curiously enough, for a planar (two-dimensional) field distribution the plasma density on the z axis has a maximum. The equation describing the behavior of the field in this case ($\partial/\partial y=0$) can be obtained from Eq. (25) by making the substitutions $(\Delta_\perp-\rho^{-2})\to d^2/dx^2$, $E_\varphi\to E_y$. The localized solution of the corresponding equation is well-known:

$$E_y(x)=\tilde{n}^{1/2}\gamma^{-1}\operatorname{sech}(k_0P\tilde{n}^{1/2}x),\tag{26}$$

where $\tilde{n}=n(0)$. We then have $p=P(1+\tilde{n}/4)$, and

$$n(x)=\tilde{n}\operatorname{sech}^2(k_0P\tilde{n}^{1/2}x).$$

In the general case, as we have already noted, we must be satisfied with the results of a numerical solution of Eqs. (12) and (17). We introduce dimensionless quantities which are convenient for the following analysis:

$$\xi=\rho\left(\frac{\alpha_0N_0}{D_{\perp 0}}\right)^{1/2},\quad \mathscr{E}_{\rho,\varphi,z}=\frac{E_{\rho,\varphi,z}}{E_p}.$$

Adopting this notation, Eqs. (12) and (17) are transformed to give

$$\frac{d^2\mathscr{E}_\varphi}{d\xi^2}+\frac{1}{\xi}\frac{d\mathscr{E}_\varphi}{d\xi}-\frac{\mathscr{E}_\varphi}{\xi^2}+\Lambda^2\left[\frac{P^4(1+n)^2}{p^2}-p^2\right]\mathscr{E}_\varphi=0,$$

$$\begin{aligned}(1+\Theta)^{1+\mu}&\left(\frac{d^2n}{d\xi^2}+\frac{1}{\xi}\frac{dn}{d\xi}\right)+(1+\mu)(1+\Theta)^\mu\frac{dn}{d\xi}\frac{d\Theta}{d\xi}\\&-\frac{(1+n)^2}{(1+\Theta)^\lambda}+\Lambda^2\frac{\omega_H^2}{\omega^2}\frac{c^2}{\beta_{e0}(T_{e0}/m)}[(\beta_i-\beta_a)\\&\times(1+n)-(\beta_{i0}-\beta_a)]+1=0,\end{aligned}\tag{27}$$

$$\begin{aligned}\Theta=|\mathscr{E}_z'|^2&+\frac{\omega^2}{\omega_H^2}|\mathscr{E}_\varphi|^2\\&\times\left[1+\frac{P^4}{p^4}(1+n)^2+4\frac{\omega}{\omega_H}\frac{P^2}{p^2}(1+n)\right],\end{aligned}$$

$$\begin{aligned}\mathscr{E}_z'&=\frac{\mathscr{E}_z}{1-i(1+\Theta)^\mu\nu_{e0}/\omega}\\&=-\frac{\omega}{\omega_H}\frac{1}{p\Lambda(1+n)\xi}\frac{d}{d\xi}(\xi(1+n)\mathscr{E}_\varphi),\end{aligned}$$

where

$$\Lambda=k_0\left(\frac{D_{\perp 0}}{\alpha_0N_0}\right)^{1/2},\quad \beta_i=\nu_iN_m^{-1},$$

$$\beta_{i0}=\nu_{i0}N_m^{-1},\quad \beta_{e0}=\nu_{e0}N_m^{-1},$$

and the remaining notation is as for Eq. (20). The localized solutions of Eqs. (27) bounded on the $\rho=0$ axis should satisfy the following conditions:

$$n\to 0,\quad \mathscr{E}_\varphi\to 0\quad(\xi\to\infty);$$

$$n\to\tilde{n}=\mathrm{const},\quad \mathscr{E}_\varphi\to 0\quad(\xi\to 0).\tag{28}$$

The state of the background plasma is characterized by the dimensionless parameters ω_{p0}/ω, ω_H/ω, ν_{e0}/ω, T_{e0}/I, and Λ.

An analysis of the system (27) shows that this has a one-parameter family of localized solutions. The parameter is the relative change in the density on the channel axis $\tilde{n}=n(0)=(N(0)-N_0)/N_0$. Figure 3 gives the curves $n(\xi)$ and $\Theta(\xi)$ obtained by solving this system numerically for three values of $\tilde{n}$: $\tilde{n}=0.5$, $\tilde{n}=0.6$, and $\tilde{n}=0.7$ for given values of $\omega_{p0}/\omega=56.5$, $\omega_H/\omega=8.8$, $\nu_{e0}/\omega=0.08$,

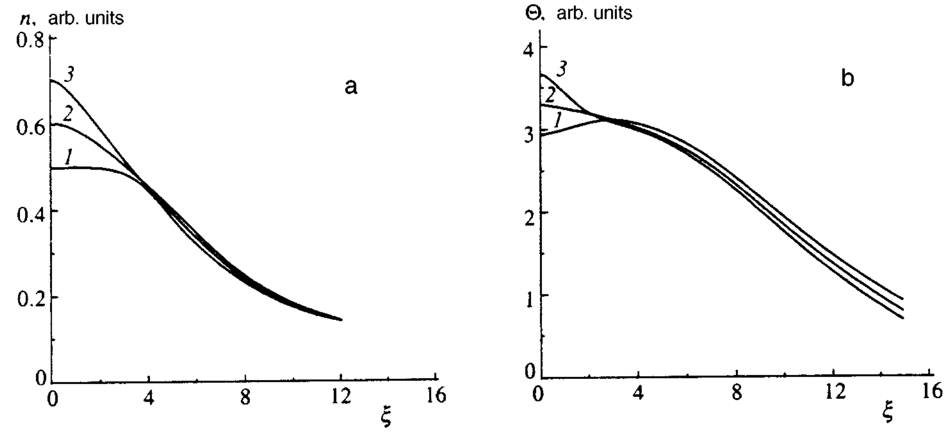

FIG. 3. Self-consistent distribution of relative perturbations of the density (a) and electron temperature (b) for $\omega_{p0}/\omega=56.5$, $\omega_H/\omega=8.8$, $\nu_{e0}/\omega=0.08$, $T_{e0}/I=0.03$, and $\Lambda=1.46\times10^{-2}$: *1*—$\tilde{n}=0.5$, *2*—$\tilde{n}=0.6$, and *3*—$\tilde{n}=0.7$.

$T_{e0}/I=0.03$, and $\Lambda=1.46\times10^{-2}$. The calculations show that for these values of $\tilde{n}$ only the dominant mode undergoes self-channeling, for which the real propagation constant takes the values $p=20.62$ ($\tilde{n}=0.5$), $p=20.68$ ($\tilde{n}=0.6$), and $p=20.76$ ($\tilde{n}=0.7$). It can be seen from these data that the characteristic transverse scale of the electron temperature distribution is appreciably greater than the corresponding scale of the plasma density variation, which justifies our neglect of the contribution of thermal diffusion to the formation of the density profile. Note that as $\tilde{n}$ increases, the width of the radial distribution of Θ decreases, so that for large enough $\tilde{n}$ ($\tilde{n}>2-3$), allowance must be made for thermal diffusion. For small $\tilde{n}$ ($\tilde{n}<0.5$), the profile $n(\xi)$ is nonmonotonic ($\tilde{n}<n_{\max}$), as noted when analyzing the simplified equation (25).

Figure 4 shows the behavior of the normalized field components $\mathscr{E}_{\rho,\varphi,z}=E_{\rho,\varphi,z}/E_p$ and $\mathscr{H}_{\rho,\varphi,z}=H_{\rho,\varphi,z}/E_p$ for values of the parameters corresponding to curves *3* in Fig. 3. Taking the results of calculations of the field components as the basis, we can easily determine the plasma channel length $L\sim1/h''$ by using the well-known energy formula[25] for the attenuation constant h'' of the waveguide mode,

$$h''=k_0p''=q/2W_\perp,$$

where $W_\perp$ is the power transferred across the channel cross section and Q is the power loss per unit length. For the distributions plotted in Fig. 3, the channel length L and its characteristic radius a are related by[4)] $L\gtrsim a\omega_H/\nu_e$, so that $L\gg a$ and

$$\left|D_\parallel\frac{\partial^2N}{\partial z^2}\bigg/D_\perp\Delta_\perp N\right|<\frac{D_{\parallel,\max}}{D_\perp}\frac{a^2}{2L^2}\lesssim0.5,$$

where $D_\parallel$ is the plasma diffusion coefficient along the external magnetic field, $D_{\parallel,\max}=T_e/m\nu_e$. The last inequality justifies the use of the simplified form (17) of the plasma density balance equation.

A comparison between the self-consistent field distributions (Fig. 4) calculated using the model with the field distributions obtained from the rigorous equations (2) for given (''frozen'') profiles $N(\rho)$ and $\nu_e(\rho)$ obtained by solving the nonlinear problem, revealed good agreement between these distributions. This can be attributed to the fact that for the results plotted in Figs. 3 and 4, the conditions of validity of the approximate description of the field based on formulas (12)–(14) are satisfied with a sufficient margin. For example, for $\tilde{n}=0.7$ ($k_0a=5.5\Lambda$, $\nu_e(0)/\omega=0.29\gg p''/p'$) we have (see Eq. (6))

$$k_0ap'\frac{\omega_H\nu_e(0)}{\omega^2}=4.25.$$

We therefore claim that these localized solutions satisfy

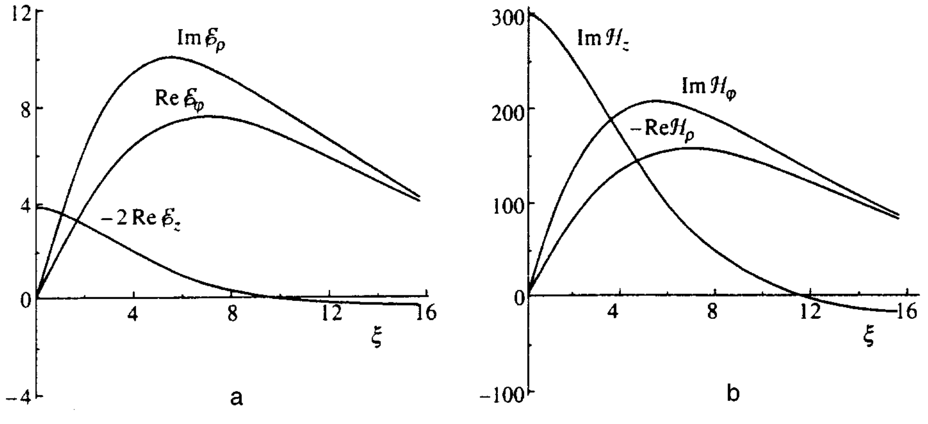

FIG. 4. Self-consistent distribution of the components of the whistler field for $\tilde{n}=0.7$. The other parameters are the same as in Fig. 3.

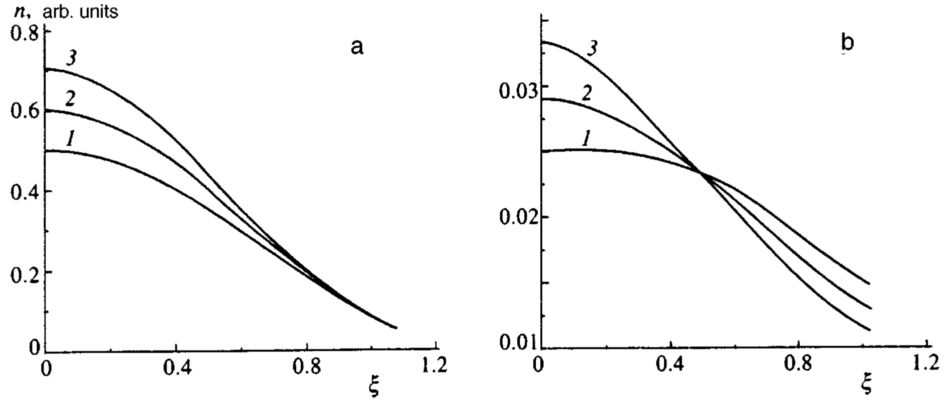

FIG. 5. Self-consistent distribution of relative perturbations of the density (a) and electron temperature (b) for $\omega_{p0}/\omega=56.5$, $\omega_H/\omega=8.8$, $\nu_{e0}/\omega=0.25$, $T_{e0}/I=0.14$, and $\Lambda=1.15$: *1*—$\tilde{n}=0.5$, *2*—$\tilde{n}=0.6$, and *3*—$\tilde{n}=0.7$.

all the constraints put forward above and describe ionization self-channeling of whistler waves in a magnetoactive collisional plasma.

These results apply to the case $T_{e0}<T_e^*$ ($q_{\text{ext}}\neq 0$). Without going into details, we discuss another possible situation where $T_{e0}=T_e^*$ ($q_{\text{ext}}=0$). Figure 5 gives self-consistent distributions $n(\xi)$ and $\Theta(\xi)$ for this case, corresponding to $\tilde{n}=0.5$ ($p=20.36$), $\tilde{n}=0.6$ ($p=20.52$), and $\tilde{n}=0.7$ ($p=20.68$) ($\nu_{e0}/\omega=0.25$, $T_{e0}/I=0.14$, $\Lambda=0.15$, and the values of ω_{p0}/ω and ω_H/ω are as before). It can be seen from Fig. 5 that the relative variations in the temperature Θ are considerably smaller than those for the similar curves plotted in Fig. 3b. The field structure is approximately the same as in the previous case $T_e<T_e^*$ (see Fig. 4) and thus is not given.

These results can be used to estimate some parameters of ionization self-channeling of whistlers in a laboratory plasma. For example, the distributions plotted in Figs. 3 and 4 correspond to $N_0=4\times 10^{10}$ cm^{-3} ($\omega_{p0}=1.13\times 10^{10}$ s^{-1}), $H_0=100$ Oe ($\omega_H=1.76\times 10^9$ s^{-1}), $\omega=2\times 10^8$ s^{-1}, $\nu_{e0}=1.6\times 10^7$ s^{-1}, and $T_{e0}=0.4$ eV (in this case the required gas pressure at room temperature is $p=8.7\times 10^{-3}$ Torr). For the case $\tilde{n}=0.7$ ($N(0)=6.8\times 10^{10}$ cm^{-3}, $T_e(0)=1.9$ eV, $\nu_e(0)=5.8\times 10^7$ s^{-1}), we obtain the following values of the main parameters: whistler wavelength $\lambda_w=2\pi/(k_0 p)\simeq 45$ cm, channel radius $a\simeq 12$ cm ($n(a)=0.5n_{\max}$), maximum field in discharge $|\mathbf{E}|_{\max}\simeq 6.4$ V/cm, and power transferred across channel cross section $W_\perp\simeq 1.5$ kW. These figures indicate that this effect can be simulated in the laboratory in modern large-scale plasma facilities (see Refs. 9 and 26).

4. CONCLUSIONS

This analysis suggests that the additional ionization of a collisional magnetoactive plasma accompanying the heating of its electrons by a fairly large-amplitude whistler field can result in the formation of plasma waveguide structures which trap and guide the whistler waves generating these structures. It can be seen from these results that the ionization self-channeling of whistlers in a collisional plasma differs substantially from their self-interaction in a collisionless plasma when the channels are formed as a result of the ponderomotive force generated by the whistler wave field.[4,5] The effect studied qualitatively here resembles the self-interaction of helicons accompanying the heating of a solid-state plasma.[27]

Finally, we note that our approach and also some specific results of this analysis are of interest not only for possible relevant experiments (both natural and model laboratory experiments) in an ionospheric plasma, but also to identify the role that ionization self-interaction of whistlers might play in bounded systems such as helicon plasma sources and facilities using these sources.

The authors would like to thank the Russian Fund for Fundamental Research for financial support (Project 95-02-05816a). The work of one of the authors (A.V.K.) was also supported by INTAS (Grant 93-2492-ext) as part of the research program of the International Center of Fundamental Physics in Moscow.

1) It should be noted that Eq. (12) essentially corresponds to the single-fluid electron magnetohydrodynamics approximation.[20] However, the derivation of expression (14) required for subsequent analysis involves going beyond this approximation.

2) For a comparatively narrow channel ($k_0ap'\sim 1$), the contribution of the E_z component to electron heating must be taken into account, as will become clear from the following reasoning.

3) See Ref. 4 in particular and the cited literature.

4) Here and subsequently the value of a corresponds to the relative density perturbation $n(a)=n_{\max}/2$.

---

[1] A. G. Litvak, Zh. Éksp. Teor. Fiz. **57**, 629 (1969) [Sov. Phys. JETP **30**, 344 (1970)].
[2] H. Washimi, J. Phys. Soc. Jpn. **34**, 1373 (1973).
[3] G. L. Markov, V. A. Mironov, A. M. Sergeev *et al.*, Zh. Éksp. Teor. Fiz. **80**, 2264 (1981) [Sov. Phys. JETP **53**, 1183 (1981)].
[4] V. I. Karpman, R. N. Kaufman, and A. G. Shagalov, J. Plasma Phys. **31**, 209 (1984).
[5] V. I. Karpman and A. G. Shagalov, Zh. Éksp. Teor. Fiz. **87**, 422 (1984) [Sov. Phys. JETP **60**, 242 (1984)].
[6] G. A. Markov, Fiz. Plazmy **14**, 1094 (1988) [Sov. J. Plasma Phys. **14**, 641 (1988)].
[7] G. Yu. Golubyatnikov, S. V. Egorov, A. V. Kostrov *et al.*, Zh. Éksp. Teor. Fiz. **96**, 2009 (1989) [Sov. Phys. JETP **69**, 1134 (1989)].
[8] Yu. N. Agafonov, V. S. Bazhanov, V. Ya. Isyakaev *et al.*, JETP Lett. **52**, 530 (1990).
[9] J. M. Urrutia and R. L. Stenzel, Phys. Rev. Lett. **67**, 1867 (1991).
[10] T. M. Zaboronkova, A. V. Kostrov, A. V. Kudrin *et al.*, Zh. Éksp. Teor. Fiz. **102**, 1151 (1992) [Sov. Phys. JETP **75**, 625 (1992)].
[11] R. A. Helliwell, *Whistlers and Related Ionospheric Phenomena*, Stanford University Press, Stanford (1965).
[12] T. M. Zaboronkova, I. G. Kondrat'ev, and A. V. Kudrin, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **37**, 887 (1994).

[13] R. W. Boswell, Plasma Phys. Controlled Fusion **26**, 1147 (1984).
[14] F. F. Chen, J. Vac. Sci. Technol. A **10**, 1389 (1991).
[15] N. F. Vorob'ev and A. A. Rukhadze, Fiz. Plazmy **20**, 1065 (1994) [Plasma Phys. Rep. **20**, 955 (1994)].
[16] K. P. Shamrai and V. B. Taranov, Plasma Sources Sci. Technol. **5**, 474 (1996).
[17] M. J. Laird and D. Nunn, Planet. Space Sci. **23**, 1649 (1975).
[18] V. L. Ginzburg, *The Propagation of Electromagnetic Waves in Plasmas*, 2nd ed., Pergamon Press, Oxford (1970).
[19] T. M. Zaboronkova, A. V. Kudrin, and G. A. Markov, Fiz. Plazmy **19**, 769 (1993) [Plasma Phys. Rep. **19**, 397 (1993)].
[20] A. S. Kingsep, K. V. Chukbar, V. V. Yan'kov *et al.*, in *Reviews of Plasma Physics*, V. V. Kadomtsev (ed.) [in Russian], Énergoatomizdat, Moscow (1987), p. 309.
[21] A. V. Gurevich and A. B. Shvartsburg, *Nonlinear Theory of Radio Propagation in the Ionosphere* [in Russian], Nauka, Moscow (1973).
[22] V. A. Rozhanskiĭ and L. D. Tsendin, *Collisional Transport in a Partially Ionized Plasma* [in Russian], Énergoatomizdat, Moscow (1988).
[23] Yu. P. Raĭzer, *Physics of Gas Discharges* [in Russian], Nauka, Moscow (1987).
[24] H. S. W. Massey and E. H. S. Burhop, *Electronic and Ionic Impact*, Clarendon Press, Oxford (1952).
[25] L. A. Vaĭnshteĭn, *Electromagnetic Waves* [in Russian], Radio i Svyaz', Moscow (1988).
[26] G. Yu. Golubyatnikov, S. V. Egorov, B. G. Eremin *et al.*, Zh. Éksp. Teor. Fiz. **107**, 441 (1995) [JETP **80**, 234 (1995)].
[27] Z. K. Yankauskas, JETP Lett. **39**, 223 (1984).

Translated by R. M. Durham

Cascade processes in a plasma oscillator

M. A. Krasil'nikov and M. V. Kuzelev

Moscow State University of Printing, 127550 Moscow, Russia

A. A. Rukhadze

Institute of General Physics, Russian Academy of Sciences, 117942 Moscow, Russia

(Submitted 25 February 1997)
Zh. Éksp. Teor. Fiz. **112**, 1299–1311 (October 1997)

A theoretical and numerical analysis is made of the dynamics of nonlinear electron-beam scattering of a wave reflected by the emitting device of a plasma oscillator. It is shown that a counterpropagating plasma wave can interact nonlinearly with other waveguide modes of the system and with charge-density beam waves, leading to changes in the operation of the oscillator. It is established by means of a numerical simulation that the generation efficiency is reduced as a result of scattering of the counterpropagating wave and stimulated emission of a strong-potential plasma wave with phase velocity $v_{ph}=\omega/k_z \ll c$. © *1997 American Institute of Physics.* [S1063-7761(97)01110-4]

1. The nonlinear dynamics of resonant stimulated Čerenkov emission in a plasma waveguide of finite length was investigated in Ref. 1. It was established that the partial reflection of a plasma wave concurrent with an electron beam leads to a change in the output signal level and substantially influences its temporal dynamics. It was observed that the counterpropagating wave can be scattered by beam electrons, which can result in the evolution of a whole series of cascade processes. Here we propose to study these processes.

We assume that a section $0<z<L$ of metallic waveguide of radius R is filled with a transversely homogeneous plasma (ω_p is the Langmuir frequency of the plasma electrons). Initially a thin electron beam with a certain rise time begins to be injected into the waveguide across the boundary $z=0$. As a result of spontaneous Čerenkov radiation, the injected electron beam excites a concurrent plasma wave at the frequency $\omega \sim \omega_0 \sqrt{\omega_p^2 - k_\perp^2 u^2 \gamma^2}$ (Ref. 2), where $k_\perp$ is the transverse wave number, and u and γ are the velocity and relativistic factor of the electron beam. An emitting device is positioned near the boundary $z=L$. Using the simplest model of a horn—an abrupt plasma–vacuum interface—the coefficient of reflection of a concurrent plasma wave is given by[1,2]

$$\kappa=\frac{n(\omega_p)-n(0)}{n(\omega_p)+n(0)}, \tag{1}$$

where $n^2(\omega_p)=1-k_\perp^2 c^2/(\omega^2-\omega_p^2)$. The reflected counterpropagating wave has the same frequency but propagates in the opposite direction. Near the boundary $z=0$ the counterpropagating wave is almost completely transformed into a concurrent wave. If the starting conditions for oscillation are satisfied for the system (in terms of beam current J_b, waveguide length L, and plasma frequency ω_p), the plasma oscillator undergoes self-excitation.[2]

In the model adopted in Ref. 1, the role of the counterpropagating wave was reduced to feedback, since the counterpropagating wave, not being in Čerenkov resonance with the beam electrons, does not, on average, interact with them. However, the counterpropagating wave can be scattered by the beam, accompanied by a change in frequency and polarization, which can lead to changes in the oscillation regimes. The main mechanism for these processes is nonlinear wave interaction.[3]

2. The oscillation spectra of a waveguide filled with a homogeneous plasma and a thin beam, located in a strong longitudinal magnetic field, can be determined from the following dispersion equations:[4]

$$D^n=k_{\perp n}^2+\left(k_z^2-\frac{\omega^2}{c^2}\right)\left(1-\frac{\omega_p^2}{\omega^2}-\frac{\omega_b^2/\gamma^3}{(\omega-k_z u)^2}G_n\right), \tag{2}$$

where $n=1,2,\ldots$ is the transverse mode number, φ_n is the eigenfunction of the empty waveguide, k_z and ω are the longitudinal wave number and the natural wave frequency, ω_b is the Langmuir frequency of the beam electrons, $G_n=S_b\varphi_n^2(r_b)/\|\varphi_n\|^2$ is the geometric factor of the beam, S_b is its cross-sectional area, and r_b is its average radius. The dispersion curves of the oscillations (2) for various n are plotted in Fig. 1. Point 0 in the figure, corresponding to the point of intersection of the plasma branch of the dispersion curve with the line $\omega=k_z u$ corresponds to exact Čerenkov resonance between the plasma wave and the beam electrons. Point 1 corresponds to a counterpropagating wave having the same frequency as the concurrent wave ($\omega_1=\omega_0$) but propagating in the opposite direction ($k_{z1}=-k_{z0}$).

In addition to electromagnetic ($\omega^2>k_z^2c^2$) and plasma ($\omega^2<\omega_p^2$) oscillations, the system also contains charge-density beam waves. For an infinitely thin beam with a transverse profile defined by a δ-function there are two of these waves: fast and slow. Their spectrum is determined by

$$\omega=k_z u \pm \Omega_b(k_z), \tag{3}$$

where Ω_b is the frequency of the beam oscillations; an accurate expression for Ω_b may be found in Ref. 2, for instance. If the Raman resonance conditions are satisfied, interaction can take place between the counterpropagating wave (point 1 in Fig. 1), another waveguide mode, and a charge-density beam wave. Raman resonance between these waves can be established if[2,3]

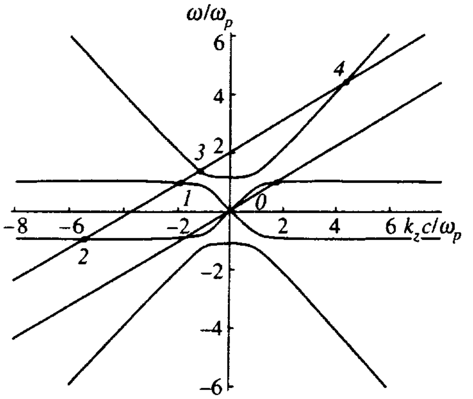

FIG. 1. Dispersion curves of a plasma waveguide (to determine the waveguide modes in Raman resonance with the charge-density beam waves).

$$\mathscr{D}_{1\alpha}=\omega_1-\omega_\alpha-(k_{z1}-k_{z\alpha})u=\pm\Omega_b, \tag{4}$$

where α is an index corresponding to the waveguide mode. Condition (4) implies that in the beam rest frame the product of the fields of the interacting waves yields beats at the oscillation frequency of the beam charge density. The plus and minus signs in Eq. (4) refer to phase matching with the fast and slow charge-density beam wave, respectively.

If the frequency Ω_b is substantially lower than all the other characteristic frequencies of the system, the frequency and wave number of the waveguide mode satisfying the resonance conditions (4) can be determined graphically simply by drawing the line $\omega=k_z u+2\omega_0$ in Fig. 1. It is easy to see that to within quantities of order Ω_b, ω_a and $k_{z\alpha}$ are the same as $\omega_{2,3,4}$ and $k_{z2,3,4}$ corresponding to the renumbered points in Fig. 1, and α is then taken as one of these subscripts.

3. We assume that t_0 is the characteristic time of variation of the interacting wave amplitudes, and $\omega^{(0)}$ and $k_z^{(0)}$ are the frequency and wave number of the charge-density beam wave, respectively. In addition to translational motion, the beam electrons also participate in many types of vibrational motion. If

$$\max(t_0^{-2},\Omega_b^2,\mathscr{D}_{1\alpha})\ll(\omega_{1,\alpha}-k_{z,1,\alpha}u)^2, \quad (k_z^{(0)}u)^2, \tag{5}$$

electron motion in the rest frame of the unperturbed beam can be divided into fast and slow components,[5] where motion in the field of the concurrent plasma and Raman waves, and in the space charge field of the beam, enters into the discharge of the slow components.

To derive nonlinear equations describing the processes in the system, we assume that the average width of the radiation spectrum is substantially less than the average frequency[6–8] and use the method of slowly varying amplitudes.[9] In addition, we analyze the plasma in the linear approximation, i.e., we describe it using linearized hydrodynamic equations.[10]

4. We describe the electromagnetic fields in a system of charged particles completely magnetized along the z axis using an equation for the scalar polarization potential Ψ (Ref. 11):

$$\frac{\partial}{\partial z}\left(\nabla^2-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\Psi=4\pi(\rho_p+\rho_b), \tag{6}$$

where ρ_p and ρ_b are the densities of the charge induced in a medium consisting of plasma and beam electrons. We seek a solution of Eq. (6) in the form

$$\Psi=\Psi_0+\Psi_1+\Psi_\alpha+\Psi_{1\alpha}, \tag{7}$$

where

$$\Psi_{0,1,\alpha}(z,t,r_\perp)=\frac{1}{2}[C_{0,1,\alpha}(z,t,r_\perp)\exp(-i\omega_{0,1,\alpha}t+ik_{z0,1,\alpha}z)+\text{c.c.}] \tag{8}$$

are the polarization potentials, where Ψ_0, Ψ_1, and Ψ_α are the polarization potentials of the concurrent plasma wave, the counterpropagating plasma wave, and a wave at one of the frequencies $\omega_{2,3,4}$, respectively. In general, Eq. (7) must take into account the contribution of all the waveguide modes, not just one. However, a comparison of the characteristic times for nonlinear scattering of the waves indicates that in practice, it is sufficient to allow only for one mode, as is done subsequently. For the polarization potential $\Psi_{1\alpha}$ of the beam space charge we have

$$\Psi_{1\alpha}(z,t,r_\perp)=\frac{1}{2}[C_{1\alpha}(z,t,r_\perp)\times\exp(-i\omega^{(0)}t+ik_z^{(0)}z)+\text{c.c.}]. \tag{9}$$

We note that C_α and $C_{1\alpha}$ are slow functions of z and t.

Assuming that the electron beam has an infinitely thin cross section, we use the phase density of the beam electrons to calculate the charge density induced in the beam[12]:

$$\rho_b(z,t)=e\,\delta(r_\perp-r_b)\sum\ \Phi(z)\,\delta[z-z_j]\,\frac{n_b S_b\lambda}{N}, \tag{10}$$

where $r_\perp$ is the coordinate in the transverse direction, r_b is the average transverse coordinate of the beam, and $\Phi(z)$ describes the envelope of the beam current in the longitudinal direction. In addition, assuming that the method of large particles[13,14] is used subsequently to model the beam, in Eq. (10) we introduced the total number of these particles over the length λ. We then have $N=n_b S_b\lambda$, where n_b is the density of the electron beam per unit length, and z_j is the coordinate of the jth particle. The perturbation of the plasma charge density ρ_p is determined from the hydrodynamic equation:

$$\frac{\partial^2\rho_p}{\partial t^2}=-\frac{\omega_p^2}{4\pi}\frac{\partial E_z}{\partial z}, \tag{11}$$

where E_z is the longitudinal component of the electric field strength, which can be expressed in terms of the polarization potential:[10]

$$E_z=\left(\frac{\partial^2}{\partial z^2}-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\Psi. \tag{12}$$

If the functions (8) and (9) have the total spatial period λ ($\lambda=(2\pi/k_{z0,1,\alpha})n_{0,1,\alpha}$, where $n_{0,1,\alpha}$ are integers denoting the number of wavelengths with the appropriate subscript within the length λ), after substituting the expression for Ψ into the field equation (6), multiplying by $\exp(i\omega_{0,1,\alpha}t-ik_{z0,1,\alpha}z)$, and averaging over z for the length λ, we obtain reduced equations for the slow amplitudes:

$$\left(\nabla_\perp^2-\frac{k^{(0)^2}}{\gamma^2}+\frac{\omega_p^2}{u^2\gamma^2}\right)C_{1\alpha}$$

$$=-i\frac{m}{e}\frac{\omega_b^2}{k_z^{(0)}}S_b\delta(r_\perp-r_b)$$

$$\times\frac{2}{N}\sum_j \exp(i\omega^{(0)}t-ik_z^{(0)}z_j)\Theta_1(z_j-z,\lambda),$$

$$\left(-iD'_{0,1,\alpha}+\frac{\partial D'_{0,1,\alpha}}{\partial\omega_{0,1,\alpha}}\frac{\partial}{\partial t}-\frac{\partial D'_{0,1,\alpha}}{\partial k_{z0,1,\alpha}}\frac{\partial}{\partial z}\right)C_{0,1,\alpha}$$

$$=\frac{m}{e}\frac{\omega_b^2}{k_{z0,1,\alpha}}S_b\delta(r_\perp-r_b)$$

$$\times\frac{2}{N}\sum_j \exp(i\omega_{0,1,\alpha}t-ik_{z0,1,\alpha}z_j)\Theta_1(z_j-z,\lambda), \tag{13}$$

where

$$\Theta_1(x,\lambda)=\left\{\begin{array}{ll}0, & x<-\lambda/2,\\ 1, & -\lambda/2\leqslant x<\lambda/2,\\ 0, & x\geqslant\lambda/2,\end{array}\right\}, \tag{14}$$

and the expression for $D'_{0,1,\alpha}$ is the same as Eq. (2) if we substitute $-k_{\perp n}^2\rightarrow\nabla_\perp^2$, neglecting the beam contribution, and $\omega=\omega_{0,1,\alpha}$ and $k_z=k_{z0,1,\alpha}$.

5. In accordance with the time scale introduced (5), we divide the vibrational motion of the electrons in the beam rest frame into fast and slow:

$$v_{zj}=u+v'_j+\tilde{v}_j,\quad z_j=ut+z'_j+\tilde{z}_j, \tag{15}$$

where $\tilde{z}_j$ and $\tilde{v}_j$ are the coordinate and velocity of fast oscillations of the jth particle in the fields of the waves with frequencies ω_1 and ω_α, and z'_j and v'_j describe the slow motion in the fields of the concurrent plasma wave, the beam space charge, and the Raman wave. If

$$|k_{z1,\alpha}\tilde{z}_j|\ll 1, \tag{16}$$

the fast oscillations are linear in terms of the amplitudes C_1 and C_α and take place within times of the order $(\omega_{1,\alpha}-k_{z1,\alpha}u)^{-1}$ (Ref. 2).

Using a standard procedure for averaging the equations of motion over time,[15–17] we obtain an equation for the slow components:

$$\frac{dz'_j}{dt}=v'_j,$$

$$\frac{dv'_j}{dt}=-\frac{ew_j}{2m\gamma^3}\left[\hat{C}_{1\alpha}\exp(ik_z^{(0)}z'_j)+\hat{C}_0\right.$$

$$\times\exp(ik_{z0}z'_j)+ik_z^{(0)}\frac{ew_j}{2m\Omega^2\gamma^3}\hat{C}_1\hat{C}_\alpha^*$$

$$\left.\times\exp(-i\mathscr{D}_{1\alpha}t+ik_z^{(0)}z'_j)+\text{c.c.}\right], \tag{17}$$

where

$$w_j=\left(1-2\gamma^2\frac{u^2}{c^2}\frac{v'_j}{u}\right)^{3/2},$$

$$\hat{C}_{0,1,\alpha}=\left(k_{z0,1,\alpha}^2-\frac{\omega_{0,1,\alpha}^2}{c^2}-2ik_{z0,1,\alpha}\frac{\partial}{\partial z}\right.$$

$$\left.-2i\frac{\omega_{0,1,\alpha}}{c^2}\frac{\partial}{\partial t}\right)C_{0,1,\alpha}, \tag{18}$$

$$\hat{C}_{1\alpha}=\left(k_z^{(0)^2}-\frac{\omega^{(0)^2}}{c^2}-2ik_z^{(0)}\frac{\partial}{\partial z}-2i\frac{\omega^{(0)}}{c^2}\frac{\partial}{\partial t}\right)C_{1\alpha},$$

$$\Omega^2\approx(\omega_1-k_{z1}u)^2\approx(\omega_2-k_{z2}u)^2\approx 4\omega_0^2.$$

After expanding the amplitude $C_{1\alpha}$ of the beam space charge wave in Eq. (13) in terms of the waveguide eigenfunctions, we substitute the coefficients of the expansion into Eq. (17). Using an expression for the frequency Ω_b of the beam oscillations,[2] we obtain an equation for the velocity of the jth particle:

$$\frac{dv'_j}{dt}=-\left[\frac{ew_j}{2m\gamma^3}\hat{C}_0\exp(ik_0z'_j)+\text{c.c.}\right]$$

$$-\frac{i}{2}\frac{\Omega_b^2}{k_z^{(0)}}w_j[\rho_{1\alpha}\exp(ik_z^{(0)}z'_j)-\text{c.c.}]$$

$$-i\left(\frac{e}{m}\right)^2\frac{k_z^{(0)}}{4\Omega^2\gamma^6}w_j^2(\hat{C}_1\hat{C}_\alpha^*$$

$$\times\exp(ik_z^{(0)}z'_j-i\mathscr{D}_{1\alpha}t)-\text{c.c.}), \tag{19}$$

where

$$\rho_{1\alpha}=\frac{2}{N}\sum_j \exp(-ik_z^{(0)}z'_j)\Theta_1(ut+z'_j-z,\lambda) \tag{20}$$

is the Raman harmonic of the beam charge density wave.

The right-hand sides of Eqs. (13) should also be averaged over time in accordance with the time scale (5), where no rapidly oscillating terms are obtained for the concurrent plasma wave because the Čerenkov resonance condition $\omega_0=k_{z0}u$ is satisfied for this wave. After averaging Eqs. (13), we obtain for the amplitude of the concurrent plasma wave C_0

$$\frac{\partial D_0}{\partial\omega_0}\frac{\partial C_0}{\partial t}-\frac{\partial D_0}{\partial k_{z0}}\frac{\partial C_0}{\partial z}+i(k_{\perp 1}^2+\nabla_\perp^2)C_0$$

$$=\frac{m}{e}\frac{\omega_b^2}{k_{z0}}S_b\delta(r_\perp-r_b)\rho, \tag{21}$$

where ρ is the fundamental harmonic of the beam charge density wave:

$$\rho=\frac{2}{N}\sum_j \exp(i\omega_0t-ik_{z0}z_j)\Theta_1(z_j-z,\lambda). \tag{22}$$

For the amplitudes of the counterpropagating wave and the waveguide mode at frequency ω_α, we obtain, after dropping the rapidly oscillating terms in Eq. (13),

$$\frac{\partial D_1}{\partial \omega_1}\frac{\partial C_1}{\partial t}-\frac{\partial D_1}{\partial k_{z1}}\frac{\partial C_1}{\partial z}+i(k_{\perp 1}^2+\nabla_\perp^2)C_1$$

$$=-\frac{i}{2}\frac{\omega_b^2}{\Omega^2\gamma^3}S_b\delta(r_\perp-r_b)\hat{\rho}_{1\alpha}\hat{C}_\alpha\exp(i\mathscr{D}_{1\alpha}t), \tag{23}$$

$$\frac{\partial D_\alpha}{\partial \omega_\alpha}\frac{\partial C_\alpha}{\partial t}-\frac{\partial D_\alpha}{\partial k_{z\alpha}}\frac{\partial C_\alpha}{\partial z}+i(k_{\perp 1}^2+\nabla_\perp^2)C_\alpha$$

$$=-\frac{i}{2}\frac{\omega_b^2}{\Omega^2\gamma^3}S_b\delta(r_\perp-r_b)\hat{\rho}_{1\alpha}^*\hat{C}_1\exp(-i\mathscr{D}_{1\alpha}t),$$

where

$$\hat{\rho}_{1\alpha}=\frac{2}{N}\sum_j w_j\exp(-ik_z^{(0)}z_j')\Theta_1(ut-z_j'-z,\lambda) \tag{24}$$

is the Raman harmonic of the beam charge density wave with allowance for the electron relativistic motion.

Since only one transverse mode (for instance, the dominant $n=1$ mode) is excited resonantly, and the other transverse modes are not initially perturbed, allowance for the nonresonant transverse modes can be made in accordance with perturbation theory,[18] using the quantity $\omega_b^2/\gamma^3\Omega^2$ as the small parameter. It can be seen from Eqs. (21) and (23) that the nonresonant amplitudes of the concurrent plasma and other waves, as well as the derivatives of the resonant amplitudes, are of the order of $\omega_b^2/\gamma^3\Omega^2$, i.e., small. Thus, in the expansion of the field equations in terms of the small parameter we confine ourselves to terms up to this order of smallness, proportional to the amplitudes of the resonant transverse harmonics. To identify the resonant transverse harmonic, we expand the amplitudes $C_{0,1,\alpha}$ in terms of the waveguide eigenfunctions φ_n:

$$C_{0,1,\alpha}=\sum_{n=1}^{\infty}A_{0,1,\alpha}^{(n)}\varphi_n. \tag{25}$$

6. For a more compact formulation of these equations we introduce the dimensionless variables:

$$x=k_z^{(0)}z,\quad \tau=\omega^{(0)}t,\quad \Delta=\frac{\mathscr{D}_{1\alpha}}{\omega^{(0)}},\quad l=\frac{k_{z0}}{k_z^{(0)}},$$

$$x_j=k_z^{(0)}z_j,\quad y_j=\frac{k_z^{(0)}v_j'}{\omega^{(0)}},\quad \kappa_{0,1,\alpha}^2=k_z^2-\frac{\omega_{0,1,\alpha}^2}{c^2}, \tag{26}$$

$$a_{0,1,\alpha}=\frac{e}{m}\frac{k_z^{(0)}}{\gamma^3}\sqrt{\frac{\kappa_{0,1,\alpha}^2\Delta}{(\omega_b^2/\gamma^3)\Omega_bG_1}\frac{\partial D_{0,1,\alpha}^{n=1}}{\partial\omega_{0,1,\alpha}}}\varphi_1(r_b)A_{0,1,\alpha}^{(1)}.$$

As a result, we obtain a system of equations for the wave amplitudes (after first replacing $iA_\alpha^{(1)}$ by $A_\alpha^{(1)}$):

$$\frac{\partial a_0}{\partial\tau}+\frac{v_{g0}}{u}\frac{\partial a_0}{\partial x}=\nu_0\rho,$$

$$\frac{\partial a_1}{\partial\tau}-\frac{v_{g0}}{u}\frac{\partial a_1}{\partial x}-i\delta_1a_1=-\nu_{1\alpha}\hat{\rho}_{1\alpha}(1-2i\hat{L}_\alpha)a_\alpha e^{i\Delta\tau}, \tag{27}$$

$$\frac{\partial a_\alpha}{\partial\tau}+\frac{v_{g\alpha}}{u}\frac{\partial a_\alpha}{\partial x}-i\delta_\alpha a_\alpha=-\nu_{1\alpha}\hat{\rho}_{1\alpha}^*(1-2i\hat{L}_1)a_1e^{-i\Delta\tau},$$

where $v_{g0,1,\alpha}$ is the group velocity of the wave $\omega_{0,1,\alpha}$:

$$v_{g0,1,\alpha}=\frac{\omega_{0,1,\alpha}}{k_{z0,1,\alpha}}\frac{\omega_p^2-\omega_{0,1,\alpha}^2}{\omega_p^2-\omega_{0,1,\alpha}^4/k_{z0,1,\alpha}^2c^2}. \tag{28}$$

Note that in deriving the second equation in the system (27), we assumed that the counterpropagating wave has the same frequency as the concurrent wave, but that it propagates in the opposite direction, i.e., $v_{g1}=-v_{g0}$. In the system (27), ρ, $\rho_{1\alpha}$, and $\hat{\rho}_{1,\alpha}$ as before express the perturbations of the charge density by the concurrent plasma and Raman waves, and can be calculated using the Eqs. (22), (20), and (24), respectively; $\hat{L}_{0,1,\alpha}$ is a differential operator describing the rearrangement of the polarization of the appropriate wave:

$$\hat{L}_{0,1,\alpha}=\frac{1}{\kappa_{0,1,\alpha}^2}\left(k_{z0,1,\alpha}\frac{\partial}{\partial z}+\frac{\omega_{0,1,\alpha}}{c^2}\frac{\partial}{\partial t}\right), \tag{29}$$

and for the coefficients ν_0 and $\nu_{1\alpha}$ we have

$$\nu_0=\frac{1}{\gamma^4}\sqrt{\frac{(\omega_b^2/\gamma^3)G_1\Delta}{\Omega_bu^2\dfrac{\partial D_0^{n=1}}{\partial\omega_0}}},$$

$$\nu_{1\alpha}=\frac{1}{2}G_1\frac{\omega_b^2/\gamma^3}{\Omega^2\omega^{(0)}}\left|\frac{1}{\kappa_1^2}\frac{\partial D_1^{n=1}}{\partial\omega_1}\frac{1}{\kappa_\alpha^2}\frac{\partial D_\alpha^{n=1}}{\partial\omega_\alpha}\right|^{-1/2}. \tag{30}$$

The expressions for δ_1 and δ_α, which determine the nonlinear frequency shift of the waveguide modes[2] have the form

$$\delta_{1,\alpha}=\frac{1}{4}\left(\frac{\omega_b^2}{\Omega^2\gamma^3}\right)^2\frac{\kappa_1^2\kappa_\alpha^2}{k_{\perp 1}\omega^{(0)}}\frac{G_1^2(G_{\rho 1}-1)}{\dfrac{\partial D_{1,\alpha}^{n=1}}{\partial\omega_{1,\alpha}}}|\hat{\rho}_{1\alpha}|^2, \tag{31}$$

where expressions for the geometric factor $G_{\rho 1}$ of the beam space charge can be found in Ref. 17.

7. Transforming to the rest frame, we obtain the following equations of motion for the beam electrons:

$$\frac{dx_j}{d\tau}=y_j,$$

$$\frac{dy_j}{d\tau}=-w_j\nu_0\gamma^4\left(\frac{l}{2}a_0+i\alpha_{p1}\frac{\partial a_0}{\partial\tau}\right)\exp[il(x_j-\tau)]$$

$$+\frac{i}{2}w_j\alpha_{b1}l\gamma\rho\exp[il(x_j-\tau)]$$

$$-\frac{1}{2}w_j\Delta^2\rho_{1\alpha}\exp[i(x_j-\tau)]$$

$$+\frac{i}{8}\delta_bw_j^2\hat{\rho}_{1\alpha}\exp[i(x_j-\tau)-i\Delta\tau]$$

$$+w_j^2\nu_{1\alpha}\left(\frac{1}{2}a_1a_\alpha^*+a_1\hat{L}_\alpha a_\alpha^*-a_\alpha^*\hat{L}_1a_1\right)$$

$$\times\exp[i(x_j-\tau)-i\Delta\tau]+\text{c.c.}. \tag{32}$$

Here the amplitudes of all the waves are taken at the location of the jth particle, i.e., at $x=x_j$. For the parameters α_{bm} and α_{pm} we have

$$\alpha_{bm}=\frac{\omega_b^2/\gamma^3}{k_{\perp m}^2 u^2 \gamma^2}, \quad \alpha_{pm}=\frac{\omega_p^2}{k_{\perp m}^2 u^2 \gamma^2}. \tag{33}$$

On the right-hand side of the second of Eqs. (32), the expressions for the forces acting on the particle contain terms corresponding to the harmonics of the concurrent plasma wave a_0 and the beam space-charge density wave ρ. These harmonics correspond to the Čerenkov resonance in the system. In addition, Eq. (32) allows for forces acting on the particle from the Raman harmonics of the waves $a_1 a_\alpha^*$ of the plasma waveguide and the corresponding harmonic of the beam space-charge density $\rho_{1\alpha}$. The value of δ_b in Eq. (32) determines the correction to the rf space-charge force of the beam[2] and ultimately gives the frequency shift of the charge-density beam wave:

$$\delta_b=\left(\frac{\omega_b^2}{\Omega^2\gamma^3}\right)^2 G_1^2(G_{\rho 1}-1)\,\frac{\kappa_1^2\kappa_\alpha^2}{k_{\perp 1}u^2}\,\frac{\Omega_b}{k^{(0)2}}$$

$$\times\left[\frac{|a_1|^2}{\dfrac{\partial D_1^{n=1}}{\partial\omega_1}}+\frac{|a_\alpha|^2}{\dfrac{\partial D_\alpha^{n=1}}{\partial\omega_\alpha}}\right]. \tag{34}$$

Equations (27) and (32) describe the nonlinear dynamics of the interaction between the counterpropagating wave (ω_1,k_{z1}) and the wave $(\omega_\alpha,k_{z\alpha})$. The counterpropagating wave was formed by reflection of the concurrent plasma wave produced by injection of a relativistic electron beam into a waveguide of finite length filled with a transversely homogeneous plasma.

For a low-density weakly relativistic beam $(w_j\sim 1)$, terms of the first approximation with respect to the parameter $\omega_b^2/\gamma^3\Omega^2$ can be taken into account in Eqs. (27) and (32). In this case, the systems (27) and (32) are slightly simplified

$$\frac{\partial a_0}{\partial\tau}+\frac{v_{g0}}{u}\frac{\partial a_0}{\partial x}=\nu_0\rho,$$

$$\frac{\partial a_1}{\partial\tau}-\frac{v_{g0}}{u}\frac{\partial a_1}{\partial x}=-\nu_{1\alpha}\rho_{1\alpha}a_\alpha e^{i\Delta\tau},$$

$$\frac{\partial a_\alpha}{\partial\tau}+\frac{v_{g\alpha}}{u}\frac{\partial a_\alpha}{\partial x}=-\nu_{1\alpha}\rho_{1\alpha}^* a_1 e^{-i\Delta\tau}, \tag{35}$$

$$\frac{dy_j}{d\tau}=-\nu_0\gamma^4\left(\frac{l}{2}a_0+i\alpha_{p1}\frac{\partial a_0}{\partial\tau}\right)\exp[il(x_j-\tau)]$$

$$+\frac{i}{2}\alpha_{b1}l\gamma\rho\,\exp[il(x_j-\tau)]-\frac{i}{2}\Delta^2\rho_{1\alpha}$$

$$\times\exp[i(x_j-\tau)]+\frac{1}{2}\nu_{1\alpha}a_1 a_\alpha^*$$

$$\times\exp[i(x_j-\tau)-i\Delta\tau]+\text{c.c.}$$

These equations are subsequently analyzed numerically. The system (35) is supplemented by the boundary conditions

$$a_1(x=k_{z0}L,\tau)=\kappa a_0(x=k_{z0}L,\tau),$$

$$a_0(x=0,\tau)=a_1(x=0,\tau), \tag{36}$$

where κ is the coefficient of conversion of the concurrent wave into the counterpropagating wave leaving the system.

To conclude our derivation of the basic equations, we clarify once again the role of the counterpropagating wave in the processes in the system. First, it performs feedback, i.e., it couples some of the output signal back into the system. Second, as a result of scattering by the electron beam, accompanied by a change in frequency to a different mode, the counterpropagating wave can alter the mechanism, or at least strongly influence the beam-plasma interaction.

8. In view of the complexity of (35), the processes in a system described by these relations subject to the boundary conditions (36) can only be described by numerical simulation.

We recall the dependence of the Čerenkov radiation dynamics on the parameters of the system when the role of the counterpropagating wave is merely reduced to feedback.[1] The main parameter influencing the processes in the oscillator when the length of the system and the reflection coefficient are fixed, is the beam current or the coupling parameter ν_0 of the beam electrons with the concurrent plasma wave. It is known that a plasma oscillator system undergoes self-excitation when the beam current exceeds a certain threshold[2] determined by the length of the system and the reflection coefficient. At electron beam currents below threshold, the perturbation of the plasma oscillations introduced by the beam front decays with time, and the oscillator does not turn on. At currents slightly above threshold, the moduli of the wave amplitudes exhibit a steady-state distribution corresponding to the optimum level of oscillation. In this case, the beam–plasma instability is stabilized near the emitting device.

Note that the saturation mechanism for homogeneous plasma filling is trapping of beam electrons by the plasma wave. As the beam current increases, the coupling between the electron beam and the plasma wave field increases, and the amplitude of this wave undergoes more rapid spatial growth. The maximum amplitude corresponding to the point of electron beam trapping is shifted inside the system and a lower output signal level is observed. The counterpropagating wave, providing feedback, leads to chaotic oscillations of the saturation point of the beam–plasma instability far inside the system, and thus results in randomization of the output signal.

The spatial distribution of the moduli of the wave amplitudes in a system with $L\approx 12$ cm and $\nu_0=0.005$ is shown in Fig. 2. Since the maximum of the modulus of the concurrent plasma wave amplitude is near the emitting device, the beam current is close to optimal in this case. After transition processes have taken place in the waveguide, a given distribution of wave amplitudes is established and will be conserved in the system, provided that any scattering of the counterpropagating wave by beam electrons is neglected. Allowance for this scattering has the result that the dynamics of the generation of a plasma wave in Čerenkov resonance with

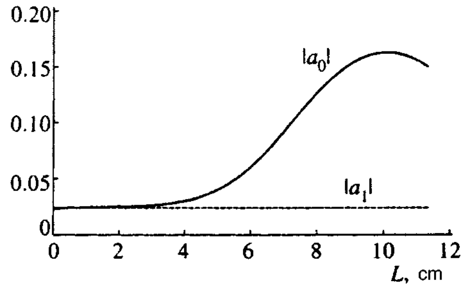

FIG. 2. Spatial distribution of the moduli of the wave amplitudes a_0 (solid curve) and a_1 (dashed curve) for $\nu_0=0.005$, and for the case where nonlinear wave interaction is neglected ($\nu_{12}=0$).

the beam is destroyed as a result of nonlinear wave interaction.

For the actual parameters of the system, the growth rates of the instability associated with nonlinear interaction between the counterpropagating wave and one of the waveguide modes $(\omega_\alpha, k_{z\alpha})$ are small compared with the Čerenkov instability growth rate. The small growth rates (or long times required for the buildup of instability associated with nonlinear interaction between the counterpropagating plasma wave and the $(\omega_\alpha, k_{z\alpha})$ wave) make it possible to observe this effect only in a limited number of cases. Since the rf waveguide modes $(\alpha=3,4)$ have fairly high group velocities (except for the case $k_{z3}\approx 0$, $v_{g3}\approx 0$ near the cutoff frequency), the resulting local perturbations of the $(\omega_\alpha, k_{z\alpha})$ wave are removed fairly rapidly from the system. The buildup of instability in this case requires long interaction lengths greater than the length of the system L.

In other words, processes involving the plasma wave (ω_2, k_{z2}) develop. Its extremely low group velocity and high coefficient of reflection from the emitting horn leads to the buildup of oscillations at the frequency ω_2, trapped in the bulk of the plasma. This may result in disruption of the oscillator operating conditions as a result of Čerenkov instability, as was observed numerically. An analysis was made of the excitation of a plasma waveguide by a beam front with possible Raman resonance between the counterpropagating plasma wave and the (ω_2, k_{z2}) plasma wave, corresponding to point 2 in Fig. 1.

Figure 3 gives the results of a numerical simulation of the processes in the system for $\nu_0=0.005$, $\nu_{12}=0.01$. Figure 3a gives the time dependence of the modulus of the amplitude of the concurrent plasma wave at the output, $|a_0(z=L,t)|$. Figure 3b gives the time dependence of the average total kinetic losses of the beam

$$\eta(t)=1-\frac{W_{\rm in}+W_{\rm out}}{W_{\rm inp}}, \tag{37}$$

where $W_{\rm inp}$, $W_{\rm in}$, and $W_{\rm out}$ are the total kinetic losses of large particles injected into the waveguide at time t, inside the system and after emergence from the system, respectively. Curves *1* in Fig. 3 are plotted for the system neglecting emission of the plasma wave (ω_2, k_{z2}), i.e., for $\nu_{12}=0$. It can be seen that after transients have ended, a steady-state distribution of wave amplitude moduli is established in the system. Curves *2* in Fig. 3 correspond to the case $\nu_{12}=0.01$ and $\Delta\approx 0$, i.e., they show how the emission regime is affected by taking into account interaction between the counterpropagating plasma wave and the (ω_2, k_{z2}) plasma wave in Raman resonance. It is observed that a steady-state distribution of the modulus of the concurrent plasma wave is established. However the output amplitude of this wave is approximately a third lower than that for the case $\nu_{12}=0$ (curve *1* in Fig. 3a), and the time taken to establish this steady-state distribution is considerably greater than the time taken to establish the steady-state distribution in the first case.

It can be seen from Fig. 3b that until steady-state generation of a wave in Čerenkov resonance with the beam electrons is established ($t<20$ ns), the kinetic losses of the beam for these cases are essentially the same. During the transient processes for the nonlinear wave interaction until a steady-state output signal is established, corresponding to curve *2* in Fig. 3a ($20\ \text{ns}<t<40\ \text{ns}$), the beam on average loses slightly less energy than that neglecting stimulated emission of the waves at frequencies ω_1 and ω_2. During this period, transient processes are observed in the system associated with redistribution of energy between the waves. At this stage, the wave amplitude a_2 begins to grow. After this growth has ceased ($t\sim 50$ ns), a dynamic equilibrium distribution of the moduli of the amplitudes a_0, a_1, and a_2 is established in the system—a steady-state amplitude distribution is observed. In this case, the kinetic losses are higher than those for the steady-state distribution of the two waves a_0 and a_1 (see Fig. 2), since additional losses result from stimulated emission of the (ω_2, k_{z2}) wave.

Thus, in the first stage, Čerenkov instability builds up in the system, the beam is modulated at the Čerenkov resonance frequency, and a counterpropagating wave appears in the resonator. In addition to the fundamental harmonic

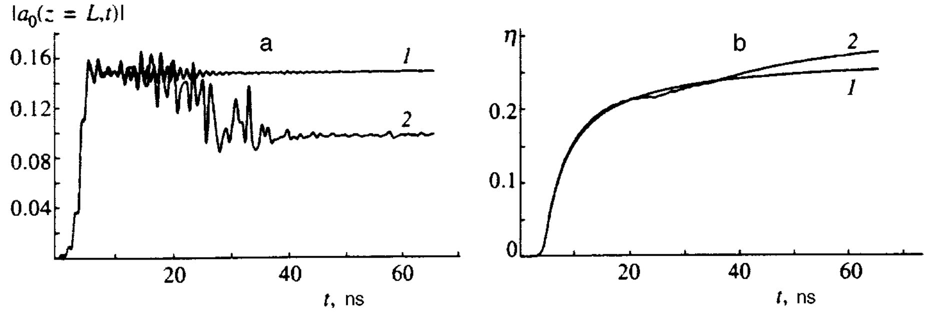

FIG. 3. Evolution of the output signal (a) and total kinetic losses of the beam (b) for $\nu_0=0.005$: *1*—$\nu_{12}=0$, *2*—$\nu_{12}=0.01$.

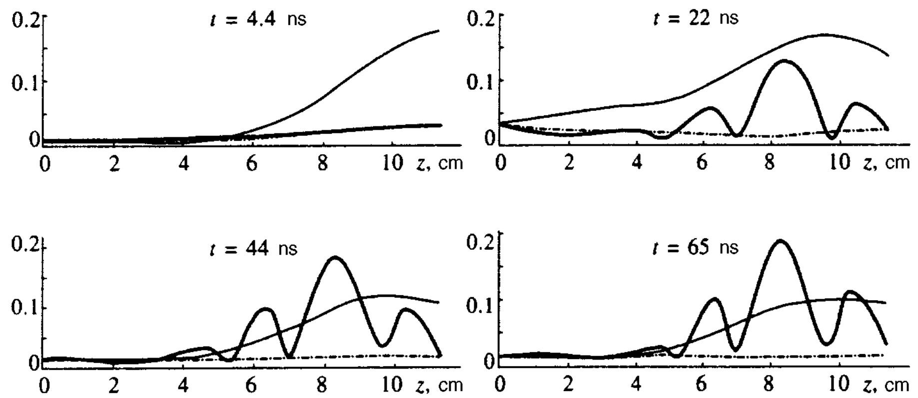

FIG. 4. Evolution of the spatial distribution of the moduli of the wave amplitudes a_0 (fine line), a_1 (dashed curve), and a_2 (heavy line) for $\nu_0=0.005$ and $\nu_{12}=0.01$.

(ω_0,k_{z0}), the density-modulated beam also contains the Raman harmonic of the beam charge density wave, which subsequently leads to stimulated emission of the (ω_2,k_{z2}) wave. The counterpropagating plasma wave, implementing feedback in the oscillator as before, also participates in this stimulated emission.

Figure 4 shows the evolution of the spatial distribution of the wave amplitudes in the system. It can be seen that after the wave amplitude a_2 has increased to a certain level, it remains almost unchanged for $t>40$ ns. Thus, in addition to the concurrent and counterpropagating plasma waves, the system also contains spatially localized oscillations at frequency ω_2 trapped in the bulk of the plasma $(v_{g2}\approx 0,\ \kappa\sim 1)$.

To conclude, it has been shown that the oscillation regime in a waveguide of finite length filled with a homogeneous plasma is strongly influenced by nonlinear interaction between the counterpropagating plasma wave and another mode of the plasma waveguide (a highly potential plasma mode, i.e., a wave with a phase velocity much lower than the velocity of light), which are in Raman resonance. A numerical simulation has been used to show that the efficiency of the output radiation of the concurrent plasma wave is reduced because some of the energy lost by the beam is transferred to the (ω_2,k_{z2}) wave trapped in the bulk of the plasma.

[1] M. A. Krasil'nikov, M. V. Kuzelev, and A. A. Rukhadze, Zh. Éksp. Teor. Fiz. **108**, 521 (1995) [JETP **81**, 280 (1995)].

[2] M. V. Kuzelev and A. A. Rukhadze, *Electrodynamics of Dense Electron Beams in Plasma* [in Russian], Nauka, Moscow (1990).

[3] J. Weiland and H. Wilhelmsson, *Coherent Nonlinear Interaction of Waves in Plasmas*, Pergamon Press, Oxford (1976).

[4] L. S. Bogdankevich, M. V. Kuzelev, and A. A. Rukhadze, Usp. Fiz. Nauk **133**, 3 (1981) [Sov. Phys. Usp. **24**, 1 (1981)].

[5] V. L. Bratman, N. S. Ginzburg, and M. I. Petelin, Zh. Éksp. Teor. Fiz. **76**, 930 (1979) [Sov. Phys. JETP **49**, 469 (1979)].

[6] I. F. Kharchenko, Ya. B. Faĭnberg, R. M. Nikolaev *et al.*, Zh. Éksp. Teor. Fiz. **38**, 687 (1960) [*sic*].

[7] R. A. Demirkhanov, A. K. Gevorkov, A. F. Popov, and G. I. Zverev, Zh. Tekh. Fiz. **30**, 306 (1982) [*sic*].

[8] M. V. Kuzelev, A. R. Maĭkov, A. D. Poezd *et al.*, Dokl. Akad. Nauk SSSR **300**, 1112 (1988) [Sov. Phys. Dokl. **33**, 444 (1988)].

[9] V. V. Migulin, V. I. Medvedev, E. R. Mustel', and V. N. Parygin, *Principles of Oscillation Theory* [in Russian], Nauka, Moscow (1988).

[10] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media*, Nauka, Moscow (1992) [second english ed., Pergamon, New York (1984)].

[11] B. Z. Katsenelembaum, *Radio-frequency Electrodynamics* [in Russian], Nauka, Moscow (1966).

[12] Yu. L. Klimontovich, *Kinetic Theory of Nonideal Gases and Nonideal Plasmas*, Nauka, Moscow, 1975 (Pergamon Press, Oxford, 1982).

[13] Yu. A. Berezin and M. P. Fedoruk, *Modeling of Time-Dependent Plasma Processes* [in Russian], Nauka, Novosibirsk (1993).

[14] Yu. A. Berezin and V. A. Vshivkov, *The Method of Particles in the Dynamics of Tenuous Plasma* [in Russian], Nauka, Novosibirsk (1980).

[15] A. G. Litvak, V. I. Petrukhina, and V. I. Trakhtenberg, JETP Lett. **18**, 111 (1973).

[16] M. V. Kuzelev and V. A. Panin, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **27**, 426 (1984).

[17] M. V. Kuzelev, V. A. Panin, A. A. Rukhadze, and D. S. Filippychev, Pis'ma Zh. Tekh. Fiz. **10**, 228 (1984) [Sov. Tech. Phys. Lett. **10**, 95 (1984)].

[18] A. N. Nayfeh, *Introduction to Perturbation Techniques*, Wiley, New York (1981).

Translated by R. M. Durham

Diffusion and transport of a magnetic field in a turbulent medium with helicity

N. A. Silant'ev

Central Astronomical Observatory, Russian Academy of Sciences, 196140 St. Petersburg, Russia;
Instituto National de Astrofisica, Optica y Electronica, Puebla, Mexico

(Submitted 15 April 1997)
Zh. Éksp. Teor. Fiz. **112**, 1312–1331 (October 1997)

An analysis is made of the relation between accurate formulas for the coefficients of turbulent diffusion D_T and the alpha effect α_T for a magnetic field in the Lagrange and Euler representations. It is shown that the quadratic term with respect to α_T in the diffusion coefficient derived by Moffatt and Kraichnan is incorrect and should be dropped. First, a numerical solution of the nonlinear equation (DIA equation) for the Green function is presented, describing the transport of a magnetic field for the case of incompressible, uniform, isotropic, steady-state turbulence possessing helicity. These solutions are used to calculate the steady-state coefficients D_T and α_T for various values of the parameters $\xi_0=u_0\tau_0/R_0$, $a=H_0/u_0^2p_0$, τ_0/τ_1, and R_0/R_1, where u_0, τ_0, and R_0 are the characteristic velocity, lifetime, and scale of the turbulent pulsations, and H_0, τ_1, and R_1 are similar values describing the helicity of the medium $h(1,2)=\langle\mathbf{u}(1)\cdot(\nabla\times\mathbf{u}(2))\rangle$, and the parameter α characterizes the degree of helicity. The DIA values of D_T and α_T and the self-consistent values of these quantities calculated using the Green tensor in the diffusion approximation are in qualitative agreement. It is shown that the coefficient of turbulent diffusion is always positive for all the types of turbulence studied. Nonsteady-state values of $D_T(t)$ and $\alpha_T(t)$ calculated by a self-consistent method are given. © *1997 American Institute of Physics.* [S1063-7761(97)01210-9]

1. INTRODUCTION

The problem of the transport and amplification of a magnetic field in turbulent conducting media has been studied in magnetohydrodynamics, astrophysics, and geophysics for many years. A detailed review of works on various aspects of this problem can be found in a number of books.[1–4] In view of the obvious complexity of the problem arising from the vector nature of the magnetic field, various plausible assumptions as to the behavior of the average magnetic field $\langle\mathbf{B}(\mathbf{r},t)\rangle$ have been used in the calculations. In particular, the diffusion approximation has been widely used, where it is assumed that $\langle\mathbf{B}(\mathbf{r},t)\rangle$ is a smooth function on the characteristic scales R_0 and correlation times τ_0 of the turbulent velocity field $\mathbf{u}(\mathbf{r},t)$ of a conducting medium. The presence of an average nonzero helicity $H_0=\langle\mathbf{u}(\mathbf{r},t)\cdot(\nabla\times\mathbf{u}(\mathbf{r},t))\rangle\neq 0$ results in amplification of the initial magnetic field (the so-called α effect). Nonzero helicity of the turbulent motion may be caused by the general rotational motion of the conducting medium (the rotation of the Earth, the sun, and so on). The kinematic formulation of the problem, where the magnetic field is assumed to be fairly weak, and does not influence the known turbulent motion of the liquid or gas, is only suitable for a turbulent medium with helicity over the time interval before the energy of the growing magnetic field becomes equal to the kinetic energy of the moving gas elements.

In the diffusion approximation, the evolution of $\langle\mathbf{B}(\mathbf{r},t)\rangle$ in an isotropic turbulent medium is described by the well-known equation

$$\left(\frac{\partial}{\partial t}-D_m\nabla^2\right)\langle\mathbf{B}(\mathbf{r},t)\rangle=\nabla\times[\alpha_T(\mathbf{r},t)\langle\mathbf{B}(\mathbf{r},t)\rangle]-\nabla\times[D_T(\mathbf{r},t)\langle\mathbf{B}(\mathbf{r},t)\rangle], \tag{1}$$

where $D_m=c^2/4\pi\sigma$ is the molecular diffusion coefficient, σ is the conductivity of the medium, D_t is the coefficient of turbulent diffusion, and α_T is a coefficient describing the amplification of $\langle\mathbf{B}(\mathbf{r},t)\rangle$ as a result of the α effect. The estimates $D_T\cong u_0R_0$ and $\alpha_T\cong u_0$ are usually used in applications, leaving a high degree of arbitrariness in specific calculations of the average magnetic field evolution. Attempts to obtain more rigorous calculations of D_T and α_T for given ensembles of the turbulent velocity $\mathbf{u}(\mathbf{r},t)$ in the Euler representation are therefore of interest. The first accurate value $D_T=u_0\tau_0/3$ was obtained in Ref. 5 for a short-range time-correlated velocity field (the correlation function is $\langle u_i(\mathbf{r},t)u_j(\mathbf{r}',t')\rangle\propto\delta(t-t')$). A more general analysis of the problem of calculating D_T, also allowing for the compressibility of the medium, was presented in Ref. 6.

The molecular diffusion coefficient is usually $D_m\ll D_T$ and may be neglected when finding D_T and α_T. In this case, it is easy to obtain accurate expressions for D_T and α_T by using the Lagrange representation for the velocity field $\mathbf{v}(\mathbf{a},t)$, where $\mathbf{a}$ is the coordinate of a liquid particle at zero time $t=0$. The transport equations for a scalar impurity with concentration $n(\mathbf{r},t)=\langle n\rangle+\mathbf{n}'$ $(\langle\mathbf{n}'\rangle=0)$ and the equation for the magnetic field induction $\mathbf{B}(\mathbf{r},t)=\langle\mathbf{B}\rangle+\mathbf{B}'$ $(\langle\mathbf{B}'\rangle=0)$ in the case $D_m=0$ have the form

$$\partial n(\mathbf{r},t)/\partial t=\nabla\cdot[\mathbf{u}(\mathbf{r},t)n(\mathbf{r},t)], \tag{2}$$

$$\partial\mathbf{B}(\mathbf{r},t)/\partial t=\nabla\times[\mathbf{u}(\mathbf{r},t)\mathbf{B}(\mathbf{r},t)],\quad \nabla\cdot\mathbf{B}=0,\quad \nabla\cdot\mathbf{u}=0. \tag{3}$$

The solution of these equations for a given Lagrange velocity field $\mathbf{v}(\mathbf{a},t)$ has the form[1,4]

$$n(\mathbf{r},t)=n_0(\mathbf{r}-\mathbf{X}(\mathbf{a},t))/D(\mathbf{a},t)\equiv n_0(\mathbf{a})/D(\mathbf{a},t). \tag{4}$$

$$B_i(\mathbf{r},t)=\frac{\partial x_i}{\partial a_j}B_j^{(0)}(\mathbf{r}-\mathbf{X}(\mathbf{a},t))\equiv\frac{\partial x_i}{\partial a_j}B_j^{(0)}(\mathbf{a}), \tag{5}$$

where $\mathbf{r}=(x_1,x_2,x_3)$ and

$$\begin{aligned}
&\mathbf{r}=\mathbf{a}+\mathbf{X}(\mathbf{a},t)\equiv\mathbf{a}+\int_0^t d\tau\,\mathbf{v}(\mathbf{a},\tau),\\
&\partial\mathbf{r}/\partial t\equiv\mathbf{u}(\mathbf{r},t)\equiv\mathbf{v}(\mathbf{a},t), && (6)\\
&D_{ij}(\mathbf{a},t)=\partial x_i/\partial a_j, \quad D(\mathbf{a},t)=\det D_{ij},\\
&\dot D_{ij}\equiv\partial D_{ij}/\partial t, \quad \dot D/D=\nabla\cdot\mathbf{u}(\mathbf{r},t),\\
&D_{sp}\dot D_{pq}^{-1}=-\frac{\partial v_s}{\partial a_q}, \quad d\mathbf{r}=D(\mathbf{a},t)d\mathbf{a}. && (7)
\end{aligned}$$

As usual, summation is implied over repeated indices. The values $n_0(\mathbf{a})$ and $\mathbf{B}^{(0)}(\mathbf{a})$ denote the concentration and magnetic field at time zero. Equation (4) expresses conservation of the number of impurity particles ($n(\mathbf{r},t)d\mathbf{r}=n_0(\mathbf{a})d\mathbf{a}$) as these are transported along a flow tube of variable cross section. Expression (5) describes the variation in the magnetic field frozen into an element of liquid in an incompressible medium as a result of convective transport, and also as a result of the deformations and rotation of an element of liquid or gas.

Defining a random field of Lagrange velocities $\mathbf{v}(\mathbf{a},t)$ determines an ensemble of realizations of turbulent flows over which the solutions (4) and (5) must be averaged to obtain equations for the average $\langle n(\mathbf{r},t)\rangle$ and $\langle\mathbf{B}(\mathbf{r},t)\rangle$. For fixed $\mathbf{r}$ and t the value $\mathbf{a}\equiv\mathbf{r}-\mathbf{X}(\mathbf{a},t)$ is a random quantity and the problem reduces to averaging the known functions $n_0(\mathbf{a})$ and $\mathbf{B}^{(0)}(\mathbf{a})$ of the random argument $\mathbf{a}$ with the random weighting factors $1/D(\mathbf{a},t)$ and $D_{ij}(\mathbf{a},t)$. The transition to the diffusion approximation presupposes that $n_0(\mathbf{a})$ and $\mathbf{B}^{(0)}(\mathbf{a})$ are smooth over the characteristic lengths $R_{(0)}\cong\sqrt{\langle X^2\rangle}$ and involves using a Taylor expansion as a series in $\mathbf{X}(\mathbf{a},t)$ (Refs. 7 and 8). As a result, we obtain accurate expressions for D_T and α_T.

For a scalar impurity field in an isotropic medium we have

$$D_T=\frac{1}{3}\int_0^t d\tau\,\langle\mathbf{v}(\mathbf{a},t)\cdot\mathbf{v}(\mathbf{a},\tau)/D(\mathbf{a},t)\rangle \tag{8}$$

where the angle brackets denote averaging over the ensemble of velocity realizations. Equation (8) was obtained for an incompressible medium ($D(\mathbf{a},t)\equiv 1$ by Taylor[9] as early as 1921. Equations for D_T and α_T for a magnetic field were obtained much later:[7,8]

$$\begin{aligned}
D_T(t)=&\frac{1}{3}\int_0^t d\tau\,\langle\mathbf{v}(\mathbf{a},t)\cdot\mathbf{v}(\mathbf{a},\tau)\rangle\\
&+\frac{1}{6}\int_0^t d\tau\int_0^t d\tau'\left[\left\langle\mathbf{v}(\mathbf{a},t)\cdot\mathbf{v}(\mathbf{a},\tau)\,\frac{\partial v_i(\mathbf{a},\tau')}{\partial a_i}\right\rangle\right.\\
&\left.-\left\langle v_i(\mathbf{a},t)\,\frac{\partial}{\partial a_i}(\mathbf{v}(\mathbf{a},\tau)\cdot\mathbf{v}(\mathbf{a},\tau'))\right\rangle\right]\\
&+\alpha_T(t)\int_0^t d\tau\,\alpha_T(\tau),
\end{aligned} \tag{9}$$

$$\begin{aligned}
\alpha_T(t)&=-\frac{1}{3}e_{ijm}\int_0^t d\tau\left\langle v_i(\mathbf{a},t)\,\frac{\partial v_j(\mathbf{a},\tau)}{\partial a_m}\right\rangle\\
&\equiv-\frac{1}{3}\int_0^t d\tau\,\langle\mathbf{v}(\mathbf{a},t)\cdot(\nabla\times\mathbf{v}(\mathbf{a},\tau))\rangle.
\end{aligned} \tag{10}$$

Here $e_{ijn}=-e_{jin}=-e_{inj}$ is a unit pseudotensor ($e_{123}=1$, and so on).

The first term in expression (9) is the same as D_T for a scalar impurity, and is due to pure convective transport of the impurity together with an element of liquid or gas. The second term in expression (9) determines the contribution to the diffusion made by the deformations and rotations of an element of the medium. Particular mention should be made of the meaning of the third term in expression (9), which is proportional to α_T^2. This term appears as a result of allowance for the increase in the average magnetic field caused by the α effect. It was derived using the solution of Eq. (1), where the diffusion terms were dropped (terms containing D_m and D_T). This approximate solution is only suitable at early times, when $\alpha_T\propto t$ and $D_T\propto t$, i.e., the term containing α_T^2 gives a negligible correction $\propto t^3$. We shall show subsequently that at late times this term does not exist. The formal use of this term for $t\to\infty$ causes D_T to increase without bound ($D_T\propto\alpha^2 t\to\infty$), i.e., produces a physically absurd result.

We also note that Parker[4] neglected the dependence of $\mathbf{B}^{(0)}(\mathbf{a})$ on $\mathbf{X}(\mathbf{a},t)$ when averaging formula (5) and obtained, for an isotropic medium where $\langle\partial x_i/\partial a_j\rangle=\delta_{ij}$, the relation $\langle\mathbf{B}(\mathbf{r},t)\rangle=\langle\mathbf{B}^{(0)}(\mathbf{a})\rangle$, which is exactly the same as the expression $\langle n(\mathbf{r},t)\rangle=\langle n_0(\mathbf{a})\rangle$ for a scalar impurity in an incompressible medium. This led to the incorrect conclusion that D_T for a magnetic field in an isotropic medium should be exactly the same as D_T for a scalar impurity. General reasoning suggests that averaging the more complex expressions contained in the second term in formula (9) will yield a smaller contribution to D_T than the first convective term. That is to say, we expect that the coefficient of turbulent diffusion of a magnetic field (in the absence of helicity) will be close to D_T for a scalar impurity.

Equations (9) and (10) were used by Kraichnan[8,10] to calculate $D_T(t)$ and $\alpha_T(t)$ for two models of turbulence with a narrow energy spectrum. A numerical simulation of liquid particle trajectories was performed up to times $t=4t_0$, where $t_0=R_0/u_0=1/u_0p_0$ is the rotation time of the liquid vortex. The calculations confirmed that D_T is close to the coefficient of turbulent diffusion of a scalar impurity for parameter values $\xi_0=u_0\tau_0/R_0=u_0\tau_0p_0\leqslant 1$. For the case of frozen and helical turbulence, where $\tau_0\gg t_0$ ($\xi_0=\tau_0t_0\to\infty$), the values of D_T were found to be negative from $t\cong 3t_0$. Our calculations (see below) performed in the Euler representation always give positive D_T. The case $\xi_0\to\infty$ is the most difficult to simulate numerically and this appears to be why this incorrect result was obtained in Ref. 10.

Equations (9) and (10) require a knowledge of the Lagrange velocities $\mathbf{v}(\mathbf{a},t)$ in a turbulent medium, whereas the correlation functions of the Euler velocity components $\mathbf{u}(\mathbf{r},t)$ are usually measured. Transforming from $\mathbf{u}(\mathbf{r},t)$ to $\mathbf{v}(\mathbf{a},t)$ involves solving a system of differential equations, and this problem has yet to be solved for the general case. Thus, we need to have accurate equations to calculate D_T and α_T in the Euler representation. Such equations were first obtained for steady-state, uniform, and isotropic turbulence in Refs. 11 and 12.

For the coefficient of turbulent diffusion D_T of a scalar impurity in a compressible medium we have[12]

$$D_T(t_1)=\frac{1}{3}\int d\mathbf{R}\int_0^{t_1}dt_2[\langle u_i(1)G(1,2)u_i(2)\rangle - \langle \mathbf{R}\cdot\mathbf{u}(1)G(1,2)\nabla\cdot\mathbf{u}(2)\rangle]. \tag{11}$$

Here and subsequently, we shall use the convenient notation

$$f(1)=f(\mathbf{r}_1,t_1),\quad f(1-2)=f(\mathbf{r}_1-\mathbf{r}_2,t_1-t_2),$$

$$dn=d\mathbf{r}_n dt_p,\quad \mathbf{R}=\mathbf{r}_1-\mathbf{r}_2,\quad \tau=t_1-t_2$$

and so on. Here $G(1,2)=G(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)$ is the Green function of Eq. (3). Its formal expression in Lagrange notation has the form

$$G(1,2)=\frac{D(\mathbf{a},t_2)}{D(\mathbf{a},t_1)}\,\delta\left(\mathbf{r}_1-\mathbf{r}_2-\int_{t_2}^{t_1}d\tau\,\mathbf{v}(\mathbf{a},\tau)\right). \tag{12}$$

Substituting this expression and the relation $\nabla\cdot\mathbf{u}(\mathbf{r},t)=\dot{D}(\mathbf{a},t)/D(\mathbf{a},t)$ into Eq. (11), we again obtain Eq. (8) for the coefficient of turbulent diffusion D_T of a scalar impurity in a compressible medium.

Expressions for D_T and α_T in the Euler representation were derived in Ref. 12:

$$\begin{aligned}D_T(t_1)=\frac{1}{6}\int d\mathbf{R}\cdot\int_0^{t_1}dt_2[&\langle u_i(1)G_{ss}(1,2)u_i(2)\rangle\\ &-\langle u_s(1)G_{is}(1,2)u_i(2)\rangle+\langle R_s u_n(1)G_{in}(1,2)\\ &\times(\nabla_i u_n(2))\rangle-\langle R_s u_i(1)G_{sn}(1,2)\\ &\times(\nabla_i u_n(2))\rangle],\end{aligned} \tag{13}$$

$$\begin{aligned}\alpha_T(t_1)=e_{ipt}\frac{1}{3}\int d\mathbf{R}\cdot\int_0^{t_1}dt_2&\langle u_i(1)G_{pn}(1,2)\\ &\times(\nabla_t u_n(2))\rangle.\end{aligned} \tag{14}$$

The Green tensor function of the induction equation (4) may be represented in the form

$$\begin{aligned}G_{ij}(1,2)=&D_{is}(\mathbf{a},t_1)D_{sj}^{-1}(\mathbf{a},t_2)\\ &\times\delta\left[\mathbf{r}_1-\mathbf{r}_2-\int_{t_2}^{t_1}d\tau\,\mathbf{v}(\mathbf{a},\tau)\right].\end{aligned} \tag{15}$$

Substituting this expression into Eqs. (13) and (14), and using the relation

$$D_{sp}^{-1}(\mathbf{a},t)\partial v_p(\mathbf{a},t)/\partial a_j=-\dot{D}_{sj}^{-1}(\mathbf{a},t)$$

(see Eq. (7)), we again obtain Eq. (9) (but without the last term!) and Eq. (10).

Thus, in order to calculate D_T and α_T for an ensemble of Euler velocities $\mathbf{u}(\mathbf{r},t)$, we need to know the Green function $G_{ij}(1,2)$ in the Euler representation and its correlation functions with the velocity components. The tensor $G_{ij}(1,2)$ is a functional of $\mathbf{u}(\mathbf{r},t)$ and, unlike the Lagrange representation (15), cannot be found in explicit analytic form. A procedure for renormalizing the integral equation for G_{ij} was proposed in Refs. 11 and 12. Substituting the iterations of the renormalized equation for G_{ij} into the basic equations (13) and (14) gives asymptotically convergent series for D_T and α_T. Averaging the series of iterations for the right-hand term in the induction equation (4) gives a hierarchy of nonlinear equations directly for the average Green function $\langle G_{ij}(1,2)\equiv g_{ij}(1-2)H(t_1-t_2)$, where $H(\tau)=1$ for $\tau>0$ and $H(\tau)=0$ for $\tau<0$. The first of the equations in the hierarchy, possessing quadratic nonlinearity, has been described in the literature as the equation in the Direct Interaction Approximation—the DIA equation. A DIA equation for a scalar impurity was proposed in Ref. 13. An investigation of the scalar DIA equation (see Refs. 14–17) showed that this can be used to calculate D_T for all values of the parameter $\xi_0=u_0\tau_0/R_0$ $(0\leqslant\xi_0<\infty)$. The difference between the DIA value of D_T and the accurate value increases monotonically from 0% for $\xi_0\to 0$ to $\cong$10% for $\xi_0\to\infty$ (frozen turbulence). The DIA equation for the tensor $\langle G_{ij}(1,2)\rangle$ in an isotropic medium reduces to a system of two nonlinear equations. Kraichnan[8] gave a pessimistic assessment of the possibility of obtaining a numerical solution of this nonlinear equation.

In the present paper we solve this system of equations numerically for the first time and we obtain DIA expressions for the steady-state values $(t\to\infty)$ of the coefficients D_T and α_T. Our values of D_T are always positive (the possibility of negative D_T arises in the DIA solutions for purely compressible potential turbulence, where $u_i(\mathbf{r},t)=\nabla_{i\varphi}(\mathbf{r},t)$). At present, the calculations of D_T and α_T using the DIA equation are the most accurate and reliable. We shall also make a detailed investigation of the dependence of D_T and α_T on the degree of helicity and its spatial and temporal scales. These relations will enable us to make a more accurate choice of D_T and α_T values, even when the correlation functions of the velocity of the medium are only known approximately and it is even difficult to select the characteristic parameters R_0 and τ_0.

2. HIERARCHY OF NONLINEAR EQUATIONS FOR THE AVERAGE GREEN TENSOR AND TRANSITION TO THE DIFFUSION APPROXIMATION

Using the method of obtaining the nonlinear equations for the Green functions $G(1,2)$ described in Refs. 11, 12, and 17, we obtain for the first three terms of the hierarchy (recall that $\langle G_{ij}(1,2)\rangle\equiv G_{ij}(1-2)$)

$$\begin{aligned}&\left(\frac{\partial}{\partial t_1}-D_m\nabla_1^2\right)G_{ij}(1-2)\\ &\quad=\delta_{ij}\delta(\mathbf{R})\delta(\tau)+\int d3\langle\hat{L}_{in}(1)G_{nm}(1-3)\end{aligned}$$

$$
\begin{aligned}
&\times \hat{L}_{ms}(3)\rangle G_{sj}(3-2)+\int d3\int d4\langle \hat{L}_{in}(1)G_{nm}(1-3)\\
&\times \hat{L}_{ms}(3)G_{sr}(3-4)\hat{L}_{rt}(4)\rangle G_{tj}(4-2)\\
&+\int d3\int d4\int d5[\langle \hat{L}_{in}(1)G_{nm}(1-3)\\
&\times \hat{L}_{ms}(3)G_{sr}(3-4)\hat{L}_{rt}(4)G_{tq}(4-5)\\
&\widehat{\times L}_{qf}(5)\rangle-\langle \hat{L}_{in}(1)G_{nm}(1-3)\langle \hat{L}_{ms}(3)G_{sr}(3-4)\\
&\times \hat{L}_{rt}(4)\rangle G_{tq}(4-5)\hat{L}_{qf}(5)\rangle-\langle \hat{L}_{in}(1)G_{nm}(1-3)\\
&\times \hat{L}_{ms}(3)\rangle G_{sr}(3-4)\langle \hat{L}_{rt}(4)G_{tq}(4-5)\\
&\times \hat{L}_{qf}(5)\rangle]G_{fj}(5-2)+\ldots. \qquad (16)
\end{aligned}
$$

Here the tensor operator $\hat{L}_{mn}(3)\equiv\nabla_m^{(3)}u_n(3)-\delta_{nm}\nabla_s^{((3)}u_s(3)$ operates on all functions of $\mathbf{r}_3$ and t_3 to its right. For Gaussian ensembles $\mathbf{u}(\mathbf{r},t)$ the second integral term vanishes and the third reduces to a single term with the cross correlation function (in time), where $L_{in}(1)$ is averaged with $L_{rt}(4)$ and $L_{ms}(3)$ is averaged with $L_{qf}(5)$. Dropping all the terms of the hierarchy except the second integral, we obtain the DIA equation. Using the equality

$$
\langle B_i(\mathbf{r},t)\rangle=\int d\mathbf{r}'\langle G_{ij}(\mathbf{r}-\mathbf{r}';t)\rangle B_j^{(0)}(\mathbf{r}'), \qquad (17)
$$

we obtain from (16) an integrodifferential equation for the average magnetic field $\langle\mathbf{B}(\mathbf{r},t)\rangle$ in a steady-state, homogeneous, isotropic medium, which can be expressed in the form

$$
\begin{aligned}
&\left(\frac{\partial}{\partial t}-D_m\nabla^2\right)\langle\mathbf{B}(\mathbf{r},t)\rangle\\
&\quad=\nabla\times\int d\mathbf{R}\int_0^t d\tau[\alpha(\mathbf{R},\tau)\langle\mathbf{B}(\mathbf{r}-\mathbf{R},t-\tau)\rangle\\
&\qquad-D(\mathbf{R},\tau)\nabla\times\langle\mathbf{B}(\mathbf{r}-\mathbf{R},t-\tau)\rangle]. \qquad (18)
\end{aligned}
$$

The form of the expression is general and does not depend on the method of terminating the hierarchy of equations for $\langle G_{ij}\rangle$.

A formal solution of Eq. (18) is easy to obtain by taking the Fourier transform with respect to $\mathbf{r}$ and the Laplace transform with respect to the time t:

$$
\langle\widetilde{\mathbf{B}}(\mathbf{p},s)\rangle=\int d\mathbf{r}\int_0^\infty dt\ \exp(-st)\exp(-i\mathbf{pr})\langle\mathbf{B}(\mathbf{r},t)\rangle. \qquad (19)
$$

Assuming that the magnetic field is solenoidal $\mathbf{p}\cdot\widetilde{\mathbf{B}}^{(0)}\times(\mathbf{p})=0$, we obtain

$$
\begin{aligned}
\langle\widetilde{B}_i(\mathbf{p},s)\rangle\equiv\langle\widetilde{G}_{ij}(\mathbf{p},s)\rangle\widetilde{B}_j^{(0)}(\mathbf{p})&=\widetilde{G}_0(p,s)\widetilde{B}_i^{(0)}(\mathbf{p})\\
&-ie_{iej}p_e\widetilde{G}_1(\mathbf{p},s)\widetilde{B}_j^{(0)}(\mathbf{p}), \qquad (20)
\end{aligned}
$$

where

$$
\begin{aligned}
&\widetilde{G}_0(\mathbf{p},s)=\frac{1}{2}(\widetilde{G}_-+\widetilde{G}_+),\quad \widetilde{G}_1=\frac{1}{2p}(\widetilde{G}_+-\widetilde{G}_-),\\
&\widetilde{G}_\mp(\mathbf{p},s)=(s+\widetilde{D}(\mathbf{p},s)p^2\mp\widetilde{\alpha}(\mathbf{p},s)p)^{-1}. \qquad (21)
\end{aligned}
$$

The expressions for $\widetilde{D}(\mathbf{p},s)$ and $\widetilde{\alpha}(\mathbf{p},s)$ in the DIA approximation have the form

$$
\widetilde{D}(\mathbf{p},s)=-ip_t[\widetilde{T}_{iti}-\widetilde{T}_{tii}+ip_n(\widetilde{S}_{timn}-\widetilde{S}_{itmn})]/2p^2, \qquad (22)
$$

$$
\begin{aligned}
\widetilde{\alpha}(\mathbf{p},s)=-e_{ims}p_sp_t[&\widetilde{T}_{itm}-\widetilde{T}_{tim}+ip_n\\
&\times(\widetilde{S}_{timn}-\widetilde{S}_{itmn})]/2p^2, \qquad (23)
\end{aligned}
$$

where

$$
\begin{aligned}
&T_{itm}(\mathbf{R},\tau)=\langle u_i(1)\langle G_{tn}(1,2)\rangle\nabla_m^{(2)}u_n(2)\rangle,\\
&S_{tinm}(\mathbf{R},\tau)=\langle u_t(1)\langle G_{im}(1,2)\rangle u_n(2)\rangle. \qquad (24)
\end{aligned}
$$

The correlation functions (24) are sharply peaked functions of $\mathbf{R}=\mathbf{r}_1-\mathbf{r}_2$ and $\tau=t_1-t_2$ on the scales $\sim R_0$ and times $\sim\tau_0$ (or $\sim t_0=R_0/u_0$ for frozen turbulence where $\tau_0\gg t_0$). For an inhomogeneous turbulent medium the functions $\alpha(\mathbf{R},\tau)$ and $D(\mathbf{R},\tau)$ depend parametrically on the average position $\mathbf{r}=(\mathbf{r}_1+\mathbf{r}_2)/2$ in the medium.

Expressions (18)–(24) hold in the general case, irrespective of the DIA approximation. In this general case, $\langle G_{tn}\rangle$ in expression (24) must be replaced by the nonaverage Green function G_{tn}, and the irreducible part must be taken. We recall that the reducible expressions contain the molecular Green function, which separates out various average parts, i.e., expressions of the type (for further details see Refs. 11, 12, and 17)

$$
\int d3\int d4\langle A(1,3)\rangle G_m(3-4)\langle B(4,2)\rangle.
$$

The diffusion equation can be obtained from Eq. (18) if we assume that $\langle\mathbf{B}(\mathbf{r},t)\rangle$ is a smooth function and we expand the integrand $\langle\mathbf{B}(\mathbf{r}-\mathbf{R},t-\tau)\rangle$ in powers of $\mathbf{R}$ and τ, retaining the zeroth- and first-order terms. Finally we obtain

$$
\begin{aligned}
\left(\frac{\partial}{\partial t}-D_m\nabla^2\right)\langle\mathbf{B}(\mathbf{r},t)\rangle=&-\alpha_T(t)\nabla\times\langle\mathbf{B}(\mathbf{r},t)\rangle\\
&-D_T^{(0)}(t)\nabla\times\nabla\times\langle\mathbf{B}(\mathbf{r},t)\rangle\\
&-\beta_T(t)\nabla\times\frac{\partial\langle\mathbf{B}(\mathbf{r},t)\rangle}{\partial t}, \qquad (25)
\end{aligned}
$$

where

$$
\begin{aligned}
&D_T^{(0)}(t)=\int d\mathbf{R}\int_0^t d\tau\ D(\mathbf{R},t)\\
&\qquad\to D_T^{(0)}(\infty)(1-\exp(-t/\tau_0)),\\
&\alpha_T(t)=\int d\mathbf{R}\int_0^t d\tau\ \alpha(\mathbf{R},t)\to\alpha_T(\infty)(1-\exp(-t/\tau_0)),
\end{aligned}
\qquad (26)
$$

$$\beta_T(t)=\int d\mathbf{R}\int_0^t d\tau\ \tau\alpha(\mathbf{R},t)\rightarrow\alpha_T(\infty)[-t\ \exp(-t/\tau_0)$$
$$+\tau_0(1-\exp(-t/\tau_0))].$$

The expressions after the arrows correspond to the model where $D(\mathbf{R},t)\propto\exp(-t/\tau_0)$ and $\alpha(\mathbf{R},t)\propto\exp(-t/\tau_0)$. It is easily confirmed that $D_T^{(0)}(t)$ and $\alpha_T(t)$ are the same as the expressions (13) and (14).

The situation is more complex for the term containing $\beta_T(t)$. The term of the first order of smallness $\alpha_T(t)\nabla\times\langle\mathbf{B}(\mathbf{r},t)\rangle$ is the dominant term on the right-hand side of Eq. (25) compared with the remaining terms of the second order of smallness. Therefore Moffatt[2,7] and Kraichnan[8] replaced $\partial\langle\mathbf{B}(\mathbf{r},t)\rangle/\partial t$ on the right-hand side by the approximate expression

$$\frac{\partial\langle\mathbf{B}(\mathbf{r},t)\rangle}{\partial t}\cong\alpha_T(t)\nabla\times\langle\mathbf{B}(\mathbf{r},t)\rangle \tag{27}$$

and obtained the usual form (1) of the diffusion equation where the coefficient of turbulent diffusion is

$$D_T(t)=D_T^{(0)}(t)+\alpha_T(t)\beta_T(t). \tag{28}$$

In the Lagrange representation, which was only used by Moffatt and Kraichnan, substitution of expression (27) yields a slightly different expression:

$$D_T(t)=D_T^{(0)}(t)+\alpha_T(t)\int_0^t d\tau\ \alpha_T(t). \tag{29}$$

This formula is the same as the more general expression (28) for short times $t\ll\tau_0$ (or $t\ll t_0$ for frozen turbulence).

It can be seen from the (26) that for short times $D_T^{(0)}\approx D_T^{(0)}(\infty)t/\tau_0$, we have $\alpha_T\approx\alpha_T(\infty)t/\tau_0$ and $\beta_T(t)\approx\alpha_T(\infty)t^2/2\tau_0$, i.e., the additional term $\alpha_T(t)\beta_T(t)$ in Eq. (28) is much smaller than the principal term $D_T^{(0)}(t)$. However, at late times ($t\gg\tau_0$ or $t\gg t_0$ for the case $\tau_0\gg t_0$), the accurate solution (21) gives a diffusion equation, with the diffusion coefficient $D_T(\infty)=D_T^{(0}(\infty)$, which contradicts the approximate expression (28) (or (29)).

We can see that the diffusion equation only accurately represents the solutions of the accurate integrodifferential equation (18) if we set $\beta_T(t)\equiv 0$, i.e., to derive the correct diffusion equation we need to take $\langle\mathbf{B}(\mathbf{r}-\mathbf{R},t-\tau)\rangle\cong\langle\mathbf{B}(\mathbf{r},t)\rangle-(\mathbf{R}\cdot\nabla)\langle\mathbf{B}(\mathbf{r},t)\rangle$ in the integral term of Eq. (18). This is clearly the frequently encountered case where it is meaningful to retain only the first term of the asymptotic expansion of the function. Note that a similar expansion with respect to τ on the right-hand side of the equations in the hierarchy (16) and allowance for the term $\tau\partial G_{ij}/\partial\tau$ leads to violation of the important normalization condition for the Green function ($\widetilde{G}_{ij}(p,\tau=0)=\delta_{ij}$). Thus, the quadratic term with respect to $\alpha_T(t)$ in expressions (9) and (29) and the similar term in expression (28) must be dropped.

3. DIA-EQUATION AND CALCULATION OF THE COEFFICIENTS D_T AND α_T

The accurate equations (13) and (14) for the coefficients D_T and α_T in the Euler representation contain correlation functions of the velocity components $u_i(\mathbf{r},t)$ and the derivatives $\nabla_n u_m(\mathbf{r},t)$ with the Green tensor function $G_{nm}(1,2)$. It was noted in the Introduction that using the solutions of the DIA equation for the average Green function $\langle G(1,2)\rangle$ instead of the accurate Green function $G(1,2)$ gives good results in calculations of the coefficient of turbulent diffusion D_T of a scalar impurity. It is important to note that the DIA solutions can yield D_T for all possible values of the parameter $\xi_0=u_0\tau_0/R_0\equiv u_0p_0\tau_0$ $(0\leqslant\xi_0<\infty)$. The coefficient D_T for a scalar impurity is determined by pure convective motion of the gas or liquid, and the deformations and rotations of the liquid elements make no contribution. The diffusion of a vector magnetic field is determined not only by the convective motion but also by the deformations and rotations of the liquid elements. This motion is more local, is described by complex correlation functions of the tensor $\nabla_n u_m(\mathbf{r},t)$, and clearly makes a far smaller contribution to D_T than the long-range convective motion, i.e., in the absence of helicity the coefficients of turbulent diffusion D_T of the magnetic field and the scalar impurity should be similar. This qualitative reasoning suggests that for the case of magnetic field transport, the DIA equation will also make a major contribution to the calculations of D_T and possibly also to the calculations of the amplification coefficient α_T. Here we confine our calculations to the steady-state DIA values of the coefficients $D_T(t)$ and $\alpha_T(t)$, which corresponds to $t\gg\tau_0$ or $t\gg t_0=R_0/u_0$ for the case $\tau_0\gg t_0$.

The Fourier transform with respect to $\mathbf{R}$ and the Laplace transform with respect to τ of the function $\langle G_{nm}(1,2)\rangle\equiv H(\tau)g_{nm}(R,\tau)\rangle$ for an isotropic turbulent medium has the form

$$\widetilde{g}_{nm}(p,s)=\delta_{nm}\widetilde{g}_0(p,s)+ie_{nmk}p_k\widetilde{g}_1(p,s)$$
$$+p_np_m\widetilde{g}_2(p,s). \tag{30}$$

The function $\widetilde{g}_2(p,s)$ can be expressed in terms of $\widetilde{g}_0(p,s)$ and $\widetilde{g}_1(p,s)$ and because of the solenoidal nature of the magnetic field $(\mathbf{p}\cdot\mathbf{B}(p)=0)$, does not appear anywhere. The DIA equation for $\langle G_{nm}(1,2)\rangle$ reduces to a system of coupled nonlinear equations for $\widetilde{g}_0(p,s)\equiv(\widetilde{g}_-(p,s)+\widetilde{g}_+(p,s))/2$ and $\widetilde{g}_1(p,s)\equiv(\widetilde{g}_+(p,s)-\widetilde{g}_-(p,s))/2p$. This system can be conveniently written in the form

$$\widetilde{g}_+(p,s)=\Bigg\{s+D_mp^2+\frac{p}{4}\int_0^\infty dq\int_0^\infty d\tau\int_{-1}^1 d\mu$$
$$\times(1-\mu^2)\exp(-s\tau)[pE(q,\tau)\widetilde{g}_0(|\mathbf{p}-\mathbf{q}|,\tau)$$
$$-pE_h(q,\tau)\widetilde{g}_1(|\mathbf{p}-\mathbf{q}|,\tau)$$
$$+(p^2+q^2-pq\mu)E(q,\tau)\widetilde{g}_1(|\mathbf{p}-\mathbf{q}|,\tau)$$
$$-E_h(q,\tau)\widetilde{g}_0(|\mathbf{p}-\mathbf{q}|,\tau)]\Bigg\}^{-1}$$

$$\tilde{g}_-(p,s)=\Bigg\{s+D_m p^2+\frac{p}{4}\int_0^\infty dq\int_0^\infty d\tau\int_{-1}^{1}d\mu$$
$$\times(1-\mu^2)\exp(-s\tau)[pE(q,\tau)\tilde{g}_0(|\mathbf{p}-\mathbf{q}|,\tau)$$
$$-pE_h(q,\tau)\tilde{g}_1(|\mathbf{p}-\mathbf{q}|,\tau)$$
$$-(p^2+q^2-pq\mu)E(q,\tau)\tilde{g}_1(|\mathbf{p}-\mathbf{q}|,\tau)$$
$$+E_h(q,\tau)\tilde{g}_0(|\mathbf{p}-\mathbf{q}|,\tau)]\Bigg\}^{-1}. \qquad (31)$$

Here $\mathbf{p}\cdot\mathbf{q}=pq\mu=pq\cos\theta$ is the scalar product of the vectors $\mathbf{p}$ and $\mathbf{q}$ and the expression $\tilde{f}(p,\tau)$ denotes the Fourier transform of the function $f(\mathbf{R},\tau)$ with respect to $\mathbf{R}$. The generalized spectra of the energy $E(p,\tau)$ and helicity $E_h(p,\tau)$ are given by

$$\langle\mathbf{u}(\mathbf{r},t)\cdot\mathbf{u}(\mathbf{r},t+\tau)\rangle\equiv\int_0^\infty dp\ E(p,\tau),$$
$$\langle\mathbf{u}(\mathbf{r},t)\cdot(\nabla\times\mathbf{u}(\mathbf{r},t+\tau))\rangle\equiv\int_0^\infty dp\ E_h(p,\tau). \qquad (32)$$

For small $p\ll p_0=1/R_0$ and $s\tau\ll1$, the equations (31) yield the diffusion expressions

$$\tilde{g}_\mp^{\text{diff}}(p,s)=[s+D_m p^2+D_T^{(0)}p^2\mp\alpha_T^{(0)}p]^{-1}, \qquad (33)$$

where the DIA steady-state coefficients of turbulent diffusion $D_T^{(0)}$ and the coefficient $\alpha_T^{(0)}$ are given by

$$D_T^{(0)}=\frac{1}{3}\int_0^\infty dp\int_0^\infty d\tau\,[E(p,\tau)\tilde{g}_0(p,\tau)$$
$$-E_h(p,\tau)\tilde{g}_1(p,\tau],$$
$$\alpha_T^{(0)}=\frac{1}{3}\int_0^\infty dp\int_0^\infty d\tau\,[p^2E(p,\tau)\tilde{g}_1(p,\tau)$$
$$-E_h(p,\tau)\tilde{g}_0(p,\tau]. \qquad (34)$$

In order to find the nonsteady-state coefficients $D_T(t)$ and $\alpha_T(t)$, we need to set t as the upper limit in the integral over the time τ.

In the opposite case of short distances and times ($p\gg p_0=1/R_0$ and $s\tau_0\gg1$), we have the asymptotic formula

$$\tilde{g}_\mp=\tilde{g}_0(p,s)\pm p\tilde{g}_1(p,s)$$
$$\cong2\{s+D_m p^2+[(s+D_m p^2)^2+4(p^2u_0^2\mp pH_0)/3]^{1/2}\}^{-1}, \qquad (35)$$

where

$$u_0^2=\langle u^2(\mathbf{r},t)\rangle,\quad H_0=\langle\mathbf{u}(\mathbf{r},t)\cdot(\nabla\times\mathbf{u}(\mathbf{r},t))\rangle.$$

In the limit $s\to0$ and $D_m=0$, it follows from Eq. (35) that

$$\tilde{g}_0(p,0)\cong\sqrt{3}/u_0p,\quad \tilde{g}_1(p,0)\cong\sqrt{3}H_0/2p^3u_0^3.$$

If the turbulent medium does not possess helicity ($E_h(p,\tau)=0$), then $\tilde{g}_1(p,s)\equiv0$ and the system (31) turns into a single nonlinear equation for $\tilde{g}_0(p,s)$:

$$\tilde{g}_0(p,s)=\Bigg[s+D_m p^2+\frac{p^2}{4}\int_0^\infty dq\int_0^\infty d\tau\int_{-1}^{1}d\mu$$
$$\times(1-\mu^2)\exp(-s\tau)E(q,\tau)\tilde{g}_0(|\mathbf{p}-\mathbf{q}|,\tau)\Bigg]^{-1}, \qquad (36)$$

which is the same as the DIA equation for a scalar impurity.[11,16] In this case $\alpha_T^{(0)}\equiv0$, and $D_T^{(0)}$ is the same as the DIA coefficient of turbulent diffusion for a scalar impurity. This confirms our previous qualitative reasoning on the similarity between the magnetic value of D_T and the value of D_T for impurity particles.

Note that the DIA equation for a scalar impurity does not depend on the helicity, and the influence of the helicity on the values of D_T was determined by analyzing the term in the equation hierarchy containing fourth-order terms with respect to the velocities $\mathbf{u}(\mathbf{r},t)$ (assuming a Gaussian ensemble of velocities). The existence of helicity increases the diffusion coefficient D_T of a scalar impurity.[10,16] In a 100% helical medium (where $E_h(p,\tau)=pE(p,\tau)$) the coefficient D_T may increase by up to 50% (for $\xi_0\to\infty$) compared with the value in the absence of helicity. The existence of the same number of right and left helical motions in a medium increases the probability of the decay of this vortex motion compared with case where one type of helical motion exists. This qualitatively explains the amplification of turbulent diffusion by helicity.

Before giving the results of calculations using Eq. (34), we estimate D_T and α_T in the limiting cases $\xi_0=u_0\tau_0/R_0\ll1$ and $\xi_0\gg1$. Values of $\xi_0\ll1$ correspond to the quasilinear approximation.[4] In this case, a good approximation for the Green function is $\tilde{g}_0(p,\tau)\approx\tilde{g}_m(p,\tau)=\exp(-p^2D_m\tau)$ and $\tilde{g}_1(p,\tau)=0$. We further assume that $D_m=0$, and we substitute these functions into the Eqs. (34). Finally we obtain the following formulas for $\xi_0\ll1$:

$$D_T^{(0)}\cong\frac{1}{3}\int_0^\infty dp\int_0^\infty d\tau\,E(p,\tau)\cong\frac{u_0^2\tau_0}{3}\equiv\frac{u_0}{p_0}\,\frac{\xi_0}{3},$$
$$\alpha_T^{(0)}\cong-\frac{1}{3}\int_0^\infty dp\int_0^\infty d\tau\,E_h(p,\tau)\cong-\frac{H_0\tau_1}{3}$$
$$\equiv\left(-\frac{H_0}{u_0p_1}\right)\frac{\xi_0\eta}{3\gamma},\qquad \eta=p_1/p_0,\qquad \gamma=\tau_0/\tau_1. \qquad (37)$$

Here τ_0, p_0 and τ_1, p_1 are the characteristic lifetimes and wave numbers of the spectra $E(p,\tau)$ and $E_h(p,\tau)$.

For frozen turbulence where $E(p,\tau)\equiv E(p)$, $E_h(p,\tau)\equiv E_h(p)$ ($\tau_0\gg t_0=R_0/u_0=1/u_0p_0$ or $\xi_0\gg1$), we can use the asymptotic equation (35) for estimates:

$$D_T^{(0)}\cong\frac{1}{\sqrt{3}u_0}\int_0^\infty\frac{dpE(p)}{p}\cong\left(\frac{u_0}{p_0}\right)\frac{1}{\sqrt{3}},$$
$$\alpha_T^{(0)}=\frac{H_0}{2\sqrt{3}u_0^3}\int_0^\infty dp\,\frac{E(p)}{p}-\frac{1}{\sqrt{3}u_0}\int_0^\infty dp\,\frac{E_h(p)}{p}$$
$$\cong\text{const}(\eta)\left(-\frac{H_0}{u_0p_1}\right). \qquad (38)$$

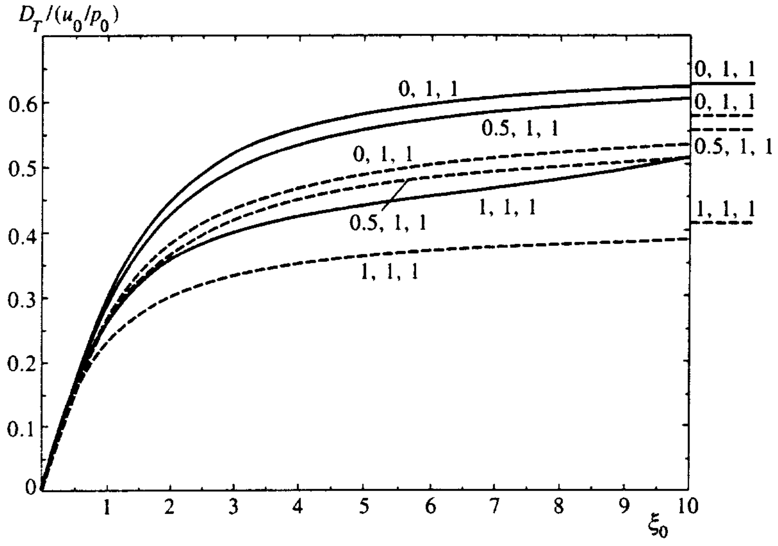

FIG. 1. Dimensionless coefficient of turbulent diffusion $\overline{D_T}=D_T(u_0/p_0)^{-1}$ for the model (41) with δ-shaped turbulent energy and helicity spectra. The solid curves give the DIA values and the dashed curves give the results of calculations using the self-consistent formula. The numbers on the curves give the values of the parameters a, γ, and η. The levels to the right give the corresponding coefficients for $\xi_0=100$. The values of $\overline{D_T}$ for the cases (0, 5, 1), (0, 1, 5), (0, 5, 5), (1, 1, 5), (1, 5. 1), and (1. 5, 5) are essentially the same as the case (0, 1, 1).

The calculations show that the function const(η) depends very weakly on η. According to the Khinchin–Bochner theorem,[3] $|E_h(p,\omega)|\leqslant pE(p,\omega)$. This implies that the dimensionless parameters a, η, and γ satisfy

$$a=\frac{H_0}{u_0^2p_0}\leqslant 1,\quad \eta=\frac{p_1}{p_0}\geqslant 1,\quad \gamma=\frac{\tau_0}{\tau_1}\geqslant 1. \tag{39}$$

Since the asymptotic forms (35) do not hold near $p=p_0$ and $p=p_1$, Eqs. (38), unlike (37), are semiquantitative. It follows from the estimates (37) and (38) that it is convenient to introduce the dimensionless variables $\overline{D_T}$ and $\overline{\alpha_T}$, defined by

$$D_T=\frac{u_0}{p_0}\,\overline{D_T}(\xi_0,a,\eta,\gamma),$$

$$\alpha_T=\left(-\frac{H_0}{u_0p_1}\right)\overline{\alpha_T}(\xi_0,a,\eta,\gamma). \tag{40}$$

We selected several turbulence models to investigate the dependence of $\overline{D_T}$ and $\overline{\alpha_T}$ on the parameters ξ_0, a, η, and γ:

$$E(p,\tau)=u_0^2\delta(p-p_0)e^{-\tau/\tau_0},$$

$$E_h(p,\tau)=H_0\delta(p-p_1)e^{-\tau/\tau_1}, \tag{41}$$

$$E(p,\tau)=\frac{u_0^2}{p_0}\frac{0.65159}{1+x^{17/3}}x^4e^{-\tau/\tau_0},$$

$$E_h(p,\tau)=au_0^2\frac{0.65159}{1+x^{17/3}}x^5e^{-\tau/\tau_1}. \tag{42}$$

Here and subsequently, $x=p/p_0$ and a is the degree of helicity. For $p\gg p_0$, the spectrum (42) has the Kolmogorov form $\propto p^{-5/3}$. The spectra (41) and (42) are examples of limiting dependences on p: a peaked dependence and slow decay for $p\rightarrow\infty$.

Calculations were also performed for the intermediate spectra:

$$E(p,\tau)=\frac{8}{3\sqrt{\pi}}\frac{u_0^2}{p_0}x^4e^{-x^2}e^{-\tau/\tau_0},$$

$$E_h(p,\tau)=au_0^2\frac{8}{3\sqrt{\pi}}x^5e^{-x^2}e^{-\tau/\tau_1}, \tag{43}$$

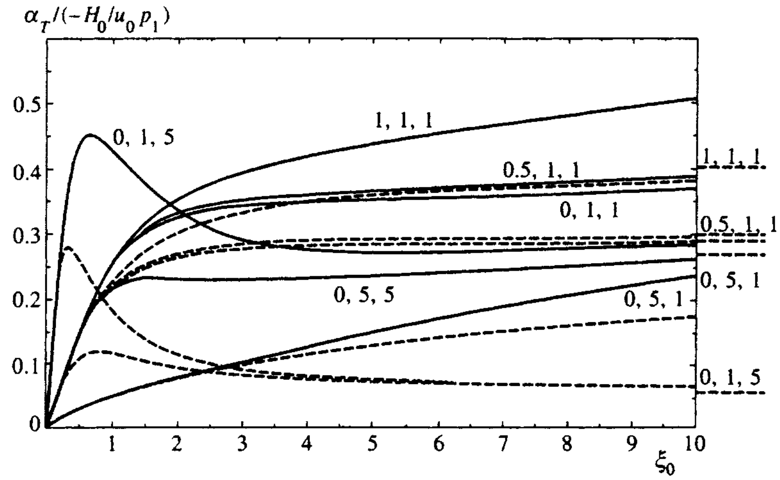

FIG. 2. Dimensionless coefficient $\overline{\alpha_T}=\alpha_T(-H_0/u_0p_1)^{-1}$ for the model (41). The notation is the same as in Fig. 1. The case (0, 1, 5) essentially coincides with (1, 1, 5), (0, 5, 1) with (1, 5, 1), and (0, 5, 5) with (1, 5, 5).

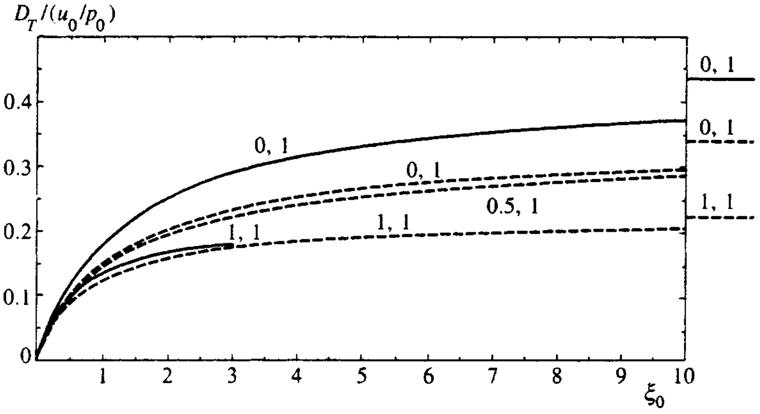

FIG. 3. Dimensionless coefficient of turbulent diffusion $\overline{D_T}$ for the model (42) with a Kolmogorov energy spectrum in the inertial range. Here $p_0=p_1$ and the numbers on the curves indicate the parameters a and γ. The solid level on the right gives $\overline{D_T}$ for $\xi_0\to\infty$ and the dashed levels correspond to $\xi_0=100$.

$$E(p,\tau)=\frac{1}{24}\,\frac{u_0^2}{p_0}x^4e^{-x}e^{-\tau/\tau_0},$$

$$E_h(p,\tau)=\frac{au_0^2}{p_0}x^5e^{-x}e^{-\tau/\tau_1}. \tag{44}$$

In addition to the calculations of the DIA values (34) of the steady-state coefficients $D_T^{(0)}(\xi_0)$ and $\alpha_T^{(1)}(\xi_0)$, we shall also give the so-called self-consistent values $D_T^{(s)}(\xi_0)$ and $\alpha_T^{(s)}(\xi_0)$. These values can be calculated from Eq. (34) if instead of the accurate DIA values for $\tilde{g}_0(p,\tau)$ and $\tilde{g}_1(p,\tau)$ we take the values of these functions in the diffusion approximation (33) with the unknown coefficients $D_T^{(s)}(\xi_0)$ and $\alpha_T^{(s)}(\xi_0)$. Equations (34) then become nonlinear equations to find the required values of $D_T^{(s)}(\xi_0)$ and $\alpha_T^{(s)}(\xi_0)$. This procedure was first proposed for a scalar field in Ref. 18.

Self-consistent values can also be found for the nonsteady-state values of $D_T^{(s)}(t)$ and $\alpha_T^{(s)}(s)$:

$$D_T^{(s)}(t)=\frac{1}{3}\int_0^\infty dp\int_0^t d\tau\,\exp\left(-p^2\int_0^\tau d\tau\,D_T^{(s)}(\tau)\right)\times\left[E(p,\tau)\cosh\left(p\int_0^\tau d\tau\,\alpha_T^{(s)}(\tau)\right)+E_h(p,\tau)\sinh\left(p\int_0^\tau d\tau\,\alpha_T^{(s)}(\tau)\right)p^{-1}\right],$$

$$\alpha_T^{(s)}(t)=\frac{1}{3}\int_0^\infty dp\int_0^t d\tau\,\exp\left(-p^2\int_0^\tau d\tau\,D_T^{(s)}(\tau)\right)\times\left[E(p,\tau)\sinh\left(p\int_0^\tau d\tau\,\alpha_T^{(s)}(\tau)\right)+E_h(p,\tau)\cosh\left(p\int_0^\tau d\tau\,\alpha_T^{(s)}(\tau)\right)\right]. \tag{45}$$

Expressions for $D_T^{(s)}(\xi_0)$ and $\alpha_T^{(s)}(\xi_0)$ corresponding to Eq. (33) are obtained from the expressions (45) if we set $t\to\infty$ and assume $D_T^{(s)}(t;\xi_0)\equiv D_T^{(s)}(\xi_0)$ and $\alpha_T^{(s)}(t;\xi_0)\equiv\alpha_T^{(s)}(\xi_0)$.

A more logical but more complex method is to calculate the nonsteady-state values (45) until those times for which constant values are obtained for $D_T^{(s)}(t)$ and $\alpha_T^{(s)}(t)$. We mainly use the first, simpler method. Using both methods to calculate $D_T^{(s)}(\xi_0)$ for a scalar impurity shows[17] that the first method gives values of $D_T(\xi_0)$ that are too low, and the second method yields overestimates; their mean very satisfactorily approximates the true value of the coefficient $D_T(\xi_0)$. In the presence of helicity, the situation is more complicated, but is qualitatively the same as before.

The calculations show that the self-consistent values of $\overline{D^s_T}(t)$ and $\overline{\alpha^s_T}(t)$ are always positive. For the limiting case of 100% helicity $|E_h(p,\tau)|=pE(p,\tau)$, where $E(p,\tau)=u_0^2\delta(p-p_0)f(\tau)$, Eqs. (45) give $\overline{D^s_T}(t)=\overline{\alpha^s_T}(t)>0$. Since

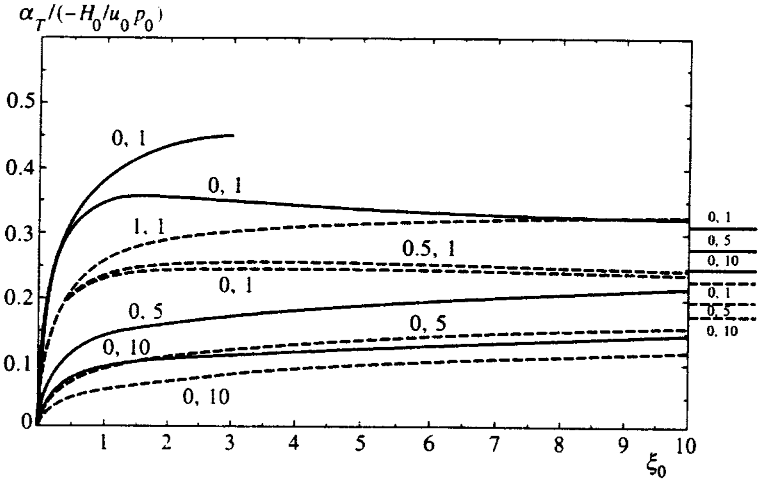

FIG. 4. Dimensionless coefficient $\overline{\alpha_T}=\alpha_T(-H_0/u_0p_0)^{-1}$ for the model (42). The notation is the same as in Fig. 3. All the levels on the right correspond to $\xi_0=50$. The case (0, 5) essentially coincides with (1, 5), and (0, 10) with (1, 10).

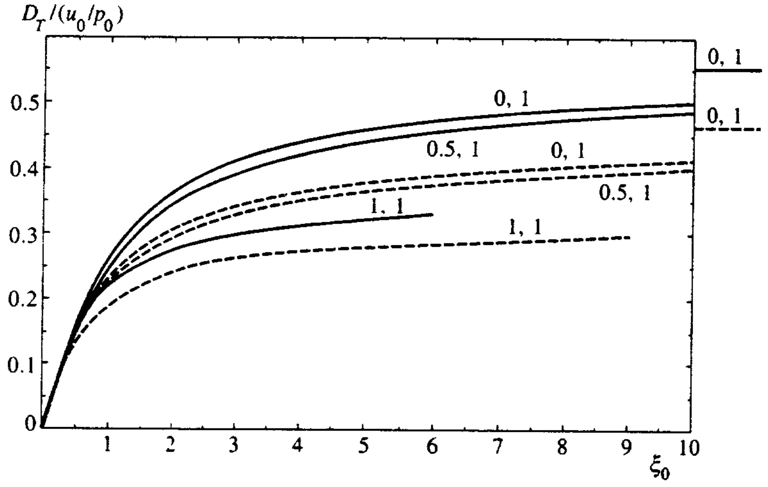

FIG. 5. Dimensionless coefficient of turbulent diffusion $\overline{D_T}$ for the model (43). The notation is the same as Fig. 3. All the levels to the right correspond to $\xi_0=100$. Case (0, 5) is essentially the same as (1, 5).

the kinetic coefficients D_T and α_T are mainly determined by large-scale turbulent motion, which is accurately described by the Green functions in the diffusion approximation, the self-consistent method gives comparatively satisfactory values of D_T and α_T which agree qualitatively with the DIA values.

The results of calculations of the steady-state DIA values $\overline{D_T}(\xi_0)$ and $\overline{\alpha_T}(\xi_0)$ are plotted in Figs. 1–8. We note above all that the presence of helicity ($a\neq 0$) reduces the diffusion coefficient D_T rather than increasing it, as in the case of scalar impurity diffusion. This effect was first noted in Ref. 10. This decrease is due to the α effect, i.e., the amplification of the magnetic field by helical motion, which slows the diffusion attenuation of the field.

A second interesting feature is the nonlinear dependence of the coefficient $\overline{D_T}$ on the degree of helicity $a=H_0/u_0^2p_0$. For $0<a<0.5$ this dependence is very weak. In practice, limiting values of $\overline{D_T}$ corresponding to $a=0$ may be used in this range of values of the parameter a. Real turbulent motion rarely has a degree of helicity $a>0.5$, so that this characteristic can reduce the number of independent parameters.

If the helical motion is small-scale ($\eta=p_1/p_0\gg 1$) or short-range correlated ($\gamma=\tau_0/\tau_1\gg 1$), the turbulent diffusion coefficient D_T is almost independent of the degree of helicity a and is the same as the DIA diffusion coefficient for a scalar impurity. In this case, the difference between the coefficients D_T for the magnetic field and the scalar impurity is determined by the different contribution to these coefficients made by the correlation functions of the fourth and higher orders.

For a scalar impurity, the maximum and invariably negative contribution of this correction (in the absence of helicity) is $<10\%$ for $\xi_0\rightarrow\infty$. For $\xi_0\rightarrow 0$ this correction tends to zero. It is very difficult to calculate the correction to D_T for magnetic field diffusion because of the very cumbersome expression for this correction. This correction may also have a low value; in any event, for $\xi_0\rightarrow 0$ it also tends to zero like the correction to D_T for a scalar impurity. In practice, for $p_1/p_0>5$ or $\tau_0/\tau_1>5$, the influence of helicity on the magnetic-field diffusion coefficient D_T can be neglected. Physically, this is quite understandable if we recall that these cases correspond to small-scale or short-range time-correlated helical motion against a background of nonhelical large-scale motion with a long correlation time.

A comparison between the DIA values of D_T and the self-consistent values of $D_T^{(s)}$ shows that these agree well for $\xi_0\leqslant 1$, and agree qualitatively for $\xi_0\gg 1$. Note that $D_T^{(s)}<D_T$ always, because for small scales and times ($p\gg p_0$, $s\tau_0\gg 1$),

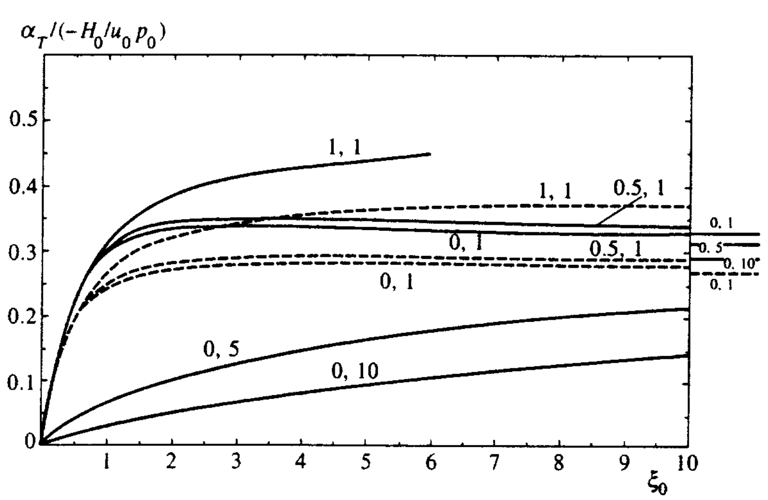

FIG. 6. Dimensionless coefficient $\overline{\alpha_T}=\alpha_T(-H_0/u_0p_0)^{-1}$ for the model (43). The notation is the same as Fig. 3. Case (0, 5) essentially coincides with (1, 5) and case (0, 10) with (1, 10).

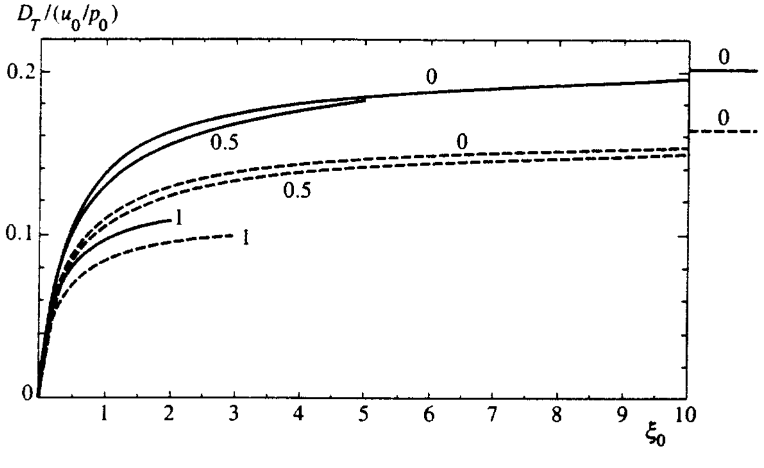

FIG. 7. Dimensionless coefficient of turbulent diffusion $\overline{D_T}$ for the model (44). The notation is the same as Fig. 3.

the real Green function $\tilde{g}_0(p,s)$ has a higher value than its approximate diffusion expression.

The dependence of the dimensionless $\overline{\alpha_T}$ coefficient on the degree of helicity a is qualitatively the same as that for $\overline{D_T}$: for $a<0.5$, the difference between the coefficients $\overline{\alpha_T}$ is very small even for $p_1 \cong p_0$ and $\tau_1 \cong \tau_0$, and for $p_1/p_0 > 5$ or $\tau_0/\tau_1 > 5$, the dependence on the parameter a essentially disappears. This implies that in these cases, the dimensional α_T coefficient depends linearly on the characteristic helicity H_0 (because of the dimensional factor in Eq. (40)).

An increase in the parameter $\gamma = \tau_0/\tau_1$ leads to a decrease in α_T for any ξ_0, and corresponds physically to a reduction in the α effect for the shorter-lived helical motion. However, the degree of reduction depends on the value of the parameter ξ_0, and for $\xi_0 \to \infty$, α_T ceases to depend on γ, which is consistent with the estimate (38).

The dependence of $\overline{\alpha_T}$ on the parameter $\eta = p_1/p_0$ was investigated using the model (41), and is plotted in Fig. 2. Attention is drawn to the peaked behavior of $\overline{\alpha_T}$ for $\xi_0<1$ for the case $\eta=5$ and $\gamma=1$. The maximum of the peak corresponds to $\xi_0=\xi_*$, which may be estimated qualitatively from the condition that over its lifetime τ_1, the helix has time to decay, i.e., $D_T\tau_1 \approx R_1^2$ for $D_T \cong u_0^2\tau_0/3$, which is satisfied for $\xi_0<1$ (see the estimate (37)). This estimate gives

$$\xi_* \cong \text{const}\sqrt{\tau_0\tau_1}(p_0/p_1) \equiv \text{const}\sqrt{\gamma}/\eta$$

provided that $\xi_*<1$. Substituting this value of ξ_* into the estimate (37), we find that $\overline{\alpha_T}_{\max} \cong (1/3)\ \text{const}/\sqrt{\gamma}$, and does not depend on the parameter η. The condition $\xi_*<1$ (or $\sqrt{\gamma}/\eta<1$) is satisfied for comparatively long-lived, small-scale helical motion ($\eta \gg 1$), i.e., when these helical formations may be considered to be a particular type of impurity particles.

It would seem that a similar estimate of ξ_* might be obtained for the other limiting case $\xi_0 \gg 1$, where $D_T \cong \text{const}(u_0/p_0)$. This would yield $\xi_* \cong \text{const}\ \gamma/\eta^2$, subject to the condition $\xi_* \gg 1$. However the calculated values of $\overline{\alpha_T}(\xi_0)$, both in the DIA approximation and the self-consistent values, do not reveal any peaks for $\xi_0 \gg 1$. In reality, such an estimate is meaningless for the region $\xi_0 \gg 1$, since the nature of the turbulent motion with $\xi_0 \gg 1$ is not consistent with the helical formations being considered as some sort of impurity exposed to diffusion damage within its lifetime τ_1.

For $\xi_0 \gg 1$ the turbulent motion is of a vortex nature with the rotation time $t_0 \cong 1/u_0p_0$ much shorter than the vortex lifetime τ_0: $\xi_0 \equiv \tau_0/t_0 \gg 1$. In this case, the helical motion is an inseparable part of these vortices. For $\tau_1 \ll t_0$ the helical

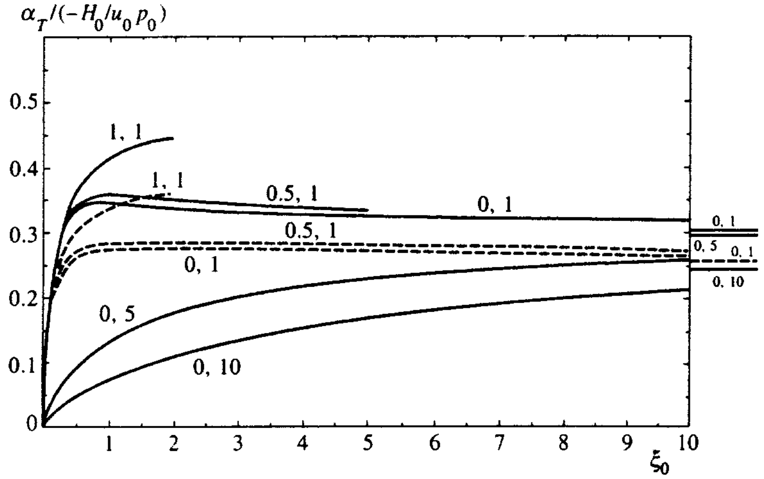

FIG. 8. Dimensionless coefficient $\overline{\alpha_T} = \alpha_T(-H_0/u_0p_0)^{-1}$ for the model (44). The notation is the same as in Fig. 3.

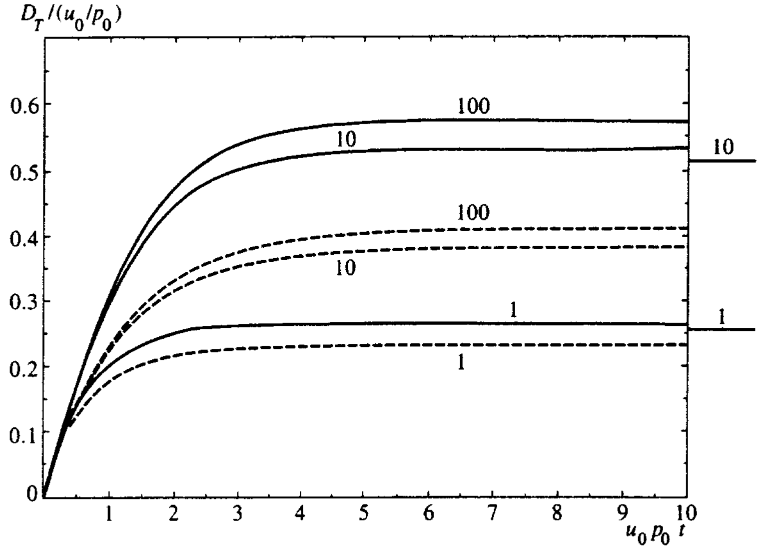

FIG. 9. Nonsteady-state dimensionless coefficients $\overline{D_T}=\overline{\alpha_T}$ for the model (41) with $a=1$, $\gamma=1$, and $\eta=1$. The solid curves give the calculations using the self-consistent formula (45) and the dashed curves give those using a simplified variant of this formula where constant values of the coefficients $D_T^{(s)}$ and $\alpha_T^{(s)}$ are set in the integrands. The numbers on the curves give the values of the parameter ξ_0. The levels on the right give the DIA values of $\overline{D_T}$ for $\xi_0=10$.

formations undergo repeated reversal in the vortex. In this case, the average α effect ceases to depend on the parameter $\gamma=\tau_0/\tau_1$. The dependence of α_T on $\eta=p_1/p_0$ is mainly determined by the dimensional factor in formula (40): $\alpha_T \propto 1/\eta$.

It can be seen from the figures that the self-consistent values of $\overline{\alpha_T^{(s)}}$ are numerically similar to the DIA values of $\overline{\alpha_T}$ for $\xi_0 \ll 1$, and are much lower for $\xi_0 \gg 1$. As the parameter $\eta=p_1/p_0 \equiv R_0/R_1$ increases, this difference becomes even greater. Note that the self-consistent method gives substantially better results when used to calculate α_T, compared with the calculations of the turbulent diffusion coefficient D_T.

Calculations of the correction $\alpha_T^{(1)}$, which depends on four-point correlation functions of the velocity (for a Gaussian ensemble and the spectrum (41)), showed that these corrections increase from 0% for $\xi_0 \to 0$ to $\leqslant 4\%$ for $p_0=p_1$ and $\tau_0=\tau_1$ for $\xi_0=10$. For small-scale or short-range correlated helicity, we have 20% instead of 4% for $\gamma=5$, 25% for $\eta=5$, and 30% for $\gamma=5$, $\eta=5$. The correction is almost independent of the degree of helicity a, and is negative for $\xi_0 \leqslant 1$ and positive for $\xi_0 \gg 1$.

4. CALCULATION OF NONSTEADY-STATE $D_T(T)$ AND $\alpha_T(T)$ BY SELF-CONSISTENT METHODS

Figure 9 gives results of calculations of $D_T(t)$ and $\alpha_T(t)$ for the model (41) with 100% helicity ($|E_h(p,\tau)|=pE(p,\tau)$). In this case, we find $\gamma=1$, $\eta=1$, $a=1$, and $\overline{D_T^{(s)}}(t)=\overline{\alpha_T^{(s)}}(t)>0$. The solid curves show the results of the calculations using Eq. (45), and the dashed curves give those using the simplified variant of this equation, where the steady-state values $D_T^{(s)}(\infty)$ and $\alpha_T^{(s)}(\infty)$, also calculated using the simplified variant of Eq. (45), were taken as $D_T^{(s)}(\tau)$ and $\alpha_T^{(s)}(\tau)$ in the integral terms. The dashed curves give low values and the solid curves give slightly high values. This case was calculated in Ref. 10 using the Lagrange equations (9) and (10) as a result of a numerical simulation of trajectories. In this study, negative values of the diffusion coefficient $D_T(t)$ were obtained for $t>3t_0$, and positive values were obtained for α_T. Our calculations are in good agreement with the steady-state DIA values for D_T and α_T, and reliably indicate that the coefficient of turbulent diffusion of the magnetic field for an incompressible medium is positive.

5. CONCLUSIONS

We present the main conclusions obtained from an analysis of these results.

1. The dimensionless coefficient of turbulent diffusion $\overline{D_T}=D_T(u_0/p_0)^{-1}$ for a magnetic field in an incompressible medium is positive and increases monotonically with increasing parameter $\xi_0=u_0p_0\tau_0$, asymptotically approaching its limiting value at $\xi_0 \to \infty$. The presence of helicity in the medium reduces D_T, and for a low degree of helicity $a=H_0/u_0^2p_0<0.5$ this effect is very weak. Small-scale or short-range correlated helicity with $\eta=p_1/p_0 \geqslant 5$ or $\gamma=\tau_0/\tau_1 \geqslant 5$ negligibly alters D_T, even for $a=1$.

2. The dimensionless coefficient describing the α effect, $\overline{\alpha_T}=\alpha_T(-H_0/u_0p_1)^{-1}$, increases monotonically with increasing parameter ξ_0 for large-scale ($\eta \cong 1$) helicity. For $a<0.5$, the value of $\overline{\alpha_T}$ is essentially independent of the degree of helicity.

3. For small-scale and short-range correlated helicity ($\eta=p_1/p_0 \geqslant 5$ or $\gamma=\tau_0/\tau_1 \geqslant 5$), the coefficient $\overline{\alpha_T}$ is essentially independent of a.

4. For small-scale helicity with $\sqrt{\gamma}/\eta<1$, $\overline{\alpha_T}(\xi_0)$ has a peaked profile with a maximum at $\xi_* \cong \text{const}\sqrt{\tau_0/\tau_1}(p_0/p_1) \equiv \text{const}\sqrt{\gamma}/\eta$ and $\overline{\alpha_T}_{\max} \cong (1/3) \times \text{const}/\sqrt{\gamma}$.

5. For $\xi_0 \to \infty$, $\overline{\alpha_T}(\xi_0)$ tends to a limiting value which depends weakly on η and does not depend on γ.

[1] P. H. Roberts, *An Introduction to Magnetohydrodynamics*, Elsevier, New York (1967).

[2] H. K. Moffatt, *Magnetic Field Generation in Electrically Conducting Fluids*, Cambridge University Press, Cambridge (1978).

[3] F. Krause and K.-H. Raedler, *Mean-Field Magnetohydrodynamics and Dynamo Theory*, Academie Verlag, Berlin (1980).

[4] E. N. Parker, *Cosmical Magnetic Fields*, Clarendon Press, Oxford (1979).
[5] A. P. Kazantsev, Zh. Éksp. Teor. Fiz. **53**, 1806 (1967) [Sov. Phys. JETP **26**, 1031 (1968)].
[6] A. P. Kazantsev, A. A. Ruzmaĭkin, and D. D. Sokolov, Zh. Éksp. Teor. Fiz. **88**, 487 (1985) [Sov. Phys. JETP **61**, 285 (1985)].
[7] H. K. Moffatt, J. Fluid Mech. **65**, 1 (1974).
[8] R. H. Kraichnan, J. Fluid Mech. **75**, 657 (1976).
[9] G. I. Taylor, Proc. London Math. Soc. A **20**, 196 (1921).
[10] R. H. Kraichnan, J. Fluid Mech. **77**, 753 (1976).
[11] N. A. Silant'ev, Zh. Éksp. Teor. Fiz. **101**, 1216 (1992) [Sov. Phys. JETP **74**, 650 (1992)].
[12] A. Z. Dolginov and N. A. Silant'ev, Geophys. Astrophys. Fluid Dyn. **63**, 139 (1992).
[13] P. H. Roberts, Fluid Mech. Res. **11**, 257 (1961).
[14] R. H. Kraichnan, Phys. Fluids **13**, 22 (1970).
[15] R. H. Kraichnan, J. Fluid Mech. **81**, 385 (1977).
[16] N. A. Silant'ev, Geophys. Astrophys. Fluid Dyn. **75**, 183 (1994).
[17] N. A. Silant'ev, Zh. Éksp. Teor. Fiz. **111**, 871 (1997) [JETP **84**, 479 (1997)].
[18] R. Phythian and W. D. Curtis, J. Fluid Mech. **89**, 241 (1978).

Translated by R. M. Durham

Integro-differential equation method in the hydrodynamics of an incompressible viscous fluid

A. B. Petrin*)

Institute of High Temperatures, Russian Academy of Sciences, 127412 Moscow, Russia

(Submitted 18 February 1997)
Zh. Éksp. Teor. Fiz. **112**, 1332–1339 (October 1997)

An integro-differential equation method is proposed to describe the motion of an incompressible viscous fluid. The method uses an analogy between the hydrodynamic equations for an incompressible viscous fluid and the magnetostatic equations. An analysis is made of the flow of an incompressible viscous fluid round an object as a specific application. The solution automatically satisfies the boundary conditions at the surface of the object and at infinity. © *1997 American Institute of Physics.* [S1063-7761(97)01310-3]

1. INTRODUCTION

A well-known description of the motion of an incompressible viscous fluid is provided by a system of differential equations for the velocity field and the vorticity field.[1–3] In Ref. 3 an analogy was noted between some of the differential hydrodynamic equations for an incompressible viscous fluid and the differential magnetostatic equations but this analogy was not developed.

Here this analogy is developed, refined, and used to solve problems in hydrodynamics.

2. FORMULATION OF THE PROBLEM

The motion of a fluid is described by the Navier–Stokes equation, which may be expressed in the form[3]

$$\frac{\partial \mathbf{v}}{\partial t}+(\mathbf{v}\cdot\nabla)\mathbf{v}=-\frac{\nabla p}{\rho}-\nabla\varphi+\frac{\mathbf{f}_v}{\rho}, \tag{1}$$

where $\mathbf{v}$ is the velocity field, p is the scalar pressure field, ρ is the fluid density, φ is the potential of the external conservative forces per unit mass, and $\mathbf{f}_v$ are the forces of viscosity acting per unit volume of fluid.

If the fluid is incompressible, its density is constant and the velocity field should satisfy the volume conservation equation:

$$\nabla\cdot\mathbf{v}=0.$$

The expression for the viscosity force is then given by $\mathbf{f}_v=\eta\nabla^2\mathbf{v}$, where η is the first coefficient of viscosity.

Thus, the Navier–Stokes equation for an incompressible fluid can be simplified, and may be written as[2]

$$\frac{\partial \mathbf{v}}{\partial t}+(\mathbf{v}\cdot\nabla)\mathbf{v}=-\frac{\nabla p}{\rho}-\nabla\varphi+\frac{\eta}{\rho}\nabla^2\mathbf{v}. \tag{2}$$

Introducing the vector vorticity field $\mathbf{\Omega}=\nabla\times\mathbf{v}$ and recalling the vector identity

$$(\mathbf{v}\cdot\nabla)\mathbf{v}=(\nabla\times\mathbf{v})\times\mathbf{v}+\frac{1}{2}\nabla(v^2),$$

we rewrite the previous equation in the form

$$\frac{\partial \mathbf{v}}{\partial t}+\mathbf{\Omega}\times\mathbf{v}=-\frac{1}{2}\nabla(v^2)-\frac{\nabla p}{\rho}-\nabla\varphi+\frac{\eta}{\rho}\nabla^2\mathbf{v}.$$

Taking the curl of this equation, we obtain

$$\frac{\partial \mathbf{\Omega}}{\partial t}+\nabla\times(\mathbf{\Omega}\times\mathbf{v})=\frac{\eta}{\rho}\nabla^2\mathbf{\Omega}. \tag{3}$$

Let D be the characteristic dimension of the problem and V be the characteristic velocity. We transform to dimensionless variables, making the substitution

$$x=x'D,\quad y=y'D,\quad z=z'D,\quad t=t'(D/V). \tag{4}$$

We then obtain an equation for the dimensionless fields:

$$\frac{\partial \mathbf{\Omega}'}{\partial t'}+\nabla'\times(\mathbf{\Omega}'\times\mathbf{v}')=\frac{1}{R}\nabla'^2\mathbf{\Omega}',$$

where $R=\rho VD/\eta$ is the Reynolds number for the given problem.

To be specific, we analyze the problem of exterior flow of an incompressible fluid around an object. We also analyze the problem in terms of dimensionless variables. Then, omitting primes from hereon, we obtain a system of three equations:

$$\nabla\cdot\mathbf{v}=0, \tag{5}$$

$$\nabla\times\mathbf{v}=\mathbf{\Omega}, \tag{6}$$

$$\frac{\partial \mathbf{\Omega}}{\partial t}+\nabla\times(\mathbf{\Omega}\times\mathbf{v})=\frac{1}{R}\nabla^2\mathbf{\Omega} \tag{7}$$

with the boundary conditions

$$\mathbf{v}=\mathbf{0}\quad\text{on the surface of the object}, \tag{8}$$

$$\mathbf{v}=\mathbf{e}\quad\text{far from the object} \tag{9}$$

($\mathbf{e}$ is the unit vector determining the direction of the velocity far from the object).

We assume that the vorticity field $\mathbf{\Omega}$ is known over all space; then Eqs. (5) and (6) can be used to find the velocity over all space. In fact, Eqs. (5) and (6) are similar to the magnetostatic equations that determine the magnetic field $\mathbf{B}$ in terms of the current density $\mathbf{j}$:

$$\nabla\cdot\mathbf{B}=0.$$

$$\nabla\cdot\mathbf{B}=\mu_0\mathbf{j}.$$

1063-7761/97/100724-04$10.00

The solution of the magnetostatic equations for given currents is given by the Biot–Savart law,

$$\mathbf{B}(1)=\frac{\mu_0}{4\pi}\int_V \frac{\mathbf{j}(2)\times\mathbf{r}_{12}}{r_{12}^3}dV_2,$$

where the integration extends over all points in the volume V at which currents exist. The symbol 1 refers to the coordinates of the point at which the field $\mathbf{B}$ is determined, the symbol 2 indicates the coordinates of points over which integration is performed, and $\mathbf{r}_{12}$ is the position vector from the point of integration 2 to point 1; its length is

$$r_{12}=\sqrt{(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2}.$$

If we replace $\mathbf{B}$ by $\mathbf{v}$ and $\mu_0\mathbf{j}$ by $\mathbf{\Omega}$ in the magnetostatic equations, we obtain Eqs. (5) and (6). Thus, the solution of these equations for a given vorticity field $\mathbf{\Omega}$ is

$$\mathbf{v}(1,t)=\frac{1}{4\pi}\int_V \frac{\mathbf{\Omega}(2,t)\times\mathbf{r}_{12}}{r_{12}^3}dV_2.$$

Since the vorticity field only differs appreciably from zero near the object, this solution tends to zero with increasing distance from the object.

Note that Eqs. (5) and (6) are a linear system of equations, so that its solution is also the velocity field

$$\mathbf{v}(1,t)=\mathbf{e}+\frac{1}{4\pi}\int_V \frac{\mathbf{\Omega}(2,t)\times\mathbf{r}_{12}}{r_{12}^3}dV_2. \tag{10}$$

The velocity field calculated according to Eq. (10) not only satisfies Eqs. (5) and (6) but also the boundary condition (9) far from the object.

Thus, if the distribution of the vorticity field is known at a certain time, this field can be found at the next instant via the integrodifferential equation

$$\frac{\partial\mathbf{\Omega}(1,t)}{\partial t}+\nabla\times\left[\mathbf{\Omega}(1,t)\times\left(\mathbf{e}+\frac{1}{4\pi}\int_V\frac{\mathbf{\Omega}(2,t)\times\mathbf{r}_{12}}{r_{12}^3}dV_2\right)\right]$$
$$=\frac{1}{R}\nabla^2\mathbf{\Omega}(1,t), \tag{11}$$

where the nabla operator acts on the coordinates denoted by the index 1.

However, the problem is that integration in Eq. (11) is performed over all space, and if the vorticity field inside the object is assumed to be zero, the vorticity field distribution obtained from Eq. (11) will automatically fail to satisfy the boundary condition at the surface of the object around which the fluid flows.

3. SOLUTION OF THE PROBLEM

To satisfy this boundary condition, either the problem must be formulated over all space, or the boundary conditions must be determined for the vorticity field.

To automatically allow for the boundary conditions at the surface of the object, we assume that the object consists of the surrounding incompressible fluid which is exposed to the action of an external, fairly large, effective bulk frictional force $\mathbf{f}_{fr}=-\gamma(\mathbf{r})\mathbf{v}$ inside the object around which the fluid flows. The coefficient $\gamma(\mathbf{r})$ takes into account the spatial distribution of this force, and can be represented in the form $\gamma(\mathbf{r})=kw(\mathbf{r})$, where k is modulus of the force and the function $w(\mathbf{r})$ is assumed to be zero outside the solid, unity inside the solid, and continuously integrable in the vicinity of its boundary. If the coefficient k is large enough, the velocity of the fluid inside the solid will be negligible. It will then be assumed that in the limit where k tends to infinity, the fluid velocity near the surface and inside the object will be zero, and the boundary conditions at the surface of the object will automatically be satisfied. Physically it is clear that the incoming stream of fluid will flow around this retarded, almost stationary fluid as if it were a solid.

Using the frictional force introduced above, Eq. (1) can be written over all space in the form:

$$\frac{\partial\mathbf{v}}{\partial t}+(\mathbf{v}\cdot\nabla)\mathbf{v}=-\frac{\nabla p}{\rho}-\nabla\varphi+\frac{\mathbf{f}_v}{\rho}-\frac{k}{\rho}w(\mathbf{r})\mathbf{v}. \tag{12}$$

As before, we take the curl of Eq. (12) and introduce the vorticity field; we then obtain the differential equation

$$\frac{\partial\mathbf{\Omega}}{\partial t}+\nabla\times(\mathbf{\Omega}\times\mathbf{v})=\frac{\eta}{\rho}\nabla^2\mathbf{\Omega}=\frac{k}{\rho}\nabla\times(w(\mathbf{r})\mathbf{v}). \tag{13}$$

As in Sec. 2, we transform to dimensionless primed variables in accordance with Eq. (4). We then obtain an equation similar to (7):

$$\frac{\partial\mathbf{\Omega}'}{\partial t'}+\nabla'\times(\mathbf{\Omega}'\times\mathbf{v}')=\frac{1}{R}\nabla'^2\mathbf{\Omega}'=K\nabla'\times(w(\mathbf{r}')\mathbf{v}'), \tag{14}$$

where $K=kD/\rho V$, and $R=\rho VD/\eta$ is, as before, the Reynolds number for the given problem.

Omitting the primes, we finally obtain

$$\nabla\cdot\mathbf{v}=0, \tag{15}$$

$$\nabla\times\mathbf{v}=\mathbf{\Omega}, \tag{16}$$

$$\frac{\partial\mathbf{\Omega}}{\partial t}+\nabla\times(\mathbf{\Omega}\times\mathbf{v})=\frac{1}{R}\nabla^2\mathbf{\Omega}-K(\nabla\times(w\mathbf{v})+w\mathbf{\Omega}). \tag{17}$$

The first term in parentheses in Eq. (17) is nonzero near the boundary of the object around which fluid flows, since the vector field ∇w is nonzero near this boundary and is directed into the object. The integral of ∇w along any curve beginning outside the object and ending inside the object will be unity. Thus, this term is determined by the tangential component of the velocity at the boundary.

The second term in parentheses in Eq. (17) is only nonzero inside the object, and is zero outside.

The solution of Eqs. (15) and (16) that satisfies the boundary condition at infinity is given by Eq. (10), as in Sec. 2. Thus, the integrodifferential equation for the problem has the form

$$\frac{\partial\mathbf{\Omega}(1,t)}{\partial t}+\nabla\times\left[\mathbf{\Omega}(1,t)\times\left(\mathbf{e}+\frac{1}{4\pi}\int_V\frac{\mathbf{\Omega}(2,t)\times\mathbf{r}_{12}}{r_{12}^3}dV_2\right)\right]$$
$$=\frac{1}{R}\nabla^2\mathbf{\Omega}(1,t)-K\left[\nabla w\times\left(\mathbf{e}+\frac{1}{4\pi}\int_V\frac{\mathbf{\Omega}(2,t)\times\mathbf{r}_{12}}{r_{12}^3}dV_2\right)\right.$$

$$\left. + w\mathbf{\Omega}(1,t) \right]. \tag{18}$$

The solution of Eq. (18) with allowance for Eq. (10), when the coefficient K tends to infinity and the size of the region of variation of the function ∇w near the boundary tends to zero, will tend to the solution of Eqs. (5)–(7) with the boundary conditions (8) and (9).

4. TWO-DIMENSIONAL PROBLEM

We consider two-dimensional flow around a cylinder of arbitrary cross section. We direct the z axis of a rectangular Cartesian coordinate system along the generatrix of the cylinder; the velocity field will then only have components in the x and y directions, and the vorticity field will only have a z component which will be a function of x and y. Then for this component of the vorticity field, Eq. (18) has the form

$$\frac{\partial \Omega_z}{\partial t} + v_x \frac{\partial \Omega_z}{\partial x} + v_y \frac{\partial \Omega_z}{\partial y} = \frac{1}{R} \nabla^2 \Omega_z - K[(\nabla w \times \mathbf{v})_z + w\Omega_z], \tag{19}$$

where the nabla operator only acts on the x and y coordinates (two-dimensional operator), and the velocity field is

$$\mathbf{v}(x_1, y_1, t) = \mathbf{e} + \frac{1}{2\pi} \int \frac{[\Omega_z(x_2, y_2, t)\mathbf{r}_{12}]}{r_{12}^2} dx_2 dy_2, \tag{20}$$

which can be obtained from Eq. (10) after integrating over the z coordinate between plus and minus infinity, where the position vector in formula (20) is two-dimensional (only has x and y components) and integration is performed over the entire xy plane.

Equation (19), with Eq. (20) inserted, yields the solution of the two-dimensional problem in the limit, i.e., the distribution of the vorticity field and the velocity field at arbitrary time.

5. TWO-DIMENSIONAL FLOW AROUND A CIRCULAR CYLINDER

Flow around a circular cylinder is taken as a specific example of a two-dimensional problem. Assuming that the coordinate origin is located on the symmetry axis of the cylinder, we take the function $w(\mathbf{r})$ in the form

$$w(\mathbf{r}) = \frac{1}{\exp\left(\frac{1}{\delta}\left(r - \frac{D}{2}\right)\right) + 1}, \tag{21}$$

where D is the cylinder diameter, $r = \sqrt{x^2 + y^2}$, and δ is approximately equal to the distance from the surface of the cylinder, on which the function is essentially unity inside the cylinder and zero outside the cylinder (similar to the Fermi–Dirac distribution, whose properties are well-known). Cast in dimensionless coordinates, the function (21) has the form

$$w(x,y) = \frac{1}{\exp\left(\frac{1}{\sigma}\left(\sqrt{x^2 + y^2} - \frac{1}{2}\right)\right) + 1}, \tag{22}$$

where $\sigma = \delta / D$ is the fraction of the diameter over which the function varies rapidly.

The gradient of the function $w(\mathbf{r})$ is then given by

$$\nabla w(\mathbf{r}) = -\frac{1}{\sigma} \frac{\exp\left(\frac{1}{\sigma}\left(r - \frac{1}{2}\right)\right)}{\left(\exp\left(\frac{1}{\sigma}\left(r - \frac{1}{2}\right)\right) + 1\right)^2} \frac{\mathbf{r}}{r}. \tag{23}$$

Substituting expressions (22) and (23) into the system (19) and (20), we obtain an integrodifferential equation for the approximate problem of two-dimensional flow around a cylinder. Solving this equation and making the coefficient K tend to infinity and the coefficient σ tend to zero, we obtain a rigorous solution of this problem.

Knowing the distribution of the velocity field from the solution of Eqs. (19) and (20), we can determine the dimensionless drag and lift coefficients of a circular cylinder. In fact, the force acting on the cylinder from the fluid is equal to the modulus of the resultant effective frictional force introduced earlier. If the fluid velocity at infinity is directed along the x axis, the dimensionless drag coefficient C_d and the dimensionless lift coefficient C_a can be expressed in terms of the dimensionless velocity:

$$C_d = \frac{F_d}{\rho V^2 DL/2} = 2K \int w(x,y) v_x(x,y) dx\, dy, \tag{24}$$

$$C_a = \frac{F_a}{\rho V^2 DL/2} = 2K \int w(x,y) v_y(x,y) dx\, dy, \tag{25}$$

where F_d is the drag and F_a is the lift acting on a section of the cylinder of length L. When the coefficient K tends to infinity and the coefficient σ tends to zero, we obtain coefficients corresponding to a rigorous solution of the problem.

6. NUMERICAL RESULTS

A numerical simulation was made of two-dimensional flow around a circular cylinder (19), (20). The function $w(\mathbf{r})$ and its gradient were taken in the form (22) and (23), respectively. The Reynolds number was taken to be $R = 400$.

It was found that the parameters σ and K of the function $w(\mathbf{r})$, which determine the magnitude and spatial distribution of the effective drag, generally influence the numerical solution. This influence has the result that the effective drag acts not only inside the cylinder but outside as well. In the calculations this leads to a slight overestimate of the dimensionless drag coefficient C_d compared with its experimental value, since the layers of fluid outside the volume are additionally retarded. This additional retardation is observed up to distances of several σ from the cylindrical surface, so that as σ decreases, the coefficient C_d should approach its accurate value.

The calculations showed that this is in fact the case. The calculations were performed on a uniform grid with $\sigma = 0.005$, $K = 40$, and $R = 400$. The drag coefficient obtained in the calculations was approximately $C_d = 1.8$. The value of C_d varied slightly with time (less than by 0.1) as a result of the periodic detachment of vortices from the cylinder and the formation of a Kármán vortex street beyond the cylinder.

When σ was halved for the same parameters K and R, the coefficient C_d was reduced to ≈ 1.4. However, the value of C_d obtained from experimental investigations of flow around a cylinder is approximately 1.3 (Ref. 1, p. 30).

The results of the numerical calculations on a uniform mesh show that in order to obtain a solution with a lower effective retardation of the fluid outside the cylinder, a nonuniform grid concentrated near the boundary must be used for the calculations and σ reduced so that artificial retardation of the fluid flow outside the cylinder is essentially eliminated.

It should be noted that a small perturbation of the incoming flow must be introduced in the numerical simulation to form a Kármán vortex street beyond the cylinder. This small perturbation, developing into a pair of vortices beyond the cylinder and formed immediately after the onset of the fluid motion, leads to axisymmetric systematic separation and in a system of alternately separating vortices—a Kármán vortex street. Such a small initial perturbation is always present in a real incoming flow of fluid at a cylinder and is responsible for the instability of a real flow. The time evolution of the vortices beyond the cylinder obtained numerically passes through the same stages as the experimental evolution (Ref. 1, p. 395). In particular, the numerical model correctly predicts the flow separation location.

7. CONCLUSIONS

The integrodifferential equation method can reduce the problem of the hydrodynamics of an incompressible viscous fluid to an integrodifferential equation for a single vector field—the vorticity field. This field only differs substantially from zero near the object around which the fluid flows and in the vortex wake, so that the vortex field need only be approximated in this bounded spatial region to obtain a numerical solution of these equations. The velocity field can be determined using the known vorticity field over all space by using the integral expressions derived.

An important advantage of the proposed method is that the region of approximation of the function describing the flow of fluid around the object is reduced in size and bounded.

The integrodifferential equation method has been applied here to study external two-dimensional flow around a circular cylinder. However, this method can be successfully applied to problems of internal flow, and can also be used to express the equations of motion in integrodifferential form in problems of forced and free convection and other problems of heat and mass transfer, both two-dimensional and three-dimensional.

*)e-mail: bit@termo.msk.su

---

[1] H. Schlichting, *Boundary Layer Theory*, 6th ed., McGraw-Hill, New York (1968).
[2] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics*, 2nd ed., (Pergamon Press, Oxford (1987).
[3] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics,* Addison-Wesley, Reading, Mass. (1963), Vol. 2, Ch. 40.

Translated by R. M. Durham

Rotation of a droplet in a viscous fluid

K. I. Morozov

Institute of Mechanics of Continuous Media, Ural Department of the Russian Academy of Sciences, 614013 Perm', Russia

(Submitted 12 May 1997)
Zh. Éksp. Teor. Fiz. **112**, 1340–1350 (October 1997)

Unique capabilities for modeling the bulk motion of one liquid in another arise from the use of droplets made of a magnetic liquid. In this paper the low-frequency rotational motion of a magnetic droplet suspended in a viscous liquid is investigated. In this frequency range, the shape of the droplet does not depend on the field frequency and is determined only by its amplitude. An analytic solution has been found in the Stokes approximation to the problem, which generalizes the classic problem of Jeffrey to the case of a liquid ellipsoidal particle. This solution makes it possible to determine the velocity field inside and outside the liquid particle, the moments of the viscous forces acting on the droplet, its coefficient of rotational mobility. © *1997 American Institute of Physics.* [S1063-7761(97)01410-8]

1. INTRODUCTION

In a classic paper by Jeffrey,[1] which appeared at the end of the first quarter of this century, the problem of slow (Stokes) motion of a solid ellipsoidal particle in a viscous liquid was solved analytically. Subsequently, the results of Ref. 1 have been used many times to describe various dispersed systems and liquid crystals (see, e.g., Refs. 2 and 3). The study of liquids containing deformable particles, such as polymer solutions, suspensions of liquid droplets, and blood, came much later. In Ref. 4, Roscoe investigated the rheology of a polymer solution by modeling it as a suspension of viscoplastic spheres in a viscous Newtonian liquid. In Ref. 5, Keller *et al.* determined the behavior of an erythrocyte in a shear flow, treating the cell as a liquid particle bounded by an easily deformable shell. The basic problem addressed in these papers is how to determine the mechanism by which a liquid particle interacts with the surrounding liquid and how it influences the structure and properties of the system under study. It is noteworthy that in the papers cited so far the droplet is assigned a passive role, i.e., its shape and the character of its motion, including internal motion, are determined ultimately by the external flow of the viscous liquid. This circumstance hinders the design and interpretation of experimental studies.

The interesting possibility of avoiding these difficulties arises when magnetic liquids—i.e., colloidal dispersions of magnets in an ordinary liquid—are used as the droplet material.[6] The basic advantage of this approach is that the liquid particle can be controlled by external magnetic fields. This control includes setting it in motion and creating various stable droplet configurations. This approach was taken in Ref. 7, where a droplet of magnetic liquid was suspended in an electrolyte solution. In the absence of a magnetic field the droplet is, of course, a sphere, in our case with a diameter ~ 1 mm. In a constant magnetic field, the droplet elongates. In this case, the shape of the droplet, which is determined by the competition between magnetic and surface forces, can be approximated to high accuracy by a ellipsoid of rotation.[8] If the field is now made to rotate, the droplet will also rotate along with it. In Ref. 7, Lebedev *et al.* investigated the rotation of this field at low frequencies ($f \leqslant 1$ Hz). At these frequencies the droplet takes part in a simple forced motion in which its long axis rotates with angular velocity $\Omega = 2\pi f$. In this case, no change is observed in the shape of the droplet compared to the case of a time-independent field.

Thus, by varying the amplitude of the magnetic field we can investigate the rotational motion of a liquid ellipsoidal particle. Outwardly, this motion looks like rigid-body rotation (since the magnetic liquid is opaque); however, there is a considerable difference between the rotation of a solid particle and a liquid droplet, since motion within the droplet can occur in the latter case. In this paper we also discuss the low-frequency rotational motion of a droplet. When the droplets are sufficiently small, the approximation of small Reynolds number can be used in the calculations. Formulated in this way, the problem generalizes the problem of Jeffrey[1] to a liquid ellipsoidal particle.

The structure of this paper is as follows. In Sec. 2 ranges of fluid parameters and values of the external magnetic field are determined for which the droplet remains ellipsoidal, and equations of motion are formulated and solved. In Sec. 3 the motion within the droplet is calculated. In Secs. 4 and 5 a balance equation is derived for the moments of the forces and the rotational mobility of the droplet is determined. In the Appendix it is shown that the energy balance equation in the system is equivalent to an integral condition for the viscous tangential stresses at the droplet surface.

2. EQUATIONS OF MOTION AND THEIR SOLUTION

Rather than immediately writing down a system of equations of motion for the droplet, let us first estimate several parameters of the problem in order to clarify the reasons why low-frequency rotation of the magnetic field does not destabilize the droplet shape. In this problem two physical mechanisms determine the shape of the liquid magnetic particle—the magnetic field and the flow that arises inside and outside the droplet as it rotates. The effectiveness of each of these mechanisms is gauged by the value of a corresponding dimensionless parameter. In the first case this is the magnetic

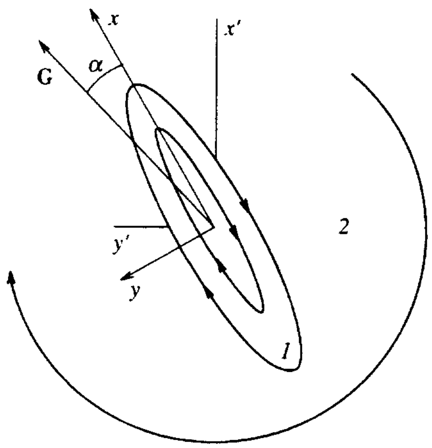

FIG. 1. Ellipsoidal magnetic droplet (*1*) in a viscous liquid (*2*) in a rotating magnetic field **G**; xyz and $x'y'z$ are axes of the rotating and laboratory coordinate systems.

Bond number $\mathrm{Bo}=G^2R/\sigma$, where G is the amplitude of the external magnetic field, R is the zero-field radius of the droplet, and σ is the surface tension at the boundary. In the second case it is the dimensionless frequency $\Omega\tau$, where $\tau=\eta_1 R/\sigma$ and η_1 is the viscosity of the liquid within the droplet. Using typical values (see data from the experiments of Ref. 7) for the quantities that enter into Bo and $\Omega\tau$ ($\eta_1\sim 10$ cp, $\Omega\leqslant 10\ \mathrm{sec}^{-1}$), we find that the contributions of both mechanisms are the same order of magnitude only for extremely weak fields ($G\sim 1$ Oe), which only slightly exceed the magnetic field of the Earth. Due to the quadratic dependence of the Bond number on G, the first mechanism becomes dominant even in a field of ~ 10 Oe. In what follows we assume that $\mathrm{Bo}\gg\Omega\tau$.

The quantity $\Omega\tau$, which is the ratio of the characteristic time τ for the droplet shape to relax to the period of the field, is also assumed to be small, i.e., $\Omega\tau\ll 1$. Actually, when we substitute a value of $\sigma\sim 10$ dyne/cm, which is typical of liquids, into the definition of τ, we find that for a droplet with $R\leqslant 1$ mm, $\Omega\tau\leqslant 10^{-2}$. The latter inequality is equivalent to saying that the long axis of the droplet lags behind the (moving) direction of the external magnetic field (see Sec. 4), but that this lag is so slight that the projection of the field on the long axis of the droplet essentially coincides with its stationary value (i.e., when $f=0$). Thus, when the frequency range and parameters of the droplet satisfy the conditions $\Omega\tau\ll 1$ and $\Omega\tau\ll\mathrm{Bo}$, the shape of the droplet is independent of Ω and is determined only by one parameter—the Bond number. In this range of parameters we can treat the droplet of magnetic liquid as an ellipsoid of rotation with semiaxes $a>b=c$, which corresponds to what is observed.[7]

Figure 1 shows an ellipsoidal magnetic droplet and the xyz coordinate system attached to it. In a constant field, the long axis of the droplet x coincides with the field direction, while in a rotating uniform magnetic field **G** it lags behind the latter by an angle α. The vector **G** rotates relative to the laboratory coordinate system $x'y'z$ with angular velocity $\boldsymbol{\Omega}$ directed along the z axis, i.e., counterclockwise. The z axis is common to both systems of coordinates.

When the droplet is acted on by the moment of the magnetic forces ($\alpha\neq 0$), it becomes entrained, i.e., the axes of the coordinate system attached to the droplet rotate relative to the laboratory coordinate system with the same angular velocity $\boldsymbol{\Omega}$ as the field. The droplet of magnetic liquid (region *1*) is surrounded by an unbounded volume of viscous liquid (region *2*) which is at rest at infinity. In the rotating system of coordinates xyz its shape is fixed, i.e., the velocity at the droplet surface has no normal components. Nevertheless, the tangential components of the velocity are not equal to zero: because the ellipsoid viewed in the rotating coordinate system is in an external flow with vorticity $-\boldsymbol{\Omega}$ (shown in Fig. 1 by the arrow in region *2*), a flow arises within it with opposite vorticity (clockwise in Fig. 1).

In the rotating coordinate system the equations of motion have the following form in the Stokes approximation:

$$\eta_{1,2}\nabla^2\mathbf{v}^{(1,2)}=\nabla p_{1,2}. \tag{1}$$

Here $\mathbf{v}^{(1,2)}$, $p_{1,2}$, and $\eta_{1,2}$ are the velocity fields, pressures, and viscosities of the two regions under discussion, i.e., inside and outside the droplet. The absence of bulk magnetic forces on the right side of Eq. (1) is a consequence of the uniformity of the magnetic field within the ellipsoid. The boundary conditions at the droplet surface

$$v_\nu^{(1)}=v_\nu^{(2)}=0,\quad v_\tau^{(1)}=v_\tau^{(2)}, \tag{2}$$

express the absence of normal and equality of tangential components of the velocity at the liquid boundary. At infinity the velocity field should be determined only by the external flow:

$$\mathbf{v}^{(2)}(r\to\infty)=-\boldsymbol{\Omega}\times\mathbf{r}. \tag{3}$$

Since the shape of the droplet is fixed in our treatment, the formulation of exact conditions for the stress components at the boundary between liquids overdetermines the problem. Instead, we supplement the condition of continuity of normal components of the stress with an integral relation—the condition that to zero order in $\Omega\tau$ the sum of magnetostatic and surface energies of the ellipsoid be an extremum. This case, which corresponds to a droplet of magnetic liquid in a constant field, was discussed in detail by Bacri *et al.* in Ref. 8, who obtained the following expression for the eccentricity of the ellipsoid $e=\sqrt{1-(b/a)^2}$ as a function of the magnetic Bond number:

$$\mathrm{Bo}=4\pi\left(\frac{1}{\mu-1}+n_1\right)^2\times\frac{(1-e^2)^{1/2}(3-2e^2)-(3-4e^2)(\arcsin e)/e}{(1-e^2)^{1/6}[(3-e^2)n_1-1+e^2]}. \tag{4}$$

Here μ is the magnetic permeability of the droplet and n_1 is the demagnetizing factor along the x axis.

The boundary condition for tangential stresses at the surface of the boundary will also be replaced by an integral relation that follows from energy balance in the system (see Sec. 3 and the Appendix).

Equation (1) with the supplementary conditions (2) and (3) for a liquid ellipsoid can be reduced to the Jeffrey problem for a solid particle. In fact, we look for solutions in the interior and exterior regions of the form

$$v_x^{(1)}=-\zeta ya/b,\quad v_y^{(1)}=\zeta xb/a,\quad v_z^{(1)}=0,\quad p_1=\text{const}, \tag{5}$$

$$v_x^{(2)}=u_J-\zeta ya/b,\quad v_y^{(2)}=v_J+\zeta xb/a,$$

$$v_z^{(2)}=\omega_J,\quad p_2=p_J, \tag{6}$$

where ζ is a parameter to be determined that characterizes the vorticity within the droplet. This leads to an equation analogous to (1) for the velocity field $\mathbf{v}_J=(u_J,v_J,\omega_J)$ and pressure p_J in region *2*, and the trivial solution $\mathbf{v}_J=0$ at the surface of the ellipsoid. Consequently $\mathbf{v}_J$ and p_J are the velocity field and pressure from the Jeffrey problem[1] for a motionless solid ellipsoid in an external flow with a tensor of velocity gradients $\partial v_1^{\infty}/\partial x_k\equiv\gamma_{ik}+\omega_{ik}$ prespecified at infinity. It follows from Eqs. (3) and (6) that the nontrivial components of the symmetric γ_{ik} and antisymmetric ω_{ik} parts of this tensor are

$$\gamma_{12}=\gamma_{21}=\frac{\zeta}{2}\left(\frac{a}{b}-\frac{b}{a}\right),\quad \omega_{12}=-\omega_{21}=\Omega+\frac{\zeta}{2}\left(\frac{a}{b}+\frac{b}{a}\right). \tag{7}$$

Expressions (5) and (6) constitute a formal solution to the system Eqs. (1)–(3) in terms of the quantities $\mathbf{v}_J$, p_J, and a single undetermined parameter ζ that describes the uniform and negative vorticity within the ellipsoid. We determine ζ in Sec. 3.

We will not give the explicit but extremely awkward expressions for $\mathbf{v}_J$ and p_J from Ref. 1 here, since only the value of the force $\mathbf{F}_J$ per unit area acting on the solid particle is important for our subsequent discussion. According to Ref. 1 (see also Ref. 3), the quantity $\langle F_J\rangle_i\equiv\sigma_{ik}\nu_k$ (where σ_{ik} is the stress tensor, and ν_k are components of the outward normal vector with respect to the ellipsoid) equals

$$(F_J)_i=-p\nu_i+\eta A_{ik}\nu_k. \tag{8}$$

Here p is a certain constant and η is the viscosity of the liquid surrounding the solid particle. The matrix quantities A_{ik} are linear combinations of components of the velocity gradient tensor of the external flow. In our problem, Eq. (7) implies that only the elements A_{12} and A_{21} are nonzero, where

$$A_{12}=2\frac{n_1\gamma_{12}+b^2n_3'\omega_{12}}{(a^2n_1+b^2n_2)n_3'},\quad n_3'=\frac{n_2-n_1}{a^2-b^2}. \tag{9}$$

The values of A_{21} are obtained from A_{12} by the simple replacements $1\leftrightarrow 2$, $a\leftrightarrow b$. The quantity n_2 entering into Eq. (9) is the demagnetizing factor of the ellipsoid along the y axis.

Combining Eqs. (6) and (8), we find the force $\mathbf{F}$ exerted by the nonmagnetic liquid (region *2*) per unit surface of the droplet:

$$F_i=-p\nu_i+\eta_2(A_{ik}-2\gamma_{ik})\nu_k. \tag{10}$$

3. CALCULATION OF THE MOTION WITHIN THE DROPLET

In order to determine the parameter ζ, which characterizes the rate of motion within the droplet, we will use the following integral condition:

$$\int v_i\sigma_{ik}'^{(1)}\nu_k dS=\int v_i\sigma_{ik}'^{(2)}\nu_k dS. \tag{11}$$

Here the integration is carried out over the surface of the ellipsoid and σ_{ik}' is the tensor of viscous stresses.

In our problem the liquid particle is assumed to be ellipsoidal, and the integral relation (11) replaces the exact condition of continuity of tangential components of the stresses (according to Eq. (2) $v_\nu=0$ at the droplet surface) at the boundary between liquids. In the Appendix we show that Eq. (11) is the form taken by the energy balance equation in the system.

Let us carry out the calculation using Eq. (11). From Eq. (10), the contraction $\sigma_{ik}'^{(2)}\nu_k$ is given by

$$\sigma_{ik}'^{(2)}\nu_k=\eta_2(A_{ik}-2\gamma_{ik})\nu_k. \tag{12}$$

According to the definition of the viscous stresses,[9] and taking into account Eqs. (5) and (7), we find for the interior region that

$$\sigma_{ik}'^{(1)}\nu_k=-2\eta_1\gamma_{ik}\nu_k. \tag{13}$$

Substituting Eqs. (12) and (13) into Eq. (11) and integrating, we find an equation that determines ζ:

$$\frac{a}{b}A_{12}-\frac{b}{a}A_{21}=\zeta\left(\frac{a}{b}-\frac{b}{a}\right)^2\left(1-\frac{\eta_1}{\eta_2}\right). \tag{14}$$

It follows from Eq. (14) that ζ is a linear function of the angular velocity Ω of rotation of the long axis of the particle. We will not present its explicit and awkward expression here (see Eqs. (7) and (9), which determine A_{12} and A_{21}), but instead limit ourselves to writing out the quantity ζ/Ω in two limiting cases—small eccentricity ($e\rightarrow 0$, or a quasispherical droplet) and large eccentricity ($e\rightarrow 1$, or a droplet like an elongated cylinder):

$$\frac{\zeta}{\Omega}=\begin{cases}-1+e^4\left(\dfrac{1}{8}+\dfrac{\eta_1}{6\eta_2}\right), & e\rightarrow 0\\[2ex] \dfrac{8\sqrt{1-e^2}}{\ln(1-e)}\,\dfrac{\eta_2}{\eta_1+\eta_2}, & e\rightarrow 1\end{cases}. \tag{15}$$

As expected, ζ is everywhere negative. In Fig. 2 we show the results of these calculations based on Eq. (14) for the modulus $|\zeta|/\Omega$ as a function of eccentricity e of the droplet, plotted for several values of the ratio η_1/η_2 (labeling the corresponding curves). Of course, the limit $\eta_1/\eta_2\rightarrow\infty$ corresponds to no motion within the droplet $(\zeta=0)$, in which case the problem reduces to the Jeffrey problem for a solid particle (see Eqs. (5) and (6)).

4. BALANCE EQUATION FOR TORQUES

Let us compute the torque $\mathbf{L}$ due to viscous friction forces acting on the ellipsoidal particle during its rotation. From considerations of symmetry it is clear that the only nontrivial component of $\mathbf{L}$ will be $L_z\equiv L$, which equals

$$L=\int(xF_y-yF_x)dS, \tag{16}$$

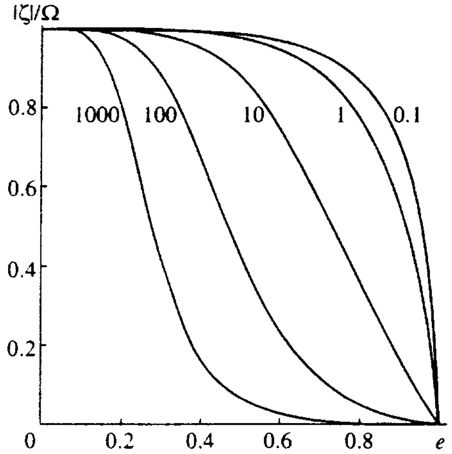

FIG. 2. Absolute value of the vorticity parameter within the droplet in units of Ω as a function of eccentricity for six values of η_1/η_2, which label the corresponding curves.

where dS is an element of surface area of the ellipsoid. Substituting the quantity F_i into Eq. (16) from Eq. (10), and carrying out the integration, we find

$$L=\eta_2(A_{21}-A_{12})V, \tag{17}$$

where V is the volume of the droplet.

The torque due to viscous friction forces Eq. (17) counterbalances the torque L_m due to external (magnetic) forces: $L+L_m=0$. We could calculate L_m by analogy with L, i.e., by considering the magnetic (Maxwell) stresses at the ellipsoidal surface. However, the results for the torque due to magnetic forces acting on an ellipsoid in a uniform external field are well known[10]:

$$L_m=(\mu-1)^2(n_2-n_1)H_xH_yV/4\pi. \tag{18}$$

Here by μ we mean the same static magnetic permeability as we used previously in Eq. (4), since the dispersion of μ at low frequencies is negligibly small.[11] The components H_x and H_y of the uniform magnetic field within the ellipsoid are related to the magnitude of the external field G by[10]

$$H_x=G\cos\alpha/(1+(\mu-1)n_1),$$

$$H_y=G\sin\alpha/(1+(\mu-1)n_2). \tag{19}$$

Finally, the balance condition on the torques due to viscous and magnetic forces acting on a liquid particle can be written in the form

$$4\pi\eta_2(A_{12}-A_{21})=(\mu-1)^2(n_2-n_1)H_xH_y. \tag{20}$$

This equation determines the lag angle α of the long axis of the ellipsoid (the x axis) behind the direction of the external field $\mathbf{G}$. In general it is impossible to calculate the angle α for arbitrary eccentricities of the droplet, because as e increases, which means that the external field G increases as well (see Eq. (4)), it is necessary to take into account the dependence of the magnetic permeability on the internal field H.[6] Here we write only the characteristic value of $\sin 2\alpha$ for a marginally elongated droplet ($e\rightarrow 0$), for which the dependence $\mu=\mu(H)$ is negligible:

$$\sin 2\alpha\rightarrow\frac{5}{4}\left(2+3\frac{\eta_2}{\eta_1}\right)\Omega\tau,\quad e\rightarrow 0. \tag{21}$$

As we already mentioned in Sec. 2, the magnitude of the angle by which the droplet shape lags the external field is determined by the small parameter $\Omega\tau$.

5. ROTATIONAL MOBILITY OF THE DROPLET

By analogy with well-known problems of flow around solid bodies in the Stokes approximation,[2] we consider the mobility B associated with apparent rotational motion of the droplet relative to the direction perpendicular to its axis of symmetry, i.e., relative to the z axis in our treatment. Using the definition of mobility as the ratio of the angular velocity of rotation of the droplet Ω to the external magnetic torque L_m and Eqs. (7) and (20), we obtain

$$B=3B_{HS}\frac{a^2n_1+b^2n_2}{a^2+b^2+2ab\zeta/\Omega}. \tag{22}$$

Here $B_{HS}=1/6\eta_2V$ is the rotational mobility of a solid spherical particle. If we set $\zeta=0$, then Eq. (22), of course, reduces to Jeffrey's result for the mobility of a solid ellipsoidal particle.[1–3] In the present case of a liquid droplet with nontrivial internal motion, the ratio ζ/Ω should be taken from the solution to Eq. (14). In the limit of a quasispherical particle it follows from Eqs. (22), (14), (9), and (7) that

$$B=\frac{12\eta_2}{2\eta_1+3\eta_2}\frac{B_{HS}}{e^4},\quad e\rightarrow 0. \tag{23}$$

This result appears paradoxical. Actually, unless the ellipsoid is solid, for which $B\rightarrow B_{HS}$ as $e\rightarrow 0$, the rotational mobility of the droplet will increase without bound for any finite value of η_1/η_2. This implies that even an infinitely weak field is sufficient to cause rotation of the droplet with finite velocity Ω. However, the latter result, at first glance, violates energy conservation. This ''paradox'' has a simple explanation: by the ''rotation'' of the droplet we mean its apparent motion, i.e., the rotation of the long axis of the droplet in the laboratory coordinate system $x'y'z$ (see Fig. 1). The velocities $\mathbf{v}'$ of material particles in the laboratory system of coordinates are related to the velocities $\mathbf{v}$ in the rotating system of coordinates by the usual kinematic relation $\mathbf{v}'=\mathbf{v}+\mathbf{\Omega}\times\mathbf{r}$, and it follows from Eqs. (5), (6) and (15) that as $e\rightarrow 0$ these quantities decrease to zero.

In the limit of a marginally elongated particle, the mobility of the droplet B approaches its rigid-body value because the internal motion disappears with increasing e (see Fig. 2). In Fig. 3 we plot the rotational mobility of the droplet B calculated using Eqs. (22), (14), (9), and (7) in units of B_{HS} as a function of the eccentricity of the ellipsoid. The mobility is calculated for five values of η_1/η_2 (labeling the corresponding curves). The limit $\eta_1/\eta_2\rightarrow\infty$ corresponds to a solid ellipsoid. Of course, the differences in character of rotation of the liquid and solid ellipsoids are associated with the motion within the droplet.

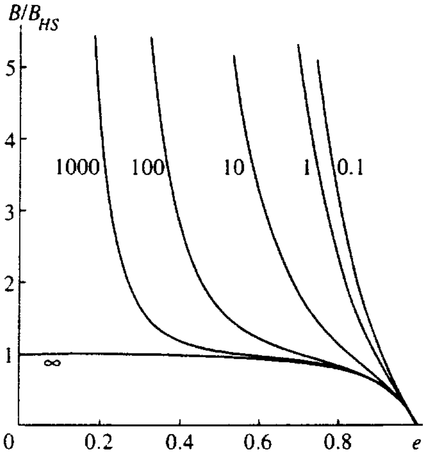

FIG. 3. Mobility of the liquid droplet B relative to the mobility B_{HS} of a solid spherical particle as a function of eccentricity for the same values of η_1/η_2 as in Fig. 2.

6. CONCLUSION

In this paper we have touched on only one aspect of the rotational motion of a droplet suspended in a viscous liquid, specifically the way the motion within the droplet interacts with an external flow, assuming that the shape of the droplet remains unchanged. The applicability of our representation of the droplet as an undeformed particle is limited to low frequencies. The experiments of Ref. 7 show that the shape of the droplet is unchanged up to a certain frequency f_1. At $f=f_1$ the droplet discontinuously elongates severalfold, transforming into a ''dumbbell.'' As the frequency of the field increases further, up to a value of f_* (which, however, differs only very slightly from f_1), dynamic equilibrium is established in the system, in which the droplet–dumbbell splits into two equal particles with the same shape as the original droplet. These individual droplets execute from one to several half-turns and once more reunite into a single droplet–dumbbell, which after several rotations once more splits into two, etc. Subsequent increases in the frequency shift this dynamic equilibrium more and more towards a configuration with two droplets, until a frequency f_2 is reached at which the ''return'' to the droplet–dumbbell shape does not occur. The droplets that form are similar to the original droplet, and as the frequency increases they also undergo division.[7]

For droplets with a radius from one to several millimeters, the frequencies f_1 at which large deformations begin are 1–3 Hz.[7] At these frequencies the Reynolds numbers for the flow within the droplet $\mathrm{Re}=\rho\Omega R^2/\eta_1$ are no longer small. Hence, the behavior of a droplet at high frequencies must be studied outside the framework of the Stokes approximation, and the formulation and solution of the stability problem requires a separate discussion.

The author is sincerely grateful to A. F. Pshenichnikov for useful comments and discussion of the results of this work. This work was carried out with the financial support of the Russian Fund for Fundamental Research (Project 95-01-00408).

APPENDIX

We show that the integral relation (11) is equivalent to the condition of energy balance in the system. Let W_1 and W_1 be the energies dissipated per unit time inside and outside the droplet. Denoting by A the power generated by external magnetic forces, let us write the energy balance equation in the system as

$$A=W_1+W_2. \tag{A1}$$

From Eqs. (17) and (18) we have

$$A=L_m\Omega=\eta_2(A_{12}-A_{21})\Omega V. \tag{A2}$$

We now compute W_1 and W_2. According to Ref. 9 the energy dissipated in the droplet is

$$W_1=\frac{\eta_1}{2}\int_E\left(\frac{\partial v_i^{(1)}}{\partial x_k}+\frac{\partial v_k^{(1)}}{\partial x_i}\right)^2 dV. \tag{A3}$$

Here the integration is carried out over the droplet volume, a fact that is denoted by the letter E below the integral sign. Using the equation of motion (1), we reduce the volume integral in Eq. (A3) to a surface integral in the standard way[2,9]:

$$W_1=\int_E v_i(\sigma_{ik}'^{(1)}-p_1\delta_{ik})\nu_k dS. \tag{A4}$$

The upper index on the velocity field in Eq. (A4) is not specified: by virtue of the boundary conditions Eq. (2) we can use $\mathbf{v}$ to denote the value of the velocity common to both liquids at the surface of the droplet. Since $v_\nu=0$ (see Eq. (2)), the dissipation of energy in the droplet can be written entirely in terms of the viscous stress tensor $\sigma_{ik}'^{(1)}\equiv\eta_1(\partial v_i^{(1)}/\partial x_k+\partial v_k^{(1)}/\partial x_i)$:

$$W_1=\int_E v_i\sigma_{ik}'^{(1)}\nu_k dS. \tag{A5}$$

By analogy with W_1 we write the dissipation of energy in the liquid surrounding the droplet in the form of two integrals— one over the surface of a sphere with large radius Λ $(\Lambda\gg a)$ and one over the surface of the ellipsoid:

$$W_2=\int_\Lambda v_i^{(2)}(\sigma_{ik}'^{(2)}-p_2\delta_{ik})\nu_k dS-\int_E v_i\sigma_{ik}'^{(2)}\nu_k dS. \tag{A6}$$

The minus sign in front of the second term is connected with the fact that the symbol ν_k in Eqs. (A4) and (A5) denotes the outward normal to the ellipsoid and the sphere.

In order to calculate W_2 we need to know the behavior of the fields $\mathbf{v}^{(2)}$ and p_2 far from the ellipsoid. Using Eqs. (6) and (7) and Jeffrey's analytic solution,[1] we find the following asymptotic dependence of the velocity and pressure up to quantities of order r^{-2} and r^{-3} respectively:

$$v_x^{(2)}=\Omega y-C\,\frac{x^2y}{r^5}-D\,\frac{y}{3r^3}, \tag{A7}$$

$$v_y^{(2)}=-\Omega x-C\,\frac{xy^2}{r^5}+D\,\frac{x}{3r^3}, \tag{A8}$$

$$v_z^{(2)} = -C\,\frac{xyz}{r^5}, \tag{A9}$$

$$p_2 = -2\,\eta_2 C\,\frac{xy}{r^5}, \tag{A10}$$

where

$$C = R^3(A_{12}+A_{21})/2, \quad D = R^3(A_{12}-A_{21})/2. \tag{A11}$$

The first terms on the right side of Eqs. (A7) and (A8) give the unperturbed velocity field in the rotating system of coordinates; all the remaining terms in Eqs. (A7)–(A10) are perturbations due to the droplet.

The integral in Eq. (A6) over the distant sphere (in what follows we will call this W_Λ) is conveniently calculated in spherical coordinates $r,\ \varphi,\ \theta$, in which the velocity field and pressure can be rewritten in the form

$$v_r^{(2)} = -\frac{C}{r^2}\sin^2\,\theta\,\sin\,\varphi\,\cos\,\varphi, \tag{A12}$$

$$v_\varphi^{(2)} = \left(-\Omega r + \frac{D}{3r^2}\right)\sin\,\theta, \tag{A13}$$

$$v_\theta^{(2)} = 0, \quad p_2 = 2\,\eta_2\,\frac{v_r}{r}. \tag{A14}$$

As is clear from Eqs. (A7)–(A9) or (A12)–(A14), the components of the viscous stress tensor are of order r^{-3}. Since we will take the limit $\Lambda \to \infty$ after we have calculated W_Λ, we need only keep the terms of order r^{-2} in the integral over the distant sphere. W_Λ can then be expressed in terms of a single component of the viscous stress tensor:

$$W_\Lambda = \int_\Lambda v_\varphi^{(2)} \sigma_{\varphi r}'^{(2)} dS. \tag{A15}$$

We now calculate $\sigma_{\varphi r}'^{(2)}$. According to Eq. (9),

$$\sigma_{\varphi r}'^{(2)} = \eta_2\left(\frac{\partial v_\varphi^{(2)}}{\partial r} + \frac{1}{r\,\sin\,\theta}\,\frac{\partial v_r^{(2)}}{\partial\varphi} - \frac{v_\varphi^{(2)}}{r}\right).$$

Hence, substituting the values of the velocity components from Eqs. (A12)–(A14), we obtain

$$\sigma_{\varphi r}'^{(2)} = -\frac{\eta_2\,\sin\,\theta}{r^3}\,(D + C\,\cos\,2\varphi). \tag{A16}$$

The integration in Eq. (A15), taking into account Eqs. (A13) and (A16), yields

$$W_\Lambda = \frac{8\pi}{3}\,\eta_2 D\Omega = \eta_2(A_{12}-A_{21})\Omega V. \tag{A17}$$

Comparison of the latter result with Eq. (A2) shows that $W_\Lambda = A$. As a result, expressions (A1), (A5), and (A6) finally lead to the integral condition (11).

[1] G. B. Jeffrey, Proc. R. Soc. London, Ser. A **102**, 161 (1922).
[2] J. Happel and H. Brenner, *Low Reynolds Number Hydrodynamics*, Prentice-Hall, Englewood Heights (1965).
[3] V. N. Pokrovskiĭ, *Statistical Mechanics of Dilute Suspensions* [in Russian], Nauka, Moscow (1978).
[4] R. Roscoe, J. Fluid Mech. **28**, 273 (1967).
[5] S. R. Keller and R. Skalak, J. Fluid Mech. **120**, 27 (1982).
[6] M. I. Shlomis, Usp. Fiz. Nauk **112**, 427 (1974) [Sov. Phys. Usp. **17**, 153 (1974)].
[7] A. V. Lebedev and K. I. Morozov, JETP Lett. **65**, 160 (1997).
[8] J.-C. Bacri and D. Salin, J. Phys. (France) Lett. **43**, 649 (1982).
[9] L. D. Landau and E. M. Lifshitz, *Fluid Mech.*, 2nd ed., Pergamon, New York (1987).
[10] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media*, 2nd ed., Pergamon, New York (1984).
[11] A. F. Pshednichnikov and A. V. Lebedev, Zh. Éksp. Teor. Fiz. **95**, 869 (1989) [Sov. Phys. JETP **68**, 4988 (1989)].

Translated by Frank J. Crowne

Phase-transition kinetics with the formation of topological defects in superconductors with a multicomponent order parameter

A. S. Zel'tser, A. V. Radievskiĭ, and A. É. Filippov

Donetsk Physicotechnical Institute, Ukrainian National Academy of Sciences, 340114 Donetsk, Ukraine

(Submitted 17 June 1996)
Zh. Éksp. Teor. Fiz. **112**, 1351–1373 (October 1997)

This paper discusses the kinetics of phase transitions to superconductivity with a multicomponent order parameter in zero external field. It is shown that as it approaches equilibrium the superconductor passes through an intermediate vortex-like state containing domain walls, single-quantum, and multiquantum axially nonsymmetric vortices and antivortices. The energy and other parameters of the domain walls are derived. Rigid superconducting bubbles are discussed and criteria are established for their local stability. © *1997 American Institute of Physics.* [S1063-7761(97)01510-2]

1. INTRODUCTION

Investigation of the kinetics of formation and stable existence of nontrivial localized structures in media with various physical properties (elementary particles, condensed media, cosmic physics, etc.) is a vigorously evolving area of physics with a large number of publications (see, for example, Refs. 1–13, and the references cited therein).

At the present time it is well established that although a variety of nonuniform states can nucleate as a result of spontaneous symmetry breaking, the only objects that can exist without instability are those with nontrivial topological structure (i.e., which cannot be transformed in a continuous way into a uniform structure), i.e., structures that possess topological charge. Therefore, the study of nonuniform structures begins with the elucidation of possible topological classes and subclasses to which the defects in a given nonequilibrium medium may belong. In particular, there are many examples of topologically stable defects (deviations from local equilibrium) in condensed matter: line singularities (vortices in superconductors and superfluid He^3, vertical Bloch lines in magnets, etc.); point singularities (monopoles in liquid crystals and superfluid He^3, Bloch points in magnets, etc.); or singularities in the form of surfaces (domain walls in superfluid He^3, magnets, multicomponent superconductors, etc.).

The existence of conservation laws for topological charge determines not only the stability of these structures relative to weak external forces and dissipation, but also in many cases determines the character of the interaction between the defects. A clear example of this is the phenomenon of topological confinement. In condensed media, two types of confinement are observed: either one defect serves as a boundary for another defect of higher dimensionality, or a defect can only exist inside another object, from which it cannot escape. The latter type of confinement is commonly encountered in magnets and superfluid He^3, and as we will show below, may be encountered in multicomponent superconductors as well.

The most complicated problem in the study of topological defects is their formation kinetics, which involves their motions, interactions between them, and their interconversion, merging, and annihilation. Significant progress in dealing with these phenomena has been achieved only by studying systems with simple order parameters and low spatial dimensions, such as vortices in superfluid He^3 and ordinary superconductors[4,13–16] and domain walls and vortices in magnets.[17]

Increasing the number of components of the order parameter, for example in liquid crystals, greatly complicates the problem.[2,3,9] For such interesting physical systems as He^3, the number of components of the order parameter is as large as $n=18$, which leads to a variety of superfluid phases[1,18,19] that combine the properties of magnets, liquid crystals, ferroelectrics, and multicomponent superconductors. Investigations of the kinetics of He^3 meet with colossal difficulties.[1,12] Rigorous analysis of the possible topological singularities in the He^3 system using the method of homotopic groups[18,19] postulates configurations in the form of domain walls, vortices, particle-like excitations, etc. Topological objects in He^3 are characterized not only by topological charge but also by elements of the symmetry group, while among the defects with a mantle (vortices), we must distinguish singular and nonsingular.[1,11] The existence of most of these defects, and also complex combinations of them (for example, chains of vortices in layers), has been confirmed experimentally.[10,12,20]

The exotic ''heavy fermion'' superconductors, in which multicomponent order parameters have been experimentally identified, approach He^3 in complexity[21,22] (the number of components of the order parameter can reach $n=10$; see Ref. 23). In light of the important analogy between multicomponent superconductors and superfluid He^3,[18,21] we expect these superconductors to exhibit topological singularities with configurations analogous to those observed in He^3. To the best of our knowledge, the current literature contains no studies of the formation kinetics of vortex states in these superconductors; however, the multicomponent nature of their order parameters and the presence of gauge vector potentials (whose influence is especially important in the kinetics) should cause these states to differ from well-known states in ordinary superconductors in highly nontrivial ways.[14–16]

The complexity of these systems arises from the fact that seed formation and generation of a vortex state cannot be separated in them, but instead form essentially a single transition process. Because a complete picture of this process is unavailable at the present time, it is our hope that direct numerical simulation of a phase transition in a multicomponent superconductor will clarify our understanding of these questions somewhat. This clarification is the goal of the present paper.

2. MODEL

The exotic superconductors can be investigated phenomenologically without analyzing the microscopic pairing mechanism.[21,22] It is customary to assume that the phase transition to the superconducting state involves a lowering of the original full symmetry of the superconducting crystal $G=G_0\times R\times U(1)$, where G_0 is the point group of macroscopic crystal symmetry, R is the time reversal operator, and $U(1)$ is the group of gauge transformations. When the transition is accompanied only by a breaking of gauge invariance, we speak of ordinary superconductivity. In this case the order parameter Δ is a complex scalar (with two components, i.e., $n=2$). If, however, in addition to the breaking of gauge invariance we also observe breaking of point-group symmetry or time-reversal symmetry R in the superconducting state, we speak of unusual (exotic) superconductivity. In the latter case, the order parameter can be expanded in basis functions belonging to one of the irreducible representations of the crystal point group:

$$\Delta(\mathbf{k})=\Sigma\,\eta_j\Phi_j(\mathbf{k}). \tag{1}$$

If the dimension of the corresponding representation is greater than unity, then the order parameter is multicomponent ($n>2$) and the Ginzburg–Landau functional is obtained by expanding the energy near T_c in powers of η_j in the form of combinations that are invariant under all the symmetry operations of the system.

Symmetry analysis of the superconducting classes of heavy-fermion systems[21,22,24] shows that they can be related to systems with anisotropic d-type pairing. Under certain restrictions, this allows us to write the simplest Ginzburg–Landau functional in the form (see also Refs. 25–32)

$$\begin{aligned} F=\frac{1}{2}\int dv\biggl\{ & a\,\boldsymbol{\eta}\cdot\boldsymbol{\eta}^*+\frac{\beta_1}{2}(\boldsymbol{\eta}\cdot\boldsymbol{\eta}^*)^2+\frac{\beta_2}{2}|\boldsymbol{\eta}\cdot\boldsymbol{\eta}|^2 \\ & +\beta_3(|\eta_x|^4+|\eta_y|^4)+K_1D_i^*\eta_j^*D_i\eta_j \\ & +K_2D_i^*\eta_i^*D_j\eta_j+K_3D_i^*\eta_j^*D_j\eta_i+K_4D_z^*\eta_j^*D_z\eta_j \\ & +\gamma(\nabla\times\mathbf{A})^2\biggr\}. \end{aligned} \tag{2}$$

where $D_j=\partial_j-igA_j$; $j=x,y$; $\mathbf{A}$ is the vector potential; $\boldsymbol{\eta}=\{\eta_x,\eta_y\}$ is an order parameter with complex vector coefficients $\eta_x=\eta_1+i\eta_2$, $\eta_y=\eta_3+i\eta_4$; $a=\alpha(T-T_c)$; $\beta_1,\beta_2,\beta_3,K_j$ are phenomenological constants; and $\gamma=1/8\pi$, $g=2e/\hbar c$. The condition of magnetic stability[22,26,28] and the requirement that the fourth-order form in Eq. (2) be positive definite impose the following restrictions on β_1, β_2, β_3, and K_j:

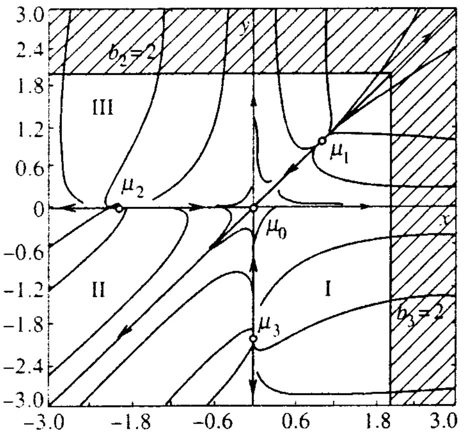

FIG. 1. Phase diagram and phase portrait of the equations of the renormalization group. The boundaries of positive definiteness of the fourth-order form are hatched.

$$\begin{gathered} K_1>|K_2|,\quad K_1+K_2+K_3>|K_2|,\quad K_4>0, \\ \beta_1+\beta_2+\frac{\beta_3}{2}+\min\left\{\frac{\beta_3}{2},-\frac{\beta_2+|\beta_2|}{2}\right\}>0. \end{gathered} \tag{3}$$

All of the new superconductors are good type II superconductors (i.e., the Landau–Ginzburg parameter $\kappa\gg 1$).

3. PHASE DIAGRAM AND RENORMALIZATION GROUP ANALYSIS

A large number of papers[25–32] have been devoted to studying functionals like Eq. (2), including the construction of phase diagrams. In this section we discuss fluctuation-induced corrections to the results of Landau theory.

In the fluctuation regime, the parameters of the free-energy expansion Eq. (2) are renormalized by the evolving critical fluctuations. These renormalizations are described by the equations of the renormalization group, which in the first ε-approximation have the form[26]

$$\begin{aligned} u'&=u-5u^2-w^2-v^2-2u(w+2v)\delta, \\ v'&=v-4uv-2v^2-4w^2-2(v^2+2wv+2u^2)\delta, \\ w'&=w[1-2u-4v]-(u^2+v^2+5w^2)\delta, \end{aligned} \tag{4}$$

where $u=(\beta_1+\beta_2+\beta_3)/6$, $v=\beta_1/6$, and $w=\beta_2/6$. The constant δ is proportional to the square of the anisotropy γ^2 of the gradient part of the functional:

$$\delta F=\gamma\int d^dk\;k_xk_y(\eta_x\eta_y^*+\eta_y\eta_x^*)$$

and is small to the extent that the anisotropy is small.

The corresponding phase portrait of the system of Eqs. (4) is shown in Fig. 1, plotted using axes $x=(u-w-v)/u(=\beta_3/(\beta_1+\beta_2+\beta_3)=b_3)$ and $y=(u+w-v)/u(=(2\beta_2+\beta_3)/(\beta_1+\beta_2+\beta_3)=b_2)$. This phase portrait shows the boundary of positive definiteness and the boundaries separating regions of energetic advantage for mean-field theory phases.[21,26] It is easy to show that the latter boundaries coincide with separatrices of the renormalization group, i.e., curves that separate regions in which the phase trajectories are attracted to various stability boundaries. We have the following three symmetrically distinct phases and conditions for their realization:

$$\text{Sector I:} \quad \beta_2<0, \quad \beta_3>0, \quad \text{phase } \boldsymbol{\eta}=(1,1); \tag{5a}$$

$$\text{Sector II:} \quad 2\beta_2+\beta_3<0, \quad \beta_3<0,$$

$$\text{phase } \boldsymbol{\eta}=(1,0) \text{ or } (0,1)); \tag{5b}$$

$$\text{Sector III:} \quad 2\beta_2+\beta_3>0, \quad \beta_2>0, \quad \text{phase } \boldsymbol{\eta}=(1,i). \tag{5c}$$

The usual ideology of the renormalization group says that the crossing of stability boundaries by phase trajectories (the dashed curves in Fig. 1) implies fluctuation-induced discontinuous phase transitions to structures III, II, and I respectively. As a result, these structures turn out to be the same as those predicted by Landau theory. The only difference is that the phase transition changes from second-order to first-order. However, in this case the system immediately makes a transition (for the corresponding β_j and weak quadratic anisotropy) to the low-temperature phases I or III, bypassing intermediate states like II, as happens in mean-field theory.

Thus, in simulating a phase transition in our system numerically it is sufficient to investigate three sets of parameters of the Ginzburg–Landau functional, corresponding to the three symmetrically distinct phases Eqs. (5a)–(5c).

Previous work has shown that the functional Eq. (2) belongs to the class of universal Ginzburg–Landau functionals with interacting vector fields[30] $\boldsymbol{\phi}_1=(\eta_1,\eta_2)$, $\boldsymbol{\phi}_2=(\eta_3,\eta_4)$. In the context of the present paper, it is convenient to use this class to clarify the topology. In particular, when $\beta_2>0$ a phase with $\boldsymbol{\phi}_1\perp\boldsymbol{\phi}_2$ is realized, while $\beta_2<0$ corresponds to the phase $\boldsymbol{\phi}_1\parallel\boldsymbol{\phi}_2$. The phase difference $(\varphi_1-\varphi_2)$ has the sense of the angle between the vectors $\boldsymbol{\phi}_1$ and $\boldsymbol{\phi}_2$. Moreover, in Ref. 26 Ivanchenko *et al.* show that the functional Eq. (2) has an additional symmetry under simultaneous exchange of the vertices $(-\beta_3)\leftrightarrow 2\beta_2$ and the rotations $(\eta_1\pm\eta_4)/2^{1/2}=\varphi_{1,4}$, $(\eta_2\pm\eta_3)/2^{1/2}=\varphi_{2,3}$, an operation that transforms the structures II and III into each other. Furthermore, the replacement $(-\beta_3)\leftrightarrow 2\beta_2$ inverts sectors II and III about the line that separates them, $b_2=0$. The existence of symmetry transformations that permute these phases among themselves permits us to investigate (at least theoretically) only one (either) of the sectors II or III. However, in the presence of strong nonuniformity or for samples of finite dimensions, direct application of this principle is not always justified, since the study of either of these phases requires numerical conversion of masses of data and boundary conditions. Hence, it is convenient to simulate numerically each of these sectors independently and compare the results. The fact that the kinetics of appearance of topological singularities in phases II and III are the same both qualitatively and quantitatively for the set of functional parameters given above indirectly confirms the correctness of the calculations.

The renormalization group analysis shows that the phase transition to superconductivity in systems with anisotropic pairing should be a first-order fluctuation-induced phase transition.[26,30] Consequently, it should be accompanied by seed-forming processes. In Ref. 33, Ivanchenko *et al.* attempt to explain the phenomenology of this process by including specific long-lived nonlinear excitations of soliton type, claiming that these excitations, which exist above T_c, are good candidates to play the role of seeds of the superconducting phase as the temperature of the system is lowered. In the present paper we discuss a possible role for topological singularities in this process.

4. NUMERICAL SIMULATION

Vortices are fundamentally mesoscopic formations, which should be generated spontaneously in the kinetics of ordering as dissipative attractors of the fluctuating system.[15] We can show that an ordinary superconductor, which is ordered in zero external field, unavoidably passes through an intermediate vortex-like state where the average mesoscopic magnetic field is zero, $\langle\mathbf{h}\rangle=0$, but the mean squared field is nonzero, $\langle\mathbf{h}^2\rangle\neq 0$. In this case, topological configurations that correspond to pairs of vortices arise from amorphous fluctuations at an arbitrary position.[13,14] For a two-component field, vortices form at points where lines of zeroes for both components of $\Delta(x,y)$ intersect. Fluctuations of both components of the order parameter field Δ also interact with the gauge electromagnetic field $\mathbf{A}$, which also generates intrinsic magnetic vortices.

When the number of components of the order parameter Δ is large, the simultaneous intersection of lines of zeroes for all the components is improbable, so that the simplest kinetic scenarios are no longer sufficient to describe the formation and structure of vortices. Since there is no doubt about the existence of a vortex state in systems with heavy fermions,[25,29] we may infer that a kinetic mechanism should exist that leads to spontaneous formation of axially asymmetric two-quantum vortices. In this case there is no requirement that all the components of the order parameter vanish at a single point.

In the presence of fluctuation noise $f(\mathbf{r},t)$, the evolution of the order parameter $\boldsymbol{\eta}$ and gauge field $\mathbf{A}$ can be described by a well-known modification of the Landau–Khalatnikov equations[13] in the form

$$\frac{\partial\eta_j}{\partial t}=-\Gamma\,\frac{\delta H}{\delta\eta_j}+f_j(\mathbf{r},t),$$

$$\frac{\partial\mathbf{A}}{\partial t}=-\Gamma_{\mathbf{A}}\,\frac{\delta H}{\delta A}+f_{\mathbf{A}}(\mathbf{r},t). \tag{6}$$

where Γ and $\Gamma_{\mathbf{A}}$ are positive relaxation constants, while the correlators for the noise fluctuations have the form

$$\langle f(\mathbf{r},t)\rangle=0,$$

$$\langle f_j(\mathbf{r},t)f_i(\mathbf{r}',t')\rangle=D\,\delta(\mathbf{r}-\mathbf{r}')\,\delta(t-t')\,\delta_{ji}. \tag{7}$$

Taking the anisotropy of these materials into account (and also for simplicity), we show that all the quantities depend weakly on the coordinate z (i.e., we solve Eq. (6) numerically in a slice of the superconductor in the plane perpendicular to the c axis). In addition, as a first approximation we take K_2, $K_3\ll K_1$. As a result, we have

$$\frac{\partial \eta_j}{\partial t} = \Delta \eta_j - (-1)^j (2\mathbf{A}\cdot\nabla\eta_k + \eta_k \nabla\cdot\mathbf{A}) - \eta_j[\mathbf{A}^2 - 1$$

$$+ S] + b_2 M \frac{\partial M}{\partial \eta_j} + b_3 L \frac{\partial L}{\partial \eta_j} + f_j,$$

$$(j = 1,...,4, \quad k = j - (-1)^j),$$

$$\vartheta_{\mathbf{A}}^{-1} \frac{\partial \mathbf{A}}{\partial t} = -\nabla\times(\nabla\times\mathbf{A}) + \frac{1}{\kappa^2}\{[\eta_1 \nabla \eta_2 - \eta_2 \nabla \eta_1]$$

$$+ [\eta_3 \nabla \eta_4 - \eta_4 \nabla \eta_3] - \mathbf{A}S\} + f_{\mathbf{A}}. \tag{8}$$

Equation (8) is reduced to dimensionless form in the standard way:

$$\eta_j = \eta_j \left(\frac{\beta_1 + \beta_2 + \beta_3}{a} \right)^{1/2}, \quad \vartheta_{\mathbf{A}} = \frac{\Gamma_{\mathbf{A}} \Gamma^{-1}}{8\pi\alpha\xi^2} \sim 1,$$

$$b_2 = \frac{2\beta_2 + \beta_3}{\beta_1 + \beta_2 + \beta_3}, \quad b_3 = \frac{\beta_3}{\beta_1 + \beta_2 + \beta_3}, \quad \xi = \left(\frac{K_1}{a} \right)^{1/2}$$

so that κ in Eq. (8) is the Landau–Ginzburg parameter, $\kappa \gg 1$. Furthermore, in order to shorten the calculations we introduce the notation

$$S = \boldsymbol{\eta} \cdot \boldsymbol{\eta}^* \equiv |\eta_x|^2 + |\eta_y|^2 \equiv \sum \eta_j^2,$$

$$M = (\eta_x^* \eta_y - \eta_x \eta_y^*)/2i$$

$$\equiv i\mathbf{e}_z \cdot (\boldsymbol{\eta} \times \boldsymbol{\eta}^*)/2 \equiv [\eta_1 \eta_4 - \eta_2 \eta_3],$$

$$L = (\eta_x^* \eta_y + \eta_x \eta_y^*)/2 \equiv \eta_1 \eta_3 + \eta_2 \eta_4,$$

$$P = |\eta_x|^2 - |\eta_y|^2 \equiv \eta_1^2 + \eta_2^2 - \eta_3^2 - \eta_4^2. \tag{9a}$$

It is easy to prove that M and L are related by the expressions

$$\eta_x \eta_y^* = L - iM, \quad \eta_x^* \eta_y = L + iM. \tag{9b}$$

Moreover, there is an identity which will be important in what follows:

$$S^2 \equiv P^2 + (2M)^2 + (2L)^2. \tag{9c}$$

The derivatives of the Ginzburg–Landau functional Eq. (2) can also be written in terms of S, L, M, and P (see the Appendix).

The equations in (8) are solved by the numerical method used in Refs. 16, 34, 35, 40. In particular, the authors of these papers observed that superconductors of the first kind exhibit (kinetic) instability of a planar domain boundary separating a superconducting phase from a normal phase, whereas in superconductors of the second kind this boundary is stabilized by the absorption of vortex excitations.[34] They also discuss the problem of boundary conditions, which is especially important for large-scale numerical simulation of a relatively small volume (in our case 256×256 computational points). In particular, the boundary conditions for Eqs. (8) include the continuity of all components of the magnetic induction vector $\mathbf{B} = \nabla \times \mathbf{A}$ at the boundary of the superconductor, and a certain condition (not entirely uniquely defined) on the order parameter $\boldsymbol{\eta}$. For superconductors with a small coherence length we can limit ourselves to the simplest requirement $\boldsymbol{\eta}|_S = 0$ (see Ref. 36). Physically meaningful results are obtained when cyclic boundary conditions are used, which are also simple.

4.1. The phase $\boldsymbol{\eta} = (1, i)$

As mentioned above, the formation and evolution of topological singularities in all phases of a multicomponent superconductor look qualitatively very similar. Therefore, we illustrate the results of simulation and their analysis in detail only for the phase $\boldsymbol{\eta} = (1, i)$ (sector III), which is the sector of greatest practical interest for the system UPt$_3$, and then very briefly touch on distinctive features of the behavior in the phases of sectors I and II.

For UPt$_3$, experiment gives us the estimate $\beta_1 \sim \beta_2$.[27,29] In this case, in the absence of an external field a phase is realized with broken time-reversal invariance, such that the ground state of the superconductor is doubly degenerate with respect to directions of the vector $i\boldsymbol{\eta} \times \boldsymbol{\eta}^*$ (corresponding to degeneracy with respect to the phase difference $(\varphi_1 - \varphi_2) = \pm \pi/2$). It is significant that these states are separated by an energy barrier. Note that $i\boldsymbol{\eta} \times \boldsymbol{\eta}^*$ is proportional to the orbital angular momentum of a Cooper pair,[21] and the corresponding uniform superconducting state is magnetic.

A fragment of a system containing typical configurations of the order parameter modulus $\boldsymbol{\eta}$ and the magnetic field $\mathbf{h} = \nabla \times \mathbf{A}$ (in zero external magnetic field) at intermediate stages of its evolution, with initial conditions in the form of fluctuations at an arbitrary position, is shown in Fig. 2 (a and b respectively). Clearly visible are ''punctures'' and ''channels'' of width Δ on the surface $\boldsymbol{\eta}(x, y)$ with different field directions (''vortices'' and ''antivortices''). The emergence of domain walls, as predicted in Ref. 21, in which the order parameter is lower than its equilibrium value but still nonzero, constitutes a fundamental quantitative novelty of the ordering kinetics of superconductors with a multicomponent order parameter.

In the intermediate stages of the transition to superconductivity, domain walls together with vortex-like formations generate a complex, and sometimes extremely tangled, configuration. However, as the system approaches equilibrium (as $t \to \infty$), out of all sets of nonuniformities there ''survive'' only those domain walls attached to the boundaries (for instance, at grain boundaries). The remaining nonuniformities, vortex–antivortex pairs and closed domain walls, all collapse with time. Invoking the transparent analogy with magnetic bubbles, we can refer to regions of the superconductor within closed domain walls as ''superconducting bubbles.''

In the initial stages, a closed domain wall acquires a ringlike shape as it gradually contracts. Then it either transforms to the uniform state or generates a two-quantum vortex $\Phi = 2\Phi_0$, at the center of which all the components of the order parameter reduce to zero. In the latter case, the domain wall—a superconducting bubble—has a ''topological charge'' from the outset, due to the phase excursion of 2π generated by transporting each of the vectors $\boldsymbol{\phi}_1$ and $\boldsymbol{\phi}_2$ along its external boundary. The wall traps a magnetic flux $\Phi = 2\Phi_0$ nonuniformly distributed along it in previous stages, and gradually shapes it into an ''ordinary'' two-quantum axially asymmetric vortex. The typical initial topol-

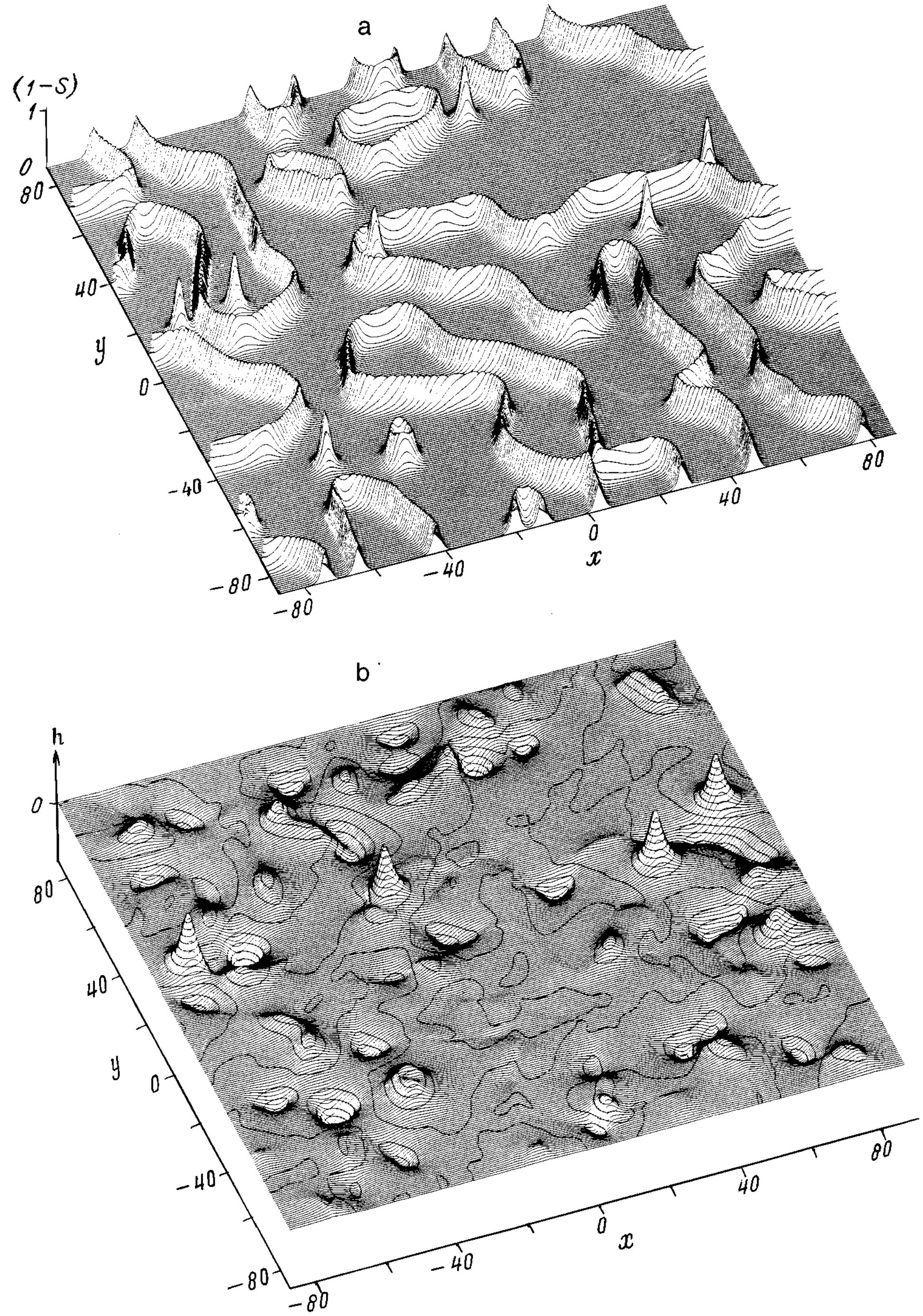

FIG. 2. The phase $\boldsymbol{\eta}=(1,i)$. A fragment of the system in the intermediate phase of its evolution for the following parameters of the Ginzburg–Landau functional: $b_2=1$, $b_3=-0.5, \kappa=10$: a—order parameter configuration $\boldsymbol{\eta}(x,y)$ (for clarity this is shown in the form of the function $(1-S)$, where $S=\boldsymbol{\eta}\cdot\boldsymbol{\eta}^*$); b—the same for the magnetic field **h**. The mutual correspondence between the distribution of $\mathbf{h}(x,y)$ and nonuniformities in the surface $(1-S)$ shown in Fig. 2a is evident. Two-quantum vortices are formed here as a result of compression of a domain wall with trapped magnetic flux.

ogy and evolution of a superconducting bubble as it becomes a two-quantum vortex is shown in Fig. 3.

The numerical observation of domain walls makes it desirable to predict them analytically, if only approximately. To do so, we must reduce the initial system of equations in such a way as to best take advantage of known results of numerical simulation, preserving their nontrivial properties as much as possible.

Above all, we note that a good approximation can already be obtained from a one-dimensional version of Eqs. (8), at least where the curvature of the domain wall is small. We can also verify that the solutions we want for η_j are only slightly distorted by the gauge field **A** $(\kappa \gg 1)$. Finally, numerical simulation shows that the rates of evolution of all the fluctuating fields of the system in the vicinity of domain walls are anomalously (logarithmically) slow compared to their rate of relaxation toward these domain walls. This is typical of attractor trajectories, which are the ''slowest'' path of the system to equilibrium. In turn, collapsing domain walls are so-called ''precursors'' for still lower-dimensional singular vortices (as are the singularities of superconducting bubbles that possess topological charge).

In light of all this, a sufficient description of slowly relaxing domain walls can be obtained by studying the following static reduced system for the field $\boldsymbol{\eta}$:

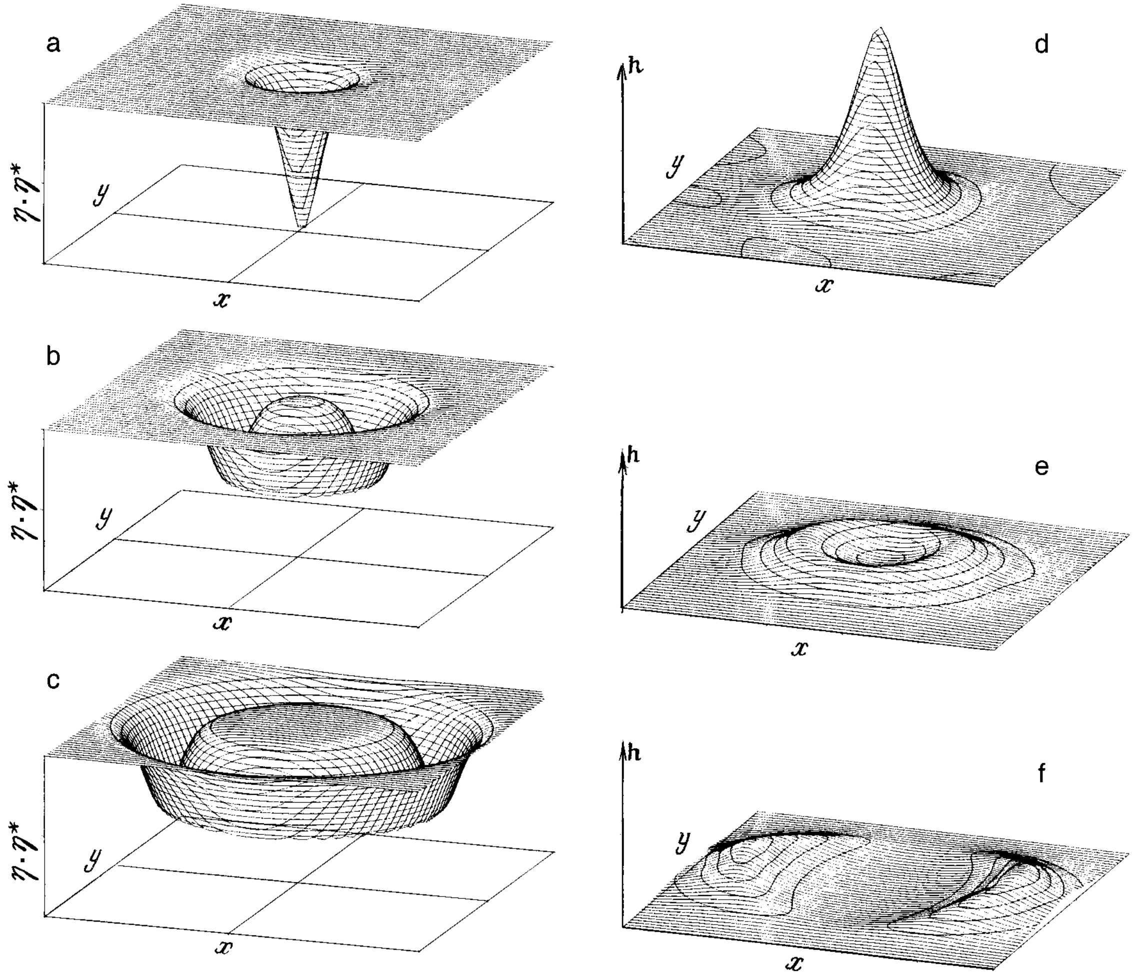

FIG. 3. Initial topology and evolution of a superconducting bubble as it becomes a two-quantum vortex.

$$\frac{d^2S}{dx^2}-2\left[\sum_j\left(\frac{d\eta_j}{dx}\right)^2\right]-2S(S-1)+4b_2M^2+4b_3L^2=0, \tag{10a}$$

$$\frac{d^2M}{dx^2}-2\left[\frac{d\eta_1}{dx}\frac{d\eta_4}{dx}-\frac{d\eta_2}{dx}\frac{d\eta_3}{dx}\right]-2M(S-1)+b_2SM=0, \tag{10b}$$

$$\frac{d^2L}{dx^2}-2\left[\frac{d\eta_1}{dx}\frac{d\eta_3}{dx}+\frac{d\eta_2}{dx}\frac{d\eta_4}{dx}\right]-2L(S-1)+b_3SL=0, \tag{10c}$$

$$\frac{d^2P}{dx^2}-2\left[\left(\frac{d\eta_1}{dx}\right)^2+\left(\frac{d\eta_2}{dx}\right)^2-\left(\frac{d\eta_3}{dx}\right)^2-\left(\frac{d\eta_4}{dx}\right)^2\right]-2P(S-1)=0, \tag{10d}$$

and then determining the field **h** using the results, as necessary. However, even in this form the system is still very complex. We solved it by using the method of verifiable smallness described recently in Ref. 13. In brief, this method essentially works as follows: we first solve the system (10) numerically, and then identify which terms of the equations (or their combinations) are small compared to the remaining terms over a wide range of parameter values.

In particular, we observed that in phase III, $L\rightarrow 0$ for $b_3<0$. Furthermore, the terms in Eqs. (10b) in square brackets are small, so that by neglecting them and using the identity (9c) we can reduce the system (10) to equations that contain only the pair of independent invariants S and M. Then, using various projections of the phase portrait of the system in the space (S, M, dS/dx, and dM/dx), we can establish approximate relations between these variables in the vicinity of the domain wall, after which we solve the remaining single equation.

After implementing this procedure we found the approximate relation $S(M)\simeq 1+b_2(2-b_2)M^2$ (in reality this expression is accurate to 10^{-2} against a background of unity!), for which the equation for M reduces to the simple form of the equations of Refs. 5, 7, 17:

$$\frac{d^2M}{dx^2}+b_2M-b_2(2-b_2)^2M^3=0. \tag{11}$$

With the boundary conditions $M\rightarrow -M_0\equiv -1/(2-b_2)$ as $x\rightarrow -\infty$, $M\rightarrow +M_0$ as $x\rightarrow +\infty$, we obtain the well-known solution:

$$M(x)=\eta_1\eta_4-\eta_2\eta_3=Q\tanh\frac{x}{\Delta}. \tag{12}$$

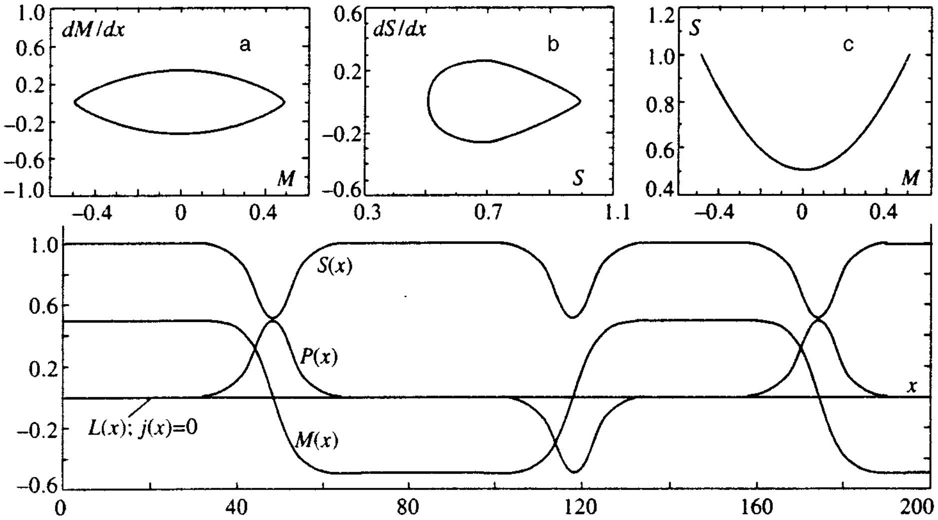

FIG. 4. The phase $\boldsymbol{\eta}=(1,i)$. Numerically evaluated distribution of the quantities $S(x)$, $M(x)$, $L(x)$, and $P(x)$ and current density $j(x)$ for $\xi=5$, $\kappa=10$, $b_2=1$, $b_3=-1$. Figures (a—c) show projections of the phase portrait onto the subspaces $(M,dM/dx)$, $(S,dS/dx)$, and (M,S), respectively. The analytic estimates of the nonuniformities (of order $\pm\Delta$ in the vicinity of each of them) are essentially indistinguishable from the numerical curves and are not shown on this figure.

Here

$$\Delta=\xi(2/b_2)^{1/2}=\xi\left(\frac{\beta_1+\beta_2}{\beta_2}\right)^{1/2} \tag{13}$$

is the width of the wall, and $Q=(M_{+\infty}-M_{-\infty})/2$ is its topological charge. Accordingly, we obtain an analytic approximation for $S(x)$ in the form of a so-called ''dark soliton'':

$$S(x)=\Sigma\,\eta_i^2=1+b_2\left(\tanh\frac{x}{\Delta}\right)^2\Big/(2-b_2). \tag{14}$$

Substituting Eqs. (12), (14) into Eq. (9c), we easily find the explicit dependence $P(x)$:

$$P(x)=\pm\left\{\left(1-\left(\tanh\frac{x}{\Delta}\right)^2\right)\left[(2-b_2)^2 - b_2^2\left(\tanh\frac{x}{\Delta}\right)^2\right]\right\}^{1/2}\Big/(2-b_2) \tag{15}$$

i.e., a ''soliton'' that vanishes at infinity. It follows from Eq. (15) that for any $b_2\leqslant 1$ the solutions describe a distribution of physical quantities at arbitrary distances from the center of the wall, whereas for $b_2>1$ the solution is applicable only at distances from the wall center of order $x\leqslant\Delta\times\operatorname{arccoth}\{(2-b_2)/b_2\}$.

The discussion above applies whenever $b_3<0$, whereupon $L\to 0$ in the superconductor. If we choose $b_3>0$ (ensuring, of course, that we stay in sector III as before, i.e., $b_2>b_3>0$), it turns out that $P(x)\to 0$ within the wall, while now it is $L(x)$, that is nonzero. In this case, in order to obtain the solution $L(x)$ it is once more sufficient to use the identity (9c), which leads to the replacement $P\to 2L$ in Eq. (15).

These results are summarized in Fig. 4, from which it is clear that S has essentially its equilibrium value throughout most of the bulk. Against this background there is a certain region of width Δ where the amplitude S decreases and varies within the limits $S_0>S>\mathrm{const}\neq 0$. In those places where the modulus of the order parameter has a minimum, M changes sign and reduces to a constant, equal in order of magnitude to the equilibrium $\pm M_0$ over the same distances ($\propto\Delta$) on both sides of the minimum, whereas P differs from zero only at the center of the wall, and vanishes over the same distances.

Projections of the phase portrait are shown in Figs. 4a, 4b, and 4c for the subspaces $(S,dS/dx)$, $(M,dM/dx)$, and (S,M) respectively. These projections illustrate the connection between these quantities, which we alluded to while obtaining the analytic estimates Eqs. (12)–(15). The fact that disagreement between these estimates and the numerical results (for b_2 in the range from 0 to 1.85) does not increase the noise level argues in favor of the approximations we have made.

The additional phase shift $(\varphi_1-\varphi_2)=\theta(x)$ in the soliton Eq. (12) reaches $\pm\pi$ and is determined by the relation

$$\sin(\theta(x))\left(1-b_2+b_2\left(\tanh\frac{x}{\Delta}\right)^2\right)^{1/2}=\tanh\frac{x}{\Delta}, \tag{16}$$

which is correct for $b_2\leqslant 1$. In the limit $b_2\to 0$ it reduces to the expression $\sin(\theta(x))=\tanh(x/\Delta)$, which coincides with the corresponding expression in Ref. 21 obtained in the limit of large anisotropy $\beta_3\to\infty$. In the other very important limit $b_2\to 1$, a jump in phase by $\pm\pi$ takes place at the center of the domain wall $\sin(\theta(x))=\mathrm{sign}(\tanh(x/\Delta))$; this also happens for $1<b_2<2$. We can directly confirm that the total electric current transverse to the wall vanishes.

Thus, this wall lies between two equilibrium states with different signs of M (the direction of magnetization), states that are separated by an energy barrier from one another; therefore it is a stable topological formation. The existence of the domain wall implies the following finite positive correction to the free energy of the superconductor:

$$\Delta F=\int(F(x)-F_0)\,dV=\sigma S. \tag{17}$$

Here S is the area of the wall (yz), and σ is the energy per unit area expended to create it (without including the magnetic field energy),

$$\sigma=H_c^2 b_2(3-b_2)\Delta/12\pi(2-b_2)>0, \tag{18}$$

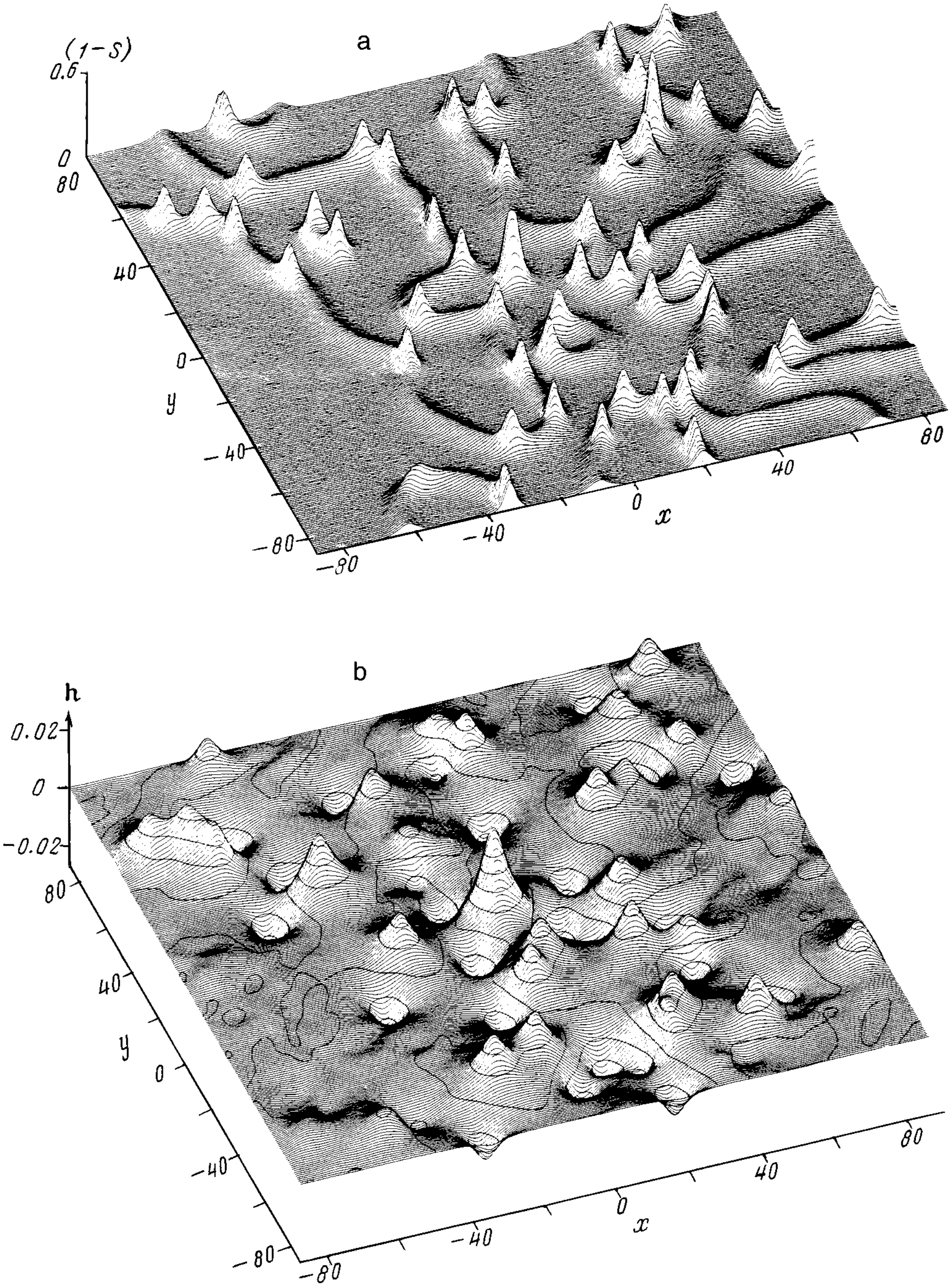

FIG. 5. The same as in Fig. 2 for low surface energy. The parameters of the Ginzburg–Landau functional are $b_2=0.1$ (in Fig. 2, $b_2=1$), $b_3=-0.5$, $\kappa=10$: a—typical configuration of the quantity $(1-S)\equiv 1-\boldsymbol{\eta}\cdot\boldsymbol{\eta}^*$, b—magnetic field **h**.

where H_c is the critical thermodynamic field in the phase $\boldsymbol{\eta}=(1,i)$, and

$$H_c^2=8\pi a^2/(2-b_2)(\beta_1+\beta_2+\beta_3). \qquad (19)$$

Since $\sigma>0$, the state of the superconductor is metastable and can exist only if the domain wall is pinned at the boundaries. We also cannot rule out states in which the entire sample is divided up into a band-like domain structure, a configuration that can lead to lowering of the total energy (wall energy+demagnetization energy of the magnetic domains). The surface energy of the walls depends monotonically on the parameter $b_2=(2\beta_2+\beta_3)/(\beta_1+\beta_2+\beta_3)$. As $b_2\rightarrow 0$, the energy tends to zero, $\sigma\rightarrow 0$, while as b_2 increases the energy increases.

As $b_2\rightarrow 2$, i.e., near the stability boundary (see Fig. 1) the energy of the wall formally tends to infinity, $\sigma\rightarrow\infty$. In this limit the estimates (12)–(18) are no longer valid. Nevertheless, an increase in σ near the stability boundary is in itself quite typical of the phenomenology of fluctuation-induced weak first-order phase transitions.[13] In this region the system must expend considerable energy to form extended domain walls; in the final analysis, this is also the reason for threshold-like seed formation.

The qualitative picture of the formation and evolution of topological singularities for large σ ($b_2\rightarrow 2$) differs only slightly from the case discussed already ($b_2=1$). New and interesting details of the behavior of the system appear only at low wall energies ($b_2\rightarrow 0$); see Sec. 5. In particular, it turns out that localized one-quantum nonsingular vortices are now present in the system (easily distinguishable punctures in the order parameter), arranged in chains within the wall. In some cases we encounter unclosed domain walls at whose end a vortex is found. For higher well energies, analogous nonsingular vortices within a domain wall are easily distinguished only in the pattern of magnetic field, whereas punctures in the order parameter are essentially invisible (Figs. 2 and 3).

In the outer regions of such a vortex only two out of the four components vanish, and at its center the modulus of the order parameter is approximately $S_0/2$. However, the order parameter in the wall itself determined by Eq. (14), can be considerably smaller at higher energies: $S(x=0) \ll S_0/2(=1/(2-b))$. Therefore, in the latter case the shape of a vortex is to a large extent determined by the function $S(x)$, and as a result the vortex seemingly merges with the wall. It is clear that an arbitrary choice of parameters of the functional could lead to noncircular one-quantum vortices in the wall, i.e., the vortices tend to elongate along the wall as the surface energy increases.

Yet another interesting detail of the behavior of the superconductor for small σ is the anomalously slow relaxation of the configuration observed in numerical experiments when a single closed wall contains a chain of one-quantum vortices with m vortices and n antivortices, $m \neq n$. In experiments, for example, it is not uncommon to encounter cases where $m \gg n$ or conversely, although the most typical situations are those described above, i.e., either $m=n$—a superconducting bubble without topological charge relaxing to the uniform state—or $m=2$, $n=0$ (or conversely)—a superconducting bubble with topological charge relaxing to an axially asymmetric two-quantum vortex (antivortex). Note, however, that for small values of σ, vortices located in the bulk of the superconductor are significantly transformed—they degenerate into two closely spaced $\sim\Delta$ single-quantum nonsingular vortices (Fig. 6a). The distribution of magnetic field of this configuration is quite close to circular (Fig. 6b). The case $m \gg n$ is closely associated with the problem of local stability of a superconducting bubble and will be discussed below. We emphasize that the magnitude of the total magnetic flux through the surface of the sample was monitored throughout the numerical experiment and remained unchanged, i.e., $\Phi=0$. In other words, the total topological charge of the entire sample equals zero.

At small σ, the pattern obtained from the numerical solutions recalls the experimental situation in rotating He^3,[10,12] where so-called vortex layers with vortex chains are observed. This similarity is by no means accidental, but rather is due both to the common origin (spontaneous symmetry breaking) and similarity of the order parameters of He^3 and multicomponent superconductors. In both cases, the order parameter in the domain wall is already decreasing, and as a result it is energetically more favorable for the system to locate topological defects that possess a mantle (vortex) not in the bulk (of the superconductor, for example) but within or along a domain wall.

Interestingly enough, experimental visualizations of the magnetic field in high-temperature superconductors[37–41] also reveal chains of one-quantum vortices,[39–41] since a functional of type Eq. (2) was proposed for the high-temperature superconductor system Y–Ba–Cu–O as far back as Ref. 24, based on analysis of its possible superconducting classes. Most contemporary experimental data, however, suggest that the symmetry of the order parameter in high-temperature superconductors is probably not lower than the symmetry of the crystal lattice, and that time reversal is not a broken symmetry. Nevertheless, the problem of spatial anisotropy of the order parameter (i.e., $\Delta=\Delta(k)$) remains open.[42] There is an abundance of evidence both against d-pairing and in favor of it, and even in favor of $(s+d)$-pairing.[42,26] In the latter two cases the effective order parameter turns out to be significantly multicomponent, which leads to a functional of Ginzburg–Landau type as in Eq. (2).[26] At this time the situation remains in flux and has not been finally resolved.

In any case, by comparing our numerical results with experiment,[39–41] we have confirmed that if high-temperature superconductors have anisotropic pairing, the parameters $b_2=(2\beta_2+\beta_3)/(\beta_1+\beta_2+\beta_3)$, $b_3=\beta_3/(\beta_1+\beta_2+\beta_3)$ of the Ginzburg–Landau functional should be small, $|b_2|$, $|b_3| \ll 1$. However, we also note that in the traditional theory (s-pairing), the explanation for the existence of vortex chains in high-temperature superconductors is a very complicated problem that requires additional assumptions to justify it.[43] The multicomponent order parameter generates these structures automatically.

4.2. The phase $\boldsymbol{\eta}=(1,1)$

In sector I a phase $\boldsymbol{\eta}=(1,1)$ is realized in Fig. 1 with a phase difference $(\varphi_1-\varphi_2)=0,\pi$. Here the ground state is doubly degenerate with respect to the sign of the quantity $L=\boldsymbol{\phi}_1\cdot\boldsymbol{\phi}_2=(\eta_1\eta_3+\eta_2\eta_4)$. In this case the states are also separated by an energy barrier. For the one- dimensional case with the corresponding substitution $b_2\rightarrow b_3$, $M\rightarrow L$ we have solutions whose form coincides completely with that given in Eqs. (12)–(18). The wall separates phases in which the vectors $\boldsymbol{\phi}_1$ and $\boldsymbol{\phi}_2$ are parallel and antiparallel.

The formation kinetics of vortex structures in a phase of sector I do not differ significantly from the picture in sector III. The results of numerical simulation in the later stages of evolution are qualitatively the same as in Figs. 3, 4, 5, and 6.

4.3. The phase $\boldsymbol{\eta}=(1,0)$ (or $\boldsymbol{\eta}=(0,1)$)

In sector II the phases $\boldsymbol{\eta}=(1,0)$ or $\boldsymbol{\eta}=(0,1)$ are realized. These states are also separated by an energy barrier, and are degenerate with respect to the sign of the quantity $P=\boldsymbol{\phi}_1^2-\boldsymbol{\phi}_2^2$. Here we also have a domain wall and singular and nonsingular vortices. The distribution of the quantities $S(x)$, $P(x)$, $M(x)$, and $L(x)$ in phase II for the one-dimensional case are qualitatively the same as for the other phases. Namely, the solution for the order parameter $P(x)$ is a kink, and for the order parameter $S(x)$ a dark soliton. In an interphase wall, depending on the ratio of the parameters of the functional, either $M\neq 0$ for $\beta_3<2\beta_2+\beta_3$ or $L\neq 0$ for $\beta_3>2\beta_2+\beta_3$. Equations (12)–(18) also maintain their validity, and we need only take into account the correct nor-

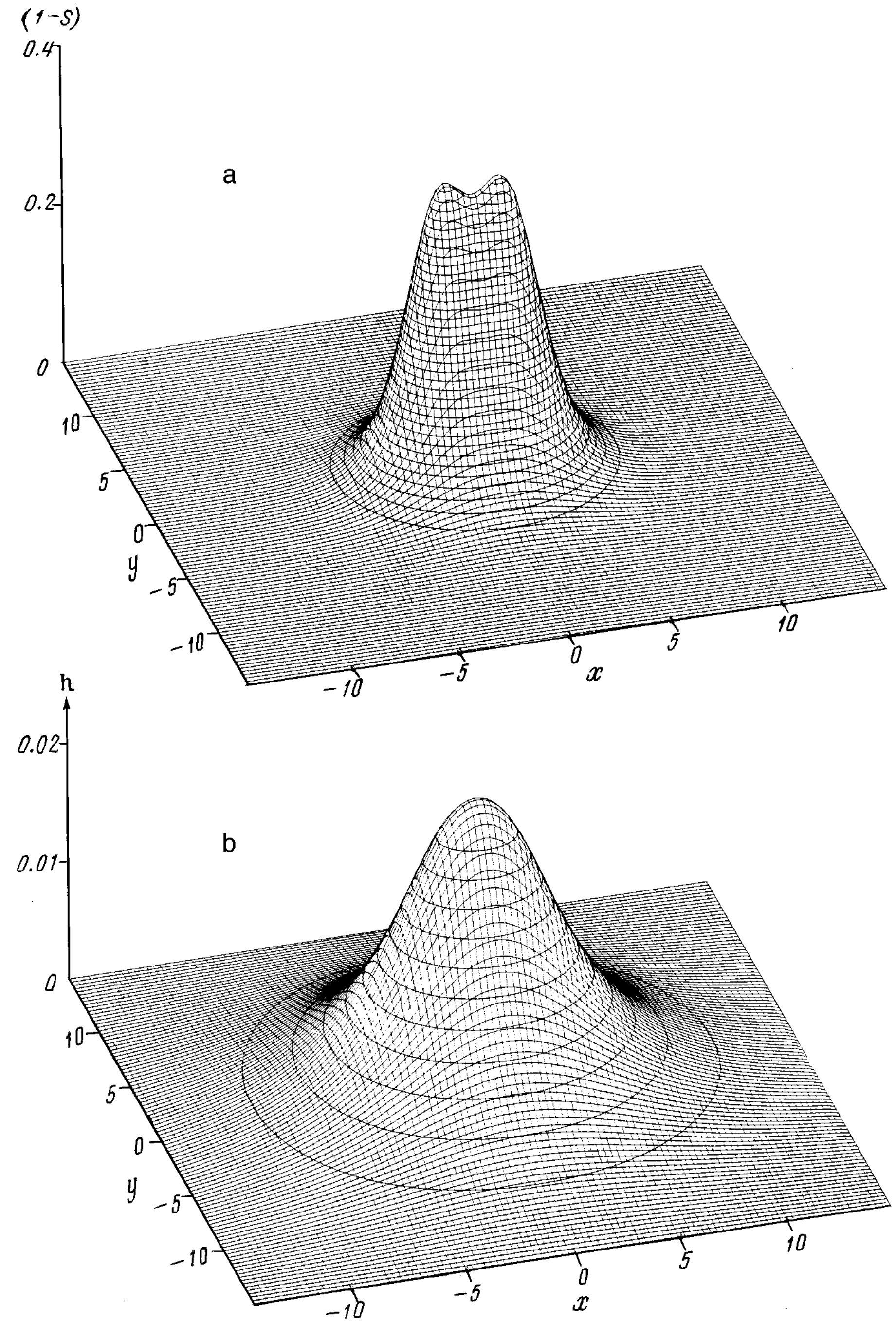

FIG. 6. Results for the evolution of a superconducting bubble with low surface energy. The two-quantum vortex splits.

malization of S_0 in this phase and replace Δ by the quantity $\Delta = \xi((2-b)/-b)^{1/2}$, where $(-b) = \min\{|b_2|;|b_3|\}$.

5. LOCAL STABILITY OF BUBBLES

In all phases, we encounter anomalously long-lived closed walls with a chain of similar vortices at low energy σ (the configuration mentioned above when $m \gg n$)—i.e., rigid superconducting bubbles. In ordinary superconductors, the energy per unit length of a single-quantum vortex is $\varepsilon_1 = (\Phi_0/4\pi\lambda)^2 \ln(\kappa+0.08)$. For the energy of a hypothetical n-quantum vortex we obtain $\varepsilon_n \approx n^2\varepsilon_1$, which is n times the total energy of n noninteracting single-quantum vortices, so that the formation of multiquantum vortices is energetically unfavorable; moreover, there are no forces that could do the work required to form them.

In multicomponent superconductors the situation is otherwise. It is energetically favorable for one-quantum vortices to be located in walls. At the same time, the system tends to shrink the wall length (the energy of the wall is positive; see Eq. (18)). Hence, we are dealing here with an interesting concrete realization of the topological confinement of objects with differing spatial dimensions mentioned in the Introduction. This type of confinement is also encountered in superfluid He^3 (Ref. 20) and in several liquid crystals. The stability of configurations of the ‘‘vortex+wall’’ type is caused by the same factors as in field theory,[5–8] and strongly recalls the situation with quarks, which are not observed individually (this analogy was pointed out in Ref. 1 in connection with He^3). When vortices are separated, their energy of separation increases proportional to the distance between them, as in the separation of quarks. A domain wall in this case plays the role of a ''gluon string,'' which in our case joins single-quantum vortices. If necessary, the analogy can be extended much farther, e.g., n-quantum vortices are hadrons, the magnetic field gives photons, etc. (see Ref. 1).

Let us discuss, for example, a configuration consisting of a closed domain wall with n similar vortices. An analogous structure for rotating superfluid He^3 was proposed in Ref. 10 and discussed in detail in Ref. 12. In our case the wall energy σ competes with the magnetic repulsion of the vortices. If $\sigma \to \infty$, then the configuration described tends to contract to an ideal n-quantum point vortex. For finite values of the surface energy σ, the configuration of superconducting bubbles is frozen, and if σ is very low, the equilibrium diameter of a bubble made up of n similar vortices can turn out to be quite large.

To estimate this diameter, consider a simple model with symmetrically placed vortices on a ring-like domain wall of radius $R \gg \xi$. The energy of this configuration can be written in the form

$$U(R)=\frac{1}{2}\sum_{i\neq j}^{N}\frac{\Phi_0^2}{8\pi^2\lambda^2}\ln\frac{\lambda}{r_{ij}}+2\pi R\sigma, \tag{20}$$

where the first term is the magnetic energy due to the repulsive interaction of one-quantum vortices at a distance of r_{ij} between them (for simplicity $\xi < r_{ij} < \lambda$), while $2\pi R\sigma$ is the surface tension energy of the wall. The curvature of the wall is considered small, in which case the contribution to the energy from the curvature is logarithmically small and can be neglected, as shown in Ref. 12. The condition of equality of the forces $F = dU/dR = 0$ is satisfied for

$$R_c = n(n-1)\Phi_0^2/32\pi^3\lambda^2\sigma. \tag{21}$$

Yet another characteristic length scale appears in this system, which must be compared with ξ and λ. Specifically, in order that the value of the order parameter at the center of the domain be close to equilibrium, the distance between vortices in the domain wall should be $l = 2R_c \sin(\pi/n) \gg 2\xi$. Using the explicit form of σ from Eq. (18) (in phase III for definiteness) and the fact that in order of magnitude we have $\Phi_0 \sim 2^{1/2}\lambda\xi H_c$ (in classical superconductors this ratio is an identity), the condition $R_c \sin(\pi/n) \gg \xi$ can be rewritten in expanded form:

$$3n(n-1)(2-b_2)\sin(\pi/n)/(3-b_2)(2b_2)^{1/2} \gg 1. \tag{22}$$

The inequality (22) is satisfied for essentially any $b_2 \ll 1$ and $n > 2$. Numerical simulation shows that for the parameters of our system, ''frozen'' multiquantum superconducting bubbles of large size can actually exist; see Fig. 7. For $n = 2$ the curvature of the domain wall cannot be neglected and Eq. (22) loses its validity.

As σ increases, the superconducting bubbles, for example with $n = 4$, degenerate into the configuration shown in Fig. 8. At the center of this ''spot,'' the value of the order parameter is less than its equilibrium value, $S \ll S_0$, but is not equal to zero. The order parameter distribution has well-defined fine structure with characteristic size $\sim \xi$ and concomitant fine magnetic field structure. Note that at large distances from the center of this n-quantum vortex ''spot,'' the magnetic field created by the vortices becomes isotropic, and should show up in experiments as a single washed out n-quantum vortex.

From this section it follows that ideal axially symmetric vortices in whose core the order parameter equals zero do not exist even when $\sigma \to \infty$ and the gradient parts in Eq. (2) are isotropic, i.e., K_2, $K_3 \ll K_1$. Fluctuations of the order parameter for $T \neq 0$ split up the multiquantum vortex, leading to spatial anisotropy of its superconducting and magnetic properties. A lattice made up of such vortices in an external field will have other than hexagonal symmetry. A detailed study of this question is beyond the scope of this paper. If the fluctuations of the order parameter are sufficient to overwhelm the energy barrier that confines the vortex in the wall (for example, in the immediate vicinity of T_c), then a $2n$-quantum vortex should dissociate into n two-quantum vortices. In any case, there is no topological hindrance to this process.

6. CONCLUSION

A phase transition to superconductivity is accompanied by the formation of an intermediate vortex-like state (a mesoscopic magnetic field with $\langle \mathbf{h} \rangle = 0$, $\langle \mathbf{h}^2 \rangle \neq 0$). In multicomponent superconductors, the nonuniform intermediate state can contain various topological singularities of the order parameter. Our numerical experiments have established that the most typical singularities are domain walls, multiquantum and single-quantum vortices (and antivortices), and superconducting bubbles, including rigid bubbles.

A multiquantum vortex is formed by the collapse of a rigid bubble with the corresponding topological charge, which in turn arise spontaneously from amorphous fluctuations at an arbitrary location. For $T \neq 0$, fluctuations in the order parameter split these vortices into n-quantum vortex spots.

When certain relations hold between the parameters of the Ginzburg–Landau functional, some of the topological singularities mentioned above, i.e., domain walls attached to the boundary of the sample and rigid bubbles, are anomalously long-lived in multicomponent superconductors. The ground state of a multicomponent superconductor in zero external magnetic field is metastable and nonuniform, and its order parameter, together with the associated magnetic properties, exhibits domain structure. The latter circumstance can affect the magnetic characteristics of the superconductor in a serious way when a magnetic field is switched on.

The nonlinear excitations described above play an important role in the processes leading to fluctuation-induced first-order transitions in these systems, so that their dynamic behavior qualitatively recalls the phenomenology of fluctuation-induced processes in anisotropic systems described in Ref. 13.

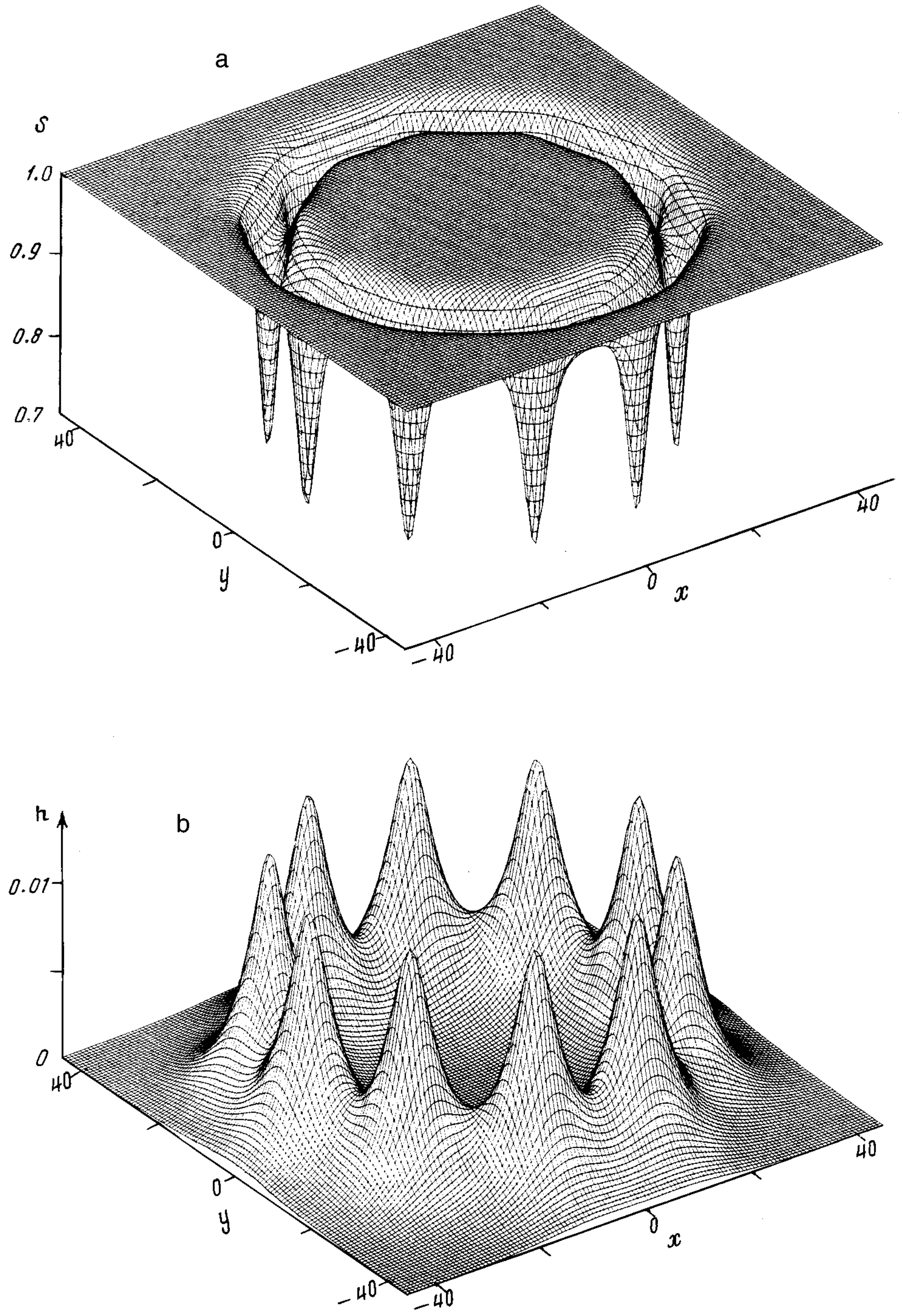

FIG. 7. Stable superconducting bubble with $n=10$ in the phase $\boldsymbol{\eta}=(1,i)$: a—distribution of order parameter $S=\boldsymbol{\eta}\cdot\boldsymbol{\eta}^*$, b—magnetic field.

Finally, we mention that these results may be of particular interest both to the theory of He^3 and to gauge field theory, since the model we have investigated can be regarded as statistically analogous to the corresponding field theory models of Refs. 5–8.

A number of papers published in recent months,[44–50] have pointed out the great relevance of investigations of vortex states in superconductors with d-pairing.

One of the authors (A.F.) is grateful to V. M. Loktev, whose interest in the problem and deep questions have had an important influence on our understanding of the nature of these processes. A.F. is also grateful to H. M. Plakida, who alerted us to the newest results in the area of high-temperature superconductivity, which has stimulated our return to this theme in the context of d-pairing. Furthermore, A.F. is grateful to A. Geim for the hospitality of the University of Nijmegen (the Netherlands), whose work with A.F. and valuable discussions enabled us to improve our understanding of the kinetics of the superconducting transition.

We are also grateful to Alan T. Dorsey for generously providing us with several prepublication results of a study of the dynamics of interphase boundaries in superconducting systems.

This work was supported by the International Science

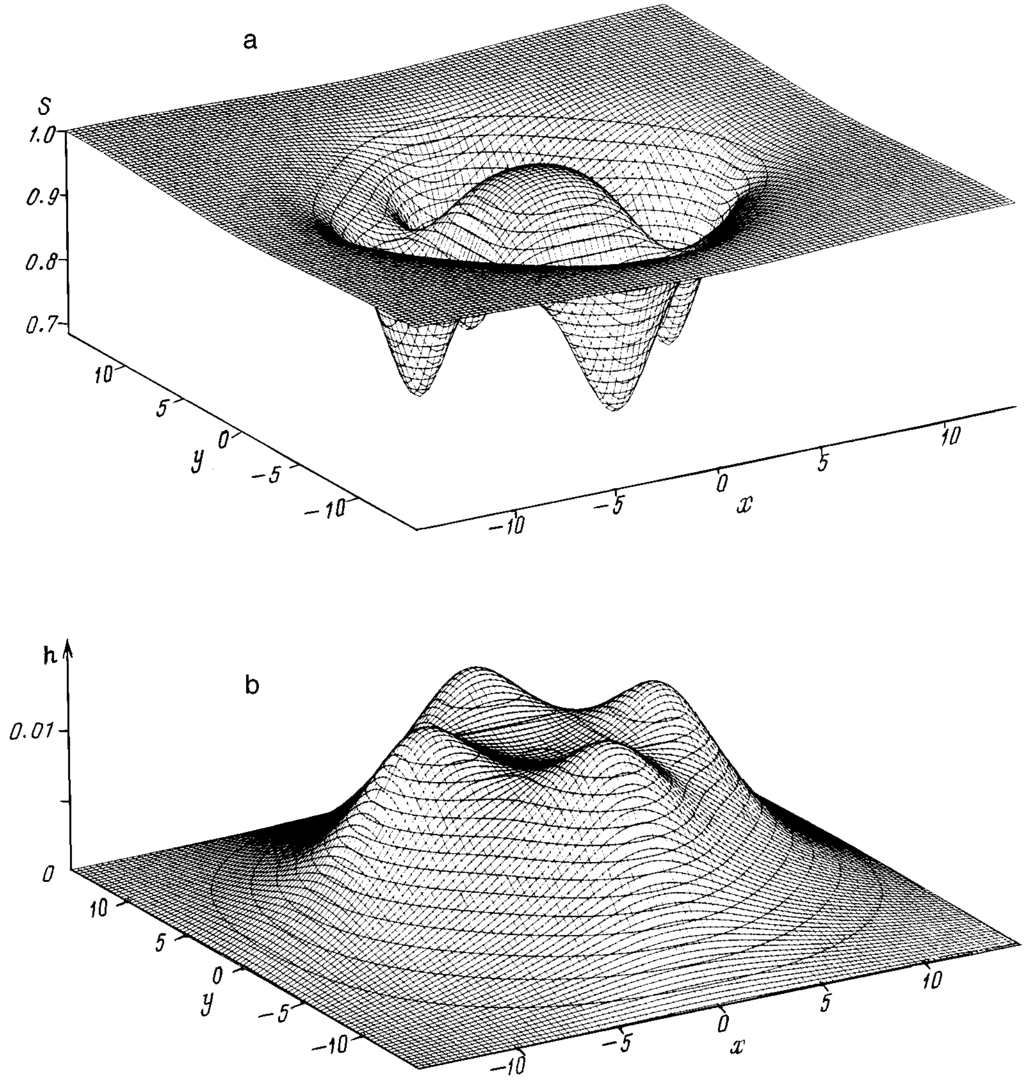

FIG. 8. n-vortex spot formed by the collapse of an n-quantum superconducting bubble in the phase $\boldsymbol{\eta}=(1,i)$ for $n=4$: a—distribution of order parameter $S=\boldsymbol{\eta}\cdot\boldsymbol{\eta}^*$, b—magnetic field.

Foundation (Grant K58100) and also by Competitive Projects No. 91011 and No. 9.01.01 of the State Committee on Science and Engineering of the Ukraine.

APPENDIX

Ginzburg–Landau functional in effective variables

Using the new variables Eq. (9), the Ginzburg–Landau functional Eq. (2) can be completely rewritten in these variables. For a more symmetric description we introduce the vector $\mathbf{S}=(S_1,S_2,S_3)$, whose components are respectively $S_1=P$, $S_2=2M$, $S_3=2L$. We have

$$\mathscr{F}=\int dV\left\{\left[\sum_j (\nabla S_j)^2/4+\mathbf{J}_s^2\right]\Big/|\mathbf{S}|-|\mathbf{S}|+\mathbf{S}^2/2 -b_2S_2^2/4-b_3S_3^2/4+\chi(\nabla\times\mathbf{A})^2\right\}. \tag{A1}$$

Of course, the number of variables in the functional (A1) remains the same as in Eq. (2), since information abort the phase of the superconducting order parameter $\boldsymbol{\eta}$ is contained in the term $\mathbf{J}_s$, which has the form

$$\mathbf{J}_s=[\eta_1\nabla\eta_2-\eta_2\nabla\eta_1]+[\eta_3\nabla\eta_4-\eta_4\nabla\eta_3]-\mathbf{A}|\mathbf{S}|, \tag{A2}$$

i.e., it coincides exactly with the expression for the superconducting current. We note that functionals of this type for $\mathbf{J}_s=0$ and $|\mathbf{S}|=\text{const}$ are encountered in the physics of magnetic phenomena when investigating the statics and dynamics of domain boundaries.[17]

[1] G. E. Volovik, in *Nonlinear Waves. Dynamics and Evolution*, A. V. Gaponov-Grekhov (ed.) [in Russian], Nauka, Moscow (1989).
[2] M. Hindmarsh, Phys. Rev. Lett. **75**, 2502 (1995).
[3] I. Chuang, R. Durrer, N. Turok, and B. Yurke, Science **251**, 1336 (1991).
[4] P. C. Hendry, N. S. Lawson, R. A. M. Lee, P. V. E. McClintock, and C. D. H. Williams, Nature (London) **368**, 315 (1994).
[5] T.-P. Cheng and L.-F. Li, *Gauge Theory of Elementary Particle Physics*, Oxford Univ. Press, Oxford (1984).
[6] A. S. Shvarts, *Quantum Field Theory and Homology* [in Russian], Nauka, Moscow (1989).
[7] R. Rajaraman, *An Introduction to Solitons and Instantons in Quantum Field Theory*, North-Holland, Amsterdam (1984).
[8] A. Vilenkin, Phys. Rep. **121**, 263 (1985).
[9] M. Bowick, L. Chandar, E. A. Schiff, and A. M. Srivastava, Science **263**, 943 (1994).
[10] U. Parts, E. V. Thyneberg, G. E. Volovik, J. H. Koivuniemi, V. M. H.

Ruutu, M. Heinila, J. M. Karimaki, and M. Krusius, Phys. Rev. Lett. **72**, 3839 (1994).

[11] T. Sh. Misiripashev and G. E. Volovik, Physica B **210**, 338 (1995).

[12] M. Heinila and G. E. Volovik, Physica B **210**, 300 (1995).

[13] A. S. Zel’tser and A. É. Filippov, Zh. Éksp. Teor. Fiz. **106**, 1117 (1994) [JETP **79**, 605 (1994)].

[14] A. É. Filippov, A. V. Radievskiĭ, and A. S. Zel’tser, Phys. Lett. A **192**, 131 (1994).

[15] A. É. Filippov, A. V. Radievskiĭ, and A. S. Zel’tser, Fiz. Nizk. Temp. **20**, 1017 (1994) [Low Temp. Phys. **20**, 801 (1994)].

[16] R. Kato, Y. Enomoto, and S. Maekawa, Phys. Rev. B **47**, 8016 (1993).

[17] R. Dodd, J. Eilbeck, J. Gibbon, and H. Morris, *Solitons and Nonlinear Wave Equations*, Academic Press, London (1982).

[18] G. E. Volovik and V. P. Mineev, Zh. Éksp. Teor. Fiz. **72**, 2256 (1977) [Sov. Phys. JETP **45**, 1186 (1977)].

[19] G. E. Volovik and V. P. Mineev, Zh. Éksp. Teor. Fiz. **73**, 767 (1977) [Sov. Phys. JETP **46**, 401 (1977)].

[20] Y. Kondo, J. S. Korhonen, M. Krusius, V. V. Dmitriev, E. V. Thuneberg, and G. E. Volovik, Phys. Rev. Lett. **68**, 3331 (1992).

[21] G. E. Volovik and L. P. Gor’kov, Zh. Éksp. Teor. Fiz. **88**, 1412 (1985) [Sov. Phys. JETP **61**, 843 (1985)].

[22] M. E. Zhitomirskiĭ, Zh. Éksp. Teor. Fiz. **97**, 1346 (1990) [Sov. Phys. JETP **70**, 760 (1990)].

[23] P. Brusov, N. Brusova, and P. Brusov, Czech. J. Phys. **46**, Suppl. 52, 1041 (1996).

[24] G. E. Volovik, Physica Scripta **38**, 502 (1988).

[25] M. E. Zhitomirskiĭ and I. A. Luk’yanchuk, Zh. Éksp. Teor. Fiz. **101**, 1954 (1992) [Sov. Phys. JETP **74**, 1046 (1992)].

[26] Yu. M. Ivanchenko, A. A. Lisyanskiĭ, and A. É. Filippov, Fiz. Tverd. Tela (Leningrad) **31**, 51 (1989) [Sov. Phys. Solid State **31**, 1681 (1989)]; Fiz. Tverd. Tela (Leningrad) **31**, 204 (1989) [Sov. Phys. Solid State **31**, 1767 (1989)].

[27] V. P. Mineev and K. V. Samokhin, Zh. Éksp. Teor. Fiz. **105**, 747 (1994) [JETP **78**, 401 (1994)].

[28] M. Palumbo, P. Muzikar, and J. A. Sauls, Phys. Rev. B **42**, 2681 (1990).

[29] A. S. Mel’nikov, Zh. Éksp. Teor. Fiz. **101**, 1978 (1992) [Sov. Phys. JETP **74**, 1059 (1992)].

[30] Yu. M. Ivanchenko, A. É. Filippov, and A. V. Radievskiĭ, Sverkhprokhod. **4**, 494 (1991).

[31] R. Joint, V. P. Mineev, G. E. Volovik, and M. E. Zhitomirsky, Phys. Rev. B **42**, 2014 (1990).

[32] M. Palumbo, C. H. Choi, and P. Muzikar, Physica B **165/166**, 1095 (1990); Physica B **165/166**, 2681 (1990).

[33] Yu. M. Ivanchenko, A. É. Filippov, and A. V. Radievskiĭ, Fiz. Nizk. Temp. **18**, 957 (1992) [Sov. J. Low Temp. Phys. **18**, 672 (1992)].

[34] H. Frahm, S. Ullah, and A. T. Dorsey, Phys. Rev. Lett. **66**, 3067 (1991).

[35] J. C. Osborn and A. T. Dorsey, Phys. Rev. B **50**, 15961 (1994).

[36] L. N. Bulaevskiĭ, V. L. Ginzburg, and A. A. Sobyanin, Zh. Éksp. Teor. Fiz. **94**, 335 (1988) [Sov. Phys. JETP **67**, 1499 (1988)].

[37] *Physical Properties of High-Temperature Superconductors*, D. M. Ginsberg (ed.), World Scientific, Singapore (1989).

[38] M. Leghissa, L. A. Gurevich, M. Kraus, G. Saemann-Ischenko, and L. Ya. Vinnikov, Phys. Rev. B **48**, 1341 (1993).

[39] C. A. Bolle, P. L. Gammel, D. G. Grier, C. A. Murray, D. J. Bishop, D. B. Mitzi, and A. Kapitulnik, Phys. Rev. Lett. **66**, 112 (1991).

[40] P. L. Gammel, D. J. Bishop, J. P. Rice, and D. M. Ginsberg, Phys. Rev. Lett. **68**, 3343 (1992).

[41] P. L. Gammel, C. A. Duran, D. J. Bishop, V. K. Kogan, M. Ledvij, A. Yu. Simonov, J. P. Rice, and D. M. Ginsberg, Phys. Rev. Lett. **69**, 3808 (1992).

[42] V. M. Loktev, Fiz. Nizk. Temp. **22**, 3 (1996) [Low Temp. Phys. **22**, 1 (1996)].

[43] A. Yu. Martynovich, Zh. Éksp. Teor. Fiz. **105**, 912 (1994) [JETP **78**, 489 (1994)].

[44] H. Wona and K. Maki, Phys. Rev. B **53**, 5927 (1996).

[45] M. Franz *et al.*, Phys. Rev. B **53**, 5795 (1996).

[46] M. Ichioka *et al.*, Phys. Rev. B **53**, 15317 (1996).

[47] M. Franz, C. Kallin, and A. J. Berlinsky, Phys. Rev. B **54**, R6897 (1996).

[48] R. Heeb *et al.*, Phys. Rev. B **54**, 9385 (1996).

[49] Q. Wang and Z. D. Wang, Phys. Rev. B **54**, R15643 (1996).

[50] J. J. V. Alvarez, G. C. Buscaglia, and C. A. Balsiero, Phys. Rev. B **54**, R16168 (1996).

Translated by Frank J. Crowne

Dynamics of the intermediate state and domain walls in an external oscillating magnetic field

V. S. Gerasimchuk

Donbas Engineering and Construction Academy, 339023 Makeevka, Ukraine

A. L. Sukstanskiĭ*)

Donetsk Physicotechnical Institute, Ukrainian National Academy of Sciences, 340114 Donetsk, Ukraine

(Submitted 29 October 1996)
Zh. Éksp. Teor. Fiz. **112**, 1374–1385 (October 1997)

We study the dynamics of 90-degree domain walls in the intermediate state of antiferromagnets, the state being realized in a first-order spin-flop transition in an external magnetic field. We show that an additional oscillating external magnetic field leads to a drift in the domain walls and find the dependence of the drift velocity on the amplitude, frequency, and polarization of the oscillating field. Finally, we discuss the possibility of the domain structure drifting as a whole. © *1997 American Institute of Physics.* [S1063-7761(97)01610-7]

1. INTRODUCTION

It is well known that first-order spin-reorientational phase transitions induced by an external magnetic field or by temperature variations in a magnetically ordered crystal can generate an intermediate state, i.e., a state in which domains of two competing phases coexist in the magnetic material. The intermediate state was first studied theoretically in Refs. 1–3, which was followed by experimental verification that such a state actually exists.[4–8] These same papers also deal with regions of stability of the intermediate state and high-frequency properties (in particular, antiferromagnetic resonance). A consistent general approach to analyzing thermodynamically stable intermediate states is developed in Refs. 9 and 10, and at present the idea that two-phase domain structures constitute the intermediate state of a magnetic material is generally accepted (see, e.g., Refs. 11–13).

The domain structure of the intermediate state is thermodynamically stable and its energy is lower than that of the constituent homogeneous magnetic phases. The domains in the intermediate state have various symmetries and are separated by domain walls. These walls are 90-degree in a spin-flop transition in an antiferromagnet in an external magnetic field directed along the easy-magnetization axis, and in a temperature phase transition of the Morin type between antiferromagnetic and weakly ferromagnetic states.

The dynamical properties of domain walls (and intermediate states as a whole), i.e., the laws governing the behavior of such walls in external fields, have not been studied as thoroughly as the analogous properties of 180-degree domain walls and the corresponding domain structures.[14–17] Ivanov[14] analyzes the dynamics of the 90-degree domain wall separating the collinear and ''flopped'' phases of an antiferromagnet in a spin-flop transition. He examines the steady-state motion of a domain wall in an external constant magnetic field, allowing for magnetic-dipole interaction and phenomenologically taking into account relaxation processes. He notes that the dynamic properties of a 90-degree domain wall differ considerably from those of 180-degree walls; in particular, he finds that the highest possible velocity of steady-state motion of a 90-degree domain wall is determined by relativistic interactions, while the highest velocity of 180-degree domain walls in antiferromagnets is determined solely by exchange interactions (see, e.g., Ref. 18).

Soboleva *et al.*[16] study the uniform motion of a 90-degree domain wall formed in a temperature phase transition and find that the velocity of such walls is determined by the balance between ''pressure,'' emerging as a result of the deviation of the system from the phase equilibrium position, and ''friction,'' due to dissipative processes. Stationary and oscillatory movements of a domain wall in rhombic Seignette-ferromagnets in an external electric field are also studied by Soboleva *et al.*[17]

The characteristic features of the dynamical properties of domain walls in dysprosium orthoferrite in a pulsed magnetic field with short buildup times were detected in the experiments of Gnatchenko *et al.*[19–21] The features are manifested in the nonlinear dependence of the wall velocity on the field's amplitude, and in the asymmetry of this dependence under reversal of the pulsed field. These studies mark the beginnings of the experimental investigation of the dynamical properties of the domain walls that exist in the intermediate state, subject to an external alternating magnetic field.

The goal of our theoretical study is to establish the dynamical properties of single 90-degree domain walls in the intermediate states and the domain structure as a whole when an external alternating magnetic field is applied to the system. By way of an example, we examine the domain walls in the intermediate state under the conditions of one of the most thoroughly studied first-order spin-reorientational phase transitions in which an intermediate state of the magnetic material occurs: a spin-flop transition in an antiferromagnet located in an external magnetic field that is parallel to the easy-magnetization axis.[1,2,9–13] The domain structure corresponding to this state consists of alternating domains of two phases, the collinear $\Phi_\parallel$ and the ''flopped'' $\Phi_\perp$, where the antiferromagnetism vector is directed, respectively, along the easy-magnetization axis and perpendicular to it. The intermediate state is realized within a finite range of magnetic field strength, $H_{\perp}<H<H_{\|}$, where $H_{\perp}$ and $H_{\|}$ are, respectively, the lability fields of the ''flopped'' and collinear phases. The spin-flop transition field H_p, i.e., the field in which the energies of the two phases are equal, lies within this range, whose length is, in order of magnitude, $\Delta H = H_{\|} - H_{\perp} \approx 4\pi H_p/\delta$, where δ is the constant of homogeneous exchange between the sublattices. Within this range, the intermediate state is thermodynamically stable, and the phase concentration is a linear function of the external magnetic field.

We start with the dynamics of a single 90-degree domain wall to which an oscillating external field is applied. We then discuss the motion of the domain structure in the intermediate state as a whole.

As is known from the example of 180-degree domain walls in different magnetically ordered crystals, a domain wall oscillates in an oscillating magnetic field at the field frequency and also drifts, i.e., acquires a constant velocity component.[22,23] Below we show that these types of motion are also typical of 90-degree domain walls.

2. BASIC EQUATIONS

According to Refs. 24 and 25 (see also Refs. 18 and 26), the dynamics of a double-sublattice antiferromagnet can be explained using an effective Lagrangian L written in terms of the unit ferromagnetism vector $\mathbf{l} = \mathbf{L}/|\mathbf{L}|$, where $\mathbf{L} = \mathbf{M}_1 - \mathbf{M}_2$, with $\mathbf{M}_1$ and $\mathbf{M}_2$ the magnetization vectors of the sublattices. For an antiferromagnet with rhombic magnetic anisotropy located in an external magnetic field $\mathbf{H}(t)$, the effective Lagrangian density can be written as

$$L = M_0^2 \left\{ \frac{\alpha}{2} \left[\frac{1}{c^2} \dot{\mathbf{l}}^2 + (\nabla \cdot \mathbf{l})^2 \right] - \frac{2}{\delta M_0^2} (\mathbf{l} \cdot \mathbf{H})^2 - \frac{1}{2} (\beta_1 l_x^2 + \beta_2 l_y^2) \right.$$
$$\left. - \frac{1}{4} (\beta_1' l_x^4 + \beta_2' l_x^2 l_y^2 + \beta_3' l_y^4) + \frac{4}{\delta g M_0^2} (\mathbf{H} \cdot (\dot{\mathbf{l}} \times \mathbf{l})) \right\}, \qquad (1)$$

where $M_0 = |\mathbf{M}_{1,2}|$ is the modulus of the sublattice-magnetization vectors, $c = g M_0 (\alpha\delta)^{1/2}/2$ is a characteristic velocity, which coincides with the minimum phase velocity of spin waves in the absence of a magnetic field, α and δ are, respectively, the homogeneous and inhomogeneous exchange constants, g is the gyromagnetic ratio, β_1 and β_2 are the second-order anisotropy constants, and β_1', β_2', and β_3' are the fourth-order anisotropy constants; a superior dot denotes a time derivative. This is a valid description of the dynamics of an antiferromagnet when the external field is much weaker than the exchange field $H_e = \delta M_0$. Here the weak-ferromagnetism vector $\mathbf{M} = \mathbf{M}_1 + \mathbf{M}_2$ is related to $\mathbf{l}$ by

$$\mathbf{M} = \frac{2}{\delta} (\mathbf{l} \times (\mathbf{H} \times \mathbf{l})) + \frac{2}{\delta g M_0} (\dot{\mathbf{l}} \times \mathbf{l}).$$

The dynamical stopping of a domain wall, for which various relaxation processes are responsible, will be taken into account by using a dissipative function Q:

$$Q = \frac{\lambda M_0}{2g} \dot{\mathbf{l}}^2, \qquad (2)$$

where λ is the relaxation constant.

In terms of two independent angle variables that parametrize the unit vector $\mathbf{l}$,

$$l_z + i l_x = \sin\theta \exp i\varphi, \quad l_y = \cos\theta, \qquad (3)$$

the equations of motion with allowance for the dissipative terms can be written as

$$\alpha \left(\Delta\theta - \frac{1}{c^2} \ddot{\theta} \right) + \sin\theta \cos\theta \left[\alpha \left(\frac{1}{c^2} \dot{\varphi}^2 - (\nabla\varphi)^2 \right) \right.$$
$$- \beta_1 \sin^2\varphi + \beta_2 - \beta_1' \sin^2\theta \sin^4\varphi + \beta_3' \cos^2\theta$$
$$\left. - \frac{\beta_2'}{2} \cos 2\theta \sin^2\varphi \right] + \frac{4}{\delta g M_0^2} [\dot{H}_x \cos\varphi$$
$$- \dot{H}_z \sin\varphi - 2\dot{\varphi} \sin^2\theta (H_x \sin\varphi + H_z \cos\varphi)$$
$$- H_y \dot{\varphi} \sin 2\theta] - \frac{4}{\delta M_0^2} (H_y \cos\theta + H_x \sin\theta \sin\varphi$$
$$+ H_z \sin\theta \cos\varphi)(H_x \cos\theta \sin\varphi - H_y \sin\theta$$
$$+ H_z \cos\theta \cos\varphi) = \frac{\lambda}{g M_0} \dot{\theta}, \qquad (4)$$

$$\alpha \nabla (\sin^2\theta (\nabla\varphi)) - \frac{\alpha}{c^2} (\sin^2\theta \dot{\varphi})$$
$$- \sin^2\theta \sin\varphi \cos\varphi \left(\beta_1 + \frac{\beta_2'}{2} \cos^2\theta \right.$$
$$\left. + \beta_3' \sin^2\theta \sin^2\varphi \right) + \frac{4}{\delta g M_0^2}$$
$$\times \left[- \frac{1}{2} \sin 2\theta (\dot{H}_x \sin\varphi + \dot{H}_z \cos\varphi) + \dot{H}_y \sin^2\theta \right.$$
$$\left. + 2\dot{\theta} \sin^2\theta (H_x \sin\varphi + H_z \cos\varphi) + H_y \dot{\theta} \sin 2\theta \right]$$
$$- \frac{4 \sin\theta}{\delta M_0^2} (H_y \cos\theta + H_x \sin\theta \sin\varphi$$
$$+ H_z \sin\theta \cos\varphi)(H_x \cos\varphi - H_z \sin\varphi)$$
$$= \frac{\lambda}{g M_0} \dot{\varphi} \sin^2\theta. \qquad (5)$$

If both β_1 and β_2 are positive, then in an external magnetic field $\mathbf{H}_0$ parallel to the z axis and weaker than the spin-flop transition field $H_p = [\delta(2\beta_1 + \beta_1')/8]^{1/2}$, the unit vector $\mathbf{l}$ in the ground state is collinear with the easy-magnetization axis z. The magnetic material in this case can have two types of 180-degree domain walls: in one the vector $\mathbf{l}$ rotates in the xz plane, and in the other, in the yz plane. If $\beta_2 > \beta_1 > 0$, the domain wall in which $\mathbf{l}$ rotates in the xz plane is stable. This domain wall corresponds to the angle $\theta = \theta_0 = \pi/2$, and the angle variable $\varphi = \varphi_0(y)$ satisfies the equation

$$\alpha \varphi_0'' - \left(\beta_1 - \frac{4H_0^2}{\delta} \right) \sin\varphi_0 \cos\varphi_0 - \beta_1' \sin^3\varphi_0 \cos\varphi_0 = 0 \qquad (6)$$

(we assume that the magnetization distribution in the domain wall is not uniform along the y axis; a prime denotes a derivative with respect to this coordinate).

Note that although we study an antiferromagnet here with rhombic magnetic anisotropy, our results can be generalized to antiferromagnets with a different type of anisotropy in the xy plane—hexagonal, for example—the only condition being that the anisotropy in the xy plane is high (see below). Our theory cannot be applied to a purely uniaxial antiferromagnet, in which the orientation of $\mathbf{L}$ in the ''flopped'' phase is not fixed, with the result that the plane of rotation of $\mathbf{L}$ in the domain wall is not defined. Analysis of the nonlinear dynamics of a domain wall in a uniaxial antiferromagnet merits a separate investigation, which is outside the scope of the present work.

It is known (see, e.g., Refs. 1–3) that if the fourth-order anisotropy constant β_1' is negative, in the spin-flop transition field, i.e., at $H_0=H_p$, a first-order phase transition occurs in which the antiferromagnetism vector $\mathbf{l}$ is reoriented perpendicular to the z axis and at $\beta_2>\beta_1$ is collinear with the x axis. At the phase transition point, the two phases (the collinear and the ''flopped'') can coexist, with a 90-degree domain wall separating them. As Eq. (6) implies (with allowance for the boundary conditions $\varphi_0(-\infty)=0$, $\varphi_0(+\infty)=\pi/2$, and $\varphi_0'(\pm\infty)=0$), this wall can be described by the equations

$$\varphi_0'=\frac{1}{2y_0}\sin 2\varphi_0=\frac{1}{2y_0}\operatorname{sech}\frac{y}{y_0},$$

$$\cos 2\varphi_0=-\tanh\frac{y}{y_0}, \tag{7}$$

where $y_0=\sqrt{2\alpha/|\beta_1'|}$ is the thickness of the 90-degree wall. Note that the wall is much thicker than ordinary 180-degree walls in view of the smallness of the fourth-order anisotropy constant as compared to the second-order anisotropy constants ($y_0\gg y_{180}=\sqrt{\alpha/\beta}$).

Now, if we apply an additional magnetic field along the z axis to the already existing domain wall, the wall begins to move in the direction of a more energetically favored phase with a constant velocity, whose value is determined by the difference between the total field and the spin-flop transition field.[16] Here the highest velocity of motion is determined by the maximum strength of the additional field, the reason being that the total field strength must remain within the lability region of the phases separated by the wall. If a varying field is applied to the domain wall, the wall's motion becomes time-dependent; in particular, in a oscillating field, the wall oscillates at the field frequency and, as we will show, moves with a certain velocity. Moreover, the presence of an additional oscillating field distorts the shape of the domain wall.

Below we examine the dynamics of a domain wall located in an additional oscillating magnetic field $\tilde{\mathbf{H}}(t)$, all three components of which are nonzero and, in the general case, the phase shifts between the components are arbitrary:

$$\tilde{H}_z=\tilde{H}_{0z}\cos\omega t,$$

$$\tilde{H}_x=\tilde{H}_{0x}\cos(\omega t+\chi), \tag{8}$$

$$\tilde{H}_y=\tilde{H}_{0y}\cos(\omega t+\chi_1).$$

We will see that the domain-wall velocity strongly depends on χ and χ_1.

3. LINEAR OSCILLATIONS OF A DOMAIN WALL

Assuming that the additional field is very weak, we employ, following Refs. 22 and 23, one of the versions of perturbation theory used for solitons to analyze domain-wall dynamics. To this end we introduce a collective variable $Y(t)$, which has the meaning of the position of the domain wall at time t, and seek a solution of Eqs. (4) and (5) in the form

$$\theta=\frac{\pi}{2}+\vartheta(\xi,t),\quad \varphi=\varphi_0(\xi)+\psi(\xi,t), \tag{9}$$

where $\xi=y-Y(t)$. The function $\varphi_0(\xi)$ describes the motion of an undistorted domain wall (the structure of $\varphi_0(\xi)$ is the same as that of $\varphi_0(y)$ in the static solution (7)), and the functions $\psi(\xi)$ and $\vartheta(\xi)$ represent the distortion of the domain wall. The velocity of domain-wall drift is defined as the instantaneous velocity $V(t)=\dot{Y}(t)$ averaged over the oscillation period, $V_{\mathrm{dr}}=\overline{V(t)}$, where the bar indicates averaging over the oscillation period of the external field.

We seek the functions $\psi(\xi)$ and $\vartheta(\xi)$, which describe the distortion of the domain wall, and the domain-wall velocity $V(t)$ in the form of series expansions in powers of the field amplitude, bearing in mind that we are interested only in the forced motion of the wall:

$$\vartheta(\xi,t)=\vartheta_1(\xi,t)+\vartheta_2(\xi,t)+\cdots,$$

$$\psi(\xi,t)=\psi_1(\xi,t)+\psi_2(\xi,t)+\cdots, \tag{10}$$

$$V=V_1+V_2+\cdots,$$

where the subscripts 1, 2, etc. denote the order of smallness of the quantity in relation to the field amplitude, and $\psi_n,\vartheta_n,V_n\sim h^n$.

We substitute the expansions (10) into Eqs. (4) and (5) and separate the terms of different orders of smallness. Obviously, in the zeroth approximation we obtain Eq. (6), which describes a domain wall at rest.

The first-order perturbation equations can be written in the form

$$(\hat{L}+\hat{T})\psi_1-2\frac{\omega_0}{\tilde{\omega}_0^2}\dot{\vartheta}_1\cos\varphi_0=\frac{\sin 2\varphi_0}{2y_0\tilde{\omega}_0^2}(\dot{V}_1+\omega_rV_1)$$

$$+\frac{gM_0}{\tilde{\omega}_0^2}[\dot{h}_y+\omega_0(h_z\sin 2\varphi_0-h_x\cos 2\varphi_0)], \tag{11}$$

$$(\hat{L}'+\hat{T}+\sigma)\vartheta_1+2\frac{\omega_0}{\tilde{\omega}_0^2}\dot{\psi}_1\cos\varphi_0$$

$$=\frac{\omega_0V_1}{y_0\tilde{\omega}_0^2}\cos\varphi_0\sin 2\varphi_0+\frac{gM_0}{\tilde{\omega}_0^2}[\dot{h}_x\cos\varphi_0$$

$$-\dot{h}_z \sin \varphi_0 + \omega_0 h_y \cos \varphi_0)], \tag{12}$$

Here we have introduced the notation $\mathbf{h}(t) = \widetilde{\mathbf{H}}(t)/M_0$, $\sigma = (\varkappa^{-1} - 1)\omega_0^2/\widetilde{\omega}_0^2$, with $\varkappa = \beta_1/\beta_2$, and

$$\hat{T} = \frac{1}{\widetilde{\omega}_0^2} \frac{d^2}{dt^2} + \frac{\omega_r}{\widetilde{\omega}_0^2} \frac{d}{dt},$$

where $\omega_0 = gM_0\sqrt{\beta_1 \delta}/2$ and $\widetilde{\omega}_0 = c/y_0 = \omega_0\sqrt{|\beta_1'|/2\beta_1} \ll \omega_0$ are the frequencies of activation of the bulk branch of spin waves in the magnetic material at $H_0 = 0$ and $H_0 = H_p$, respectively, and $\omega_r = \lambda \delta g M_0/4$ is the characteristic relaxation frequency.

The operator $\hat{L}$ in Eq. (11) has the form of a Schrödinger operator with a nonreflecting potential:

$$\hat{L} = -y_0^2 \frac{d^2}{d\xi^2} + 1 - \frac{2}{\cosh^2(\xi/y_0)}. \tag{13}$$

The spectrum and the wave functions of the operator (13) are well known. The spectrum has one discrete level $\lambda_0 = 0$ corresponding to a localized wave function,

$$f_0(\xi) = \left[\sqrt{2y_0} \cosh \frac{\xi}{y_0}\right]^{-1}, \tag{14}$$

and a continuous part $\lambda_k = 1 + (ky_0)^2$ corresponding to the eigenfunctions

$$f_k(\xi) = \frac{1}{b_k\sqrt{L}}\left[\tanh \frac{\xi}{y_0} - iky_0\right] e^{ik\xi}, \tag{15}$$

where $b_k = \sqrt{1 + (ky_0)^2}$, and L is the crystal length.

In contrast to $\hat{L}$, the operator

$$\hat{L}' = \hat{L} + \frac{5}{4\cosh^2(\xi/y_0)} + \left(1 + \frac{\beta_2'}{|\beta_1'|}\right) \frac{\exp(\xi/y_0)}{2\cosh(\xi/y_0)} \tag{16}$$

is not nonreflecting, and there is no way in which we can find its spectrum and eigenfunctions, which makes the analysis of the dynamics of a 90-degree domain wall much more complicated than the analysis of the similar problem for 180–degree walls (for all models of magnetic materials considered in Refs. 22, 23, 27 and 28 the first-approximation equations contain only the nonreflecting operator $\hat{L}$). Hence, in solving the first-approximation equations we (a) use the fact that $\widetilde{\omega}_0 \ll \omega_0$ and (b) restrict our study to low-frequency external fields, i.e., $\omega \ll \widetilde{\omega}_0$. Such restrictions are justified by experimental considerations (see Refs. 19–21), with the frequencies of the alternating field being of order 10^6–10^7 s^{-1}, while for typical antiferromagnets $\omega_0 \sim 10^{11}$ s^{-1} and $\widetilde{\omega}_0 \sim 10^9$–$10^{10}$ s^{-1}. We also assume that in our antiferromagnet anisotropy in the xy plane is high, i.e., $(\beta_2 - \beta_1)/\beta_1 \gg (\widetilde{\omega}_0/\omega_0)^2 \sim |\beta_1'|/\beta_1$. This condition is not too stringent, since usually fourth-order anisotropy constants are much smaller than second-order anisotropy constants, by a factor of 10 or even by several orders of magnitude. Hence, a 10% difference in the constants (which is typical of most antiferromagnets) is enough for the approximation to work.1) Here the parameter σ in Eq. (12) is much greater than unity, with the result that in this equation we can ignore the term $(\hat{L}' + \hat{T})\vartheta_1$ in comparison with $\sigma\vartheta_1$ and express the function $\vartheta_1(\xi,t)$ in terms of $\psi_1(\xi,t)$:

$$\vartheta_1 = \frac{1}{\sigma}\left\{\frac{\omega_0 V_1}{y_0\widetilde{\omega}_0^2} \cos \varphi_0 \sin 2\varphi_0 + \frac{gM_0}{\widetilde{\omega}_0^2}[\dot{h}_x \cos \varphi_0 \right.$$
$$\left. - \dot{h}_z \sin \varphi_0 + \omega_0 h_y \cos \varphi_0] - 2\frac{\omega_0}{\widetilde{\omega}_0^2}\dot{\psi}_1 \cos \varphi_0\right\}. \tag{17}$$

Plugging the expression for ϑ_1 into (11) and ignoring small terms in the parameter $\omega/\widetilde{\omega}_0 \ll 1$, we arrive at an equation for the function $\psi_1(\xi,t)$:

$$(\hat{L} + \hat{T})\psi_1 = \frac{\sin 2\varphi_0}{2y_0\widetilde{\omega}_0^2}(\dot{V}_1 + \omega_r V_1) + \frac{gM_0}{\widetilde{\omega}_0^2}$$
$$\times[(1 + 2\kappa \cos^2 \varphi_0)\dot{h}_y$$
$$+ \omega_0(h_z \sin 2\varphi_0 - h_x \cos 2\varphi_0)]. \tag{18}$$

We seek a solution of Eq. (18) in the form of an expansion in the eigenfunctions of the operator (13), which comprise a complete orthonormal set:

$$\psi_1(\xi,t) = \mathrm{Re}\left\{\left[\sum_k d_k f_k(\xi) + d_0 f_0(\xi)\right] e^{i\omega t}\right\}. \tag{19}$$

The expansion coefficients d_k and d_0 in (19) can be found in the usual way by multiplying the right-hand side of Eq. (18) by f_k^* and f_0^*, respectively, and integrating the products with respect to ξ.

One important remark is in order here. The first-approximation equations (11) and (12) describe the excitation of linear spin waves against a domain-wall background. The last term in the expansion (19) corresponds to the shear (Goldstone) mode, which describes the domain-wall motion as a whole. The corresponding degree of freedom of the system, however, has already been taken into account by introducing the collective coordinate $Y(t)$ into the definition of the variable ξ. The Goldstone mode in the expansion (19) must therefore be dropped, i.e., we must set the corresponding expansion coefficient to zero, $d_0 = 0$ (see the discussion of this problem in Rajaraman's book[29]).

For a monochromatic field of frequency ω we obtain

$$\psi_1(\xi,t) = \frac{gM_0}{\widetilde{\omega}_0^2}[a_1 \cos 2\varphi_0 + a_2 G(\xi)],$$

$$\vartheta_1(\xi,t) = \frac{gM_0}{\widetilde{\omega}_0^2}[\cos \varphi_0(a_3 \cos 2\varphi_0 + a_4 \sin 2\varphi_0)$$
$$+ a_5 \cos \varphi_0 + a_6 \sin \varphi_0]. \tag{20}$$

Here we have introduced the notation

$$a_0 = \kappa\dot{h}_y - \omega_0 h_x, \quad a_2 = (1 + \kappa)\dot{h}_y, \quad a_3 = 2\dot{h}_x,$$

$$a_4 = -\frac{2\widetilde{\omega}_0^2\omega_r}{\omega^2 + \omega_r^2}h_z, \quad a_5 = \left(\frac{\widetilde{\omega}_0}{\omega_0}\right)^2(\dot{h}_x + \omega_0 h_y),$$

$$a_6=-\left(\frac{\tilde{\omega}_0}{\omega_0}\right)^2\dot{h}_z,$$

$$G(\xi)=\frac{y_0}{2}\int_{-\infty}^{\infty}dk\,\frac{\tanh(\xi/y_0)\sin k\xi-ky_0\cos k\xi}{(1+k^2y_0^2)^{3/2}\sinh(\pi ky_0/2)}.$$

The requirement that there be no Goldstone mode in (19) is equivalent to the requirement that the right-hand side of Eq. (18) be orthogonal to the function f_0, and this determines the equation for the velocity $V_1(t)$ of a domain wall in the approximation that is linear in the field strength:

$$\dot{V}_1+\omega_rV_1=-y_0gM_0[2\omega_0h_z(t)+\pi(1+\kappa)\dot{h}_y(t)]. \quad (21)$$

Solving Eq. (21) with a monochromatic field of type (8) is fairly simple:

$$\begin{aligned}V_1(t)=&\mu_zh_{0z}[\omega_r\cos\omega t+\omega\sin\omega t]\\&+\mu_yh_{0y}[\omega_r\sin(\omega t+\chi_1)-\omega\cos(\omega t+\chi_1)],\end{aligned} \quad (22)$$

where

$$\mu_z=-\frac{2gM_0y_0\omega_0}{\omega^2+\omega_r^2},\quad \mu_y=\frac{\pi gM_0(1+\kappa)y_0\omega}{\omega^2+\omega_r^2}.$$

The expression (22) describes the oscillations of a domain wall in an oscillating external field. Clearly, it does not lead to wall drift: $\overline{V_1(t)}=0$. Note that at $\omega=0$, i.e., in a static field, μ_z remains finite (in contrast to μ_y, which vanishes) and describes domain-wall motion with a constant velocity.[16]

4. SECOND APPROXIMATION. DOMAIN-WALL DRIFT

We now analyze the equations of the second approximation in the amplitude of the external magnetic field.

Here we will not write the system of equations of the second approximation in general form. Instead, we give only the equation that follows from Eq. (4):

$$\hat{L}\psi_2=\frac{\sin 2\varphi_0}{2y_0\tilde{\omega}_0^2}(\dot{V}_2+\omega_rV_2)+N(\xi,t), \quad (23)$$

where the function $N(\xi,t)$ is defined as follows:

$$\begin{aligned}N(\xi,t)=&\frac{1}{\tilde{\omega}_0^2}(\dot{V}_1+\omega_rV_1)\psi_1'+\sin 2\varphi_0(5\cos^2\varphi_0-1)\psi_1^2\\&-\frac{1}{4}\left(\frac{V_1}{c}\right)^2\sin 4\varphi_0-y_0\sin 2\varphi_0\vartheta_1\vartheta_1'\\&-\vartheta_1^2\left(\sin^2\varphi_0-\frac{\beta_2'}{2|\beta_1'|}\right)\sin 2\varphi_0\\&-\frac{gM_0}{\tilde{\omega}_0^2}[(\dot{h}_x\sin\varphi_0+\dot{h}_z\cos\varphi_0)\vartheta_1\\&+2(h_x\sin\varphi_0+h_z\cos\varphi_0)\dot{\vartheta}_1]\\&-\left(\frac{gM_0}{\tilde{\omega}_0}\right)^2\left[h_xh_z\cos 2\varphi_0+\frac{1}{2}(h_x^2\right.\\&\left.-h_z^2)\sin 2\varphi_0\right]-2\frac{\omega_0}{\tilde{\omega}_0^2}(V_1\vartheta_1'\cos\varphi_0\\&+\psi_1\dot{\vartheta}_1\sin\varphi_0)+\frac{gM_0\omega_0}{\tilde{\omega}_0^2}[2\psi_1(h_z\cos 2\varphi_0\\&+h_x\sin 2\varphi_0)-\vartheta_1h_y\sin\varphi_0].\end{aligned} \quad (24)$$

The second equation in the system, which follows from Eq. (5) and determines the function $\vartheta_2(\xi,t)$, has a similar structure but contains no second-order term in the expansion of the domain-wall velocity, V_2, so that it is of no interest to us.

A solution of Eq. (23) can also be sought in the form of an expansion in the eigenfunctions of $\hat{L}$ similar to (19). Here too, as in the first-approximation equation, we must require that there be no shear mode in the expansion of the function $\psi_2(\xi,t)$, i.e., the right-hand side of Eq. (23) must be orthogonal to the function $f_0(\xi)$. This leads to an equation that determines the second-order term V_2 in the expansion of the domain-wall velocity:

$$\dot{V}_2+\omega_rV_2=-\tilde{\omega}_0^2\int_{-\infty}^{\infty}d\xi\,N(\xi,t)\sin 2\varphi_0. \quad (25)$$

Plugging the functions $\psi_1(\xi,t)$ and $\vartheta_1(\xi,t)$ calculated in Sec. 3 into (23) and averaging over the oscillation period and integrating in (25), we arrive at an expression for the drift velocity $V_{\mathrm{dr}}=\overline{V_2(t)}$,

$$V_{\mathrm{dr}}=\sum_{ij}\nu_{ij}(\omega;\chi,\chi_1)\tilde{H}_{0i}\tilde{H}_{0j}. \quad (26)$$

We will call the coefficients ν_{ij}, which are functions of the field frequency and phase shifts, the nonlinear mobilities of a domain wall:

$$\begin{aligned}&\nu_{xx}=-\nu_0\left(\frac{\omega_0}{\tilde{\omega}_0}\right)^2,\quad \nu_{yy}=\nu_0\kappa,\\&\nu_{zz}=-\nu_0\left[1+\frac{4}{3}\kappa^2\left(\frac{\beta_2'}{|\beta_1'|}-\frac{2}{5}\right)\left(\frac{\tilde{\omega}_0}{\omega_r}\right)^2\right],\\&\nu_{xz}=2\pi\nu_0\left(\frac{\omega_0}{\tilde{\omega}_0}\right)^2\left[\cos\chi-\kappa\frac{\omega_0}{\omega_r}\sin\chi\right],\\&\nu_{xy}=\nu_0\frac{\omega\omega_0}{\tilde{\omega}_0^2}[2\kappa+(1+\kappa)\eta]\sin(\chi-\chi_1),\\&\nu_{yz}=2\pi\nu_0\kappa\left[\frac{\omega\omega_0}{\tilde{\omega}_0^2}\sin\chi_1\right.\\&\qquad\left.-\frac{\kappa}{16}\left(1-\frac{4\beta_2'}{|\beta_1'|}\right)\frac{\tilde{\omega}_0^2}{\omega_0\omega_r}\cos\chi_1\right],\end{aligned} \quad (27)$$

where $\nu_0=y_0g^2/2\omega_r$, and η is a numerical factor of order unity. In deriving (27) we employed the fact that the characteristic relaxation frequency in an antiferromagnet is very high ($\omega_r\sim10^{10}\ \mathrm{s}^{-1}$), since the relaxation constant λ in (27) is exchange-amplified, so that at frequencies commonly used in experiments (10^6–$10^7\ \mathrm{s}^{-1}$) we have $\omega\ll\omega_r$.

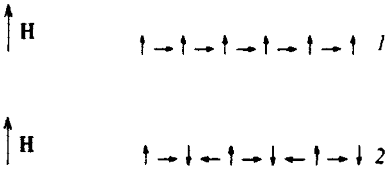

FIG. 1. Two possible types of orientation of the antiferromagnetism vector in a domain structure that is in the intermediate state; only the structure *2* can drift.

Note that in contrast to the drift of 180-degree domain walls studied in Refs. 22, 23, 27, and 28, the drift of 90-degree domain walls is possible even if the oscillating external magnetic field has only one component (in the case of 180-degree domain walls only the off-diagonal components of the nonlinear mobility tensor are nonzero).

To make a numerical estimate of the components of the nonlinear mobility tensor we use the following typical values of the parameters: $y_0 \sim 10^{-5}$ cm, $g \sim 10^7\ \mathrm{s}^{-1}\,\mathrm{Oe}^{-1}$, and $\omega_r \sim 10^{10}\ \mathrm{s}^{-1}$. Here ν_0 proves to be of order $10^{-1}\ \mathrm{cm\,s}^{-1}\,\mathrm{Oe}^{-2}$. Assuming also that $\kappa \sim 1$, $\omega_0 \sim 10^{11}\ \mathrm{s}^{-1}$, and $\tilde{\omega}_0 \sim 10^{10}\ \mathrm{s}^{-1}$, we find from the expressions for the components of the nonlinear mobility tensor that at the frequencies used in the experiments, $\omega \sim 10^7\ \mathrm{s}^{-1}$, the following estimates hold: $\nu_{xx} \sim 10^2 \nu_0$, $\nu_{yy} \sim \nu_0$, $\nu_{zz} \sim 10^2 \nu_0$, $\nu_{xz} \sim 10^2 \nu_0$ at $\chi = 0$, $\nu_{xz} \sim 10^3 \nu_0$ at $\chi \neq 0$, $\nu_{yz} \sim 10^{-1} \nu_0$ at $\chi_1 \neq \pi/2$, $\nu_{yz} \sim 10^{-2} \nu_0$ at $\chi = \pi/2$, $\nu_{xz} \sim 10^{-2} \nu_0$ at $\chi \neq \chi_1$, and $\nu_{xy} = 0$ at $\chi = \chi_1$.

Note that some of the components of the nonlinear mobility tensor strongly depend on the phase shift between the corresponding components of the external field; in particular, ν_{xz} increases by a factor of ten when $\chi \neq 0$. It is in this case, i.e., in the presence of x- and z-components of the oscillating field and a nonzero phase shift between these components, that the domain-wall drift velocity is highest (with the amplitude of the external field of about one oersted, the velocity may be as high as 1 m/s).

5. DRIFT OF A TWO-PHASE DOMAIN STRUCTURE IN THE INTERMEDIATE STATE

The intermediate state in question emerges as a result of a spin-flop transition and consists of alternating domains of the ''easy-axis'' (antiferromagnetic) phase in which the antiferromagnetism vector **l** is either parallel or antiparallel to the z axis (and $\mathbf{m} = 0$) and the ''flopped'' (weakly ferromagnetic) phase in which **l** is either parallel or antiparallel to the x axis (and $\mathbf{m} \neq 0$). A typical domain structure in the intermediate state is depicted in Fig. 1.

Clearly, in this state there can be eight types of 90-degree domain walls separating the domains with different orientations of **l** and, in addition, having different directions of rotation of this vector when we go from the left domain to the right one. Naturally, the domain structure in the intermediate state drifts as a whole only if all the domain walls in the structure drift in one direction and have the same drift velocity.

In Sec. 4 we discussed the drift of a single domain wall with a well-defined orientation of the antiferromagnetism vector in the left and right domains; specifically, in solving Eq. (6), which determines the domain-wall structure, we used the boundary conditions $\varphi_0(-\infty) = 0$ and $\varphi_0(+\infty) = \pi/2$, i.e., we assumed that **l** is parallel to the z axis in the left domain ($y \to -\infty$) and parallel to the x axis in the right domain ($y \to +\infty$).

Analysis of the dynamics of all other domain walls that can exist in the current structure shows that in a given oscillating external magnetic field, all the walls drift with the same absolute value of the drift velocity but with different directions of drift, which depend on the mutual orientation of **l** in domains separated by a wall. It turns out that all domain walls in which the sign of the derivative φ_0' is the same, i.e., walls with the same direction of rotation of the vector **l**, move in the same direction. For example, all walls in which $\varphi_0' > 0$, i.e., in which **l** rotates clockwise in the transition from the left domain to the right one, move in the same direction as the specific domain wall studied above. Such a structure is shown schematically in Fig. 1 (type *2*).

Generally speaking, in the majority of experiments in which the spin-flop transition was studied, the intermediate state with such a domain structure does not emerge—instead, a structure emerges in which the adjacent domain walls are characterized by different directions of rotation of **l**, which means that for such structures, drift as a whole is impossible in our model. However, the required structure can still be formed by special means, and the principles of such formation are discussed in Refs. 30 and 31.

This work was partially supported by a grant from the International Soros Science Education Program (ISSEP) (Grant No. APU 062018).

[*] E-mail: sukstan@purr.dipt.donetsk.ua

[1] As noted earlier, the adopted approximation does not work with uniaxial antiferromagnets, in which $\beta_1 = \beta_2$ and $\sigma = 0$.

---

[1] V. G. Bar'yakhtar, A. E. Borovik, and V. A. Popov, JETP Lett. **9**, 391 (1969).

[2] V. G. Bar'yakhtar, A. E. Borovik, and V. A. Popov, Zh. Éksp. Teor. Fiz. **59**, 1299 (1970) [Sov. Phys. JETP **32**, 709 (1971)].

[3] V. G. Bar'yakhtar, A. E. Borovik, and V. A. Popov, Zh. Éksp. Teor. Fiz. **62**, 2233 (1972) [Sov. Phys. JETP **35**, 1169 (1972)].

[4] V. G. Bar'yakhtar, A. A. Galkin, S. N. Kovner, and V. A. Popov, Zh. Éksp. Teor. Fiz. **58**, 494 (1970) [Sov. Phys. JETP **31**, 264 (1970)].

[5] K. L. Dudko, V. V. Eremenko, and V. M. Fridman, Zh. Éksp. Teor. Fiz. **61**, 678 (1971) [Sov. Phys. JETP **34**, 362 (1972)].

[6] A. King and D. Paguette, Phys. Rev. Lett. **30**, 662 (1972).

[7] N. F. Kharchenko, V. V. Eremenko, and S. L. Gnatchenko, JETP Lett. **20**, 280 (1974).

[8] V. V. Eremenko, A. V. Klochko, and V. M. Naumenko, Zh. Éksp. Teor. Fiz. **89**, 1002 (1985) [Sov. Phys. JETP **62**, 576 (1985)].

[9] V. G. Bar'yakhtar, A. N. Bogdanov, and D. A. Yablonskiĭ, Fiz. Nizk. Temp. **12**, 43 (1986) [Sov. J. Low Temp. Phys. **12**, 24 (1986)].

[10] V. G. Bar'yakhtar, A. N. Bogdanov, and D. A. Yablonskiĭ, Usp. Fiz. Nauk **156**, 47 (1988) [Sov. Phys. Usp. **31**, 810 (1988)].

[11] S. V. Vonsovskii, *Magnetism*, 2 vols., Wiley, New York (1974).

[12] K. P. Belov, A. K. Zvezdin, A. M. Kadomtseva, and R. Z. Levitin, *Orientational Transitions in Rare-Earth Magnetic Substances* [in Russian], Nauka, Moscow (1979).

[13] V. V. Eremenko, N. F. Kharchenko, and Yu. G. Litvinenko, *Magneto-optics and Spectroscopy of Antiferromagnets* [in Russian], Naukova Dumka, Kiev (1989).

[14] B. A. Ivanov, Zh. Éksp. Teor. Fiz. **79**, 581 (1980) [Sov. Phys. JETP **52**, 293 (1980)].

[15] A. K. Zvezdin and A. A. Mukhin, Kratk. Soobshch. Fiz. **6**, 11 (1985).
[16] T. K. Soboleva, E. P. Stefanovskiĭ, and A. L. Sukstanskiĭ, JETP Lett. **42**, 68 (1985).
[17] T. K. Soboleva, E. P. Stefanovskiĭ, and A. L. Sukstanskiĭ, Fiz. Tverd. Tela (Leningrad) **26**, 2725 (1984).
[18] V. G. Bar'yakhtar, B. A. Ivanov, and M. V. Chetkin, Usp. Fiz. Nauk **146**, 417 (1985) [Sov. Phys. Usp. **28**, 563 (1985)].
[19] S. L. Gnatchenko, N. F. Kharchenko, A. B. Chizhik, V. V. Eremenko, and R. Shimchak, Fiz. Nizk. Temp. **12**, 1111 (1986) [Sov. J. Low Temp. Phys. **12**, 628 (1986)].
[20] S. L. Gnatchenko, A. B. Chizhik, and N. F. Kharchenko, Fiz. Nizk. Temp. **15**, 304 (1989) [Sov. J. Low Temp. Phys. **15**, 170 (1989)].
[21] S. L. Gnatchenko, A. B. Chizhik, and N. F. Kharchenko, JETP Lett. **51**, 324 (1990).
[22] V. G. Bar'yakhtar, Yu. I. Gorobets, and S. I. Denisov, Zh. Éksp. Teor. Fiz. **98**, 1345 (1990) [Sov. Phys. JETP **71**, 751 (1990)].
[23] V. S. Gerasimchuk and A. L. Sukstanskiĭ, Zh. Éksp. Teor. Fiz. **103**, 151 (1993) [JETP **76**, 82 (1993)].
[24] I. V. Bar'yakhtar and B. A. Ivanov, Fiz. Nizk. Temp. **5**, 759 (1979) [Sov. J. Low Temp. Phys. **5**, 361 (1979)].
[25] V. G. Bar'yakhtar, B. A. Ivanov, and A. L. Sukstanskiĭ, Zh. Éksp. Teor. Fiz. **78**, 1509 (1980) [Sov. Phys. JETP **51**, 757 (1980)].
[26] A. M. Kosevich, B. A. Ivanov, and A. S. Kovalev, *Nonlinear Magnetization Waves. Dynamic and Topological Solitons* [in Russian], Naukova Dumka, Kiev (1983).
[27] V. S. Gerasimchuk and A. L. Sukstanskiĭ, Fiz. Nizk. Temp. **20**, 142 (1994) [Low Temp. Phys. **20**, 114 (1994)].
[28] V. S. Gerasimchuk and A. L. Sukstanskii, J. Magn. Magn. Mater. **146**, 323 (1995).
[29] R. Rajaraman, *Solitons and Instantons: An Introduction to Solitons and Instantons in Quantum Field Theory*, North-Holland, Amsterdam (1982).
[30] N. F. Kharchenko and V. A. Bedarev, JETP Lett. **56**, 345 (1992).
[31] N. F. Kharchenko and V. A. Bedarev, Fiz. Nizk. Temp. **19**, 72 (1993) [Low Temp. Phys. **19**, 52 (1993)].

Translated by Eugene Yankovsky

Anomalous behavior of the specific heat and upper critical magnetic field in the superconducting single crystal $La_{1.85}Sr_{0.15}CuO_4$

M. N. Khlopkin, G. Kh. Panova, N. A. Chernoplekov, A. A. Shikov, and A. V. Suetin

Kurchatov Institute, 123182 Moscow, Russia

(Submitted 24 December 1996)
Zh. Éksp. Teor. Fiz. **112**, 1386–1395 (October 1997)

The specific heat and resistive upper critical magnetic field of the single crystal $La_{1.85}Sr_{0.15}CuO_4$ are investigated in the temperature range 2–50 K in magnetic fields up to 8 T for two directions of the magnetic field, parallel and normal to the ab crystalline plane. For both orientations a nonlinear (close to square root) magnetic field dependence of the mixed-state specific heat and a positive curvature of the temperature dependence of the upper critical magnetic field are observed. Neither of these anomalies is described by standard theories of superconductivity. Within the framework of the thermodynamic relations it is shown that in a type-II superconductor a relationship exists between the temperature dependence of the critical magnetic field and the field dependence of the specific heat. The anomalies observed in these phenomena are interrelated. *© 1997 American Institute of Physics.* [S1063-7761(97)01710-1]

1. INTRODUCTION

Many properties of high-temperature superconductors (HTSC's), discussed within the framework of various theoretical models,[1–6] different significantly from the properties of ordinary superconductors. An upper critical magnetic field $H_{c2}(T)$ with positive curvature of its temperature dependence is frequently observed in oxide (including high-temperature) superconductors[7–13] and indicates a possible difference between the superconductivity mechanism in these systems and the standard mechanism (standard superconductors are characterized by negative curvature of the $H_{c2}(T)$ curve with saturation at low temperatures).

The electronic specific heat is of special interest. For one of the models of nonstandard pairing (with lines where the energy gap on the Fermi surface vanishes), Volovik[6] has predicted an anomalous influence of the magnetic field on the specific heat of the superconductor near zero temperature, specifically a nonlinear dependence of the specific heat for the superconductor in the mixed state on the magnetic field (in standard superconductors with an isotropic energy gap this dependence is nearly linear[14–16]). Later a similar anomaly was observed experimentally in the superconductor $YBa_2Cu_3O_7$ (Refs. 17–20). Note, however, that in Refs. 17–20 the effect of the magnetic field on the specific heat was observed against the background of a large parasitic magnetic contribution to the specific heat, which introduced significant uncertainty in the inferred field dependence of the specific heat of the superconductor proper.

The indicated anomalies are observed in various phenomena and in different temperature regions (the first—in the critical magnetic field near T_c, and the second—in the specific heat in the low-temperature region), and were previously considered to be independent.

The goal of the present work was to experimentally study the specific heat of the high-purity single crystal $La_{1.85}Sr_{0.15}CuO_4$ in the mixed state, and to determine its upper critical magnetic field for the magnetic field oriented parallel and perpendicular to the ab plane.

2. SAMPLE PREPARATION AND CHARACTERISTICS

The $La_{1.85}Sr_{0.15}CuO_4$ single crystal for study was obtained by the floating zone melting method with optical radiation heating on a URN-2-ZM setup[21] at a crystallization rate of 0.5–0.7 mm/h. The samples were nearly cylindrical, with a diameter of 7 mm and length of 16 mm, and were black with metallic lustre.

The structural parameters and degree of crystalline perfection of the samples were monitored by neutron and x-ray diffractometers.

According to the room-temperature x-ray diffraction data, the sample had tetragonal structure (F4/mmm) with lattice parameters $a=b=3.773(1)$ Å and $c=13.233(2)$ Å. The misalignment was on the order of 8°. The lattice parameters of the sample are in good agreement with the data of Chou *et al.*[22] for a sample of $La_{1.84}Sr_{0.16}CuO_4$.

The superconducting transition temperature T_c and its width ΔT_c were determined to be $T_c=39.2$ K and $\Delta T_c=0.5$ K from the magnetic susceptibility measured at constant current and from the resistance, and $T_c=38$ K and $\Delta T_c=2$ K from the specific heat. A jump was observed in the specific heat temperature dependence, corresponding to the superconducting transition with amplitude $\Delta C/T_c=9\ \mathrm{mJ/mol\cdot K^2}$. The resistance ratios $R(296\ \mathrm{K})/R(45\ \mathrm{K})$ were 1.4 for the current perpendicular to the ab plane and 1.8 for the current in the ab plane for the corresponding absolute values of the room-temperature resistivity $\rho_\perp=30\ \mathrm{m\Omega\cdot cm}$ and $\rho_\parallel=6\ \mathrm{m\Omega\cdot cm}$.

3. EXPERIMENTAL TECHNIQUE

The specific heat of the sample in magnetic fields up to 8 T was determined by the adiabatic method with pulsed

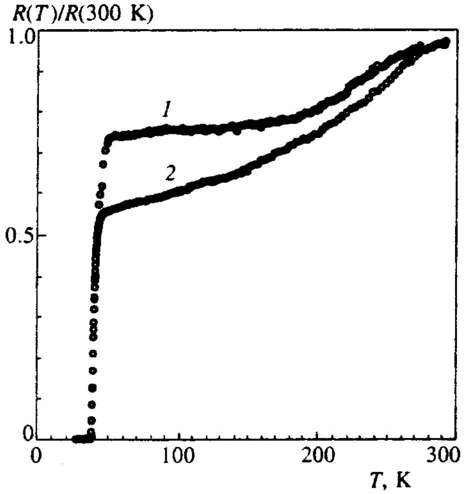

FIG. 1. Temperature dependence of the resistance in zero magnetic field for two current directions: (*1*) perpendicular and (*2*) parallel to the *ab* plane.

heating.[23] The experimental error was about 2% in the temperature range 1.5–4 K, 1% in the range 4–10 K, and 0.2–0.5% in the range 10–50 K.

The temperature dependence of the resistance in a magnetic field was measured in a separate experiment by the standard constant-current four-contact method with the current density not exceeding 0.5 A/cm^2. In the first case the ab plane was perpendicular to the field, and the current was directed in this plane. In the second case the ab plane was parallel to the field, and the current was perpendicular to this plane.

4. EXPERIMENTAL RESULTS

The main results of our resistance and specific heat measurements of the single crystal $La_{1.85}Sr_{0.15}CuO_4$ are shown in Figs. 1–6.

The temperature dependence of the relative resistance $R(T)/R(300\,K)$ is plotted in Fig. 1 for the two directions of the current at which the measurements were made, parallel and perpendicular to the ab plane. As can be seen from this figure, both current directions are characterized by ''metallic behavior'' of the resistance with the value of $\partial R/\partial T$ falling as the temperature is lowered.

The temperature dependence of the resistance in magnetic fields up to 8 T in the region of the superconducting transition is shown in Fig. 2. The superconducting transition temperature, defined as the point at which the resistance fell to 50% of its normal-state value, was found to be 39.2 K in zero field. Here as R_{norm} we took the resistance at $T=46$ K, where the magnetic field is no longer observed to have an effect on the resistance.

As can be seen from Fig. 2, the superconducting transition observed in the temperature dependence of the resistance is noticeably smeared on the high-temperature side. As the magnetic field is increased, the curves of the superconducting transition are shifted toward lower temperatures with an increase in their width ΔT_c. A strong anisotropy in the broadening of the superconducting transition is also observed, which is typical of all single-crystal HTSC systems and which, as a rule, can be explained within the framework of the fluctuation theory.[24,25] In particular, for $\mu_0 H=8$ T ($\mu_0=4\pi\cdot 10^{-7}$ H/m is the permeability) the anisotropy of broadening of the superconducting transition $\Delta T_c(\perp)/\Delta T_c(\parallel)$ at the 10% and 90% levels (of R_{norm}) reaches 2.6.

The experimental results allowed us to determine the temperature dependence of the upper critical field $H_{c2}(T)$ for the two given directions of the magnetic field, which is plotted in Fig. 3. These curves display a nonlinear dependence with positive curvature. The initial slope $\mu_0[-dH_{c2}(T)/dT]_{T_c}$, determined from the linear part of the dependence $H_{c2}(T)$, is equal to 1.45 T/K for the magnetic field perpendicular to the ab plane, and 4.37 T/K for the magnetic field in the ab plane.

Regarding the problem of estimating $H_{c2}(0)$ in the $La_{1.85}Sr_{0.15}CuO_4$ system, note that for low-temperature superconductors, measuring $H_{c2}(T)$ does not present any special difficulties since their values lie in the range of magnetic fields accessible to most laboratories. Similar measurements for high-temperature superconductors, as a rule, are limited to the temperature region near T_c, since the values of $H_{c2}(T)$ significantly exceed the accessible magnetic fields. Therefore, to estimate $H_{c2}(0)$ use is usually made of the following expression, taken from the WHH theory:[26]

$$H_{c2}(0)=0.7T_c(dH_{c2}/dT)_{T_c}, \tag{1}$$

where $(dH_{c2}/dT)_{T_c}$ is found from the experimental data. Estimates of the upper critical field from the initial slope dH_{c2}/dT in this relation yield $\mu_0 H_{c2}(0)=40$ T for H perpendicular to the ab plane and $\mu_0 H_{c2}(0)=120$ T for H parallel to the ab plane.

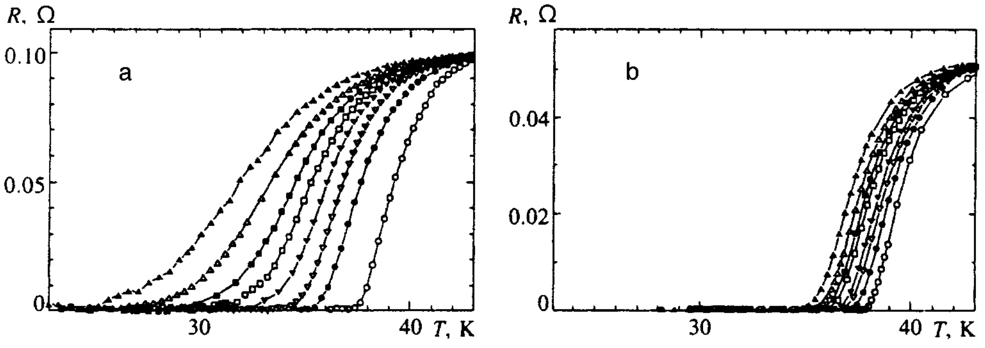

FIG. 2. Temperature dependence of the resistance in various magnetic fields (for H perpendicular (a) and parallel (b) to the ab plane) at various values of $\mu_0 H$: ○—0, ●— 0.5, ▽—1, ▼—2, □—3, ■—4, △—6, ▲—8 T.

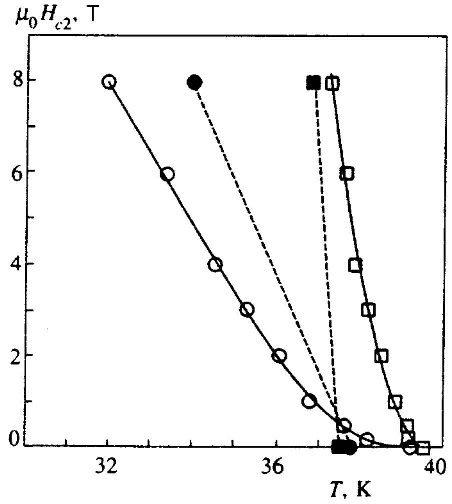

FIG. 3. Temperature dependence of the upper critical magnetic field $H_{c2}(T)$ for the field perpendicular (circles) and parallel (squares) to the ab plane. The points denoted by empty symbols were inferred from the resistance curves of the transition, and those denoted by filled symbols—from the midpoint of the specific heat discontinuity.

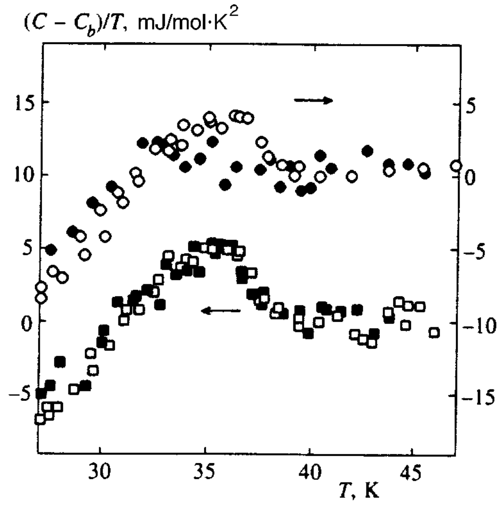

FIG. 4. Temperature dependence of the specific heat near the superconducting transition in magnetic fields $\mu_0 H = 0$ (the points denoted by empty symbols) and $\mu_0 H = 8$ T (the points denoted by filled symbols) for the field perpendicular (circles) and parallel (squares) to the ab plane.

However, the unusual temperature dependence of $H_{c2}(T)$ observed in a number of single crystals differs significantly from the dependence predicted by Eq. (1) and cannot be explained within the scope of this theory. Therefore, Eq. (1) can lead to substantial errors in the estimates of $H_{c2}(0)$, as was shown experimentally[12,13,27] for the system $Ba_{1-x}K_xBiO_y$.

Figure 4 plots the temperature dependence of the specific heat of the $La_{1.85}Sr_{0.15}CuO_4$ sample near T_c for the two field directions after subtracting the three-parameter regular part of the specific heat C_b, represented in the form of a polynomial:

$$C_b = \sum_{n=1}^{3} A_n T^n. \tag{2}$$

The coefficients of the polynomial A_n were determined by least-squares fitting to ensure minimum deviation of the regular part of the specific heat from the experimental points in the region above T_c. Jumps corresponding to the superconducting transitions are visible in the specific heat temperature dependence. The jumps are 2–7 K wide, and the temperatures at their midpoints are in good agreement with the resistance and susceptibility measurements. At temperatures below T_c the magnetic field affects the specific heat, leading to a decrease in the size of the jump and shifting the transition toward lower temperatures, while above T_c the magnetic field does not affect the specific heat.

Figure 5 plots the experimental low-temperature specific heat data in the coordinates C/T vs T^2 in magnetic fields of 0, 1, 2, 4, and 8 T for the magnetic field parallel and perpendicular to the ab plane. As can be seen from the figure, a small magnetic contribution is observed in the specific heat in the vicinity of 2 K, whose magnitude depends on the magnetic field.

To correctly determine the electronic specific heat coefficient $\gamma^*(H)$ in the mixed state, we represented the experimental dependence as a sum of three terms

$$C(T,H) = \gamma^*(H)T + \beta T^3 + nC_{\mathrm{Sch}}(\mu_B H/k_B T), \tag{3}$$

where the first term describes the electronic component of the specific heat, the second describes the phonon component, and the third describes the magnetic component corresponding to the two-level Schottky anomaly, which has the form

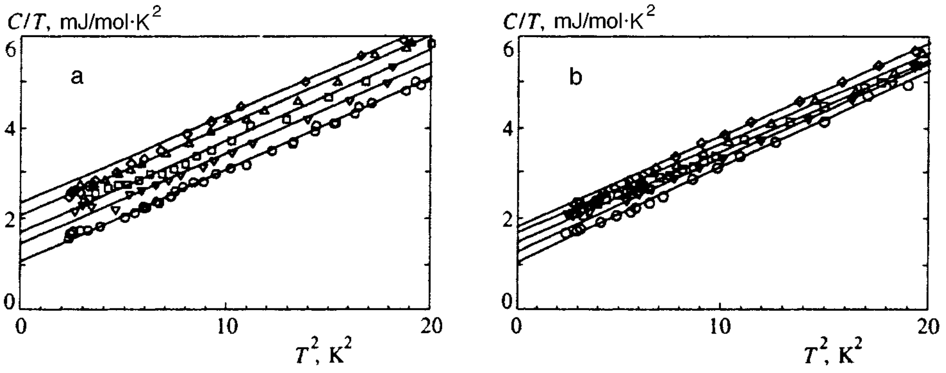

FIG. 5. Temperature dependence of the specific heat of $La_{1.85}Sr_{0.15}CuO_4$ in various magnetic fields (○—0, ▽—1, □—2, △—4, ◇—8 T) for the field H perpendicular (a) and parallel (b) to the ab plane.

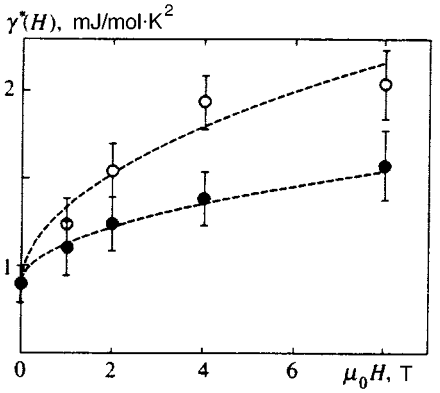

FIG. 6. Magnetic field dependence of the coefficient γ^* for the field H perpendicular (○) and parallel (●) to the ab plane.

$$C_{\mathrm{Sch}}(x)=\frac{x^2e^x}{(1+e^x)^2}. \tag{4}$$

Field-and-temperature dependences of the specific heat $C(T,H)$ were measured in perpendicular and parallel magnetic fields for $H=0$, 1, 2, 4, and 8 T. These series of dependences were described with the help of Eqs. (3) and (4) without any assumptions regarding the functional dependence of $\gamma^*(H)$. The coefficient β, determined in zero field, was fixed for all other magnetic fields. As a result, we obtained the following values of the variable parameters: $\beta=0.202$ mJ/mol·K ($\Theta_D=407$ K), $n=0.2\pm0.1$ mJ/mol·K (it was difficult to determine n more accurately due to the smallness of the magnetic contribution). The maximum magnetic contribution $nC_{\mathrm{Sch}}(\mu H/k_BT)$ in the vicinity of 2 K did not exceed 5% of the total. The value of $\gamma^*(H)$ was determined by least squares in the 3–5-K region, where the magnetic contribution amounted to not more than 1.3% and did not exceed the random error.

In the investigated samples the coefficient $\gamma^*(H)$ in zero field was not large: $\gamma^*(0)=1.0$ mJ/mol·K^2, which can be explained by the presence of a small quantity of metallic nonsuperconducting phase. Besides the above-described technique for determining the amount of superconducting phase in the sample, this quantity can also be estimated by another technique, e.g., by assuming that the electronic specific heat of the nonsuperconducting phase, characterized by the Sommerfeld coefficient γ_n, is equal to the electronic specific heat of $\mathrm{La_{1.85}Sr_{0.15}CuO_4}$ in the normal (nonsuperconducting) state. According to the estimates of Balbashev *et al.*,[21] γ_n is approximately 9 mJ/mol·K^2. Estimation of the fraction of superconducting phase from the relation $\eta=1-\gamma^*(0)/\gamma_n$ also yields a value of η in the range 85–90% for the investigated samples of $\mathrm{La_{1.85}Sr_{0.15}CuO_4}$.

It can be seen from Fig. 5 that the presence of a magnetic field leads to an increase in $\gamma^*(H)$, and the effect of the magnetic field is greater when the field is oriented perpendicular to the ab plane.

Figure 6 plots the field dependence of the electronic specific heat coefficient $\gamma^*(H)$ for two directions of the magnetic field. As can be seen from the figure, for both field orientations this dependence is substantially nonlinear: in the weak-field region $\gamma^*(H)$ grows rapidly with increasing field, while in large fields it varies weakly with the field. Note that this dependence holds for both field directions.

The observed influence of the magnetic field on the specific heat in the low-temperature region is at variance with the predictions of the Abrikosov theory.[16] According to this theory, the specific heat of a superconductor in the mixed state contains a term that is linear in the temperature, $\gamma^*(H)T$. The coefficient $\gamma^*(H)$ grows with the field, reaching at $H=H_{c2}(0)$ the Sommerfeld value γ_n, which is proportional to the electronic density of states at the Fermi level. This theory predicts a nearly linear dependence of $\gamma^*(H)$ and negative curvature of the temperature dependence of the upper critical field H_{c2}.

In the investigated sample, as pointed out above, a nonlinear dependence of $\gamma^*(H)$ is observed, where the temperature dependence of the upper critical field has positive curvature. Note also that both the positive curvature of $H_{c2}(T)$ and the nonlinear dependence of $\gamma^*(H)$ are observed in the same range of magnetic fields. It is of interest to investigate the interrelationship of these two anomalies.

Let us analyze the temperature dependence of the upper critical field $H_{c2}(T)$ and the effect of the magnetic field on the specific heat of a superconductor in the mixed state. Here we consider only the thermodynamic relations; specifically, we note that the transition from the mixed state to the normal state is a second-order transition, the latent heat of the transition is equal to zero, and the entropy S of the superconductor, expressed in terms of the integral of the specific heat

$$S(T')=\int_0^{T'}\frac{C}{T}\,dT$$

at or above the critical temperature T_c, does not depend on the magnetic field. Therefore, the variation of the entropy δS with varying magnetic field vanishes at the superconducting transition temperature:

$$\delta S(T_c)=\int_0^{T_c}\delta\left(\frac{C}{T}\right)dT+\frac{\Delta C(H)}{T_c(H)}\,\delta T_c(H)=0. \tag{5}$$

Here δS, $\delta(C/T)$, and δT_c are the variation of the entropy, specific heat, and critical temperature, due to variation of the magnetic field by δH, and the quantities $\Delta C(H)$ and $T_c(H)$ are respectively the specific heat jump at the transition from the mixed state to the normal state in a magnetic field H and the transition temperature.

Denoting by $\langle d(C/T)/dH\rangle$ the temperature-averaged value of $d(C/T)/dH$,

$$\left\langle\frac{d(C/T)}{dH}\right\rangle=\frac{1}{T_c}\int^{T_c}\frac{d(C/T)}{dH}\,dT, \tag{6}$$

and noting that $\delta(C/T)=\delta H d(C/T)/dH$ and $\delta T_c(H)=\delta H/(dH_{c2}/dT$, after some transformations we obtain

$$\left\langle\frac{d(C/T)}{dH}\right\rangle\frac{dH_{c2}}{dT}=\frac{-\Delta C(H)}{T_c^2(H)}. \tag{7}$$

Introducing the dimensionless parameter p relating the quantities $\langle d(C/T)/dH\rangle$ and $d\gamma^*(H)/dH$, the field derivative of the linear term in the specific heat, by

$$\frac{d\gamma^*(H)}{dH}=p\left\langle\frac{d(C/T)}{dH}\right\rangle, \tag{8}$$

we obtain

$$\frac{d\gamma^*(H)}{dH}\frac{dH_{c2}}{dT}=-p\,\frac{\Delta C(H)}{T_c^2(H)}. \tag{9}$$

Since we used only thermodynamic arguments in the derivation of this relation, it should be valid for any system, and consequently a relationship exists between the anomalies of the temperature dependence of the upper critical field and the field dependence of the specific heat.

If $\Delta C(H)/T_c^2$ and p depend smoothly on the magnetic field, then the abrupt anomalous increase of the temperature derivative of the upper critical field with increasing external magnetic field should be accompanied by an abrupt decrease of the slope of the curve $\gamma^*(H)$. Theoretical models of superconductivity with an anomalous temperature dependence of the critical magnetic field,[2–4,28,29] as a rule, should therefore lead to a nonlinear dependence of the low-temperature specific heat on the field and *vice versa*. Thus, models describing the nonlinear dependence of $\gamma^*(H)$ (see, for example, Volovik[6]), as a rule, should lead to an anomalous temperature dependence of the critical magnetic field.

5. CONCLUSIONS

We have investigated the specific heat in the superconducting single crystal $La_{1.85}Sr_{0.15}CuO_4$ in the mixed state, and also the upper critical magnetic field for two orientations of the field relative to the crystallographic axes, parallel and perpendicular to the ab plane. For both field orientations the temperature dependence of the upper critical field exhibits positive curvature, and the specific heat of the mixed state in the low-temperature region depends nonlinearly on the magnetic field. Neither of these phenomena is described by the traditional theory of the effect of a magnetic field on a superconductor.

We have analyzed the relationship between these two anomalies within the framework of the thermodynamic relations without resort to microscopic or phenomenological theories. We have shown that the temperature dependence of the critical magnetic field and the field dependence of the specific heat are interrelated by the condition of entropy balance.

The abrupt increase in the temperature derivative of the upper critical field with increasing external magnetic field, as a rule, should be accompanied by an abrupt decrease in the slope of the curve $\gamma^*(H)$. The anomalous temperature dependence of the critical magnetic field, as a rule, should therefore lead to a nonlinear dependence of the low-temperature specific heat on the field and *vice versa*.

Note that the observed dependence of the coefficient of the electronic specific heat in the mixed state on the magnetic field (close to a square-root law), when $\gamma^*(H)$ is extrapolated to the normal-state value γ_n, yields a value of $H_{c2}(0)$ substantially greater than the value obtained from the WHH theory. This can lead to revised estimates of the current values of the main parameters of HTSC systems.

ACKNOWLEDGMENTS

We are grateful to A. M. Balbashov and D. A. Shulyat'ev for preparation of the $La_{1.85}Sr_{0.15}CuO_4$ single crystals.

This work was supported by the Scientific Committee on Problems of High-Temperature Superconductors and by Grant No. 96039 of the State Program on High-Temperature Superconductivity.

[1] A. Junod, in *Physical Properties of Type-II High-Temperature Superconductors*, D. Ginzberg (World Scientific, Singapore, 1990).

[2] K. R. Biagi, V. Kogan, and J. R. Clem, Phys. Rev. B **32**, 7165 (1985).

[3] S. Takahashi and M. Tachiki, Phys. Rev. B **33**, 4620 (1986).

[4] A. J. Schofield, Phys. Rev. B **52**, 11733 (1995).

[5] H. Won and K. Maki, Europhys. Lett. **30**, 421 (1995).

[6] G. E. Volovik, JETP Lett. **58**, 469 (1993).

[7] S. V. Zaĭtsev-Zotov, E. A. Protasov, and M. N. Khlopkin, Fiz. Tverd. Tela **26**, 2933 (1984) [Sov. Phys. Solid State **26**, 1772 (1984)].

[8] E. Z. Meĭkhov and V. G. Shapiro, Sverkhprovodimost: Fiz., Khim., Tekh. **4**, 1437 (1991).

[9] A. P. Mackenzie, A. Carrington, S. D. Hughes, R. S. Liu, D. C. Sinclair, S. R. Julian, and G. G. Lonzarich, Phys. Rev. Lett. **71**, 1238 (1993).

[10] A. P. Mackenzie, S. R. Julian, G. G. Lonzarich, A. Carrington, S. D. Hughes, R. S. Liu, and D. C. Sinclair, J. Supercond. **7**, 271 (1994).

[11] M. S. Osofsky, R. J. Soulen, S. A. Wolf, J. M. Bioto, H. Racoto, J. S. Ousset, G. Coffe, S. Askenazy, P. Pari, and J. Bozovic, Phys. Rev. Lett. **71**, 2315 (1993).

[12] C. Escribe-Filippini, J. Marcus, M. Afronte, H. Racoto, J. M. Broto, J. C. Ousset, and S. Askenazy, Physica C **210**, 133 (1993).

[13] V. F. Gantmakher, L. A. Klinkova, N. V. Barkovskii, G. E. Tsydynzhapov, S. Wiegers, and A. K. Geim, Phys. Rev. B **54**, 6133 (1996).

[14] C. J. Gorter, H. Van Beelen, and R. De Bruyn Outboter, Phys. Lett. **8**, 13 (1964).

[15] K. Maki, Phys. Rev. **139**, A702 (1965).

[16] A. A. Abrikosov, *Fundamentals of the Theory of Metals*, North-Holland, Amsterdam (1988).

[17] K. A. Moler, D. J. Baar, J. S. Urbach, Ruixing Liang, W. H. Hardy, A. Kapitulnik, Phys. Rev. Lett. **73**, 2744 (1994).

[18] K. A. Moler, D. J. Baar, Ruixing Liang, W. H. Hardy, A. Kapitulnik, J. Supercond. **8**, 571 (1995).

[19] R. A. Fisher, D. A. Wright, J. P. Emerson, B. F. Woodfield, N. E. Phillips, J. E. Gordon, S. F. Reklis, and E. M. McCarron, Physica C **252**, 237 (1995).

[20] B. Revaz, A. Junod, A. Mirmelstein, A. Erb, J.-Y. Genoud, and G. Triscone, Czech. J. Phys. **46**, Suppl. S3, 1205 (1996).

[21] A. M. Balbashev, D. A. Shulyatev, G. Kh. Panova, M. N. Khlopkin, N. A. Chernoplekov, A. A. Shikov, and A. V. Suetin, Physica C **256**, 371 (1996).

[22] F. C. Chou, J. H. Cho, and D. C. Johnston, Physica C **197**, 303 (1992).

[23] M. N. Khlopkin, N. A. Chernoplekov, and P. A. Cheremnykh, Preprint IAÉ-3549/10, Moscow (1982).

[24] T. Tsuneto, J. Phys. Soc. Jpn. **57**, 3499 (1988).

[25] I. Ikeda, T. Ohmi, and T. Tsuneto, J. Phys. Soc. Jpn. **58**, 1377 (1989).

[26] N. R. Werthammer, E. Helfand, and P. C. Hohenberg, Phys. Rev. **147**, 295 (1966).

[27] G. Kh. Panova, A. A. Shikov, B. I. Savel'ev, A. P. Zhernov, N. V. Anshukova, A. I. Golovashkin, A. I. Ivanova, and A. P. Rusakov, Zh. Éksp. Teor. Fiz. **103**, 605 (1993) [JETP **76**, 302 (1993)].

[28] A. S. Alexandrov, J. Ranninger, and S. Robaszkewicz, Phys. Rev. B **33**, 4526 (1986).

[29] A. S. Alexandrov and J. Ranninger, Solid State Commun. **81**, 403 (1992).

Translated by Paul F. Schippnick

On resistive nonlocal Josephson electrodynamics

V. P. Silin*)

P. N. Lebedev Physics Institute, Russian Academy of Sciences, 117924 Moscow, Russia

(Submitted 14 January 1997)
Zh. Éksp. Teor. Fiz. **112**, 1396–1408 (October 1997)

New solutions of one-dimensional nonlocal Josephson electrodynamics are proposed that describe the steady and nonsteady Abrikosov–Josephson vortex states of the resistive model; these solutions are based on the superposition principle of the vortex structures whose properties are determined by the nonlinear interaction of the vortices. The stability of the current–voltage characteristic (1.13) is shown and the relaxation-oscillation mode of establishing the corresponding state is investigated. The laws governing the annihilation and dispersal of the interacting Abrikosov–Josephson vortices are examined.
[S1063-7761(97)01810-6]

1. The electrodynamics of Josephson junctions are a vigorously developing area of nonlinear physics. The theory of large-scale Josephson vortex structures with a characteristic size much greater than the London length λ is usually based on the sine-Gordon equation or on its extension, which takes dissipation into account (see, for example, Refs. 1 and 2). Powerful general mathematical methods can be used to investigate the sine–Gordon equation (see, for example, Ref. 3). This makes it possible to obtain abundant information about large-scale Josephson structures. Many different numerical investigations have been carried out and approximate approaches have also been developed[4] to take account of dissipative processes and other perturbations in such structures.

Josephson electrodynamics of junctions with a comparatively high critical current density j_c, when the inequality[5]

$$j_c > j_0[\mathrm{A/cm}^2] = \hbar c^2(16\pi|e|\lambda^3)^{-1} \sim 10^4\lambda^{-3}[\mu\mathrm{m}], \tag{1.1}$$

is satisfied, have been attracting particular attention recently. The current density j_0 is less than the dissociation current of Cooper pairs for superconductors with a large Ginzburg–Landau parameter κ. For the conditions when the inequality (1.1) is satisfied Josephson electrodynamics is not only nonlinear but also nonlocal. The existence of this nonlocality was pointed out as far back as in Ref. 6 (see also Ref. 7). Following Ref. 8 for a Josephson junction that is infinite in the (y,z) plane when the phase difference of Cooper pairs on opposite sides of the junction on the coordinate z is one-dimensional, we have the following equation:

$$\sin\varphi + \frac{\beta}{\omega_j^2}\frac{\partial\varphi}{\partial t} + \frac{1}{\omega_j^2}\frac{\partial^2\varphi}{\partial t^2} = \frac{l}{\pi}\frac{\partial}{\partial z}\int_{-\infty}^{\infty} dz' K_0\left(\frac{|z-z'|}{\lambda}\right)\frac{\partial\varphi(z',t)}{\partial z'}. \tag{1.2}$$

Here $K_0(z)$ is a modified Bessel function,

$$\frac{1}{\omega_j^2} = \frac{\hbar C_s}{2|e|j_c} = \frac{\hbar\varepsilon}{16\pi|e|dj_c},$$

$$\frac{\beta}{\omega_j^2} \equiv t_0 = \frac{\hbar}{2|e|j_cR_s} = \frac{\hbar\sigma}{4|e|dj_c},$$

$$l = \frac{\lambda_j^2}{\lambda} = \frac{\hbar c^2}{16\pi|e|j_c\lambda^2}, \tag{1.3}$$

$\hbar$ is the Planck constant, c is the velocity of light, e is the electron charge, $2d$ is the junction width, which is assumed to be small compared with the London length, C_s and R_s are, respectively, the capacitance and resistance of a unit area of the tunnelling junction, and ε and σ are, respectively, the dielectric constant and the conductivity.

In the nonlocal theory, unlike the usual local Josephson electrodynamics, no general mathematical approach has been formulated. The description of only some vortex states has been obtained thus far (see below). Therefore, it is important to analyze the general properties of the results of nonlocal electrodynamics that have been obtained thus far, thereby permitting the possibility of searching for new vortex states. A nonlinear superposition principle is formulated below, according to which the expressions for the Cooper-pair phase difference φ can be represented in the form of a linear superposition of spatially nonuniform and uniform states, the temporal evolution and steady-state properties of which are determined by the nonlinear interaction of the vortices.

In this paper we will consider the consequences following from the case of a negligibly small capacitance when the second derivative with respect to time in Eq. (1.2) can be ignored. This is the so-called resistive model,[1] which corresponds in the English-language literature to the term ''resistively shunted junction (RSJ) model.''[7] This model is most widely applicable at temperatures near the superconducting phase transition. In addition, we will be interested in the conditions (1.1) of a Josephson junction with a large critical current density when

$$\lambda \gg \lambda_j. \tag{1.4}$$

In this case, according to Ref. 6 (see also Ref. 9), the asymptotic form

$$K_0\left(\frac{|z-z'|}{\lambda}\right) \simeq -\ln\frac{|z-z'|}{2\lambda}, \tag{1.5}$$

can be used in Eq. (1.2). Then we write the following equation:

$$\sin\varphi + \frac{\partial\varphi}{\partial\tau} - \frac{l}{\pi}\int_{-\infty}^{\infty}\frac{dz'}{z'-z}\frac{\partial\varphi(z',t)}{\partial z'} = \gamma. \tag{1.6}$$

Here the integral is understood in the sense of the principal Cauchy value, $\tau = t/t_0$ (1.3), and the term $\gamma = j/j_c$, where j is the density of the current that is shorting the tunnelling junction, is added to the right side of Eq. (1.6) compared with (1.2). Below it will be assumed that j does not depend on the spatial coordinate. A time-independent solution of this equation has already been given in Ref. 10 for $\gamma=0$, which describes a single Abrikosov–Josephson vortex:

$$\varphi = \pi + 2\,\mathrm{arctanh}(z/l). \tag{1.7}$$

The length l characterizes the size of the Abrikosov–Joesphson vortex. This solution was obtained independently later in Ref. 9. The relaxation process for the establishment of such a vortex in time (for $\gamma=0$) has been described in Ref. 11. The solution of Eq. (1.6) for $\gamma\neq 0$ was obtained in Ref. 12, and it describes a single Abrikosov–Josephson vortex moving with a constant velocity:

$$\varphi(z,t) = \pi + \sin^{-1}\gamma + 2\tan^{-1}\left(\frac{z+vt}{l/\sqrt{1-\gamma^2}}\right), \tag{1.8}$$

where $v = (\gamma l/t_0\sqrt{1-\gamma^2})$. The relaxation process for establishing such a nonlinear state has been described in Ref. 13. The relaxation processes for the establishment of the periodic structure of Abrikosov–Josephson vortices with a nonzero average magnetic field and for the decay of a vortex structure with a zero average magnetic field were also considered in Ref. 14. Finally, a time-dependent solution of Eq. (1.6) was obtained in Ref. 15 that describes an infinite chain of Abrikosov–Josephson vortices with a nonzero average magnetic field, traveling with a constant velocity:

$$\varphi(z,t) = \theta_0 + \pi + 2\tan^{-1}\left(\frac{\tan[(z+vt)/2L]}{\tanh[\alpha_0/2]}\right), \tag{1.9}$$

where

$$\sin\theta_0\cosh\alpha_0 = \gamma, \quad \cos\theta_0\sinh\alpha_0 = (l/L), \tag{1.10}$$

$$v = \frac{\omega_j^2 L}{\beta}\left\{\left(\frac{1}{4}\left[\gamma^2 + \frac{l^2}{L^2} - 1\right]^2 + \frac{l^2}{L^2}\right)^{1/2} + \frac{1}{2}\left[\gamma^2 - \frac{l^2}{L^2} - 1\right]\right\}^{1/2}. \tag{1.11}$$

The solution (1.9) is also realized for large j values, unlike the case of a steady-state moving single Abrikosov–Josephson vortex, which is able to exist for $\gamma^2<1$, i.e., when the current density is less than the critical value j_c. The current–voltage (I–V) characteristic

$$\frac{V^2}{R_s^2} \equiv j_r^2 = j_c^2\left\{\left(\frac{1}{4}\left[\frac{j^2}{j_c^2} + \frac{l^2}{L^2} - 1\right]^2 + \frac{l^2}{L^2}\right)^{1/2} + \frac{1}{2}\left[\frac{j^2}{j_c^2} - \frac{l^2}{L^2} - 1\right]\right\} \tag{1.12}$$

or in a different form

$$j^2 = j_r^2 + \frac{L^2 j_c^2 j_r^2}{L^2 j_r^2 + l^2 j_c^2}, \tag{1.13}$$

corresponds to Eqs. (1.9) and (1.11). A comparison of the analytically obtained I–V characteristic (1.12), (1.13) with the results of numerical calculations in Ref. 10 makes it possible not only to discern a certain similarity but also to establish the closeness of the results of Ref. 10. The second term of Eq. (1.13) gives an extremely simple analytical expression for the excess superconducting current in the ohmic region, where $j>j_r$. Finally, since the average magnetic field $\overline{H}$, which corresponds to a traveling chain of vortices (1.9), is related to the parameter L by the expression

$$\overline{H} = \frac{\Phi_0}{4\pi\lambda L} \equiv \frac{\hbar c}{4|e|\lambda L}, \tag{1.14}$$

Eqs. (1.12) and (1.13) give the analytical dependence of the voltage and current on the magnetic field. However, the question of the stability of the traveling chain of vortices (1.9) is still not answered.

Observations on the solutions of Eq. (1.6) and, in particular, those which describe time-dependent resistive states,[11–15] make it possible to assume that there exists that class of solutions among the possible nonlinear solutions of the resistive nonlocal Josephson electrodynamics which can be said to be superpositional. In this case the nonlinear state, which is characterized by the phase difference of the Cooper pairs, is described by a linear combination of phases, each of which is outwardly similar to the separate nonlinear solution of Eq. (1.6). At the same time the parameters that characterize such separate states turn out to be interdependent and are determined by the interaction of the vortex states. This interaction determines, in particular, the evolution of the nonlinear system of interacting vortices with time. The implementation of such a superpositional approach is demonstrated below. The relaxation-oscillation law for establishing a steady traveling chain of vortices (1.9) shows up most clearly within the framework of the resistive (or RSJ) model.

Another question associated with the dynamics of Abrikosov–Josephson vortices, which is considered below, is related to an investigation of the interaction of a vortex and an antivortex—a fluxon and an antifluxon. (Compare Ref. 16 in the case of local Josephson electrodynamics.) For the case when the current j is absent ($\gamma=0$) the possibility of the annihilation of the vortices and the possibility of their dispersal in the presence of a steady-state phase come out of the linear superposition of the solutions describing the interacting vortices. The presence of a spatially uniform and constant current j ($\gamma=$const), as shown below, allows stationary states to exist, corresponding to a linear superposition of a fluxon and antifluxon, as well as a linear superposition of an infinite chain of vortices and antivortices. These new stationary spatially nonuniform current states correspond to the well-known portion of the I–V characteristic of a Josephson junction, for which there is a nonzero superconducting current with less than critical density $j<j_c$ when the voltage is equal to zero ($V=0$). In other words, these spatially nonuni-

form solutions must be considered along with the spatially uniform solutions of Eq. (1.6) (and more precisely (1.2)) $\varphi=\theta(t)$:

$$\sin\theta+\frac{d\theta}{d\tau}=\gamma, \tag{1.15}$$

when the known solution

$$\theta=\sin^{-1}\gamma, \quad \gamma^2<1, \tag{1.16}$$

exists in the steady-state case corresponding to zero voltage $(V=0)$.

We consider the dynamics of the exact solutions of the resistive model, representing a linear superposition of the spatially uniform phase of a fluxon and antifluxon which obeys Eq. (1.15). The nonlinearity of the resistive model appears in the nonlinear relationship of the phase $\theta(t)$ with the coordinates of the centers of the fluxon and antifluxon $z_0(t)$ and with the characteristic vortex size $\rho(t)$. Just as in the absence of a current, both the vortex annihilation and the vortex dispersal situations are possible. The possibility of a similar linear superposition of solutions is also established for vortex chains. The law describing the trend with time of the annihilation process of the vortex and antivortex chains, which is accompanied by a comparatively slow dispersal of the individual vortices of the chains, is also established.

2. We begin out presentation of the results of the resistive dynamics of Abrikosov–Josephson vortices with the case of the relaxation solution of Eq. (1.6), describing the establishment of a traveling chain of vortices (1.9). The corresponding solution of Eq. (1.6) is

$$\varphi(z,\tau)=\theta(\tau)+\pi+2\tan^{-1}\left[\frac{\tan\left(\dfrac{z+z_0(\tau)}{2L}\right)}{\tanh(\alpha(\tau)/2)}\right]. \tag{2.1}$$

If the following Hilbert transformation is used here ($\alpha>0$, $A>0$):[14]

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{dx}{x-y}\,\frac{\sinh\alpha}{\cosh\alpha-\cos Ax}=\frac{\sin Ay}{\cos Ay-\cosh\alpha}, \tag{2.2}$$

then it is easy to see that Eq. (2.1) satisfies Eq. (1.6) if the functions θ, z_0 and α satisfy the following ordinary nonlinear differential equations:

$$\frac{d\theta}{d\tau}=\gamma-\sin\theta\cosh\alpha. \tag{2.3}$$

$$\frac{d\alpha}{d\tau}=\frac{l}{L}-\cos\theta\sinh\alpha. \tag{2.4}$$

$$\frac{1}{L}\frac{dz_0}{d\tau}=\sin\theta\sinh\alpha. \tag{2.5}$$

The system of Eqs. (2.3), (2.4) defines the interaction of the spatially uniform phase θ and finite vortex chain, described by an arctangent function; this is expressed in the nonlinear differential relationship of θ and α. In addition to this, Eq. (2.5) also characterizes the nonlinear action of the field of the spatially uniform phase on the movement of the vortex chain as time passes $z_0(\tau)$. The traveling chain of Abrikosov–Josephson vortices of Ref. 15 is obtained for the condition γ=const. In this case the system of Eqs. (2.3) and (2.4) corresponds to an autonomous dynamic system. This makes it possible to find the law governing the evolution of the nonsteady-state vortex chain (2.1) with time by analyzing the phase portrait of the equation

$$\frac{d\alpha}{d\theta}=\frac{(l/L)-\sinh\alpha\cos\theta}{\gamma-\cosh\alpha\sin\theta}, \tag{2.6}$$

in the (α,θ) plane. It is easy to see that in the (α,θ) phase plane Eq. (2.6) has only one singular (or stationary) point, corresponding to Eq. (1.10). In other words, Eq. (1.9) is the only steady-state solution of the autonomous system (2.3), (2.4). The solution of Eq. (2.5) corresponding to the stationary point (1.10) $z_0=v\tau$, where $v=L\sinh\alpha_0\sin\theta_0$, gives the expression of the constant velocity (1.11) after converting to the dimensional time $t=\tau t_0$.

Near the stationary point (1.10), when we have $\alpha=\alpha_0+\Delta\alpha$, $\theta=\theta_0+\Delta\theta$, and $\Delta\alpha$ and $\Delta\theta$ are small, Eq. (2.6) reduces to the following:

$$\frac{d\Delta\theta}{d\Delta\alpha}=\frac{\sin\theta_0\sinh\alpha_0\Delta\alpha+\cos\theta_0\cosh\alpha_0\Delta\theta}{\cos\theta_0\cosh\alpha_0\Delta\alpha-\sin\theta_0\sinh\alpha_0\Delta\theta}. \tag{2.7}$$

In accordance with the usual classification of singular points (see Refs. 17, 18) the stationary point (1.10) is a focus according to Eq. (2.7). The fact that the focus is stable can be seen directly from the dependence of $\Delta\alpha$ and $\Delta\theta$ on time using the relation

$$\exp(-\tau\cos\theta_0\cosh\alpha_0)\cos(\tau\sin\theta_0\sinh\alpha_0+\text{const}). \tag{2.8}$$

Consequently, it can be proved that within the framework of the resistive model the solution (1.9) is stable, and the relaxation process for establishment of a chain of Abrikosov–Josephson vortices traveling with a constant velocity is described by the time-dependent solution (2.1). In accordance with Eqs. (1.10) we have the following expressions for the frequency $\Omega=t_0^{-1}\sin\theta_0\sinh\alpha_0$ and damping rate $\Gamma=t_0^{-1}\cos\theta_0\cosh\alpha_0$:

$$\Omega=\frac{1}{t_0}\left\{\frac{1}{2}\left[\gamma^2-1-\frac{l^2}{L^2}\right]+\left(\frac{1}{4}\left[\gamma^2+\frac{l^2}{L^2}-1\right]^2+\frac{l^2}{L^2}\right)^{1/2}\right\}^{1/2}. \tag{2.9}$$

$$\Gamma=\frac{1}{t_0}\left\{\frac{1}{2}\left[1-\gamma^2+\frac{l^2}{L^2}\right]+\left(\frac{1}{4}\left[\gamma^2+\frac{l^2}{L^2}-1\right]^2+\frac{l^2}{L^2}\right)^{1/2}\right\}^{1/2}, \tag{2.10}$$

which characterize relaxation oscillations. Remembering the I–V characteristic in the form (1.12), it is easy to see that Eq. (2.9) can be written as

$$\Omega=\frac{2|e|V}{\hbar}. \tag{2.11}$$

This relationship has been confirmed in an experiment with the AC Josephson effect.[1] Moreover, it must be emphasized that it is only in the limit.

$$\gamma^2 - 1 \gg (l/L)^2 \tag{2.12}$$

that Eq. (2.9) becomes the well-known formula of Aslamazov and Larkin,[19]

$$\Omega = \frac{2|e|R_s}{\hbar}\sqrt{j^2 - j_c^2}. \tag{2.13}$$

In this case we have

$$\Gamma = \frac{c^2 R_s}{8\pi\lambda^2 L} = \frac{c R_s \overline{H}|e|}{2\pi\lambda\hbar}, \tag{2.14}$$

for the damping rate. Equation (2.11) is obviously general, corresponding to an energy $\hbar\Omega = 2|e|V$ of a Cooper pair in the field of a static potential difference V.

According to Eqs. (2.9) and (2.10) the frequency and damping rate are interrelated by the simple expression

$$\Gamma^2 = \Omega^2 - \left(\gamma^2 - 1 - \frac{l^2}{L^2}\right)\frac{1}{t_0^2}. \tag{2.15}$$

It is seen from this, in particular, that for $(j/j_c)^2 = \gamma^2 = 1 + (l/L)^2$ the damping rate is equal to the frequency. For larger values of the current density j the damping rate is less than the frequency. The following general dependence of the damping rate on voltage can be written:

$$\Gamma = \frac{R_s c^2}{8\pi\lambda^2 L}\left\{1 + \left[\frac{l^2}{L^2} + \frac{V^2}{R_s^2 j_c^2}\right]^{-1}\right\}^{1/2}. \tag{2.16}$$

This makes it possible to obtain the following simple relation:

$$\frac{\Gamma}{\Omega} = \frac{R_s j_c}{V}\frac{l}{L}\sqrt{1 + \left[\frac{l^2}{L^2} + \frac{V^2}{R_s^2 j_c^2}\right]^{-1}}. \tag{2.17}$$

In particular, if

$$V \gg R_s j_c l/L, \tag{2.18}$$

then

$$\frac{\Gamma}{\Omega} = \frac{R_s j_c}{V}\frac{l}{L}\sqrt{1 + \frac{R_s^2 j_c^2}{V^2}}. \tag{2.19}$$

This last formula allows one to see the possibility of realizing weakly damped relaxation oscillations and, therefore, it indicates the possibility for the effective excitation of such oscillations with a frequency exceeding $\Omega_0 = (2|e|R_s j_c/\sqrt{l/L}\hbar)$ in a Josephson junction with a traveling chain of Abrikosov–Josephson vortices (1.9).

In the limit opposite to Eq. (2.18) the damping rate does not depend on the voltage $\Gamma(V=0)$. If $l \ll L$ in this case, then $\Gamma = (2|e|/\hbar)R_s j_c$ and the damping rate turns out to be much greater than the frequency.

3. Proceeding now to a consideration of states corresponding to a linear superposition of vortices and antivortices let us note first of all the case when there is no shorting current (γ=0). In this case we write the solution of Eq. (1.6) as

$$\varphi(z,\tau) = 2\tan^{-1}\left[\frac{z + z_0(\tau)}{\rho(\tau)}\right] - 2\tan^{-1}\left[\frac{z - z_0(\tau)}{\rho(\tau)}\right]. \tag{3.1}$$

Using a Hilbert transformation (Ref. 20, p. 173, Eq. 15.2.10)

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{dx}{x-y}\,\frac{1}{x^2+y^2} = -\frac{y}{a(y^2+x^2)}, \quad a>0, \tag{3.2}$$

it is easy to prove that Eq. (3.1) is the solution of Eq. (1.6) if the functions $z_0(\tau)$ and $\rho(\tau)$ obey the equations

$$\frac{dz_0}{d\tau} = -\frac{\rho^2 z_0}{\rho^2 + z_0^2}, \quad \frac{d\rho}{d\tau} = l - \frac{\rho z_0^2}{\rho^2 + z_0^2}. \tag{3.3}$$

Equation (3.1) corresponds to the superposition of a fluxon (2π-kink) and antifluxon (or antikink), carrying single magnetic flux quanta with opposite signs and separated from each other by a distance $2z_0(\tau)$. The function $\rho(\tau)$ characterizes the vortex size. The magnetic field corresponding to the solution (3.1) (compare Ref. 21) is a linear superposition of two vortices:

$$H_y(x,z,\tau) = \frac{\hbar c\rho(\tau)}{2\pi e\lambda^2}\int_{-\infty}^{\infty}\frac{d\xi}{\xi^2 + \rho^2(\tau)} \times\left\{K_0\left(\frac{1}{\lambda}\sqrt{[z + z_0(\tau) - \xi]^2 + x^2}\right) - K_0\left(\frac{1}{\lambda}\sqrt{[z - z_0(\tau) - \xi]^2 + x^2}\right)\right\}. \tag{3.4}$$

At distances much greater than the size $\rho(\tau)$ of the vortex core, $(z \pm z_0)^2 + x^2 \gg \rho^2$,

$$H_y(x,z,t) = \frac{\hbar c}{2e\lambda}\left\{K_0\left(\frac{1}{\lambda}\sqrt{(z+z_0)^2 + x^2}\right) - K_0\left(\frac{1}{\lambda}\sqrt{(z-z_0)^2 + x^2}\right)\right\}, \tag{3.5}$$

this represents the field of an Abrikosov vortex and antivortex and conforms to a corresponding combination of a vortex and Abrikosov–Josephson vortex at sufficiently large distances. Conversely, if $(z \pm z_0)^2 + x^2 \ll \rho^2$, we have from Eq. (3.4)

$$H_y(x,z,t) = \frac{\hbar c}{4e\lambda^2}\ln\frac{(|x| + \rho)^2 + (z - z_0)^2}{(|x| + \rho)^2 + (z + z_0)^2}. \tag{3.6}$$

It follows from Eq. (3.3) that with an increase in time z_0^2 decreases; this corresponds, according to Eqs. (3.5) and (3.6), to a decrease of the magnetic field. Here if $z_0^2 \gg \rho^2$ holds, then the characteristic scale (or core) of the vortex approaches the value l so that $\rho - l \sim \exp(-\tau)$, and z_0^2 decreases in accordance with the linear law $2l^2\tau$. If, however, the kink and antikink are sufficiently close together, so that $z_0^2 \ll \rho^2$, then according to Eq. (2.3) the characteristic scale of the vortex increases in accordance with the linear law $\rho \sim l\tau$, and the distance between the vortex and antivortex decreases rapidly in accordance with the exponential law $\exp(-\tau)$. These relations correspond to the annihilation of two Abrikosov–Josephson vortices, the magnetic field of which vanishes as z_0 goes to zero according to Eq. (3.4).

Let us now give another solution of Eq. (1.6) for γ=0:

$$\varphi(z,\tau)=\pi+2\tan^{-1}\frac{z+z_0(\tau)}{\rho(\tau)}-2\tan^{-1}\frac{z-z_0(\tau)}{\rho(\tau)}, \tag{3.7}$$

which, like the solution (3.1), describes a fluxon and antifluxon. Here the magnetic field of the linear superposition (3.7) is described just as in the case of Eq. (3.1), by the formula (3.4). The only difference between the solutions (3.7) and (3.1) is the presence of the constant phase π in Eq. (3.7). This difference, however, leads to a nonlinear variation of the ordinary differential equations describing the time evolution of the distance between vortices in Eq. (3.7) and the time evolution of the core $\rho(\tau)$ of the vortices. Substitution of Eq. (3.7) into Eq. (1.6) and use of the transformation (2.2) yield the following equations, which are different from Eq. (3.3):

$$\frac{dz_0}{d\tau}=\frac{z_0\rho^2}{z_0^2+\rho^2}, \quad \frac{d\rho}{d\tau}=l+\frac{z_0^2\rho}{z_0^2+\rho^2}. \tag{3.8}$$

It obviously follows from this that the distance between kink and antikink in Eq. (3.7) increases with an increase in time. In other words, the solution (3.7), unlike the annihilation solution (3.1), describes a dispersal of two vortices with their simultaneous spreading.

4. Continuing the discussion of the nonlinear behavior of the properties of vortices in the properties of the solutions of Eq. (1.6), which represent linear superpositions of vortices, let us now turn to the case of a nonzero shorting current ($\gamma\neq 0$). Here, first of all, let us point out a time-independent solution that describes an fluxon (2π-kink) and antifluxon (or anti-kink) pair at rest:

$$\varphi(z)=\sin^{-1}\gamma+2\tan^{-1}\left(\left[\frac{z}{l}+\frac{1}{\gamma}\right]\sqrt{1-\gamma^2}\right)$$
$$-2\tan^{-1}\left(\left[\frac{z}{l}-\frac{1}{\gamma}\right]\sqrt{1-\gamma^2}\right). \tag{4.1}$$

The integral transformation (2.2) easily makes it possible to prove that Eq. (4.1) satisfies Eq. (1.6). On the other hand, the solution (4.1) is completely analogous to the corresponding known solution of the static analog of Eq. (1.6), which was obtained in the Peierls–Nabarro theory of dislocations.[22] The magnetic field corresponding to an vortex-antivortex pair (4.1) is given by Eqs. (3.4)–(3.6), in which the characteristic core size of the vortex is $\rho=(l/\sqrt{1-\gamma^2})$, and the distance between vortices is $2z_0=2l/\gamma$. With a decrease in current the distance between steady-state vortices given by (4.1) increases, and the structure of the magnetic field of each vortex approaches the structure of an isolated vortex since ρ approaches l.

The nonsteady-state extension of the solution (4.1), corresponding to the nonlinear state of two interacting Abrikosov–Josephson vortices for $\gamma\neq 0$, can be represented by

$$\varphi(z,t)=\theta(\tau)+2\tan^{-1}\left[\frac{z+z_0(\tau)}{\rho(\tau)}\right]-2\tan^{-1}\left[\frac{z-z_0(\tau)}{\rho(\tau)}\right]. \tag{4.2}$$

Here the function $\theta(\tau)$ satisfies Eq. (1.15). Therefore, Eq. (4.2) can be interpreted as describing the linear superposition of two vortices superimposed on a spatially uniform current state. The magnetic field is again described by Eqs. (3.4)–(3.6). However, the functions ρ and z_0 now obey the following system of nonlinear ordinary differential equations:

$$\frac{d\rho}{d\tau}=l-\frac{\rho z_0}{z_0^2+\rho^2}\left[\rho\sin\theta+z_0\cos\theta\right], \tag{4.3}$$

$$\frac{dz_0}{d\tau}=\frac{\rho z_0}{z_0^2+\rho^2}\left[z_0\sin\theta-\rho\cos\theta\right]. \tag{4.4}$$

In the special cases of $\theta=0$ and $\theta=\pi$, which correspond to the steady-state solution of Eq. (1.15), we have the formal transformation of these equations into Eqs. (3.3) and (3.8), respectively.

In the case of a constant current $\gamma^2<1$ holds and the solution (1.16) of Eq. (1.5) exists, the system of equations corresponds to an autonomous second-order dynamic system. Correspondingly, it follows from Eqs. (4.3) and (4.4) that

$$\frac{d\rho}{dz_0}=\frac{z_0^2(l-\rho\cos\theta)+\rho^2(l-z_0\sin\theta)}{\rho z_0(z_0\sin\theta-\rho\cos\theta)}. \tag{4.5}$$

This equation makes it possible to see one singular (or stationary) point of the dynamic system on the (ρ,z_0) phase plane

$$\rho\cos\theta=l, \quad z_0\sin\theta=l. \tag{4.6}$$

This point corresponds to the stationary state of the two interacting vortices described by Eq. (4.1). Near the stationary point, when

$$\rho=(l/\cos\theta)+\Delta\rho, \quad z_0=(l/\sin\theta)+\Delta z_0,$$

and $\Delta\rho$ and Δz_0 are small, Eq. (4.5) yields

$$\frac{d(\Delta\rho\cos\theta)}{d(\Delta z_0\sin\theta)}=\frac{\Delta z_0\sin\theta+\cot^2\theta(\Delta\rho\cos\theta)}{\Delta\rho\cos\theta-\Delta z_0\sin\theta}. \tag{4.7}$$

According to the usual classification of singular points on a phase plane,[17,18] the stationary point (4.6), as follows from Eq. (4.7), is a saddle point. The stationary point (4.7) turns out to be unstable. Two qualitatively different conditions can be seen far from the singular point. First of all, if the vortex and antivortex are sufficiently close to each other so that $\rho\sin\theta\gg z_0\cos\theta$ and $\rho\cos\theta\gg z_0\sin\theta$, then the further evolution of the approach of the vortices occurring in this case is given by the law $z_0\sim\exp(-\tau\sqrt{1-\gamma^2})$. In this case if $z_0\sin\theta<l$, a comparatively slow separation of the vortices occurs in accordance with the law $\rho\sim l\tau$. It is seen that the presence of a nonzero current ($\gamma\neq 0$) delays vortex annihilation. Secondly, when the vortex and antivortex are sufficiently far apart ($z_0\cos\theta\gg\rho\sin\theta$, $\rho\cos\theta\ll z_0\sin\theta$), further dispersal of the vortices occurs. In this case the evolution of the interacting vortex and antivortex is determined by the fact that the size of their cores approaches $(l/\sqrt{1-\gamma^2})$, and the distance between them varies as $(2\gamma l/\sqrt{1-\gamma^2})\tau$. In other words, the asymptotic dispersal mode corresponds to

$$\varphi_{ass}(z,t)=\sin^{-1}\gamma+2\tan^{-1}\left(\frac{z+vt}{l}\sqrt{1-\gamma^2}\right)$$

$$-2\tan^{-1}\left(\frac{z-vt}{l}\sqrt{1-\gamma^2}\right), \tag{4.8}$$

where the asymptotic dispersd rate of the vortices is equal to

$$v=\frac{2|e|R_s jl}{\hbar\sqrt{1-(j/j_c)^2}}=\frac{c^2R_s}{8\pi\lambda^2}\frac{j}{\sqrt{j_c^2-j^2}}, \tag{4.9}$$

which corresponds to the velocity of the uniform motion of a single Abrikosov–Josephson vortex (1.8).

5. Let us consider another class of solutions of Eq. (1.6) for $\gamma\neq 0$, which makes it possible to see a linear superposition of spatial structures whose quantitative characteristics as well as the evolution with time exhibit a highly nonlinear character. Note, first of all, the static solution corresponding to two superposed periodic chains of vortices and antivortices superposed on a spatially uniform phase (1.16), when the condition $\gamma^2<1$ determined by the current is satisfied,

$$\begin{aligned}\varphi(z)={}&\sin^{-1}\gamma\pm 2\tan^{-1}\left\{\left[\frac{L}{l}\sqrt{l-\gamma^2}\right.\right.\\&\left.\left.+\sqrt{\frac{L^2}{l^2}(1-\gamma^2)-1}\right]\tan\left(\frac{z}{2L}+\frac{1}{2}\tan^{-1}\frac{1}{\gamma L}\right)\right\}\\&\mp 2\tan^{-1}\left\{\left[\frac{L}{l}\sqrt{l-\gamma^2}+\sqrt{\frac{L^2}{l^2}(1-\gamma^2)-1}\right]\right.\\&\left.\times\tan\left(\frac{z}{2L}-\frac{1}{2}\tan^{-1}\frac{l}{\gamma L}\right)\right\}.\end{aligned} \tag{5.1}$$

Note that this current solution, just like the steady-state solution (4.1) and the solution (1.16), corresponds to zero voltage on the I–V characteristic of the Josephson junction. The solutions (5.1) and (4.1), unlike (1.16), are highly spatially nonuniform. Equation (5.1) can be considered as a linear superposition; this follows from the form of the magnetic field. At the same time, the spatially uniform phase exerts a highly nonlinear effect on the properties of the spatially nonuniform periodic vortex chains.

The nonlinear influence of the interacting vortices is also obvious in the case of the time dependent analog of the solution (5.1):

$$\begin{aligned}\varphi(z,\tau)={}&\theta(\tau)+2\tan^{-1}\left(\frac{\tan\left[\dfrac{z+z_0(\tau)}{2L}\right]}{\tanh[\alpha(\tau)/2]}\right)\\&-2\tan^{-1}\left(\frac{\tan\left[\dfrac{z-z_0(\tau)}{2L}\right]}{\tanh[\alpha(\tau)/2]}\right).\end{aligned} \tag{5.2}$$

Here the function $\theta(\tau)$ is defined by Eq. (1.15). Therefore, it can be stated that the chains of vortices and antivortices, corresponding to the other two terms of Eq. (5.2), have no effect on the spatially uniform phase θ. Conversely, the influence of the phase θ on the spatially periodic vortex chains is significant and is nonlinear; this is seen from the equations which are obeyed by the functions z_0 and α:

$$\begin{aligned}&\frac{1}{L}\frac{dz_0}{d\tau}\\&=\frac{\sinh\alpha\sin\dfrac{z_0}{L}\left[\sin\theta\cosh\alpha\sin\dfrac{z_0}{L}-\cos\theta\sinh\alpha\cos\dfrac{z_0}{L}\right]}{\cosh^2\alpha-\cos^2(z_0/L)},\end{aligned} \tag{5.3}$$

$$\begin{aligned}\frac{d\alpha}{d\tau}={}&\frac{l}{L}-\frac{\sinh\alpha\sin(z_0/L)}{\cosh^2\alpha-\cos^2(z_0/L)}\left[\sin\theta\sinh\alpha\cos\frac{z_0}{L}\right.\\&\left.+\cos\theta\cosh\alpha\sin\frac{z_0}{L}\right].\end{aligned} \tag{5.4}$$

If the current j does not depend on time and $\gamma^2<1$ holds, then Eqs. (5.3) and (5.4) correspond to an autonomous dynamic system, an analysis of the properties of which can be carried out on the (α,z_0) phase plane using the equation

$$L\frac{d\alpha}{dz_0}=\frac{\dfrac{l}{L}\left[\cosh^2\alpha-\cos^2\dfrac{z_0}{L}\right]-\sinh\alpha\sin\dfrac{z_0}{L}\left[\sin\theta\sinh\alpha\cos\dfrac{z_0}{L}+\cos\theta\cosh\alpha\sin\dfrac{z_0}{L}\right]}{\sinh\alpha\sin\dfrac{z_0}{L}\left[\sin\theta\cosh\alpha\sin\dfrac{z_0}{L}-\cos\theta\sinh\alpha\cos\dfrac{z_0}{L}\right]}. \tag{5.5}$$

It follows from this, in particular, that a stationary (or singular) point exists for

$$\tan\alpha=\frac{l}{L\sqrt{1-\gamma^2}},\quad \frac{z_0}{L}=\tan^{-1}\frac{l}{\gamma L}, \tag{5.6}$$

which corresponds to the steady-state solution (5.1).

On the other hand, the system of Eqs. (5.3) and (5.4) makes it possible to trace the annihilation process of the chain of vortices and antivortices, which corresponds to a decrease of the distance $2z_0$ between chains with an increase in time. In particular, when the vortex chains turn out to be quite close, $z_0\ll L$, Eqs. (5.3) and (5.4) become

$$\frac{dz_0}{d\tau}=-z_0\cos\theta,\quad \frac{d\alpha}{d\tau}=\frac{l}{L}-\frac{z_0}{L}\sin\theta. \tag{5.7}$$

These equations are like those describing the evolution of a vortex–antivortex pair. An exponentially rapid decrease of z_0 with an increase in time follows from them, and a slower decay of the vortex occurs for $z_0\sin\theta\ll l$ when the size of its core increases linearly.

6. In summarizing all that has been presented above, it is first necessary to emphasize that the observation of the previously known solutions of the resistive model for both separate Abrikosov–Josephson vortices (1.7) and (1.8) and also for chains of these vortices (1.9) makes it possible to represent the superposition property used in our paper in the form of the spatially uniform and nonuniform phases in the case of Eqs. (1.7)–(1.9) and (2.1) as well as that of the nonuniform phases (3.1), (3.7), (4.1), (4.2), (4.8), (5.1), (5.2). In all of the cases considered the nonlinear character of the vortex structures is determined by those nonlinear relationships which determine the spatial dimensions of the structures and which determine in steady-state cases the evolution of the interacting vortices with time.

Among the relationships presented let us mention the annihilation of Abrikosov–Josephson vortices, their dispersal and spreading, both in the case of individual vortices and also in the case of periodic structures. Let us also mention the stability of a traveling vortex chain (1.9) with the I–V characteristic (1.12), (1.13), which was shown above, and also the relaxation-oscillation properties for the establishment of such a traveling chain.

Let us make the following comment with regard to the possible experimental manifestation of the properties of nonlocal Josephson electrodynamics. While the condition (1.4) (or the condition (1.1) in expanded form corresponding to it) can require extreme values of the Josephson critical current density for its realization in the case of a tunnelling junction between massive superconductors, the situation is simplified for a Josephson junction in a thin film. Actually, if the film thickness D turns out to be less than the London length ($D \ll \lambda$), $\lambda_e = \lambda^2/D$ plays the role of the effective Meissner shielding depth. Then, as is known (see, for example, Refs. 9 and 10), the condition $\lambda_j \ll \lambda_e$ must be satisfied in order to realize the clearly exhibited limit of nonlocal Josephson electrodynamics. This leads to a decrease of the critical Josephson current density, which requires using the nonlocal approach, compared with the j_0 of Eq. (1.1) by a factor of D/λ; (Ref. 9) this broadens the area of experimental study of the effects in question.

However, even without such an extension to the case of thin films, our consideration of the properties of junctions between massive superconductors opens up the possibility of describing the situation of fairly strong magnetic fields. Thus, the I–V characteristic (1.12), (1.13) in accordance with Eq. (1.14) is applicable, in particular, when $\overline{H} > H_0 = \phi_0/4\pi\lambda^2$, i.e., when the scale of the periodic nonuniformity of the traveling chain of Abrikosov–Josephson vortices turns out to be less than the London length. If the latter amounts to 10^{-5} to 10^{-4} cm, then $H_0 \sim 160-1.6$ 0e. In other words, for superconductors with sufficiently long London lengths manifestation of the relationships of nonlocal Josephson electrodynamics is possible at comparatively small magnetic fields, which, however, turn out to be so strong that they bar using the usual local electrodynamics of Josephson junctions. The corresponding magnetic fields can be even weaker in the case of films.

This work was performed with the support of the VTSP Scientific Council (Project ''AD'' No. 95008) and within the framework of Project No. 96-02-17303 of the Russian Fund for Fundamental Research.

*) E-mail: silin@sci.lpi.ac.ru

---

[1] A. A. Abrikosov, *Fundamentals of the Theory of Metals* (North-Holland Elsevier, Groningen, 1988; Nauka, Moscow, 1987).

[2] I. O. Kulik and I. K. Yanson, *Josephson Effect in Superconducting Tunnelling Structures* (Halsted, 1972; Nauka, Moscow, 1976).

[3] G. B. Whitham, *Linear and Nonlinear Waves* (Wiley Interscience, New York, 1974; Mir, Moscow, 1977).

[4] Yu. S. Kivshar and B. A. Malomed, Rev. Mod. Phys. **61**, 763 (1989).

[5] Yu. M. Aliev, G. L. Alfimov, K. N. Ovchinnikov, V. P. Silin, and S. A. Uryupin, J. Low Temp. Phys. **22**, 477 (1996).

[6] G. M. Lapir, K. K. Likarev, L. A. Maslova, and V. K. Semenov, Fiz. Nizk. Temp. **1**, 1235 (1975) [Sov. J. Low Temp. Phys. **1**, 590 (1975)].

[7] K. K. Likarev, Rev. Mod. Phys. **51**, 101 (1979).

[8] Yu. A. Aliev, V. P. Silin, and S. A. Uryupin, Sverkhprovodimost': Fiz., Khim., Tekh. **5**, 228 (1992).

[9] A. Gurevich, Phys. Rev. B **46**, 3187 (1992).

[10] M. Yu. Kupriyanov, K. K. Likarev, and A. K. Semenov, Fiz. Nizk. Temp. **2**, 706 (1976) [Sov. J. Low Temp. Phys. **2**, 346 (1976)].

[11] Yu. A. Aliev, V. P. Silin, and S. A. Uryupin, JETP Lett. **57**, 193 (1993).

[12] A. Gurevich, Phys. Rev. B **48**, 12857 (1993).

[13] A. Gurevich, Physica C **243**, 191 (1995).

[14] G. L. Alfimov and V. P. Silin, Zh. Eksp. Teor. Fiz. **106**, 671 (1994) [JETP **79**, 369 (1994)].

[15] V. P. Silin, JETP Lett. **60**, 462 (1994).

[16] B. A. Malomed and A. V. Ustinov, Phys. Rev. B **49**, 13024 (1994).

[17] V. V. Stepanov, *Course of Differential Equations* (GONTI, Moscow-Leningrad, 1939).

[18] N. N. Bautin and E. A. Leontovich, *Methods and Procedures of the Qualitative Investigation of Dynamic Systems on a Plane* (Nauka, Moscow, 1976).

[19] L. G. Aslamazov and A. I. Larkin, JETP Lett. **9**, 87 (1969).

[20] A. Erdelyi, ed., *Higher Transcendental Functions* (California Institute of Technology H. Bateman MS Project), Vol. 2 (McGraw-Hill, New York, 1953; Nauka, Moscow, 1963).

[21] Yu. M. Aliev and V. P. Silin, Zh. Éksp. Teor. Fiz. **104**, 2526 (1993) [JETP **77**, 142 (1993).

[22] A. Seeger, in *Handbuch der Physik*, Vol. 7, Pt. 1, Kristallphysik (Springer-Verlag, Berlin-Göttingen-Heidelberg, 1955), p. 383.

Translated by Eugene R. Heath

Physical manifestations of valence-bond structures in correlated systems

A. A. Ovchinnikov and M. Ya. Ovchinnikova

N. N. Semenov Institute of Chemical Physics, Russian Academy of Sciences, 117977 Moscow, Russia

(Submitted 17 January 1997)
Zh. Éksp. Teor. Fiz. **112**, 1409–1429 (October 1997)

Using the Hubbard model and the corresponding $t-J$ model, we study the properties of correlated states with valence-bond structures. Mixed states with such structures and with antiferromagnetic spin ordering can be constructed by means of local unitary transformations of uncorrelated states. The latter turn out to have lower energy than mean-field antiferromagnetic solutions. Spin correlations for various degrees of doping $\delta=n-1$ are in good agreement with the results of exact calculations for finite systems. In contrast to mean-field solutions, allowance for valence-bond correlations leads to a reasonable value of the critical δ, at which long-range antiferromagnetic order disappears. A calculation of the spectral functions that describe photoemission reveals typical behavior in two bands of effective hole (and electron) excitations, and energy transport in bands as the quasimomentum varies from (0,0) to (π,π), consistent with calculations in finite systems. We construct a homogeneous correlated state of fluctuating valence bonds (the band-model analog of states of fluctuating valence bonds), and demonstrate that its energy is lower than that of valence-bond alternant structures. *© 1997 American Institute of Physics.* [S1063-7761(97)01910-0]

1. INTRODUCTION

Two factors facilitate progress in studying the electronic properties of high-temperature superconductors (HTSC).[1] The first is ongoing improvement in experimental methods, e.g., in angle-resolved photoemission spectroscopy (ARPES)[2–8] and spin correlations via neutron scattering,[9] and so forth. The second is the development of numerical methods for calculating the energy and dynamical properties of strongly correlated finite systems within the scope of the basic Hubbard model or the equivalent $t-J$ model (exact diagonalization methods or the quantum Monte Carlo method).[1,10,11] The results of these and other studies need to be interpreted in the language of simpler solutions for these same models, employing one-particle representations in treating a correlated state.

Solutions for states with a variety of electronic structure—antiferromagnetic, spiral,[12,13] and others—have been constructed in studies of the Hubbard model via the mean-field method. Analogs of these states or more complicated spin structures—chiral states,[11,14] for example—have been proposed and studied with the $t-J$ model. A number of authors[1,10,15–20] have compared experimentally derived characteristics of photoemission spectra, Fermi surfaces, spin susceptibility, and so on, while numerically modeling finite systems and antiferromagnetic mean-field solutions with the generalized Hartree–Fock method.

In the present paper, we wish to continue such a comparative investigation, extending it to other kinds of correlated states—specifically, to dimer structures of valence bonds and four-site valence bonds, which we call $Q4$ structures. Valence-bond states were originally introduced by Anderson.[21,22] Studies of alternant valence-bond structures can be found in Refs. 23 and 24.

To construct correlated states with valence-bond structure, we make use of the variational method of local unitary transformations,[24] which enables one to determine the order parameter responsible for valence-bond dimer structure, and to work at an arbitrary doping level. To assess the contribution of some structure to an exact ground state, it is important to know not just the energy of the variational correlated state, but its physical characteristics, such as the shape of the Fermi surface, its photoemission spectrum, and spin correlations. Our aim here is to study the physical manifestations of valence-bond structures, and to discuss in a qualitative way how they relate to the results of exact calculations of finite clusters.

We deal here with the Hubbard model and the corresponding $t-J$ model. The Hamiltonian of the former is

$$H=-t\sum_{\langle nm\rangle,\sigma}(c^{\dagger}_{n\sigma}c_{m\sigma}+\text{h.c.})+\sum_{n}Un_{n\uparrow}n_{n\downarrow}-\mu\sum_{n}n_{n}. \tag{1}$$

Modeling of a CuO$_2$ plane[25,26] in the electron representation corresponds to $t>0$ and $n=n_e<1$ for a hole-doped plane.

In the subspace of functions with only singly-occupied and vacant sites, the model (1) corresponds to the $t-J$ model, which is described by the Hamiltonian

$$H_{t-J}=H^S+T,\quad H^S=J\sum_{\langle nm\rangle}\left(\mathbf{S}_n\cdot\mathbf{S}_m-\frac{1}{4}n_nn_m\right), \tag{2}$$

$$T=-t\sum_{\langle nm\rangle,\sigma}(c^{\dagger}_{1n\sigma}c_{1m\sigma}+\text{h.c.})+O\left(\frac{t^2}{U}\right), \tag{3}$$

$$J=4t^2/U,\quad c_{1n\sigma}=c_{n\sigma}(1-n_{n-\sigma}).$$

We restrict our attention to leading terms in the operator T, discarding three-site hopping terms[11,27] $\sim t^2/U$ and a constant that depends on the total number of particles.

1063-7761/97/100767-11$10.00

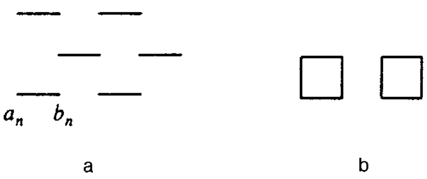

FIG. 1. a) Dimer structure in the x orientation of valence bonds DM_x, and b) the structure of four-site clusters $Q4$.

2. SPIN STATES OF THE UNDOPED $t-J$ MODEL

We begin by elucidating the symmetry of correlated states (which constitute the object of study here) using the undoped $t-J$ model.

For the $n=1$ state, the $t-J$ model with one electron or hole localized per site is specified by the spin Hamiltonian (2). We wish to compare the energies $\langle H_{t-J}\rangle$ of possible localized states Ψ_{loc}, which display a variety of spin structure. These correspond to various solutions (see Sec. 3) of Hubbard-model correlated states with finite U/t and $n\neq 1$. Among these states are the following four.

1) The Néel state of a perfect antiferromagnetic

$$\Psi_{\mathrm{loc}}^{AF}=\prod_n |c_{n\sigma_n}^\dagger\rangle,\quad \sigma_n=\frac{1}{2}(-1)^{n_x+n_y}\tag{4}$$

is the limit as $U\to\infty$, $n\to 1$ of the antiferromagnetic mean-field solution.

2) The dimer spin state is a product of the singlet components of the complexes $\{a_n,b_n\}$,

$$\Psi_{\mathrm{loc}}^{DM}=\prod_n \Psi_n^S(ab),$$

$$\Psi_n^S(ab)=\frac{1}{\sqrt{2}}\sum_\sigma \frac{\sigma}{|\sigma|}|a_{n\sigma}^\dagger b_{n,-\sigma}^\dagger\rangle,\tag{5}$$

and corresponds to a paramagnetic correlated state of the Hubbard model with valence-bond dimer structure. Here n labels the dimers.

3) A mixed state with dimer and antiferromagnetic spin structure is obtained by mixing triplet components into the spin functions of the complexes,

$$\Psi_{\mathrm{loc}}^{DM+AF}=\prod_n \{\cos\gamma\,\Psi_n^S(ab)+\sin\gamma\,\Psi_n^T(ab)\},\tag{6}$$

and has energy

$$N^{-1}\langle H_{t-J}\rangle^{DM+AF}=J\left\{-\frac{5}{8}-\frac{1}{2}\cos 2\gamma-\frac{3}{8}\sin^2 2\gamma\right\}.\tag{7}$$

4) Finally, valence-bond structures of four-site complexes $\{a_n,b_n,c_n,d_n\}$, which are depicted in Fig. 1b ($Q4$ structures), correspond to the paramagnetic or mixed state

$$\Psi_{\mathrm{loc}}^{Q4}=\prod_n \phi_n^{L=0},\tag{8}$$

$$\Psi_{\mathrm{loc}}^{Q4+AF}=\prod_n\left(\sum_{L=0}^{2} c_L\phi_n^L\right).\tag{9}$$

Here n labels the four-site clusters, while the cluster components

$$\phi_n^L=\sum_m C_{1m;1,-m}^{L,0}T_m(ac)T_{-m}(bd),\tag{10}$$

consisting of triplets $T(ac)$ and $T(bd)$ of the diagonally opposite cluster sites, are eigenfunctions of S_{ac}^2, S_{bd}^2, and the total spin $\mathbf{L}^2=(S_a+S_b+S_c+S_d)^2$ with vanishing projection in the z direction, $L_z=0$.

It is not difficult to show that the spin state (8) is the lowest-lying eigenstate of the intracluster part of the Hamiltonian in the $t-J$ model,

$$h_{Q4}=J(\mathbf{S}_a\cdot\mathbf{S}_b+\mathbf{S}_b\cdot\mathbf{S}_c+\mathbf{S}_c\cdot\mathbf{S}_d+\mathbf{S}_d\cdot\mathbf{S}_a)$$
$$=\frac{1}{2}J(\mathbf{L}^2-\mathbf{S}_{ac}^2-\mathbf{S}_{bd}^2).$$

If we assume additional antiferromagnetic order, the mean field of the spins of neighboring clusters results in mixing of the states of clusters with $\mathbf{L}\neq 0$, i.e., to the mixed state (9) with energy

$$N^{-1}\langle H_{t-J}\rangle^{Q4+AF}=\frac{1}{4}J\{-3-c_0^2+2c_2^2(\sqrt{2}c_0+c_1)^2\},\tag{11}$$

$$c_0^2+c_1^2+c_2^2=1.$$

The mean calculated energy per lattice site for states (4–6, 8, and 9) is shown in Table I. For the $DM+AF$ and $Q4+AF$ mixed states, the minina occur at $\gamma\neq 0$ and $c_1,c_2\neq 0$, which confirms that valence-bond and antiferromagnetic order structures can coexist. Table I also lists the mean spin $S_0=(-1)^{n_x+n_y}\langle S_z\rangle$ per site for states with antiferromagnetic order.

The data in Table I imply a number of properties:

1) stabilization of the state as a result of the organization of the valence-bond structure, which is consistent with the phase diagram calculated[24] for the paramagnetic DM state and the Hubbard-model ($DM+AF$) state. The energy is re-

TABLE I. Calculated mean energy per lattice site for states (4–9) of the undoped $t-J$ model with various valence-bond structures or antiferromagnetic order. Spin S_0 shown for the latter. Last column pertains to the ferromagnetic state.

	AF	DM	$DM+AF$	$Q4$	$Q4+AF$	F
$\langle H\rangle/J$	-1	-0.875	-1.042	-1	-1.084	0
S_0	0.5		0.475		0.434	

duced in a number of states (4–6, 8, and 9), and all of them have lower energy than the chiral phase for this model, where it is $-0.92t$ (Refs. 11 and 14);

2) a high degree of degeneracy: the mean energies of all dimer structures that completely cover a two-dimensional lattice are identical. This same degeneracy is also found in the $Q4+AF$ structures if the sign of the spin is transferred to the parity of the site in the original lattice;

3) a small energy difference between the paramagnetic and antiferromagnetic solutions when valence-bond structure is present. This presumes a lower Néel temperature than the elevated values given by mean-field calculations;

4) a reduced value of the local spin S_0 at magnetic sublattice sites in mixed antiferromagnetic states with valence-bond structures, as compared with the state of a rigid antiferromagnetic.

3. PROPERTIES OF BAND SOLUTIONS IN THE HUBBARD MODEL WITH ALTERNANT DIMER VALENCE-BOND STRUCTURE

In Ref. 24, we proposed a variational method for constructing states with alternant valence-bond structure at finite U/t for arbitrary doping. The Hubbard-model correlated state wave function is represented as resulting from local unitary transformations of an uncorrelated state Φ given by the Hartree–Fock approximation, or by the generalized Hartree–Fock method:

$$\Psi = W\Phi, \quad W=\prod_n W_n . \tag{12}$$

The local unitary operators W_n refer to disjoint elements of the structure in question (i.e., clusters). Then the original problem (1) with Hamiltonian H in the basis of correlated states Ψ is rigorously equivalent to the problem with transformed Hamiltonian

$$\widetilde{H} = W^{\dagger} H W \tag{13}$$

in the basis of functions Φ. Specifying the local operator $W_n(\alpha)$ with variational parameters α (essentially the order parameters of the new structures) and examining the new problem by means of the mean-field method makes it possible to study new types of correlated states at arbitrary doping levels. In contrast to the nonunitary local transformations of Gutzwiller[28], one can find for (13) an explicit (albeit cumbersome) expression for $\widetilde{H}$ in terms of the Fermi operators of the system, and carry out an exact energy calculation for Hartree–Fock-type functions.

For the simplest valence-bond dimer structure (Fig 1a), a cluster consists of the two centers $\{a_n b_n\}=\{c_n c_{n+e_x}\}$ that form the bond, while the operator with a single variational parameter is

$$W_n(\alpha)=\exp(\alpha[g^{\dagger}_{n\sigma}g^{\dagger}_{n-\sigma}u_{n-\sigma}u_{n\sigma}-\text{h.c.}])=\mathbf{I}$$

$$+W_{ij}X^{n}_{2i,2j}; \quad W_{ij}=\begin{pmatrix} C_{\alpha} & s_{\alpha} \\ -s_{\alpha} & C_{\alpha}\end{pmatrix}_{ij}, \tag{14}$$

$$g_{n\sigma}=\frac{1}{\sqrt{2}}(a_{n\sigma}+b_{n\sigma}), \quad u_{n\sigma}=\frac{1}{\sqrt{2}}(a_{n\sigma}-b_{n\sigma}),$$

$$C_{\alpha}=\cos\alpha-1, \quad s_{\alpha}=\sin\alpha. \tag{15}$$

Here the $X^{n}_{2i,2j}$ are Hubbard operators for states of the dimer $\{a_n b_n\}$ with a total of two electrons $(n_a+n_b=2)$:

$$|2i\rangle=\{|g^{\dagger}_{n\sigma}g^{\dagger}_{n-\sigma}\rangle,|u^{\dagger}_{n\sigma}u^{\dagger}_{n-\sigma}\rangle\}_i, \quad i=1,2. \tag{16}$$

Thus, W_n induces a rotation of the singlet two-hole components (16) of the nth cluster.

As a result, the effective Hamiltonian (13) takes the form

$$\widetilde{H}=\sum_n \widetilde{h}_n - t\sum_{n,\sigma}\sum_{l}[\widetilde{b}^{\dagger}_{n\sigma}\widetilde{a}_{n+l,\sigma}+\text{h.c.}]. \tag{17}$$

Here n labels the dimers (even sites of the original lattice), l runs through values $l=e_x\pm e_y$, $2e_x$, and the transformed operators

$$\widetilde{Q}_n=\{\widetilde{h}_n,\widetilde{b}_{n\sigma},\widetilde{a}_{n\sigma}\}=W^{\dagger}_n Q_n W_n, \tag{18}$$

like the original operators $Q_n=\{h_n, b_{n\sigma}, a_{n\sigma}\}$, affect the nth cluster.

As a result of the transformation, the intracluster part of the Hamiltonian (1)

$$h_n=U[n^{a}_{n\uparrow}n^{a}_{n\downarrow}+n^{b}_{n\uparrow}n^{b}_{n\downarrow}]-t(a^{\dagger}_{n\sigma}b_{n\sigma}+\text{h.c.})$$

takes the form

$$\widetilde{h}_n=h_n+h_{2i',2j'}[W_{i'i}W_{j'j}-\delta_{i'i}\delta_{j'j}]X_{2i,2j}. \tag{19}$$

Here $i',i,j',j=1,2$, $h_{2i',2j'}=\langle 2i'|h_n|2j'\rangle$ are matrix elements of h_n in the basis of cluster states (16), and the $W_{i,j}$ are defined in (14). The Hubbard operators $X_{2i,2j}$ can easily be expressed in terms of the original Fermi operators.[24] Similarly, $\widetilde{a}_{n\sigma}$ and $\widetilde{b}_{n\sigma}$ can be expressed in terms of

$$\widetilde{g}_{\sigma}=g_{\sigma}+C_{\alpha}g_{\sigma}\{n^{g}_{-\sigma}(1-n^{u})+n^{u}_{\sigma}n^{u}_{-\sigma}\}+s_{\alpha}g^{\dagger}_{\sigma}u_{-\sigma}u_{\sigma},$$

$$\widetilde{u}_{\sigma}=u_{\sigma}+C_{\alpha}u_{\sigma}\{n^{u}_{-\sigma}(1-n^{g})+n^{g}_{\sigma}n^{g}_{-\sigma}\}-s_{\alpha}u^{\dagger}_{\sigma}g_{-\sigma}g_{\sigma}. \tag{20}$$

We have omitted the cluster label here; C_{α} and s_{α} are defined in (15).

We thus have an explicit expression for $\widetilde{H}$ in terms of Fermi operators. We previously derived a self-consistent approximate solution[24] for the transformed Hamiltonian $\widetilde{H}(\alpha)$ at fixed α via the generalized Hartree–Fock method, with subsequent minimization over α. Note that in the basis of single-determinant functions $\Phi=|\Pi_k\chi^{+}_{k\sigma}\rangle$ of the generalized Hartree–Fock approximation, the mean energy can be expressed exactly[24] in terms of a set of one-electron means in the basis $\{\Phi\}$ defined in Appendix 1. The correlated state can then be represented in the form

$$\Psi=W\Phi=\left|\prod_{k,\sigma}\widetilde{\chi}^{\dagger}_{k\lambda\sigma}\right\rangle, \quad \widetilde{\chi}^{\dagger}_{k\lambda\sigma}=W^{\dagger}\chi^{\dagger}_{k\lambda\sigma}W. \tag{21}$$

The product over k here applies to quasimomenta inside the Fermi surface, and the one-electron functions $\chi^{\dagger}_{k\lambda\sigma}$ are eigenstates of the linearized Hamiltonian $(\widetilde{H})_L=\Sigma_{k\sigma}h_{k\sigma}+\text{const}$, with eigenvalues $E_{k\lambda}$, $\lambda=1,2$. The expressions for $h_{k\sigma}$,

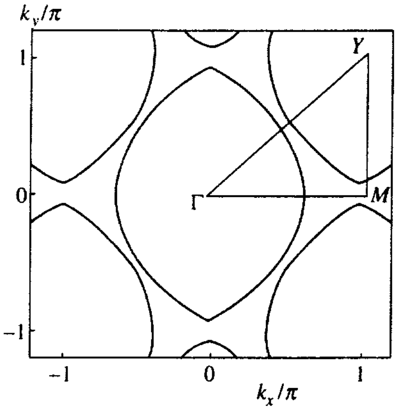

FIG. 2. Fermi boundary for states with dimer structure in the x-oriented valence bonds (DM_x) for the Hubbard model with $U/t=8$ and $n=0.8$.

$E_{k\lambda}$, and $\chi_{k\lambda\sigma}$, as well as the procedure for ensuring self-consistency developed in Ref. 24, are briefly detailed in Appendix 1.

To assess the role of valence-bond structures in actual correlated systems, we calculate a number of physical characteristics of these states, and broadly discuss their correspondence with exact numerical calculations for finite systems.

We begin with the shape of the Fermi surface and the density of states

$$n_k=\frac{1}{2}\sum_{\sigma}\langle c^{\dagger}_{k\sigma}c_{k\sigma}\rangle_{\Psi}=\frac{1}{2}\sum_{\sigma}\langle\Phi W^{+}c^{\dagger}_{k\sigma}c_{k\sigma}W\Phi\rangle. \quad (22)$$

Appendix 1 sets out the computational scheme (22) for a correlated state with wave functions given by (12) and (14). The energy eigenvalues of the linearized Hamiltonian $\tilde{H}_L$ form two bands $\lambda=1,2$, with a band gap $\sim U$ at $n=1$. For x-oriented dimer structure (denoted DM_x; Fig. 1a), the very flat minima of the upper band and maxima of the lower band correspond to coordinates $k_x=\pm(\pi-\delta_x)$, $k_y=0$ in the first Brillouin zone $|k_x\pm k_y|\leqslant\pi$. For an undoped system, the band occupation functions $f_{\lambda k}=\langle\chi^{\dagger}_{k\lambda\sigma}\chi_{k\lambda\sigma}\rangle_{\Phi}$ are 1 and 0, respectively. In accordance with the electron–hole symmetry of the Hubbard model, the Fermi boundaries, which are specified by $E_{k\lambda}=\mu$, are the same when the electron and hole doping are equal in absolute magnitude. Figure 2 shows the Fermi surface in the extended Brillouin zone for $n=0.8$, $U/t=8$. The height of the maximum is $E^{\max}_{1k}-\mu=0.092t$.

Inasmuch as real HTSC single crystals (apart from YBaCuO) are symmetric under the interchange $k_x\leftrightarrow k_y$, it makes sense to calculate a density of states n_k that is averaged over the two dimer structures of valence bonds with x and y orientation,

$$\overline{n}_k=\frac{1}{2}[n_k(DM_x)+n_k(DM_y)].$$

In Fig. 3, we plot profiles of the mean density of states $\overline{n_k}$ along the loop $\Gamma YM\Gamma$ for $n=1.2$, 1, and 0.8, as well as contour maps of $\overline{n_k}(k_x,k_y)$ for $DM+AF$ mixed states. For $n=0.8$, we also plot the density n_k for an antiferromagnetic state of the generalized Hartree–Fock method with no valence-bond structure, but with a gap along the boundary $|k_x\pm k_y|=\pi$ of the first Brillouin zone. The dotted and dashed curves correspond to structures with x and y orientation.

Discontinuities or regions of rapid variation of n_k (black areas in Fig. 3b) are associated, albeit not so directly, with the observable signal limits in the photoemission spectrum. In particular, n_k is the integral over ω of the spectral density measured in forward and backward photoemission (ARPES, IPES). Note that the boundary of rapid variation in n_k exterior to the point Γ (0,0) can, to a gross approximation, be associated with the Fermi "boundary" established via ARPES measurements. Likewise, the interior boundary can be associated with the Fermi "boundary" that might be observed via IPES, if it were possible to do so with angular resolution. Hence, the Fermi "boundary" that would be observed for a state with the structure in question forms an arc about the point Y (π,π) for both electron and hole doping.

With such a definition of the Fermi "boundary," Lattinger's theorem would not hold, in direct conflict with experimental observations. For hole doping ($n_h>1$), the region of phase space outside the "boundary" constitutes less than half the entire space, while according to the theorem, it should comprise $n_h/2>0.5$. This inconsistency was also noted in calculations of the Fermi surface based on projection operators.[29]

With regard to the arc-shaped Fermi surface about the point Y (π,π), the very same shape can be observed for most HTSC single crystals—both electron-doped[3,4] and hole-doped.[5–8] This shape is usually explained by empirical models that incorporate next-nearest-neighbor (diagonal) interactions t' between centers.[30–33] A direct derivation[25,26] of the generalized Hubbard-model parameters from the three-band Emery model yields lower values of t' than the empirical values obtained from ARPES data. Our own original model (1) contains no such interactions. Nevertheless, dimer structures in the correlated state under investigation will give rise to non-nearest-neighbor interactions. Interaction terms between sites n and m with $|n-m|=\sqrt{2},2$ appear in the explicit expression for $\tilde{H}(\alpha)$.

We now discuss spin correlations in mixed $DM+AF$ states with dimer structure. In Fig. 4, we plot the calculated static correlation function

$$\chi(l)=\frac{1}{3}N^{-1}\sum_{m}(-1)^{l_x+l_y}\langle 4\mathbf{S}_m\mathbf{S}_{m+l}\rangle_{\Psi}$$

for $U/t=8$ and various doping levels. Here m and l label lattice sites, rather than dimers. It is this quantity, rather than $\chi_z\sim(-1)^{l_x+l_y}\langle 4(S_m)_z(S_{m+l})_z\rangle_{\Psi}$, that must be compared with $\chi(l)$, which was calculated in Refs. 17 and 19 for finite clusters. The ground state of finite systems with equivalent axes x,y,z must be put into correspondence with the set of $DM+AF$ states in the two-dimensional plane with various polarizations of local spins. This is equivalent to calculating $\chi(l)$ instead of $\chi_z(l)$ for a $DM+AF$ state with one (z) polarization. The existence of a plateau in the curves of $\chi(l)$, which disappears as the doping level rises, corresponds to

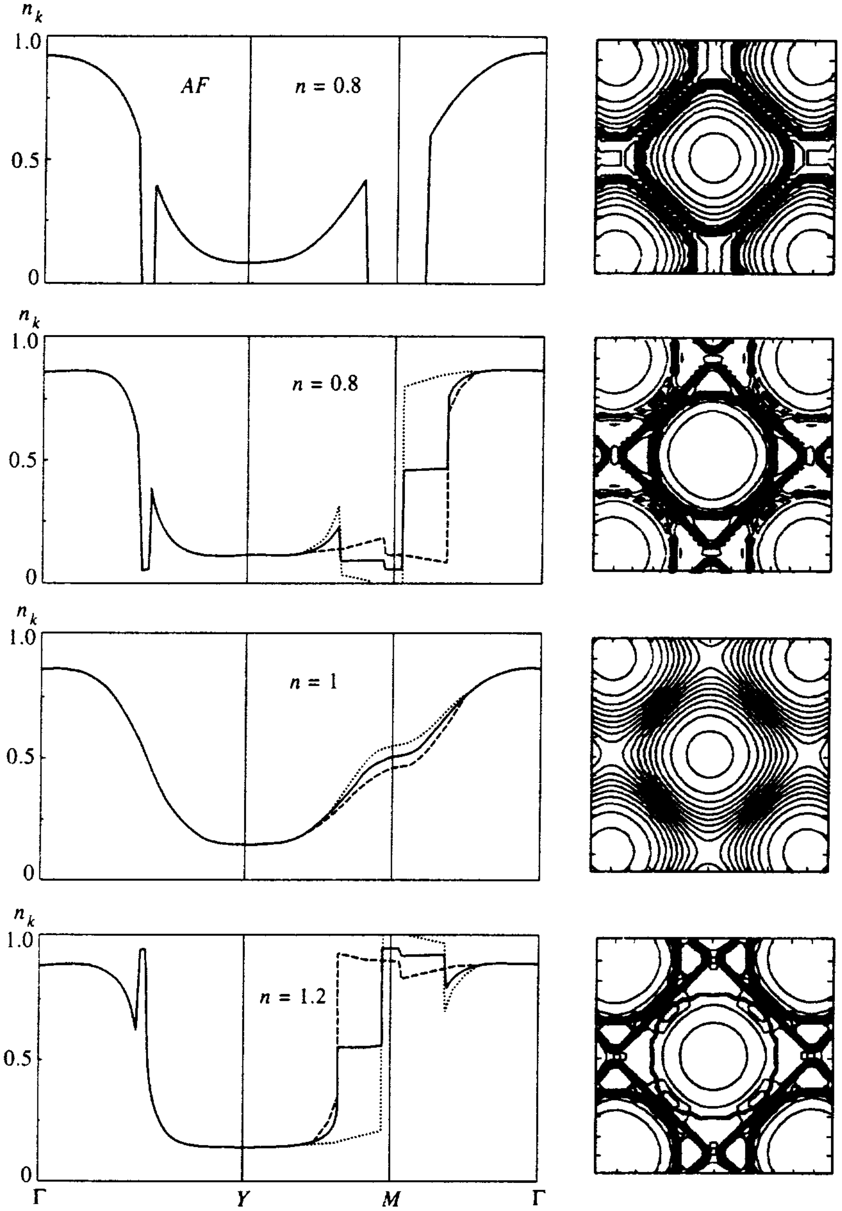

FIG. 3. Profiles of the density of states $\overline{n_k}$ averaged over two structures along the loop $\Gamma YM\Gamma$ (solid curves), and contour maps of $\overline{n_k}(k_x,k_y)$ for mixed $DM+AF$ states in the region $|k_{x(y)}|\leqslant 1.2\pi$. The uppermost plots pertain to an antiferromagnetic state of the mean field with no valence-bond structure. Dotted and dashed curves correspond to x and y orientations.

long-range antiferromagnetic order with $\chi_z(l)\sim 3\chi(l)$, and is consistent with the previously found phase diagram.

The derived behavior of $\chi(l)$ closely resembles results for $\chi(l)$ obtained by Haas *et al.*[17] for the $t-J$ model (8×8 cluster, $J=4t^2/U=0.4$) via exact diagonalization. The similar results in Ref. 17 and Fig. 4 signify that the plateau in χ_z at $l\geqslant\sqrt{2}$ in Ref. 17 is indicative of long-range antiferromagnetic order in finite systems. Quantum Monte Carlo calculations[19] of $\chi(l)$ for finite systems in the Hubbard model were carried out at high temperatures ($kT=t/4$), so that long-range antiferromagnetic order disappears at lower doping levels than in our calculations and those of Haas *et al.*[17]

Note that similar curves of $\chi(l)$ with a plateau at $l\geqslant\sqrt{2}$ are typical of states (6) and (9) of the $t-J$ model with a variety of valence-bond structures, such as dimer or $Q4$ symmetry. Figure 4 shows the result for the latter at $n=1$. This means that the behavior of $\chi(l)$ is not particularly sensitive to the structure of the valence bonds. In the mean-field method with no valence-bond structure, however, i.e., for $\alpha=0$ in $W(\alpha)$, both antiferromagnetism and the plateau in $\chi(l)$ disappear at very high doping levels ($\delta\sim 0.45$, as opposed to $\delta\sim 0.25$ for $AF+DM$ states with $U/t=8$).

4. EXCITATION BANDS OBSERVED IN PHOTOEMISSION SPECTRA

We now attempt to calculate the photoemission spectra responsible for the correlated ground state with valence-bond dimer structure; for simplicity, we neglect antiferromagnetic order. In angle-resolved forward and backward photoemission (ARPES, IPES), the measured signal is proportional to the spectral functions, which in turn specify the quasimomentum k and energy ω (in the electron representation):[1]

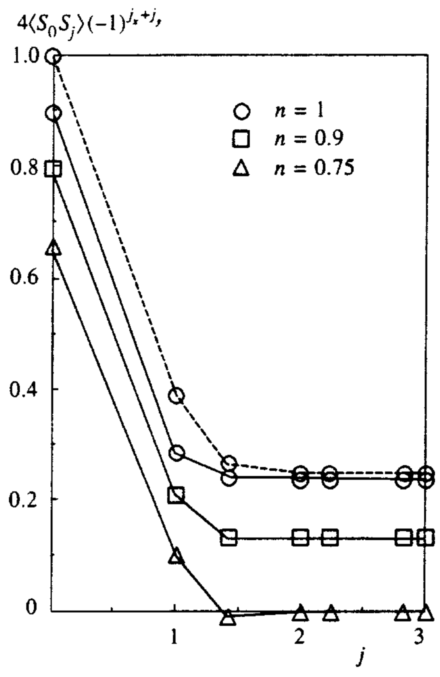

FIG. 4. Spin correlation function as a function of distance between spins for $DM+AF$ states of the Hubbard model with $U/t=8$, for various levels of doping. The dashed line represents the spin correlation for the localized state (9) of a $Q4+AF$ structure in the undoped $t-J$ model.

$$A^-(k,\omega)=\sum_\nu |\langle\Psi_\nu^{N_e-1}|c_{k,\sigma}|\Psi_0^{N_e}\rangle|^2\delta \times(\omega-\mu+E_\nu^{N_e-1}-E_0^{N_e}),\quad \omega\geqslant 0, \tag{23}$$

$$A^+(k,\omega)=\sum_\nu |\langle\Psi_\nu^{N_e+1}|c^+_{k,\sigma}|\Psi_0^{N_e}\rangle|^2\delta \times(\omega-\mu-E_\nu^{N_e+1}+E_0^{N_e}),\quad \omega\leqslant 0. \tag{24}$$

In these equations, the $\Psi_\nu^{N_e\pm1}$ are excited states of the system with an extra electron or hole and quasimomentum k.

The projection operator method of Unger and Fulde[29] facilitates the self-consistent approximate calculation of the spectral functions. Ignorance of the ground state in Ref. 28 is made up for by the use of sum rules and splitting of complicated operators. Here, in constrast to Unger and Fulde,[29] we start with the variational ground state of a certain structure—specifically, a valence-bond dimer structure. Our goal is to identify manifestations of this structure in photoemission spectra. Its relationship to the projection operator methods of Ref. 29 consists in the use of a finite basis set of operators $\{Q_{\lambda k\sigma}\}$ to construct the excited states $\Psi_\nu^{N_e-1}$ with quasimomentum k, spin projection σ, and one extra hole above the given ground state $\Psi_0=W\Phi_0$.

We seek excited states with quasimomentum k in the form

$$\Psi_{k\sigma}^\nu=W\sum_\lambda D_\lambda^\nu\mathbf{Q}_{\lambda k\sigma}\Phi_0,\quad \lambda=1,\ldots,L_0. \tag{25}$$

For given k and an arbitrary set of odd operators $\{q_{ik\sigma}\}$ that reduce the number of particles by 1, we can selectively construct a smaller set of "orthogonal" operators $\{Q_{\lambda k\sigma}\}$ with nonvanishing norm (see Appendix 2), such that

$$S_{\lambda\lambda'}=\langle\Phi_0Q^+_{k\lambda\sigma}Q_{k\lambda'\sigma}\Phi_0\rangle=S_\lambda\delta_{\lambda\lambda'},\quad S_\lambda\neq0,$$
$$\lambda,\lambda'=1,\ldots,L_0(k). \tag{26}$$

This is precisely the set used in (25).

A discrete set of effective excited states that mimics the multiparticle hole excitation continuum can then be derived from the equation

$$[S_\lambda^{-1/2}H_{\lambda\lambda'}S_{\lambda'}^{-1/2}]D_{\lambda'}^\nu=E_k^\nu D_\lambda^\nu, \tag{27}$$

where

$$H_{\lambda\lambda'}=\langle\Phi_0Q^+_{k\lambda\sigma}\widetilde{H}Q_{k\lambda'\sigma}\Phi_0\rangle. \tag{28}$$

The photoemission signal amplitude $A(k,\omega)$, which is responsible for exciting effective hole states E_k^ν from a correlated state with x-oriented valence-bond dimer structure, can be written in the form

$$A(k,\omega)=\sum_\nu |M^\nu(k_e)|^2\frac{\Gamma}{(\omega-E_k^\nu)^2+\Gamma^2}, \tag{29}$$

$$M^\nu(k_e)=\sum_\lambda (D_\lambda^\nu)^*\langle\Phi_0Q^+_{k\lambda\sigma}|\widetilde{c}_{k\sigma}\Phi_0\rangle. \tag{30}$$

Here k_e is the momentum of the ejected electron, and $\widetilde{c}_{k\sigma}=W^+c_{k\sigma}W=\widetilde{a}_{k\sigma}+e^{ik_x}\widetilde{b}_{k\sigma}$ can be obtained by applying $c_{k\sigma}$ to $\Psi_0=W\Phi_0$ and invoking the definition (25); $\delta(\omega-E_k^\nu-E_0)$ in (23) has been replaced by a Lorentzian with width Γ, and the D_λ^ν are eigenvectors of the matrix equation (27). In subsequent calculations, we treat Γ as an adjustable parameter equal to $0.1t$. In principle, it should be possible to evaluate the widths Γ^ν for any effective level $E^\nu(k)$ in terms of the commutators $[[Q_{\lambda k\sigma},H],H]$, but these are too tedious to calculate. Similar expressions can be obtained for $A^+(k,\omega)$ $(\omega>0)$ by extracting the orthogonal set $\{Q^+_{\lambda k\sigma}\}$ from the operator basis $\{q^+_{ik\sigma}\}$. It can be shown that for undoped systems, $A^+(k,\omega)|_{n=1}=A^-(\widetilde{k},-\omega)$, where $\widetilde{k}=(\pi\pi)-k$.

For an alternant DM_x+AF structure with two sites per elementary cell, the minimal reference basis for the operators $\{q_{ik\sigma}\}$ consists of the two operators $\{a_{k\sigma},b_{k\sigma}\}_i$. For $n=1$, only one of the combinations $\chi_{1k\sigma}=c_aa_{k\sigma}+c_bb_{k\sigma}$—the solution to the linearized problem $(\widetilde{H})_L$, which is responsible for the lower Hubbard band $\lambda=1$—produces hole excitations with nonvanishing norm $(\chi_{2k\sigma}\Phi=0)$. In this sort of one-particle approximation, the photoemission signal would correspond to an isolated band $E_{\lambda k\sigma}$—the lower $(\lambda=1)$ or upper $(\lambda=2)$, respectively, for forward or backward photoemission. One encounters such behavior in the antiferromagnetic ground state of the generalized Hartree–Fock method with structureless valence bonds: all strong (multiparticle) excited states are inaccessible from the single-determinant ground state, since they cannot be produced by the one-particle operator $c_{k\sigma}$, which figures in the interactions between light and electrons.

Meanwhile, in calculations of finite clusters via exact diagonalization in the $t-J$ model[17] or the quantum Monte

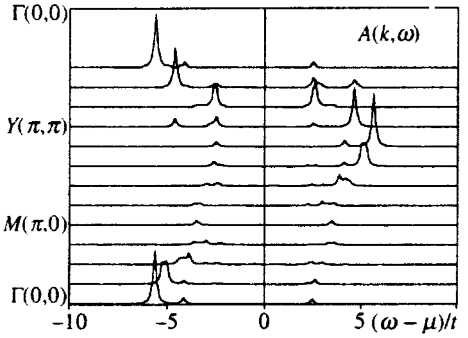

FIG. 5. Spectral functions calculated with Eqs. (30) and (31), $\Gamma=0.1t$, for a state with valence-bond dimer structure (no antiferromagnetism) in the Hubbard model with $U/t=8$, $n=1$, averaging over the two structures with x and y orientation. The curves (top to bottom) correspond to $k=(0,0),(\pi/4,\pi/4),\ldots,(0,0)$ along the loop $\Gamma YM\Gamma$ of Fig. 2.

Carlo method for the Hubbard model,[20] one can readily trace the two-band structure of each of the spectral functions (24) and (25) responsible for forward and backward photoemission. Using the expanded basis of operators $q_{ik\sigma}$, our calculation with Eqs. (29) and (30) shows that allowance for correlations like those that produce valence-bonds leads to splitting of the forward and backward photoemission bands into two subbands. The basis operators $q_{ik\sigma}$ and the computational scheme and approximations are detailed in Appendix 2.

In Fig. 5, we plot the spectral functions $A(k,\omega)$ for the ground state with valence-bond dimer structure with no antiferromagnetic order, calculated with Eqs. (29) and (30), $\Gamma=0.1t$. Two bands are easily seen, with both $\omega<0$ and $\omega>0$, displaying typical interband energy transport as the quasimomentum k varies from the point Γ (0,0) to M (π,0) and on to Y (π,π). This is the very sort of energy transport shown by exact calculations of finite systems.[20]

The emergence of two hole (or electron) excitation bands for states with valence-bond dimer structure becomes understandable in light of their approximate description in the undoped $t-J$ model. The state (5), with localized spins and dimer spin structure in the $\{a_n b_n\}$ bonds, can be identically rewritten in the form

$$\Psi_{DM}^{t-J}=W\left(\alpha=\frac{\pi}{4}\right)\Phi_0^{t-J},\quad \Phi_0^{t-J}=\left|\prod_{k,\sigma} g_{k\sigma}^{\dagger}\right\rangle$$

$$=\left|\prod_{n,\sigma} g_{n\sigma}^{\dagger}\right\rangle,\quad W=\prod_n W_n\left(\alpha=\frac{\pi}{4}\right).$$

Here k occupies the unit cell F_0 of the Brillouin zone, n labels the dimers, the operator W_n transforms the corresponding component of the function Φ to the singlet dimer state with spin localization:

$$W_n\left(\frac{\pi}{4}\right)|g^+_{n\sigma}g^+_{n,-\sigma}\rangle=\frac{1}{\sqrt{2}}\,|a^+_{n\uparrow}b^+_{n\downarrow}+b^+_{n\uparrow}a^+_{n\downarrow}\rangle.$$

We can write the general state of the $t-J$ model, with quasimomentum k and one hole above the background of the state (5) with the structure of Fig. 1a, in the form

$$\Psi^h_{k\sigma}=W\left(\frac{\pi}{4}\right)\sqrt{\frac{2}{N}}\sum_n e^{ikx}(c_g g_{n\sigma}$$

$$+c_u u^+_{n,-\sigma}g_{n\sigma}g_{n,-\sigma})\Phi_0^{t-J}.$$

The energy of the two hole-excitation bands, up to terms $\sim t/U$, is then

$$E_k^{1(2)}=\mp t\{5/4-\cos k_x\cos k_y-\cos 2k_x+\cos^2 k_y\}$$

$$+O(t^2/U).$$

Thus, two hole (and analogous electron) excitation bands are associated with the two dimer states with $n_e=1$ (or $n_e=3$), which are mixed when $k\neq(0,0)$ and $k\neq(\pi,\pi)$. Note that photoemission of an electron with spin σ from the ground state (4) of a perfect antiferromagnetic can produce both sublattice sites with spin polarization σ and sublattice sites with the opposite polarization, but with the neighboring spin flipped in the process. Splitting and dispersion of the resulting hole excitations are of order $J=4t^2/U$, instead of being of order t like the splitting of excitations from the localized states (5) and (6) with valence-bond dimer structure.

The photoemission intensity at the corresponding frequencies $\omega=E_k^{1(2)}-\mu$ from any localized state of the $t-J$ model, however, in no way reproduces the behavior shown in Fig. 5 and analogous results computed for finite systems.[20] For the $t-J$ model with $n=1$, we have the symmetry

$$A(k,-\omega)=A(k,\omega),\quad A(\tilde{k},\omega)=A(k,\omega),\quad \tilde{k}=(\pi,\pi)-k. \tag{31}$$

In contrast to Eqs. (31) for the $t-J$ model, in Hubbard-model calculations for finite U there is an explicit asymmetry in $A(k\omega)$ for forward and reverse photoemission, which reflects the band nature of the correlated ground state. In particular, a property typical of noninteracting electrons is explicit: the intensity of forward ($\omega>0$) or reverse ($\omega<0$) photoemission is nonzero for quasimomenta both inside and outside the Fermi surface.

The explanation set forth here, however, does not prove that doubling of the photoemission bands is necessarily a result of the emergence of exclusively dimer structures in the valence bonds. The simplest calculation of hole excitations against the background of the localized ground state (8) with $Q4$ symmetry in the valence-bond structure for the $t-J$ model shows that this structure can also mimic band doubling. Out of nine bands correlated with the levels of an individual four-site complex with $n=3$, i.e., with one hole, and with total spin 1/2, only two pairs of close-lying bands turn out to be active in photoemission, which with allowance for broadening can mimic two excitation bands.

5. HOMOGENEOUS STATE OF FLUCTUATING VALENCE BONDS

Correlated states with alternant dimer structure of the valence bonds with a doubled unit cell were studied above and in Ref. 24. These were constructed with the aid of local operators that operate on disjoint clusters—the dimers $\{a_n,b_n\}$. The local operators chosen in this way commute with one another, and they are responsible for the inequivalence of the bonds within the dimers and the bonds with neighboring sites that belong to other dimers. The self-consistent Hartree–Fock solution for the transformed Hamiltonian $\widetilde{H}$ reinforces a ''structure''—a self-maintaining inequivalence of the bonds—that resembles spin-wave density waves.

One can, however, imagine a correlated state of fluctuating valence bonds, constructed from the uncorrelated state Φ via the unitary transformation

$$\Psi = W\Phi, \quad W = e^{\alpha Z}, \quad Z = \sum_{\langle nm \rangle} Z_{nm}. \tag{32}$$

Here the local antihermitian operator Z_{nm}, which operates on two nearest-neighbor sites, is a perfect analog of the generator in the transformation operator (14) within the dimer (in contrast to (14), n and m in (32) label the sites, rather than the dimers):

$$Z_{nm} = [g^{\dagger}_{nm,\sigma} g^{\dagger}_{nm,-\sigma} u_{nm,-\sigma} u_{nm,\sigma} - \text{h.c.}]$$

$$\equiv -\frac{1}{2} \sum_{\sigma} j_{nm\sigma} \Delta_{nm,-\sigma}, \tag{33}$$

$$j_{nm\sigma} = c^{\dagger}_{n\sigma} c_{m\sigma} - c^{\dagger}_{m\sigma} c_{n\sigma}, \quad \Delta_{nm,-\sigma} = n_{n-\sigma} - n_{m-\sigma},$$

where g and u denote even and odd combinations of orbits at neighboring centers n,m.

The summation in (32) takes place over all nearest-neighbor bonds. Z_{nm} operates on the singlet components of neighboring sites with two particles per bond in the total function Φ. Such components interact with one another probabilistically in Φ. One can therefore speak of random formation of valence bonds, i.e., the formation of singlets from localized spins at pairs of sites following charge fluctuations at those pairs in the single-determinant function Φ.

In view of the noncommutativity of the local operators, we cannot find (as we did previously) an effective Hamiltonian $\widetilde{H} = W^{\dagger} H W$ to all orders in the transformation parameter α. It is possible, however, to find $\widetilde{H}$ up to terms $\sim \alpha^2$,

$$\widetilde{H} \approx H + \alpha[H,Z] + \frac{\alpha^2}{2}[[H,Z],Z], \tag{34}$$

and to find a fully homogeneous, self-consistent Hartree–Fock solution for $\widetilde{H}$, neglecting terms $\sim \alpha^3$ and higher. The smallness of the optimal value of α, which minimizes the energy $\widetilde{H} = \langle \widetilde{H}(\alpha) \rangle_{\Phi}$, justifies the expansion in α and makes it possible to assess the accuracy of the approximation.

The mean energy for such a problem, in the single-determinant state Φ of the generalized Hartree–Fock method, depends on nine one-electron expectation values over Φ:

$$\langle \widetilde{H} \rangle_{\Phi} = \widetilde{H}(y_{\nu}), \quad y_{\nu} = \{\rho_l, \delta_l\}_{\nu},$$

$$\rho_l = \frac{1}{2} \sum_{\sigma} \langle c^{\dagger}_{n\sigma} c_{n+l,\sigma} \rangle, \quad \delta_l = \frac{1}{2} \sum_{\sigma} \frac{\sigma}{|\sigma|} \langle c^{\dagger}_{n\sigma} c_{n+l,\sigma} \rangle, \tag{35}$$

$$l = (0,0),(1,1),(0,2), \quad (0,1),(1,2),(0,3).$$

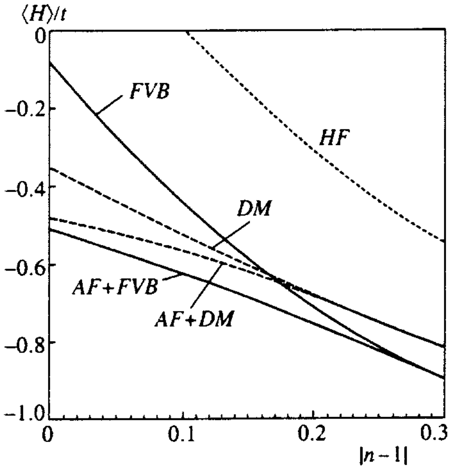

FIG. 6. Dependence of energy on the doping level for correlated homogeneous states with fluctuating valence bonds (solid curve) and states with alternant dimer structure (dashed curve), with antiferromagnetism either present or absent. The upper curve corresponds to the energy of the Hartree–Fock state. For an electron-doped system ($n > 1$), the vertical axis corresponds to $[\langle H \rangle - U(n-1)]/t$.

In these means, $\delta_l \equiv 0$ for odd l. The ground state of the Hamiltonian $\widetilde{H}$ in the Hartree–Fock approximation,

$$\Phi = \left| \prod_{k\sigma}^{F} \chi^{\dagger}_{\lambda k \sigma} \right\rangle,$$

is defined by the eigenfunctions $\chi_{\lambda k \sigma}$ and spectrum $E_{\lambda k}$ of the linearized Hamiltonian $\widetilde{H}_L = \Sigma_{\nu} \{\partial \widetilde{H} / \partial y_{\nu}\} (\mathbf{y}_{\nu} - y_{\nu}) + \widetilde{H}$. In the latter equation, $\mathbf{y}_{\nu}$ is the operator corresponding to the mean y_{ν}.

In Fig. 6, we plot the calculated mean energy per site as a function of doping level for a homogeneous state with fluctuating valence bonds and spins with antiferromagnetic order ($\delta_l \neq 0$), as well as the same quantities in the absence of antiferromagnetic order ($\delta_l = 0$), compared with the energies of states with alternant structure of the valence bonds for the Hubbard model with $U/t = 8$. Among antiferromagnetic solutions, states with a homogeneous distribution of valence bonds turn out to be lower in energy than states with alternant structure DM_x. This may well be related to the quadrupling of the number of bonds on which optimization of two-hole components takes place.

Figure 7 shows the optimal transformation parameter $\alpha(n, U/t)$ as a function of the doping $\delta = |n-1|$ at fixed $U/t = 8$, or as a function if U/t with $n = 1$, for a homogeneous $FVB + AF$ state of the valence bonds, compared with the corresponding parameter α for the alternant dimer structure $DM_x + AF$. For $U/t \leqslant 9$, α is indeed small for $FVB + AF$, $\alpha < 0.2$, which justifies the quadratic (in α) representation of the effective Hamiltonian (34). At $U/t \sim 12$, however, there is an abrupt increase in the optimal α for the FVB state, while the function $\widetilde{H}(\alpha)$ becomes quite flat near minimum. It is unclear whether this behavior signals a qualitative change in the ground state at $U/t \sim 12$; the issue cannot be resolved in the context of the expansion (34).

The preceding comments have to do with antiferromagnetic solutions. At low doping, paramagnetic FVB states are

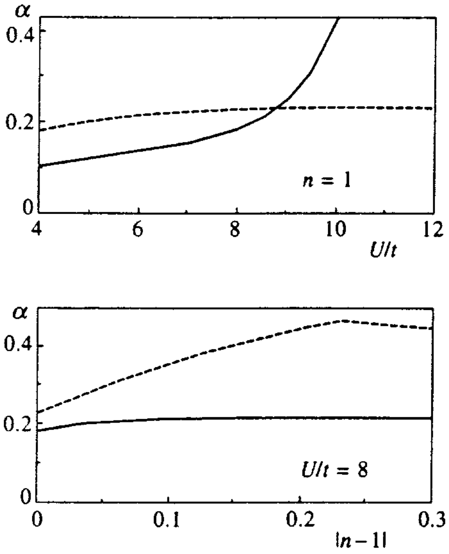

FIG. 7. Dependence of optimal transformation parameter $\alpha(n,U/t)$ on U/t at $n=1$ (upper panel), and on $\delta=|n-1|$ at $U/t=8$ (lower panel) for the $FVB+AF$ (solid curve) and $DM+AF$ states (dashed curve).

significantly higher in energy than states with valence-bond dimer structure, and $AF+FVB$ states with a doubled unit cell (see Fig. 6) as a result of the emergence of a gap in the spectrum of the latter. In accordance with the results obtained by Dagotto *et al.*[34] for finite systems, that gap gives rise to a jump in the chemical potential $\mu(n)$ at $n=1$. In Fig. 8 we plot $\mu(n)$ for the paramagnetic FVB state and the $AF+FVB$ state for the Hubbard model with $U/t=8$. The behavior of $\mu(n)$ for the former is not much different from $\mu(n)$ for the Hartree–Fock solution (dashed curve).

We now note the distinction at the Fermi surface between periodic structures (antiferromagnetic or valence-bond dimer structure) with two sites per unit cell and the homogeneous paramagnetic FVB state. In the former with $n\neq 1$, the Fermi surface encloses the lobe along the nesting line $|k_x\pm k_y|=\pi$ of the undoped system, so that the outer and inner parts of the Fermi surface enclose phase volumes $V_{1(2)}=4\pi^2(1\pm|n-1|)$. For the homogeneous FVB state in the paramagnetic case (i.e., with no magnetic doubling of the unit cell), only one of these boundaries remains, enclosing a phase volume $V=4\pi^2 n$ in accordance with Lattinger's theorem, as confirmed by ARPES experiments.[2–8] In contrast to the work reported in Ref. 2–8, the shape of the Fermi surface for the homogeneous state FVB can be described phenomenologically by a linear model with hopping interactions between sites (n,m) with odd $l=m-n$. In such a model, there are no next-nearest-neighbor interactions between sites with $l=m-n=(\pm 1,\pm 1)$.

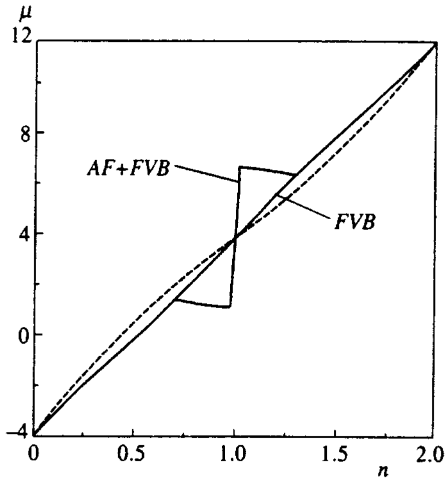

FIG. 8. Chemical potential $\mu(n)$ as a function of electron number density for the antiferromagnetic and paramagnetic solutions FVB. Solid and dashed curves have the same significance as in Fig. 7.

The homogeneous FVB state has a Fermi surface shaped much like that of the two-dimensional Hubbard model with no interaction. This surface does not display the typical experimentally observed shape of the Fermi surface (for NdNiCuO,[2–4] YBaCuO,[5] and BiBaCuO[6–8])—an arc about the point Y ($\pi\pi$), possibly for a variety of reasons. The correlations described by the transformations (32) and (33) may not exhaust the electron correlations; possibly, the original model ought to include nearest-neighbor hopping interactions, as in Refs. 31–33. Nevertheless, the state (32) is important, in that it yields a paradigm for the explicit construction of a correlated state with fluctuations.

6. CONCLUSIONS

In the $t-J$ and Hubbard models, the formation of dimer or more complicated valence-bond structures stabilizes both antiferromagnetic and paramagnetic states. In contrast to a macroscopically inhomogeneous state with phase differentiation, which emerges from a mean-field solution only in doped systems,[13] states with microscopically inhomogeneous valence-bond structures can coexist perfectly with antiferromagnetic order both in undoped systems and in doped systems, over a significant range of both $\delta=|n-1|$ and U/t. Local unitary transformations with a variational parameter provide a convenient means of constructing states with valence-bond structure.

The good agreement between the spin correlation $\chi(l)$ as a function of distance l between the spins and the doping level δ obtained for states with valence-bond dimer structure and in exact calculations of finite systems[1,15–20] makes it possible to interpret short-range correlations as a consequence of the formation of valence-bond structures, and long-range order as a consequence of antiferromagnetism, which vanishes at the critical doping level δ_c. When correlations like those that produce valence-bond structures are taken into consideration, the critical value $\delta_c\sim 0.25$ for the transition of the antiferromagnetic state into the paramagnetic state turns out to be much less than $\delta\sim 0.45$, the value predicted by mean-field theory (these numbers are obtained for $U/t=8$).

States with valence-bond structure exhibit a two-band pattern of effective (multiparticle) hole (and electron) excitations with dispersion and splitting $\sim t$. This is precisely the pattern obtained in exact calculations of finite clusters.[20] Simpler solutions for the $t-J$ model with valence-bond structure make it possible to interpret them as a consequence of two (even and odd) dimer states with $n=1$ or $n=3$ particles that mix values of k other than (0,0) and (π,π). In contrast to localized spin states, calculations for the bands of correlated states with valence-bond structure also reproduce the interband energy transport when the quasimomentum changes from (0,0) to (π,π), as observed in exact calculations of finite systems, and typical of uncorrelated states.

The high degeneracy of states with varied valence-bond structure (as observed in the simpler solutions of the $t-J$ model) make only one particular structure unlikely. This also confirms the homogeneous variational correlated state of the valence bonds that we constructed, whose fluctuations follow the charge fluctuations in the uncorrelated state. The energy of such a state with fluctuating valence bonds turns out to be lower than that of a state with alternant structure of the valence bonds. FVB states provide an example of the explicit representation of a correlated state exhibiting one of the various types of fluctuations, in terms of the formation of nearest-neighbor valence bonds.

The problem of the shape of the Fermi surface in HTSC is more complicated. It is unclear to what degree a generalization to other types of correlations—in particular, those involving non-nearest neighbors—or expansion of the Hubbard model itself, or a subtler treatment of the photoemission data, can eliminate the existing discrepancies in the shape of the Fermi surface.

This research was supported by Grant No. 015-943 from the International Science and Technology Center, and by the Russian Fund for Fundamental Research (Project 97-03-33727A). We also thank V. Ya. Krivnov for useful discussions of our results, and E. Plekhanov for assistance in checking the calculations.

APPENDIX 1

The mean energy $\overline{H(y_i)}=\langle\Psi H\Psi\rangle=\langle\Phi\widetilde{H}\Phi\rangle$ for states $\{a_n,b_n\}$ with valence-bond dimer structure (Fig. 1a) can be calculated exactly, and depends on the following one-electron means over the single-determinant state Φ:

$$\{y_i\}=\{\rho_0,\delta_0,\gamma_0,\rho_1,\delta_1,\gamma_1,\beta_1,\rho_2,\delta_2,\gamma_2,\beta_2\}_i\,, \tag{36}$$

$$\rho_\nu-\zeta_\sigma\delta_\nu=\langle a^\dagger_{n\sigma}a_{m\sigma}\rangle_\Phi=\langle b^\dagger_{n,-\sigma}b_{m,-\sigma}\rangle_\Phi\,,$$

$$\gamma_\nu+\beta_\nu=\langle a^\dagger_{n\sigma}b_{m\sigma}\rangle_\Phi\,,$$

$$\gamma_\nu-\beta_\nu=\langle b^\dagger_{n\sigma}a_{m\sigma}\rangle_\Phi\,,\qquad \nu=1,2,3. \tag{37}$$

Here $\mathbf{m}=\mathbf{n}+\mathbf{l}_\nu$, $\mathbf{l}_\nu=(0,0)$, (1,1), or (2,0) for $\nu=0,1,2$. The parameter $\beta_0\equiv 0$ is not incorporated into $\{y_i\}$, and $\zeta_\sigma=\sigma/|\sigma|=\pm 1$.

From the dependence of $\overline{H(y_i)}$, we find the linearized Hamiltonian

$$\widetilde{H}_L=\sum_{k\sigma}\hat{h}_{k\sigma}+\overline{H}-\sum_\nu\frac{\partial\overline{H}}{\partial y_\nu}y_\nu\,, \tag{38}$$

$$\hat{h}_{k\sigma}=A_k\delta_{i,j}+\begin{pmatrix}-\zeta_\sigma\Delta_k & V'_k+iV''_k\\ V'_k-iV''_k & \zeta_\sigma\Delta_k\end{pmatrix}_{ij}\begin{pmatrix}a^\dagger_{k\sigma}\\ b^\dagger_{k\sigma}\end{pmatrix}_i\begin{pmatrix}a_{k\sigma}\\ b_{k\sigma}\end{pmatrix}_j, \tag{39}$$

$$\{A_k,\Delta_k,V'_k\}_\mu=\frac{1}{4}\sum_{\nu=0}^{2}\left\{\frac{\partial\overline{H}}{\partial\rho_\nu},\frac{\partial\overline{H}}{\partial\delta_\nu},\frac{\partial\overline{H}}{\partial\gamma_\nu}\right\}_\mu \times\cos k_xl_x\cos k_yl_y\,,$$

$$V''_k=\frac{1}{4}\sum_{\nu=0}^{2}\frac{\partial\overline{H}}{\partial\beta_\nu}\sin k_xl_x\cos k_yl_y\,.$$

In the summands, $l=l(\nu)=(0,0)$, (1,1), or (2,0) for $\nu=0,1,2$. The eigenfunctions and energies of the linearized problem (38) and (39)

$$\chi_{\lambda k\sigma}=\frac{1}{\sqrt{2}}\begin{pmatrix}c_k+\zeta_\sigma s_k & c_k-\zeta_\sigma s_k\\ c_k-\zeta_\sigma s_k & -c_k-\zeta_\sigma s_k\end{pmatrix}_{ij}\begin{pmatrix}e^{i\theta}a^\dagger_{k\sigma}\\ e^{-i\theta}b^\dagger_{k\sigma}\end{pmatrix}_j, \tag{40}$$

$$E_k^{1(2)}=A_k\mp[\Delta^2+V_0^2]^{1/2},\qquad V_0^2=(V'_k)^2+(V''_k)^2, \tag{41}$$

$$c_k=\cos\phi_k\,;\qquad s_k=\sin\phi_k\,;\qquad \tan 2\phi_k=-\Delta_k/V_0\,;$$

$$\tan 2\theta_k=V''_k/V'_k$$

in turn make it possible to calculate the means (36), and complete the self-consistency procedure.

The density of states $\langle n_k\rangle_\Psi=\langle\widetilde{n}_k\rangle_\Phi$ for a correlated state with valence-bond dimer structure can be calculated in the same way as the mean energy. If we limit consideration to the linear approximation in the nonlocal one-electron means $\rho_l,\delta_l,\gamma_l,\beta_l$, $l\neq 0$, but take full account of all orders in the local means $(l=0)$, we obtain

$$\langle\widetilde{n}_k\rangle_\Phi=\rho_0+\cos k_xG_0+\sum_{\mu=1}^{4}G_\mu\{\rho_k,\delta_k,\gamma_k,\beta_k\}_\mu\,. \tag{42}$$

Here the $G_\mu=G_\mu(\rho_0,\delta_0,\gamma_0)$, with $\mu=0,\ldots,4$ are explicit, albeit cumbersome, functions of the local means (36), and the remaining quantities in the summand are determined by the occupation functions $f_{\lambda k}=f_F(E_{\lambda k}-\mu)$ of the energy bands in the linearized problem:

$$\rho_k=\frac{1}{2}(f_{1k}+f_{2k}),\qquad \delta_k=\frac{1}{2}(f_{1k}-f_{2k})\sin 2\phi_k\,,$$

$$\gamma_k+i\beta_k=\frac{1}{2}(f_{1k}-f_{2k})e^{2i\theta_k}\cos 2\phi_k\,.$$

The equations defining ϕ_k and θ_k are given above.

APPENDIX 2

In calculating the effective hole excitations above the ground-state background, where the latter has valence-bond dimer structure, we take advantage of the smallness of these states' nonlocal one-electron means (36) $(l\neq 0)$ in comparison with the local intradimer means $(l=0)$. In constructing the effective hole excitations, we therefore base the initial set of operators $q_{jk\sigma}$ on local operators belonging to the nth dimer:

$$q_{jk\sigma}=\sqrt{\frac{2}{N}}\sum_n e^{ikn}q_{jn\sigma}\,,$$

$$q_{jn\sigma}=\{g_{n\sigma},u_{n\sigma},g_{n\sigma}n^{g}_{n,-\sigma}(1-n^{u}_{n\sigma})(1-n^{u}_{n-\sigma}),\quad u_{n\sigma}n^{u}_{n,-\sigma}(1-n^{g}_{n\sigma})(1-n^{g}_{n,-\sigma}),$$
$$u^{\dagger}_{n,-\sigma}g_{n,-\sigma}g_{n\sigma}(1-n^{u}_{n\sigma}),\quad g^{\dagger}_{n,-\sigma}u_{n,-\sigma}u_{n\sigma}(1-n^{g}_{n\sigma})\}_{j}. \tag{43}$$

The operators $q_{jn\sigma}$ with $j>2$ are generated from the one-particle operators $g_{n\sigma},u_{n\sigma}$ by commutation with the transformation operator W.

The matrices

$$s_{ij}=\langle\Phi_0 q^{\dagger}_{in\sigma}|q_{jn\sigma}\Phi_0\rangle,$$

$$H_{ij}=\langle\Phi_0 q^{\dagger}_{in\sigma}|\widetilde{H}|q_{jn\sigma}\Phi_0\rangle,$$

can be calculated in the linear approximation in the nonlocal means (for example, with $n=1$, $U/t=8$ for the local ($l=0$) and nonlocal ($l\neq 0$) means y_i in (36), we have (0.5, 0.399, 0.227) and (0, -0.036, 0.042, -0.06, 0, -0.022, 0.049, 0.056)). In this approximation, the matrix elements s_{ij} and H_{ij} can be expressed in the form of expansions analogous to (42) for $\langle n_k\rangle$, with appropriate coefficients in place of G_{μ}. For every pair (i,j), these coefficients are explicit functions of the local means. As a result, s_{ij} and H_{ij} are completely known from the solution of the linearized problem $\widetilde{H}_L$.

Since $\rho_k,\delta_k,\gamma_k,\beta_k$ in expansions like (42) are proportional to the Fermi occupation functions $f_{\lambda k}=f_F(E_{\lambda k}-\mu)$ of the one-electron levels for $|\Phi_0\rangle$, the number of nonvanishing eigenvalues of the hermitian matrix s_{ij} calculated in such an approximation depends on the location of **k** relative to the Fermi surface. In particular, for $n=1$ we have three nonvanishing eigenvalues S_{λ}.

Steps in the calculation include:

1) calculating the matrices s_{ij} and H_{ij};

2) finding the eigenvalues and eigenvectors of the hermitian, positive-definite matrices

$$s_{ij}u_{j\lambda}=u_{i\lambda}S_{\lambda};$$

3) identifying among the latter the eigenvectors numbered $\lambda=1,\dots,L_0(k)$ with nonvanishing norm $S_{\lambda}\neq 0$;

4) constructing from these vectors the basis for the operators

$$Q_{\lambda k\sigma}=\sum_i q_{ik\sigma}u_{i\lambda},\quad \lambda=1,\dots,L_0(k)$$

with nonvanishing norm (26), which can be used to construct hold excitations;

5) calculating rank-L_0 matrices (28), equal to

$$S_{\lambda}^{-1/2}H_{\lambda\lambda'}S_{\lambda'}^{-1/2}=S_{\lambda}^{-1/2}u^{*}_{i\lambda}h_{ij}u_{j\lambda'}S_{\lambda'}^{-1/2},$$

the diagonalization of which specifies the vectors (25) and energies $E_{k,\nu}$ of effective hole (multiparticle) excitations for given k and a given initial set of operators $q_{jk\sigma}$. In similar fashion, one can also calculate the amplitudes $M^{\nu}(k_e)$ in the spectral functions (29).

Note that in the projection-operator method,[29] the analogs of the matrices s_{ij} and H_{ij} are

$$s_{ij}(k)=\langle[q^{\dagger}_{ik\sigma},q_{jk\sigma}]_{+}\rangle_{\Psi},\quad F_{ij}=\langle[q^{\dagger}_{ik\sigma},[H,q_{jk\sigma}]_{-}]_{+}\rangle_{\Psi}$$

instead of our matrices

$$s^{h}_{ij}=\langle\widetilde{q}^{\,\dagger}_{ik\sigma}\widetilde{q}_{jk\sigma}\rangle_{\Psi},\quad s^{e}_{ij}=\langle\widetilde{q}_{ik\sigma}\widetilde{q}^{\,\dagger}_{jk\sigma}\rangle_{\Psi},$$

which are separately used for hole and electron excitations. In the method used by Unger and Fulde,[29] effective electron and hole excitations are simultaneously determined from the poles of the function $\langle\{S^{-1/2}FS^{-1/2}-\omega\}^{-1}\rangle$.

[1] E. Dagotto, Rev. Mod. Phys. **66**, 763 (1994).
[2] J. W. Allen, R. Claessen, and R. O. Anderson *et al.*, in *The Physics of the Hubbard Model*, D. K. Campbell, J. M. P. Carmelo, and F. Guinea (Eds.), Plenum, New York (1994).
[3] R. O. Anderson, R. Claessen, J. W. Allen *et al.*, Phys. Rev. Lett. **70**, 3163 (1993).
[4] D. M. King, Z.-X. Shen, D. S. Dessau *et al.*, Phys. Rev. Lett. **70**, 3159 (1993).
[5] K. Gofron, J. C. Campuzano, H. Ding, C. Gu *et al.*, J. Phys. Chem. Solids **54**, 1193 (1993).
[6] D. M. King, Z.-X. Shen, D. S. Dessau *et al.*, Phys. Rev. Lett. **73**, 3289 (1993).
[7] D. S. Dessau, Z.-X. Shen, D. M. King *et al.*, Phys. Rev. Lett. **71**, 2781 (1993).
[8] H. Ding, J. C. Campuzano, A. F. Bell *et al.*, Phys. Rev. Lett. **74**, 2284 (1995); Phys. Rev. B **54**, 9678 (1996).
[9] T. E. Mason, G. Aeppli, S. M. Hayden *et al.*, Phys. Rev. Lett. **71**, 919 (1993).
[10] G. Senator and N. H. March, Rev. Mod. Phys. **66**, 445 (1994).
[11] Yu. A. Izyumov, M. I. Katsnel'son, and Yu. N. Skryabin, *Magnetism of Collective Electrons* [in Russian], Nauka, Moscow (1994).
[12] M. Hamada, H. Shimahara, and H. Mori, Phys. Rev. B **51**, 11597 (1995).
[13] F. Hu, S. K. Sarker, and C. Jayarakash, Phys. Rev. B **50**, 17901 (1994).
[14] X.-G. Wen, F. Wilczek, and A. Zee, Phys. Rev. B **39**, 11413 (1989).
[15] E. Dagotto, A. Moreo, F. Ortolani *et al.*, Phys. Rev. Lett. **67**, 1918 (1991).
[16] E. Dagotto, A. Nazarenko, and M. Boninsegni, Phys. Rev. Lett. **73**, 728 (1994).
[17] S. Haas, A. Moreo, F. Ortolani *et al.*, Phys. Rev. Lett. **74**, 4281 (1995).
[18] P. Aebi *et al.*, Phys. Rev. Lett. **72**, 2757 (1994).
[19] D. Duffy and A. Moreo, Phys. Rev. B **75**, 11882 (1995).
[20] R. Preuss, W. Hanke, and W. von der Linden, Phys. Rev. Lett. **75**, 1344 (1995).
[21] P. W. Anderson, Science **235**, 1196 (1987).
[22] E. J. Mele, Phys. Rev. B **38**, 8940 (1988).
[23] I. I. Ukrainskii, Int. J. Quantum Chem. **52**, 413 (1994).
[24] A. A. Ovchinnikov and M. Ya. Ovchinnikova, Zh. Éksp. Teor. Fiz. **110**, 342 (1996) [JETP **83**, 184 (1996)].
[25] H. B. Schuttler and A. J. Fedro, Phys. Rev. B **45**, 7588 (1992).
[26] J. H. Jefferson, H. Eskes, and L. F. Feiner, Phys. Rev. B **45**, 7959 (1992).
[27] J. E. Hirsch, Phys. Rev. Lett. **54**, 1317 (1985).
[28] M. C. Gutzwiller, Phys. Rev. A **137**, 1726 (1965).
[29] P. Unger and P. Fulde, Phys. Rev. B **48**, 16607 (1993).
[30] R. J. Radke and M. R. Norman, Phys. Rev. B **50**, 9554 (1995).
[31] M. R. Norman, M. Randeria, H. Ding, and J. C. Campuzano, Phys. Rev. B **52**, 615 (1995).
[32] T. Dahm, Phys. Rev. B **54**, 10150 (1996).
[33] F. Onufrieva, S. Petit, and Y. Sidis, Phys. Rev. B **54**, 12464 (1996).
[34] E. Dagotto, A. Mareo, F. Ortolani *et al.*, Phys. Rev. Lett. **67**, 1918 (1991).

Translated by Marc Damashek

Transformation of spin-wave resonance spectra in multilayer films upon passage through the Curie point of the locking layer

A. M. Zyuzin and A. G. Bazhanov

Mordovian State University, 430000 Saransk, Russia

(Submitted 24 January 1997)
Zh. Éksp. Teor. Fiz. **112**, 1430–1439 (October 1997)

A technique has been developed for controlling the degree of spin locking in two- and three-layer ferrite–garnet films. Variation of the degree of locking is achieved by varying the temperature and by passing one of the layers of the film through its Curie point. A fundamental difference in the nature of the transformation of the spin-wave resonance spectra taking place near the Curie point of the locking layer in films with initially symmetric and asymmetric boundary conditions is identified. © *1997 American Institute of Physics.* [S1063-7761(97)02010-6]

1. INTRODUCTION

A study of the effect of spin locking on the salient characteristics of spin-wave resonance spectra is undoubtedly of current interest. Essentially, all studies of spin-wave resonance start out from an analysis of boundary conditions. Despite the large number of publications dedicated to this question (see, e.g., Refs. 1–5), the potential of techniques for controlling the degree of spin locking and thereby varying the boundary conditions remains far from being realized. At the same time, such techniques allow one to obtain richer and more detailed information about the peculiarities of excitation and the salient characteristics of spin-wave resonance spectra in thin films.

In a number of works (see, e.g., Refs. 6 and 7) definite boundary conditions were created by depositing additional surface layers, whose magnetization M or effective field H_{eff} differs either on the high side or the low side from the values of these parameters in the main layer (the excitation layer of standing spin-wave modes). However, such an approach is not easily reproducible. In this case, control of the parameters of very thin surface layers is extremely difficult as a rule.

One way of controlling the degree of spin locking in one or both surfaces of the excitation layer consists in gradually reducing the thickness of the locking layer (spatial factor) by layered etching.[8]

Another technique for controlling the degree of spin locking that is at the same time reversible and smoothly controllable is possible, based on the following property. Let the Curie temperature T_C in the locking layer be substantially lower than in the excitation layer. As the temperature is raised and the Curie point of the locking layer is approached, there takes place a gradual weakening followed by complete destruction of magnetic ordering. The exchange interaction constant A and the magnetization M in this case fall to zero. The transition from the magnetically ordered to the paramagnetic state should lead to the disappearance of the spin locking mechanisms that exist below the Curie temperature. Of course, increasing the temperature will also alter the values of the physical parameters in the excitation layer. But since these parameters are easily controlled, such a technique can be used to vary the degree of spin locking and study the effect of the latter on the spin-wave resonance spectra. To provide greater accuracy in the interpretation of the results, it is necessary that the Curie temperature of the locking layer be substantially lower than the Curie temperature of the excitation layer of the spin-wave modes. Despite the clear potential of this approach, we know of no studies which have implemented such an approach for varying the degree of spin locking.

The aim of the present paper is to investigate the transformation of spin-wave resonance spectra accompanying the transition through the Curie point of the locking layer in multilayer films with initially symmetric and asymmetric boundary conditions.

2. EXPERIMENT

We investigated two- and three-layer single-crystal films of ferrite–garnets in which the dominant spin-locking mechanism is dissipation.[9] The films were grown by liquid-phase epitaxy on gadolinium–gallium garnet substrates with orientation (111) by successive submersion in two or three different melts. Simultaneously, as a control we grew single-layer analogs of each of the layers of the multilayer films on clean substrates. The first layer (the locking layer) of the two-layer film with composition $(\mathrm{YSmLuCa})_3(\mathrm{FeGe})_5\mathrm{O}_{12}$ and thickness $h=2.1\ \mu$m had Curie temperature $T_C=215$ °C, saturation magnetization $4\pi M=470$ G, effective uniaxial anisotropy field $H_{\mathrm{eff}}=1210$ Oe, and Gilbert decay parameter $\alpha=\Delta H\gamma/\omega=0.15$. Here ΔH is the half-width of the absorption line, γ is the gyromagnetic ratio, ω is the angular frequency of the microwave field. The second layer had composition $\mathrm{Y}_{2.98}\mathrm{Sm}_{0.02}\mathrm{Fe}_5\mathrm{O}_{12}$. This was the layer in which standing harmonic spin-wave modes were excited and it had the following parameters: $T_C=280$ °C, $4\pi M=1740$ G, $H_{\mathrm{eff}}=-1715$ Oe, $\alpha=0.003$, and $h=0.5-0.85\ \mu$m.

The three-layer films were prepared by growing a third layer onto part of the two-layer film which had the composition $(\mathrm{SmEr})_3\mathrm{Fe}_5\mathrm{O}_{12}$, $T_C=280$ °C, $4\pi M=1330$ G, $H_{\mathrm{eff}}=96$ Oe, $\alpha=0.2$, and $h=1.1\ \mu$m. Three-layer films were also grown in which the lower and upper locking layers

1063-7761/97/100778-06$10.00

had the same composition, $(SmEr)_3Fe_5O_{12}$, and $T_c = 280$ °C. The saturation magnetization was measured by the technique described in Ref. 10. The thickness was measured on the single-layer analogs by the interference method, and for the thin films it was also estimated from the total etching time. The spin-wave resonance spectra were recorded on a radiospectrometer (resonator with TE_{011} mode) with perpendicular and parallel orientations of the constant magnetic field relative to the plane of the film. Spin-wave resonance was excited by a linearly polarized microwave field with frequency 9.34×10^9 Hz. The resonance fields of the spin-wave modes were measured with an NMR magnetometer. The Curie temperature was determined from the temperature of disappearance of the homogeneous resonance peak of the corresponding layer. To create the required sample temperatures, a thermally insulated, thin-walled quartz tube with diameter 8 mm was placed inside the resonator. Air that was passed through a heater was blown onto the sample. As the sensor in the temperature stabilization system we used a platinum resistance thermometer. To monitor the temperature in front of and behind the sample, two platinum/platinum–rhodium thermocouples were placed near the resonance cavity in the tube. From the mean value of the thermocouple voltage we determined the sample temperature, which was kept constant to an accuracy no worse than ± 2 °C.

Since the spin-wave resonance spectra were recorded at constant frequency ω, the dispersion curves were constructed as a dependence of the difference of the resonance fields of the zeroth and nth modes on $(2n+1)^2$ for the three-layer films and $(n+1/2)^2$ for the two-layer films. In such films the boundary conditions are respectively symmetric and asymmetric.

3. EXPERIMENTAL RESULTS

Temperature studies of the spin-wave resonance spectra of two- and three-layer films revealed the following. At room temperature, eight (including the zeroth) spin-wave modes were clearly recorded in the spectrum of the two-layer film, which consisted only of a locking layer with $T_C = 215$ °C and an excitation layer (Fig. 1a). The dispersion curves $H_0 - H_n = f((n+1/2)^2)$ obeyed a quadratic law, by virtue of which they remained linear over a wide temperature interval. With increasing temperature, starting roughly at 140 °C, the number of spin-wave modes gradually decreased (Fig. 1b). Modes successively disappeared, starting with the largest mode number, and the amplitudes of the remaining modes decreased. In the interval from 215 to 280 °C in the spectrum of this film only one mode (the zeroth) remained (Fig. 1c).

The transformation of the spectra of the three-layer films, taking place with increasing temperature, had a fundamentally different character and was accompanied not by a decrease, but an increase of the number of spin-wave modes. Such films differed from the two-layer films by the presence of a third layer, which like the first was a locking layer but had the composition $(SmEr)_3Fe_5O_{12}$ with $T_C = 280$ °C.

The transformation proceeded as follows. Starting at the same temperature $T \simeq 140$ °C, intermediate peaks appeared between the peaks corresponding to the high-order modes. Increasing the temperature led to growth of the amplitudes of the peaks that had already appeared and to subsequent appearance of peaks corresponding to intermediate modes of lower order (Fig. 2b). As follows from the figure, in this temperature region the distribution of intensities of the absorption peaks of the spin-wave modes ceases to be monotonic. For $T \geqslant 215$ °C the total number of modes roughly doubled, and remained unchanged up to $T \sim 260$ °C (Fig. 2c). Increasing the temperature to $T = 280$ °C, the Curie temperature of the excitation layer, led to a decrease in the number of modes and to total disappearance of the entire spectrum. As can be seen from Fig. 3, the variation of the number of modes $N(T)$ as the Curie temperature of the excitation layer is approached is more abrupt than for approach to the Curie temperature of the locking layer. It can also be seen that a pronounced correlation occurs in the behavior of the dependences $N(T)$ for the two- and three-layer films.

One of the characteristic differences in the transformation of the spin-wave resonance spectra of the three-layer films is that in the temperature region from 140 to 215 °C a knee appears in the dispersion curve, which with growth of the temperature and the appearance of new intermediate modes, shifts into the region of small n. This is illustrated in Fig. 4, which presents dispersion curves for different temperatures of a three-layer film with thickness of the excitation layer $h = 0.52$ μm. For $T \geqslant 215$ °C, i.e., after the appearance of all the intermediate modes, the knee in the dispersion curve disappeared, but the slope angle of the dispersion curve differed by roughly a factor of two from that of the curve corresponding to room temperature. At temperatures below 140 °C and above 215 °C the dispersion curves were straight lines. Only their slope angles decreased, due to the faster (in comparison with the magnetization M) decrease of the exchange interaction constant A with temperature in the excitation layer (Fig. 5).[11]

For comparison, we also investigated the effect of temperature on the spectra of the three-layer films in which the first and third layers had the composition $(SmEr)_3Fe_5O_{12}$ and the same Curie temperature T_C, equal, as in the excitation layer, to 280 °C. The number of modes in the spectra of such films remained unchanged over a wide temperature interval (see Fig. 3, curve *3*). Only with approach to T_C did their number gradually decrease, and for $T > T_C$ the entire spin-wave resonance spectrum disappeared.

As the temperature was lowered, the transformation of the spin-wave resonance spectra of the investigated films took place in the reverse order.

4. CALCULATION OF SPIN-WAVE RESONANCE SPECTRA

The dominant spin locking mechanism in the investigated films was the dissipative mechanism,[9] which in contrast to the dynamic mechanism[12] renders the localization region of the standing harmonic spin waves independent of the orientation of the external magnetic field **H** relative to the film. Such a locking mechanism arises for excitation of variable magnetization in two- or three-layer films with highly

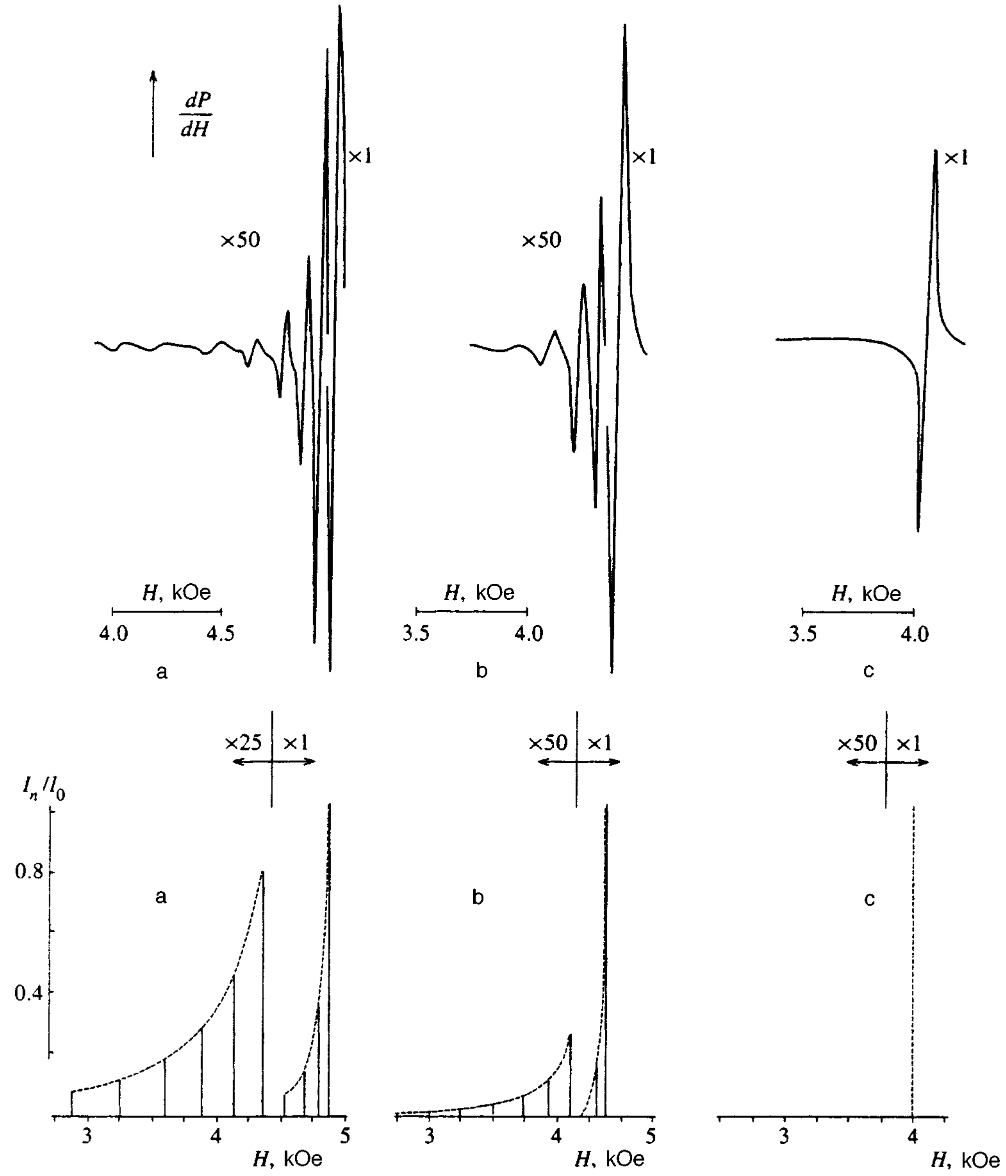

FIG. 1. Above—experimental spin-wave resonance spectra of a two-layer film (perpendicular orientation) at 20 °C (a), 185 °C (b), and 240 °C (c). Below—calculated spectra and intensities of the spin-wave modes for various values of β/k_1h: a) 5×10^{-8} cm, b) 5×10^{-7} cm, c) 5×10^{-5} cm.

different decay parameters in the layers. The presence of exchange coupling between the layers, and also the fact that even under conditions of homogeneous resonance, the deviation of the magnetization vector **M** from its equilibrium value in the layer with large decay parameter α_d is α_d/α_e times smaller than in the layer with the small decay parameter α_e, leads to the appearance of a node of the standing spin wave near the interlayer boundary. Note also that the dispersion or reactive (elastic) properties of the layers with large α have an additional effect on spin locking.[13,14]

We based our treatment and explanation of the temperature-induced transformation of spin-wave resonance spectra of multilayer films on the following assumptions.

Standing spin waves, excited by a uniform microwave field in the layer with small decay parameter α, are harmonic.

In layers with strong absorption (the locking layers) the spin waves are excited by waves localized in the layer with small α, and are decaying. The depth of penetration (mean free path) of these waves is related to the wave number of the harmonic modes in the excitation layer.

Noting the harmonic time dependence of the variable magnetization $m(z,t)$, its spatial configuration in the excitation layer for perpendicular configuration relative to the film can be described by the equation[12,15]

$$\frac{d^2m}{dz^2}-\frac{M}{2A}\left(H-\frac{\omega}{\gamma}+H_k^{\text{eff}}-\frac{2}{3}H_{k1}\right)m=0, \tag{1}$$

where H is the external constant field and H_{k1} is the cubic anisotropy field. This equation is valid for a medium with small decay. This condition is satisfied by the parameters of the excitation layer.

The solutions of Eq. (1) have the form

$$m_1(z)=B_1\sin k_1z+B_2\cos k_1z. \tag{2}$$

where k_1 is the wave number and B_i are constants.

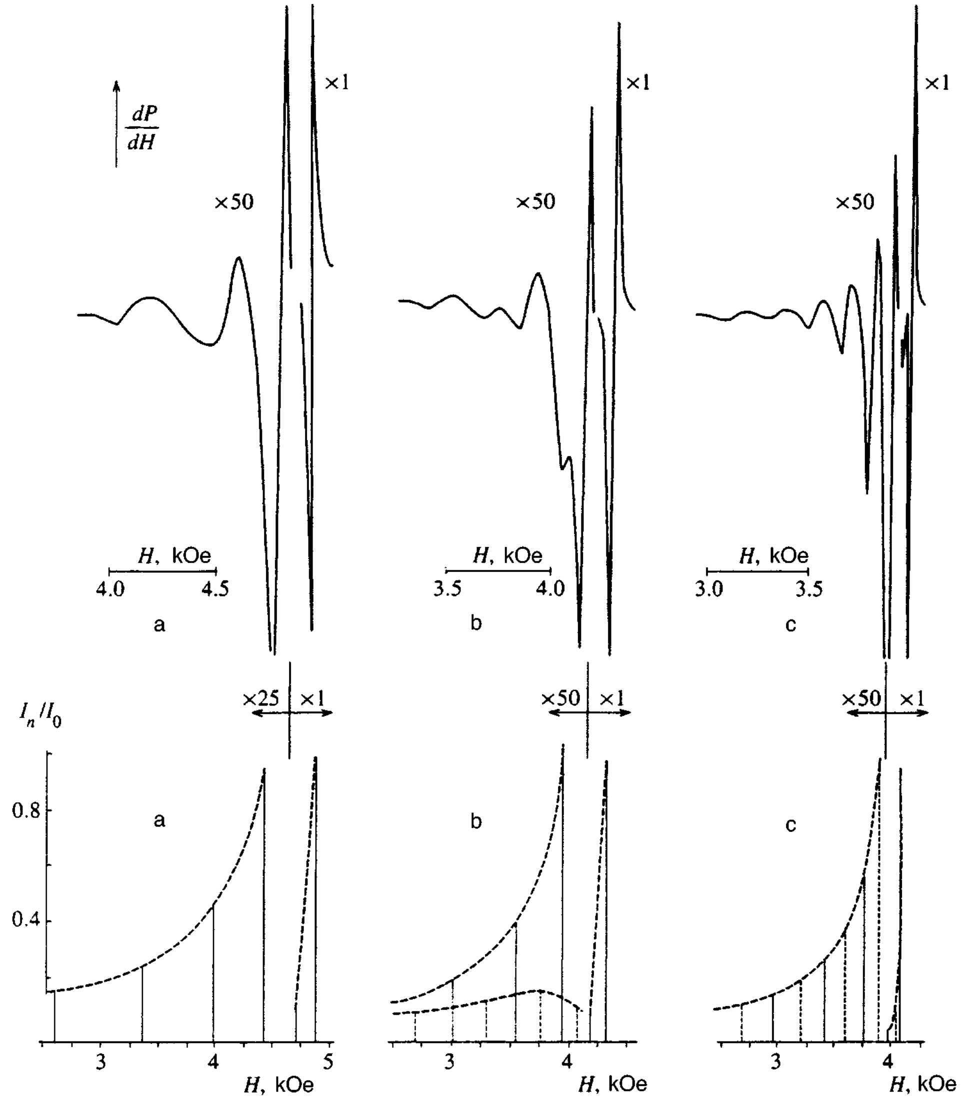

FIG. 2. The same as in Fig. 1, but for a three-layer film (perpendicular orientation). The vertical dashed lines correspond to intermediate modes.

In the locking layers we approximate the spin wave as being exponentially decaying with distance from the interlayer boundary:

$$m_2(z)=C_1e^{k_2''z},\qquad z<0,$$

$$m_2(z)=C_2e^{-k_2''(z-h)},\qquad z>h, \tag{3}$$

where $z=0$, h corresponds to the boundaries of the excitation layer. Here $k_2''=1/l$, where l is the penetration depth of the spin wave in the locking layer, which, as was assumed, grows linearly with k_1 (Ref. 16). To determine the relations between the constants B_i, C_i and to find acceptable values of the wave numbers, we used the exchange boundary conditions on m (Ref. 15), which must be satisfied by the magnetization and its first derivative on one boundary of the excitation layer $(z=0)$, where locking is large and hardly varies up to $T\simeq 280$ °C, and on the other boundary $(z=h)$, where the degree of locking varies substantially with approach to $T=215$ °C and then completely disappears:

$$\frac{m_1}{M_1}=\frac{m_2}{M_2},\qquad \frac{A_1}{M_1}\frac{dm_1}{dz}=\frac{A_2}{M_2}\frac{dm_2}{dz}. \tag{4}$$

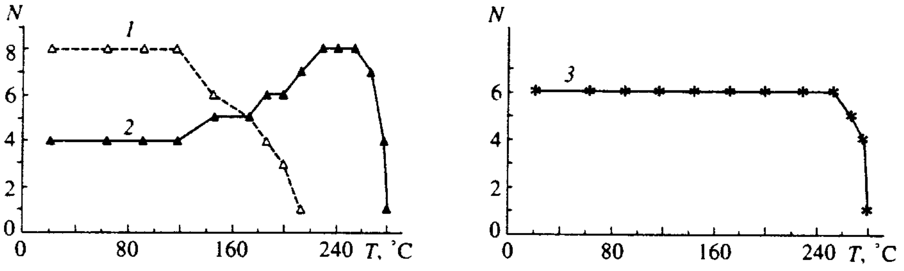

FIG. 3. Temperature dependence of the number of spin-wave modes: *1*) in a two-layer film, *2*) in a three-layer film, *3*) in a three-layer film in which both locking layers have the composition $(SmEr)_3Fe_5O_{12}$ with $T_C=280$ °C.

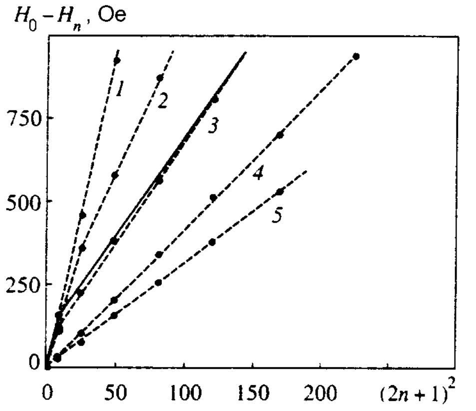

FIG. 4. Dispersion curves $H_0 - H_n = f((2n+1)^2)$ of a three-layer film at 20 °C (*1*), 185 °C (*2*), 198 °C (*3*), 240 °C (*4*), 267 °C (*5*). The points correspond to experiment, the solid curve *3*—to calculation for $\beta/k_1h = 5\times10^{-7}$ cm.

Here the indices 1 and 2 denote respectively the parameters of the excitation and locking layer.

Taking into account that locking at one of the boundaries of the excitation layer remained strong up to $T \simeq 280$ °C, to simplify the calculations we assumed the presence of a node on this boundary, i.e., $m_1|_{z=0} = 0$. This approximation does not have a noticeable effect on the calculated values of the peak intensities of the spin-wave modes or their resonance fields. The equation, obtained with the help of Eqs. (2)–(4) and defining the set of acceptable values of k_1, has the form

$$\tan(k_1h) = -\frac{A_1}{A_2}\frac{k_1}{k_2''} = -\beta k_1, \qquad (5)$$

where $\beta = A_1/A_2k_2''$ is a parameter defining the degree of spin locking, which, noting the above assumption that $k_2'' \propto k_1^{-1}$, is seen to be proportional to k_1.

The solutions of Eq. (5) allow us to find the eigenvalues H of Eq. (1), i.e., the resonance fields of the possible spin-wave modes and to determine the wave numbers and configuration of the latter.

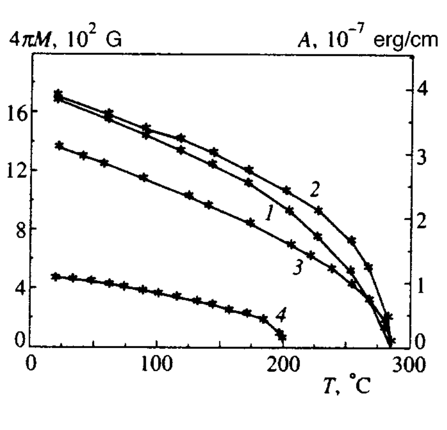

FIG. 5. Temperature dependence of A (curve *1*) and $4\pi M$ (curve *2*) for the excitation layer. Curves *3* and *4* are the dependence of $4\pi M(T)$ in the locking layers with composition $(SmEr)_3Fe_5O_{12}$ and $(YSmLuCa)_3(FeGe)_5O_{12}$, respectively.

To determine those spin-wave modes which can be excited by a uniform microwave field from among those found to satisfy the exchange boundary conditions, we calculated the intensities I_n of the corresponding peaks. In the case of a linearly polarized microwave field directed along the x axis, we have[6,15]

$$I_n = K_0 \frac{\left(\int m_{xn}dz\right)^2}{\int \frac{\alpha}{2\gamma M}(m_{xn}^2 + m_{yn}^2)dz}, \qquad (6)$$

where K_0 is a constant. In the calculation the integration is limited to the thickness of the excitation layer. It can be shown that the main contribution to the interaction integral between the microwave field and the magnetization is determined by precisely this layer. This is due, first of all, to the smallness of the exponential decay region of the spin waves in the layers with large α and, second, to the fact that the phase of the magnetization oscillations in this region is shifted relative to the phase of the magnetization oscillations in the excitation layer. At the same time, it should be remarked that despite its smallness the indicated region, due to the large value of its decay parameter α, is an additional energy dissipation channel of the standing harmonic waves and affects the width of the absorption peaks of the spin-wave modes, especially the higher-order modes.

5. DISCUSSION

Results of our calculation of the intensities and resonance fields of the spin-wave modes of a three-layer film for various values of the parameter β/k_1h are presented in Fig. 2. In the case of strong locking at both boundaries, which corresponds to symmetric boundary conditions, only modes with an odd number of half-waves over the thickness of the excitation layer are excited. The intensity of the modes with even number, although they are allowed by the boundary conditions, is equal to zero. Calculation confirms this conclusion (Fig. 2a). An increase in the temperature and a faster decrease in A and M of the layer with $T_C = 215$ °C in comparison with the other layers destroys the equality between the degrees of locking at the boundaries and thereby the symmetry of the boundary conditions. An increase in the parameter β alters the roots of Eq. (5). This is accompanied by a change in the spatial phase of the standing spin wave at this boundary. In turn, a change in the phase has the result that the total variable magnetic moment of the modes that under symmetric conditions had an even number of half-waves becomes nonzero and corresponding intermediate peaks appear in the spin-wave resonance spectrum. As follows from Fig. 2b, which displays both the experimental and calculated spectra corresponding to one of the intermediate situations, the distribution law of the intensities of the intermediate modes differs from the distribution law of the mode intensities for symmetric or asymmetric boundary conditions: $I_n \propto (1/n^2 \ldots 1/n^4)$ (Ref. 16). As can be seen from Fig. 2b, the intermediate modes between the previously observed zeroth mode and first mode, and also between the first mode and the second mode, have very small relative intensity. This has the result that intermediate modes are recorded in the spectrum, starting only at some mode number. This is the reason for the appearance of the salient point in the dispersion curve, which with a decrease in the degree of locking is shifted toward spin-wave modes with lower number. If the modes whose intensities amount, for example, to less than 10% of the intensity of the subsequent mode were not taken into account in the calculated spectra, this would also lead to the appearance of a knee in the dispersion curve (curve *3* in Fig. 4). Note that such a knee, explained by fluctuations in the exchange constant and the magnetization in amorphous and polycrystalline films, was observed in Refs. 17 and 18. Passing the layer with $T_C = 215$ °C through its Curie point leads to the disappearance of spin locking at the corresponding boundary. The boundary conditions become asymmetric: on one boundary the spins are locked, on the other they are free. In this case all modes are excited which have an odd number of quarter waves over the thickness of the excitation layer, not half-waves as is the case with symmetric boundary conditions. As a result, over the same interval of wave numbers the number of excited spin-wave modes roughly doubles. The intensity distribution becomes monotonic (Fig. 2c). The set of wave numbers corresponding to these modes begins to correspond to the formula $k=(\pi/h)(n+1/2)$, where $n=0,1,2...$ is the mode number. Such is the mechanism of transformation of the spin-wave resonance spectrum taking place with increasing temperature in the three-layer film.

A decrease in the degree of locking at the boundary of the excitation layer of a two-layer film leads to a gradual decrease in the efficiency of excitation of spin-wave modes, where, to start with, this decrease affects first the higher-order modes, as a consequence of which they cease to be recorded in the spectrum. Results of our calculation of the spin-wave resonance spectra of the two-layer film for different degrees of locking, which corresponds to different temperatures, are also found to be in good agreement with the experimental data (Fig. 1).

Note that the nature of the transformation of the spin-wave resonance spectra for perpendicular and parallel orientations was analogous. For both orientations there was a gradual appearance of intermediate spin-wave modes in the spectrum of the three-layer film. At the same time, for the perpendicular orientation, the high degree of locking for the modes with small mode number was preserved up to higher temperatures. This has to do with the fact that for this orientation the locking layer continues to be a reactive (elastic) medium with strong dissipation all the way to T_C, while for parallel orientation it is dispersive.

To summarize, in this paper we have shown that

—varying the temperature and passing one or more layers of a multilayer film through the Curie point can be used as a means of controlling the degree of spin locking;

—the disappearance of spin locking at one of the boundaries of a three-layer film roughly doubles the number of observed spin-wave modes. This has to do with making the transition from symmetric to asymmetric boundary conditions and with the excitation of previously forbidden intermediate modes. The correlation observed between the temperature dependences of the number of spin-wave modes in two- and three-layer films confirms this conclusion;

—both a growth and a decrease in the number of spin-wave modes can be observed, depending on the initial symmetry of the boundary conditions as the temperature is raised and the Curie point of the locking layer is approached;

—the non-equivalence of the degree of locking at the surfaces of the excitation layer and its dependence on mode number together are one of the reasons for the appearance of a knee in the dispersion curves.

[1] Yu. A. Korchagin, R. G. Khlebopros, and N. S. Chistyakov, Fiz. Tverd. Tela **14**, 2121 (1972) [Sov. Phys. Solid State **14**, 1826 (1972)].
[2] K. Vayhinger and H. Kronmuller, J. Magn. Magn. Mater. **72**, 307 (1988).
[3] J. Barnas, J. Magn. Magn. Mater. **102**, 319 (1991).
[4] S. L. Vysotskiĭ, G. T. Kazakov, and M. L. Kats, Fiz. Tverd. Tela **35**, 1190 (1993) [Phys. Solid State **35**, 606 (1993)].
[5] B. Hillebrands, Phys. Rev. B **41**, 530 (1990).
[6] N. M. Salanskiĭ and M. Sh. Erukhimov, *Physical Properties and Application of Magnetic Films* [in Russian], Nauka, Novosibirsk (1975).
[7] R. S. Iskhakov, M. M. Brushtunov, and A. S. Chekanov, Fiz. Tverd. Tela **29**, 2699 (1987) [Sov. Phys. Solid State **29**, 1553 (1987)].
[8] A. M. Zyuzin and A. G. Bazhanov, Zh. Éksp. Teor. Fiz. **111**, 1667 (1997) [JETP **84**, 912 (1997)].
[9] A. M. Zyuzin, N. N. Kudel'kin, V. V. Randoshkin, and R. V. Telesnin, Pis'ma Zh. Tekh. Fiz. **9**(3), 177 (1983) [Sov. Tech. Phys. Lett. **9**, 78 (1983)].
[10] A. M. Zyuzin, V. N. Van'kov, and V. V. Radaĭkin, Pis'ma Zh. Tekh. Fiz. **17**(23), 65 (1991) [Sov. Tech. Phys. Lett. **17**, 848 (1991)].
[11] A. M. Zyuzin and A. G. Bazhanov, JETP Lett. **63**, 555 (1996).
[12] B. Hoekstra, R. P. van Stapele, and J. M. Robertson, J. Appl. Phys. **48**, 382 (1977).
[13] F. S. Crawford, *Waves*, McGraw-Hill, New York (1968).
[14] M. A. Gintsburg, Fiz. Tverd. Tela **11**, 913 (1961) [*sic*].
[15] C. H. Wilts and S. Prasad, IEEE Trans. Magn. **MAG-17**, 2405 (1981).
[16] A. G. Gurevich, *Magnetic Resonance in Ferrites and Antiferromagnets* [in Russian], Nauka, Moscow (1973).
[17] V. A. Ignatchenko and R. S. Iskhakov, Zh. Éksp. Teor. Fiz. **74**, 1386 (1978) [Sov. Phys. JETP **47**, 725 (1978)].
[18] R. S. Iskhakov, A. S. Chekanov, and L. A. Chekanova, Fiz. Tverd. Tela **32**, 441 (1990) [Sov. Phys. Solid State **32**, 255 (1990)].

Translated by Paul F. Schippnick

Spin relaxation of Mn ions in (CdMn)Te/(CdMg)Te quantum wells under picosecond optical pumping

M. G. Tyazhlov, A. I. Filin, A. V. Larionov, and V. D. Kulakovskiĭ

Institute of Solid State Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

D. R. Yakovlev

A. F. Ioffe Physico-Technical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia

A. Waag and G. Landwehr

Physikalisches Institut der Universität Würzburg, Würzburg 97074, Germany

(Submitted 30 January 1997)
Zh. Éksp. Teor. Fiz. **112**, 1440–1463 (October 1997)

Spin relaxation of Mn ions in a $Cd_{0.97}Mn_{0.03}Te/Cd_{0.75}Mg_{0.25}Te$ quantum well with photogenerated quasi-two-dimensional electron–hole plasma at liquid helium temperatures in an external magnetic field has been investigated. Heating of Mn ions by photogenerated carriers due to spin and energy exchange between the hot electron–hole plasma and Mn ions through direct sd-interaction between electron and Mn spins has been detected. This process has a short characteristic time of about 4 ns, which leads to appreciable heating of the Mn spin subsystem in about 0.5 ns. Even under uniform excitation of a dense electron–hole plasma, the Mn heating is spatially nonuniform, and leads to formation of spin domains in the quantum well magnetic subsystem. The relaxation time of spin domains after pulsed excitation is measured to be about 70 ns. Energy relaxation of excitons in the random exchange potential due to spin domains results from exciton diffusion in magnetic field $B=14$ T with a characteristic time of 1 to 4 ns. The relaxation time decreases with decreasing optical pump power, which indicates smaller dimensions of spin domains. In weak magnetic fields ($B=2$ T) a slow down in the exciton diffusion to 15 ns has been detected. This slow down is due to exciton binding to neutral donors (formation of bound excitons) and smaller spin domain amplitudes in low magnetic fields. The optically determined spin–lattice relaxation time of Mn ions in a magnetic field of 14 T is 270 ± 10 and 16 ± 7 ns for Mn concentrations of 3% and 12%, respectively. © *1997 American Institute of Physics.* [S1063-7761(97)02110-0]

1. INTRODUCTION

Manganese atoms contain five unpaired electron spins in the $3d^5$ shell, and form a dilute magnetic subsystem in semimagnetic semiconductors.[1] In Mn atoms d-electrons efficiently interact with carriers in the valence band (pd-ineraction) and conduction band (sd-interaction) of the semiconductor, which leads to such effects as giant spin splitting of band states, giant Faraday rotation of the plane of polarized light, formation of magnetic polarons,[1,2] etc. Interaction among d-electrons of different Mn atoms within the magnetic subsystem results in the formation of antiferromagnetic clusters of nearest neighbors.[3]

Spin–lattice relaxation in the magnetic subsystem of bulk semiconductors has been intensely studied.[4–6] It is found that at liquid-helium temperatures, isolated Mn atoms interact weakly with phonons, resulting in long spin–lattice relaxation times in the millisecond range.[4] The presence of antiferromagnetic Mn clusters in semimagnetic semiconductor crystals increases the spin–lattice relaxation rate by several orders of magnitude.[5,6] Spin exchange between isolated Mn atoms and clusters is efficient owing to the spin–spin interaction. Clusters, in turn, interact more efficiently than isolated atoms with phonons. In semimagnetic semiconductors, the spin–lattice relaxation time decreases considerably with increasing Mn density and temperature. At liquid helium temperatures and Mn concentration of about 3%, the spin–lattice relaxation time is in the microsecond range.[5,6]

The subject of this publication is spin and energy exchange among the magnetic, electron, and phonon subsystems in a semimagnetic quantum well filled with high-density photogenerated electron–hole ($e-h$) magnetoplasma. Attention is focused on effects due to the direct spd-exchange interaction between electron and Mn spins. Owing to the small quantum well width, a high concentration of photocarriers can be generated without substantial lattice heating, which is impossible in bulk semiconductors.

The possibility of direct spin and energy exchange between Mn spins and spins of photogenerated carriers in semimagnetic semiconductors is obvious, and was studied theoretically long ago.[7] The experimental data concerning this topic, however, are scarce. Investigations have been performed only in bulk semimagnetic semiconductors, and have not yielded positive results,[8] since the density of photogenerated electrons was insufficient.

Significant recent progress in molecular beam epitaxy techniques has allowed manufacture of high-quality hetero-

1063-7761/97/100784-13$10.00

structures based on II–VI compounds and opened new avenues for studying phenomena at high densities of photoexcited carriers in quantum wells with isoelectronic magnetic impurities.[9–11]

In studying quantum wells with magnetic impurities at high densities of interband optical pumping, we detected significant heating of Mn when $e-h$ plasma of high density was generated in a short time of about 0.5 ns,[11] which led to a blue shift of the exciton recombination X-line in a magnetic field. Moreover, we found that the relaxation time of the X-line position after a powerful pump pulse depended on the technique used to measure the exciton spectrum. The relaxation time during luminescence decay immediately after a powerful optical pulse was $\simeq 4$ ns, whereas the relaxation time of the X-line position measured via low-power nonheating probe pulses, with a long delay following the pump pulse, was two orders of magnitude longer (about 300 ns). This long time is in agreement with the Mn spin relaxation time averaged over the quantum well due to spin–lattice coupling. To account for the short spin–lattice relaxation time, we suggested that the Mn spin temperature after a powerful pump pulse becomes nonuniform, that spin domains are formed in the magnetic subsystem of the quantum well, and that the energy relaxation of excitons takes place in the resulting random exchange potential.

In the present work, we have continued studying physical properties of spin domains formed in a $Cd_{1-x}Mn_xTe$ quantum well. We have used $Cd_{0.97}Mn_{0.03}Te/Cd_{0.75}Mg_{0.25}Te$ quantum wells 10 nm wide with a low concentration of Mn magnetic impurities and studied in detail the mechanism of spin–spin heating of the Mn subsystem, the effect of phonons on spin heating of magnetic impurities, and the energy relaxation of carriers in the exchange field due to spin domains.

2. EXPERIMENTAL TECHNIQUE

We measured an undoped $Cd_{0.97}Mn_{0.03}Te/Cd_{0.75}Mg_{0.25}Te$ heterostructure grown by molecular beam epitaxy on a CdTe substrate with the (001) orientation. The sample consisted of a set of isolated quantum wells, with widths $L_z = 1.8$, 4.5, 6, and 10 nm in order of increasing distance from the surface. The quantum wells were separated by (CdMg)Te barriers with a width of 100 nm. In order to obtain a sample with a single quantum well, the initial chip was chemically etched in a diluted solution of bromine in methanol, $CH_3OH:Br_2$. The resulting sample contained a single quantum well separated from the surface by a barrier with a width of ~ 100 nm. Some of the measurements were performed on a single $CdTe/Cd_{0.88}Mn_{0.12}Te$ quantum well with $L_z = 10$ nm. The sample was placed in a liquid helium cryostat equipped with a superconducting solenoid. The quantum well plane was perpendicular to the magnetic field. Photoluminescence spectra were recorded at $T = 4.2$ K in a magnetic field $B = 0-14$ T using a grating monochromator. Nonequilibrium carriers were generated by a pulsed dye-laser. The laser wavelength was 585 nm, the pulse repetition rate 0.8 MHz, and the pulse width 30 ps. For conducting laser light to the sample and collecting luminescence radiation, we used a quartz optical fiber with a diameter of 0.6 mm. The fiber end was placed close to the sample surface in order to generate uniform $e-h$ plasma and to collect the photoluminescence signal only from the excited region of the sample.[9–11] The photogenerated carrier density was derived from the shape of the luminescence line of the $e-h$ plasma taking into account the number of filled Landau levels.[9] The maximum carrier density was $\simeq 2\times 10^{12}$ cm^{-2} at a pulsed pump power density $P = 3$ μJ/cm^2.

Photoluminescence signal was detected by a time-correlated photon counting system with a time resolution of 0.3 ns build around a cooled photomultiplier tube. For studying Mn spin–lattice relaxation after delays longer than the carrier lifetime, we used additional weak probing photoexcitation. The carrier density in this case was at most 3×10^9 cm^{-2}. For this purpose, the laser was operated in a special mode. The dye-laser was pumped by the second harmonic of a Nd:YAG laser operated at a repetition rate of 82 MHz, and this sequence of pulses was diluted by a factor of 1:100. In order to generate hf probe pulses, the laser was run in a special mode with partial suppression of driving pulses. As a result, the dye-laser output was a sequence of probe pulses with a repetition rate of 82 MHz (12.2 ns between pulses) uniformly distributed between pump pulses generated at a rate of 0.8 MHz (1220 ns between pulses). The ratio between the intensities of the probe and pump pulses was about 1/1000, so that the total energy of probe pulses was at most 10% of the pump pulse energy.

3. Mn SPIN AVERAGED OVER THE QUANTUM WELL. SPIN–LATTICE RELAXATION

3.1. Photoluminescence spectra under cw pumping

Figure 1 shows photoluminescence spectra of a single $Cd_{0.97}Mn_{0.03}Te/Cd_{0.75}Mg_{0.25}Te$ quantum well with a width of 10 nm due to low-power over-barrier nonresonance cw excitation by an Ar^+ laser for various magnetic fields. The spectra were recorded at zero time delay, $\tau_d = 0$, and a temperature of the liquid helium bath $T_{\mathrm{bath}} = 4.2$ K. In the absence of magnetic field, only one relatively wide line with FWHM of ~ 4 meV due to recombination of excitons bound at neutral donors (D_0X complex) is detected.[12] The line notably shifts to the low-energy side by $\simeq 20$ meV as the magnetic field increases, and it becomes more narrow. Small shifts to the low-energy side are related to polarization of Mn magnetic impurity spins in applied magnetic field B and are caused by the spd-interaction in the conduction and valence bands of the quantum well. The shift ΔE is directly proportional to the average Mn spin $\langle S_z\rangle$, effective concentration of magnetic impurities N_0x^*, and the exchange integrals β and α in the valence and conduction bands, respectively[1,13]:

$$E(B) - E(0) = -\frac{1}{2}(|\beta| + |\alpha|)N_0x^*\langle S_z\rangle. \tag{1}$$

Here N_0 is the number of lattice cells per unit volume and x^* is the effective mole content of Mn. The parameter x^* is always smaller than the Mn content x because of formation of antiferromagnetic Mn clusters, which contribute little to

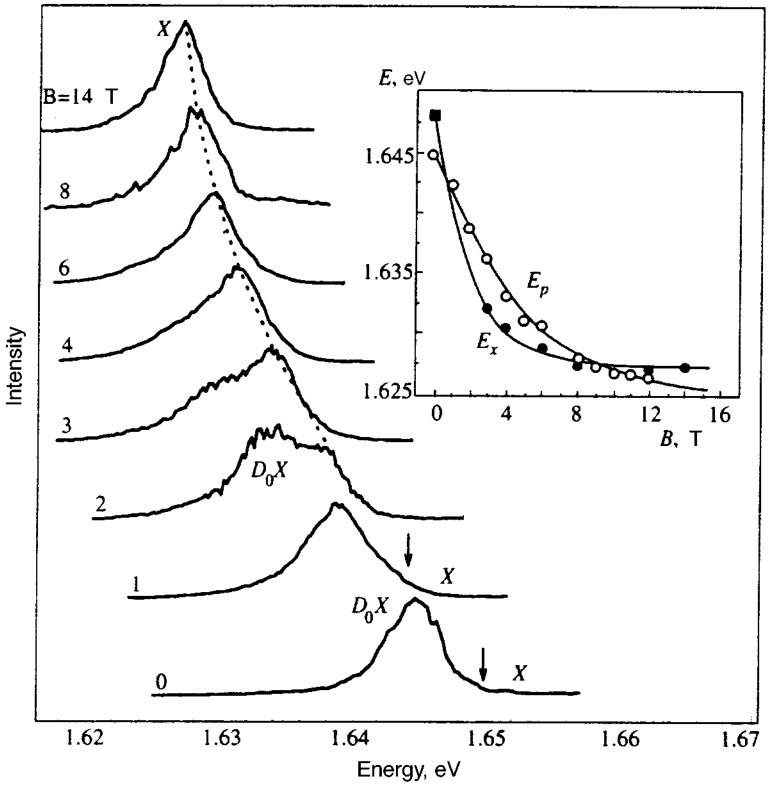

FIG. 1. Photoluminescence spectra from an isolated $Cd_{0.97}Mn_{0.03}Te/Cd_{0.75}Mg_{0.25}Te$ quantum well with a width of 10 nm under cw pumping by an Ar^+ laser at various magnetic fields ($T_{\mathrm{bath}}=4.2$ K). The insert shows the energy of the excitonic optical transition (E_x) and of the transition in the $e-h$ plasma (E_p) at $P=3\ \mu\mathrm{J/cm}^2$, $\tau_d=0$ ns, $T_{\mathrm{bath}}=4.2$ K as functions of magnetic field. E_p is defined as the spectral energy corresponding to 3/4 of the spectral peak intensity $I_{\max}$ on the low-energy edge of the $e-h$ plasma spectral line. The solid lines are fits based on Eqs. (1) and (2). The exciton position in zero magnetic field (black square) is derived from photoluminescence excitation spectra.

the sample magnetization in magnetic fields $B<15T$.[3] The factor 1/2 is the eigenvalue of the Pauli spin operator for electrons and holes.

The average Mn spin is determined by the population of the Mn Zeeman multiplet in applied magnetic field. In a homogeneous system under conditions of thermal equilibrium within the magnetic subsystem, $\langle S_z\rangle$ is described by a Brillouin function for spin 5/2 in external magnetic field B[1,13]:

$$\langle S_z\rangle=B_{5/2}\frac{\mu_B g_{\mathrm{Mn}}B}{k(T_{\mathrm{Mn}}+T_0)}, \tag{2}$$

where μ_B is the Bohr magneton, g_{Mn} is the Mn g-factor, k is the Boltzmann constant, T_{Mn} is the Mn spin temperature, and T_0 is the effective antiferromagnetic temperature due to a weak long-range interaction in Mn-subsystem. In the case of a low Mn concentration, the latter parameter is small, and $T_0\sim 1$ K for $x=0.03$.[13]

The photoluminescence line becomes narrow in high magnetic fields because of ionization of D_0X complexes due to the exchange field.[12] In the studied quantum well at a temperature of 4.2 K, the complexes are ionized in a magnetic field of 2–2.5 T. Thus, in magnetic field $B>4$ T, the X-line due to recombination of excitons weakly localized on fluctuations of the quantum well random potential dominate in the spectra.

At a low pump power densities, the line shift in magnetic field is in a good agreement with the Brillouin function corresponding to the effective temperature $T_{\mathrm{Mn}}\simeq T_{\mathrm{bath}}$, where $T_{\mathrm{bath}}=4.2$ K is the liquid helium bath temperature (with due account of the additional line shift due to ionization of exciton complex). In a saturating magnetic field, the X-line shift with respect to its position at zero magnetic field is 24 meV (here $\simeq 20$ meV is the shift with respect to the D_0X-line at zero field and $\simeq 4$ meV is the D_0X-complex binding energy).[12] The magnetic field of 14 T fully saturates the magnetic subsystem at a low Mn spin temperature ($T_{\mathrm{Mn}}\simeq 4.2$ K). In Fig. 1 the exciton transition energy changes little in magnetic fields higher than $\simeq 6$ T.

Equation (1) allows us to determine the Mn spin $\langle S_z\rangle$ directly from the line position in the photoluminescence spectrum, and Eq. (2) enables determination of T_{Mn} through $\langle S_z\rangle$ at a fixed external magnetic field. Thus, the spectral position of the exciton luminescence line at $B=14$ T for a low excitation level corresponds to $\langle S_z\rangle=5/2$.

3.2. Photoluminescence spectra under pulsed photoexcitation. Probe pulses as a means of determining the volume-averaged $\langle S_z\rangle$ as a function of time. Spin–lattice relaxation

Figure 2 shows a photoluminescence spectrum of the $Cd_{0.97}Mn_{0.03}Te$ quantum well recorded during the lifetime of the $e-h$ plasma at delay time $\tau_d=0$ (the time gate width is 0.3 ns), $T_{\mathrm{bath}}=4.2$ K, $B=14$ T, and $P=3\ \mu\mathrm{J/cm}^2$. For comparison, luminescence spectra under cw excitation for $B=0$ and 14 T are given in the graph. The spectrum contains a broad line due to luminescence of the $e-h$ plasma from the lowest size-quantized subband of electrons and holes ($n_z=1$).[10] In a magnetic field of 14 T, an additional feature due to quantization of both electron and hole states in strong magnetic fields is resolved. Allowed optical transitions la-

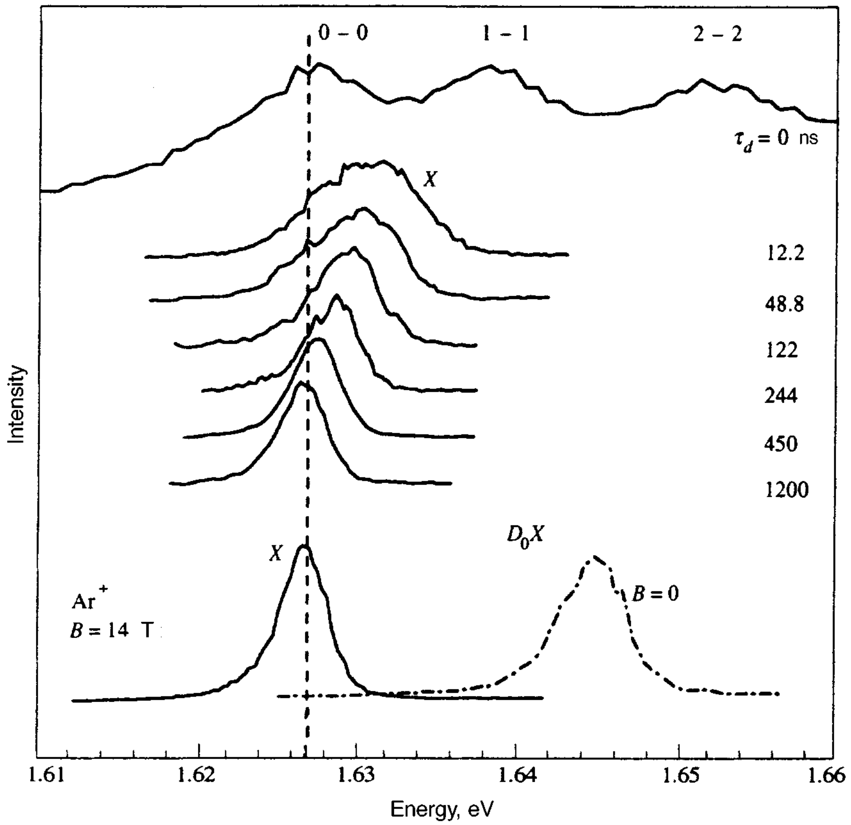

FIG. 2. Photoluminescence spectra of the $Cd_{0.97}Mn_{0.03}Te/Cd_{0.75}Mg_{0.25}Te$ quantum well with a width of 10 nm due to additional photoexcitation by probe pulses recorded after a powerful pump pulse ($P=3\ \mu J/cm^2$) at the initial moment ($T_{bath}=4.2$ K). The signal was detected with delays corresponding to peaks of the probe pulses with a time gate width of 0.3 ns. For comparison, the graph shows photoluminescence spectra under cw excitation at $B=0$ and $B=14$ T. The vertical dashed line shows the energy E^* of the excitonic X-line in a magnetic field $B=14$ T under cw excitation.

beled by $j-j$ ($j=0,1,\ldots$) between Landau levels of electrons and holes with equal quantum numbers ($j_e=j_h$) are clearly seen (Fig. 2).

As in the case of low pump power densities, the luminescence line of the $e-h$ plasma shifts to the low-energy side in magnetic fields of up to 14 T. The shift rate, however, is notably smaller in this case. The insert to Fig. 1 shows the optical transition energies for excitons (E_x) at low cw pump powers and for the $e-h$ plasma (E_p) at $P=3\ \mu J/cm^2$ and $\tau_d=0$ ns as functions of magnetic field for $B<10$ T and $T_{bath}=4.2$ K. In this range of magnetic fields, Landau levels are not resolved, and E_p is defined as the energy corresponding to a level of 3/4 of the maximum luminescence intensity I_{max} on the low-energy tail of the $e-h$ plasma line. By defining E_p in this manner, we approximately take into account the renormalization of the bandgap in the $e-h$ plasma.[10] Since the shape of the low-energy edge of the $e-h$ plasma line changes little in magnetic fields ranging to 10 T, the uncertainty in the E_p shift is $\simeq 1$ meV, even though the $e-h$ plasma line is very broad ($\simeq 40$ meV). The measured shift of E_p versus magnetic field is accounted for by spin temperature $T_{Mn}\simeq 10\pm 0.5$ K in accordance with Eqs. (1) and (2).

The Mn spin temperature is different from both T_{bath} and the electron temperature T_e of the $e-h$ plasma in the quantum well. T_e rapidly grows with the optical pump power.[10] This parameter is determined by analyzing shapes of luminescence spectra.[14] At the maximum pump power density $P=3\ \mu J/cm^2$ the temperature T_e rises to 200–300 K. Thus, the spin temperature is considerably lower than the electron temperature, $T_{Mn}\ll T_e$.

Comparison of plasma luminescence spectra corresponding to different pump power densities indicates that Mn spin heating is important only at $P\geqslant 0.3\ \mu J/cm^2$, with T_{Mn} (or the average Mn spin $\langle S_z\rangle$) being a function of the time delay following the pump pulse.

The variation in T_{Mn} during the luminescence decay time immediately after the pump pulse is discussed in Sec. 4. For testing the magnetic state of the quantum well at long delays after the excitation of carriers, we have used photoluminescence spectra due to additional photoexcitation by probe pulses (Fig. 2). The optical signal was detected at time delays corresponding to peaks of probe pulses with a time gate width of 0.3 ns. The power density in a single probe pulse is $P\simeq 3\ nJ/cm^2$, and the filling factor of Landau levels at peaks of probe pulses was at most $\nu\sim 0.01$.

The X-line due to excitons dominates in the photoluminescence spectra. Figure 2 shows that the X-line is blue shifted by ~ 4 meV with respect to its position under the low cw power density during the first probe pulse ($\tau_d=12.2$ ns). This shift reflects a decrease in the average spin $\langle S_z\rangle$ or an increase in the Mn spin temperature (see Eqs. (1) and (2)) due to a high-power pump pulse at $\tau_d=0$. We performed a dedicated experiment to check that the probe pulse energy was too small to heat the magnetic subsystem. As was noted above, no Mn heating was detected even at a pump power density two orders of magnitude higher, $P\simeq 0.3\ \mu J/cm^2$.

In recording photoluminescence spectra using probe pulses, we detected the signal at the moment when the pump pulse was on so that $\langle S_z\rangle$ derived from these spectra could be treated as a Mn spin averaged over the quantum well volume, $\langle S_z\rangle_V$.

One can see in Fig. 2 that as the delay increases, the

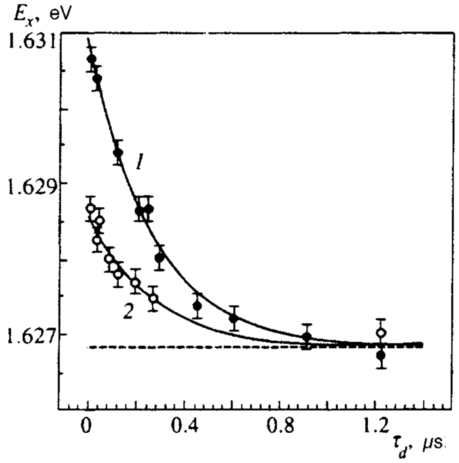

FIG. 3. Delay dependence of the spectral position E_x of the exciton recombination line under weak additional photoexcitation for two values of the pump power density, $P=3$ μJ/cm^2 (curve 1) and $P=2$ μJ/cm^2 (curve 2) in magnetic field $B=14$ T. The $Cd_{0.97}Mn_{0.03}Te/Cd_{0.75}Mg_{0.25}Te$ quantum well has width $L_z=10$ nm; $T_{bath}=4.2$ K. The horizontal dashed line shows the exciton line position E^* under cw excitation. The solid lines are exponential least-squares fits of Eq. (3).

X-line slowly shifts to lower energies, eventually reaching its equilibrium value E^*. This shift reflects the energy relaxation of $\langle S_z\rangle_V$ to its equilibrium value $\langle S_z\rangle=5/2$, or the relaxation of the spin temperature T_{Mn} averaged over the quantum well volume to the liquid helium bath temperature $T_{bath}=4.2$ K. Since the concentration of free carriers in these experiments is low, we attribute this behavior to the energy relaxation of the Mn spins due to interaction with phonons.

The exciton transition energy versus time, $E_x(t)$, measured using probe pulses at $B=14$ T and $T_{bath}=4.2$ K, is plotted in Fig. 3 for two pump power densities, $P=3$ and 2 μJ/cm^2. The horizontal dashed line shows the position of the X-line under cw pumping, $E_x=E^*=1.6268$ eV, when the Mn spin temperature equals the liquid helium bath temperature, $T_{Mn}=T_{bath}=4.2$ K. The solid lines are least-squares exponential fits based on the formula

$$\Delta E(t)=E_x(t)-E^*=\Delta E(0)\exp(-t/\tau_{Mn}). \tag{3}$$

The quantity $\Delta E(0)$ is a function of the pump energy density. It equals 4.2 and 1.8 meV for $P=3$ and 2 μJ/cm^2, respectively. Although the initial value of $\Delta E(0)$ characterizing the spin disorder was different, the relaxation time in both cases was the same, $\tau_{Mn}=270\pm10$ ns.

The equal values of τ_{Mn} at different $\Delta E(0)$ indicate that the Mn spin subsystem is cooled down owing to the spin–lattice relaxation, and the phonon subsystem temperature, at least after delays $\tau_d\geqslant 12$ ns, is close to the bath temperature. Otherwise the spin relaxation could not be described by an exponential because the spin–lattice relaxation time would be a function of the phonon temperature.[6] The derived spin relaxation time $\tau_{Mn}=270$ ns is in good agreement with investigations of the spin–lattice relaxation of bulk (CdMn)Te crystals in a magnetic field $B\simeq14$ T at liquid helium temperatures.[5]

In order to have an additional test of the statement that $\langle S_z\rangle_V$ relaxation in the range of long delays is due to interaction with phonons, we measured τ_{Mn} as a function of temperature and Mn concentration. Experiments were carried out in magnetic field $B=14$ T. As the bath temperature increased from 4.2 to 10 K, the spin relaxation time in the quantum well with $x=0.03$ decreased from 270 to 70 ns. Measurements of the spin relaxation time in similar CdTe/(CdMn)Te quantum wells with a higher Mn concentration $x=0.12$ indicate that the relaxation time τ_{Mn} decreases to 16 ± 7 ns at $T_{bath}=4.2$ K.[1)] These data are also in agreement with published data on the spin–lattice relaxation time as a function of T and x.[5,6]

One can also see in Fig. 2 that in the range $\tau_d=10-100$ ns the exciton X-line is notably broader than in spectra recorded under cw pumping. This broadening cannot be caused by the luminescence of the exciton–impurity complex since (see Fig. 1) the exciton is in the ground state by the time the magnetic field is 2–3 T, at an exciton energy of 1.635 eV, whereas the X-line is about 4 meV lower on the energy scale at $\tau_d=10-100$ ns. This line broadening decreases with increasing time delay, and at $\tau_d\sim100$ ns it is negligible.

We presume that X-line broadening at $\tau_d<100$ ns is related to spin inhomogeneities in the magnetic subsystem of a quantum well. The decrease in the X-line width reflects the emergence of equilibrium due to spin–spin interaction. The characteristic time in which equilibrium is established, estimated from measurements of the X-line half width, is $\tau_{ss}\sim70$ ns. Presently no published data are available on spin–spin relaxation times in semimagnetic semiconductors with $x=0.03$, which were studied in the present work. A fit to the previously reported[15,16] experimental curve of $\tau_{ss}(x)$ in the range of $0.05<x<0.1$ to the region of $x\simeq0.03$ yields $\tau_{ss}\simeq1$ ns, which is much less than our measured value. This difference can be accounted for in terms of the macroscopic character of spin inhomogeneities in the Mn subsystem in our experiments.

4. HEATING OF MAGNETIC SYSTEM DUE TO SPIN–SPIN INTERACTION WITH *e–h* PLASMA

4.1. Direct heating of Mn spins by photogenerated carriers. Experimental results

As was noted above, the spin temperature T_{Mn} of the magnetic subsysetem remains close to the liquid helium bath temperature at an optical pump power density up to $P=0.3$ μJ/cm^2. At higher P, the spin temperature T_{Mn} is a function of the time delay τ_d between the pump pulse and the moment when the luminescence signal is detected. This can be seen in Fig. 4, which shows photoluminescence spectra of the $Cd_{0.97}Mn_{0.03}Te$ quantum well recorded at various delays τ_d during the lifetime of the $e-h$ plasma at $T_{bath}=4.2$ K and $B=14$ T for two pump power densities, $P=3$ μJ/cm^2 and 0.3 μJ/cm^2. For comparison, the graph also shows photoluminescence spectra under cw excitation for $B=0$ and $B=14$ T.

Figure 4 shows that at $P=3$ μJ/cm^2 and $\tau_d\leqslant1.3$ ns, $e-h$ plasma luminescence dominates the spectrum. The maximum carrier concentration derived from the filling factor of Landau levels[10] is $n_{eh}=1.7\times10^{12}$ cm^{-2}. The carrier

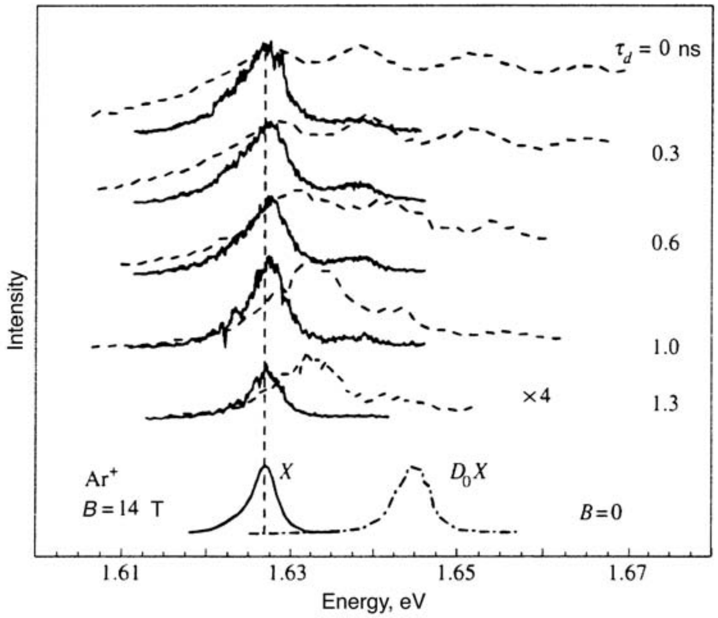

FIG. 4. Photoluminescence spectra of the $Cd_{0.97}Mn_{0.03}Te$/$Cd_{0.75}Mg_{0.25}Te$ quantum well with a width of 10 nm recorded during the $e-h$ plasma lifetime at various delays $\tau_d<1.3$ ns, $T_{\text{bath}}=4.2$ K, $B=14$ T, $P=3$ and 0.3 μJ/cm^2 (dashed and solid lines, respectively). For comparison, photoluminescence spectra due to cw excitation at $B=0$ and 14 T are given. The vertical dashed line marks the spectral position of the X-line in magnetic field $B=14$ T under cw excitation.

concentration decreases with increasing time delay owing to carrier recombination, and the luminescence line due to transitions between the lowest Landau levels (0–0) gradually transforms to the X-line of magnetoexcitons at $\tau_d\simeq 1.5$ ns. The energy of the 0–0 transitions in the $e-h$ plasma, like the exciton transition energy, is a function of the average Mn spin (see Eq. (1) and Refs. 10 and 17). Therefore, in testing the magnetic state of the system during the $e-h$ plasma lifetime, one can also measure the energy of 0–0 transitions, E_{0-0}.

Figure 4 indicates that in the range $\tau_d=0-1.3$ ns the energy E_{0-0} increases with τ_d. The maximum shift is about 6 meV. It occurs approximately at the time when excited Landau levels are emptied (filling factor $\nu\simeq 1$). There are two reasons for changes in the energy of the 0–0 transition with τ_d. First, this is a decrease in the attractive interparticle interaction in the plasma, ΔE_{e-h}, due to the decrease in the plasma density n_{eh}. Second, the exchange interaction energy ΔE_{exc} decreases owing to the increase in the Mn spin temperature. Previously[10] we argued that the renormalization of E_{0-0} due to the interparticle interaction in the $e-h$ plasma is small in a CdMnTe quantum well ($\Delta E_{e-h}<3$ meV for $n_{eh}\leqslant 2\times 10^{12}$ cm^{-2}). Consequently, the change in ΔE_{exc} due to the higher Mn spin temperature is at least 3 meV for $\tau_d=1-1.3$ ns.

A similar value of ΔE_{exc} is obtained by directly comparing photoluminescence spectra at various τ_d recorded at two values of the optical pump power density, $P=3$ and 0.3 μJ/cm^2, but at time delays differing by about 1 ns (Fig. 5). A delay of about 1 ns was selected to make the spectra correspond to equal carrier densities. In this case, the difference between the spectral positions of the luminescence lines

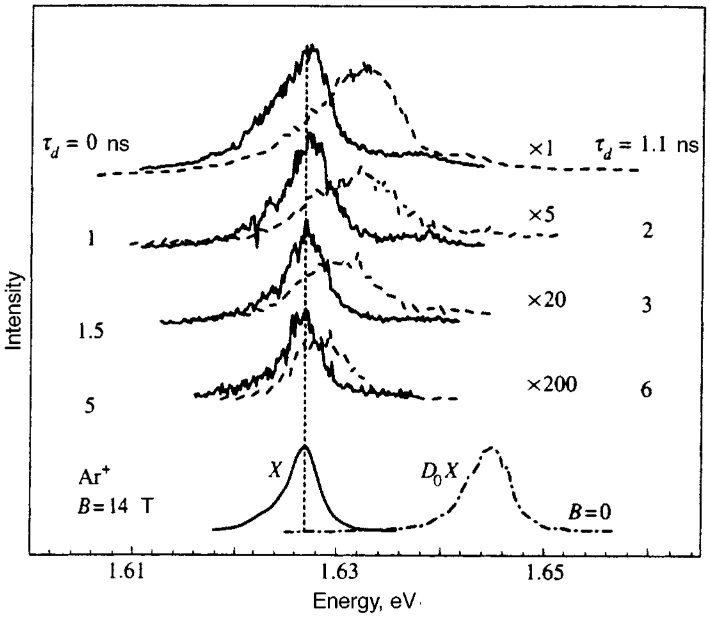

FIG. 5. Comparison between photoluminescence spectra of the $Cd_{0.97}Mn_{0.03}Te/Cd_{0.75}Mg_{0.25}Te$ quantum well 10-nm wide at $B=14$ T, $T_{\text{bath}}=4.2$ K, and $P=3$ and 0.3 μJ/cm^2 (dashed and solid lines, respectively). The delays are selected so that the carrier densities in both cases are equal.

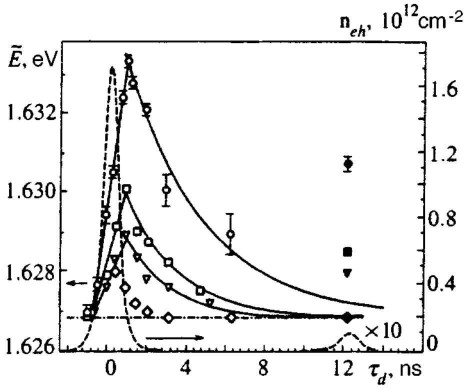

FIG. 6. Transition energies E_{0-0} and E_x ($\tilde{E}$, see the discussion in Sec. 5) in the photoluminescence spectra of the $Cd_{0.97}Mn_{0.03}Te/Cd_{0.75}Mg_{0.25}Te$ quantum well ($L_z = 10$ nm, $T_{bath} = 4.2$ K) in magnetic field $B = 14$ T at $P = 3$, 2, 1, and 0.3 μJ/cm^2 (plotted by circles, squares, triangles, and rhombi, respectively) versus time in the range $\tau_d < 16$ ns. The spectra were recorded either without the additional optical probe pulses during photoluminescence decay (open symbols) or under additional excitation by the first probe pulse at its peak (full symbols). Within the time interval $\tau_d < 1.5$ ns the corrections for the interparticle interaction in the $e-h$ plasma were taken into account. The dashed line plots the concentration n_{eh} of two-dimensional electrons and holes derived from the number of occupied Landau levels at $P = 3$ μJ/cm^2. The dash-dotted horizontal line shows the position of the exciton line under cw excitation.

is determined solely by the difference between the Mn spin temperatures, since the renormalization due to the interparticle interaction ΔE_{e-h} in these cases is the same. Figure 5 demonstrates that the difference between the luminescence line positions reaches 5 meV at $\tau_d \simeq 1$ ns. By substituting $\Delta E_{exc} \simeq 5$ meV into Eqs. (1) and (2) and using the value $\Delta E = 24$ meV for the X-line shift in the saturating magnetic field, we obtain $T_{Mn} \simeq 25$ K for $P = 3$ μJ/cm^2 at $\tau_d \sim 1$ ns.

In Fig. 6 we plot the transition energy E_{0-0} between the lowest Landau levels for a magnetic field $B = 14$ T, $T_{bath} = 4.2$ K, and $P = 3$, 2, 1, and 0.3 μJ/cm^2. At $\tau_d < 1$ ns ($\nu > 1$), we have corrected for interparticle interaction in the $e-h$ plasma. At $\tau_d > 1$ ns ($\nu < 1$), these corrections are so small as to be negligible. Figure 6 also shows a fit to the deviation $E_{0-0} - E^*$ during heating of Mn spins at pump power density $P \geqslant 1$ μJ/cm^2. When heating the spin system, $\tau_d \leqslant 1.5$ ns, and the deviation from equilibrium is linear in time to within the experimental errors, $E_{0-0} - E^* \propto vt$. The heating rate v decreases with pump power, and is 3.7, 1.9, and 1.1 meV/ns at $P = 3$,2, and 1 μJ/cm^2. Heating typically ends simultaneously with the transition from the plasma to exciton regime of the charge carriers.

There are two feasible mechanisms for heating the Mn subsystem on a subnanosecond time scale. The first is spin–spin spd-interaction between charge carriers and Mn; the second is, the electron–phonon interaction, provided that the optical pump pulse heats the lattice to a temperature of order 100 K, at which the spin–lattice relaxation time drops to several nanoseconds.^{6}

In order to estimate the effect of lattice heating on the Mn spin temperature, we measured a sample that contained, in addition to a quantum well with a width $L_z = 10$ nm, a partially ‘‘etched off’’ quantum well with $L_z = 6$ nm. Figure 7 shows photoluminescence spectra of such a sample at $T_{bath} = 4.2$ K, $B = 14$ T under conditions of both weak cw pumping by an Ar$^+$ laser and a high power pulsed pump ($P = 2$ μJ/cm^2). Under cw pumping, the X-line from the 6-nm quantum well is a factor of about 50 weaker than that

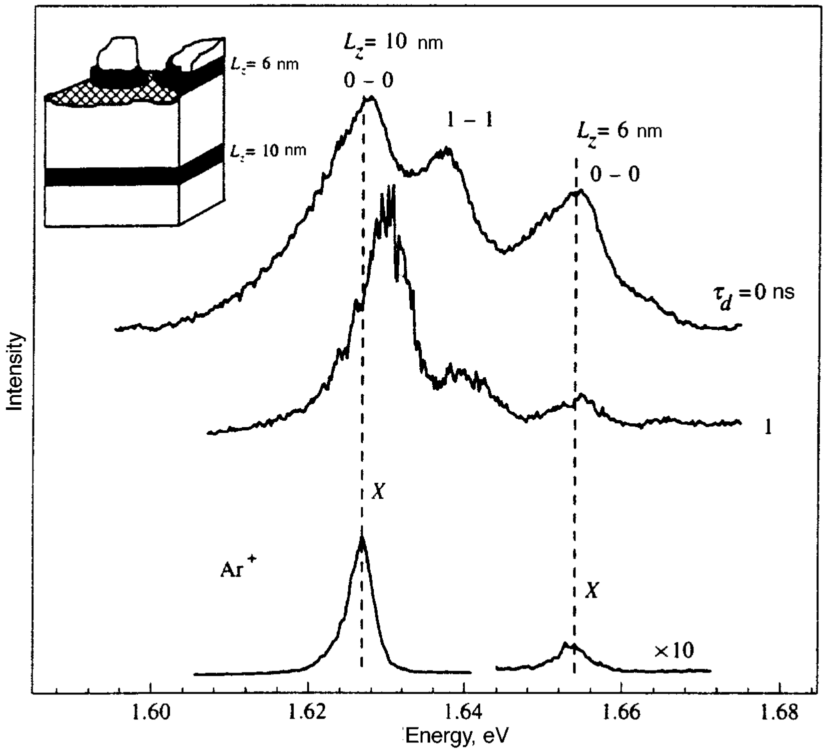

FIG. 7. Photoluminescence spectra of the $Cd_{0.97}Mn_{0.03}Te/Cd_{0.75}Mg_{0.25}Te$ sample containing two quantum wells with widths of 10 nm and 6 nm under pulsed excitation ($P = 2$ μJ/cm^2 and $\tau_d = 0$ and 1 ns), and under low-power cw pumping (Ar$^+$ laser). $B = 14$ T, $T_{bath} = 4.2$ K.

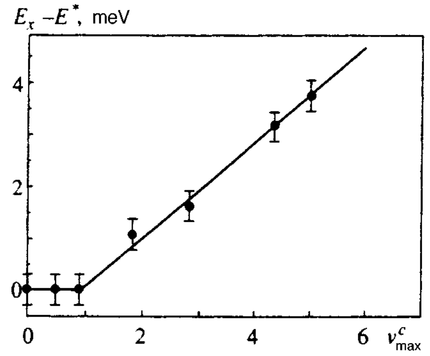

FIG. 8. Amplitude of the energy (E_x) deviation from its equilibrium value E^* measured at the peak of the first probe pulse as a function of the maximum filling factor of the conduction band, $\nu^c_{\max}=\nu(\tau_d=0)$.

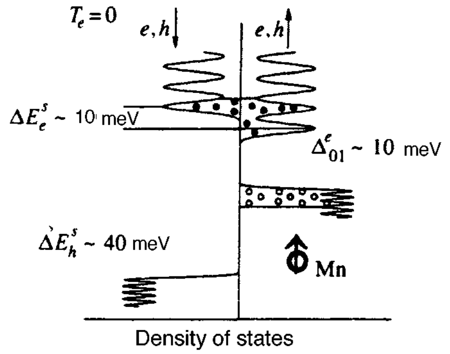

FIG. 9. Diagram of spin splitting of the density of states near the band gap edge in a saturating magnetic field. The situation when the width of the $e-h$ plasma photoluminescence spectra is about 30 meV is illustrated.

from the 10-nm well, and is shifted by about 30 meV to the high-energy side owing to the size quantization.

Under pulsed pumping, several Landau levels are filled in the 10-nm quantum well at the moment $\tau_d=0$, whereas in the 6-nm quantum well only one level is occupied. The higher filling factor in the first well is caused in this case by a more efficient collection of photoexcited carriers from the bulk to the wider quantum well, and additional surface recombination in the partly ''etched-off'' quantum well (see the insert to Fig. 7).

The separation between the first and second well is only 100 nm, so a large difference between the lattice temperatures in these wells is out of the question.[18] Nonetheless, heating of the Mn spins in the 10-nm quantum well is clearly seen in Fig. 7, whereas no heating is detected in the 6-nm well. Consequently, the difference between the Mn spin temperatures in these wells is associated with the difference between the densities of photoexcited carriers.

Thus, we can conclude that the short time of Mn spin depolarization in a photoexcited quantum well is related to direct interaction between the photoexcited carriers and the magnetic subsystem.

4.2. Spin–spin *sd*-exchange between Mn and *e–h* plasma. Discussion of experimental results

In the previous section we demonstrated that the heating of the magnetic subsystem in the quantum well under optical pumping develops during a very short time interval of about 1 ns, and is terminated when only one Landau level is occupied, which roughly corresponds to the transition from the plasma mode to excitonic mode of the nonequilibrium carriers existence.

Figure 8 shows the amplitude of the energy (E_x) deviation from its equilibrium value measured at the maximum of the first probe pulse, $E_x(12.2\text{ ns})-E^*$, as a function of the maximum filling factor of the Landau levels in the conduction band, $\nu^c_{\max}=\nu(0)$. The parameter $E_x(12.2\text{ ns})-E^*$ is approximately linear with $\nu^c_{\max}$, and this curve has a cut-off at $E_x=E^*$ near a filling factor $\nu^c_{\max}\simeq 1$. Thus, we conclude that heating of the magnetic subsystem as a function of the filling factor of electron states has a threshold.

The presence of the cut-off point in Fig. 8 indicates that the Mn spin heating does not depend on the total concentration of carriers in the quantum well, but is instead a function of the filling of the upper spin subband of electrons. Figure 9 shows a diagram of spin splitting of the density of states near the band gap in a strong magnetic field. Under equilibrium conditions at 4 K, Mn spins are fully polarized and aligned with the magnetic field (states marked by the up-arrow in Fig. 9). The band gap edges corresponding to allowed optical transitions in σ^+ and σ^- circular polarizations are formed by electrons and holes with parallel spins, $e\uparrow$, $h\uparrow$ and $e\downarrow$, $h\downarrow$, respectively. The total spin splitting of the band states, $\Delta E^s \simeq 50$ meV, is shared between electron and hole states proportionally to the respective exchange integrals, namely $\Delta E^s_e \simeq 10$ meV in the conduction band and $\Delta E^s_h \simeq 40$ meV in the valence band.[13] The electron Fermi energy E^F_e, on the contrary, is higher than the hole Fermi energy E^F_h, owing to the smaller electron effective mass, $m_h/m_e > 2$.[19] One can see in Fig. 9 that even at the maximum accessible Fermi energies ($\Delta E^s_e + \Delta E^s_h \sim 60$ meV, which corresponds to $n_e \sim 2\times 10^{12}$ cm^{-2}), the holes are totally spin polarized. The onset of filling of the electron spin-split level $e\downarrow$, on the contrary, occurs at $E^F_e \simeq 10$ meV. The states $e\downarrow$ are optically inactive because of the absence of $h\downarrow$ holes, and cannot be seen in photoluminescence spectra. In a magnetic field of B=14 T the spin splitting in the conduction band is approximately equal to the cyclotron splitting between the electron Landau levels, $\Delta E^s_e \sim \Delta^e_{01}$, which provides a convenient qualitative criterion for the onset of the electron spin-split state $e\downarrow$ filling, namely, the filling factor $\nu=\nu^c$ is approximately equal to unity.

The absence of Mn spin heating at $\nu<1$ (Fig. 8) and the termination of heating at $\nu\sim 1$ in the plots of $E_{0-0}(t)$ in Fig. 6 indicate that Mn spin heating is determined by the filling of the upper electron spin subband and is due to the spin–spin sd-interaction.

The most likely mechanism of the energy transfer is exchange scattering of electrons with a transition from the $e\uparrow$ subband to the $e\downarrow$ subband accompanied by a change in the Mn spin, i.e., so-called spin-flip scattering.

The rate of paramagnetic impurity spin relaxation, $1/\tau_e$, in bulk semiconductors due to the contact spin–spin interaction with free Fermi carriers is calculated by Abraham[20] for spin 1/2 under conditions when in which Landau quantization can be neglected. This calculation considers resonant spin-flip scattering of band electrons. The total energy of electrons and Mn spins is conserved, and the change in the electron energy is $\mu_B g_{\mathrm{Mn}} B$. The electron quasimomentum is not conserved because of the short range of the exchange potential action. Given a sufficiently high concentration of electrons with a Fermi distribution function, the rate of spin-flip scattering is independent of the conduction band spin splitting ΔE_e^s.

A calculation of the spin-flip scattering rate for two-dimensional carriers and impurities with spin 5/2 in a symmetrical quantum well (see Appendix) shows that in the two-dimensional case, the Mn spin relaxation rate is proportional to the coupling constant squared, the density of states squared, and the electron temperature, which is in qualitative agreement with the result for bulk semiconductors.[20] The numerical coefficient, however, is a factor of twenty higher owing to a change in the spin matrix element and size quantization of the electron motion along the z-axis. Moreover, in a quantum well the relaxation rate depends on the distance between the Mn atoms and heterojunction, because the electron density is modulated in the direction perpendicular to the quantum well plane.

In the structure studied here, the characteristic rate of Mn spin relaxation is less than a nanosecond (see Refs. 15 and 16, and the discussion in Sec. 3.2). This rate determines the characteristic time in which a local equilibrium is established in the magnetic subsystem. Therefore we use in our estimates the spin relaxation rate averaged over the quantum well profile (Eq. (A12)) and ignore oscillations in the electron density of states D due to Landau quantization. The density of states D is estimated to be $D = n_e / E_e^F = 5\times 10^{10}\ \mathrm{meV}^{-1}\,\mathrm{cm}^{-2}$ ($n_e = n_{eh} = 1.7\times 10^{12}$ cm^{-2}, $E_e^F = 40$ meV). By substituting into Eq. (A12) D, $\alpha = 1.5\times 10^{-23}\ \mathrm{eV}\cdot\mathrm{cm}^3$ ($N_0\alpha = 0.22$ eV),[13] and $T_e = 200$ K, we obtain the characteristic spin relaxation time $\tau_e \simeq 4$ ns for the maximum pump power densities.

As stated in Sec. 3.2, the temperature of photoexcited electrons at maximum pump power densities is very high, $T_e \simeq 200$ K. Therefore the process of Mn spin relaxation through interaction with photoexcited electrons constitutes spin heating to the temperature T_e of thermal equilibrium between Mn spins and electrons, with characteristic time τ_e. The Mn spin heating is described by the equation

$$\Delta E(t) = E(t) - E^* = \Delta E_{\max}[1 - \exp(-t/\tau_e)], \tag{4}$$

where $\Delta E_{\max}$ corresponds to thermal equilibrium between electrons and Mn spins. Taking $\Delta E_{\max} \simeq 20$ meV ($T_{\mathrm{Mn}} \approx T_e \approx 200$ K) and expanding Eq. (4) in terms of t/τ_e, we have in the range of small t/τ_e the Mn spin heating rate $v \approx \Delta E_{\max}/\tau_e \simeq 5$ meV/ns, which is in fair agreement with our measurements ($v = 3.7$ meV/ns).

Note that the parameters of our experimental conditions are beyond the range of applicability of Eq. (A12). In a magnetic field of 14 T, separate Landau levels are resolved in the photoluminescence spectra, which indicates that the energy distribution of electrons has notable oscillations. Since the density of states is included in the final formula in the form of a convolution (Eq. (A9)), the quantization of the density of states should lead to an increase in $1/\tau_e$ in the case when the spin splitting is a multiple of the cyclotron energy in the conduction band ($\Delta E_e^s = 0,\ \Delta_{01}^e, \dots, \Delta_{0i}^e$).

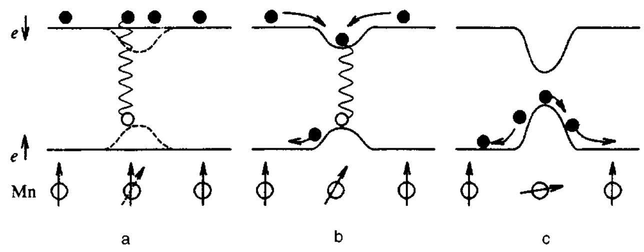

FIG. 10. Diagram illustrating spatial separation in the Mn subsystem caused by the spin-flip scattering of conductance electrons. The situation of the totally polarized Mn subsystem at the initial moment is shown. The diagram shows two electron subbands ($e\uparrow$ and $e\downarrow$) split by the sd-exchange. Electrons are shown only in the upper $e\downarrow$ subband. At the initial moment (a) both the magnetic and electronic subsystems are homogeneous. Then (b) a fraction of Mn ions change their spin states, and the electron spin splitting in heated regions is reduced. The probability of spin-flip relaxation increases in these regions. Later the spin-flip relaxation occurs largely in the heated regions. When (c) all electrons are in the lower $e\uparrow$ subband, diffusion of electrons from cold regions is driven by the gradient of the exchange potential due to the Mn magnetization.

The relaxation rate as a function of the homogeneous width Γ of Landau levels should be approximately described by the relation $1/\tau_e \propto \int D^2 dE \propto 1/\Gamma$. On the other hand, the matrix element of the spin-flip scattering should drop with the increasing difference between Landau level numbers and, accordingly, with ΔE_e^s. Thus, the spin relaxation rate $1/\tau_e$ in a quantizing magnetic field should be a decreasing oscillating function of ΔE_{e^s} with an absolute maximum at $\Delta E_e^s = 0$. The oscillation amplitude should increase with decreasing Γ, resulting in a total suppression of resonant spin-flip scattering at certain ratios among Γ, Δ_{ij}^e, and ΔE_e^s. Under these conditions, the nonresonant three-particle processes of electron spin-flip scattering with participation of phonons will be dominant.

Agreement between measurements of the heating rate v and theoretical calculations with Eq. (A12) is probably due to a notable homogeneous broadening of lines in $e-h$ plasma and the mutual cancellation of corrections to the density of states and matrix element of the spin-flip scattering.2)

As we just noted, the Mn spin relaxation rate in a quantizing magnetic field of $B = 14$ T should drop with increasing ΔE_e^s. The increase in the spin temperature leads to a lower splitting in the conduction band, which in turn leads to positive feedback and results in spatial inhomogeneities in the magnetic subsystem of the quantum well. Figure 10 illustrates this process. At the initial moment, both the magnetic and electronic subsystems are homogeneous (Fig. 10a), and the probability W of spin-flip scattering is independent of position. As a result, a fraction of all Mn ions change their spin states (Fig. 10b). At a fixed magnetic field, W is larger in heated regions owing to the smaller electron spin splitting ΔE_e^s. Thus, it is natural to expect that the Mn spin heating should be inhomogeneous and lead to the separation of the magnetic subsystem in the quantum well into hot and cold regions.

Large inhomogeneities in the spin density should result in additional inhomogeneous broadening of the X-line. Figure 2 demonstrates that this broadening really occurs in the range of time delays $\tau_d = 10-100$ ns. On the other hand, inhomogeneities in the magnetic subsystem should give rise to new features in the relaxation of photoexcited excitons. These features will be discussed in detail in the next section.

5. SPIN DOMAINS IN THE MAGNETIC SUBSYSTEM. ENERGY RELAXATION OF EXCITONS IN A RANDOM EXCHANGE POTENTIAL

The behavior of the excitonic subsystem during the photoluminescence decay after a pump pulse strongly depends on the laser pulse energy. This is clearly seen in Fig. 5, which shows spectra of decaying photoluminescence due to pulses with two different pump power densities ($P = 3$ and $0.3\ \mu$J/cm^2) recorded at moments with a delay of about 1 ns between them. As we noted above, this delay was selected so that the carrier concentrations in both cases were equal.

In the time delay range $\tau_d = 2-10$ ns, the shift of the X-line to the low-energy side is about 1 meV in the case of the lower pump power density $P = 0.3\ \mu$J/cm^2. Hence, when T_{Mn} is close to the liquid helium bath temperature, the loss in the exciton energy due to the relaxation in the random potential is small ($\leqslant 1$ meV), whereas in the case of the higher pump power density $P = 3\ \mu$J/cm^2, when T_{Mn} reaches 25 K, the X-line rapidly shifts to the low-energy side by about 4 meV. This shift indicates a fast (within several nanoseconds) relaxation of Mn spins, which is in disagreement with the spin lattice relaxation time of 270 ns measured using probe pulses (Sec. 3). This shift also cannot be attributed to formation of a magnetic polaron, since the Mn concentration in the sample is too low.$^{3)}$

In order to describe the evolution of features in the photoluminescence spectrum throughout the studied time interval, it is convenient to introduce a band gap edge energy $\tilde{E}$. We define $\tilde{E} = E_{0-0}$ during the lifetime of the $e-h$ plasma and $\tilde{E} = E_x$ when the carriers are bound to excitons. Since $E_{0-0} = E_x$ when the Landau level filling factor is $\nu \leqslant 1$ and $E_{0-0} \leqslant E_x$ for $\nu > 1$,[10,17] the Mn spin $\langle S_z \rangle$ can be directly derived from $\tilde{E}$ using Eqs. (1) and (2) throughout the studied range of carrier densities n_{eh} with due account of the renormalization of E_{0-0} caused by interparticle interaction in the $e-h$ plasma.

The energy $\tilde{E}$ is plotted in Fig. 6 as a function of time delay for $P = 3$, 2, 1, and $0.3\ \mu$J/cm^2 at a magnetic field $B = 14$ T and $T_{\mathrm{bath}} = 4.2$ K. We first discuss the case of the maximum pump power density $P = 3\ \mu$J/cm^2. The dashed line in Fig. 6 shows the concentration n_{eh} of two-dimensional electrons and holes for $P = 3\ \mu$J/cm^2. At higher power densities n_{eh} was derived from the number of occupied Landau levels.[14] In the range $\nu < 1$ the density of carriers was assumed to be proportional to the intensity of the X-line in the luminescence spectra. The lifetime of nonequilibrium carriers in the studied quantum well is within the time resolution of the light detection device, which equals 300 ps. Nonetheless, the FWHM of the $n_{eh}(t)$ curve is about 1 ns. Hence follows a conclusion that the accumulation of carriers generated in the barriers in the quantum well is important during a fairly long time of about 1 ns.

In the time delay range 2–10 ns, when the $e-h$ system is composed of excitons, the Mn spin system relaxes to the equilibrium state. The carrier concentration drops during this process by several orders of magnitude. Exponential fits describing the cooling of Mn spins are shown in Fig. 6 by solid lines. These exponentials are in satisfactory agreement with experimental points. The characteristic time τ_{loc} of spin relaxation during the photoluminescence decay proved to be 4 ± 1 ns. As will be shown below, this time does not reflect a real, fast process of Mn spin relaxation, but is controlled by specific features of energy losses by excitons and can be described in terms of the exciton diffusion model discussed below.

Measurements in the range $\tau_d < 10$ ns were performed during the photoluminescence decay without an additional photoexcitation. In this case, a local Mn spin $\langle S_z \rangle_x$ averaged over the exciton volume is measured. The term ''local spin'' means in this case that an exciton can migrate across the sample for a long time, lose its energy in both crystal and random exchange potentials, form a magnetic polaron, etc. Therefore the measured value of $\langle S_z \rangle_x$, generally speaking, is different from the average over the quantum well volume $\langle S_z \rangle_V$.

Figure 6 also shows energies of excitonic transitions measured using additional photoexcitation at the maximum of the first probe pulse ($\tau_d = 12.2$ ns). This graph demonstrates that at shorter delays ($\tau_d = 3-10$ ns) the value of $\langle S_z \rangle_x$ is larger than $\langle S_z \rangle_V$ measured after longer delays time ($\tau_d = 12.2$ ns). It is obvious that such behavior would be impossible in a homogeneous spin system. The difference between measurements of the same parameter $\langle S_z \rangle$ by different methods is a clear indication that the magnetic subsystem of the quantum well contains microscopic spin inhomogeneities, i.e., spin domains.

When the pump power density is reduced to $P = 1\ \mu$J/cm^2, the behavior of $\tilde{E}(t)$ is essentially the same. It can be described in terms of heating and cooling to the bath temperature during the photoluminescence decay ($\tau_d < 6$ ns). The maximum in the spin temperature lags the maximum in the electron and hole density. The Mn spin relaxation time for $P = 2$ and $1\ \mu$J/cm^2 and $\tau_d = 1-6$ ns is adequately described by an exponential function with typical times $\tau_{\mathrm{loc}} = 2.7$ and 2.3 ns. At the peak of the probe pulse ($\tau_d = 12.2$ ns) the X-line has a higher energy than at $\tau_d = 4-6$ ns.

At lower pump power densities, the behavior of $\tilde{E}(t)$ is qualitatively different. Starting at $P = 0.3\ \mu$J/cm^2, it is identical to that of $n_{eh}(t)$, so that the maximum energy shift occurs in the region $\tau_d \simeq 0$. If the pump power density is reduced further, the $e-h$ system devolves into the exciton regime, the shape of the $\tilde{E}(t)$ curve does not change, and the maximum of the difference $\tilde{E} - E^*$ tends to zero. In this regime, heating of the Mn spin subsystem, on average, is curtailed. Starting at $P = 0.3\ \mu$J/cm^2, the deviation of E_x from its equilibrium value measured at the first probe pulse is zero. An important point is that a laser pulse of a high

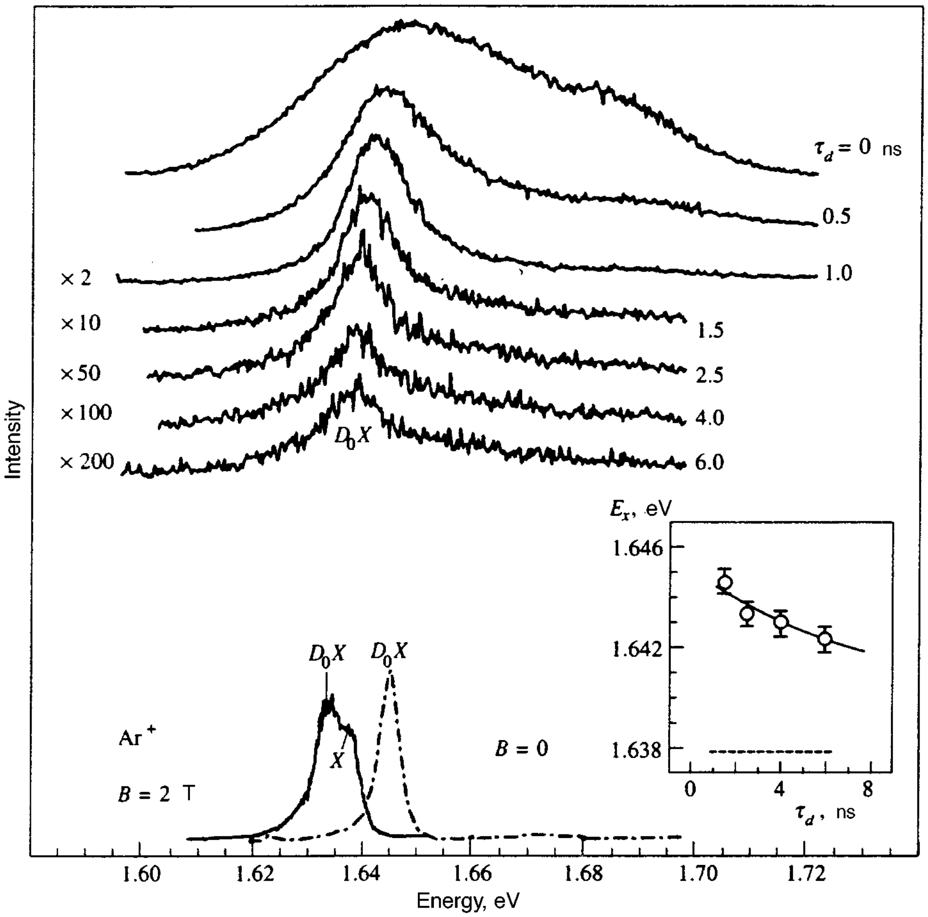

FIG. 11. Photoluminescence spectra of the $Cd_{0.97}Mn_{0.03}Te/Cd_{0.75}Mg_{0.25}Te$ quantum well ($L_z = 10$ nm) recorded at different moments τ_d at $T_{\text{bath}} = 4.2$ K, $P = 3$ μJ/cm^2, and $B = 2$ T). The graph also includes spectra recorded at a low density of nonequilibrium carriers for $B = 0$ and 2 T. The insert shows the energy E_x as a function of time at a magnetic field of 2 T, which was derived from the position of the D_0X line at delays $\tau_d = 1-6$ ns. The exciton binding energy at a neutral donor ($\simeq 4$ meV) was derived from photoluminescence spectra under cw excitation at the magnetic field of 2 T. The dashed horizontal line in the insert shows the X-line position under cw excitation.

enough intensity to fully populate the first Landau level does not heat the Mn spin system. Therefore the deviation $\widetilde{E}(0) - E^*$, about 1 meV at zero delay and $P = 0.3$ μJ/cm^2, relates to the population of localized states of the lowest Landau level and the effective shift of the photoluminescence line peak to the region of maximal density of states.

We ascribe the short Mn spin relaxation time (several nanoseconds) to the diffusion of excitons in a highly nonuniform exchange potential. The diffusion process is schematically shown in Fig. 10c. The decrease in τ_{loc} with decreasing pump power density indicates an increase in the exchange potential gradient, i.e., a relative decrease in the spatial dimensions of spin domains dimension as compared to their amplitude (the difference between $\langle S_z \rangle$ inside and outside the domain). Hence follows the inverse statement. In the process of the domain evolution, the amplitude of a spin domain first increases, and then the domain grows.

In order to detect a suppression of the exciton diffusion, we performed an experiment in small magnetic fields, when excitons are bound on impurities and their diffusion is impeded or even inhibited completely. Moreover, in small magnetic fields, Mn spin heating should be more uniform (see the discussion in Sec. 4.2), which should reduce the amplitude of spin domains and also slow exciton diffusion.

Figure 11 shows photoluminescence spectra of a $Cd_{0.97}Mn_{0.03}Te$ quantum well recorded at different moments of time τ_d at $T_{\text{bath}} = 4.2$ K, $P = 3$ μJ/cm^2, and magnetic field $B = 2$ T. This plot also gives spectra recorded at a low carrier density under cw pumping at $B = 0$ and 2 T. These photoluminescence spectra evolve in a manner similar to that of the spectra of Figs. 4 and 5. In the magnetic field of 2 T in the time range $\tau_d = 0-1.3_{ns}$, the spectra are mainly due to the $e-h$ plasma luminescence, as in the magnetic field of 14 T. When the delay increases to 1.5–2 ns and the carrier density drops, the luminescence line gradually transforms to the line of exciton recombination bound to neutral donors (see the discussion in Sec. 3.1 and Ref. 12). The X-line of exciton recombination should have an energy about 4 meV higher. The insert to Fig. 11 gives the energy E_x in a magnetic field of 2 T versus time. The energy E_x was derived from the transition energy of the D_0X complex, and the binding energy of the D_0X complex ($\simeq 4$ meV) was directly determined from the separation between the X and D_0X lines recorded under cw excitation in a field of 2 T.[12] In the time delay range $\tau_d = 1.5-6$ ns, E_x weakly depends on time (compare to Fig. 6). The spin relaxation time τ_{loc} is approximately 15 ns. The increase in τ_{loc} from 4 to 15 ns as the magnetic field drops from 14 to 2 T reflects the curtailment of exciton diffusion in weak magnetic fields and is associated, on the one hand, with a decrease in the amplitude of spin domains, while on the other, it reflects the localization of excitons at neutral donors. In order to separate these two contributions to the exciton diffusion, additional experiments are needed for an independent measurement of the spin domain amplitude as a function of magnetic field (see footnote 2).

6. CONCLUSIONS

It follows from our measurements that the main contribution to Mn spin heating in a quantum well doped with magnetic impurities under powerful pulsed excitation comes from free carriers, which generate spin domains in strong magnetic field. To conclude this paper, we give a schematic description of the sequence of events in the quantum well.

In the first stage ($\tau_d = 0-1$ ns) the quantum well contains dense two-dimensional $e-h$ plasma (Fig. 10a). In this stage electrons from the spin-split subband heat the Mn system through spin-flip scattering. This heating is spatially nonuniform (Fig. 10a,b). This effect is due to positive feedback in the form of a higher spin-flip scattering intensity caused by lower spin splitting in the conduction band. In the first stage the magnetic subsystem of the quantum well is divided into regions with higher and lower Mn spin temperatures, i.e., magnetic domains are formed. The presence of the spin density inhomogeneities is indicated by the notable broadening of the X-line in measurements with probe pulses with time delays up to $\tau_d \simeq 100$ ns.

In the second stage ($\tau_d \geqslant 1$ ns) the existence of carriers in the form of $e-h$ plasma has terminated, carriers are present in the quantum well in the form of excitons, and their density gradually drops as a result of recombination. This stage is characterized by the diffusion of excitons driven by the exchange field gradient from hotter to colder regions (Fig. 10b). The exciton diffusion manifests itself through the difference between the local spins $\langle S_z \rangle_x$ measured during the photoluminescence decay and the Mn spin averaged over the quantum well volume, $\langle S_z \rangle_V$, obtained using probe pulses (Figs. 2 and 3). We have observed in the experiments how $\langle S_z \rangle_x$ deviates from $\langle S_z \rangle_V$ to higher values corresponding to lower spin temperatures and a more ordered state of Mn spins in the range $\tau_d = 3-10$ ns. The fact that the local Mn spin deviates from the average value is direct evidence in favor of microscopic spin inhomogeneities in the magnetic subsystems in the quantum well.

During the second stage, local equilibrium in the magnetic subsystem is established through the spin–spin interaction, and Mn spins concurrently relax to the equilibrium state through interaction with acoustic phonons.

The authors are indebted to V. B. Timofeev and S. I. Gubarev for fruitful discussions. This work was supported by the Volkswagen Fund (Germany), Russian Nanostructures Program, and Russian Fund for Fundamental Research (grant No. 97-02-17697). One of the authors (D. R. Ya.) would like to express gratitude to Deutsche Forschungsgemeinschaft (SFB 410) for financial support.

APPENDIX

Mn spin relaxation via Fermi-distributed quasi-two-dimensional free electrons

The spin relaxation rate due to interaction with Fermi-distributed three-dimensional electrons was calculated by Abraham[20] for spin 1/2 and low magnetic fields. We now perform similar calculations for two-dimensional carriers. Suppose that there is thermal equilibrium between spin subbands. We assume that the magnetic field is sufficiently low so that the Landau quantization of electron orbital motion could be ignored.

Let the magnetic field be aligned with the z-axis, perpendicular to the quantum well plane. The well occupies the space between $z=0$ and L_z, and a Mn atom is at distance a from the heterojunction. Suppose for simplicity that the walls of the quantum well are infinitely high. A finite height of the walls should lead to inessential corrections that can be ignored in our case.

The Hamiltonian of the sd-exchange between spins of band electrons and Mn spin has the form[1]

$$H = \alpha \vec{\sigma} \vec{S} \delta(\mathbf{r}) = \alpha \delta(\mathbf{r}) \left[\sigma_z S_z + \frac{1}{2} (\sigma_+ S_- + \sigma_- S_+) \right]. \tag{A1}$$

Here α is the exchange integral for conduction electrons, and spin operators $\vec{S}$ and $\vec{\sigma}$ act on the Mn and free electron spins, respectively. The notation in this equation is identical to that in Eqs. (1) and (2).

The spin wave function of Mn varies as a result of an electron transition from one spin subband to another simultaneously with a change in the spin projection of the magnetic impurity by a unit. The elementary probability of such a transition is given by the expression[23]

$$w = \frac{2\pi}{\hbar} |\langle i|H|j\rangle|^2 \delta(E_i - E_j). \tag{A2}$$

Consider a situation when the magnetic field is sufficiently strong so that Mn has in the initial state the spin projection on the magnetic field direction $S_z = 5/2$. Hamiltonian (A.1) is responsible for transitions from state 5/2 to state 3/2. We consider electrons from the lowest size-quantized subband. Then the wave functions of the initial and final states have the form

$$|i\rangle = \sqrt{2}\, \frac{\sin(\pi a/L_z)}{\sqrt{L_z}} |\mathbf{k}\rangle \left| \frac{5}{2}, -\frac{1}{2} \right\rangle, \tag{A3}$$

$$|j\rangle = \sqrt{2}\, \frac{\sin(\pi a/L_z)}{\sqrt{L_z}} |\mathbf{k}'\rangle \left| \frac{3}{2}, \frac{1}{2} \right\rangle. \tag{A4}$$

Here $\mathbf{k}$ and $\mathbf{k}'$ are wave vectors of Bloch electron functions in the initial and final states in the quantum well plane:

$$|\mathbf{k}\rangle = \mathbf{u}_\mathbf{k}(\mathbf{r}) \frac{e^{i\mathbf{kr}}}{\sqrt{L_x L_y}}. \tag{A5}$$

Neglecting the Zeeman energy of the Mn spin in applied magnetic field,

$$E_i - E_j = E_k - E_{k'} + \mu_B g_{\mathrm{Mn}} B \approx E_k - E_{k'}, \tag{A6}$$

and calculating the spin matrix element

$$\left| \left\langle \frac{5}{2}, -\frac{1}{2} \right| \frac{1}{2} \sigma_- S_+ \left| \frac{3}{2}, \frac{1}{2} \right\rangle \right|^2 = \frac{5}{4}, \tag{A7}$$

we obtain the elementary probability of the Mn spin flip:

$$w=\frac{10\pi}{\hbar}\alpha^2\;\frac{\sin^4(\pi a/L_z)}{L_z^2}\,|\langle\mathbf{k}|\delta(\mathbf{r})|\mathbf{k}'\rangle|^2\,\delta(E_k-E_{k'}). \tag{A8}$$

The total probability W is obtained by integrating the elementary probability w over all occupied initial states and all vacant final states of band electrons:

$$W=\int_0^\infty wN(E)N(E')f(E)[1-f(E')]\;dE \tag{A9}$$

(hereinafter $E=E_k$, $E'=E_{k'}$). Integration is performed over the spin subband $\sigma_z=-1/2$, $f(E)=\{\exp[(E-E_e^F)/kT_e]+1\}^{-1}$ is the Fermi distribution function, T_e is the electron temperature, and $N(E)$ is the total number of states with a given energy in one spin subband. Since the density of states of two-dimensional carriers is independent of their energy, $D(E)=N(E)/\sqrt{L_xL_y}=D$, it can be taken out of the integrand. As a result, we have

$$\begin{aligned}W&=\frac{10\pi}{\hbar}\sin^4\left(\frac{\pi}{L_z}a\right)\alpha^2\left(\frac{D}{L_z}\right)^2\int_0^\infty f(E)[1-f(E')]\;dE\\&=\frac{10\pi}{\hbar}\sin^4\left(\frac{\pi}{L_z}a\right)\alpha^2\left(\frac{D}{L_z}\right)^2\frac{kT_e}{\exp(-E_e^F/kT_e)+1}.\end{aligned} \tag{A10}$$

In degenerate electron systems, the exponent is usually negligible in comparison to unity. Then the Mn spin relaxation rate $1/\tau_e=2W$ (Ref. 20) is described by the expression

$$\frac{1}{\tau_e}=2W\approx\frac{20\pi}{\hbar}\sin^4\left(\frac{\pi}{L_z}a\right)\alpha^2\left(\frac{D}{L_z}\right)^2kT_e. \tag{A11}$$

Usually the relaxation rate $1/\tau_e$ measured experimentally is averaged in the z direction. Thus, we derive from Eq. (A11) the formula

$$\left\langle\frac{1}{\tau_e}\right\rangle_z=\frac{15\pi}{2\hbar}\,\alpha^2\left(\frac{D}{L_z}\right)^2kT_e \tag{A12}$$

for the quantum spin relaxation rate averaged over the quantum well profile.

---

[1] Our experimental technique does not allow us to change the 12.2-ns time interval between probe pulses. For this reason, the relative uncertainty in τ_{Mn} in the quantum well with $x=0.12$ is notably larger than in samples with $x=0.03$.

[2] A final conclusion about direct effect of magnetic field on the relaxation rate of Mn spin due to free electrons demands a dedicated experimental and theoretical investigation. The main experimental difficulty is the presence of the D_0X line in the luminescence spectra of the studied quantum well at low pumping densities, which makes difficult identification of lines in the magnetic field range of 4–10 T. Further progress in this field is probably impossible without a notable improvement in the quality of investigated structures.

[3] In similar undoped Cd$_{1-x}$Mn$_x$Te heterostructures, a magnetic polaron is detected only at $x\geqslant 0.1$.[21,22] In order to check whether a magnetic polaron is formed in the quantum well, we studied the time evolution of the bound exciton line at zero magnetic field. We have not detected a notable shift in the spectral line position, which could be an indication that a magnetic polaron is formed in the studied quantum well.

[1] J. K. Furdyna, J. Appl. Phys. **64**, R29 (1988).
[2] P. A. Wolf, in *Semiconductors and Semimetals*, J. K. Furdyna and J. Kossut (eds.), Academic Press, London (1988), v. 25, p. 439.
[3] Y. Shapira, S. Foner, D. H. Rigley *et al.*, Phys. Rev. B **30**, 4021 (1984).
[4] I. Lambe and C. Kikuchi, Phys. Rev. **119**, 1256 (1960).
[5] T. Strutz, A. M. Witowski, and P. Wyder, Phys. Rev. Lett. **68**, 3912 (1992).
[6] D. Scalbert, J. Cernogora, and C. Benoit a la Guillaume, Solid State Commun. **66**, 571 (1988).
[7] Yu. G. Semenov, Zh. Éksp. Teor. Fiz. **81**, 1498 (1981) [Sov. Phys. JETP **54**, 794 (1991].
[8] W. Farah, D. Scalbert, and M. Nawrocki, Phys. Rev. B **53**, R10461 (1996).
[9] M. Tyazhlov, M. V. Lebedev, V. D. Kulakovskii *et al.*, Phys. Status Solidi B **188**, 565 (1995).
[10] V. D. Kulakovskii, M. G. Tyazhlov, A. F. Dite *et al.*, Phys. Rev. B **54**, 4981 (1996).
[11] V. D. Kulakovskii, M. G. Tyazhlov, A. I. Filin *et al.*, Phys. Rev. B **54**, R8333 (1996).
[12] S. I. Gubarev, V. D. Kulakovskii, M. G. Tyazhlov *et al.*, Ann. de Phys. **20**, C2-137 (1995).
[13] J. A. Gaj, R. Planel, and G. Fishman, Solid State Commun. **29**, 435 (1979).
[14] V. D. Kulakovskii, E. Lach, A. Forchel *et al.*, Phys. Rev. B **40**, 8087 (1989).
[15] H. A. Sayad and S. Bhagat, Phys. Rev. B **31**, 591 (1985).
[16] R. E. Kremer and J. K. Furdyna, Phys. Rev. B **32**, 5591 (1985).
[17] Yu. A. Bychkov and E. I. Rashba, Phys. Rev. B **44**, 6212 (1991).
[18] D. V. Kazakovtsev, I. B. Levinson, Zh. Éksp. Teor. Fiz. **88**, 2228 (1985) [Sov. Phys. JETP **61**, 1318 (1985)].
[19] L. C. Andreani, A. Pasquarello, and F. Bassani, Phys. Rev. B **36**, 5887 (1987).
[20] A. Abraham, *Nuclear Magnetism* [Russian translation], Izd. Inostr. Literatury, Moscow (1963), p. 334.
[21] G. Mackh, W. Ossau, D. R. Yakovlev *et al.*, Phys. Rev. B **49**, 10248 (1994).
[22] D. R. Yakovlev and K. V. Kavokin, Comments Condens. Matter Phys. **18**, 51 (1994).
[23] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, Pergamon Press, New York (1977).

Translation was provided by the Russian Editorial office.

Distinctive features of Raman–Nath scattering of light by sound in easy-plane antiferromagnets

E. A. Turov*

Institute of Metal Physics, Ural Department of the Russian Academy of Sciences,
620219 Ekaterinburg, Russia

(Submitted 5 February 1997)
Zh. Éksp. Teor. Fiz. **112**, 1464–1475 (October 1997)

This paper discusses the acoustic diffraction of light in the presence of an exchange-enhanced photoelastic interaction of aniferromagnetic origin resulting from acoustic modulation of the dielectric permittivity due to oscillations in the antiferromagnetism vector **L**. In the ''easy-plane'' type of antiferromagnet these oscillations arise from antiferromagnetoelastic interactions, and can be so large that the photoelastic interaction corresponding to them can be comparable in value to (or even exceed) the interaction in well-known nonmagnetic crystals actually used in acoustooptic devices. The advantage of antiferromagnets lies in the fact that both the diffraction angle and the amplitude of the diffracted light can depend on the magnitude and direction of a magnetic field in these materals. Here the Raman–Nath diffraction regime is discussed, which is probably more favorable from an experimental point of view for the antiferromagnets in question. It is shown that for these materials, the usual mechanism of photoelastic interaction associated with acoustic modulation of the index of refraction is accompanied by an additional mechanism arising from modulation of the polarization of the optical modes. Qualitative estimates are given for $FeBO_3$. © *1997 American Institute of Physics.* [S1063-7761(97)02210-5]

1. INTRODUCTION

In a previous paper by the author,[1] it was shown that in the ''easy-plane'' type of antiferromagnet the photoelastic interaction associated with the antiferromagnetic part of the dielectric permittivity tensor $\varepsilon_{\alpha\beta}$ is subject to the so-called exchange enhancement,[2–4] and can reach values that are at least comparable to those of well-known nonmagnetic materials (e.g., lithium niobate or sapphire) actually used in acoustooptic devices. Furthermore, antiferromagnets also have the additional advantage that both the angle of diffraction and the amplitude of the diffracted light in these materials can depend strongly on an applied magnetic field. This makes it possible to further control the acoustic diffraction (and, in particular, modulate it with a magnetic field).

Usually acoustooptic diffraction is discussed theoretically in two limiting regimes[2,3]—Bragg diffraction and Raman–Nath diffraction, depending on the value of the parameter

$$Q=\frac{2\pi d\lambda}{n\Lambda^2}, \tag{1}$$

where d is the width of the acoustic beam traversed by light, λ and Λ are the wavelengths of the light and sound, and n is the index of refraction of the light. For $Q\gg 1$ we have Bragg diffraction, while for $Q<1$ we have Raman–Nath diffraction. However, according to Ref. 4, an acoustooptic cell schematically shown in Fig. 1 in fact operates in the Raman–Nath regime even for $Q<10$.

In Ref. 1 mentioned above, applications of Bragg diffraction in the antiferromagnets $\alpha-Fe_2O_3$ (hematite) and $FeBO_3$ (iron borate) were discussed. However, it was discovered that the Raman–Nath regime is more favorable from the point of view of experimental feasibility, at least for these antiferromagnets. Thus, for example, in $FeBO_3$ for $d=0.1$ cm and an acoustic frequency $\Omega/2\pi=100$ MHz we obtain from Eq. (1) a value $Q=0.6$, which rapidly decreases with decreasing Ω. (Here and everywhere below, the required parameters of the material are taken from the article Ref. 1, where they are collected in an Appendix according to the original papers.) We mention only that for $FeBO_3$ the issue is the wavelength $\lambda=514.5$ nm for which the crystal is transparent (in blue light).

We will not pause to discuss the various physical pictures of acoustooptic diffraction in the Bragg and Raman–Nath regimes, referring the reader to well-known books on these topics.[2,3] The present article is devoted to Raman–Nath diffraction, for which the interaction length d of light with an acoustic beam is sufficiently small (e.g., in a thin plate). Discussion of this case usually involves only acoustic modulation of the index of refraction, which leads to modulation of the phase of light waves that cross the sonic beam. In agreement with Fig. 1 we treat the case of normally incident light with wave vector $\mathbf{k}\parallel z$ (normal to the plate). Sound with wavelength $\mathbf{q}\parallel x$ propagates in the plane of the plate. The z axis is directed along the trigonal axis of symmetry; for the x axis we will consider two alternatives.

If only this mechanism for photoelastic interaction existed (modulation of the index of refraction), we could find the antiferromagnetic diffraction component in the Raman–Nath regime by simply inserting the components of the tensor found in Ref. 1 into the corresponding expressions of Refs. 2, 3 or other papers. However, for the case of easy-plane antiferromagnets, the sound also modulates the optical anisotropy (birefringence) associated with the vector **L**, which lies in this plane. As a result, an additional mechanism

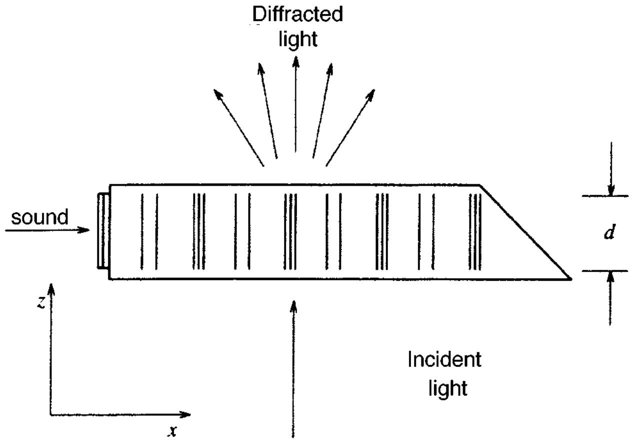

FIG. 1.

for Raman–Nath diffraction arises, due to modulation of the polarization of light passing through the acoustic beam. It is likely that this polarization mechanism was observed previously in the modulation of light by intrinsic low-frequency vibrations (0.5 MHz) in hematite.[6]

As we will see below, this latter mechanism sometimes turns out to be decisive.

2. DIELECTRIC PERMITTIVITY AND ITS ACOUSTIC MODULATION

For exchange magnetic structures of the form $\overline{1}(-)3_z(+)2_x(-)$ (see Ref. 6), which are typical of the antiferromagnet $FeBO_3$ of interest to us, the components of the dielectric permittivity tensor $\tilde{\varepsilon}_{\alpha\beta}$ affected by the light $(\mathbf{k}\|z)$, taking into account the antiferromagnetic contribution, have the form[7]

$$\begin{gathered}\tilde{\varepsilon}_{xx}=\varepsilon_\perp+b_1L_x^2+b_2L_y^2+c_1H_xL_y-c_3H_zL_x,\\ \tilde{\varepsilon}_{yy}=\varepsilon_\perp+b_2L_x^2+b_1L_y^2-c_2H_xL_y+c_3H_zL_x.\\ \tilde{\varepsilon}_{zy}=\tilde{\varepsilon}_{yx}=(b_1-b_2)L_xL_y-\frac{1}{2}(c_1+c_2)H_xL_x+c_3H_zL_y.\end{gathered}\tag{2}$$

Here $\varepsilon_\perp$, b_i, and c_i are constants determined experimentally. The coordinate system is chosen so that $x\|2$, i.e., the second-order symmetry axis, while $z\|3$, i.e., the trigonal axis.

The modulation of $\tilde{\varepsilon}_{\alpha\beta}$ arises from acoustically induced oscillations in $\mathbf{L}$. The most effective acoustic modes (for $\mathbf{q}\|x$) are those that contain the strains e_{xz} and e_{xy}, which give rise to rotation of $\mathbf{L}$ in the easy plane through a comparatively large angle

$$\delta\varphi\approx\sin\,\delta\varphi=-\delta L/L\approx-2U_\alpha e_{x\alpha},\quad \alpha=y,\,z,\tag{3}$$

due to the very small anisotropy in this plane. Here U_α is the coefficient of exchange enhancement of the photoelastic interaction introduced in Ref. 1 (Eqs. (25) and (16)). In weak fields $H_\perp=50$ Oe (the projection of $\mathbf{H}$ on the easy plane), which in the case of $FeBO_3$ are sufficient to overcome the basal anisotropy and destroy the domain structure, these coefficients can reach $U_\alpha\approx10^4$.

We further assume that $H_z\gg H_\perp$, and that H_z is limited only by the condition $H_z\ll H_E$ (where H_E is the exchange field). In this case we assign to the component $H_\perp$ the role of creating uniformly oriented states with $\mathbf{L}\perp\mathbf{H}_\perp$. Actually, in what follows we neglect those terms in Eq. (2) and in subsequent expressions that are proportional to $H_\perp$ (in Eq. (2) $H_\perp\equiv H_x$) compared with terms of the form H_zL_α and L_α^2. On the other hand, it is important to note that the enhancement coefficient U_α in Eq. (3) is actually independent of H_z ($\mathbf{H}\|z$ does not prevent rotation of $\mathbf{L}$ in the xy plane due to the strain $e_{x\alpha}$).

To simplify the theoretical discussion and interpretation of experiment, it is desirable that these acoustic deformations be referred to the normal (intrinsic) magneto-acoustic modes of the problem. Hence, we limit ourselves to discussing two alternatives here.

$\mathbf{H}_\perp\|\mathbf{M}_\perp\|x\|2$, $\mathbf{L}\|y$. We use the axes in which the component $\tilde{\varepsilon}_{\alpha\beta}$ of Eq. (2) are written. In this case, we must set $\delta L\equiv L_x$ $[L_y\approx L(1-L_x^2/2L^2)\approx L]$ and $\alpha=z$ in Eq. (3). It is true that oscillations of e_{xz} alone do not constitute a pure magnetoacoustic mode; however, the other component participating in the pure mode $e_{xy}\approx0.15e_{xz}$, i.e., it is relatively small. The phase velocity of these waves is $v_z\approx4\cdot10^5$ cm/s.

Let us split $\tilde{\varepsilon}_{\alpha\beta}$ into two parts—a static part and a dynamic part induced by the sound:

$$\tilde{\varepsilon}_{\alpha\beta}=\varepsilon_{\alpha\beta}+\Delta\varepsilon_{\alpha\beta}(x,t).\tag{4}$$

The latter is associated with L_x and will be written in the linear approximation with respect to $L_x\propto e_{x\alpha}$. From Eq. (2) we obtain

$$\begin{gathered}\varepsilon_{xx}=\varepsilon_\perp+b_2L_y^2,\quad \varepsilon_{yy}=\varepsilon_\perp+b_1L_y^2,\\ \varepsilon_{xy}=\varepsilon_{yx}=c_3H_zL_y,\quad L_y\approx L,\end{gathered}\tag{5}$$

and furthermore

$$\begin{gathered}\Delta\varepsilon_{xx}=-\Delta\varepsilon_{yy}=2\varepsilon_{xy}U_\alpha e_{x\alpha},\\ \Delta\varepsilon_{xy}=-2U_ze_{xz}(\varepsilon_{xx}-\varepsilon_{yy}).\end{gathered}\tag{6}$$

$\mathbf{H}_\perp\|\mathbf{M}_\perp\|x\perp2$, $\mathbf{L}\|y$. Here the coordinate axes are chosen differently from the first alternative, so that expression (2) for $\tilde{\varepsilon}_{\alpha\beta}$ must be transformed by rotating the x and y axes by $\pi/2$ about the z axis (we keep the notation as before). Here we note a well-known fact pointed out in Ref. 7: when $H_z=0$, the tensor $\tilde{\varepsilon}_{\alpha\beta}$ is invariant under rotation through any angle at all, not just $\pi/2$. Thus, all of the change arises from terms with $H_z\neq0$. Under these conditions (recall that $\mathbf{q}\|x$) there exists a pure magnetoacoustic mode with polarization $\mathbf{u}\|y$. In this case even in the new system of coordinates we again have $\delta L\equiv L_x$ and the component of the strain e_{xy} is nonzero. This means we must set $\alpha=y$ in Eq. (3). The corresponding phase velocity $v_y\approx6\cdot10^5$ cm/s.

The final expressions for the static and dynamic components in Eq. (4) take the form

$$\begin{gathered}\varepsilon_{xx}=\varepsilon_\perp+b_2L_y^2+c_3L_yH_z,\quad \varepsilon_{yy}=\varepsilon_\perp+b_1L_y^2-c_3L_yH_z,\\ \varepsilon_{xy}=0,\end{gathered}\tag{7}$$

$$\Delta\varepsilon_{xx}=\Delta\varepsilon_{yy}=0,\quad \Delta\varepsilon_{xy}=2U_y e_{xy}[(b_1-b_2)L_y^2+c_3L_yH_z]. \tag{8}$$

For both alternatives we assume that

$$e_{x\alpha}=a_\alpha\sin(qx-\Omega t),\quad \alpha=z\ \text{ or }\ y. \tag{9}$$

3. EQUATIONS OF THE ELECTROMAGNETIC FIELD AND THEIR SOLUTION

When $q\ll k$, we can show that $E_z\approx 0$ in the Raman–Nath regime, and the equations for the electric vector of the light passing through the acoustic beam have the form

$$c^2\frac{\partial^2E_x}{\partial z^2}-\frac{\partial^2}{\partial t^2}(\tilde\varepsilon_{xx}E_x+\tilde\varepsilon_{xy}E_y)=0.$$

$$c^2\frac{\partial^2E_y}{\partial z^2}-\frac{\partial^2}{\partial t^2}(\tilde\varepsilon_{yy}E_y+\tilde\varepsilon_{yx}E_x)=0. \tag{10}$$

For $\Delta\varepsilon_{\alpha\beta}=0$ these equations yield the intrinsic optical modes that propagate along the z axis, with index of refraction n_1 and n_2, and directions of the polarization vector in the xy plane. Under conditions of Raman–Nath diffraction, i.e., small width d $(Q<1)$, we seek a solution to the system (10) perturbed by the light in the form

$$E_\alpha(x,z,t)=A_\alpha(x,t)\exp\{i[\tilde k(x,t)z-\omega t]\},\quad \alpha=x,\ y. \tag{11}$$

We now substitute Eq. (11) into Eq. (10), taking into account Eqs. (5), (6) or (7), (8), and Eq. (9). In calculating the time derivatives we must take into account that $\partial A_\alpha/\partial t\sim\partial\Delta\varepsilon_{\alpha\beta}/\partial t\sim\Omega\ll\omega$ (in our case, $\Omega/\omega\lesssim 10^{-7}$). Thus, of all the terms with these derivatives we need only keep those with ω^2. This also applies to terms with the derivative $\partial\tilde k/\partial t$, although to a lesser degree. These terms can be eliminated when the condition

$$\left|d\,\frac{\partial\tilde k}{\partial t}\right|\ll\omega \tag{12}$$

holds. Once we have found the eigenvalues $\tilde k$, we will prove that this inequality holds (see below).

This in fact means that when we calculate the derivatives with respect to t in Eq. (10), the functions A_α, $\Delta\varepsilon_{\alpha\beta}$, and $\tilde k$ can be treated as constants. Furthermore, in the Raman–Nath regime (i.e., small width d) the amplitudes A_α are taken to be independent of z. It is this fact in particular that distinguishes the Raman–Nath regime from the Bragg regime, for which the evaluation of the function $A_\alpha(z)$ is most of the problem.[1–3] Here, however, the problem reduces firstly to including the effect of acoustic strain on the phase of the transmitted light waves (through $\tilde k(x,t)$), and secondly its effect on the optical polarization (through the ratio of amplitudes A_x and A_y).

With these approximations, Eqs. (10) lead to a system of wave equations for $A_\alpha(x,t)$ that determine both the index of refraction of the optical modes $\tilde n_{1,2}(x,t)$ and the polarization, given by the ratio of amplitudes of each of the modes. Both of these quantities take into account perturbations caused by sound. The total dielectric permittivity (4) enters into the resulting expressions:

$$\tilde n^2_{1,2}=\frac{\tilde\varepsilon_{xx}+\tilde\varepsilon_{yy}}{2}\pm\sqrt{\left(\frac{\tilde\varepsilon_{xx}-\tilde\varepsilon_{yy}}{2}\right)^2+\tilde\varepsilon^2_{xy}}, \tag{13}$$

$$\left(\frac{A_y}{A_x}\right)_1=-\left(\frac{A_x}{A_y}\right)_2=\frac{\tilde\varepsilon_{xy}}{\tilde n^2_1-\tilde\varepsilon_{yy}}\equiv\tilde G(x,t). \tag{14}$$

These expressions are suitable for both alternatives proposed above if we choose coordinate axes corresponding to each of them and expressions for $\varepsilon_{\alpha\beta}$ and $\Delta\varepsilon_{\alpha\beta}$ given in Sec. 2 (taking Eq. (4) into account). We discuss each alternative separately.

4. ACOUSTOOPTIC DIFFRACTION FOR $\mathbf{H}_\perp\|2\|x$ (CONFIGURATION 1)

In order to describe the light waves in the medium it is necessary to add to Eqs. (11) and (14) a boundary condition at the input $(z=0)$. Let this be

$$A_x\equiv A_{1x}+A_{2x}=E_0,\quad A_y\equiv A_{1y}+A_{2y}=0. \tag{15}$$

The solution to Eqs. (15) combined with Eq. (14) gives the wave amplitudes, while knowledge of the indices of refraction $\tilde n_1$ and $\tilde n_2$ from Eq. (13) allows us to find the corresponding phase advance in the medium, and consequently the components E_x and E_y of the resulting field where it leaves the acoustic beam $(z=d)$. In this case it is necessary to separate the part unperturbed by the sound (for $\Delta\varepsilon_{\alpha\beta}=0$) from the waves scattered by the sound.

To this end, we specify the phase advance $\tilde\Phi_i=\tilde k_i d$ corresponding to $\tilde n_i=(c/\omega)\tilde k_i(x,t)$ in the form

$$\tilde\Phi_i=\Phi_i+\Delta\Phi_i\sin(qx-\Omega t). \tag{16}$$

$$\Phi_i=\frac{\omega n_i d}{c},\quad \Delta\Phi_i=(-1)^i\Phi_i\frac{3\varepsilon_{xy}U_za_z}{n_i^2\sqrt{1+s^2}}, \tag{17}$$

$$s=\frac{2\varepsilon_{xy}}{\varepsilon_{xx}-\varepsilon_{yy}}, \tag{18}$$

where quantities without tildes ($\varepsilon_{\alpha\beta}$, n_i, etc.) denote unperturbed values of these quantities, i.e., in the absence of sound. Analogously, in Eq. (14) we also expand $\tilde G(x,t)$ in static and dynamic parts:

$$\tilde G=G_0+G_1\sin(qx-\Omega t), \tag{19}$$

where

$$G_0=\frac{s}{1+\sqrt{1+s^2}},\quad G_1=2U_za_z\frac{s^2-2}{\sqrt{1+s^2}(1+\sqrt{1+s^2})}. \tag{20}$$

In the resulting expressions for $E_x(d)$ and $E_y(d)$ it is necessary, as usual,[2,3] to expand an exponential of the form $\exp[\Delta\Phi_i\sin(qx-\Omega t)]$ in a Fourier series, whose coefficients are Bessel functions of the first kind $J_p(\Delta\Phi_i)$ (integer p).

The final expressions for the so-called diffraction indices

$$R_\alpha=\mathrm{Re}(E_\alpha/E_0),\quad \alpha=x,y$$

are given in the Appendix in view of their awkwardness (Eqs. (A1) and (A2)). Here we discuss the basic physical results obtained by using them.

Above all, we prove the inequality (12) used above by estimating the quantity

$$d\partial\widetilde{k}/\partial t=-\Delta\Phi_i\Omega\,\cos(qx-\Omega t).$$

According to Eq. (17),

$$|\Delta\Phi_i|\Omega=\frac{3d\Omega}{4cn_i}\,(\varepsilon_{xx}-\varepsilon_{yy})\,\frac{s}{\sqrt{1+s^2}}\,(2U_z a_z)\omega. \qquad (21)$$

Using the overestimates (see the Appendix of Ref. 1) $d=1$ cm, $\varepsilon_{xx}-\varepsilon_{yy}\lesssim 10^{-3}$, $2U_z a_z\lesssim 0.5$, and keeping in mind that the maximum value of $s/\sqrt{1+s^2}\approx 1$ (for $s\gg 1$), we find from Eq. (21) that for an acoustic frequency $\Omega/2\pi=10^6$ Hz

$$|\Delta\Phi_i|\Omega<10^{-5}\omega.$$

Consequently, the inequality (12) is easily satisfied.

The awkwardness of the general expressions for R_α given in the Appendix relates to the fact that the parameters s (Eq. (18)) and $\Delta\Phi_i$ (Eq. (17)), as yet undetermined by experiment, can in principle vary over a very wide range (principally due to the field H_z). This implies that the physical results for R_α can vary markedly depending on s and $\Delta\Phi_i$, not only quantitatively but qualitatively as well.

In this situation (Configuration 1), we can identify three types of effects:

1. Because there is optical anisotropy in the basal plane associated with the antiferromagnetism (the vector **L**) determined by the components Eq. (5) of the tensor $\varepsilon_{\alpha\beta}$, even in the absence of sound ($\Delta\varepsilon_{\alpha\beta}=0$) there exists a linear birefringence (the Cotton–Mouton effect), which leads to a rotation of the plane of polarization and ellipticity of light at the output that was linearly polarized at the input (see, for example, Refs. 7 and 8). In the Appendix this effect is given by the first lines of each of the expressions for R_x and R_y (Eqs. (A1) and (A2)).

A discussion of this effect for waves that pass through the plate without scattering would divert attention from the primary goal of this paper, which is to discuss diffraction in the Raman–Nath regime (although it nonetheless must be taken into account in the corresponding experiments). Therefore, we add only one more result to what was said above involving birefringence. Namely, we present a curve that describes the end of the vector **E** (more precisely, only its relative projections in the xy, plane $R_x\equiv x$ and $R_y\equiv y$). From the expressions in the Appendix (the first lines in the expressions for R_x and R_y), excluding any temporal factors, we find that this curve is an ellipse. We will not write out the general expression (i.e., for arbitrary values of the parameter G_0, ranges from $s/2$ to 1 depending on the magnitude of H_z); rather, we give only the characteristics of this ellipse in the limit $G_0\rightarrow 1$ (when $s\gg 1$). This will be a canonical ellipse $(x/a)^2+(y/b)^2=1$, whose semiaxes are

$$a=2J_0(\Delta\Phi)\cos(\Delta kd),\quad b=2J_0(\Delta\Phi)\sin(\Delta kd). \qquad (22)$$

Here we introduce

$$\Delta k=(k_1-k_2)/2\equiv(\omega/2c)(n_1-n_2),$$

which determines the phase difference of the optical eigenmodes, and also set

$$\Delta\Phi_1\cong-\Delta\Phi_2\equiv\Delta\Phi\approx-\frac{3\pi d}{2n\lambda}\left(\frac{s}{\sqrt{1+s^2}}\right)(\varepsilon_{xx}-\varepsilon_{yy})\times(2U_z a_z), \qquad (23)$$

since (see Eq. (17)) they differ only because of small differences in n_1 and n_2 (of order 10^{-3}–10^{-4}) due to the antiferromagnetism. Consequently, depending on the argument Δkd of the cosine and sine, the curve in question in the xy plane can change from the straight line $x=0$, when the field $\mathbf{E}(d)$ is linearly polarized along the y axis ($\cos(\Delta kd)=0$), to the straight line $y=0$ when $\mathbf{E}(d)$ is polarized along the x axis ($\sin(\Delta kd)=0$). In the intermediate case, when

$$\cos^2(\Delta kd)=\sin^2(\Delta kd)=1/2,$$

circular polarization is observed at the output.

In what follows we discuss the role played by Δkd as we investigate the diffracted waves. Here we add only that for other values of $G_0^2<1$, the polarization ellipse can be rotated relative to the x and y axes by a certain angle that depends on the field H_z.

2. A second effect is ordinary Raman–Nath diffraction associated with acoustic modulation of the phase Eq. (16). Due to the anisotropy of $\varepsilon_{\alpha\beta}$ we are dealing with so-called anisotropic Raman–Nath diffraction: for the boundary condition $\mathbf{E}\|x$ ($z=0$), diffracted waves appear at the output ($z=d$) with both components E_x and E_y. Due to multiple scattering, diffracted beams appear with wave vectors $\mathbf{k}_i\pm p\mathbf{q}$ and corresponding frequencies $\omega\pm p\Omega$ ($p=1,2,3,...$). Our retention of the frequency $\Omega\ll\omega$ in the latter expression is purely symbolic (in calculating the amplitude from solutions to Eq. (10) we have neglected terms of order Ω/ω), in order to illustrate the quantized (discrete) nature of the scattering process when momentum and energy are conserved.

Complete expressions Eqs. (A1) and (A2) for diffracted waves are given in the Appendix; here we demonstrate the peculiarities of Raman–Nath diffraction arising from antiferromagnetism by discussing first-order diffraction ($p=1$) with $\mathbf{k}_{1,2}\pm\mathbf{q}$. Essentially we again discuss only the limiting cases $s\ll 1$ ($G_0\approx s/2$) and $s\gg 1$ ($G_0\approx 1-1/s$).

Accordingly, for these two cases we obtain (the second line of Eq. (A1)):

$$G_0\ll 1:\ R_{x1}(\Delta\Phi)=J_1(\Delta\Phi)\{\cos[(\mathbf{k}_1+\mathbf{q})\mathbf{r}-(\omega+\Omega)t]-\cos[(\mathbf{k}_1-\mathbf{q})\mathbf{r}-(\omega-\Omega)t]\}, \qquad (24)$$

$$G_0\approx 1:\ R_{x1}(\Delta\Phi)=J_1(\Delta\Phi)\sin(\Delta kd)\times\{\sin[(\mathbf{k}+\mathbf{q})\mathbf{r}-(\omega+\Omega)t]-\sin[(\mathbf{k}-\mathbf{q})\mathbf{r}-(\omega-\Omega)t]\}, \qquad (25)$$

where we again invoke the condition Eq. (13) and introduce $\mathbf{k}=(\mathbf{k}_1+\mathbf{k}_2)/2$ along with $\Delta\mathbf{k}=(\mathbf{k}_1-\mathbf{k}_2)/2$. Recall also that $r_x\equiv x$ and $r_z\equiv d$. Note that in the first case only the first optical mode is diffracted (with index of refraction $n=n_1$). In fact, Eq. (24) implies simple modulation of the amplitude of this wave by the sound wave sin($qx-\Omega t$). For the second case both modes (with $n=n_1$ and $n=n_2$) are diffracted, and in this case will interfere.

In the first case, the amplitudes of the diffracted waves are determined by the function $J_1(\Delta\Phi)$, which is a maximum when $\Delta\Phi=1.85$. In the second case, for a maximum it is also necessary that

$$\Delta kd=(2p+1)\,\frac{\pi}{2}, \quad p=0,\,1,\,2,\dots . \tag{26}$$

This corresponds to the fact that the phase difference of two optical modes will be an odd multiple of $\pi/2$. Note that this is exactly the condition for a wave passing through the beam without scattering (see Eq. (22)) to be linearly polarized along the y axis.

Let us now turn to the component $R_{y1}(\Delta\Phi)$. Here we can write for an arbitrary value of G_0 (from the second line of Eq. (A2))

$$\begin{aligned} R_{y1}(\Delta\Phi)=-\,&\frac{G_0 J_1(\Delta\Phi)}{1+G_0^2}\,\sin(\Delta kd)\\ &\times\{\sin[(\mathbf{k}+\mathbf{q})\mathbf{r}-(\omega+\Omega)t]\\ &-\sin[(\mathbf{k}-\mathbf{q})\mathbf{r}-(\omega-\Omega)t]\}. \end{aligned}\tag{27}$$

From this expression it is clear that under the same conditions as in Eq. (26) the amplitude is a maximum as $G_0\to 1$ and decreases proportional to s when $s\ll 1$.

3. A third effect is additional Raman–Nath diffraction due to acoustic modulation of the polarization of the optical waves, determined by Eq. (14). However, in this case, as in the Bragg regime, there is only first-order diffraction. The corresponding expressions that determine the amplitude of the modulation G_1 in Eq. (20) are given by the next two lines of Eqs. (A1) and (A2):

$$\begin{aligned} R_x(G_1)=\,&\frac{2G_0G_1J_0(\Delta\Phi)}{(1+G_0^2)^2}\,\sin(\Delta kd)\\ &\times\{\cos[(\mathbf{k}+\mathbf{q})\mathbf{r}-(\omega+\Omega)t]\\ &-\cos[(\mathbf{k}-\mathbf{q})\mathbf{r}-(\omega-\Omega)t]\}, \end{aligned}\tag{28}$$

$$\begin{aligned} R_y(G_1)=\,&\frac{2(1-G_0^2)G_1}{(1+G_0^2)^2}\,J_0(\Delta\Phi)\sin(\Delta kd)\\ &\times\{\sin[(\mathbf{k}+\mathbf{q})\mathbf{r}-(\omega+\Omega)t]\\ &-\sin[(\mathbf{k}-\mathbf{q})\mathbf{r}-(\omega-\Omega)t]\}. \end{aligned}\tag{29}$$

The ratio of their amplitudes is $s/2$, while their phases are shifted by $\pi/2$. Thus, when $s\ll 1$ we have $R_y\gg R_x$, while for $s\gg 1$ we have the opposite, $R_x\gg R_y$. Recall that s in Eq. (18) is directly proportional to H_z according to Eq. (5); therefore, this ratio of R_x and R_y can be changed by changing the field H_z.

It is clear from the expressions given above that the amplitudes of the scattered waves depend very strongly on the field H_z (through the chain of relations (20), (18), and (5)), and also on the amplitudes of the acoustically induced phase shifts $\Delta\Phi$ (see Eq. (23)). Equations (28) and (29) still contain the amplitude of the acoustic modulation of the polarization G_1 Eq. (20) as well. Both of these quantities can vary over wide limits. Thus, for experimentally attainable values of the parameters, $\Delta\Phi$ can vary by three orders of magnitude (from 10 to 10^{-2}), while the amplitude G_1 can even change sign: $G_1\approx -2U_za_z$ for $s\ll 1$ and $G_1\approx +2U_za_z$ for $s\gg 1$; for $s=2$ it reduces to zero.

The quantity $\Delta\Phi$ affects the amplitudes, since it appears in the arguments of the Bessel functions (J_0 and J_1 in the case under discussion here). The function $J_0(\Delta\Phi)\approx 1$ for $\Delta\Phi\ll 1$ and falls off rapidly as $\Delta\Phi$ decreases, changing sign for $\Delta\Phi\approx 2.4$. The function $J_1(\Delta\Phi)\approx(1/2)\Delta\Phi$ for $\Delta\Phi\ll 1$, and reaches a maximum for $\Delta\Phi=1.85$. Both of these special cases can be achieved experimentally.

We now turn our attention once more to the factor sin(Δkd), which is present in all the expressions except Eq. (24). According to Eq. (26) it reduces to 1 for

$$(2\pi d/\lambda)(n_1-n_2)=(2p+1)\pi/2, \quad p=0,\,1,\,2,\dots .$$

From this we may estimate the width d corresponding to maximum amplitude of the diffracted beams under discussion:

$$d=d_{\max}=(2p+1)\,\frac{\lambda}{4(n_1-n_2)}\approx(2p+1)\cdot 0.1\ \text{cm} \tag{30}$$

(usually $n_1-n_2\approx 10^{-4}$). We can vary this width by varying the field H_z, since n_1 and n_2 are obtained from Eq. (13) when $\varepsilon_{\alpha\beta}=0$ due to the dependence of ε_{xy} on H_z; see Eq. (5).

Finally, it is of interest to estimate the ratio of the amplitudes of the diffracted beams, which are related to modulation of the polarization G_1 and the index of refraction $\Delta\Phi$, respectively. For example, let us estimate the ratio of amplitudes $R_{y1}(\Delta\Phi)$ from Eq. (27) and $R_{y1}(G_1)$ from Eq. (29). For the special case $s\ll 1$ (so that $|\Delta\Phi|\ll 1$ as well) we find

$$R_{y1}(\Delta\Phi)/R_{y1}(G_1)\approx s^2.$$

In this case, modulation of the polarization gives a larger first-order contribution to the diffracted amplitude than modulation of the index of refraction (phase).

5. ACOUSTIC DIFFRACTION FOR $\mathbf{H}_\perp \| x \perp 2$ (CONFIGURATION 2)

We now turn to Configuration 2, for which $\mathbf{H}_\perp \| x \perp 2$. In this case the quantities $\varepsilon_{\alpha\beta}$ and $\Delta\varepsilon_{\alpha\beta}$ are given by Eqs. (7) and (8). According to Eqs. (13) and (4), these expressions clearly imply that there is no linear modulation of the refractive indexes $\tilde{n}_{1,2}$ in general; conversely, the function $\tilde{G}(x,t)$ that determines the polarization ratio (14) has no unperturbed part ($G_0=0$), but does contain a nonzero modulated part ($G_1\neq 0$). In the end, it turns out that

$$G_0=\Delta\Phi=s=0. \tag{31}$$

However, according to Eqs. (7) and (8) G_1 is given by

$$G_1=\frac{\Delta\varepsilon_{xy}}{\varepsilon_{xx}-\varepsilon_{yy}}=\frac{2U_ya_y[(b_1-b_2)L_y^2+c_3L_yH_z]}{(b_1-b_2)L_y^2+2c_3L_yH_z}. \tag{32}$$

This yields $G_1 \approx 2U_y a_y$ or $G_1 \approx U_y a_y$ respectively, depending on which of the terms—spontaneous or field-induced—dominates in the difference $\varepsilon_{xx} - \varepsilon_{yy}$.

It is easy to see that the diffraction indices R_x and R_y for Configuration 2 can be obtained from expressions for the previous configuration as a special case, taking into account Eqs. (31) and (32). The results will be the following (they can be obtained directly from Eqs. (A1) and (A2)):

$$R_x = \cos(k_1 d - \omega t).$$

$$\begin{aligned} R_y = {} & 2G_1 \sin\left(\frac{k_1 - k_2}{2}\, d\right) \\ & \times \left\{ \cos\left[\left(\frac{\mathbf{k}_1 + \mathbf{k}_2}{2} + \mathbf{q} \right) \mathbf{r} - (\omega - \Omega) t \right] \right. \\ & \left. - \cos\left[\left(\frac{\mathbf{k}_1 + \mathbf{k}_2}{2} - \mathbf{q} \right) \mathbf{r} - (\omega - \Omega) t \right] \right\}, \end{aligned} \tag{33}$$

where

$$k_1 = \frac{\omega}{c} \sqrt{\varepsilon_{xx}}, \quad k_2 = \frac{\omega}{c} \sqrt{\varepsilon_{yy}},$$

while ε_{xx}, ε_{yy}, and G_1 are taken from Eqs. (7) and (32) respectively.

Thus, in this case there are only two symmetrically disposed diffraction beams (which are first-order only), with polarization vectors perpendicular to the polarization of the incident light. Ordinary birefringence does not appear here, because the incident light is one of the characteristic optical modes with index of refraction $n_1 = \sqrt{\varepsilon_{xx}}$.

The amplitude R_y from Eq. (33) is a maximum under conditions analogous to Eq. (26) (or Eq. (30)), and is determined by the quantity G_1 (Eq. (32)), which for an acoustic power flux of $1-10\ \mathrm{W/cm^2}$ can reach 0.15–0.5. Equation (32) also determines the field dependence of the amplitude R_y.

6. CONCLUSION

The goal of this paper is not so much to obtain results as to discuss them, and to obtain only those estimates that are necessary to guide the corresponding experiments. Nevertheless, the author is inclined to add a few comments with regard to the estimates. The fact is that in applications to $\mathrm{FeBO_3}$ for Configuration 1 we do not in fact know one of the parameters s in Eq. (18), more precisely the constant $c_3 L$ in the expression for ε_{xy} (Eq. (5)). The same may be said for Configuration 2, where we also need to know the terms with $c_3 L$ in Eqs. (7) and (8).

For both cases, it is probably sufficient to perform one of two possible magnetooptic experiments in order to determine the quantities $c_3 L$: either measure $\varepsilon_{xy}(H_z)$ for Configuration 1, or the field-dependent terms in the difference $\varepsilon_{xx} - \varepsilon_{yy}$ for Configuration 2.

As we have already noted, in order to obtain high intensity for the contributions to Raman–Nath diffraction given by Eq. (24), (25), and (27) associated with phase modulation, it is desirable to have $\Delta\Phi \approx 1.85$. This can be achieved by taking, e.g., $d = 2$ mm, $\varepsilon_{xx} - \varepsilon_{yy} \approx 10^{-3}$ and $s = 1$. Then from Eq. (23) we have $|\Delta\Phi| = 10 U_z a_z$, which yields the desired magnitude for $a_z < 10^{-5}$ by fitting U_z to the changes in the field $H_\perp$.

As for contributions to Raman–Nath diffraction connected with modulation of the polarization (with the parameter G_1 from Eq. (20)), here according to Eqs. (28) and (29) it is convenient to have $|\Delta\Phi| \ll 1$ for Configuration 1. In this case, under the condition $s < 1$ the quantity $R_y(G_1)$ will have the larger amplitude of the two contributions (28) and (29). Its value is determined by the product $U_z a_z$, which can give a relative amplitude of the diffracted wave of several tens of percent. For Configuration 2 the quantity G_1 of Eq. (32), which is in order of magnitude equal to $|U_z a_y|$, also give tens of percent amplitudes for the Raman–Nath diffraction amplitude in R_y Eq. (29).

Finally, the angle of first-order diffraction equals

$$\Theta_{1,2} \approx \sin \Theta_{1,2} = \pm \frac{q}{k}. \tag{34}$$

Taking into account that $q = \Omega / v$ and $k = 2\pi n / \lambda$, and choosing $v \approx 5 \cdot 10^5$ cm/s (the average of the two sound velocities mentioned above) (Configurations 1 and 2) we find for $\mathrm{FeBO_3}$ from Eq. (34) that

$$|\Theta_{1,2}| \approx 5 \cdot 10^{-3}\ \mathrm{rad} \approx 0°20'.$$

It is still necessary to keep in mind that in several cases the diffractive wave is polarized perpendicular to the incident wave.

Equation (33) probably applies to the simplest case for the first experiment, corresponding to Configuration 2, where there is no phase modulation (to linear order in the acoustic deformation) and where there is no undiffracted wave with $\mathbf{E} \| y$ for the given boundary conditions Eq. (15), so that the diffracted wave with $\mathbf{E} \| y$ can be observed against zero background. The geometry of the experiment, which is described in the text (for both configurations), requires samples in the form of plates with thickness of order 1 mm and a normal parallel to the trigonal axis of symmetry.

The author thanks M. I. Kurkina for discussing the results of the work and stimulating remarks, and V. V. Nikolaev for valuable help.

This work was supported by the Russian Fund for Fundamental Research (Grant N. 96-02-16489).

APPENDIX

The general expressions for the diffraction indices for configuration $\mathbf{H}_\perp \| 2 \| x$ are

$$\begin{aligned} R_x = \sum_{i=1,2} & \left\{ \frac{\delta_{i1} + G_0^2 \delta_{i2}}{1 + G_0^2} \left[J_0(\Delta\Phi_i) \cos(k_i d - \omega t) \right. \right. \\ & \left. + \sum_{p=1}^{\infty} J_p(\Delta\Phi_i) (\cos(ip+) + (-1)^p \cos(ip-)) \right] \\ & + \frac{G_0 G_1}{(1 + G_0^2)^2} (-1)^i J_0(\Delta\Phi_i) \\ & \times [\sin[(\mathbf{k}_i + \mathbf{q})\mathbf{r} - (\omega + \Omega) t] \end{aligned}$$

$$-\sin[(\mathbf{k}_i-\mathbf{q})\mathbf{r}-(\omega-\Omega)t]]\Big\}. \tag{A1}$$

$$\begin{aligned}R_y=\sum_{i=1,2}(-1)^{i-1}\Bigg\{&\frac{G_0}{1+G_0^2}\Bigg[J_0(\Delta\Phi_i)\cos(k_i d-\omega t)\\&+\sum_{p=1}^{\infty}J_p(\Delta\Phi_i)(\cos(ip+)+(-1)^p\cos(ip-))\Bigg]\\&+\frac{(1-G_0^2)G_1}{(1+G_0^2)^2}J_0(\Delta\Phi_i)[\sin[(\mathbf{k}_i+\mathbf{q})\mathbf{r}-(\omega+\Omega)t]\\&-\sin[(\mathbf{k}_i-\mathbf{q})\mathbf{r}-(\omega-\Omega)t]]\Bigg\}.\end{aligned} \tag{A2}$$

Here

$$(ip\pm)=(\mathbf{k}_i\pm p\mathbf{q})\mathbf{r}-(\omega\pm p\Omega)t.$$

$$\mathbf{k}_i\|z,\quad \mathbf{q}\|x,\quad r_x=x,\quad r_z=d,\quad k_i=\frac{\omega}{c}n_i.$$

[*] E-mail: theormag@ifm.e-burg.su

---

[1] E. A. Turov, Zh. Éksp. Teor. Fiz. **98**, 655 (1990) [Sov. Phys. JETP **71**, 365 (1990)].

[2] A. Yariv and P. Yeh, *Optical Waves in Crystals: Propagation and Control of Laser Radiation*, Wiley, New York (1984).

[3] E. Dieulesaint and D. Royer, *Elastic Waves in Solids*, Wiley, New York (1981).

[4] J. N. Lee and E. Vanderlugt, Proc. IEEE **77**, 1528 (1989).

[5] N. I. Evtikheev, V. V. Moshkin, V. L. Preobrazhenskiĭ, and N. A. Economov, JETP Lett. **35**, 38 (1982).

[6] E. A. Turov, Usp. Fiz. Nauk **164**, 325 (1994) [Phys. Usp. **37**, 303 (1994)].

[7] E. A. Turov, *Kinetic, Optical, and Acoustic Properties of Antiferromagnets* [in Russian], Ural Dept., USSR Acad. Sci., Sverdlovsk (1990).

[8] G. A. Smolenskiĭ, R. V. Pisarev, and I. G. Siniĭ, Usp. Fiz. Nauk **116**, 231 (1975) [Sov. Phys. Usp. **18**, 410 (1975)].

Translated by Frank J. Crowne

Dynamics of domain walls in ultrathin magnetic films

A. L. Sukstanskiĭ

Physicotechnical Institute, National Academy of Sciences of Ukraine, 310114 Donetsk, Ukraine

V. V. Tarasenko

Institute of Physics, Warsaw University Bialystok Branch, 15-424 Bialystok, Poland

(Submitted 7 February 1997)
Zh. Éksp. Teor. Fiz. **112**, 1476–1489 (October 1997)

The domain walls in ultrathin ferromagnetic films with uniaxial magnetic anisotropy are investigated theoretically. It is shown that taking account of the magnetodipole and magnetoelastic interactions leads to the appearance of an effective anisotropy with respect to the direction of the normal to the plane of the wall. The existence of a new type of domain walls—''corner'' walls, at which the magnetization vector is rotated in the plane making a certain angle, which depends on the film parameters, with the plane of the domain wall and the static and dynamic properties of these walls are investigated. The dependence of the limiting velocity of the domain walls on the film thickness is found. © *1997 American Institute of Physics.* [S1063-7761(97)02310-X]

1. INTRODUCTION

A vast number of both theoretical and experimental papers has been devoted to investigating magnetoelastic interaction in magnetically ordered crystals; various problems have been examined in these papers that are associated with this interaction: magnetoacoustic resonance, the appearance of a magnetoelastic gap in the spin-wave spectrum, etc. The effect of the magnetoelastic interaction on the static and dynamic properties of large-scale inhomogeneities of the domain wall type in the magnetization distribution has been analyzed in a number of papers; the fundamental value, which the relationship between the elastic and magnetic subsystems of a crystal has on the structure and dynamic properties of domain walls, was demonstrated in this case. For example, it is known that the formation of 180-degree domain walls in a ferromagnetic with cubic magnetic anisotropy is caused precisely by magnetoelastic interaction.[1]

The role of magnetoelastic interaction is especially important in antiferromagnetics and weak ferromagnetics, in which this interaction is exchange-amplified. The effect of magnetoelastic interaction on the dynamics of the main walls in antiferromagnetics was investigated in Ref. 2, where it was shown that the experimentally observed features in the field dependence of the steady-state velocity of the walls in weak ferromagnetics[3,4] are related to Cerenkov radiation of sound waves, which arises when the walls attain the velocity of sound.

The role of magnetoelastic interaction in ferromagnetics in the formation and the dynamic properties of domain walls is significantly smaller since the effect of this interaction on the properties of the domain walls is masked by the much stronger magnetodipole, for example, interactions. Moreover, the velocities of the walls in ferromagnetics are limited by the so-called Walker limit, the value of which is caused by relativistic interactions (rhombic anisotropy, magnetodipole interaction). The limiting velocity of the domain walls in ferromagnetics is usually much less than the velocity of sound, so that resonance effects such as the Cerenkov radiation of sound do not appear.

A situation exists, however, when the role of the magnetoelastic interaction in the formation of domain walls (and domain structures) in ferromagnetics can be exceptionally important. It is a question of an analysis of the static and dynamic properties of domain walls in ultrathin magnetic films with uniaxial magnetic anisotropy. In a ferromagnetic with uniaxial magnetic anisotropy the motion of domain walls is generally impossible unless the magnetodipole interaction is into account, in view of the presence of an additional constant motion of the system—the total projection of the magnetization on the anisotropy axis[5]—since ignoring this interaction leads to a zero value of the limiting velocity of the domain walls[6] and a formally infinitely large mass of the domain walls. As will be shown below, in an ultrathin ferromagnetic film the magnetodipole interaction in the basic small-parameter approximation—the ratio of the film thickness to the characteristic dimension in the magnetization distribution (or thickness of the domain wall)—leads only to a renormalization of the uniaxial anisotropy constant and does not determine the mass of the wall and its limiting steady-state velocity.[7] In this case, of course, the role of other interactions increases, which can affect these characteristics of the domain wall; this is primarily the magnetoelastic interaction.

This paper is also devoted to a theoretical analysis of the problem of the ''magnetoelastic'' mass of domain walls in an ultrathin ferromagnetic film with uniaxial magnetic anisotropy.

Let us consider a ferromagnetic film of thickness $2h$, lying in the xy plane, with the normal oriented along the anisotropy axis (or z axis). The energy of this system with the elastic subsystem and magnetoelastic interaction taken into account can be represented in the form

1063-7761/97/100804-08$10.00

$$W=W_m+W_e+W_{me},$$

$$W_m=\int d\mathbf{r}\left\{\frac{\alpha}{2}(\nabla\mathbf{M})^2-\frac{\beta}{2}M_z^2-\frac{1}{2}\mathbf{M}\mathbf{H}_d\right\}, \tag{1}$$

$$W_e=\int d\mathbf{r}\left\{\frac{\lambda}{2}u_{ii}^2+\mu u_{ij}^2\right\},\quad W_{me}=\int d\mathbf{r}\; bM_iM_ju_{ij},$$

where $\mathbf{M}$ is the magnetization vector, α and β are the nonuniform exchange interaction and uniaxial magnetic anisotropy constants, respectively ($\beta>0$), $\mathbf{H}_d$ is the demagnetizing field, u_{ij} is the strain tensor, λ and μ are the elastic moduli of the crystal. The integration in Eq. (1) is carried out over the volume of the crystal. For simplicity we have written the energy of the elastic subsystem of the crystal and the magnetoelastic interaction energy in isotropic form since taking account of the more complicated structure of these interactions does not lead to any fundamentally new results but only complicates the calculations.

The dynamics of the system being considered is described on the basis of the standard equations of motion for the magnetization (the Landau–Lifshitz equations) and the equations of elasticity theory $\rho\partial^2u_i/\partial t^2=\partial\sigma_{ij}/\partial x_j$, where σ_{ij} is the stress tensor and ρ is the density of the material.

In the general case it is impossible to solve the complete system of magnetoelastic equations; therefore, let us proceed in the following manner: let us first analyze the contribution of the magnetodipole interaction, then for an arbitrary (one-dimensional) magnetization distribution $\mathbf{M}(x)$ we will express the strain tensors on the basis of the equations of elasticity theory in terms of the Fourier components of this distribution. This will make it possible to eliminate the strain tensors from the energy of the magnetic substance and to write the latter only in terms of the distribution $\mathbf{M}(x)$ or its Fourier components (Sec. 2). We show in Sec. 3 that for the solution of the domain wall problem in ultrathin magnetic films the contribution of both the magnetodipole and magnetoelastic interactions leads to the appearance of some effective magnetic anisotropy in the plane of the film, which gives rise to the existence of so-called corner domain walls. The dynamic properties of domain walls in ultrathin magnetic films are considered in Sec. 4, in particular, the mass and limiting steady-state velocity of the walls are calculated.

2. EFFECTIVE ENERY OF MAGNETIC SUBSTANCE

The demagnetizing field $\mathbf{H}_d$ is defined by the equations of magnetostatics

$$\text{div}\,\mathbf{H}_d=-4\pi\,\text{div}\,\mathbf{M},\quad \text{curl}\,\mathbf{H}_d=0. \tag{2}$$

Here, the magnetodipole energy W_{md}, corresponding to some arbitrary magnetization distribution $\mathbf{M}(\mathbf{r})$, is equal to

$$W_{md}=-\frac{1}{2}\int d\mathbf{r}\cdot\mathbf{M}(\mathbf{r})\mathbf{H}_d(\mathbf{r})$$

$$=\frac{1}{8\pi}\int d\mathbf{r}\int d\mathbf{r}'\,\frac{(\nabla\cdot\mathbf{M}(\mathbf{r}))(\nabla\cdot\mathbf{M}(\mathbf{r}'))}{|\mathbf{r}-\mathbf{r}'|}. \tag{3}$$

Converting to the Fourier components of the unit magnetization vector $\mathbf{m}=\mathbf{M}/M_0$, $M_0=|\mathbf{M}|$ in the xy plane

$$\mathbf{m}(\mathbf{k},z)=\int d\mathbf{r}_\perp\,\mathbf{m}(\mathbf{r})\exp(-i\mathbf{k}\cdot\mathbf{r}_\perp), \tag{4}$$

where $\mathbf{r}_\perp=(x,y)$, $\mathbf{k}=(k_x,k_y)$, we obtain the following expression for the magnetodipole interaction energy:

$$W_{md}=\frac{1}{8\pi^2S}\int d\mathbf{k}\left\{\int dz\; m_z(\mathbf{k},z)m_z(-\mathbf{k},z)\right.$$

$$+\int\int dz\,dz'\,\exp(-|\mathbf{k}||z-z'|)[|\mathbf{k}|m_z(\mathbf{k},z)$$

$$\times m_z(-\mathbf{k},z')-2im_z(\mathbf{k},z)(\mathbf{k}\cdot\mathbf{m}(-\mathbf{k},z'))$$

$$\left.\times\text{sign}(z-z')+|\mathbf{k}|^{-1}(\mathbf{k}\cdot\mathbf{m}(\mathbf{k},z))(\mathbf{k}\cdot\mathbf{m}(-\mathbf{k},z'))]\right\}, \tag{5}$$

where S is the film surface area.

In the ultrathin magnetic films of interest to us the magnetization distribution $\mathbf{m}(\mathbf{r})$ can be considered to be uniform over the film thickness, i.e., it does not depend on the z coordinate. In this case the integrals with respect to z and z' in Eq. (5) are easily calculated, and as a result the magnetodipole energy for an arbitrary (in the xy plane) magnetization distribution in an ultrathin film can be written as

$$W_{md}=2\pi M_0^2\left\{\int d\mathbf{r}\;m_z^2+\frac{h^2}{4\pi^2}\int d\mathbf{k}[|\mathbf{k}|^{-1}(\mathbf{k}\cdot\mathbf{m}(\mathbf{k}))\right.$$

$$\left.\times(\mathbf{k}\cdot\mathbf{m}(-\mathbf{k}))-|\mathbf{k}|m_z(\mathbf{k})m_z(-\mathbf{k})]\right\}. \tag{6}$$

The first term in Eq. (6), as is easily seen, leads to a simple renormalization of the uniaxial anisotropy constant β in the energy of the magnetic subsystem W_m in Eq. (1): $\beta\to\beta_*=\beta-4\pi$. The second term in Eq. (6) is proportional to the square of the film thickness and, consequently, is small. Nevertheless, it plays a very important role in the dynamics of the domain walls since in a uniaxial ferromagnetic (ignoring magnetoelastic interaction) it is precisely this term that determines the mass of the domain wall, and if it is ignored, the effective magnetic anisotropy remains uniaxial and, as already pointed out, the mass of the wall turns out to be formally infinite; therefore, the domain wall cannot move. Specifically, the role of the magnetodipole interaction in the statics and dynamics of domain walls in the problem being considered will be discussed below.

Let us now turn to an analysis of the magnetoelastic interaction. The equations of elasticity theory with the interaction with the magnetic subsystem taken into account are

$$\rho\,\frac{\partial^2u_i}{\partial t^2}=\mu\,\frac{\partial^2u_i}{\partial x_j^2}+(\lambda+\mu)\,\frac{\partial^2u_j}{\partial x_i\partial x_j}+bM_0^2\,\frac{\delta(m_im_j)}{\partial x_j}, \tag{7}$$

where b is the magnetoelastic interaction constant.

We will assume that the film surface is free of stresses and, therefore, the wall conditions at the surface, i.e., for $z=\pm h$, have the form $\sigma_{iz}(\pm h)=0$.

In the problem we are treating of a domain wall in a ferromagnetic film the velocity of the domain walls is much less than the velocity of sound and, therefore, as an analysis shows, the time derivatives of the components of the displacement vector in Eq. (7) can be dropped out; this corresponds to passing to the elastostatic limit.

We will assume, as before, that the magnetization distribution is uniform over the film thickness. Moreover, taking account of the fact that the system being considered is isotropic in the (xy) film plane, with no loss of generality we can choose the Cartesian axes x and y in such a manner that the x axis is oriented along the normal to the plane of the domain wall. Then the magnetization distribution and the elastic fields will not depend on the y coordinate: $\mathbf{m}=\mathbf{m}(x)$, $\mathbf{u}=\mathbf{u}(x,z)$; this greatly facilitates solving the system of Eqs. (7) since in such a coordinate system the equation for the u_y component of the displacement vector turns out to be unrelated to the equations for the u_x and u_z components.

Determining the Fourier components $u_i(k,z)$

$$u_i(k,z)=\int_{-\infty}^{\infty} dx\ u_i(x,z)\exp(-ikx), \tag{8}$$

we rewrite the equation for the u_y component as

$$\mu(u_y''-k^2u_y)=-ikbM_0^2 s_{xy}(k), \tag{9}$$

where the notation

$$s_{ij}(k)=\int_{-\infty}^{\infty} dx\ m_i(x)m_j(x)\exp(-ikx), \tag{10}$$

has been introduced, and the prime denotes a derivative with respect to the coordinate z.

Since the right side of Eq. (9) does not depend on z, the solution of this equation, which satisfies the free wall conditions

$$\sigma_{yz}(z=\pm h)=[\mu u_y'+bM_0^2 s_{yz}(k)]_{z=\pm h}=0 \tag{11}$$

at the film surface, is simple to find:

$$u_y(k,z)=\frac{bM_0^2}{\mu k}\left[is_{xy}(k)-s_{yz}(k)\ \frac{\sinh k(z-h)}{\cosh kh}\right]. \tag{12}$$

Let us now consider the equations for the components u_x and u_z. In the Fourier components introduced with Eq. (8), these equations are

$$\begin{gathered}\mu u_x''-(\lambda+2\mu)k^2u_x+ik(\lambda+\mu)u_z'=-ikbM_0^2 s_{kk}(k),\\ (\lambda+2\mu)u_z''-\mu k^2u_x+ik(\lambda+\mu)u_x'=-ikbM_0^2 s_{xz}(k).\end{gathered} \tag{13}$$

The left side of the system of Eqs. (13) has the doubly degenerate eigenvalue $q=\pm k$. Therefore, the solution of the inhomogeneous Eqs. (13) is

$$\begin{gathered}u_x(k,z)=(a_1+b_1z)\cosh kz+(c_1+d_1z)\sinh kz+B_x,\\ u_z(k,z)=(a_2+b_2z)\cosh kz+(c_2+d_2z)\sinh kz+B_z.\end{gathered} \tag{14}$$

where $a_{1,2}$, $b_{1,2}$, $c_{1,2}$, $d_{1,2}$, B_x, B_z are some constants. Substituting Eq. (14) into (9), we find that the coefficients $a_{1,2}$, $b_{1,2}$, $c_{1,2}$, $d_{1,2}$ are interrelated by the expressions

$$\begin{gathered}a_2=\frac{i}{k}(\nu b_1-kc_1),\quad b_2=-id_1.\\ c_2=\frac{i}{k}(\nu d_1-ka_1),\quad d_2=-ib_1,\end{gathered} \tag{15}$$

and the z-independent terms are equal to

$$B_x=\frac{ibM_0^2}{k(\lambda+2\mu)}\ s_{kk}(k),\quad B_z=\frac{ibM_0^2}{\mu k}\ s_{kz}(k). \tag{16}$$

Then, substituting Eqs. (14) and (16) into the wall conditions at the film surface

$$\begin{gathered}\sigma_{kz}(z=\pm h)=[\mu(u_x'+iku_z)+bM_0^2 s_{kz}(k)]_{z=\pm h}=0,\\ \sigma_{zz}(z=\pm h)=[(\lambda+2\mu)u_z'+ik\lambda u_k)+bM_0^2 s_{zz}(k)]_{z=\pm h}=0,\end{gathered} \tag{17}$$

we find that $a_2=d_2=b_1=c_1=0$,

$$\begin{gathered}d_1=\frac{ibM_0^2}{2}\ \frac{[(\lambda+2\mu)s_{zz}(k)-\lambda s_{xx}(k)]}{\mu(\lambda+2\mu)}\ \frac{2\sinh(kh)}{\sinh(2kh)+2kh},\\ b_2=-id_1,\quad a_1=\frac{d_1}{2k}(\nu-1-2kh\ \coth(kh)),\\ c_2=\frac{id_1}{2k}(\nu+1+2kh\ \coth(kh)),\end{gathered} \tag{18}$$

where $\nu=(\lambda+3\mu)/(\lambda+\mu)$.

By substituting the expressions (12) and (14) found for $u_i(k,z)$ into the terms W_e and W_{me}, which describe the elastic subsystem of the crystal and magnetoelastic interaction, and integrating with respect to the z coordinate, we obtain after simple but rather lengthy calculations

$$\begin{aligned}W_1\equiv W_e+W_{me}=-\frac{(bM_0^2)^2S}{\pi\mu}\int_0^{\propto}dk\Bigg\{&|s_{xy}(k)|^2+|s_{xz}(k)|^2\\ &+\frac{\mu}{\lambda+2\mu}|s_{xx}(k)|^2+\frac{\tanh(kh)}{kh}|s_{yz}(k)|^2\\ &+\frac{\sinh^2(kh)}{kh[\sinh(2kh)+2kh]}\ \frac{|(\lambda+2\mu)s_{zz}(k)-\lambda s_{xx}(k)|^2}{(\lambda+2\mu)(\lambda+\mu)}\Bigg\}.\end{aligned} \tag{19}$$

It is important to point out that for small film thicknesses the energy W_1 turns out to be proportional to the thickness, and its density is finite, unlike the portion of the magnetodipole energy W_{md} that is not associated with the renormalization of the uniaxial anisotropy (the second term in Eq. (6)), the density of which itself becomes proportional to the film thickness and, therefore, small as $h\rightarrow 0$. This result is also precisely the reason why the role of magnetoelastic interaction is so important in ultrathin films.

3. DOMAIN BOUNDARIES IN ULTRATHIN MAGNETIC FILMS

Thus, by adding the energy of the magnetic subsystem of the crystal to the energy W_1, we obtain the total energy of the system, written only in terms of the magnetization distribution (or its Fourier components):

$$W=2M_0^2S\left\{\int_{-\infty}^{\infty}dx\left[\frac{\alpha}{2}\left(\frac{d\mathbf{m}}{dx}\right)^2-\frac{\beta_*}{2}\,m_z^2\right]\right.$$
$$+2h\int_0^{\propto}dk|k||m_x(k)|^2-\frac{(bM_0)^2}{2\pi\mu}\int_0^{\infty}dk\left\{|s_{xy}(k)|^2\right.$$
$$+|s_{xz}(k)|^2+\frac{\mu}{\lambda+2\mu}\,|s_{xx}(k)|^2+\frac{\tanh(kh)}{kh}\,|s_{yz}(k)|^2$$
$$\left.\left.+\frac{\sinh^2(kh)}{kh[\sinh(2kh)+2kh]}\,\frac{|(\lambda+2\mu)s_{zz}(k)-\lambda s_{xx}(k)|^2}{(\lambda+2\mu)(\lambda+\mu)}\right\}\right\}. \qquad (20)$$

The solution of the equations of motion for the magnetization in a system described by such an energy is impossible to find even for the case of a single domain wall. Therefore, in order to describe the static and dynamic properties of a wall in ultrathin magnetic films we employ a method analogous to that used in Ref. 7 to analyze the dynamic properties of domain walls in ultrathin magnetic films when only the magnetodipole interaction is taken into account. Actually, this method is a variation of the well-known approach to describing the dynamics of a domain wall on the basis of the Slonczewski equations.[8]

The dynamics of a ferromagnetic can be described on the basis of the Lagrangian L, which is

$$L=\frac{M_0}{g}\int dV\left\{\varphi\,\sin\,\theta\,\frac{\partial\theta}{\partial t}\right\}-W. \qquad (21)$$

where g is the gyromagnetic ratio, W is the energy of the ferromagnetic, and the angular variables θ and φ parametrize the magnetization vector $\mathbf{M}$,

$$\mathbf{M}=M_0\,(\sin\,\theta\,\cos\,\varphi,\ \sin\,\theta\,\sin\,\varphi,\ \cos\,\theta). \qquad (22)$$

We will assume that the magnetization distribution in a plane 180-degree domain wall moving along the x axis with velocity V is described by the usual relations (see, for example, Ref. 8):

$$\cos\,\theta(x)=-\tanh\left[\frac{x-Vt}{\Delta}\right],\qquad \varphi=\mathrm{const}, \qquad (23)$$

where Δ is the effective thickness of the domain wall.

The Fourier components of the magnetization $m_x(k)$ and the quantities $s_{ij}(k)$, entering into the energy (20) and corresponding to the magnetization distribution (23), are

$$\begin{aligned}
&m_x(k)=\frac{\pi\Delta}{\cosh(\pi k\Delta/2)},\qquad s_{xx}(k)=R_1(k)\cos^2\varphi,\\
&s_{xy}(k)=R_1(k)\sin\,\varphi\,\cos\,\varphi,\qquad s_{xz}(k)=R_2(k)\cos\,\varphi,\\
&s_{yz}(k)=R_2(k)\sin\,\varphi,\qquad s_{zz}(k)=2\pi\delta(k)-R_1(k),\\
&R_1(k)=\frac{\pi k\Delta^2}{\sinh(\pi k\Delta/2)},\qquad R_2(k)=\frac{i\pi k\Delta^2}{\cosh(\pi k\Delta/2)}.
\end{aligned} \qquad (24)$$

The quantity Δ and the angle φ are variational parameters of the problem and are determined from the Euler–Lagrange equations for the effective Lagrangian $L_{\mathrm{eff}}(\varphi,\Delta)$, which is obtained when the magnetization distribution (23) and the Fourier components (24) are substituted into the Lagrangian (21) and integrated with respect to the x coordinate:

$$\begin{aligned}
&L_{\mathrm{eff}}(\varphi,\Delta)=\frac{4hM_0V}{g}\,\varphi-E(\varphi,\Delta),\\
&E(\varphi,\Delta)=2M_0^2h\left[\frac{\alpha}{\Delta}+\Delta(\tilde{\beta}-\rho_1\,\sin^2\varphi+\rho_2\,\sin^4\varphi)\right].
\end{aligned} \qquad (25)$$

The notations

$$\begin{aligned}
&\tilde{\beta}=\beta_*-\frac{(bM_0)^2}{\mu},\qquad \rho_1=\frac{8h\,\ln 2}{\Delta},\\
&\rho_2=\frac{(bM_0)^2}{6\mu}\,\frac{3\lambda+2\mu}{\lambda+\mu},
\end{aligned} \qquad (26)$$

have been introduced here, and the quantity $E(\varphi,\Delta)$ has the meaning of the energy of a unit length of the domain wall. In the derivation of Eq. (26) we make use of the fact that the width Δ of the domain wall in ultrathin magnetic films is considerably greater than the film thickness $2h$.

Thus, we see that both the magnetodipole and magnetoelastic interactions lead, primarily, to a renormalization of the uniaxial anisotropy constant and, secondly, to the appearance of some effective anisotropy in the (xy) plane. It is important to emphasize that this effective anisotropy is not related to the crystallographic axes but to the normal to the plane of the domain wall, the direction of which in the (xy) plane is arbitrary, generally speaking. It turns out in this case that the magnetodipole interaction determines the ‘‘second-order’’ constant ρ_1 while the magnetoelastic interaction determines the ‘‘fourth-order’’ constant ρ_2.

By varying $L_{\mathrm{eff}}(\varphi,\Delta)$, we obtain equations determining the parameters φ and Δ:

$$\begin{aligned}
&\alpha-(\tilde{\beta}+\rho_2\,\sin^4\varphi)\Delta^2=0,\\
&\frac{V}{gM_0}-(\rho_1-2\rho_2\,\sin^2\varphi)\Delta\,\sin\,\varphi\,\cos\,\varphi=0.
\end{aligned} \qquad (27)$$

From Eqs. (27) it is easy to obtain explicit expressions for the domain wall parameters φ and Δ, which are fairly lengthy in the general case. Therefore, below we will limit ourselves to the most crucial case and will take into account that the uniaxial anisotropy constant is much greater than the effective rhombic anisotropy constants ρ_1 and ρ_2: the constant ρ_1 is small, scaling like the film thickness, while the constant ρ_2 (as with the renormalization of the uniaxial anisotropy constant caused by the magnetoelastic interaction) scales with the parameter $(bM_0)^2/\mu$. Here, as is easy to see, the parameter Δ depends weakly on the quantities ρ_1 and ρ_2 and is equal to $\Delta=(\alpha/\beta_*)^{1/2}$. It is easy to prove, starting from the second of Eqs. (27), that the angle φ, which determines the direction at which the magnetization vector leaves the plane of the domain wall, has three possible values in a stationary domain wall: 1) $\varphi=0$, which corresponds to a Néel domain wall (the rotation plane of the magnetization vector is perpendicular to the plane of the domain wall); 2) $\varphi=\pi/2$, which corresponds to a Bloch domain wall (the magnetization vector lies in the plane of the domain wall); 3) $\varphi=\varphi^*=\arcsin(\rho_1/2\rho_2)^{1/2}$. The magnetization distribution,

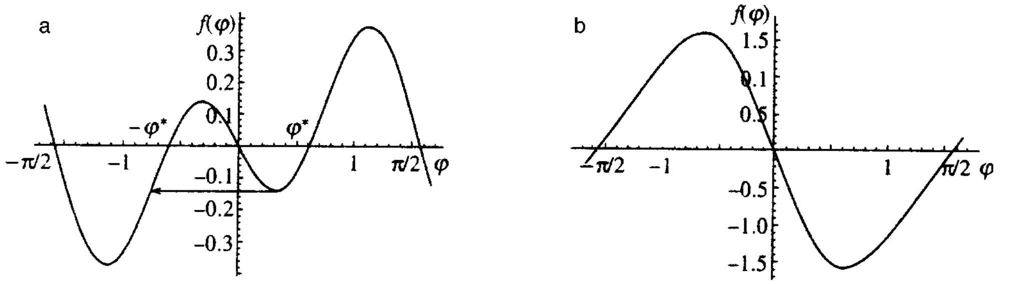

FIG. 1. The function $f(\varphi,p)$ for $p<1$ (a) and $p>1$ (b).

corresponding to the last solution, describes a domain wall in which the rotation plane makes the angle $\varphi=\pi/2-\varphi^*$ with the plane of the wall and, therefore, below we will call such a wall a corner domain wall. For the existence of a corner domain wall it is necessary that the inequality $\rho_1<2p_2$ be satisfied (of course, in this situation one of four equivalent orientations of the rotation plane of the magnetization vector in the domain wall can be realized: $\pm\varphi^*$ and $\pi\pm\varphi^*$).

An analysis of the stability of these solutions shows that the state corresponding to the Néel domain wall ($\varphi=0$) is unstable for any value of the effective anisotropy constants ρ_1 and ρ_2. The Bloch domain wall, in which the magnetization vector is turned around in the plane of the wall ($\varphi=\pi/2$), is stable if the inequality $\rho_1>2\rho_2$ is satisfied, which is opposite to the condition for the existence of the corner domain wall. If, however, the latter condition is satisfied, i.e., $\rho_1<2\rho_2$, then it is precisely the corner domain wall that turns out to be stable.[1)]

By using these expressions for the parameters ρ_1 and ρ_2, we find the criterion for the existence of a corner domain wall: the latter exists and is stable if the film thickness h is less than some critical value h_0:

$$h_0=\Delta\,\frac{(bM_0)^2}{24\ln 2}\,\frac{3\lambda+2\mu}{\mu(\lambda+\mu)}. \tag{28}$$

The quantity h_0 naturally has meaning if it amounts to at least several lattice constants, i.e., is $\geqslant 10^{-7}$ cm. However, the thickness of the domain wall Δ is usually of the order of 10^{-5} to 10^{-6} cm; therefore, one can speak of the critical thickness h_0 and the existence of corner domain walls only in magnetic substances with a sufficiently large magnetostriction.

4. DYNAMICS OF DOMAIN BOUNDARIES IN ULTRATHIN MAGNETIC FILMS

Let us now consider the dynamic properties of domain walls. The movement of domain walls is usually associated with the action of an external magnetic field, oriented in such a manner that one of the domains separated by a wall becomes energetically more advantageous: in the geometry being considered by us this is a field directed along the anisotropy axis z. In the approximation being considered, in which the thickness of the domain wall is a constant (which is valid in the usual situation for a large value of the figure of merit of the magnetic substance $Q=\beta/4\pi\gg 1$), the external magnetic field H, oriented along the anisotropy axis, is related to the velocity of the wall by the simple expression (see, for example, Ref. 8)

$$V=gH\Delta/\gamma, \tag{29}$$

where γ is the relaxation constant.

According to the second equation of the system (27) the velocity of the domain wall for $\Delta=\text{const}$ is related to the angle φ by the expression

$$V=V_0 f(\varphi),\quad f(\varphi)=\sin 2\varphi(\sin^2\varphi-p), \tag{30}$$

where the parameter $p=\rho_1/2\rho_2$ and the characteristic velocity of the domain wall[2)] $V_0=\rho_2(g,M_0)\Delta$ have been introduced.

The function $f(\varphi,p)$ is shown in Fig. 1. For $p<1$ (Fig. 1a) this function has two sets of extrema: one sets corresponds to values of the angle φ defined by the condition $\sin^2\varphi=\sin^2\varphi_+$ and the second to the condition $\sin^2\varphi=\sin^2\varphi_-$, where

$$\sin^2\varphi_\pm=\frac{1}{8}\,[3+2p+((2p-1)^2+8)^{1/2}]. \tag{31}$$

If, however, we have $p>1$ (Fig. 1b), then the extrema corresponding to the "+" sign in Eq. (31) vanish and the extrema corresponding to the "−" sign remain. The values of the angles $\varphi_\pm$ as a function of the parameter p are shown in Fig. 2, while the values $|f(\varphi,p)|$ at these points are shown in Fig. 3. These last values determine the limiting steady-state velocity V_c of the domain wall and, according to Eq. (29), the limiting value of the external field $H_c=\gamma V_c/g\Delta$, at which such motion is possible.

Let us first consider the case when the parameter $p<1$, and the corner domain wall turns out to be stable in the static case. As already pointed out, in this case four equivalent

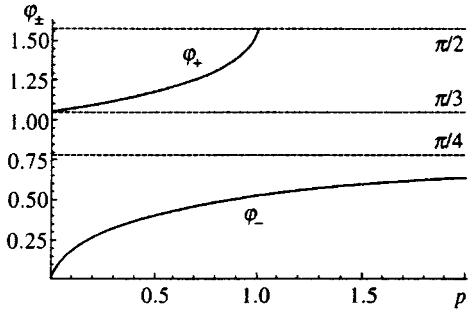

FIG. 2. Values of the critical angle $\varphi_\pm$ as a function of the parameter p.

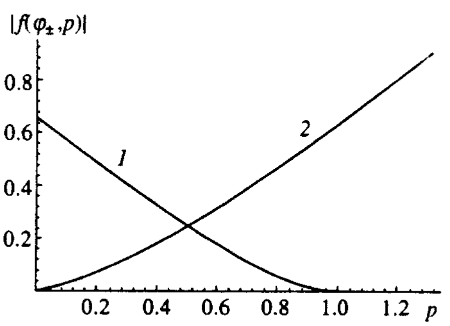

FIG. 3. Dependence of the values of the function $|f(\varphi,p)|$ at the extremal points $\varphi=\varphi_\pm$ on the parameter p.

values of the angle φ are possible, which determine the rotation plane of the magnetization vector in a static domain wall. Furthermore, to be specific we will assume that this angle is equal to $\varphi^*=\arcsin(p^{1/2})$.

For $1/2<p<1$ we have $\varphi^*>\pi/4$ and $|f(\varphi_+,p)|<|f(\varphi_-,p)|$; if, however, $p<1/2$ holds, then we have $\varphi^*<\pi/4$ and $|f(\varphi_+,p)|>|f(\varphi_-,p)|$ (this is precisely the situation depicted in Fig. 1a). For $p=1/2$ the angle satisfies $\varphi^*=\pi/4$, and the extrema become identical, $|f(\varphi_+,p)|=|f(\varphi_-,p)|$.

The asymmetrical magnetization distribution with respect to the positive and negative x axis in a static corner domain wall causes the domain wall velocity to depend on the direction of the external magnetic field, i.e., a unique ''nonreciprocity'' exists: $|V(H)|\neq|V(-H)|$.

If the external field is directed along the positive direction of the z axis $(H>0)$, then the wall moves in the positive x direction $(V>0)$ and for $p<1/2$, as is clearly seen from Fig. 1a, the angle φ increases with velocity from its static value $\varphi=\varphi^*$ to the value $\varphi=\varphi_+$, which corresponds to a velocity $V_+=V_0 f(\varphi_+,p)$, and the limiting field value, for which steady-state motion of the domain wall is possible, is equal to $H_+=\gamma V_+ g\Delta$. If, however, we have $H<0$ and the domain wall moves in the negative direction of the x axis $(V<0)$, then with an increase in velocity the angle φ decreases from the initial value $\varphi=\varphi^*$ to $\varphi=\varphi_-$, which corresponds to the velocity $V_-=V_0 f(\varphi_-,p)$, and $|V_-|<V_+$.

The velocity V_- is achieved for a field $H_-=\gamma V_-/g\Delta$, which is smaller in absolute magnitude than H_+, $|H_-|<H_+$. With a further increase in the negative field the solution, corresponding to the branch of the multivalued function $\varphi=\varphi(V,p)$ being considered, vanishes and therefore a transition to the other branch occurs unavoidably, on which the velocity V_- is no longer an extremum (this transition is shown by the arrow in Fig. 1(a). Thus, for $H=H_-$ the rotation plane of the magnetization vector in the domain wall changes suddenly: the value of the angle φ changes abruptly from the extremum value φ_- to some value φ', which lies in the interval $(-\varphi_+,-\varphi^*)$, and with a further increase in the field the velocity of the domain wall continues to increase (in absolute magnitude) up to V_+.

If we have $1/2<p<1$ and $|f(\varphi_-,p)|>|f(\varphi_+,p)|$, the situation is reversed: if $H<0$ holds, then, starting from the static value of the angle $\varphi=\varphi^*$, the domain wall reaches the limiting velocity $V_c=V_0|f(\varphi_-,p)|$ as the field increases without any sudden changes of the rotation plane, while for $H>0$ the value $V_0|f(\varphi_+,p)|$ is reached first, and then for $H=H_+$ a jump occurs in the angle φ to the other branch, and as the field increases further the velocity of the domain wall reaches its limiting value $V_0|f(\varphi_-,p)|$.

Starting from the expression for $E(\varphi,\Delta)$ (25) and Eq. (30), which determines the angle φ as a function of the velocity V, one can find in principle the energy of a moving domain wall as a function of its velocity (or the value of the external magnetic field) and the parameter p. In general form this expression is quite involved; however, for small domain wall velocities the calculations can be greatly simplified if, starting from Eqs. (25) and (30), we write the energy density $\sigma(V)$ of a moving corner domain wall in the form

$$\sigma^*(V)=\sigma^*(0)+2M_0^2\rho_2\Delta\left(\frac{V}{V_0}\right)^2\frac{1}{\sin^2 2\varphi},$$

$$\sigma^*(0)=4M_0^2\beta^*\Delta\left(1-\frac{\rho_1^2}{8\beta\rho_2}\right), \tag{32}$$

where $\sigma(0)$ is the effective energy density of the static corner domain wall, and the second term has the meaning of the kinetic energy density of the domain wall.

For small values of the external field the velocity V of the domain wall is also small, and the angle φ differs only slightly from its equilibrium value in the static corner domain wall: $\varphi=\varphi^*+\psi$, $\psi\ll 1$. From Eq. (30) linearized with respect to ψ we find

$$\psi=\frac{V}{V_0}\frac{1}{4p(1-p)}. \tag{33}$$

By substituting this value into Eq. (32), we obtain the effective energy density of a moving corner domain wall to order $(V/V_0)^3$:

$$\sigma(V)=\sigma(0)+\frac{mV^2}{2}+q\left(\frac{V}{V_0}\right)^3. \tag{34}$$

$$m=\frac{2}{g^2\Delta}\frac{\rho_2}{\rho_1(2\rho_2-\rho_1)},\quad q=\frac{M_0^2\Delta}{8}\frac{2p-1}{[p(1-p)]^{5/2}}, \tag{35}$$

where the quantity m can be interpreted as the mass density of the corner domain wall.

Let us recall that Eqs. (34) and (35) are obtained for a corner domain wall, in which the static value of the angle φ is equal to φ^*. An analogous calculation of the energy of a corner domain wall, in which the angle φ is equal to $-\varphi^*$ in the static condition, leads to exactly the same expression for $\sigma(V)$, but with q replaced by $-q$. Thus, one can conclude that despite the asymmetry in the dynamics of corner domain walls, having different initial (or static) values of the angle φ, their masses (i.e., the coefficients of the squared velocity in the kinetic energy of the domain walls) are identical; however, the coefficients of the higher powers of the expansion in terms (V/V_0) are different: they differ in sign. Consequently, for one and the same velocity, corner domain walls existing on different branches of the function $\varphi=\varphi(V,p)$ have different energy.

As follows from Eqs. (34) and (35), in the case considered above $(p<1/2,\ V<0)$ $qV^3>0$ holds and, therefore, for one and the same value of the external field $H<0$ which determines the velocity of the domain wall, the energy of a corner domain wall with angle φ in the interval (φ_-,φ^*) turns out to be greater than the energy of a domain wall with angle φ the interval $(-\varphi_+,-\varphi^*)$. Consequently, the rotation plane of the magnetization vector in a moving corner domain wall can undergo a sudden reorientation, i.e., the angle φ can pass to the other branch of the function $\varphi=\varphi(V,p)$, not only at the extremal velocity V_-, as described above, but also at a lower (in absolute magnitude) velocity, since such a transition is energetically advantageous. For $|V|<V_-$ this process can be induced by thermal fluctuations in the magnetic or elastic subsystems of the film or by the interaction of the moving domain wall with lattice defects; in this case the transition to the energetically more advantageous state must be accompanied, of course, by the release of energy in the form of spin-wave radiation and sound.

If $1>p>1/2$ holds, then we have $qV^3>0$ for $V>0$; in this case the corner domain walls, for which $\varphi=-\varphi^*$ holds in the static condition, are energetically more advantageous and a reorientation of the domain wall can occur in positive fields $H>0$.

Let us now consider the case $p>1$, in which the static Bloch domain wall with a static value of the angle $\varphi=\pm\pi/2$ is stable. Both states are equivalent, and to be specific we will assume that $\varphi=\pi/2$ holds in the static condition.

The dependence of the angle φ on the velocity of the domain wall in this case is described by Eq. (30) as before; however, as already pointed out, for $p>1$ the function $f(\varphi,p)$ has extrema only for $\sin^2\varphi=\sin^2\varphi_-$ (see Fig. 1b). The nonreciprocity of the motion of the domain wall described above is absent in this case, $V(H)=-V(-H)$, and sudden jumps of the rotation plane of the magnetization vector in the domain wall are also absent: as the field rises the velocity of the domain wall increases linearly according to Eq. (29), achieving its limiting value $V_c=V_0|f(\varphi,p)|$ for $H_c=\gamma V_c/g\Delta$.

If the effect of the magnetoelastic interaction on the dynamics of the domain wall is negligibly small compared with the effect of the magnetostatic interaction, i.e., in the limit $p\gg 1$, then, as follows from Eq. (30), the limiting emergence angle of the magnetization vector from the plane of the domain wall is close to $\pi/4$ (this is typical for a Bloch domain wall in a ferrormagnetic), $|f(\varphi_-,p)|\approx p$, and the limiting domain wall velocity is equal to $V_c\approx pV_0=4hgM_0\ln 2$. With a decrease of the parameter p the limiting velocity of the domain wall also decreases; as $p\rightarrow 1$, the limiting angle φ_- is close to $\pi/6$, while the limiting velocity of the Bloch domain wall being considered is equal to $V_c\approx 0.63V_0$.

The mass of the Bloch domain wall is calculated analogously to the mass of the corner domain wall and is equal to

$$m=\frac{2}{g^2\Delta}\,\frac{1}{\rho_1-2\rho_2}. \tag{36}$$

It is interesting to point out that in the limit $p\gg 1$ the mass $m_1=2hm$ of a unit length of a domain wall is independent of all material constants of the film except the gyromagnetic ratio, $m_1=[2g^2\ln 2]^{-1}$.[7]

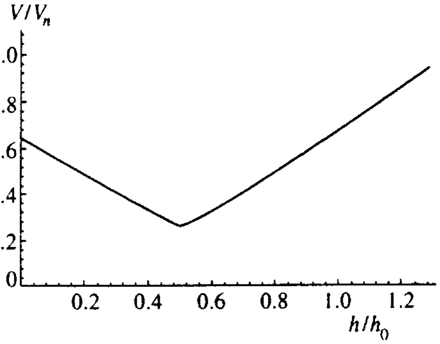

FIG. 4. Dependence of the limiting steady-state velocity of a domain wall on the film thickness.

Note also that as $\rho_1\rightarrow 2\rho_2(\rho\rightarrow 1)$ the masses of both the Bloch (36) and corner (35) domain walls increase without limit and for $\rho_1=2\rho_2$ they formally go to infinity.

This does not mean, however, that the domain wall cannot move: the fact of the matter is that the value $p=1$, which separates the existence and stability regions of the two types of domain walls, is isolated, and the dependence of the kinetic energy of the domain wall on the velocity for $p=1$ turns out to be nonquadratic. As it is easy to see from Eq. (30), for $p=1$ an expansion of the function $f(\varphi)$ for small deviations of the angle φ from its equilibrium value of $\pi/2$ starts out with a term that is cubic in ψ rather than a linear term, and in place of the $\psi\sim V$ dependence (33) we have $\psi=(V/2V_0)^{1/3}$. Moreover, for $p=1$ the expansion of the kinetic energy of the domain wall for small velocities starts out with ψ^4 rather than ψ^2, and as a result we arrive at the following nonstandard dependence of the kinetic energy of the domain wall on the velocity:

$$\sigma(V)=\sigma(0)+\frac{M_0^2\rho_1\Delta}{2}\left(\frac{V}{2V_0}\right)^{4/3}. \tag{37}$$

5. CONCLUSION

The analysis conducted above attests to the important role which the magnetoelastic interaction plays in ultrathin magnetic films. It is the magnetoelastic interaction which results in the appearance of a fourth order effective energy of rhombic anisotropy $(\rho_2\neq 0)$ in the energy of a magnetic substance, the presence of which leads to extremely important features in the structure and dynamics of domain walls in ultrathin magnetic films: if this interaction is sufficiently strong and the inequality $2\rho_2>\rho_1$ is satisfied, then corner domain walls exist and are stable in the static condition, and their behavior differs in a nontrivial manner from the behavior of the usual Bloch domain walls.

Taking account of the magnetoelastic interaction implies that the limiting steady-state velocity of domain walls in ultrathin magnetic films becomes nonmonotonically dependent on the film thickness. As was shown, this velocity is equal to $V_c=V_{0\,\max}\{|f(\varphi_-,p)|,\ |f(\varphi_+,p)|\}$, and therefore for $p>1/2$ $(h>h_0/2)$ we have $V_c=V_0|f(\varphi_-,p)|$ and it decreases with a decrease in film thickness, whereas for $p<1/2$ $(h<h_0/2)$ we have $V_c=V_0|f(\varphi_+,p)|$ and it increases with a decrease in thickness. The dependence of the limiting velocity on the

film thickness is shown in Fig. 4. For large thicknesses, when the effect of the magnetoelastic interaction is small compared with the magnetodipole interaction and $p \gg 1$ holds, the quantity V_c is directly proportional to the thickness, $V_c \approx pV_0 = 4hgM_0 \ln 2 \propto h$. The minimum value of the limiting velocity of the domain wall V_c is reached for $h = h_0/2$ and is equal to $V_c^{\min} = V_c(h = h_0/2) = \rho_2 g M_0 \Delta/2$, and as the film thickness approaches zero we have $V_c(h \to 0) = 3\rho_2 g M_0 \Delta/4 = 1.5 V_c^{\min}$.

One of the authors (A. L. Sukstanskiĭ) is sincerely grateful to Prof. M. Svenskii (Institute of Physics, Warsaw University, Bialystok Branch) for his hospitality and support during the performance of this work.

---

1)We note that a unique analog of corner domain walls exists in some weak ferromagnetics, in which the magnetization vectors of the sublattices are rotated in the ac plane of easy magnetization, and the axis along which the magnetization distribution varies does not coincide with the *b* axis of the magnetic substance.

2)In massive ferromagnetics the characteristic domain wall velocity is the Walker velocity V_W (the limiting steady-state velocity of a domain wall).[6] The characteristic velocity V_0 we have introduced does not have such a simple physical interpretation and is simply a convenient dimensional parameter of the problem.

[1] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* Pergamon, Oxford (1960).

[2] V. G. Bar'yakhtar, B. A. Ivanov, and A. L. Sukstanskiĭ, Zh. Éksp. Teor. Fiz. **75**, 2183 (1978) [Sov. Phys. JETP **48**, 1100 (1978)].

[3] M. V. Chetkin, A. N. Shalygin, and A. de la Campa, Zh. Éksp. Teor. Fiz. **75**, 2345 (1978) [Sov. Phys. JETP **48**, 1184 (1978)].

[4] V. G. Bar'yakhtar, B. A. Ivanov, and M. V. Chetkin, Usp. Fiz. Nauk **146**, 417 (1985) [Sov. Phys. Usp. **28**, 563 (1985)].

[5] B. A. Ivanov, Fiz. Nizk. Temp. **4**, 352 (1978) [Sov. J. Low Temp. Phys. **4**, 171 (1978)].

[6] L. R. Walker (unpublished), quoted by J. F. Dillon in *Magnetism*, Vol. 3 (Academic Press, New York, 1963).

[7] A. Stankevich and V. V. Tarasenko, ''*Proceedings 6th European Conf. on Magnetism*'' (Poznan, Poland, 1996).

[8] A. P. Malozenoff and J. C. Slonczewski, *Magnetic Domain Walls in Bubble Materials* (Academic Press, 1979).

Translated by Eugene R. Heath

Nonlinear response of photorefractive lithium tantalate and niobate at acoustic frequencies

P. A. Prudkovskiĭ, O. V. Skugarevskiĭ, and A. N. Penin

M. V. Lomonosov Moscow State University, 119899 Moscow, Russia

(Submitted 9 April 1997)
Zh. Éksp. Teor. Fiz. **112**, 1490–1498 (October 1997)

An amplification of the intensity of pump oscillations is observed experimentally at frequencies from 100 Hz to 1 kHz during photoinduced light scattering and holographic-type parametric scattering in photorefractive lithium tantalate and niobate. Possible ways are analyzed for explaining the existence of a photorefractive response in these crystals over times of 10^{-2}–10^{-3} s, which are five orders of magnitude shorter than the Maxwell time. © *1997 American Institute of Physics.* [S1063-7761(97)02410-4]

1. INTRODUCTION

The phenomenon of photorefraction was observed in the 1960's,[1,2] but interest in photorefractive media continues up to the present owing to the enormous variety of their properties. For example, a holographic grating inscribed on lithium niobate can be kept for months,[3] while the characteristic photorefractive response time in barium titanate is nano- or picoseconds.[4] According to the generally accepted theory of photorefraction proposed by Glass,[5] when photorefractive crystals are illuminated a macroscopic current develops in them and rearranges the electric charge density. The electrostatic field of the separated charges changes the refractive index of the medium owing to an electrooptical effect. This theory assumes the existence of a single temporal parameter which characterizes the entire photorefractive process, the Maxwell time,

$$\tau_M = \varepsilon_{st}/4\pi\sigma, \tag{1}$$

where ε_{st} is the dielectric permittivity of the medium and $\sigma=\sigma_T(1+I/I_s)$ is its conductivity, whose value under illumination varies over a scale length determined by the parameter I_s.[6]

From time to time, however, papers are published which indicate the existence of a photorefractive response with a characteristic time very different from the Maxwell time.[4,7,8] In our work we have also observed a photorefractive response in copper-doped lithium tantalate and niobate at characteristic frequencies $f \sim 0.1-1$ kHz, whereas the Maxwell times for these crystals are tens of minutes.

2. DESCRIPTION OF THE EXPERIMENT

When a coherent light beam strikes a photorefractive medium, wide-angle scattering that is degenerate in frequency may take place over a characteristic time. This scattering is usually associated with optical damage and efforts are made to avoid it, as it greatly reduces the pump intensity. This is photoinduced light scattering, which has been explained in terms of the recording and self-consistent amplification of a large number of holographic gratings.[9] When certain conditions are met in directions which form a distinct cone whose shape is determined by the four-wave spatial synchronization condition $2\mathbf{k}_p=\mathbf{k}_{s_1}+\mathbf{k}_{s_2}$ (where the subscripts p and s denote the pump mode and scattered light, respectively), in anisotropic crystals it is possible to observe a sudden rise in the scattered light intensity. This effect arises because in these directions two, rather than one, scattered light modes interact on a single holographic grating and is referred to as holographic-type parametric scattering.[10]

In a study of the temporal characteristics of the photorefractive response, we have examined $o-e$-mode photoinduced light scattering and $oo-ee$-mode holographic-type parametric scattering in $LiTaO_3$:Cu and $ee-oo$-mode holographic-type parametric scattering in $LiNbO_3$:Cu. The pump intensity was modulated at one or two frequencies in the range from 100 Hz to 1 kHz with a small (less than 10%) depth of modulation. The magnitude of the photorefractive response of the crystals at these frequencies was determined from the change in the depth of modulation in the intensity of the light during the scattering process.

The experimental apparatus is shown schematically in Fig. 1. An ordinary polarized beam from a helium-cadmium laser ($\lambda=442$ nm), for lithium tantalate, or an extraordinary polarized beam from an argon laser ($\lambda=488$ nm), for lithium niobate, was passed through an electrooptical modulator and was incident on an $X-Z$-cut crystal. Two photodetectors detected the pump and scattered light intensities. A computer connected to the measurement apparatus through a CAMAC crate processed the data. The depth of modulation in the intensity of the light in both channels was determined from the magnitudes of the Fourier components at the modulation frequencies, and in order to isolate the regular modulation more precisely from the noise, we used the spectrum of the fluctuations in the intensity (rather than the spectrum of the signal), which is the Fourier transform of the autocorrelation function of the luminous intensity, normalized to the square of the average value. Figure 2 shows a typical plot of the autocorrelation function of the pump and scattered light intensities modulated at the two frequencies, together with their Fourier spectra. A coefficient characterizing the magnitude of the photorefractive response at these frequencies was obtained from the ratio of the amplitudes of the Fourier components at the modulation frequencies:

$$K(f)=\sqrt{G_s(f)/G_p(f)}. \tag{2}$$

where

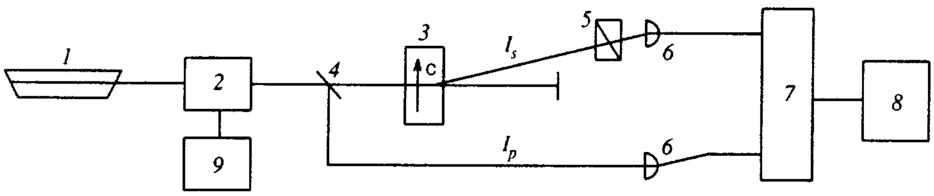

FIG. 1. A sketch of the experimental apparatus: (*1*) argon or helium-cadmium laser; (*2*) electrooptical modulator; (*3*) photorefractive crystal of copper-doped lithium niobate or tantalate (**c** is the optical axis); (*4*) semitransparent mirror; (*5*) analyzer; (*6*) photodetectors; (*7*) CAMAC crate; (*8*) computer; (*9*) electrical pulse generator.

$$G(f)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}b(\tau)e^{2\pi i f\tau}d\tau$$

is the spectrum of the fluctuations and

$$b(\tau)=\langle I(t)I(t+\tau)\rangle/\langle I(t)\rangle^2$$

is the normalized autocorrelation function of the luminous intensity. $K(f)=1$ corresponds to equality of the relative modulation amplitudes of the pump and scattered light intensities, i.e., to ordinary diffraction on a stationary holographic grating. It can be seen clearly in Fig. 2, however, that the depth of modulation for the scattered light intensity is much higher than that for the pump light, which means that the photorefractive holographic grating is able to vary as the pump intensity changes, thereby amplifying the modulation in the scattered radiation.

3. DISCUSSION OF RESULTS

Figure 3 shows the scattered intensity and conversion factor for the modulation depth in the pump intensity as

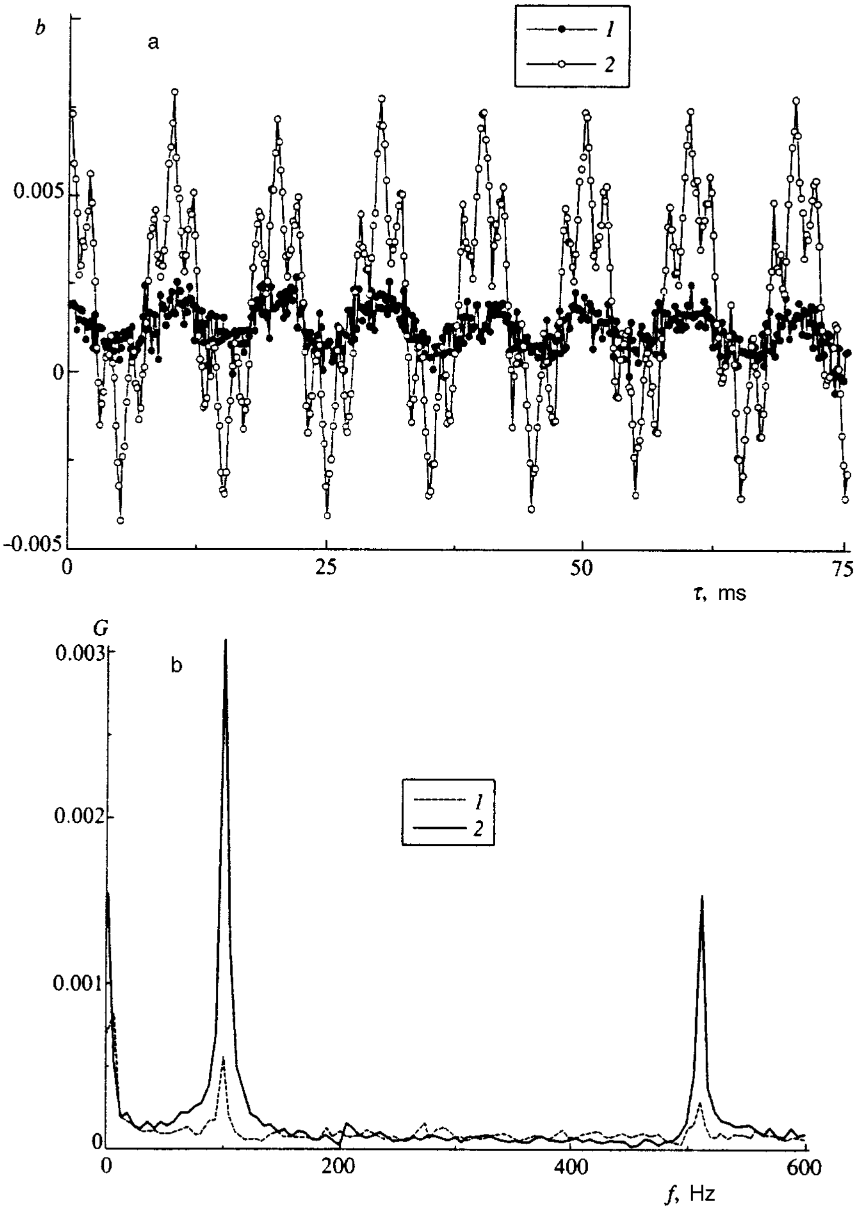

FIG. 2. A typical example of the autocorrelation function (a) and spectrum (b) of the fluctuations in the intensity of a pump modulated at frequencies of 100 Hz and 500 Hz (*1*) and in the intensity of light scattered as a result of $oo-ee$-mode holographic-type parametric scattering in LiTaO$_3$:Cu (*2*).

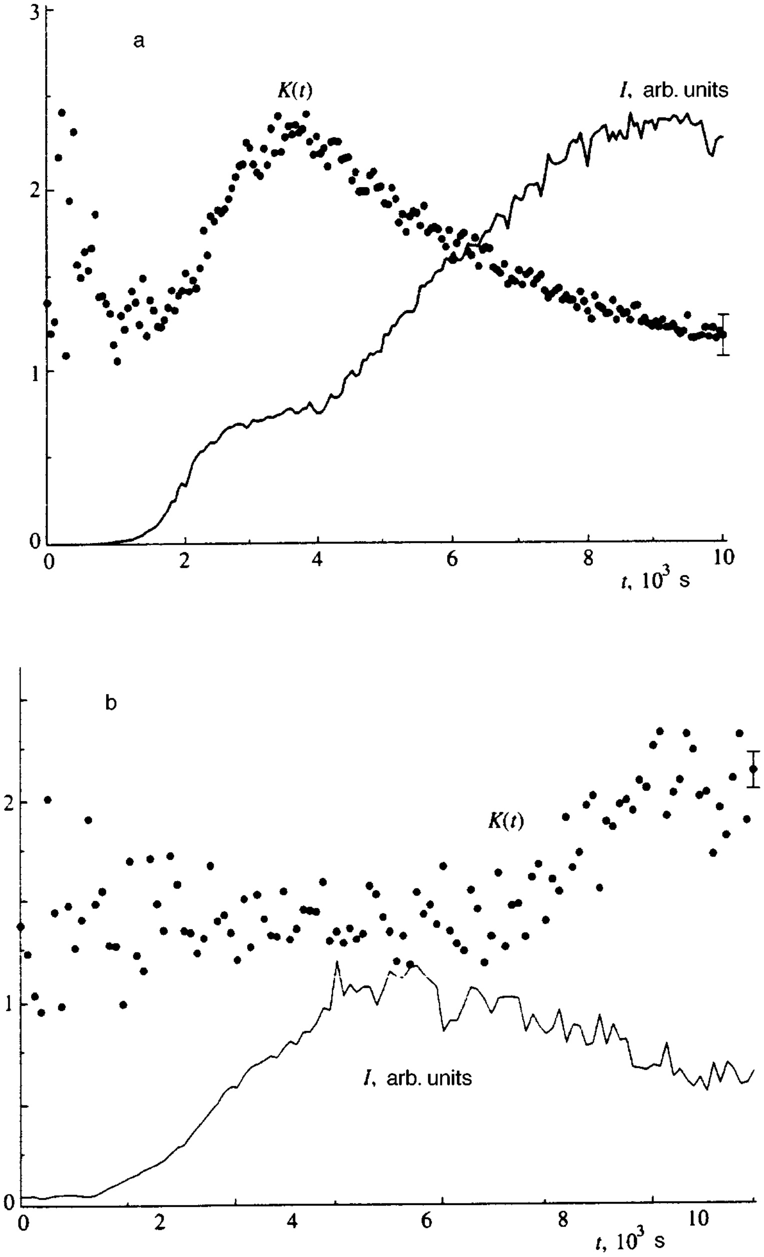

FIG. 3. Time dependence of the scattering intensity and conversion factor for the depth of modulation in the pump intensity at frequencies of $f=300$ Hz (a) and $f=100$ Hz (b) during detection of $ee-oo$-mode holographic-type parametric scattering in LiNbO$_3$:Cu (a) and $oo-ee$-mode holographic-type parametric scattering in LiTaO$_3$:Cu (b).

functions of time during the detection of holographic-type parametric scattering in lithium niobate and tantalate. By the end of the first hour for the lithium tantalate and the second hour for the lithium niobate, the scattering intensity has stopped rising and has attained an approximately constant level, i.e., the recording of the holographic gratings saturates with the effect of the electrostatic field of the separated charges balancing the photogalvanic current. The time over which the process of recording the holographic-type parametric scattering reaches saturation is of the same order as the Maxwell time $\tau_M \sim 10^3$ s. At the same time, however, the holographic grating is able to amplify pump oscillations at frequencies on the order of a hundred hertz, i.e., for a oscillation period, $T \sim 10^{-2}$ s, which is five orders of magnitude shorter than the Maxwell time.

The time dependence of the conversion factor for the depth of modulation is different for lithium niobate and tantalate. In Fig. 3a for LiNbO$_3$:Cu the coefficient $K(t)$ reaches a maximum in the middle of the dynamic regime where the scattering intensity is rising rapidly, while it falls off to unity in the saturation regime. Thus, in lithium niobate the holographic grating is able to react to the changes in the pump intensity only during the period when the grating is increasing even without this change, while in the saturation regime, as might be expected, it becomes stable. The $K(t)$ curve is substantially different for lithium tantalate, as can be seen in Fig. 3b. It has no distinct maxima, as in Fig. 3a; the conversion factor differs little from $K \simeq 1.4$ throughout the dynamic regime. Then in the saturation regime the conversion factor not only fails to drop to unity, but increases to $K \simeq 2$. Thus, even in the stationary state, a photoinduced holographic grating in lithium tantalate is able to react to changes in the pumping over times much shorter than the Maxwell time.

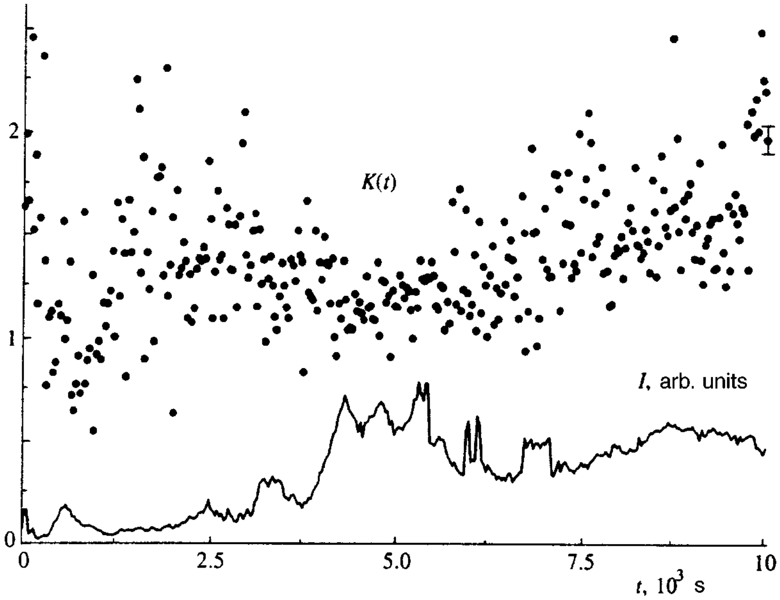

FIG. 4. Time dependence of the scattering intensity $I(t)$ and conversion factor for the depth of modulation of the pump, $K(t)$ at a frequency of $f=100$ Hz during detection of $o-e$-mode photoinduced light scattering in $LiTaO_3$:Cu.

Figure 4 shows the time dependence of the scattering intensity and conversion factor of the oscillations for photoinduced light scattering in lithium tantalate. As can be seen from this graph, it is meaningless to introduce the concepts of dynamic and saturation regimes for photoinduced light scattering, since the scattering intensity experiences rapid drops throughout the entire observation period. This is caused, first of all, by the competition between photoinduced light scattering and holographic-type parametric scattering which causes a large part of the pump intensity to be removed by holographic-type parametric scattering and, second, by the possibility of electrical breakdowns in the crystal, and, finally, by the fact that energy exchange between two interacting waves is forbidden under stationary conditions.[11] Evidently, for the same reasons the conversion factor for the depth of the intensity modulation has a large scatter, although on the average, as for holographic-type parametric scattering, it is undoubtedly greater than unity.

Let us finally examine Fig. 5 which shows the conversion factor for the modulation depth of the pump intensity for two frequencies (not multiples) at once, $f_1=100$ Hz and $f_2=430$ Hz, during holographic-type parametric scattering in $LiTaO_3$:Cu. The picture is much more complicated for two modulation frequencies: although the scattering intensity varies significantly only within the first hour after the onset of recording, as in Fig. 3b, the conversion coefficients undergo rapid changes for eight hours and only after that do they approach a roughly constant level of $K_1 \simeq K_2 \simeq 1.5$. Evidently, the amplifications of the intensity oscillations at the two frequencies are interrelated and cannot be treated independently. Otherwise it would be difficult to explain the fact that a time almost an order of magnitude longer than the time to record the gratings is necessary to attain a stationary situation, not to mention that twice during the observation period there was a time when one of the conversion factors had a local maximum while the other was less than unity (i.e., the

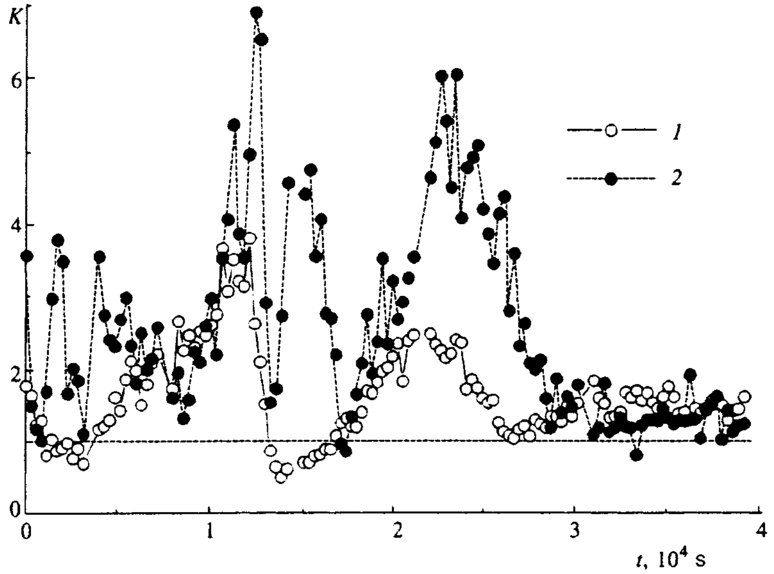

FIG. 5. Time dependence of the conversion factor for the depth of modulation of the pump intensity, $K(t)$ at frequencies of $f_1=100$ Hz (*1*) and $f_2=430$ Hz (*2*) during detection of $oo-ee$-mode holographic-type parametric scattering in $LiTaO_3$:Cu.

relative amplitude of the oscillations decreased), an effect observed in the preceding plots.

In completing our discussion of the experimental data, we must note that the results were completely irreproducible during this study of the temporal dependence of the conversion factor for the intensity modulation depth: two curves obtained under the same conditions can have substantially different features aside from the fact that the gain coefficient is almost always greater than unity. This suggests a stochastic mechanism for amplification of the intensity oscillations at these frequencies.

In sum, our results yield the following conclusions:

1. Amplification of oscillations (or fluctuations) in the pump intensity during scattering in photorefractive lithium tantalate and niobate has been observed at frequencies from 100 Hz to 1 kHz. We were unable to detect amplification of the oscillations reliably at higher frequencies.

2. Amplification of the oscillations is observed in lithium tantalate under any conditions, while in lithium niobate it occurs only during the dynamic regime, when there is a rapid increase in the holographic gratings, i.e., in the nonstationary case.

3. The amplification of oscillations during photoinduced light scattering and holographic-type parametric scattering indicates that the holographic gratings recorded in a crystal vary with changes in the pump intensity. As a result, the oscillations in the scattered radiation are determined by a combination of the oscillations in the pump intensity and of the amplitude of the holographic grating.

4. The amplification of oscillations with a period much shorter than the Maxwell time of the medium indicates the existence of a new, as yet unknown mechanism for photorefraction with a substantially shorter characteristic time.

5. The mechanism for the amplification of the oscillations appears to be stochastic in nature and leads to strong fluctuations in the gain coefficient, a lack of reproducibility in the results, and, perhaps, interactions among the amplification processes for oscillations at different frequencies.

We now consider some possible explanations for the observed effect based on earlier work.[4,7] The existence of a photorefractive response in barium titanate over times on the order of tens of picoseconds has been observed by recording holograms with picosecond pulses,[4] while at least a nanosecond is required to redistribute the electron density in a crystal of this type. This rapid photorefractive response was explained by the fact that electrons photoexcited in the conduction band of the crystal change its refractive index owing to Drude transfer and, thereby, record a grating based on free carriers.[4] This grating spreads out owing to diffusion of electrons in the conduction band, i.e., its lifetime is bounded above by the Maxwell time. The idea of the existence of a free carrier grating can be used to explain the amplification of pump fluctuations. In fact, a change in the carrier concentration in the conduction band takes place essentially immediately (over times on the order of picoseconds), i.e., a free carrier grating is capable of reacting immediately to a change in the intensity of a pump at frequencies on the order of hundreds of hertz.

Another possible explanation of the observed effect is based on data obtained in Ref. 7. There a photo-Hall effect was observed which involved the appearance of a photo-Hall current in a $LiNbO_3$:Fe crystal perpendicular to the magnetic field vector and to a photogalvanic current that is proportional to the intensity of illumination of the crystal. The mobility of the carriers forming this current was determined from the magnitude of the photo-Hall current and turned out to be three orders of magnitude higher than the ordinary carrier mobility in lithium niobate. It was proposed that the measured mobility corresponds to nonthermal electrons that have not yet been clothed in a polaron ‘‘coat’’ and, therefore, have a much lower effective mass and a higher mobility. Since the mobility of a medium is proportional to the mobility of its carriers, while the concentration of nonthermal electrons may increase under nonstationary conditions (for example, in the case of fluctuations in the pump intensity), the Maxwell time (1) corresponding to the nonthermal carriers may be much smaller than usual, which would also explain the amplification of the light intensity fluctuations during scattering at acoustic frequencies.

4. CONCLUSION

We have observed a response in photorefractive lithium tantalate and niobate over times five orders of magnitude shorter than the Maxwell time that leads to nonlinear conversion of fluctuations in the pump intensity. We have proposed two possible ways of explaining the observed effect which, in principle, are not mutually exclusive. The existing phenomenological theory of photorefraction, which successfully describes the majority of photorefractive processes (for example, the recording of holograms, the development of holographic-type parametric scattering), does not account for the existence of free electrons on gratings or of nonthermal carriers in photorefractive crystals. Thus, in order to clarify the true causes of the amplification of pump intensity oscillations, a more rigorous theory of photorefraction must be developed which includes the coupling of various processes which have very different characteristic times in photorefractive crystals: from the photoexcitation times for carriers in the conduction band (times on the order of picoseconds) to the Maxwell time for redistribution of the electron density in a crystal.

The authors thank V. V. Laptinskaya for many fruitful discussions. This work was supported by the Russian Fund for Fundamental Research (Grant No. 96-02-1633a).

[1] A. Ashkin, G. D. Boyd, J. M. Dziedzic, R. G. Smith, A. A. Ballman, J. J. Leinstein, and K. Nassau, Appl. Phys. Lett. **9**, 72 (1966).

[2] F. S. Chen, J. T. La Macchia, and D. V. Fraser, Appl. Phys. Lett. **13**, 223 (1968).

[3] R. Sommerfeldt, R. A. Rupp, H. Vormann, and E. Kratzig, Phys. Status Solidi A **99**, K15 (1987).

[4] A. L. Smirl, G. C. Valley, R. A. Mullen, K. Bokznert, C. D. Mire, and T. F. Boggess, Opt. Lett. **12**, 501 (1987).

[5] A. M. Glass, D. von der Lande, and T. J. Negran, Appl. Phys. Lett. **25**, 233 (1974).

[6] V. A. Pashkov, N. M. Solov'ev, and E. M. Uyukin, Fiz. Tverd. Tela **21**, 1879 (1979) [Sov. Phys. Solid State **21**, 1079 (1979)].

[7] A. I. Levanyuk, A. R. Pogosyan, and E. M. Uyukin, Dokl. Akad. Nauk **256**, 60 (1981) [Sov. Phys. Dokl. **26**, 43 (1981)].

[8] Y. Zhang, S. Campbell, P. Yeh, D. Shen, X. Ma, and J. Chen, Opt. Lett. **19**, 1397 (1994).
[9] V. V. Obukhovskiĭ, A. V. Stoyanov, and V. V. Demeshko, Kvant. Élektron. **14**, 113 (1987) [Sov. J. Quantum Electron. **17**, 64 (1987)].
[10] I. H. Kiseleva, V. V. Obukhovskiĭ, and S. G. Odulov, Fiz. Tverd. Tela **28**, 2975 (1986) [Sov. Phys. Solid State **28**, 1673 (1986)].
[11] V. P. Kondileno, V. B. Markov, S. G. Odulov, and M. S. Soskin, Ukr. Fiz. Zh. **23**, 2039 (1978).

Translated by D. H. McNeill

Structure of the superconducting state of superconductors near the critical field H_{c2} for values of the Ginzburg–Landau parameter κ close to unity

Yu. N. Ovchinnikov

L. D. Landau Institute of Theoretical Physics, Russian Academy of Sciences, 142432 Moscow, Russia

(Submitted 24 January 1997)
Zh. Éksp. Teor. Fiz. **112**, 1499–1509 (October 1997)

It is shown that near the transition temperature T_c the coefficients of the second and third terms in the expansion of the free energy in powers of $H_{c2}-B$ (B is the magnetic field induction inside the superconductor) go to zero simultaneously for a value of $\kappa=1$ for the Ginzburg–Landau parameter. Thereby the structure of the mixed state near H_{c2} for a value of the parameter κ close to unity is determined by the temperature correction to the coefficient for the third power and the coefficient for the fourth power in the expansion of the free energy in powers of $H_{c2}-B$. The values of these coefficients depend on the type of vortex lattice. [S1063-7761(97)02510-9]

1. INTRODUCTION

The mixed state of superconductors was investigated by Abrikosov.[1] He found the magnitude of the magnetic field inside the superconductor $H(\mathbf{r})$ and, in particular, the magnetic induction $B=\langle H(\mathbf{r})\rangle$ from the equation for the vector potential $\mathbf{A}(\mathbf{r})$ (see also Refs. 2–4). Actually, however, the induction value $\mathbf{B}$ cannot be found from equations for the order parameter Δ and the vector potential $\mathbf{A}(\mathbf{r})$, since in the thermodynamic limit this system of equations has a vast number of solutions of order eHR^2 (R is the characteristic transverse dimension of the superconductor). The induction value B is determined from the condition of minimum free energy with respect to the area of an elementary cell (or the induction B) for a given value of the external magnetic field H_0.[5] Physically, such an ambiguity is associated with the possibility of surface currents, which produce a total magnetic moment of the same order as the internal vortex currents.

We show below that near the transition temperature T_c the coefficients for the second and third terms of the expansion of the free energy in powers of $H_{c2}-B$ vanish simultaneously for a Ginzburg–Landau parameter $\kappa=1$. Consequently, there are three small parameters: $1-T/T_c$, κ^2-1 and $1-B/H_{c2}$. As a result, the structure of the mixed state depends on the temperature correction $\sim(1-T/T_c)$ to the coefficient of the third power and the value of the coefficient of the fourth power of the expansion of the free energy in powers of $H_{c2}-B$. Thus, the equation for B is a third-degree polynomial. In the general case three roots exist for it. It must be expected that the coefficient of the $(H_{c2}-B)^4$ term in the expansion of the free energy in the powers of $H_{c2}-B$ is positive. In this case the free energy as a function of B has two local minima and one maximum. Let us investigate three types of lattices: triangular with one and two flux quanta in a cell and square with one flux quantum in a cell. The two coefficients go to zero simultaneously for $\kappa=1$ for all three types of lattices.

2. SOLUTION OF THE GINZBURG–LANDAU EQUATION NEAR CRITICAL FIELD H_{c2}

Near the transition temperature T_c the free energy F_S of a superconductor in an external magnetic field H_0 can be represented by[6]

$$F_S-F_N=\nu\int d^3\mathbf{r}\left\{-\left(1-\frac{T}{T_c}\right)|\Delta|^2+\frac{\pi D}{8T_c}|\partial_-\Delta|^2+\frac{7\zeta(3)}{16\pi^2T_c^2}|\Delta|^4\right\}+\frac{1}{8\pi}\int d^3\mathbf{r}(\mathrm{curl}^2\,\mathbf{A}-2H_0\,\mathrm{curl}\,\mathbf{A}+\mathbf{H}_0^2), \tag{1}$$

where $\mathbf{A}$ is the vector potential, $\partial_-=\partial/\partial\mathbf{r}-2ie\mathbf{A}$, and $\nu=mp_0/2\pi^2$ is the density of states at the Fermi surface. The coefficient D depends on the transport mean free path l_{tr} of the electrons and is equal to

$$D=D_{\mathrm{dif}}\eta;\quad D_{\mathrm{dif}}=\frac{vl_{\mathrm{tr}}}{3};$$
$$\eta=1-\frac{8T\tau_{\mathrm{tr}}}{\pi}\left(\psi\left(\frac{1}{2}+\frac{1}{4\pi T\tau_{\mathrm{tr}}}\right)-\psi\left(\frac{1}{2}\right)\right), \tag{2}$$
$$l_{\mathrm{tr}}=v\tau_{\mathrm{tr}}.$$

In Eq. (2) v is the electron velocity at the Fermi surface and $\psi(x)$ is the Euler ψ function. We will show below that for a complete investigation of the states of a superconductor for values of the Ginzburg–Landau parameter κ close to unity it is necessary to supplement Eq. (1) for the free energy by terms of next order in $1-T/T_c$. This problem is beyond the scope of this paper and will not be considered here.

By varying the free energy (1) with respect to Δ and $\mathbf{A}$, we obtain the Ginzburg–Landau system of equations

$$\frac{\pi D}{8T_c}\partial_-^2\Delta+\left(1-\frac{T}{T_c}\right)\Delta-\frac{7\zeta(3)}{8\pi^2T_c^2}|\Delta|^2\Delta=0,$$
$$\frac{1}{4\pi}\,\mathrm{curl}\,\mathrm{curl}\,\mathbf{A}=\mathbf{j};\quad j=\frac{i\pi e\nu D}{4T_c}(\Delta\partial_+\Delta^*-\Delta^*\partial_-\Delta). \tag{3}$$

As shown in Ref. 6, at the point H in terms of the magnetic field

$$H = H_{c2} = \frac{4T_c}{\pi e D}\left(1 - \frac{T}{T_c}\right) \tag{4}$$

the linearized system of Eqs, (3) has solutions of the form

$$\Delta = \exp(2ieHx_1 y - eH(x - x_1)^2) \tag{5}$$

for an arbitrary value of the parameter x_1. The calibration is used in obtaining Eq. (5).

$$\mathbf{A} = (0;Hx;0). \tag{6}$$

In the external magnetic field $H_0 < H_{c2}$ we will look for solutions of the system of Eqs. (3) such that all physical quantities $|\Delta|^2$, $H(x,y)$, j are periodic functions of position. We assume that $\mathbf{a}_{1,2}$ are the vectors of the elementary cell, i.e.,

$$|\Delta(\mathbf{r} + N\mathbf{a}_1 + M\mathbf{a}_2)|^2 = |\Delta(\mathbf{r})|^2. \tag{7}$$

From the periodicity condition of the current density and $|\Delta|^2$ we find

$$\oint_\Gamma \left(\frac{\partial \chi}{\partial \mathbf{r}} - 2e\mathbf{A}\right) d\mathbf{l} = 0. \tag{8}$$

where Γ is a closed contour running along the edge of the elementary cell and χ is the phase of the order parameter. Since the order parameter is a single-valued function of the coordinates, Eq. (7) leads to the quantization condition of the magnetic flux ϕ with respect to the elementary cell

$$\phi = \frac{\pi}{e} N, \quad N = 1, 2... \tag{9}$$

This exact relation greatly simplifies the search for solutions of the system of equations (3).

In a magnetic field H_0 close to H_{c2} we look for the solution of the system of equations (3) in the form of a series in terms of powers of $H_{c2} - B$

$$\mathbf{A} = (0,Bx,0) + \mathbf{A}_1 + \mathbf{A}_2 + \ldots, \quad \Delta = \Delta_0 + \Delta_1 + \Delta_2 + \ldots,$$
$$\Delta_0 = \sum_N C_N \exp(2ieBNx_1 y - eB(x - Nx_1)^2), \tag{10}$$

where B is the induction of the magnetic field inside the superconductor $(B = \langle H(\mathbf{r})\rangle)$, and the vectors $\mathbf{A}_k$ have two nonzero components (1,2) and are proportional to $(H_{c2} - B)^k$. The physical quantities satisfy $|\Delta_k|^2 \sim (H_{c2} - B)^{2k+1}$, $C_N^2 \sim H_{c2} - B$. Below we will use the calibration div $\mathbf{A} = 0$. In this calibration all quantities $\mathbf{A}_k$ are periodic functions of the coordinates. In the calibration used in Ref. 1, the quantities A_k are increasing functions of x; this leads to additional difficulties in solving the system of Eqs. (3).

It follows from Eqs. (1) and (10) that the free energy density $(F_S - F_N)/V$ can be represented in the form of a series in terms of powers of $H_{c2} - B$:

$$\frac{F_S - F_N}{V} = \frac{1}{8\pi}\left\{(B - H_0)^2 + P_1(H_{c2} - B)^2 + \frac{P_2}{H_{c2}} \times (H_{c2} - B)^3 + \frac{P_3}{H_{c2}^2}(H_{c2} - B)^4 + \ldots\right\}. \tag{11}$$

The coefficients P_k in Eq. (11) are determined by the type of vortex lattice and by the value of the Ginzburg–Landau parameter κ.

The induction B is found from the condition of minimum free energy (1), (11) with respect to B for a given value H_0 of the external magnetic field,[5]

$$\frac{\partial(F_S - F_N)}{\partial B} = 0. \tag{12}$$

Our task is to find the coefficients P_1 and P_2 in the expansion (11). To do this it is necessary to find the quantities Δ_0, Δ_1 and $\mathbf{A}_1$, $\mathbf{A}_2$.

The periodicity conditions (7) impose a stringent restriction on the values of the coefficients C_N (Eq. (10)). For different types of lattices we find

$$C_N = C_0 \exp\left(-\frac{i\pi}{2}N^2\right), \tag{13a}$$

a triangular lattice with one flux quantum;

$$C_{N+1} = C_N = C_0, \tag{13b}$$

a square lattice with one flux quantum;

$$C_{N+1} = C_N = C_0, \tag{13c}$$

a triangular lattice with two flux quanta. The coefficients C_0 can be assumed real in all cases.

To construct the functions $\Delta_1, \Delta_2, \ldots$ we need a basis derived from the eigenfunctions of the operator $\hat{L}$

$$\hat{L} = -\left[\frac{\partial^2}{\partial x^2} + \left(\frac{\partial}{\partial y} - 2ieBx\right)^2\right], \quad \hat{L} f_n = \lambda_n f_n. \tag{14}$$

The operator $\hat{L}$ has been thoroughly investigated. All of its eigenvalues λ_n are given by

$$\lambda_n = 4eB(n + 1/2), \quad n = 0, 1, 2... \tag{15}$$

Each eigenvalue is infinitely degenerate

$$f_n \equiv f_n(x_1, \mathbf{r}) = \exp(2ieBx_1 y) D_n(2\sqrt{eB}\,(x - x_1)). \tag{16}$$

where x_1 is an arbitrary real number and D_n is the parabolic cylinder function:[7]

$$D_n(z) = 2^{-n/2} \exp(-z^2/4) H_n(z/\sqrt{2}). \tag{17}$$

In Eq. (17) H_n is the Hermitian polynomial. The correction Δ_1 to the order parameter can be represented by

$$\Delta_1 = \sum_{M=1}^{\infty} \sum_{N=-\infty}^{\infty} Q_N^M \exp(2ieBNx_1 y) \times D_M[2\sqrt{eB}(x - Nx_1)]. \tag{18}$$

From the periodicity conditions (7) for any type of lattice we find

$$Q_N^M = \alpha_M C_N. \tag{19}$$

where α_M is a complex number not depending on M. Below, for convenience, we assume

$$\Delta_1^M=\sum_{N=-\infty}^{\infty} C_N \exp(2ieBNx_1y)D_M[2\sqrt{eB}(x-Nx_1)],\qquad M=1,\ 2\ldots \tag{20}$$

From Eq. (20) the expression for the correction Δ_1 to the order parameter becomes

$$\Delta_1=\sum_{M=1}^{\infty}\alpha_M\Delta_1^M. \tag{21}$$

In the approximation being considered the system of equations (3) becomes

$$\begin{gathered}\frac{\pi eD}{4T_c}H_{c2}(\Delta_0+\Delta_1)-\frac{\pi D}{4T_c}\left[eB\Delta_0+2eB\sum_{M=1}^{\infty}\left(M+\frac{1}{2}\right)\right.\\ \left.\times\alpha_M\Delta_1^M+2ieA_1\partial_-\Delta_0\right]-\frac{7\zeta(3)}{8\pi^2T_c^2}|\Delta_0|^2\Delta_0=0,\\ \operatorname{curl}\operatorname{curl}(\mathbf{A}_1+\mathbf{A}_2)=4\pi\mathbf{j}.\\ j=-\frac{\pi e\nu D}{4T_c}\left(\frac{\partial}{\partial y};-\frac{\partial}{\partial x}\right)|\Delta_0|^2+\frac{i\pi e\nu D}{4T_c}[4ieA_1|\Delta_0|^2\\ +\Delta_1\partial_+\Delta_0^*+\Delta_0\partial_+\Delta_1^*-\Delta_0^*\partial_-\Delta_1-\Delta_1^*\partial_-\Delta_0],\end{gathered} \tag{22}$$

where

$$\partial_\pm=\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y}\pm 2ieBx\right). \tag{23}$$

From the system (22) with the orthogonality of the functions Δ_0 and Δ_1 taken into account we find

$$\begin{gathered}H_1(r)=-\frac{\pi^2e\nu D}{T_c}(|\Delta_0|^2-\langle|\Delta_0|^2\rangle),\qquad \mathbf{H}_1=\operatorname{curl}\mathbf{A}_1,\\ \frac{\pi eD}{4T_c}(H_{c2}-B)\langle|\Delta_0|^2\rangle+\frac{1}{4\pi\nu}\langle H_1^2(r)\rangle\\ -\frac{7\zeta(3)}{8\pi^2T_c^2}\langle|\Delta_0|^4\rangle=0,\\ \frac{\pi eBDM}{2T_c}\alpha_M\langle\Delta_1^M\Delta_1^{M*}\rangle\\ =-\frac{i\pi eD}{4T_c}\langle A_1(\Delta_1^{M*}\partial_-\Delta_0-\Delta_0\partial_+\Delta_1^{M*})\rangle\\ -\frac{7\zeta(3)}{8\pi^2T_c^2}\langle|\Delta_0|^2\Delta_0\Delta_1^{M*}\rangle,\\ \left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)\mathbf{A}_1=\frac{\pi^2e\nu D}{T_c}\left(\frac{\partial}{\partial y};-\frac{\partial}{\partial x}\right)|\Delta_0|^2.\end{gathered} \tag{24}$$

Equation (1) for the free energy density $(F_S-F_N)/V$ is reduced to a rather simple form by means of Eqs. (10), (20), (21), (22), (24):

$$\begin{gathered}\frac{F_S-F_N}{V}=\frac{1}{8\pi}(B-H_0)^2+\frac{7\zeta(3)\nu}{16\pi^2T_c^2}\left(\langle|\Delta_0|^4\rangle\left(1-\frac{1}{\kappa^2}\right)\right.\\ \left.+\frac{\langle|\Delta_0|^2\rangle^2}{\kappa^2}\right)-\frac{\pi e\nu D}{4T_c}(H_{c2}-B)\langle|\Delta_0|^2\rangle\\ +\frac{\pi e^2\nu D}{2T_c}\langle A_1^2|\Delta_0|^2\rangle\\ -\frac{\pi eB}{2T_c}\sum_{M=1}^{\infty}M|\alpha_M|^2\langle\Delta_1^M\Delta_1^{M*}\rangle.\end{gathered} \tag{25}$$

From the first two of Eqs. (24) we find

$$\langle|\Delta_0|^2\rangle=\frac{2\pi^3eDT_c}{7\zeta(3)}\frac{H_{c2}-B}{\beta-(\beta-1)/\kappa^2}, \tag{26}$$

where

$$\beta=\frac{\langle|\Delta_0|^4\rangle}{\langle|\Delta_0|^2\rangle^2},\qquad \kappa=\frac{1}{\pi^2eD}\left(\frac{7\zeta(3)}{2\pi\nu}\right)^{1/2}. \tag{27}$$

For further calculations it is convenient to convert to a Fourier series expansion of the functions $|\Delta_0|^2$, $\mathbf{A}_1,\ldots$ In an isotropic superconductor it must be expected that $|\mathbf{a}_1|=|\mathbf{a}_2|$ and the angle between the vectors $\mathbf{a}_{1,2}$ is equal to $\pi/3$ or $\pi/2$. We assume that $\mathbf{K}_{1,2}$ are the elementary vectors of the reciprocal lattice. The order parameter $|\Delta_0|^2$ in this case can be represented by

$$|\Delta_0|^2=\sum_{N,M=-\infty}^{\infty}C_{NM}\exp(i(N\mathbf{K}_1+M\mathbf{K}_2)\mathbf{r}). \tag{28}$$

Further calculations depend on the type of lattice.

For a triangular lattice with one flux quantum in a cell we find

$$\begin{gathered}\mathbf{K}_1=\frac{2\pi}{\sqrt{3}x_1}(0,1),\qquad \mathbf{K}_2=\frac{\pi}{\sqrt{3}x_1}\left(\sqrt{3};-1\right),\\ 2\sqrt{3}Bx_1^2=\pi,\qquad C_N=C_0\exp\left(-\frac{i\pi}{2}N^2\right),\\ C_{NM}=C_0^2 3^{1/4}\exp\left(-i\pi NM-\frac{\pi}{\sqrt{3}}(N^2+M^2-NM)\right).\end{gathered} \tag{29}$$

For a square lattice with one flux quantum in a cell we obtain

$$\begin{gathered}\mathbf{K}_1=\frac{2\pi}{x_1}(0,1),\qquad \mathbf{K}_2=\frac{2\pi}{x_1}(1,0),\\ eBx_1^2=\pi,\qquad C_{N+1}=C_N=C_0,\\ C_{NM}=\frac{C_0^2}{\sqrt{2}}(-1)^{NM}\exp\left(-\frac{\pi}{2}(N^2+M^2)\right).\end{gathered} \tag{30}$$

Let us also give an expression for the quantities that characterize a triangular lattice with two flux quanta in a cell

$$\begin{gathered}\mathbf{K}_1=\frac{2\pi}{\sqrt{3}x_1}(0,1),\qquad \mathbf{K}_2=\frac{\pi}{\sqrt{3}x_1}(\sqrt{3},-1),\\ \sqrt{3}Bx_1^2=\pi,\qquad C_{N+1}=C_N=C_0,\qquad C_{N,2K+1}=0,\end{gathered} \tag{31}$$

$$C_{N,2K}=\frac{C_0^2 3^{1/4}}{\sqrt{2}}\exp\left(i\pi K(N-K)-\frac{\pi}{2\sqrt{3}}\times(N^2+4K^2-2NK)\right).$$

By using Eq. (20) for the explicit form of the functions Δ_1^M, we find that

$$\langle\Delta_1^M\Delta_1^{M*}\rangle=\langle|\Delta_0|^2\rangle M! \tag{32}$$

To calculate the quantity α_M we need the following relation:

$$\Delta_1^M\partial_+\Delta_0^*-\Delta_0^*\partial_-\Delta_1^M =i\left(\frac{\partial}{\partial y};-\frac{\partial}{\partial x}\right)|\Delta_0|^2-2\sqrt{eB}M\Delta_0^*\Delta_1^{M-1}(1;-i). \tag{33}$$

The explicit form of the vector potential $\mathbf{A}_1$ is easily found from Eqs. (24) and (28):

$$\mathbf{A}_1=-\frac{i\pi^2 e\nu D}{T_c}\sum_{\mathbf{K}_{NM}\neq 0}C_{NM}((\mathbf{K}_{NM})_y;-(\mathbf{K}_{NM})_x)\exp(i\mathbf{K}_{NM}\mathbf{r})/\mathbf{K}_{NM}^2. \tag{34}$$

We now represent the function $\Delta_0^*\Delta_1^M$ in the form of a Fourier series

$$\Delta_0^*\Delta_1^M=\sum_{N_1K}C_{NK}^M\exp(i\mathbf{K}_{NM}\mathbf{r}). \tag{35}$$

The coefficients C_{NK}^M depend on the type of lattice:

$$C_{NK}^M=C_{NK}(-1)^M\begin{cases}\left(\frac{2\pi}{\sqrt{3}}\right)^{M/2}\left(N-\frac{K}{2}+\frac{i\sqrt{3}}{2}K\right)^M, \\ \text{triangular, one flux quantum,} \\ \pi^{M/2}(N+iK)^M, \\ \text{square, one flux quantum}\end{cases}$$

triangular, two flux quanta:

$$\begin{aligned}&C_{N,2K+1}^M=0,\\ &C_{N,2K}^M=C_{N,2K}(-1)^M\left(\frac{\pi}{\sqrt{3}}\right)^{M/2}(N-K+i\sqrt{3}K)^M.\end{aligned} \tag{36}$$

By using Eqs. (34)–(36), we find the following expression for the correlators $i\langle\mathbf{A}_1(1;-i)\Delta_0^*\Delta_1^{M-1}\rangle$, which determine the value of the coefficients α_M:

$$i\langle\mathbf{A}_1(1;-i)\Delta_0^*\Delta_1^{M-1}\rangle =\frac{\pi\sqrt{e}\nu D}{2T_c\sqrt{B}}\sum_{N,K}C_{NK}C_{NK}^M\begin{cases}\frac{\sqrt{3}}{2}(N^2+K^2-NK)^{-1}\\ (N^2+K^2)^{-1}\\ \sqrt{3}(N^2+K^2-NK)^{-1}.\end{cases} \tag{37}$$

In Eq. (37) the top line to the right of the brace corresponds to a triangular lattice with one flux quantum, the middle line to a square lattice with one flux quantum, and the bottom line to a triangular lattice with two flux quanta.

We find the following expression for the coefficients α_M from Eqs. (24), (26), (35)–(37):

$$\alpha_M=-\frac{(-1)^M}{2M\cdot M!}\frac{H_{c2}-B}{B(\beta-(\beta-1)/\kappa^2)}D_M; \tag{38}$$

where the coefficients D_M depend on the type of lattice:

$$D_M=\left(\frac{2\pi}{\sqrt{3}}\right)^{M/2}\sum_{N^2+K^2\neq 0}\left(\left(1-\frac{1}{\kappa^2}\right)+\frac{M\sqrt{3}}{2\pi\kappa^2}\frac{1}{N^2+K^2-NK}\right)\left(N-\frac{K}{2}-\frac{i\sqrt{3}K}{2}\right)^M\times\exp\left(-\frac{2\pi(N^2+K^2-NK)}{\sqrt{3}}\right),$$

$$D_M=\pi^{M/2}\sum_{N^2+K^2\neq 0}\left(\left(1-\frac{1}{\kappa^2}\right)+\frac{M}{\pi\kappa^2}\frac{1}{N^2+K^2}\right)\times(N-iK)^M\exp(-\pi(N^2+K^2)), \tag{39}$$

$$D_M=\left(\frac{\pi}{\sqrt{3}}\right)^{M/2}\sum_{N^2+4K^2-2NK\neq 0}\left(\left(1-\frac{1}{\kappa^2}\right)+\frac{M\sqrt{3}}{\pi\kappa^2}\frac{1}{N^2+4K^2-2NK}\right)(N-K-i\sqrt{3}K)^M\times\exp\left(-\frac{\pi(N^2+4K^2-2NK)}{\sqrt{3}}\right).$$

The arrangement of the lines in Eq. (39) is the same as in Eq. (37) in terms of lattice types.

Finally, by means of Eqs. (34)–(39) Eq. (25) for the free energy density is reduced to

$$\begin{aligned}\frac{F_S-F_N}{V}=&\frac{1}{8\pi}(B-H_0)^2-\frac{(H_{c2}-B)^2}{8\pi\kappa^2(\beta-(\beta-1)/\kappa^2)}\\ &-\frac{(H_{c2}-B)^3}{8\pi\kappa^2B(\beta-(\beta-1)/\kappa^2)^3}\sum_{M=1}^{\infty}\frac{D_M^2}{MM!}\\ &+\frac{(H_{c2}-B)^3}{8\pi\kappa^6B(\beta-(\beta-1)/\kappa^2)^3}G.\end{aligned} \tag{40}$$

The coefficient G in Eq. (40) depends on the type of lattice, and we give its value for the three types of lattices in the same order as in Eq. (37):

$$G=\frac{1}{\pi}\frac{\sqrt{3}}{2}\sum\frac{NN_1+KK_1-0.5(N_1K+K_1N)}{(N^2+K^2-NK)(N_1^2+K_1^2-N_1K_1)}\times(-1)^{N_1K+NK_1}\exp\left(-\frac{2\pi}{\sqrt{3}}(N^2+K^2+N_1^2+K_1^2-NK-N_1K_1-NN_1-KK_1+0.5(N_1K+K_1N))\right),$$

$$G=\frac{1}{\pi}\sum\frac{NN_1+KK_1}{(N^2+K^2)(N_1^2+K_1^2)}(-1)^{N_1K+NK_1}\times\exp(-\pi(N^2+K^2+N_1^2+K_1^2-NN_1-KK_1)), \tag{41}$$

TABLE I.

Lattice type	Triangular, one flux quantum	Square, one flux quantum	Triangular, two flux quanta
G	$3.661705\cdot 10^{-2}$	$5.145751\cdot 10^{-2}$	0.1803918
γ_0	$3.661705\cdot 10^{-2}$	$5.145751\cdot 10^{-2}$	0.1803918
γ_1	$4.361318\cdot 10^{-2}$	0.0687983	0.2520449
γ_2	$1.299969\cdot 10^{-2}$	$2.338158\cdot 10^{-2}$	$9.405005\cdot 10^{-2}$
β	1.15952	1.18034	1.33897

$$G=\frac{1}{\pi}\sqrt{3}\sum\frac{NN_1+4KK_1-N_1K-K_1N}{(N^2+4K^2-2NK)(N_1^2+4K_1^2-2N_1K_1)}\times(-1)^{K_1N+KN_1}\exp\left(-\frac{\pi}{\sqrt{3}}(N^2+4K^2+N_1^2+4K_1^2-2NK-2N_1K_1-NN_1-4KK_1+N_1K+NK_1)\right).$$

The sums in Eq. (37) are taken over the values N; N_1; K; K_1 such that both factors in the denominator are different from zero.

It follows from Eq. (39) that the dependence on κ^2 of the sum in the third term of Eq. (40) is given by

$$\sum_{M=1}^{\propto}\frac{D_M^2}{MM!}=\frac{\gamma_0}{\kappa^4}+\frac{\gamma_1}{\kappa^2}\left(1-\frac{1}{\kappa^2}\right)+\gamma_2\left(1-\frac{1}{\kappa^2}\right)^2. \qquad (42)$$

A direct calculation shows that for all three types of lattices considered the relation

$$G=\gamma_0 \qquad (43)$$

is satisfied. Equation (43) is obviously exact. The results of a numerical calculation for the coefficients γ_i, β and G are listed in Table I.

A triangular lattice with one flux quantum has a sixth-order symmetry axis, while the square lattice has a fourth-order axis. Symmetry with respect to the replacement $\mathbf{r}\to-\mathbf{r}$ exists in the triangular lattice with two flux quanta. As a result, only the coefficients D_M with $M=K(6;4;2)$, $K=1, 2,...$, are different from zero.

With Eqs. (42) and (43) taken into account, Eq. (40) for the free energy density is reduced to

$$\frac{F_S-F_N}{V}=\frac{1}{8\pi}[(H_{c2}-H_0)^2-2(H_{c2}-H_0)(H_{c2}-B)+(H_{c2}-B)^2]-\frac{(H_{c2}-B)^2}{8\pi\kappa^2(\beta-(\beta-1)/\kappa^2)}-\frac{(H_{c2}-B)^3}{8\pi\kappa^2B(\beta-(\beta-1)/\kappa^2)^3}\left(\frac{\gamma_1}{\kappa^2}\left(1-\frac{1}{\kappa^2}\right)+\gamma_2\left(1-\frac{1}{\kappa^2}\right)^2\right). \qquad (44)$$

We point out once more that in the Ginzburg–Landau approximation (Eq. (1) for the free energy) for a value of $\kappa=1$ the coefficients of $(H_{c2}-B)^2$ and $(H_{c2}-B)^3$ in the power expansion of the free energy $(H_{c2}-B)$ (Eq. (44)) go to zero simultaneously. With the subsequent terms in $(1-T/T_c)$ taken into account this effect vanishes, and at the point when the coefficient for $(H_{c2}-B)^2$ is equal to zero the coefficient for the term $(H_{c2}-B)^3$ will be equal to $\widetilde{C}(1-T/T_c)/8\pi H_{c2}$, $\widetilde{C}\sim 1$. Note also that the coefficients $\gamma_{0,1,2}$ are numerically small for the triangular and square lattices with one flux quantum. This is due to the high symmetry of such lattices. The coefficient for $(H_{c2}-B)^4$ will also obviously be small numerically. There is no basis for expecting the coefficient $\widetilde{C}$ to be small.

Thus, the structure of the mixed state for values of the parameter κ^2 close to unity is determined by a fourth-degree polynomial in $H_{c2}-B$. As a result, the absolute value of the parameter $|\Delta|$ and the induction B change drastically for a small change in the value of the external magnetic field H_0.

A region exists in terms of the parameters $(H_{c2}-H_0)/H_{c2}$, κ^2-1, $1-T/T_c$, in which the free energy as a function of B has three extremal points for a given type of flux lattice. One of these points corresponds to a maximum and the other two to a local minimum.

3. MAGNETIC MOMENT M_V, CREATED BY INTERNAL CURRENTS

The total magnetic moment $\mathbf{M}$ of a unit volume of a superconducting cylinder is defined by the relation

$$\mathbf{B}=\mathbf{H}_0+4\pi\mathbf{M}. \qquad (45)$$

The magnetic moment $\mathbf{M}$ is created by both internal currents and surface

$$\mathbf{M}=\mathbf{M}_V+\mathbf{M}_S. \qquad (46)$$

Both of these contributions, generally speaking, are of the same order.

By means of Eqs. (22), (26), (28) the magnetic moment M_V of a unit volume of a superconductor, created by internal currents, is written as

$$M_V=\frac{1}{2S}\int_S[\mathbf{rj}]_z d^2\mathbf{r}=\frac{\pi e\nu D}{8T_c}\frac{1}{S}\int_S d^2\mathbf{r}\left(x\frac{\partial|\Delta_0|^2}{\partial x}+y\frac{\partial|\Delta_0|^2}{\partial y}\right)=-\frac{\pi e\nu D}{4T_c}\langle|\Delta_0|^2\rangle\widetilde{\alpha}=-\frac{\widetilde{\alpha}(H_{c2}-B)}{4\pi\kappa^2(\beta-(\beta-1)/\kappa^2)}. \qquad (47)$$

The integration in Eq. (47) is carried out over the area S of an elementary cell.

The constant $\widetilde{\alpha}$ depends on the type of lattice:

$$\widetilde{\alpha}=\left(\sum_{M\neq 0}C_{0M}(-1)^{M+1}+\sum_{N\neq 0}C_{N0}(-1)^{N+1}\right)\Big/2C_{00}, \qquad (48)$$

$$\widetilde{\alpha}=\begin{cases}0.324654\text{—triangular lattice, one flux quantum}\\0.412025\text{—square lattice, one flux quantum}\\0.404058\text{—triangular lattice, two flux quanta.}\end{cases}$$

4. CONCLUSION

We have shown that in an external magnetic field H_0 the free energy of a superconductor can be expanded in a series in powers of $H_{c2}-B$, where B is the induction of the magnetic field inside the superconductor. The expected radius of convergence of the series is of the order of H_{c2} and it remains finite as $\kappa \to 1$. Because of the presence of surface currents a vast number of solutions (of the order of eHR^2, where R is the characteristic diameter of the superconducting cylinder) exists for the system of Ginzburg–Landau equations. Some of them correspond to a local minimum and some to saddle points. Surface currents create a bulk magnetic moment of the same order as the internal currents. For a superconducting cylinder with a sufficiently large diameter the surface energy is small compared with the volume and the induction B can be found by minimizing the bulk portion of the free energy with respect to the induction B in a given external magnetic field H_0. In Eq. (11) for the free energy the only place where the external magnetic field H_0 enters is the first term.

We have also shown that in the Ginzburg–Landau approximation a unique situation arises—the coefficients for the square and cubic terms of the expansion of $(F_S - F_N)/V$ in powers of $H_{c2}-B$ go to zero simultaneously at the point $\kappa=1$. Thus, in the vicinity of the point $\kappa=1$ the free energy is a fourth-degree polynomial in $H_{c2}-B$, and the coefficient for the term $(H_{c2}-B)^3$ contains the small parameter $1-T/T_c$. To calculate it it is necessary to go outside the framework of the free energy in the Ginzburg–Landau approximation.

As a result, three extremal points for the free energy as a function of the induction B exist, generally speaking, for a given type of flux lattice. Two of these, are probably, local minima and one is a maximum.

The quantities $\gamma_{0,1,2}$, which arise in a calculation of the cubic term, are numerically small for the triangular and square lattices with one flux quantum. This smallness is the result of the high symmetry of these lattices. The coefficient for the fourth power of $H_{c2}-B$ will also apparently be numerically small. A detailed investigation of the structure of the eddy lattice, arising in the parameter region

$$(\kappa^2-1) \ll 1, \quad (1-H_0/H_{c2}) \ll 1, \quad 1-T/T_c \ll 1.$$

is the subject of a separate investigation.

This work was supported by CRDF (Grant RP1-194) and by a contract with the U. S. Naval Research Lab (N00173-97-P-3488).

[1] A. A. Abrikosov, Zh. Éksp. Teor. Fiz. **32**, 1442 (1957) [Sov. Phys. JETP **5**, 1174 (1957)].
[2] R. D. Parks, ed., *Superconductivity* (Marcell Dekker, New York, 1969).
[3] P. G. de Gennes, *Superconductivity of Metals and Alloys* (W. A. Benjamin, New York-Amsterdam, 1966).
[4] A. A. Abrikosov, *Fundamentals of the Theory of Metals* (North-Holland Elsevier, Groningen, 1988; Nauka, Moscow, 1987).
[5] Yu. N. Ovchinnikov, JETP Lett. **65**, 632 (1997).
[6] V. L. Ginzburg and L. D. Landau, Zh. Eksp. Teor. Fiz. **20**, 1064 (1950).
[7] I. S. Gradshteyn and I. M. Ryzhik, eds., *Tables of Integrals, Series, and Products* (Academic Press, New York, 1965; Fizmatgiz, Moscow, 1962).

Translated by Eugene R. Heath

Exact soliton-like solutions in generalized dynamical models of a quasi-one-dimensional crystal

O. V. Gendel'man and L. I. Manevich

N. N. Semyonov Institute of Chemical Physics, Russian Academy of Sciences, 117977 Moscow, Russia

(Submitted 18 February 1997; resubmitted 13 May 1997)
Zh. Éksp. Teor. Fiz. **112**, 1510–1515 (October 1997)

We examine the dynamics of a one-dimensional crystal on a substrate with allowance for both nonlinear interaction with the substrate and intramolecular dispersion and nonlinearity. We show that such models, with the potential chosen appropriately, have soliton-like solutions with a distinct velocity. Such a potential allows for an independent selection of the curvature of the potential well and the height of the potential barrier, and in this way ''globally'' models the real potential. We show that such global modeling is correct to a first approximation. © *1997 American Institute of Physics.* [S1063-7761(97)02610-3]

1. INTRODUCTION

In certain areas of solid-state physics, a consistent description of the nonlinear dynamics of the phenomena leads to equations that contain, in addition to gradient terms representing nonlinearity and dispersion (i.e., terms that depend on the derivatives of the desired displacement field), terms that depend only on the field proper. Such equations may contain, e.g., a combination of terms that are characteristic, on the one hand, of Korteweg–de Vries and modified Korteweg–de Vries models and, on the other, of sine–Gordon, ϕ^4, and similar models.

The structure of these equations takes them outside the class of well-known models, including integrable models. Furthermore, generally speaking, even the search for solutions of the soliton type leads in these cases to an analytically insoluble problem for an ordinary fourth-order nonlinear equation. The simplest example of systems of this type is a one-dimensional crystal on a substrate.

Hence it comes as no surprise that the usual approach in specific investigations is to try to avoid this crucial difficulty by keeping only a nonlinearity and dispersion of the gradient type (and then in the continuum approximation we arrive at the ordinary models of the Korteweg–de Vries or modified Korteweg–de Vries type), or only the substrate potential combined with continuum models of the sine–Gordon or ϕ^4 type (in all such cases the problem of finding soliton-like solutions reduces to solving nonlinear second-order differential equations).

In some cases, however, such ''truncation'' is unjustified. For example, in the physics of polymer crystals the gradient nonlinear and dispersion terms are responsible for intramolecular interaction, and the nongradient terms, for intermolecular interaction. Although physically these interactions are of different origin, their strength may coincide in order of magnitude. Below we call equations that contain nonlinear and dispersion of both types generalized equations.

The present authors know of only two papers that deal with generalized models. Kovalev and Kosevich[1] give exact soliton solutions of the generalized equations

$$U_{tt}-c^2U_{zz}-6GU_zU_{zz}-FU_{zzzz}=-\frac{\partial\Psi}{\partial U} \tag{1}$$

and

$$U_{tt}-c^2U_{zz}-HU_z^2U_{zz}-PU_{zzzz}=-\frac{\partial\Phi}{\partial U} \tag{2}$$

with substrate potentials

$$\Psi=A(U^2-1)^2 \quad (\text{the } \phi^4 \text{ model}), \tag{3}$$

$$\Phi=B(1+\cos \pi U) \quad (\text{sine–Gordon model}), \tag{4}$$

where U is the scalar displacement field, G, F, H, and P are the coefficients of the nonlinear and dispersion terms, and c is the speed of sound.

In Eq. (1), the relationships between the parameters can be arbitrary, and the soliton velocity assumes only one possible value for any specific set of parameters. Equation (2) allows only for one combination of parameters, but the soliton velocity spectrum is continuous. Konno *et al.*[2] show that in the latter case Eq. (2) is exactly integrable by the inverse scattering method.

The gradient terms in Eq. (1) are the same as in the well-known Boussinesq equation, so that below we call this equation the generalized Boussinesq equation. By analogy, we say that Eq. (2) is the generalized modified Boussinesq equation. This terminology will be retained even for more complicated substrate potentials.

With the aid of a computer model of the polyethylene crystal, Balabaev *et al.*[3] show (using the molecular dynamics method) that polymer crystals exhibit a preferential localization of torsional and longitudinal nonlinear excitations along a single chain, with the gradient and nongradient terms playing comparable roles, as noted earlier. With this in mind, we can say that the longitudinal and torsional dynamics of a polymer chain in a crystal can be described (in the long-wave approximation) by the generalized and modified Boussinesq equations, respectively, which, however, contain a more realistic substrate potential.

The potentials (3) and (4) are not suited to such a description, the reason being that they are represented by functions in which the curvature of the potential well and the height of the potential barrier are intimitely related, and this relationship gives way upon passage of the solitary wave. In addition, Eq. (2) has a solution only when the dispersion and nonlinearity are rigorously related. Such restrictions are unacceptable for any realistic model of a polymer crystal.

In this paper we discuss generalized models unfettered by these restrictions, which allow for analytic soliton-like solutions.

2. EXACT SOLITON-LIKE SOLUTIONS IN GENERALIZED MODELS

We now examine the simplest possible generalizations of the substrate potentials in Eqs. (1) and (2) that lift the first of the restrictions mentioned earlier. The only thing that needs to be done is to allow for terms of the next order (for the potential (3) this is the sixth-order term and for the potential (4), the cosine of twice the angle). If we take into account the requirement that the equilibrium positions must remain unchanged, we arrive at the following expressions for the modified potentials:

$$\Psi^*=A_1(U^2-1)^2+A_2(U^2-1)^3 \quad (\text{the } \phi^4-\phi^6 \text{ model}), \tag{5}$$

$$\Phi^*=B_1+B_2+B_1\cos\pi U + B_2\cos 2\pi U \quad (\text{double sine–Gordon}). \tag{6}$$

Clearly, such a choice of potentials makes possible an independent selection of the height of the potential barrier and the curvature of the potential well. In relation to the real substrate potential, these approximations are of a ''global'' nature, i.e., they model the basic characteristics of the real potential and do not necessarily coincide with the first term in the expansion of the potential in the corresponding bases—the Taylor series and Fourier series (respectively). The validity of such an approach will be justified later (in contrast to the standard $\phi^4-\phi^6$ and double sine–Gordon models, below we allow for nonlinear gradient terms).

We seek a solution of Eq. (1) with potential (5) in the form

$$U=\tanh\kappa(z-Vt). \tag{7}$$

Substituting this into the equation leads to conditions that link the reciprocal soliton halfwidth κ and the soliton velocity V to the parameters of the problem:

$$\kappa^2(c^2-V^2)+4F\kappa^4=2A_1, \quad -2G\kappa^3+4F\kappa^4=A_2. \tag{8}$$

We seek a solution of Eq. (2) with the substrate potential (6) in the form

$$U=-1+\frac{4}{\pi}\arctan\left[\exp\left(\frac{\pi a}{2}(z-Vt)\right)\right]. \tag{9}$$

This is a true solution if the soliton parameters obey the conditions

$$a=2\left(\frac{4B_2}{2H-3\pi^2P}\right)^{1/4},$$

$$c^2-V^2=\frac{4B_1}{a^2}+\frac{a^2(H-P\pi^2)}{2}. \tag{10}$$

Obviously, a solution exists if and only if we have the following relationships between the model parameters:

$$2B_2H-3\pi^2P>0,$$

$$c^2-\frac{B_1}{\sqrt{4B_2/(2H-3\pi^2P)}}+2\sqrt{\frac{4B_2}{2H-3\pi^2P}}\,(H-P\pi^2)>0. \tag{11}$$

The two solutions correspond to the discrete velocity spectrum: (8) can yield two values for the velocity, but (10) yields only one. Depending on the parameters of the corresponding generalized Boussinesq equation, solitons specified by (7) and (8) can correspond to both expansion of a one-dimensional crystal $(\kappa<0)$ and compression $(\kappa>0)$.

In the system of the generalized modified Boussinesq equations (9) and (10), expansion and compression solitons exist for the same values of the parameters: if Eqs. (10) have a solution, the parameter a is determined to within a sign. We also note that this solution exists over a certain continuous range of dispersion-to-nonlinearity ratios, rather than just for a single ratio. Thus, the second restriction on the use of the generalized Boussinesq model is also lifted. Naturally, this system is nonintegrable.

The solution (9), (10) for $B_2=0$ was originally obtained by Beklemeshev and Klochikhin.[4]

The soliton-like solutions (7) and (9) differ from the standard solutions for systems with a discretely degenerate potential. First, these solutions can have only one or two fixed values of the soliton velocity. Such behavior of a solution for a polymer crystal model was indeed detected in the computer experiment of Balabaev *et al.*[3] and cannot be explained by the standard models, which yield a continuous velocity spectrum. Second, the speed of long-wave sound is not ''critical'' for the given type of equations, as it is in the standard models. In other words, for certain ratios of the parameters the solitons move with the speed of sound and have finite energy.

3. CORRECTNESS OF THE ''GLOBAL'' APPROXIMATION OF THE SUBSTRATE POTENTIAL

We now examine the applicability of the ''global'' approximation of the real substrate potential by employing model potentials of the $\phi^4-\phi^6$ or double sine–Gordon type.

We employ the generalized continuum model of a nonlinear chain on a substrate. The model is described by the equation

$$U_{tt}-c^2U_{zz}-t_1U_z^nU_{zz}-t_2U_{zzzz}=-\frac{\partial\mathscr{V}}{\partial U}, \tag{12}$$

where n is either 1 or 2, and the potential $\mathscr{V}$ satisfies the following conditions:

(a) $\mathscr{V}(z)=\mathscr{V}(-z)$;

(b) $z=0$ is the maximum point, $z=\pm1$ are the minimum points, and there no other extrema in the interval $(-1,1)$.

The topological soliton of Eq. (12), $u=u(z-Vt)=u(p)$, satisfies the conditions

$$u(p)=-u(-p),$$

$$p\to\infty,\quad u\to 1-d\exp(-\lambda p). \tag{13}$$

How does the soliton-like solution of Eq. (12) behave when $p\to\infty$? If we insert the second condition in (13) into Eq. (12) and ignore multiple powers of the exponential, we arrive at an expression for the decay constant of the topological soliton at infinity,

$$\lambda^2(V^2-c^2)-t_2\lambda^4=-\omega, \tag{14}$$

where ω is the curvature of the potential well near $U=1$. Equation (14) suggests that the extent of decay of the solution at infinity is determined by the curvature of the potential well of the substrate.

What is the behavior of the same solution near $p=0$? Since the solution is antisymmetric (the first condition in (13)), it can be expanded in a Taylor series in odd powers of p. The formula for calculating the coefficients of this series can be obtained by multiplying Eq. (12) by U_p and integrating the product once:

$$\frac{1}{2}(V^2-c^2)U_p^2-\frac{t_1}{n+2}U_p^{n+2}$$

$$+\frac{1}{2}t_2(U_{pp}^2-2U_pU_{ppp})=\mathscr{H}-\mathscr{V}. \tag{15}$$

The constant of integration $\mathscr{H}-\mathscr{V}(0)$ for the desired solution (which ''links'' the equilibrium positions) coincides with the height of the potential barrier of the substrate. Clearly, if in Eq. (15) we allow only for the lowest-order terms in the series expansion of the solution near zero, the right-hand side of Eq. (15) contains only this quantity.

The asymptotic procedure for finding the solutions of Eq. (12) consists in calculating the expansion in exponentials at infinity via this equation and expanding in a Taylor series at zero via Eq. (15). The two expansions can then be matched by Padé approximation. According to the general theorems of Padé approximation theory, successive diagonal approximations will then converge to the exact solution.[5]

Let us use Eq. (1) with the potential (5) to show how a soliton-like solution of the equation describing the dynamics of the generalized model can be obtained by the Padé approximation method. We seek a solution in the form $U(x)=U(z-Vt)$, and for the sake of convenience change variables: $q(U)=U_x^2$.

As a result, Eq. (1) with the potential (5) becomes

$$\frac{1}{2}(V^2-c^2)q'-3Gp^{1/2}q'-\frac{F}{2}\left(q'''q+\frac{1}{2}q'q''\right)$$

$$=-4A_1U(U^2-1)-6A_2U(U^2-1)^2. \tag{16}$$

Integrating once with respect to U, we obtain

$$-(V^2-c^2)q+4Gq^{3/2}+F\left(q''q-\frac{1}{4}(q')^2\right)$$

$$=\mathscr{V}_1+2(A_1(U^2-1)^2+A_2(U^2-1)^3). \tag{17}$$

The solution of Eq. (17) is soliton-like if the function q and its first derivative vanish at $U=\pm 1$. Clearly, these conditions uniquely determine the choice of the constant of integration: $\mathscr{V}_1=0$. Equation (17) ''suggests'' the following form for the expansion of the solution near the limit points:

$$U\to 1\, q\sim A_0(1-U)^2,\quad U\to -1\, q\sim A_0(1+U)^2. \tag{18}$$

Matching these two expansions via the two-point Padé approximation yields

$$q\sim\frac{1}{4}A_0(1-U^2)^2. \tag{19}$$

Equations (17) and (19) contain two unknown parameters, A_0 and V. Inserting (19) near the points $U\to 1$ and $U\to 0$ into (17), we arrive at two equations for these parameters,

$$(c^2-V^2)A_0+FA_0^2=8A_1, \tag{20}$$

$$(c^2-V^2)\frac{1}{4}A_0+\frac{1}{2}GF_0^{3/2}-\frac{1}{4}A_0^2F=2(A_1-A_2). \tag{21}$$

The right-hand side of Eq. (20) is the curvature of the potential well, and on the right-hand side of Eq. (21) only the overall height of the barrier from the potential remains. Putting $A_0=4\kappa^2$, we can easily see that these two equations coincide with Eqs. (8) for the parameters of the exact soliton-like solution of Eq. (1) with the potential (5).

Of course, calculating higher-order terms in the expansions requires detailed information about the substrate potential. Nevertheless, as we have shown, to calculate the first terms in the expansions near zero and at infinity we need only know the height of the potential barrier and the curvature of the potential well. In this sense any substrate potential in which these parameters are properly ''adjusted'' can serve as a satisfactory approximation. Thus, the generalized models introduced in Sec. 2 make it possible to find, to a first approximation, soliton-like solutions for all realistic substrate potentials that exhibit the necessary symmetry, provided that dispersion and nonlinearity are taken into account to leading order.

This work was supported financially by the Russian Fund for Fundamental Research (Grant No. 95-03-09026).

[1] A. S. Kovalev and A. M. Kosevich, Solid State Commun. **12**, 763 (1973).
[2] K. Konno, W. Kameyama, and H. Sanuki, J. Phys. Soc. Jpn. **37**, 171 (1974).
[3] N. K. Balabaev, O. V. Gendelman, and L. I. Manevich, submitted to Phys. Rev. Lett.
[4] S. A. Beklemeshev and V. L. Klochikhin (private communication).
[5] G. Baker and P. Graves-Morris, *Padé Approximations,* Part I: *Basic Theory,* Addison-Wesley, Reading, Mass. (1981).

Translated by Eugene Yankovsky

Effect of Pyragas feedback on the dynamics of a Q-switched laser

N. A. Loĭko, A. V. Naumenko, and S. I. Turovets

Institute of Physics, Belarus Academy of Sciences, 22072, Minsk, Belarus

(Submitted 29 December, 1996)
Zh. Éksp. Teor. Fiz. **112**, 1516–1530 (October 1997)

A linear analysis is made of the stability of the limit cycle in a Q-switched laser with continuous delayed feedback which controls the onset of instability. The Stokes generalization of the Floquet theory for the functional equations was used to derive a general equation to determine the Andronov–Hopf bifurcations and also saddle-node and subharmonic bifurcations. Conditions for stabilization of T-periodic regimes were determined and possible destabilizing effects introduced by the feedback were identified. © *1997 American Institute of Physics.* [S1063-7761(97)02710-8]

1. INTRODUCTION

Dynamic chaos, inherent in most nonlinear systems, is usually an undesirable effect in practice. Thus, the possibility of using a suitable method of dynamic control to stabilize various periodic orbits embedded within a chaotic attractor has recently been attracting close attention among researchers. Several years ago, Ott, Grebogi, and Yorke[1] proposed a general method of stabilization based on applying small feedback-controlled pulse perturbations to available system parameters. Since then, several modifications of this method have been successfully applied to various experimental systems.[2] However, it has been shown that all of these methods are sensitive to noise, mainly as a result of the time discreteness. The pulsed nature of the action also imposes fairly stringent constraints on the spectral pass band of the feedback loop. In consequence, none of these pulse techniques has been implemented so far in fast systems with comparatively high oscillation frequencies.

Pyragas[3] proposed a new method of control based on continuous delayed feedback, which was then generalized by using information on many previous states of the system.[4] In the simplest case of continuous delayed feedback, the control signal is proportional to the difference between the output signals taken at a given time and at some preceding time. The main advantages of this method are that real-time control of fast systems can be achieved, the method is stable against noise, and is relatively easy to implement experimentally. The effectiveness of continuous delayed feedback has been investigated numerically and demonstrated experimentally in various electronic devices,[5–7] in the dynamics of chemical reactions,[8,9] in the ferromagnetic resonance effect,[10] and in lasers.[11] However, this method has not been sufficiently well-studied analytically and the physical mechanisms responsible for the stabilization of unstable orbits have not been fully clarified. In many cases, the treatment of the results is based more on intuitive semiphenomenological concepts than on a mathematical analysis of stability. The present paper aims to fill this gap.

For the analysis we selected a model of a class B Q-switched laser.[12] It is well known that a Q-switched laser demonstrates a wide range of nonlinear phenomena, including a cascade of period-doubling bifurcations leading to chaos.[12–16] The difficulty of obtaining analytical results for controlling chaos in this model arises mainly from the relative mathematical complexity of the problem, which requires generalizing the Floquet theory for the functional equations. Nevertheless, the first results obtained by us earlier[17,18] showed good agreement with the experimental data on the stabilization of a T-periodic cycle (T is the loss modulation period) in a modulated CO_2 laser with continuous delayed feedback.[11] Another study[19] appeared at the same time, containing a theoretical analysis performed for the first period-doubling bifurcation using a multiscale perturbation method. This method yielded an equation for the slowly varying amplitude and phase of the approximate solution near the parametric resonance (ratio of modulation frequency to relaxation frequency$\simeq$2). From knowledge of where the $2T$-periodic solutuion exists, it was concluded that the bifurcation diagram underwent a shift.

In order to identify possible bifurcations and effects introduced by continuous delayed feedback, we propose to develop further the approach used by us in Refs. 17 and 18, i.e., direct analysis of the stability of the T-periodic regime in a system with a delay using the Floquet theory. The T-periodic solution may be obtained by using one of the well-known asymptotic methods for nonlinear laser equations.[14,15] Particular attention is focused on obtaining analytic estimates for the boundaries of saddle-node, Andronov–Hopf, and period-doubling bifurcations, and identifying the physical mechanisms responsible for each of these bifurcations. We examine the influence of the technical delay in the feedback circuit, which is unavoidable in real experimental systems, on the stabilization process. It is shown that for specific parameters additional feedback-induced instabilities may occur.

2. DESCRIPTION OF MODEL

The mathematical model of the laser has the form[14,15]

$$\begin{aligned}\dot{u}&=vu[y-1-\lambda\cos(wt+f)-K(t)],\\ \dot{y}&=y_0-(1+u)y,\end{aligned}\tag{1}$$

where u and y are the dimensionless intensity of the output field and the population inversion, v is the ratio of the rate of attenuation of the radiation in the cavity to the rate of relaxation of the population inversion, y_0 is the pump parameter, and λ and w are the amplitude and frequency of the periodic loss modulation. The term K describes the action of the continuous delayed feedback on the lasing process:

$$K(t)=\alpha(u_\tau-u_{\tau+T}). \tag{2}$$

Here the feedback depth is $\alpha\ll 1$, and U_τ and $u_{\tau+T}$ are the intensities at the delayed times $t-\tau$ and $t-(\tau-T)$. Unlike Refs. 3 and 11 we introduce here an additional delay τ which may be comparable with T and associated with the finite time taken for propagation of the signal in the feedback loop. It is known that this type of feedback does not influence unstable T-period orbits and does not require them to be known in advance. We postulate, as in Ref. 3, that the signal $K(t)$ is bounded so that the feedback only influences the stability of the T-periodic orbit in a certain small neighborhood and does not significantly alter the dynamics of the system far from this orbit.

3. ANALYSIS OF STABILITY

Below we shall use a logarithmic change of variables in which the solutions of system (1) corresponding to spiky behavior of the laser look smoothed for any values of the parameter v. This makes it easier to find analytic and numerical solutions of the problem.[14] It is known that the laser equations rewritten in these variables are the equations for a nonlinear oscillator with the Toda potential controlled by an external periodic force and parametrically, and may be expressed in the form

$$\begin{aligned}&\ddot{x}+\dot{x}(\epsilon_1+\epsilon_2e^x)+\dot{R}+[1+\lambda\cos(mt')]e^x+R(\epsilon_1+\epsilon_2e^x)\\&\quad=1+\lambda\epsilon_2^{-1}m\sin(mt')-\lambda\epsilon_1\epsilon_2^{-1}\cos(mt'),\end{aligned} \tag{3}$$

where $R=\alpha\epsilon_1^{-1}[\exp(x_{\tau'})-\exp(x_{\tau'+T'})]$ describes the influence of feedback and

$$x=\ln\frac{u}{y_0-1},\quad m=w\epsilon_1,\quad t'=\frac{t}{\epsilon_1},\quad \epsilon_1=\frac{1}{\sqrt{v(y_0-1)}},$$

$$\epsilon_2=\sqrt{\frac{y_0-1}{v}},\quad T'=\frac{T}{\epsilon_1},\quad \tau'=\frac{\tau}{\epsilon_1}.$$

Here the normalized modulation frequency m is expressed in units of the frequency of the relaxation oscillations ϵ_1^{-1}.

In the following equations all the primes will be omitted to simplify the notation. Unlike Eqs. (1), the harmonic balance method applied to Eq. (3) gives a fairly good approximation to the accurate T-periodic solution when the zeroth-order and first-order harmonics are used. For instance, a Fourier transformation shows that even near the main resonance subject to the condition $m>1$, the difference between the trial function for the T-periodic solution taken in the form $x_0=c+a\cos(mt)+d\sin(mt)$ and the accurate solution determined numerically is around a few percent.[14] The values of c and $y=a/r$ $(r^2=a^2+d^2)$ may be determined from the following equations:

$$e^c\,[I_0(r)+\lambda yI_1(r)]=1,$$

$$\frac{(c_0+c_1y)^2(1-y^2)}{(c_2+c_3y+c_4y^2)^2}+\frac{d_1+d_2y}{d_3+d_4y}-d_5\left[\frac{(c_0+c_1y)(1-y^2)}{c_2+c_3y+c_4y^2}\right.$$

$$\left.-d_6y\right]=0,\quad d_1=\frac{2I_1(r)}{r},$$

$$d_2=\frac{2\lambda}{r}\left[I_0(r)-\frac{I_1(r)}{r}\right],\quad d_3=I_0(r),$$

$$d_4=\lambda I_1(r),\quad d_5=\frac{\lambda}{\epsilon_2},\quad d_6=\epsilon_1,$$

$$c_0=\frac{d_3\lambda}{r\epsilon_2}-\frac{\lambda d_1}{r\epsilon_1},\quad c_1=\frac{d_4\lambda}{r\epsilon_2},\quad c_2=d_3+\frac{d_1\epsilon_2}{\epsilon_1},$$

$$c_3=d_4+\frac{d_3\lambda}{r\epsilon_1\epsilon_2},\quad c_4=\frac{d_4\lambda}{r\epsilon_1\epsilon_2}, \tag{4}$$

where $I_n(r)$ is a modified Bessel function obtained by expanding

$$\exp(r\cos t)=I_0(r)+2\sum_{k=1}^{\infty}I_k(r)\cos(kt).$$

To analyze the stability of this solution, we write the linearized equation for the deviation $\xi(t)=x(t)-x_0(t)$:

$$\begin{aligned}&\ddot{\xi}+f(t)\dot{\xi}+g(t)\xi+\alpha[h(t)(\dot{\xi}_\tau-\dot{\xi}_{\tau+T})\\&\quad+q(t)(\xi_\tau-\xi_{\tau+T})]=0,\end{aligned} \tag{5}$$

where $f(t)$, $g(t)$, $h(t)$, and $q(t)$ are T-periodic functions of time, which may be expressed in terms of the solution $x_0(t)$ as follows:

$$\begin{aligned}&f(t)=\epsilon_1+\epsilon_2e^{x_0},\quad g(t)=[\epsilon_2\dot{x}_0+1\\&\quad+\lambda\cos(mt)]e^{x_0},\quad h(t)=\epsilon_1^{-1}e^{x_\tau^0},\\&q(t)=\epsilon_1^{-1}e^{x_\tau^0}(\dot{x}_\tau^0+\epsilon_2e^{x_0}+\epsilon_1).\end{aligned}$$

Here the following notation is introduced for brevity: $x_\tau^0=x_0(t-\tau)$, $\xi_\tau=\xi(\tau)$, and $\xi_{\tau+T}=\xi(\tau+T)$.

We shall first make a linear analysis of the stability of the system (1) when the modulation depth is zero $(\lambda=0)$. In this limiting case, Eq. (5) is reduced to the following expression:

$$\ddot{\xi}+\delta\dot{\xi}+\xi+b[(\dot{\xi}_\tau-\dot{\xi}_{\tau+T})+\delta(\xi_\tau-\xi_{\tau+T})]=0, \tag{6}$$

where $b=\alpha/\epsilon_1$, $\delta=\epsilon_1+\epsilon_2$ is the damping factor for small oscillations of the unperturbed Toda oscillator (3) $(\lambda=0, b=0)$. If the solution is taken in the form $\xi(t)=e^{\mu t}$, a characteristic quasipolynomial may be obtained to find the exponent μ:

$$\mu^2+\delta\mu+1+b(\mu+\delta)(e^{-\mu\tau}-e^{-\mu(\tau+T)})=0.$$

The bifurcation boundary (or region of instability) is determined from the condition Re $\mu=0$ and has a form similar to the case of ordinary negative feedback whose signal is only proportional to a single value of the intensity at a delayed time:[20]

$$\begin{aligned}&-\chi^2+1+b(\delta C+\chi S)=0,\\&\delta\chi+b(-\delta S+\chi C)=0,\end{aligned} \tag{7}$$

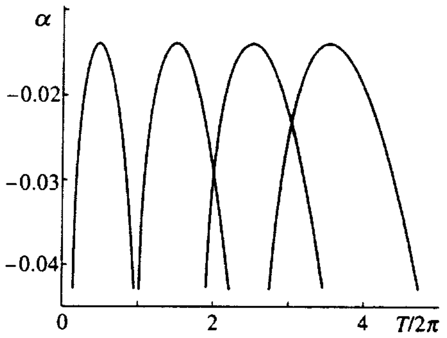

FIG. 1. Boundaries of regions of instability on the (α,T) plane without modulation $(\lambda=0)$. Parameters: $v=70$, $y_0=2$.

where $\chi=\mathrm{Im}\,\mu$ is the pulsation frequency at the bifurcation boundary,

$$C=\cos(\chi\tau)-\cos[\chi(\tau+T)],$$

$$S=\sin(\chi\tau)-\sin[\chi(\tau+T)].$$

These equations can easily be solved for $\tau=0$. Solving the linear system (7) for $\sin(\chi T)$ and $\cos(\chi T)$, squaring, and summing the resultant expressions yields the following relation:

$$b=\frac{1}{2\delta}[1-\chi^2(2-\delta^2)+\chi^4],$$

from which it follows that

$$\chi_{1,2}^2=\frac{1}{2}(2-\delta^2\pm\sqrt{D}). \tag{8}$$

For instability to exist the value of the discriminant D should be positive:

$$D=-\delta^2(4-\delta^2)+8b\delta>0.$$

Then

$$b<b^*=-\frac{\delta}{2}\left(1-\frac{\delta^2}{4}\right),$$

where b^* is the threshold value of b for the existence of an Andronov–Hopf bifurcation. The corresponding value is $\chi^*=1-\delta^2/2$. Since we have $\delta\ll 1$ for the typical parameters of a class B laser, the frequency χ of the unstable solution near the threshold is close to the resonant relaxation frequency, which is unity in this normalization.

The bifurcation value of T is determined by

$$T=\frac{1}{\chi}\left\{\arctan\frac{-2\delta\chi[1-(\delta^2+\chi^2)]}{\chi^2[1-(\delta^2+\chi^2)]-\delta^2}+2\pi n\right\}. \tag{9}$$

The value of the arctangent is selected according to the sign of the function $\sin(\chi T)$ which is proportional to $\delta^2+\chi^2-1$ for $\chi>0$. The bifurcation curves determined from formulas (8) and (9) are plotted in Fig. 1. For small δ the instability thresholds (maxima of the curves for $b=b^*$) are obtained for the following values of the delay T:

$$T^*\approx-\delta+2\pi\left(1+\frac{\delta^2}{4}\right)\left(n+\frac{1}{2}\right). \tag{10}$$

The corresponding feedback frequencies $2\pi/T^*$ causing instability at the relaxation frequency are close to $2/(2n+1)$. Thus, for a given relation between the feedback frequencies and the relaxation frequency, auto-oscillations are excited if the depth of feedback has a negative value, greater than the magnitude of the bifurcation value $|b^*|$. It is interesting to note that for these values of the external modulation frequency, period-doubling regions are observed corresponding to odd Matthieu zones.[14,15] It will be shown subsequently that for this frequency ratio the system may become destabilized (or stabilized—depending on the sign of the feedback depth) and in general $\lambda\neq 0$. At this point, it is appropriate to note that, as can be seen from the previous analysis and from Fig. 1, the physical mechanism responsible for the action of Pyragas feedback involves the introduction of effective frequency-selective dissipation into the system. In order to optimize the excitation (or suppression) of relaxation oscillations with the period 2π near the steady-state solution, the delay in the feedback loop should be selected in accordance with condition (10), and in particular should be of the order π.

We now consider the case of nonzero modulation depth λ.

It is known from the theory of linear delayed differential equations with periodic coefficients that the solution of Eq. (5) may be expressed in the form[21]

$$\xi=\sum_k^\infty \exp(\mu_k m t)\sum_{j=0}^{n_k-1} C_{kj}t^j\varphi_{kj}(t),$$

where μ_k are the characteristic exponents, n_k is the multiplicity of the exponent μ_k, $\varphi_{kj}(t)$ are T-periodic functions, and C_{kj} are constants. Assuming the nondegenerate case $(n_k=1)$ and substituting the particular solution $\xi_k=\exp(\mu_k mt)\varphi_{k0}$ into formula (5), we obtain an equation for each φ_{k0} (we subsequently omit the subscripts):

$$\ddot{\varphi}+(2\mu m+f)\dot{\varphi}+[(\mu m)^2+\mu m f+g]\varphi$$
$$+\beta[h\dot{\varphi}_\tau+(\mu m h+q)\varphi_\tau]=0, \tag{11}$$

where $\beta=\alpha\{\exp(-\mu m\tau)-\exp(-\mu m(\tau+T))\}$. We can say that Eq. (11) is an equation for the unknown functions $f(t)$, $g(t)$, $h(t)$, $q(t)$, and therefore the parameters of the system (1) if the characteristic exponent μ of Eq. (5) is fixed assuming that the function φ is T-periodic. Then the opposite is also true: Eq. (11) can be used to determine μ for specified system parameters and hence specified forms of the functions listed above, subject to the same condition for the function φ.

Using a Fourier series expansion of the function φ $(\varphi=\Sigma_k^\infty\varphi_k e^{imkt})$, as in the usual Floquet theory, we can obtain the following characteristic equation for the exponent μ:

$$\det A_{nl}=0, \tag{12}$$

where

$$A_{nl}=-\delta_{nl}(ml-i\mu m)^2+W(l,n-1),$$

$$\delta_{nl}=\begin{cases}1, & n=l,\\ 0, & n\neq l,\end{cases}$$

$W(l,n)$ are the Fourier components of the function

$$W(t|l)=\mu mf(t)+g(t)+imlf(t)+\beta[\mu mh(t)+q(t)+imlh(t)]e^{-iml\tau},$$

$$W(t|l)=\sum_{k=-\infty}^{\infty} W(l,k)e^{imkt}.$$

The coefficients $W(l,k)$ have the property $W(l,k)^{*}=W(-l,-k)|_{\mu=\mu^{*}}$.

In the absence of feedback Eq. (12) is similar to an infinite Hill determinant. At the same time, it is an analog of a quasipolynomial which has an infinite number of roots in the complex plane, in the particular case of a laser without Q-switching examined above. Equation (12) can also be reduced to the Hill determinant in the general case if the technical delay is $\tau=0$. It is known[22] that in the Hill equation, only that part of the determinant corresponding to the maximum harmonics contains the main information on the bifurcation boundary. On the basis of general concepts of continuity, we assume that this also remains valid for $\tau\neq 0$. The validity of this assumption will be confirmed subsequently by means of a numerical simulation.

We consider the following cases of loss of stability by the T-cycle: saddle-node, period-doubling, and Andronov–Hopf bifurcations.

For the case of a saddle-node bifurcation, one purely real multiplier $M=e^{2\pi\mu}$ intersects the unit circle at the point $M=1$. Assuming $\mu=0$ in order to obtain this value of the exponential, we find $\beta=0$, and Eqs. (11) and (12) to determine the bifurcation boundary are the same as the equations in the absence of feedback. Thus, a saddle-node bifurcation does not vary under the influence of Pyragas feedback (a detailed analysis of this bifurcation using a quasiconservative Lyapunov approach was reported in Ref. 14). For the same reason, continuous delayed feedback cannot give rise to a new saddle-node bifurcation boundary (or any other bifurcation which involves the real multiplier passing through unity, such as a cycle symmetry-breaking bifurcation). Thus, in principle, the stabilization of an unstable T-cycle with a multiplier greater than unity cannot be achieved by continuous delayed feedback.

To determine the period-doubling bifurcation, we can express the exponent μ in the form $\mu=\zeta+i/2$, where ζ is a real quantity. At the bifurcation boundary we find $\zeta=0$, and this can be determined using Eq. (12), substituting $\mu=i/2$. It may be postulated that μ is real and equal to ζ however, but instead of a T-periodic function we take a $2T$-periodic function for φ: $\varphi(t+2T)=\varphi(t)$ and $\varphi(t+T)=-\varphi(t)$. Then in Eq. (11) we have

$$\beta=\alpha[e^{-\zeta m\tau}+e^{-\zeta m(\tau+T)}],$$

and an equation similar to (12) may be obtained for the exponent μ, using an expansion in the form

$$\varphi=\sum_{k}^{\infty} \varphi_{2k+1}\exp\left[im\left(k+\frac{1}{2}\right)t\right].$$

Since the first harmonic has a maximum near the bifurcation boundary in this sum,[22] the new equation of the form (12) may be converted to a reduced equation for the Floquet exponent ζ:

$$|-(m/2-i\zeta m)^2+W_0|^2=|W_1|^2, \tag{13}$$

where $W_{0,1}$ are the first Fourier components of the function

$$W(t)=W\left(t\middle|l=\frac{1}{2}\right)=\zeta mf(t)+g(t)+\frac{im}{2}f(t)+\beta\left[\zeta mh(t)+q(t)+\frac{im}{2}h(t)\right]e^{-im\tau/2}.$$

The bifurcation boundary is determined from Eq. (13) where $\zeta=0$.

If we set $\tau=0$, this boundary can also be determined by a standard method[22] from the following ordinary differential equation, obtained from Eq. (11) with a delay, for the value of β redetermined above:

$$\ddot{\varphi}+(f+2\alpha h)\dot{\varphi}+(g+2\alpha q)\varphi=0. \tag{14}$$

Equation (14), like Eq. (11), is an equation for the parameters of the system, including the bifurcation parameters, provided that the function φ is $2T$-periodic. Equation (14) has two linearly independent solutions: φ_1 and $e^{-\delta t}\varphi_2$, where $\varphi_{1,2}$ are $2T$-periodic functions, and δ is a constant term in the Fourier expansion of the function $F(t)=f(t)+2\alpha h(t)$, and it determines the damping rate, which is equal to the sum of the two Lyapunov exponents ($\kappa_1=0$, $\kappa_2=-\delta$) with opposite sign. For small α these exponents are largest in the initial system (1) which, like any system with a delay, has an infinite number of characteristic exponents.

By means of the change in variables $\varphi=\exp\{-(1/2)\times\int F(\tau')d\tau'\}\eta$, we reduce Eq. (14) to a Hill equation:

$$\ddot{\eta}+G(t)\eta=0,$$

where

$$G(t)=[g(t)+2\alpha q(t)]-\frac{1}{2}[f(t)+2\alpha h(t)]_t'-\frac{1}{4}[f(t)+2\alpha h(t)]^2.$$

Introducing the new normalized time $\tau'=mt$, we obtain the equation in canonical form:

$$\ddot{\eta}+\frac{1}{m^2}\left[\theta_0+2\sum_{k=1}^{\infty}\theta_k\cos(k\tau'-\epsilon_k)\right]\eta=0,$$

where θ_k are the Fourier components of the function $G(t)$ and $\omega=\sqrt{\theta_0}/m\approx 1/m$ is the oscillation frequency of the unperturbed system. If we adopt the Whittaker method for the nth instability zone, the solution may be taken in the first approximation as

$$\eta=e^{\nu\tau'}\sin\left(\frac{n}{2}\tau'+\sigma\right).$$

Here ν, the Floquet exponent for the Hill equation, has real or purely imaginary values, and σ is a constant. We shall

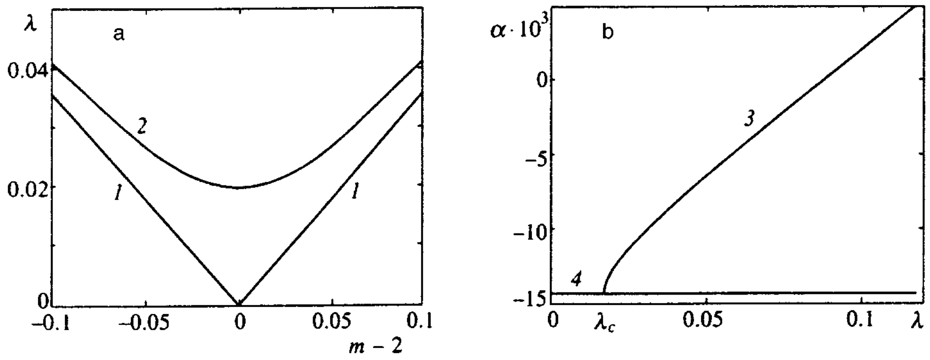

FIG. 2. Bifurcation boundaries in the planes (λ,m) and (α,λ) for small detunings $\Delta=m-2$ and an additional delay $\tau=0$: a) Curve *1*—boundary of virtual Andronov–Hopf bifurcation, curve *2*—boundary of period-doubling bifurcation for $\alpha=-0.011$. The region between these curves is the region of classical compression; b) Curves *3* and *4*—boundaries of period-doubling bifurcation for $\Delta=0.05$ and Andronov–Hopf bifurcation, respectively, and λ_c is the point of a degenerate Andronov–Hopf bifurcation.

only analyze odd zones, since the necessary condition of $2T$-periodicity is only satisfied at the boundary of these zones. Note that the odd zones $(n=2l+1)$ are determined by values of the modulation frequency $m=2/n$ close to the frequencies $2\pi/T^*$ of the positive $(\alpha<0)$ feedback in an unmodulated system $(\lambda=0)$, for which an Andronov–Hopf bifurcation may occur. We are principally interested in the main (first) instability zone. The Floquet exponents may be obtained from

$$\nu^2=-\left[\frac{\theta_0}{m^2}+\left(\frac{n}{2}\right)^2\right]+\sqrt{n^2\frac{\theta_0}{m^2}+\frac{\theta_n^2}{m^4}}.$$

The stability condition is determined by the inequality

$$\mathrm{Re}\ \kappa=\mathrm{Re}(-\delta/2\pm\nu)<0, \tag{15}$$

which reduces to the form

$$\delta>0,$$

$$\left[\theta_0-\left(\frac{mn}{2}\right)^2\right]^2+\left(\frac{\delta}{2}\right)^2\left[\theta_0+\left(\frac{mn}{2}\right)^2\right]+\left(\frac{\delta}{2}\right)^4>\theta_n^2. \tag{16}$$

Expressions (15) and (16) determine a bifurcation curve which is almost the same as that obtained from the more general equation (13). However, expressions (15) and (16) are more convenient for the analysis and allow various simple conclusions to be drawn.

Let us consider the first zone $(n=1)$. The behavior of the system near the unstable orbit at the bifurcation point is mainly determined by the two maximum Lyapunov exponents which give the following damping rate:

$$\delta=\epsilon_1+(\epsilon_2+2\alpha/\epsilon_1)e^cI_0(r).$$

Considering the feedback depth α to be a variable which is second-order in the parameter $\epsilon_{1,2}\sim\epsilon$, we can conclude that near the instability threshold at $m-2=\Delta\ll1$, the cycle amplitude r is small. Then the feedback mainly influences only the damping rate $(\delta_\Delta=\delta(\alpha)-\delta(\alpha=0)\sim\alpha/\epsilon\sim\epsilon\sim\delta$, $G_\Delta\sim\alpha\sim\epsilon^2\ll G)$. The variation of the coefficients θ_0 and θ_1, and also the Fourier components of the function G (and therefore the exponents ν) accompanying any change in α is also small. With these constraints on the system parameters, the following relation can be obtained for the cycle amplitude:[14]

$$r\approx\frac{\lambda m}{\epsilon_2(1-m^2)}.$$

In this case, we have

$$\delta=\epsilon_1+\epsilon_2+\frac{2\alpha}{\epsilon_1},\quad \theta_0\approx1-\left(\frac{\delta}{2}\right)^2,\quad \theta_1^2\approx\frac{r^2}{4}.$$

Then the period-doubling bifurcation boundary $\nu=\delta/2$ determined by inequality (15) is described by

$$\left(\frac{\lambda}{3\epsilon_2}\right)^2=\Delta^2+\delta^2, \tag{17}$$

which is the same as that derived in Ref. 19. This curve is plotted in Figs. 2a and 2b.

For $\tau\neq0$ the bifurcation boundary can only be obtained directly from Eq. (13) in a form similar to Eq. (17) under the same assumptions

$$\left(\frac{\lambda}{3\epsilon_2}\right)^2=\left(\Delta-\frac{2\alpha}{\epsilon_1}\sin\frac{m\tau}{2}\right)^2+\left(\epsilon_1+\epsilon_2+\frac{2\alpha}{\epsilon_1}\cos\frac{m\tau}{2}\right)^2. \tag{18}$$

It can be concluded from a comparison of expressions (18) and (17) that if there is an additional delay τ, the role of the damping in the system is played by $\epsilon_1+\epsilon_2+(2\alpha/\epsilon_1)\cos(m\tau/2)$ which depends periodically on τ.

Negative feedback $(\alpha>0)$ increases δ, which leads to suppression of the period-doubling bifurcation and stabilization of the T-cycle since, as can be seen from Eq. (17), for small Δ the instability threshold is proportional to the dissipation in the system:

$$\lambda_{\mathrm{th}}=3\epsilon_2\delta\left(1+\frac{\Delta^2}{2\delta^2}\right).$$

The results of the calculations using the general equation (13), plotted in Fig. 3a, show that this conclusion, reached near resonance $(m\approx2)$ is valid over a wider region. Values of $m\gg2$ are an exception. For instance, for $m>2.4$ the instability threshold in terms of λ is slightly reduced, i.e., negative feedback may lead to destabilization via a period-doubling bifurcation.

Figure 3a also shows the bifurcation boundaries obtained by a numerical integration of the system (1). It can be seen that the curves calculated analytically and numerically show good agreement for $1.6\leq m\leq2.4$ and describe the correct trend of the boundary shift for $m<1.6$ and $m>2.4$. Thus, the analytic results remain valid over a fairly wide range of m.

The shift of the bifurcation boundary also depends on the delay τ (Fig. 3b) which determines the phase shift of the

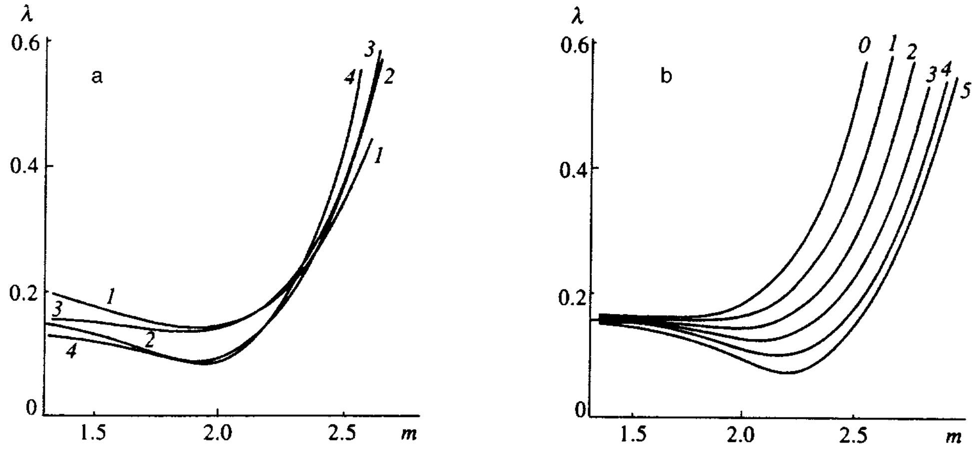

FIG. 3. Boundaries of period-doubling bifurcation on the (λ,m) plane: a—$\alpha=0$ (curves *2* and *4*), $\alpha=0.01$, $\tau/T=0.1$ (curves *1* and *3*); curves *1* and *2* give the numerical results obtained by integrating system (1), and curves *3* and *4* were obtained by analytic calculations using formula (13); b—$\alpha=1/70$, $\tau/T=0$, 0.1, 0.2, 0.3, 0.4, and 0.5, respectively for curves *0*, *1*, *2*, *3*, *4*, and *5*.

feedback signal pulsations relative to the laser pulses and therefore influences the effective feedback depth $\alpha'=\alpha\cos(m\tau/2)$ by altering the total dissipation in the system. For instance, a change in τ by the value T is equivalent to a change in the sign of α'. It follows from Eqs. (11)–(13) and (18) that the values of the bifurcation parameters are $2T$-periodic in τ. Near $m\approx 2$ the increase in the instability threshold is optimum for $\tau\approx 0$.

Positive feedback $(\alpha<0)$, on the other hand, reduces friction in the system and therefore lowers the period-doubling threshold in λ, even near resonance. In particular, it follows from Eq. (17) that for $m=2$ $(\Delta=0)$ the period-doubling bifurcation threshold approaches zero for $\alpha_0=-[(\epsilon_1+\epsilon_2)\epsilon_2]/2(\delta=0)$. For finite detunings from the exact resonance condition $(m=2)$, the minimum threshold is also determined by the constraint $\delta=0$ and is $\lambda_c=3\epsilon_2|\Delta|$ (curve *1* in Fig. 2a, intersection of curves *3* and *4* in Fig. 2b). At this boundary we find $\nu=0$.

Note that curve *1* is determined by the condition $\nu=0$ for any δ and the gap between it and the period-doubling bifurcation boundary $\nu=\delta/2$ is the region of classical compression (or anisotropic phase space).[23,24] The condition $\nu=0$ determines a curve which may be called the virtual Andronov–Hopf bifurcation boundary,[25] since the multipliers above it are real and those below it become complex-conjugate. As the modulation depth decreases to zero, these multipliers undergo virtual motion near the boundary of the unit curve, without intersecting it.

In the region where ν becomes purely imaginary, the virtual Andronov–Hopf bifurcation is converted into a real one where the positive feedback compensates for the friction in the system. Thus, when $\delta\leqslant 0$ holds for $\lambda<\lambda_c$ the system is only unstable as a result of the Andronov–Hopf bifurcation whose α threshold in a first approximation does not depend on λ and Δ in this region, and is equal to $\alpha_{\text{Hopf}}\approx\alpha_0$. This value is the minimum α threshold for the Andronov–Hopf bifurcation in an unmodulated laser (Eqs. (8) and (9), $\alpha^*=\epsilon_1 b^*$). Note that the frequency of the excited pulsations is also the same as the Andronov–Hopf frequency at $\lambda=0$. The nature of the instability in these cases is clearly the same, and is associated with the buildup of relaxation oscillations with a particular relation between their frequencies and the feedback and Q-switching frequencies. If $\delta=0$ holds, then for $\lambda=\lambda_c$ we have the case of a degenerate Andronov–Hopf bifurcation where this is the same as a period-doubling bifurcation. Above this boundary, the T-periodic solution is unstable to perturbations with the period $2T$ and below this boundary it is unstable to perturbations with the Andronov–Hopf frequency. For $\delta\geqslant 0$, the T cycle is unstable near $\lambda\leqslant\lambda_c$. It is natural to assume that a bifurcation boundary with zero damping exists far from $m\approx 2$. In this case, the value of α_{Hopf} determined from the condition $\delta=0$ also depends on m and λ and may, in principle, have positive values. Moreover, because of the complexity of the dependence $\delta(\lambda,m)$ far from resonance, the condition $\delta=0$ may be multivalued in terms of λ and m.

These relationships are confirmed by numerical integration of the system linearized near the T-periodic solution x_0 and are shown in Fig. 4. Curves *AB* (Figs. 4a and 4c) and *AC* (Fig. 4b) are the boundaries of a saddle-node bifurcation which does not depend on α. Since the initial approximation of the first harmonic of the exact T-periodic solution x_0 in the resonance region where $m\leqslant 1$ holds is rather rough, this boundary is slightly shifted relative to the real boundary determined by direct integration of the system (1). Curve *CD* is the boundary of a period-doubling bifurcation. Its dependence on α (Fig. 4a, $\alpha>0$ and Fig. 4b, $\alpha<0$) and τ (Fig. 4c) demonstrates agreement with our analytical results. When this part of the curve touches the $\lambda=0$ axis, it becomes the boundary of a virtual Andronov–Hopf bifurcation near $m\approx 2$ (section *CD* of curve *3* in Fig. 4b). Both multipliers on this curve near the minimum are close to -1. The shift of the point of contact relative to the exact resonance $(m=2)$ is associated with a nonzero value of τ and is determined from Eq. (18) by $\Delta=(2\alpha/\epsilon_1)\sin(m\tau/2)$. Below this, the multipliers of the system are complex and the stability of the T-cycle depends on δ. For $|\alpha|>|\alpha_0|$, i.e., when $\delta\leqslant 0$ holds near $m\approx 2$ (curve *4* in Fig. 4b), the T-periodic solution is unstable to the Andronov–Hopf bifurcation over a wide range of m bounded by the bifurcation curves *DB* and *FE*, where δ changes sign. For $|\alpha|<|\alpha_0|$, i.e., when $\delta\geqslant 0$ holds near $m\approx 2$ (curves *3* in Fig. 4b and curves *2–4* in Fig. 4c), Andronov–Hopf instability is observed far from resonance in regions bounded by curves *BC* (Fig. 4a) and *DE* (Fig. 4b). For instance, for $\alpha>0$ this instability is found to the left of the period-doubling bifurcation curve (Fig. 4a). As we move

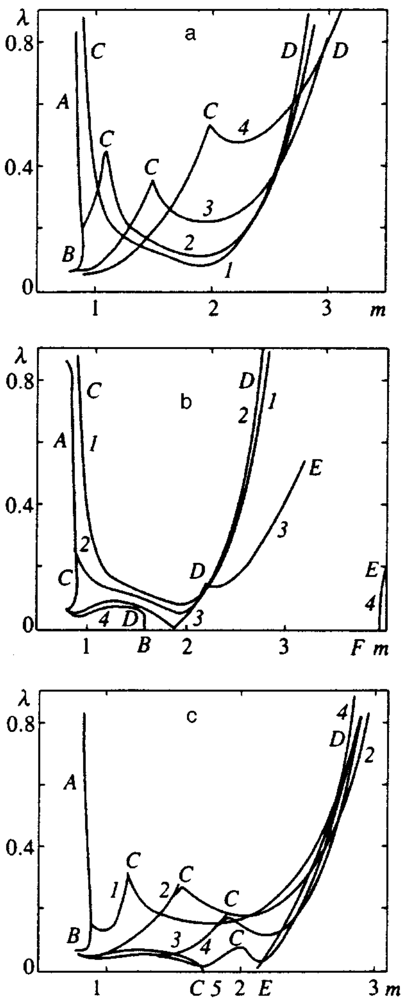

FIG. 4. Bifurcation boundaries on the (λ, m) plane obtained by integrating the linearized equation (5): a) $\tau/T=0.1$, $\alpha=0$ (*1*), 0.005 (*2*), 0.02 (*3*), and 0.04 (*4*); *AB* corresponds to a saddle-node bifurcation, *CD* is a period-doubling bifurcation, and *BC* is an Andronov–Hopf bifurcation. b) $\tau/T=0.1$, $\alpha=0$ (*1*), -0.005 (*2*), $-1/70$ (*3*), and -0.02 (*4*); *AC*—boundary of saddle-node bifurcation, *CD*—period-doubling bifurcation, *DE*, *DB*, and *FE*—Andronov–Hopf bifurcations. c) $\alpha=1/70$, $\tau/T=0$ (*1*), 0.2 (*2*), 0.5 (*3*), 0.7 (*4*), and 0.8 (*5*); *AB*—boundary of saddle-node bifurcation, *CD*—period-doubling bifurcation, *BC* and *ED*—Andronov–Hopf bifurcations.

along curve BC to point C (period-doubling bifurcation point), the multipliers tend to -1. As we approach point B (saddle-node bifurcation), the multipliers tend to 1. For $\alpha<0$ Andronov–Hopf instability exists to the right of the period-doubling bifurcation (Fig. 4b).

4. CONCLUSIONS

We have examined the possibility of controlling chaos by continuous delayed feedback in class B Q-switched lasers. For this purpose we developed a theory derived from an analysis of stability in a general form. Analytic estimates were made of the boundaries of saddle-node, period-doubling, and Andronov–Hopf bifurcations, which provide information on instability trends in the system as the parameters vary.

As a result, it was established that a system with Pyragas feedback does not alter stability cycles with real multipliers greater than unity which implies, for instance, that the boundaries of saddle-node and symmetry-breaking bifurcations do not depend on this feedback.

The physical mechanism for the suppression of a period-doubling bifurcation is associated with increased total dissipation in the system caused by the feedback effect. For certain feedback parameters, on the other hand, the dissipation in the system may decrease, leading to the excitation of instability.

An analytic investigation was made of the influence of the unavoidable technical delay in a continuous feedback system. This influence becomes particularly important for fast systems where the delay may be comparable with or even greater than the fundamental period of the oscillations which need to be stabilized, as in a semiconductor laser, for instance ($\tau\simeq 10$ ns, $T\simeq 1$ ns). In these cases, the analytic results are most useful since direct numerical calculations require a large amount of computer time and memory. The technical delay alters the threshold of a period-doubling bifurcation with the period $2T$, as was recently confirmed in independent numerical investigations.[26,27]

An Andronov–Hopf bifurcation appears in the system when the dissipation is compensated by feedback. Near resonance ($m=2$) this instability is found for the same depths of feedback as without loss modulation and is mainly attributable to the action of the feedback, which excites relaxation oscillations with a particular relation between their frequency and the feedback frequencies.

Note that our analysis was made using a harmonic balance method to determine the T-periodic solution in its simplest form where only one harmonic is taken into account for the intensity logarithm. This method gives good results for the lowest bifurcations (the first instabilities) in the modulation frequency range $m\geqslant 1.5$ where the oscillations are still fairly small. Near resonance ($m\leqslant 1$), this method is less accurate but still gives a very useful idea of the general trends in the instability behavior. The analytic estimates in this region may be improved either by using a larger number of harmonics in the harmonic balance method or by adopting the asymptotic quasiconservative Lyapunov approach to find the T-periodic solution.[14,15]

The validity of the approximations made in the stage of solving the linearized system of equations to determine the bifurcation boundaries was confirmed by numerical integration of Eqs. (5) and (1).

This work was made possible by support from the International Science Foundation (Grant No. MX5000) and the Belarus Fund for Fundamental Research (Grant No. 15006).

[1] E. Ott, C. Grebogi, and J. A. Yorke, Phys. Rev. Lett. **64**, 1196 (1990).
[2] B. T. Shinbort, Adv. Phys. **44**, 73 (1995).
[3] K. Pyragas, Phys. Lett. A **170**, 421 (1992).
[4] J. E. S. Socolar, D. W. Sukow, and D. J. Gauthier, Phys. Rev. E **50**, 3245 (1994).
[5] K. Pyragas and A. Tamaševičius, Phys. Lett. A **180**, 99 (1993).
[6] D. J. Gauthier, D. W. Sukow, H. M. Concannon, and J. E. S. Socolar, Phys. Rev. E **50**, 2343 (1994).
[7] Y. H. Yu, K. Kwak, and T. K. Lim, Phys. Lett. A **191**, 233 (1994).

[8] F. W. Schneider, R. Blittersdorf, A. Förster, T. Hauck, D. Lebender, and J. Müller, J. Phys. Chem. **97**, 12 244 (1993).

[9] A. Lekebusch, A. Forster, and F. W. Schneider, J. Phys. Chem. **99**, 681 (1995).

[10] F. E. Wigen and D. W. Peterman, J. Magn. Magn. Mater. **140**, 2074 (1995).

[11] S. Bielawski, D. Derozier, and P. Glorieux, Phys. Rev. E **49**, R971 (1994).

[12] F. T. Arecchi, R. Meucci, G. P. Piccioni, and J. R. Treduce, Phys. Rev. Lett. **49**, 1217 (1982).

[13] T. Midsvaine, D. Dangoise, and P. Glorieux, Phys. Rev. Lett. **55**, 1989 (1985).

[14] M. Samson and S. I. Turovets, Preprint No. 438 [in Russian], Institute of Physics, Academy of Sciences of the Belarussian SSR (1986); Sov. Phys. Dokl. **31**, 888 (1987).

[15] A. M. Samson, L. A. Kotomtseva, and N. A. Loĭko, *Self-Oscillations in Lasers* [in Russian], Nauka i Tekhnika, Minsk (1990).

[16] A. M. Samson, S. I. Turovets, V. N. Chizhevskiĭ, and V. V. Churakov, Zh. Éksp. Teor. Fiz. **101**, 1177 (1992) [Sov. Phys. JETP **74**, 628 (1992)]; V. N. Chizhevsky and S. I. Turovets, Phys. Rev. A **50**, 1840 (1994).

[17] N. A. Loiko, A. V. Naumenko, and S. I. Turovets, Conference Digest of NDOS'95, Rochester, NY, 1995, Paper TE 7.

[18] N. A. Loiko, A. V. Naumenko, and S. I. Turovets, in *Nonlinear Dynamics in Lasers*, edited by N. B. Abraham, Proc. SPIE **2792**, 102 (1996).

[19] S. Bielawski, D. Derozier, P. Glorieux, and T. Erneux, J. Europ. Opt. Soc. B **7**, 957 (1995).

[20] E. V. Grigorieva, S. A. Kashchenko, N. A. Loiko, and A. M. Samson, Physica D **59**, 297 (1992).

[21] L. E. El'sgol'ts and S. B. Norkin, *Introduction to the Theory of Differential Equations with a Deviating Argument*, Academic, New York (1973).

[22] C. Hayasi, *Nonlinear Oscillations in Physical Systems*, McGraw-Hill, New York (1984); reprinted by Princeton Univ. Press (1985).

[23] V. N. Chizhevsky and S. I. Turovets, Opt. Commun. **102**, 175 (1993).

[24] N. P. Pettiaux, P. Mandel, and C. Fabre, in *OSA Proceedings on Nonlinear Dynamics in Optical Systems*, edited by N. B. Abraham, E. Garmire, and P. Mandel, Vol. **7 (1990), p.** 576.

[25] K. Wiesenfeld, Phys. Rev. A **32**, 1744 (1985).

[26] S. I. Turovets, J. Dellunde, and K. A. Shore, J. Opt. Soc. Am. B **14**, 200 (1997).

[27] J. Dellunde and M. C. Torrent, Appl. Phys. Lett. **68**, 1601 (1996).

Translated by R. M. Durham