Consistent treatment of color transparency effects in diffractive electroproduction of vector mesons off nuclei

O. Benhar

INFN Sezione Sanità, Physics Laboratory, Istituto Superiore di Sanità, Viale Regina Elena, 299, 1-00161 Roma, Italy

S. Fantoni

Interdisciplinary Laboratory, SISSA, and INFN, Sezione di Trieste, Via Beirut, 4, I-34014 Trieste, Italy

N. N. Nikolaev

Institut für Theoretische Kernphysik der Universität Bonn, Nussallee 14-16, D-53115, Bonn, Germany

B. G. Zakharov

L. D. Landau Institute for Theoretical Physics, 117940 Moscow, Russia

(Submitted 11 July 1996)
Zh. Éksp. Teor. Fiz. **111**, 769–786 (March 1997)

Nuclear transparency for incoherent real and virtual photoproduction of vector mesons off nuclei depends crucially on the interplay of the initial-state and final-state interactions. We develop a consistent description of initial-state and final-state interactions based on the coupled-channel multiple-scattering theory. We present detailed predictions for incoherent production of $s\bar{s}$ and $c\bar{c}$ mesons. The onset of initial-state interactions is controlled by the production length l_P, and we find strong variations of nuclear transparency when l_P rises with energy and becomes comparable to the radius of the target nucleus R_A. For $s\bar{s}$ mesons, the regime $l_P \sim R_A$ corresponds to precisely the kinematic range of the HERA–HERMES experiment, whereas for $c\bar{c}$ mesons the same condition will be met at the ENC electron–nucleus collider at GSI. In spite of the subasymptotic values of Q^2 and ν, we find a complex pattern of color transparency-induced effects whose experimental study may provide information on several issues relevant to the understanding of QCD, both in the perturbative and nonperturbative regimes. *© 1997 American Institute of Physics.* [S1063-7761(97)00103-0]

1. INTRODUCTION

The color transparency (CT) phenomenon in diffractive exclusive virtual photoproduction of vector mesons off nuclear targets has recently attracted much attention.[1–8] The Q^2-dependence of the size of the initial $q\bar{q}$ fluctuation of the virtual photon makes the measurements of nuclear transparency in these reactions a unique tool for probing the color-dipole cross section, which plays a fundamental role in low-x physics.[9–11] Measurements of nuclear transparency in vector meson electroproduction may also provide important information on the spatial wave function of vector mesons.[4] Predictions[4,5] of CT effects at asymptotic energies have been confirmed in the E665 muon scattering experiment at FERMILAB, although the statistical accuracy of the E665 data was limited.[12] Much more accurate data on nuclear transparency in virtual photoproduction of vector mesons will be obtained in the forthcoming high-luminosity experiments at CEBAF and HERA (HERMES collaboration). In these experiments the energy will be rather low, which makes a thorough theoretical analysis of CT effects at subasymptotic energies a pertinent issue.

Virtual photons are produced in the scattering of electrons (muons), and nuclear transparency for virtual photoproduction with photon energy ν and virtuality Q^2 is defined as

$$T_A=\frac{1}{A}\,\frac{d\sigma(e+A\to e'+A^*+V)}{d\sigma(e+N\to e'+N+V)}, \tag{1}$$

where A is the mass number of the target nucleus. In incoherent production off the nucleus, the nuclear cross section is summed over all excited and nuclear break-up states A^*, excluding production of secondary particles (mesons). In the single incoherent rescattering approximation, the reaction mechanism can be viewed as a formation and propagation of the projectile wave packet, the incoherent (quasi-elastic) scattering of which off the bound nucleon produces the ejectile wave packet, which propagates through the nucleus and forms the observed final state vector meson V. The nuclear transparency can be written as

$$T_A(\mathbf{q})=\frac{1}{A}\int d^3r\; n_A(\mathbf{r})\,\frac{|\langle V|\hat{S}_f(\mathbf{r})\hat{f}(\mathbf{q})\hat{S}_i(\mathbf{r})|\gamma^*\rangle|^2}{|\langle V|\hat{f}(\mathbf{q})|\gamma^*\rangle|^2}. \tag{2}$$

Here $n_A(\mathbf{r})$ is the nuclear density, $\mathbf{q}$ is the transverse momentum transfer, $\hat{S}_i(\mathbf{r})$ and $\hat{S}_f(\mathbf{r})$ are the evolution operators describing the coherent initial-state interaction (ISI) of the projectile and final-state interaction (FSI) of the ejectile in a nucleus, respectively, and $f(\mathbf{q})$ is the scattering matrix, which describes projectile-to-ejectile diffractive transitions in interaction on a free nucleon. The evolution operators $\hat{S}_{i,f}$ and the scattering matrix $\hat{f}$ in Eq. (1) act in a space including both the hadronic states and the photon state. The physical meaning of the numerator in the integrand of (2) is obvious: it is the cross section of the one-fold incoherent interaction with a bound nucleon, modified by the intranuclear coherent ISI of the projectile and FSI of the ejectile, the denominator is the photoproduction cross section on a free nucleon. Equation (1) is valid at $q^2 \lesssim 1/B$, where B is the diffraction slope for photoproduction on a free nucleon, for which the many-fold incoherent rescatterings in a nuclear medium can be neglected. For incoherent diffractive scattering of hadrons off nuclei at subasymptotic energies, the counterpart of Eq. (2) has been derived in Ref. 13. For the first application to an evaluation of CT effects in the charge exchange of pions off nuclei at asymptotic energies, see Ref. 14.

To leading order in the fine-structure constant $\alpha_{em}=1/137$, the intermediate state decomposition of the matrix element $\langle V|\hat{S}_f(\mathbf{r})\hat{f}(\mathbf{q})\hat{S}_i(\mathbf{r})|\gamma^*\rangle$ which enters into Eq. (2) is given by

$$\langle V|\hat{S}_f\hat{f}\hat{S}_i|\gamma^*\rangle=\sum_h \langle V|\hat{S}_f|h\rangle\langle h|\hat{f}|\gamma^*\rangle+\sum_{h,h'} \langle V|\hat{S}_f|h'\rangle \times\langle h'|\hat{f}|h\rangle\langle h|\hat{S}_i|\gamma^*\rangle, \tag{3}$$

where $|h\rangle,|h'\rangle$ are the intermediate hadronic states, and we have taken into account that to leading order in α_{em}, one can set $\langle\gamma^*|\hat{S}_i|\gamma^*\rangle=1$. The first term in the decomposition (3) describes incoherent scattering at the level of production of the intermediate state $|h\rangle$ at the electromagnetic interaction vertex, and is described by the amplitude $\langle h|\hat{f}|\gamma^*\rangle$. ISI effects are present only in the second term, in which incoherent scattering takes place at the level of intermediate states $|h\rangle$ and $|h'\rangle$ and is described by the amplitude $\langle h'|\hat{f}|h\rangle$. The coherency properties of ISI of the projectile and FSI of the ejectile in virtual photoproduction of vector mesons are characterized by two different length scales. For ISI, the relevant scale is the production length of the $q\bar{q}$ pair by the virtual photon,

$$l_P\sim\frac{2\nu}{Q^2+m_V^2}, \tag{4}$$

while for FSI, the relevant scale is the formation length of the wave function of the final vector meson,

$$l_F\sim\frac{2\nu}{m_{V'}^2-m_V^2}, \tag{5}$$

where $m_{V,V'}$ are the masses of the $1S$ and $2S$ vector mesons.[1,2]

The strength of ISI effects critically depends on the ratio l_P/R_A, where R_A is the nuclear radius. At sufficiently low energy and/or large Q^2, when $l_P \ll R_A$, the off-diagonal matrix element $\langle h|\hat{S}_i|\gamma^*\rangle$ is suppressed due to the large longitudinal momentum transfer, $\sim 1/l_P$, in the coherent $\gamma^*\to q\bar{q}$ transition. As a result, the second term in Eq. (2) related to ISI can be neglected. In this case the nuclear effects are exhausted by FSI, which in turn depends crucially on the ratio l_F/R_A. In Ref. 2 this regime for $c\bar{c}$ vector meson photoproduction was considered within the Green's function approach to FSI. This technique was further used in Ref. 8 for evaluation of nuclear transparency in $s\bar{s}$ vector meson electroproduction for the energy range of future experiments at CEBAF. At $l_P \gtrsim R_A$, ISI and FSI must be taken into account on the same footing. In the limit of high energy, when $l_{P,F} \gg R_A$, ISI and FSI absorption effects can be evaluated within the frozen-size approximation for propagation of the $q\bar{q}$ fluctuation of the virtual photon through the nuclear medium.[2]

The intermediate energy region when $l_{P,F}$ are comparable with the nucleus size is the most complicated for evaluation of CT effects. In Ref. 3 an approximate formula for nuclear transparency, interpolating between the high- and low-energy regimes was obtained within the two-channel model. Because of the relatively weak nuclear attenuation in J/Ψ photoproduction, this interpolation formula works well, and the theoretical results of Ref. 3 agree with the experimental data of NMC collaboration.[15] A more rigorous treatment of CT effects at $l_{P,F}\sim R_A$ is highly desirable, because the kinematic conditions at the CEBAF and HERA–HERMES experiments on production of light vector mesons correspond precisely to this situation. In the energy range of the ENC electron–nucleus collider which is being planned at GSI, one will encounter the case $l_P\sim R_A$ in the production of $J/\Psi,\Psi'$ mesons off nuclei. (At much higher energies of HERA operating in the electron–nucleus collider mode, one will have the asymptotic situation $l_P, l_F \gg R_A$, which requires a separate analysis.) The theoretical basis for evaluation of CT effects without restrictions on the values of $l_{P,F}$ is the Glauber–Gribov coupled-channel multiple-scattering theory (CCMST).[16,17] The formalism of CCMST allows for both evolution and absorption effects in propagation of the projectile state through nuclear medium.

The interplay of ISI and FSI effects is also important in diffractive hadroproduction off nuclei. Because in hadronic interactions one always has $l_P \approx l_F$, the ISI and FSI effects cannot be separated. The virtuality of the photon, Q^2, which can easily be varied experimentally over a broad range, gives an important handle on the values of l_P and the strength of ISI, which makes the real and virtual photoproduction of vector mesons a unique testing ground of the coherency properties of diffractive production off nuclei. In the present paper we carry out a CCMST analysis of CT effects in ϕ, ϕ' and $J/\Psi,\Psi'$ electroproduction.

2. ISI AND FSI EVOLUTION OPERATORS IN THE COUPLED-CHANNEL MULTIPLE-SCATTERING FORMALISM

We use for evaluation of the ISI and FSI evolution operators $\hat{S}_{i,f}$ an extension of the formalism previously developed[18] for calculating FSI effects in nuclear transparency for quasielastic $A(e,e'p)$ scattering. As in the case of the FSI evolution operator for the ejectile state in $(e,e'p)$ scattering, the operators $\hat{S}_i$ and $\hat{S}_f$ within CCMST can be written in z-ordered operator exponential form

$$\hat{S}_i(\mathbf{r})=\hat{P}_z\,\exp\left[-\frac{1}{2}\int_{-\infty}^{z}d\xi\;\hat{\sigma}(\xi-z)n_A(\mathbf{b},\xi)\right], \tag{6}$$

$$\hat{S}_f(\mathbf{r})=\hat{P}_z\ \exp\left[-\frac{1}{2}\int_z^\infty d\xi\ \hat{\sigma}(\xi-z)n_A(\mathbf{b},\xi)\right], \tag{7}$$

where $\mathbf{r}=(\mathbf{b},z)$, and $\hat{P}_z$ is the z-ordering operator. The matrix elements of the z-dependent operator $\hat{\sigma}(z)$ are related to the diffraction scattering matrix at $\mathbf{q}=0$:

$$\langle i|\hat{\sigma}(z)|j\rangle=-i\ \exp(ik_{ij}z)\langle i|\hat{f}(\mathbf{q}=0)|j\rangle, \tag{8}$$

where k_{ij} is the longitudinal momentum transfer associated with the transition $iN\rightarrow jN$,[17]

$$k_{ij}=\frac{m_i^2-m_j^2}{2\nu}, \tag{9}$$

m_i and m_j are the masses of the states $|i\rangle$ and $|j\rangle$, and for the virtual photon $m_{\gamma^*}^2=-Q^2$. The exponential phase factor in Eq. (8) results from the difference between the phases for the plane waves describing the states $|i\rangle$ and $|j\rangle$ after propagating the distance z. It is easy to check that the full phase factors, which the operators (6) and (7) yield in the case of arbitrary sequences of intermediate states for transitions $\gamma^*\rightarrow h$ and $h\rightarrow V$, coincide with the phase factors that can be obtained by solving the set of coupled-channel wave equations.

For numerical calculations, it is convenient to treat the off-diagonal part of matrix $\hat{\sigma}(z)$ in (6) and (7) as a perturbation. Then we can represent the matrix elements of the operators $\hat{S}_{i,f}(\mathbf{r})$ in the form of ν-fold off-diagonal rescattering series:

$$\langle h|\hat{S}_i(\mathbf{r})|\gamma^*\rangle=\sum_{\nu=1}^{\infty}\ \langle h|\hat{S}_i^{(\nu)}(\mathbf{r})|\gamma^*\rangle, \tag{10}$$

$$\langle V|\hat{S}_f(\mathbf{r})|h\rangle=\sum_{\nu=0}^{\infty}\ \langle V|\hat{S}_f^{(\nu)}(\mathbf{r})|h\rangle, \tag{11}$$

where

$$\langle h|\hat{S}_i^{(1)}(\mathbf{b},z)|\gamma^*\rangle=-\frac{1}{2}\ \sigma_{h\gamma^*}\ \exp(ik_{\gamma^*h}z)\int_{-\infty}^{z}dz_1\ n_A(\mathbf{b},z_1)\times\exp\left[ik_{h\gamma^*}z_1-\frac{1}{2}\ t(\mathbf{b},z,z_1)\sigma_{hh}\right], \tag{12}$$

$$\begin{aligned}\langle h|\hat{S}_i^{(\nu)}(\mathbf{b},z)|\gamma^*\rangle=&\left(-\frac{1}{2}\right)^{\nu}\sum_{i_1,\dots i_{\nu-1}}\sigma'_{hi_{\nu-1}}\sigma'_{i_{\nu-1}i_{\nu-2}}\cdots\sigma_{i_1\gamma^*}\\ &\times\exp(ik_{\gamma^*h}z)\int_{-\infty}^{z}dz_\nu\ n_A(\mathbf{b},z_\nu)\\ &\times\exp\left[ik_{hi_{\nu-1}}z_\nu-\frac{1}{2}\ t(\mathbf{b},z,z_\nu)\sigma_{hh}\right]\\ &\times\int_{-\infty}^{z}dz_{\nu-1}\ n_A(\mathbf{b},z_{\nu-1})\\ &\times\exp\left[ik_{i_{\nu-1}i_{\nu-2}}z_{\nu-1}\right.\\ &\left.-\frac{1}{2}\ t(\mathbf{b},z_\nu,z_{\nu-1})\sigma_{i_{\nu-1}i_{\nu-1}}\right]\cdots\\ &\int_{-\infty}^{z_2}dz_1\ n_A(\mathbf{b},z_1)\exp\left[ik_{i_1\gamma^*}z_1-\frac{1}{2}\ t(\mathbf{b},z_2,z_1)\sigma_{i_1i_1}\right],\quad \nu\geqslant 2,\end{aligned}$$

$$\langle V|\hat{S}_f^{(0)}(\mathbf{b},z)|h\rangle=\delta_{Vh}\ \exp\left[-\frac{1}{2}\ t(\mathbf{b},\infty,z)\sigma_{VV}\right], \tag{14}$$

$$\begin{aligned}\langle V|\hat{S}_f^{(\nu)}(\mathbf{b},z)|h\rangle=&\left(-\frac{1}{2}\right)^{\nu}\sum_{i_1,\dots i_{\nu-1}}\sigma'_{Vi_{\nu-1}}\sigma'_{i_{\nu-1}i_{\nu-2}}\cdots\sigma'_{i_1h}\\ &\times\exp(ik_{hV}z)\int_z^\infty dz_1\ n_A(\mathbf{b},z_1)\\ &\times\exp\left[ik_{i_1h}z_1\right.\\ &\left.-\frac{1}{2}\ t(\mathbf{b},z_1,z)\sigma_{hh}\right]\\ &\times\int_{z_1}^{\infty}dz_2\ n_A(\mathbf{b},z_2)\exp\left[ik_{i_2i_1}z_2\right.\\ &\left.-\frac{1}{2}\ t(\mathbf{b},z_2,z_1)\sigma_{i_1i_1}\right]\cdots\int_{z_{\nu-1}}^{\infty}dz_\nu\\ &\times n_A(\mathbf{b},z_\nu)\exp\left[ik_{Vi_{\nu-1}}z_\nu\right.\\ &\left.-\frac{1}{2}\ t(\mathbf{b},\infty,z_\nu)\sigma_{VV}\right],\quad \nu\geqslant 1.\end{aligned} \tag{15}$$

Here $\sigma'_{ik}=\sigma_{ik}-\delta_{ik}\sigma_{ii}$, the matrix $\hat{\sigma}$ is related to the forward diffraction scattering matrix, $\hat{f}(\mathbf{q}=0)=i\hat{\sigma}$, and $t(\mathbf{b},z_2,z_1)=\int_{z_1}^{z_2}dz\ n_A(\mathbf{b},z)$ is the partial optical thickness.

The ISI and FSI evolution operators comprise two effects: the first can be called nuclear filtering of the projectile and ejectile wave packets due to the difference in attenuation factors $\exp[-(1/2)t(\mathbf{b},z_2,z_1)\sigma_{ii}]$, the second is the space-time evolution of these wave packets resulting from off-diagonal coherent rescattering in a nuclear medium. The first-order ISI term $\langle h|\hat{S}_i^{(1)}(\mathbf{b},z)|\gamma^*\rangle$ of Eq. (12) and zeroth-order FSI term $\langle V|S_f^{(0)}(\mathbf{b},z)|h\rangle$ of Eq. (11) describe the conventional Glauber ISI and FSI absorption, respectively. In this Glauber approximation, only the nuclear filtering effects are included. The terms (13) and (15) give corrections to ISI and FSI associated with the off-diagonal rescattering of the intermediate hadronic states in a nuclear medium, and describe the space-time evolution of the projectile and ejectile wave packets due to intranuclear interactions. It is precisely the oscillating exponential phase factors in Eqs. (13) and (15) that lead to suppression of the contributions of the off-diagonal inelastic intermediate states at low energies of the virtual photon. Notice that the suppression of ISI at $l_P\lesssim R_A$ discussed above is also related to the oscillating exponential phase factor $\exp(ik_{h\gamma^*}z_1)$ in the integrand in Eq. (12). Equations (2) and (10)–(15) form the basis for numerical evaluation of nuclear transparency in incoherent vector meson and electroproduction in the framework of CCMST.

3. CALCULATION OF THE DIFFRACTION SCATTERING MATRIX

The ISI and FSI operators $\hat{S}_{i,f}$ describing the evolution of the $q\bar{q}$ state as it propagates through the nucleus depend critically on the form of the diffraction matrix. In our analysis we describe the $q\bar{q}$ states in the nonrelativistic oscillator model. Then, assuming the dominance of Pomeron exchange, the diffraction matrix elements σ_{ik} can be written as

$$\langle i|\hat{\sigma}|k\rangle=\int d^2\boldsymbol{\rho}\, dz\Psi_i^*(\boldsymbol{\rho},z)\sigma(\rho)\Psi_k(\boldsymbol{\rho},z), \qquad (16)$$

where $\boldsymbol{\rho}$ is the transverse size of the $q\bar{q}$ pair, $\Psi_{i,k}(\boldsymbol{\rho},z)$ are the wave functions describing the $q\bar{q}$ states, and $\sigma(\rho)$ is the dipole cross section describing the interaction of the $q\bar{q}$ pair with a nucleon.

We need also the matrix element for the diffraction excitation of the virtual photon into a $q\bar{q}$ state on the free nucleon. Following Ref. 2, we use the perturbative $q\bar{q}$ light-cone wave function of the virtual photon[9] and write the matrix elements $\langle i|\hat{\sigma}|\gamma^*\rangle$ in the form

$$\langle i|\hat{\sigma}|\gamma^*\rangle=\lambda\int d^2\boldsymbol{\rho}\, dz\Psi_i^*(\boldsymbol{\rho},z)\sigma(\rho)K_0(\epsilon\rho) \times\exp\left(-\frac{z^2}{d^2}\right), \qquad (17)$$

where

$$\epsilon^2=m_q^2+Q^2/4, \qquad (18)$$

m_q is the quark mass, $K_0(x)$ is the modified Bessel function, $d=1/2m_q$ and λ is a normalization coefficient that is immaterial for the evaluation of nuclear transparency. Equation (14) yields the diffraction amplitude for transition of the virtual photon into a $q\bar{q}$ state with the sum of quark helicities equal to the photon helicity.[2] This amplitude dominates for production of nonrelativistic $q\bar{q}$ states and/or for the moderately large $Q^2\lesssim m_V^2$ considered in the present paper.

In our analysis we restrict ourselves to calculation of nuclear transparency at small momentum transfer. Then, we can use Eqs. (16) and (17) to calculate the scattering matrix $\hat{f}$ in the numerator and denominator of the integrand in Eq. (2). At small momentum transfer; rescatterings of the $q\bar{q}$ state on nucleons do not change the projection of its orbital momentum on the z-axis. As a result, in the oscillator model, the intermediate $q\bar{q}$ states emerging in the coupled-channel formalism are exhausted by the transverse excitations of the $q\bar{q}$ system with zero value of the azimuthal quantum number. The masses of these $q\bar{q}$ states entering into Eq. (9) for longitudinal momentum transfer can be written as $m_i=m_0+2i\omega_q$, where ω_q is the oscillator frequency. For the oscillator frequencies of the $s\bar{s}$ and $c\bar{c}$ systems we use the values $\omega_s=(m_{\phi'}-m_{\phi})/2=0.33$ GeV and $\omega_c=(m_{\Psi'}-m_{J/\Psi})/2=0.3$ GeV. For the quark masses we take $m_s=0.5$ GeV and $m_c=1.5$ GeV.

The form of the diffraction matrix is sensitive to the ρ-dependence of the dipole cross section. The available experimental data on the structure function F_2 at low x and the cross sections of vector meson electroproduction off nucleons can be described by representing the dipole cross section as a sum of the energy-dependent perturbative and the energy-independent nonperturbative components.[11]

The energy dependence of the perturbative part of $\sigma(\rho)$ generated by the higher Fock states is governed by the generalized BFKL equation.[19] The analysis of CT effects at high energy in the regime of large contribution to the dipole cross section of higher Fock states requires an accurate treatment of the nuclear filtering effects for the many-body parton components of the projectile and ejectile. The interplay of absorption effects for the different Fock states may be important to the energy dependence of nuclear transparency in the limit $l_P\gg R_A$.

In the present paper we restrict ourselves to evaluation of CT effects at relatively low energies, in which the higher Fock states do not greatly affect nuclear transparency. We use for $\sigma(\rho)$ the parametrization given by the two-gluon exchange model of the pomeron,[20,21] which yields the energy-independent dipole cross section

$$\sigma(\rho)=\frac{16}{3}\int d^2\mathbf{k} \times\frac{\alpha_S(\max(C/\rho,k))\,\alpha_S(k)[1-\exp(i\mathbf{k}\boldsymbol{\rho})][1-G_2(\mathbf{k},-\mathbf{k})]}{(k^2+\mu_g)^2}, \qquad (19)$$

where $G_2(\mathbf{k}_1,\mathbf{k}_2)=\langle N|\exp(i\mathbf{k}_1\mathbf{r}_1+i\mathbf{k}_2\mathbf{r}_2)|N\rangle$ is the two-quark form factor of the nucleon, μ_g is an effective gluon mass, and $C\simeq1.5$.[9] We calculated the dipole cross section making use of the running and freezing coupling $\alpha_S(k)$ according to the prescription of Ref. 9. The value of α_S was frozen at the momenta $k\leqslant k_{f_r}=0.7$ GeV. At higher momenta, the running α_S was calculated through the one-loop formula with $\Lambda_{\mathrm{QCD}}=0.3$. As in Ref. 22, we use for the effective gluon mass value $\mu_g=0.17$ GeV. The ρ-dependence of the dipole cross section calculated in the two-gluon model is shown in Fig. 1. Calculation of $\sigma_{\mathrm{tot}}(\pi N)=\langle\pi|\hat{\sigma}|\pi\rangle$ with our parametrization of the dipole cross section reproduces the experimental value of the pion–nucleon total cross section in the 10–100 GeV energy range.

The decrease in dipole cross section $\sigma(\rho)$ at perturbatively small ρ given by the parametrization (19) agrees well with the behavior of the dipole cross section extracted recently from experimental data on vector meson electroproduction off free nucleons for the range of dipole size $\rho\sim0.2$–1.5 fm.[23] The ρ-dependence of the dipole cross section given by Eq. (19) was previously corroborated by the analysis of CT effects in the quasielastic charge-exchange reaction $\pi^-=A\rightarrow\pi^0A'$,[14] and of the nuclear shadowing and diffraction cross section in deep inelastic scattering.[9,24]

At this point, it must be made clear that the above choice of gluon mass is oriented toward description of the nonperturbative part of the dipole cross section, and μ_g is a phenomenological parameter that must not be taken at face value. The analysis of low-x HERA data on the structure function F_2[11] within the generalized BFKL equation,[19] and the nonperturbative evaluation of the gluon correlation radius[25] yield clearcut evidence for an effective gluon mass ~0.75 GeV. The value of μ_g used in the present paper does not affect the dipole cross section at perturbatively small ρ

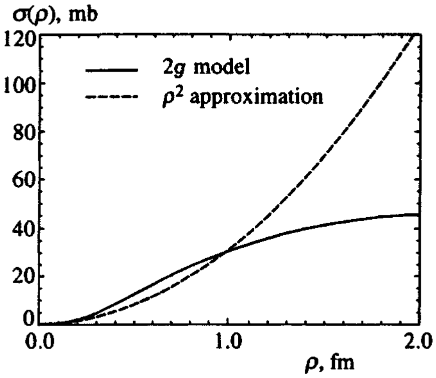


FIG. 1. The dipole–size dependence of color dipole cross sections for the parametrizations (19) (solid curve) and (21) (dashed curve).

$\ll 1/\mu_g$, but it provides a viable parametrization of the experimental information on the dipole cross section in the semiperturbative and nonperturbative ranges of larger ρ, which really cannot be described in perturbative QCD.

Making use of the parametrization (19) for $\sigma(\rho)$, with the above set of the quark masses and oscillator frequencies, we find the following total interaction cross sections for vector mesons with a free nucleon: $\sigma_{\mathrm{tot}}(\phi N) \approx 16$ mb, $\sigma_{\mathrm{tot}}(\phi' N) \approx 24.5$ mb, $\sigma_{\mathrm{tot}}(J/\Psi N) \approx 7.5$ mb, and $\sigma_{\mathrm{tot}}(\Psi' N) \approx 14.5$ mb. For the ratio $\langle V(2S)|\hat{\sigma}|\gamma^*\rangle / \langle V(1S)|\hat{\sigma}|\gamma^*\rangle$ at $Q^2=0$, we obtain 0.186 and 0.49 for the $s\bar{s}$ and $c\bar{c}$ mesons, respectively. Notice that the resulting ratio $R(2S/1S)=\sigma(\gamma^* N \to \Psi' N)/\sigma(\gamma^* N \to J/\Psi N) \approx 0.24$ predicted for $Q^2=0$ agrees well with the result of the NMC collaboration, $R(2S/1S)=0.22 \pm 0.05$,[26] and $R(2S/1S)=0.21 \pm 0.02$ from the E687 collaboration.[27]

In concluding this section it is appropriate to discuss the scanning phenomenon and the effect of the nodal structure of the spatial wave function of radially excited $2S$ vector mesons on the amplitude of the transition $\gamma^* \to V(2S)$. The transverse spatial $q\bar{q}$ wave function of the virtual photon, $\propto K_0(\epsilon\rho)$, which enters into the matrix element (17), decreases proportionally to $\exp(-\epsilon\rho)$ at large $\rho \gtrsim 1/\epsilon \sim 1/\sqrt{m_q^2+Q^2/4}$; see Eq. (19). Due to the vanishing of the dipole cross section at $\rho=0$, the typical size of the $q\bar{q}$ pair that dominates the matrix element (17), the so-called scanning radius r_S,[4] turns our to be considerably greater than the naively expected value $\sim 1/\sqrt{m_q^2+Q^2/4}$, and is given by

$$r_S \approx \frac{6}{\sqrt{Q^2+m_V^2}} \tag{20}$$

(here we put $2m_q \approx m_V$). At low Q^2, r_S is close to the radius r_V of the $1S$ state and to the position of the node, $r_n \sim r_V$, in the spatial wave function of the radially excited $2S$ vector mesons. This leads to considerable cancellation between the contributions to the amplitude $\langle V(2S)|\hat{\sigma}|\gamma^*\rangle$ coming from the regions of large ($\rho \gtrsim r_n$) and small ($\rho \lesssim r_n$) dipole size.[3,28] Precisely this cancellation is responsible for the cited suppression of the cross section of photoproduction of $2S$ vector mesons as compared to $1S$ vector mesons. Similar suppression of the production amplitude $\langle V(nS)|\hat{\sigma}|\gamma^*\rangle$ holds for higher states $V(nS)$.

The effect is extremely strong for light vector mesons. For the diffraction amplitudes $\langle V(2S)|\hat{\sigma}|\gamma^*\rangle$ for light vector

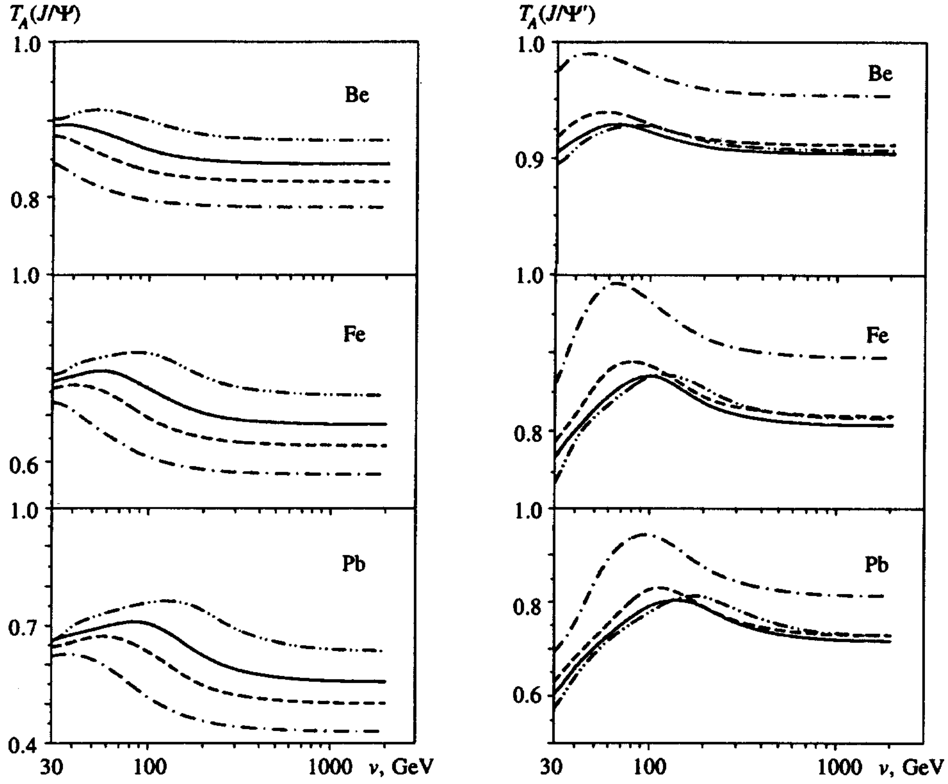


FIG. 2. Predictions for nuclear transparency in incoherent production of $J/\Psi(1S)$ and $\Psi'(2S)$ mesons off nuclei for the color dipole cross section of Eq. (19). The dot-dashed, dashed, solid, and dot-dot-dashed curves are for $Q^2=0$, 5, 10, and 20 GeV^2, respectively.

mesons at low Q^2, one can have either the undercompensation scenario, with dominance of the small-ρ contribution, or the overcompensation scenario with dominance of the large-ρ contribution. Which scenario is realized can only be decided experimentally. To this end, the nuclear filtering effects drastically affect the cancellation between the contributions of the regions of large and small interquark distances to the $2S$ vector meson photoproduction amplitudes. The analysis of ρ' photoproduction in the frozen-size approximation appropriate to the high-energy regime in Ref. 6 shows that the strong nodal effect for light $2S$ vector mesons can lead to an anomalous A-dependence of nuclear transparency.

The behavior of nuclear transparency for light $2S$ vector mesons at low Q^2 differs strongly for the undercompensation and overcompensation scenarios. In our model for the dipole cross section and for the wave functions of vector mesons, the undercompensation regime takes place in real photoproduction of the ϕ' and Ψ' on the free nucleon at $Q^2=0$. In contrast to Ψ', for the ϕ' the compensation effect is very strong, and at present the overcompensation regime for ϕ' photoproduction also cannot be ruled out. A recent analysis[23] of vector meson photoproduction on free nucleons in the light-cone wave function formalism shows that for photoproduction of light $2S$ vector mesons by longitudinal photons, relativistic effects can lead to the overcompensation scenario at low Q^2.

In order to illustrate the behavior of nuclear transparency in the overcompensation scenario, we also compute nuclear transparency, which makes for a parametrization of $\sigma(\rho)$ in the form

$$\sigma(\rho)=\rho^2\,\frac{\sigma_{\text{tot}}(\phi N)}{\langle\phi|\rho^2|\phi\rangle}, \tag{21}$$

which enhances the large-ρ contribution and leads to the overcompensation scenario for ϕ' photoproduction, although this parametrization is somewhat unrealistic, in view of the available experimental information on the dipole cross section $\sigma(\rho)$ analyzed in Ref. 23. The normalization in (21) has been so chosen as to produce approximately the same results for $T(\phi)$, while emphasizing the overcompensation effects in $T(\phi')$.

4. NUMERICAL RESULTS

We carried out our numerical calculations in the range $Q^2 \lesssim 2m_V^2$. The number of included $q\overline{q}$ resonance states and the multiplicity of the off-diagonal rescatterings used in Eqs. (7)–(12) were equal to 5 and 3, respectively. We checked that the contributions from higher excitations and rescatterings with $\nu > 3$ are negligible in the above range of Q^2, even at high energies, when the suppression of the higher mass states related to the longitudinal momentum transfer vanishes. In the above range of Q^2, predictions of the CCMST are saturated by the contribution of several lowest resonant $q\overline{q}$ states for two related-reasons: the suppression of the diffraction matrix element (16) for transitions with large difference between the masses of the initial and final states, and the similar suppression of matrix elements (17) for virtual photoexcitation of higher states $|h\rangle$ for a relatively large value of the scanning radius (20).

In the present paper, we focus on predictions for nuclear transparency in ϕ, ϕ', J/Ψ, and Ψ' virtual photoproduction on the target nuclei ^{9}Be, ^{66}Fe, and ^{207}Pb. For the nuclear matter density in the target nucleus ^{9}Be we use the oscillator shell model nuclear density, with the oscillator frequency adjusted to reproduce the experimental rms radius of the charge distribution $\langle r^2\rangle^{1/2}_{^9\text{Be}}=2.51$ fm.[29] For the target nucleus ^{56}Fe, the parametrization of the nuclear density in terms of a sum of Gaussians from Ref. 29 was used. For ^{207}Pb the Wood–Saxon parametrization of the nuclear density with parameters taken from Ref. 29 was used.

The results obtained for the dipole cross section parametrized by Eq. (19) are shown in Figs. 2 and 3. At low energies, the magnitude of nuclear transperencies $T_A(\phi)$ and $T_A(J/\Psi)$ is controlled by FSI attenuation of the ejectile. The decrease in $T_A(\phi)$ and $T_A(J/\Psi)$ at higher energies is due to the onset of attenuation for ISI effects. The energy ν at which $T_A(\phi)$ and $T_A(J/\Psi)$ start dropping is higher for larger Q^2 and for heavier vector mesons, which nicely correlates with the condition $l_P \gtrsim R_A$, i.e. $\nu \gtrsim R_A(m_V^2 + Q^2)$ for fully-developed ISI. This ISI-driven drop in nuclear transparency is preceded by a rise in $T_A(\phi)$ and $T_A(J/\Psi)$ at lower energies. We shall comment more on the origin of this rise below, when discussing the results for the Glauber approximation to the ISI and FSI operators shown below in Figs. 6 and 7. The asymptotic values of $T_A(\phi)$ and $T_A(J/\Psi)$ rise with Q^2, which is the CT effect. In the frozen size approximation, appropriate at high energies, this diminishing attenuation derives from the decrease in the scanning radius r_S with Q^2; see Eq. (20).

The preasymptotic rise in nuclear transparency is much stronger for the $2S$ vector mesons. Also for the $2S$ states, the onset of ISI is followed by a decrease in $T_A(\phi')$ and $T_A(\Psi')$ at higher energies. At small Q^2, when the scanning radius r_S is larger and the node effect is stronger, both $T_A(\phi')$ and $T_A(\Psi')$ exhibit rapid Q^2 dependence, faster than for $T_A(\phi)$ and $T_A(J/\Psi)$, respectively. At much higher Q^2, when the scanning radius r_S is substantially smaller than the radius r_V of the $1S$ states, the role of the node effect diminishes and we find $T_A(\phi')\approx T_A(\phi)$ and $T_A(\Psi')\approx T_A(J/\Psi)$ (for more discussion on this point see Ref. 4). Nuclear transparency for the $2S$ states peak in the range $\nu\sim 10-30$ GeV for ϕ' photoproduction and $\nu\sim 50-150$ GeV for Ψ' photoproduction. Our results show that despite the inequality $\sigma_{\text{tot}}(V'N) > \sigma_{\text{tot}}(VN)$, in the kinematic region studied, the counterintuitive inequality $T_A(V(2S)) > T_A(V(1S))$ is predicted. For real ϕ' photoproduction $(Q^2=0)$, we predict a considerable antishadowing effect. Of course, one must bear in mind that due to the strong node effect in ϕ' photoproduction discussed above, the theoretical predictions for $T_A(\phi')$ at low Q^2 are very model-dependent.

We consider the above predictions for nuclear transparency in ϕ' photoproduction as only an illustration of the energy- and Q^2-dependence of nuclear transparency expected for the undercompensation scenario. Notice that much of the interesting energy dependence of $T_A(\phi)$ and $T_A(\phi')$

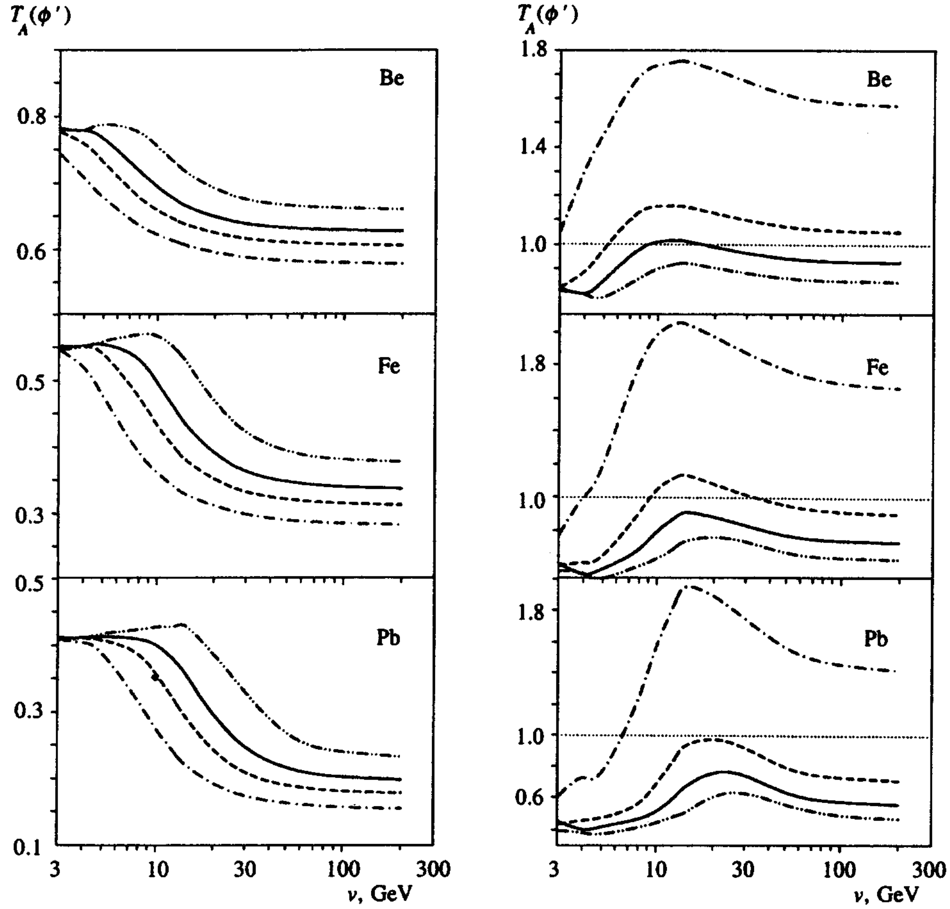


FIG. 3. Predictions for nuclear transparency in incoherent production of the $\phi(1S)$ and $\phi'(2S)$ mesons off nuclei for the color dipole cross section of Eq. (19). The dot-dashed, dashed, solid, and dot-dot-dashed curves are for $Q^2=0$, 0.5, 1.0, and 2.0 GeV^2, respectively.

takes place at $\nu \sim 10-30$ GeV, which is precisely the kinematic range of the HERA–HERMES experiment. Similarly, the interesting variations of $T_A(J/\Psi)$ and $T_A(\Psi')$ take place at $\nu \sim 30-200$ GeV, which are in the kinematic range of the ENC collider.

In Fig. 4 we compare the theoretical predictions obtained for the dipole cross section (19) with the experimental data on the ratio of nuclear transparencies for tin and carbon targets at $Q^2=0$ from the NMC collaboration.[15] The agreement between theory and experiment is good. Here we note that the results obtained in the full CCMST for the energy dependence of the ratio $T_{\mathrm{Sn}}(J/\Psi)/T_C(J/\Psi)$ turn out to be very close to the evaluation of the energy dependence from the approximate two-channel extrapolation formula for nuclear transparency suggested in Ref. 3.

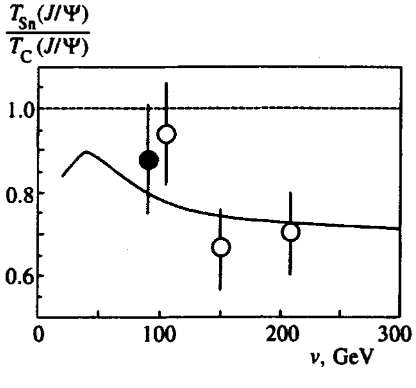


FIG. 4. Predictions for the ratio of nuclear transparencies in incoherent photoproduction of the J/Ψ off Sn and C targets for the dipole cross section of Eq. (19), in comparison with experimental data from the NMC muon scattering experiment. The open and full circles are for muon energy 280 GeV and 200 GeV, respectively.[15]

In Figs. 5 and 6 we show the nuclear transparency obtained with the quadratic parametrization of the dipole cross section (21). As stated above, this parametrization was so devised as to enforce the overcompensation scenario, and to have a negative amplitude for $\gamma^* \to \phi'$ transition at low Q^2. For Ψ' production, the undercompensation regime is retained. In the undercompensation scenario, nuclear transparency $T_A(\phi')$ decreases monotonically in the range of Q^2 considered.

In the overcompensation scenario, $T_A(\phi')$ rises with Q^2 at low Q^2, which is mostly due to the decrease in the denominator in Eq. (2) with rising Q^2. In our simplified model in which we neglect the small real part of the pomeron amplitudes, $T_A(\phi')$ even becomes infinite at a certain value of Q^2, at which the denominator in Eq. (2) vanishes and then decreases with Q^2, following the pattern for the undercompensation regime.

In a more realistic model, the infinity of $T_A(\phi')$ is removed, and one instead finds a maximum in $T_A(\phi')$ if one takes into account the nonzero real part of the amplitude for $\gamma^* \to \phi'$ transition. We shall not dwell on this, but merely emphasize that an experimental observation of such a non-

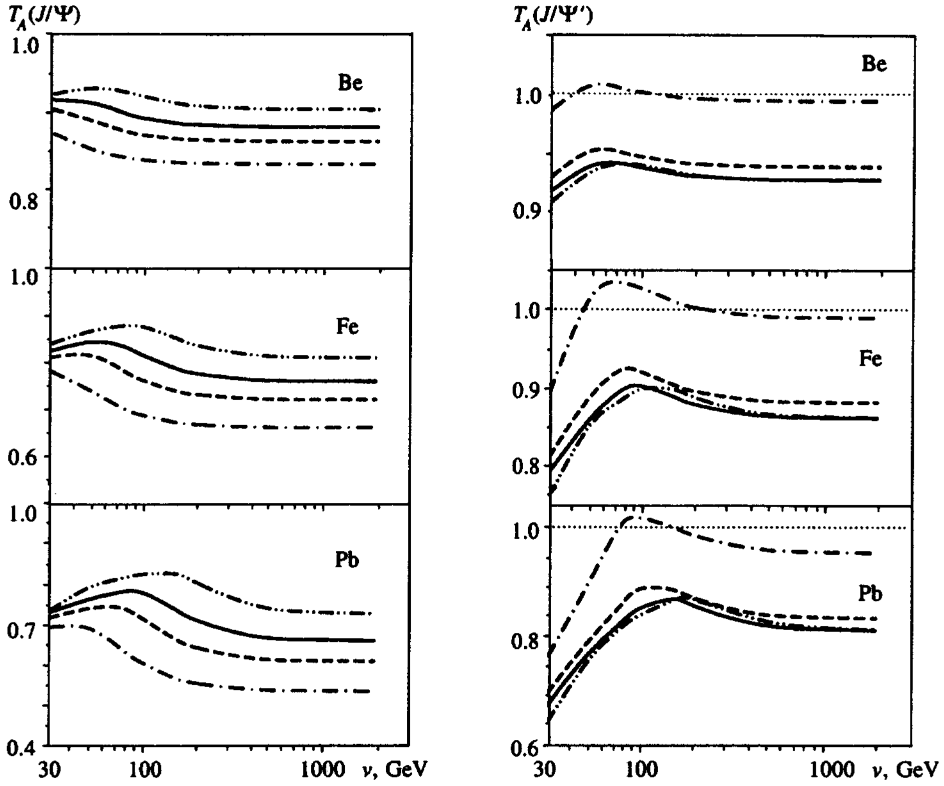


FIG. 5. The same as Fig. 2, but for the dipole cross section of Eq. (21).

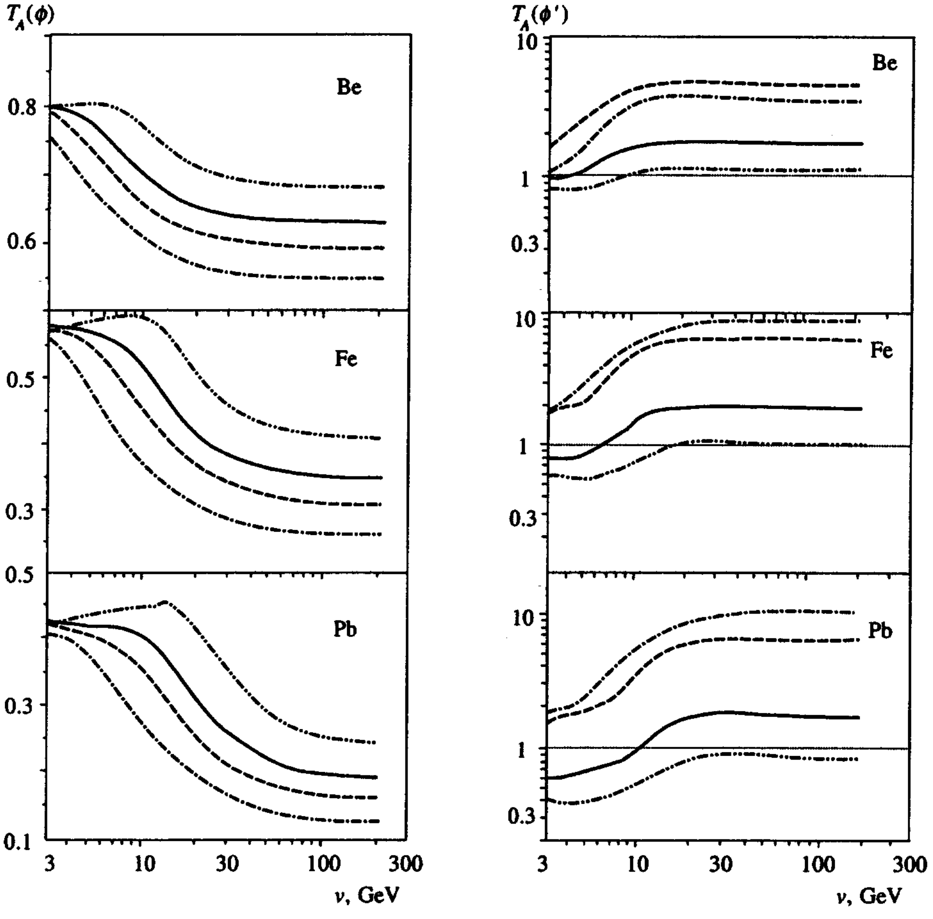


FIG. 6. The same as Fig. 3, but for the dipole cross section of Eq. (21).

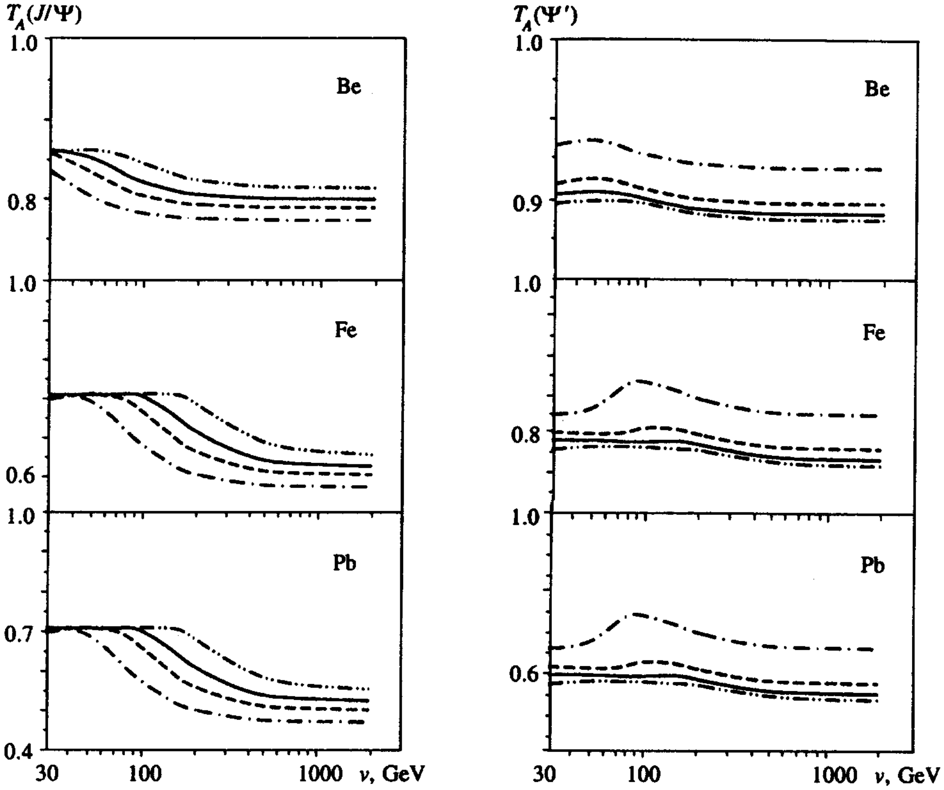


FIG. 7. The same as Fig. 2, but in the Glauber approximation as described in the text. The dipole cross section of Eq. (19) is used.

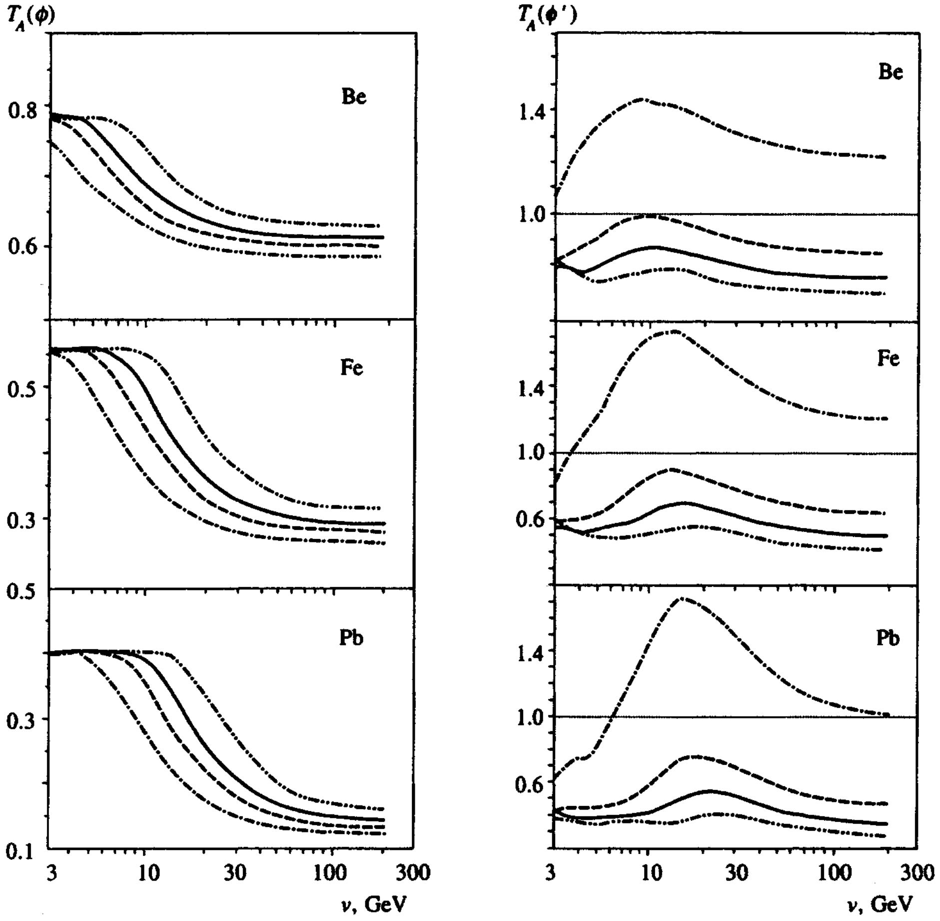


FIG. 8. The same as Fig. 3, but in the Glauber approximation as described in the text. The dipole cross section of Eq. (19) has been used.

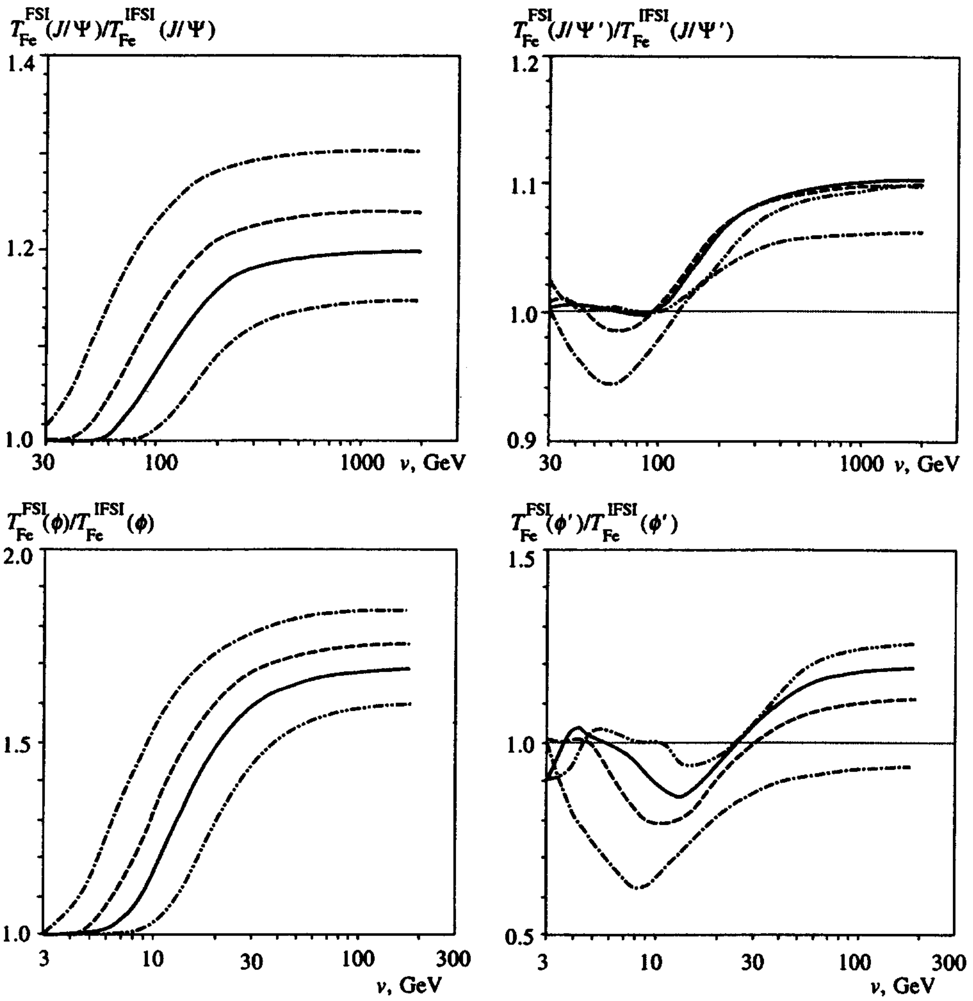


FIG. 9. The ratio of nuclear transparancies for incoherent production of the J/Ψ, Ψ', ϕ, and ϕ' mesons off nuclei, allowing only for final state interaction and neglecting the initial state interaction (FSI) and with both initial and final state interactions included (IFSI). The printed patterns for J/Ψ and Ψ' mesons are the same as in Fig. 2, and for ϕ and ϕ' mesons they are the same as in Fig. 3. The dipole cross section of Eq. (19) has been used.

monotonic Q^2-dependence of $T_A(\phi')$ would be clearcut evidence for the overcompensation scenario in ϕ' electroproduction.

For the 1*S* states, we find only marginal changes in $T_A(J/\Psi)$ and $T_A(\phi)$ from the ones shown in Figs. 3 and 4 for the dipole cross section (19). For the J/Ψ, Fig. 5 shows slightly weaker attenuation than in Fig. 3 which is obvious from the fact that the dipole cross section (21) is smaller than the parametrization (19) in the range of ρ relevant to J/Ψ production. The cross sections (19) and (21) differ substantially only at $\rho \gtrsim 1$ fm, and the changes in $T_A(\Psi)$ from one parametrization to another are also marginal. The node effect is somewhat enhanced, though, and the inequality $T_A(\Psi') > T_A(J/\Psi)$ becomes stronger, with $T_A(\Psi')$ even reaching the antishadowing regime of $T_A(\Psi')$ for $Q^2 \sim 0$. Finally, the CT effect derives from the variation of $\sigma(\rho)$ with ρ. Because this variation is stronger for the parametrization (21), the Q^2 dependence of $T_A(J/\Psi)$ and $T_A(\phi)$ in the high-energy regime of the developed ISI is somewhat stronger for the parametrization (21) than for the parametrization (19), in particular for the $T_A(\phi)$; cf. Fig. 6 and Fig. 3.

The rise of $T_A(J/\Psi)$ and $T_A(\phi)$ with Q^2 is the CT effect, which in the framework of CCMST derives from the off-diagonal transitions $h \rightarrow h' \rightarrow \ldots \rightarrow V$ in the ISI and FSI operators, and from the interference of transitions $\gamma^* \rightarrow V$ and $\gamma^* \rightarrow h$ in the photoabsorption vertex. It is interesting to assess the relative importance of difference off-diagonal transitions. In the most naive and unrealistic single-channel vector meson dominance (VMD) model, one allows only the $\gamma^* \rightarrow V$ transitions in the photoabsorption vertex and the diagonal $V \rightarrow V$ transitions in the ISI and FSI operators. The energy dependence caused by the onset of ISI is present even in VMD, but it completely fails to describe the Q^2 dependence of photoproduction on free nucleons—all the CT effects are lost, and we skip a discussion of this unrealistic approximation. A more interesting case is the one in which the ISI and FSI operators are approximated by the Glauber formulas (12) and (14). In such a simplified model, which we call here the Glauber approximation, one neglects the evolution of the ejectile and projectile wave packets caused by off-diagonal rescattering. The retention of the interference of transitions $\gamma^* \rightarrow V$ and $\gamma^* \rightarrow h$ in the photoabsorption vertex yields the Q^2-dependent projectile wave packet, and the interference of the diagonal $V \rightarrow V$ and off-diagonal $h \rightarrow V$ transitions in the incoherent rescattering vertex allow a correct description of the Q^2 dependence of the free nucleon cross section.

The importance of off-diagonal transitions in the FSI operator is best seen by comparing the energy dependence of $T_A(J/\Psi)$ in Figs. 3 and 7 with that of $T_A(\phi)$ in Figs. 4 and

8 in the low-energy regime $l_p \lesssim R_A$. In Figs. 7 and 8, nuclear transparency is flat versus energy ν, whereas in Figs. 3 and 4 nuclear transparency exhibits a growth, which is a contribution from the off-diagonal transitions to the CT effect. For the same reason, the Q^2-dependence of $T_A(V(1S))$ in the regime of developed ISI at high energies is weaker in the Glauber approximation (cf. Figs. 7, 8 and Figs. 3, 4), which is evidence for the importance of the off-diagonal transitions also in the ISI.

The difference between the full CCMST and Glauber approximation is even greater in the case of the $2S$ states. Here the Glauber model substantially underestimates the CT effects, most notably the preasymptotic-growth of $T_A(\Psi')$ with energy and the Q^2 dependence at high energies.

The high-energy region where the curves in Figs. 2 and 3 become flat corresponds to the regime in which the transverse interquark distance is frozen during propagation of the $q\overline{q}$ fluctuation of the virtual photon through the target nucleus. At the same time, this is the regime of developed ISI. In order to illustrate the energy dependence of the strength of ISI in Fig. 9 we show the ratio of nuclear transparency obtained with ISI excluded, i.e., putting $\langle h|\hat{S}_i|\gamma^*\rangle = 0$ in the decomposition (3), to nuclear transparency computed with full allowance for both ISI and FSI. As one can see, neglecting ISI leads to a systematic overestimate of nuclear transparency for $1S$ states.

For $2S$ vector mesons as more complicated interplay of ISI and FSI takes place at lower energies. After the energy upgrade, virtual photon energies $\nu \sim 5-8$ GeV will become feasible in experiments at CEBAF. Figure 9 shows that evaluations of $T_A(\phi)$ and $T_A(\phi')$ carried out in[8] neglecting the ISI have an accuracy of about 20–30%. The interpretation of the higher-energy data from the HERA–HERMES experiment requires full allowance for ISI. Still another light difference between the present calculations and those in Ref. 8 is in the parametrization of the dipole cross section. In Ref. 8 we parametrized the dipole cross section as $\sigma(\rho) = \sigma_0[1 - \exp(-\rho^2/R_0^2)]$, which reproduces well the gross features of the cross section of the model of Eq. (19).

5. CONCLUSIONS

We have presented a consistent theory of coherency effects in initial and final state interaction effects for virtual photoproduction of vector mesons off nuclei. The relative role of nuclear filtering and of the off-diagonal coherent rescatterings in the ISI and FSI for the onset of color transparency effects is elucidated. The onset of ISI with increasing photon energy is shown to have a very strong impact on nuclear transparency and on the onset of color transparency effects. For light vector mesons, the onset of ISI takes place in the kinematic domain of the HERA–HERMES experiment, and for charmonium states, the interesting energy dependence of nuclear transparency takes place in the kinematic domain of the GSI–ENC.

B. G. Z. and N. N. N. thank Prof. J. Spetch for hospitality at IKP, KFA Jülich, where part of this work was done. The work of N. N. N. was supported by DFG Grant ME864/132-1. N. N. N. is grateful to Prof. U. Meissner for hospitality at ITKP, University of Bonn.

[1] S. J. Brodsky and A. H. Mueller, Phys. Lett. B **206**, 685 (1988).
[2] B. Z. Kopeliovich and B. G. Zakharov, Phys. Rev. D **44**, 3466 (1991).
[3] O. Benhar, B. Z. Kopeliovich, C. Mariotti *et al.*, Phys. Rev. Lett. **69**, 1156 (1992).
[4] B. Z. Kopeliovich, J. Nemchick, N. N. Nikolaev, and B. G. Zakharov, Phys. Lett. B **309**, 179 (1993).
[5] B. Z. Kopeliovich, J. Nemchick, N. N. Nikolaev, and B. G. Zakharov, Phys. Lett. B **324**, 469 (1994).
[6] J. Nemchick, N. N. Nikolaev, and B. G. Zakharov, Phys. Lett. B **339**, 194 (1994).
[7] S. J. Brodsky, L. Frankfurt, J. F. Gunion *et al.*, Phys. Rev. D **50**, 3134 (1994).
[8] O. Benhar, B. G. Zakharov, S. Fantoni, and N. N. Nikolaev, Phys. Rev. Lett. **74**, 3565 (1995).
[9] N. N. Nikolaev and B. G. Zakharov, Z. Phys. C **49**, 607 (1991).
[10] N. N. Nikolaev and B. G. Zakharov, Z. Phys. C **64**, 631 (1994).
[11] N. N. Nikolaev and B. G. Zakharov, Phys. Lett. B **327**, 149 (1994).
[12] E665 Collaboration: M. R. Adams, S. Aid, P. L. Anthony *et al.*, Phys. Rev. Lett. **74**, 1525 (1995).
[13] N. N. Nikolaev, Zh. Éksp. Teor. Fiz. **81**, 814 (1981) [Sov. Phys. JETP **54**, 434 (1981)].
[14] B. Z. Kopeliovich and B. G. Zakharov, Yad. Fiz. **46**, 1535 (1987) [Sov. J. Nucl. Phys. **46**, 91 (1987)]; Phys. Lett. B **264**, 434 (1991).
[15] NMC Collaboration: M. Arneodo, A. Arvidson, B. Badelek *et al.*, Phys. Lett. B **332**, 195 (1994).
[16] R. J. Glauber, in *Lectures in Theoretical Physics*, Vol. 1, W. Brittain and L. G. Dunham (eds.) Interscience Publ. New York (1959); R. J. Glauber and G. Matthiae, Nucl. Phys. B **21**, 135 (1970).
[17] V. N. Gribov, Zh. Éksp. Teor Fiz. **56**, 892 (1969) [Sov. Phys. JETP **29**, 483 (1969)]; Zh. Éksp. Teor Fiz. **57**, 1306 (1970) [Sov. Phys. JETP **30**, 709 (1970)].
[18] O. Benhar, S. Fantoni, N. N. Nikolaev, J. Spetch, A. Usmani, and B. G. Zakharov, Zh. Éksp. Teor Fiz. **110**, 1933 (1996) [JETP **83**, 1063 (1996)].
[19] N. N. Nikolaev, V. R. Zoller, and B. G. Zakharov, Phys. Lett. B **328**, 143 (1994).
[20] F. E. Low, Phys. Rev. D **12**, 163 (1975).
[21] J. F. Gunion and D. E. Soper, Phys. Rev. D **15**, 2617 (1977).
[22] B. G. Zakharov, Yad. Fiz. **49**, 1386 (1989) [Sov. J. Nucl. Phys. **49**, 860 (1989)].
[23] J. Nemchik, N. N. Nikolaev, E. Predazzi, and B. G. Zakharov, Phys. Lett. B **374**, 199 (1996).
[24] N. N. Nikolaev and B. G. Zahkarov, Z. Phys. C **53**, 331 (1992).
[25] E. Schuryak, Rev. Mod. Phys. **65**, 1 (1993).
[26] NMC Collaboration: P. Amaudruz, M. Arneodo, A. Ervidson *et al.*, Nucl. Phys. B **371**, 553 (1992).
[27] E687 Collaboration: P. L. Frabetti, H. W. K. Cheung, I. P. Cumalat *et al.*, *Report at the Europhysics Conference on High Energy Physics*, Brussels, June 1995.
[28] N. N. Nikolaev, Nucl. Part. Phys. **21**, 41 (1992).
[29] H. de Vries, C. W. de Jaeger, and C. de Vries, Atomic Data and Nuclear Data Tables **36**, 496 (1987).

Published in English in the original Russian journal. Reproduced here with stylistic changes by the Translation Editor.

Reduction of the finite grand unification theory to the minimal supersymmetric standard model

I. N. Kondrashuk*)

Bogoliubov Laboratory of Theoretical Physics, Joint Institute for Nuclear Research, 141980 Dubna, Moscow Region, Russia

(Submitted 29 May 1996)
Zh. Éksp. Teor. Fiz. **111**, 787–795 (March 1997)

The recently proposed mechanism for reducing the finite $SU(5)$ grand unification theory (GUT) to the minimal supersymmetric standard model (MSSM) is reanalyzed and simplified. For the scalar $SU(2)\times U(1)$ invariant Higgs doublet potential that results from $SU(5)$ symmetry breaking to have no dangerous directions, a restriction on the parameters of the unified theory should be imposed. At the same time, this restriction guarantees that the scalar Higgs doublet potential has a minimum at zero at the GUT scale, and the low-energy theory appears to be exactly the MSSM. © *1997 American Institute of Physics.* [S1063-7761(97)00203-5]

1. INTRODUCTION

The supersymmetric (SUSY) field theories have remarkable properties in the ultraviolet range. The no-renormalization theorem for $N=1$ SUSY theories[1] guarantees absence of divergences in quantum corrections to the superpotential. The only possible divergences in these theories (in the background field method) are the logarithmic divergences of the two-point Green functions of gauge and chiral superfields. If the group and the multiplet contents of the $N=1$ theory are chosen in some particular way, these divergences disappear at the one-loop level too. This occurs as the result of mutual cancellation of the divergent contributions from gauge and Yukawa interactions.[2] The gauge groups and the multiplet contents of the theories for which this cancellation is possible have been classified in Ref. 3. In Refs. 2 and 4 it has been shown that one-loop finiteness guarantees two-loop finiteness of the theory without imposing new conditions, but this requirement appeared to be insufficient for the theory to be free from ultraviolet divergences at the three-loop level.[5] However, an algorithm for constructing an $N=1$ SUSY field theory finite in all orders of perturbation theory has been proposed and a finite $SU(5)$ grand unification theory (GUT) was constructed.[6] The method used there was based on fine-tuning of the Yukawa coupling constants in each new order of perturbation theory. The only requirement imposed on the theory for this algorithm to work is one-loop finiteness (and, automatically, two-loop finiteness).[6]

The idea of complete finiteness of the unified theory is very attractive, and it is not surprising that many efforts have been made to derive low-energy predictions of the finite unified theory and compare them with modern experimental data.[7] For this purpose the standard approach is used: it is assumed that $SU(5)$ symmetry is spontaneously broken at the unification scale, and the unified theory is reduced to a low-energy supersymmetric theory with the corresponding boundary conditions for the coupling constants of the low-energy Lagrangian at the GUT scale. Then, the renormalization group equation method is applied to get quantities of the Lagrangian at the electroweak scale, where spontaneous breaking of the electroweak symmetry occurs.[8,9]

Although the one-loop finiteness conditions fix the gauge groups and the multiplet contents of the finite $N=1$ SUSY theory, they allow considerable arbitrariness in the Yukawa and mass matrices.[10] In this situation the main guiding principles in choosing the finite GUT are simplicity and aesthetic attractiveness of the unified theory. In Ref. 11, the finite GUT satisfying these requirements was completely investigated in a consistent way. The model was based on the $SU(5)$ gauge group and is the simplest finite GUT compatible with the low-energy phenomenology. Its distinction from the minimal SUSY $SU(5)$ GUT model is in the extension of the Higgs sector: it contains eight Higgs superfields instead of two in the minimal $SU(5)$ GUT.[12] The Yukawa and mass parts of the Lagrangian are chosen in the most economical way. Soft supersymmetry breaking takes place at the Planck scale M_P due to the appearance of soft terms in the Lagrangian.[13,14] A universal form for these terms at the Planck scale is assumed.[14] In Ref. 11, the condition of finiteness was extended to them, which resulted in a completely finite theory between M_P and M_{GUT}. Complete finiteness in this case means that no charge or mass coupling of the theory changes in this energy range. To get the small initial mass parameters of the low-energy theory from the large mass parameters of the unified theory, the usual fine-tuning procedure was used in Ref. 11. This procedure generates the hierarchy of the mass scales in the doublet part of the $SU(5)$ superpotential which decouples from the triplet part of the latter after spontaneous breaking of the $SU(5)$ symmetry.

In the treatment of the low-energy part of the finite GUT model in Ref. 11, in addition to the matter superfields of the minimal supersymmetric standard model (MSSM), three Higgs doublets were included in the low-energy Lagrangian. To get the Higgs potential at the electroweak scale, the renormalization group equations for the parameters of the potential were used. According to the radiative symmetry-breaking scenario,[8] the parameters of the scalar Higgs potential yield nontrivial vacuum expectation values of the scalar Higgs fields at the electroweak scale. Due to the degeneracy of the Yukawa couplings with respect to the generations of the matter superfields,[11,15] the quark and lepton mass spectrum at the GUT scale in this model is completely determined by the spectrum of vacuum expectation values of the Higgs fields at the electroweak scale.

In this paper, the low-energy part of this finite theory is analyzed in a simpler and more efficient way than that of Ref. 11. Namely, the finite GUT is shown to reduce to the MSSM after spontaneous breaking of the $SU(5)$ symmetry at the unification scale. The parameters of the electroweak Lagrangian need not be evolved down to low energies in this approach.

2. FINITE GUT

In this section, a brief review of the main points of the softly broken supersymmetric finite $SU(5)$ model constructed in Ref. 11 is proposed. The multiplet contents of the model and its Lagrangian were described there. The sector of the chiral matter and Higgs superfields has the following contents (in terms of the irreducible representations of $SU(5)$):

Matter fields: $\Psi_i - \overline{5}, \quad \Lambda_i - 10, \quad i = 1,\ 2,\ 3,$

Higgs fields: $\Phi_a - 5, \quad \overline{\Phi}_a - \overline{5}, \quad \Sigma - 24,$

$a = 1,\ 2,\ 3,\ 4,$

where i and a are the generation indices of the matter and Higgs superfields, respectively.

The most general form of the superpotential for the theory having this field content is[6]

$$\begin{aligned} W = {} & A^a_{ij}\overline{\Phi}_{a\alpha}\Psi_{i\beta}\Lambda^{\alpha\beta}_j + \frac{1}{8} B^a_{ij}\Phi^\alpha_a\Lambda^{\beta\gamma}_i\Lambda^{\delta\sigma}_j\epsilon^{\alpha\beta\gamma\delta\sigma} \\ & + C_{ab}\overline{\Phi}_{a\alpha}\Phi^\alpha_b\Sigma^\alpha_\beta + \frac{1}{3} D\Sigma^\alpha_\beta\Sigma^\beta_\gamma\Sigma^\gamma_\alpha \\ & + \frac{1}{2} E^i_{ab}\overline{\Phi}_{a\alpha}\overline{\Phi}_{b\beta}\Lambda^{\alpha\beta}_i + F_{i\alpha}\Psi_{i\alpha}\Phi^\beta_a\Sigma^\alpha_\beta \\ & + \frac{1}{2} G^k_{ij}\Psi_{i\alpha}\Psi_{j\beta}\Lambda^{\alpha\beta}_k, \end{aligned}$$

where $\epsilon^{\alpha\beta\gamma\delta\sigma}$ is a completely antisymmetric tensor. The last three terms, which would violate the baryon and lepton numbers at the tree levels and lead to $(B - L)$-nonconservation, are usually ignored.

The one-loop finiteness conditions of the two-point Green's functions in the theory with the above potential are[6]

$$\begin{aligned} & 4\sum_{i,j} A^a_{ij}(A^b_{ij})^* + \frac{24}{5}\sum_e C_{ae}(C_{be})^* + 4\sum_{i,e} E^i_{ae}(E^i_{be})^* \\ & = \frac{12}{5} g^2\delta_{ab}, \quad 3\sum_{i,j} B^a_{ij}(B^b_{ij})^* + \frac{24}{5}\sum_e C_{ea}(C_{eb})^* \\ & + \frac{24}{5}\sum_i F_{ia}(F_{ib})^* = \frac{12}{5} g^2\delta_{ab}, \end{aligned}$$

$$\begin{aligned} & 4\sum_{a,k} A^a_{ik}(A^a_{jk})^* + \frac{24}{5}\sum_a F_{ia}(F_{ja})^* + 4\sum_{l,k} G^k_{il}(G^k_{jl})^* \\ & = \frac{12}{5} g^2\delta_{ij}, \quad 2\sum_{a,k} A^a_{ki}(A^a_{kj})^* + 3\sum_{a,k} B^a_{ik}(B^a_{jk})^* \\ & + 2\sum_{a,b} E^i_{ab}(E^j_{ab})^* + 2\sum_{k,l} G^i_{kl}(G^j_{kl})^* = \frac{18}{5} g^2\delta_{ij}, \\ & \sum_{a,b} C_{ab}(C_{ab})^* + \frac{21}{5} DD^* + \sum_{i,a} F_{ia}(F_{ia})^* = 5g^2. \qquad (1) \end{aligned}$$

In Ref. 11 the following simple ansatz for the Yukawa matrices of the superpotential was proposed:

$$\begin{aligned} W = {} & y_1\Psi_i K_{ij}\overline{\Phi_i}\Lambda_j + y_1'\Psi_i\overline{\Phi_4}\Lambda_i + \frac{y_2}{8}\Phi_i\Lambda_i\Lambda_i \\ & + \frac{y_2'}{8}\Phi_4\Lambda_i\Lambda_i + y_3\overline{\Phi_i}S_{ij}\Sigma\Phi_j + y_3'\overline{\Phi_4}\Sigma\Phi_4 + \frac{y_4}{3} \\ & \times\Sigma^3 + \overline{\Phi_i}M_{ij}\Phi_j + \overline{\Phi_4}M\Phi_4 + \frac{M_0}{2}\Sigma^2, \qquad (2) \end{aligned}$$

where the $SU(5)$ indices are omitted, but can easily be recovered in a covariant manner. The potential (2) is taken in this form so that each generation of the matter interacts with its pair of Higgs fields, while the fourth pair of the Higgs fields is coupled with all the generations of matter as well as the Higgs pair of the minimal SUSY $SU(5)$ GUT. In Ref. 6 it is demonstrated that the Yukawa matrices corresponding to this specific ansatz will not be changed by the quantum corrections if the Yukawa couplings in (2) satisfy the conditions of one-loop finiteness (1) and if their necessary fine tuning is performed in each order of perturbation theory step by step.

The presence of unitary K and S matrices does not contradict the finiteness conditions (1). The matrix K is necessary to create the initial mixing of the quark fields at the unification scale (that is, initial values of the Cabibbo–Kobayashi–Maskawa matrix). As for the matrix S, it contains all initial information about the hierarchy of the quark mass spectrum at this scale. This role of S will become clear below.

From the no-renormalization theorem for the superpotential[1] it follows that the mass parameters M_0, M, and M_{ij} are not fixed by the requirement of one-loop finiteness. If M_0 is negative, the unified $SU(5)$ symmetry is broken by the vacuum expectation value of Σ (Ref. 11):

$$\langle\Sigma\rangle = \begin{pmatrix} V & & & & \\ & V & & & \\ & & V & & \\ & & & -\frac{3}{2}V & \\ & & & & -\frac{3}{2}V \end{pmatrix},$$

where $V \sim M_0/y_4 \sim 10^{16}$ GeV.

After breaking of the $SU(5)$ symmetry, the Higgs quintets Φ_a and $\overline{\Phi}_a$ split into doublets and triplets. As can be seen from (2), their mass terms look like

$$y_3\overline{\Phi}_i S_{ij}\langle\Sigma\rangle\Phi_j+\overline{\Phi}_i M_{ij}\Phi_j =\overline{\Phi}_i\begin{pmatrix} y_3S_{ij}V+M_{ij} & \\ & -\frac{3}{2}y_3S_{ij}V+M_{ij}\end{pmatrix}\Phi_j, \tag{3}$$

and

$$y_3'\overline{\Phi}_4\langle\Sigma\rangle\Phi_4+\overline{\Phi}_4 M\Phi_4 =\overline{\Phi}_4\begin{pmatrix} y_3'V+M & \\ & -\frac{3}{2}y_3'V+M\end{pmatrix}\Phi_4. \tag{4}$$

All the mass parameters in these relations are on the order of the $M_{\rm GUT}$ scale. To generate light initial boundary values for the masses in the doublet part of the Higgs sector of the unified theory, which should be radiatively corrected to give the mass parameters at the electroweak scale, it is necessary to carry out a fine-tuning procedure. For this purpose, the following trick was used in Ref. 11.

First, the unitary matrix S was represented as

$$S=\overline{X}\begin{pmatrix} e^{i\theta_1} & 0 & 0\\ 0 & e^{i\theta_2} & 0\\ 0 & 0 & e^{i\theta_3}\end{pmatrix}X^T=\overline{X}DX^T,$$

$$\overline{X^T}\,\overline{X}=I,\quad X^TX=I,$$

where X and $\widetilde{X}$ are real orthogonal matrices, D is a unitary diagonal matrix, and I is a unit matrix. The solution of the one-loop finiteness conditions for the specific ansatz of the Yukawa matrices used in the superpotential (2) still has some arbitrariness,[11] which can be used to set $y_3'=0$. This allows one to absorb one common phase into the redefinition of the fields. Therefore, in what follows it is supposed that $\theta_3=0$. It is necessary to note that $y_3'=0$ implies that the fourth pair of the Higgs doublets (4) remains heavy in any case.

Second, the requirement of one-loop finiteness does not restrict the mass matrix M_{ij}. This matrix can be written as

$$M=\overline{X}(RI+T'D)X^T,$$

where R and T' are some heavy mass parameters,

$$R\sim T'\sim V.$$

Now, if the fourth pair of the Higgs doublets and its Yukawa interactions are omitted, the $SU(2)\times U(1)$ invariant superpotential at the $M_{\rm GUT}$ scale can be represented as

$$\begin{aligned} W=&\left(\sqrt{\frac{2}{5}}\,gQ_j^bK_{ij}\overline{H_i^a}D_i+\sqrt{\frac{2}{5}}\,gL_i^b\overline{H_i^a}E_i\right.\\ &\left.+\sqrt{\frac{8}{15}}\,gQ_i^bH_i^aU_i\right)\epsilon_{ab}+\{\overline{H_i^a}[\overline{X}(RI\\ &+TD)X^T]_{ij}H_j^b\epsilon_{ab}, \end{aligned}\tag{5}$$

where $a,\ b=1,\ 2$ are the $SU(2)$ indices and $\epsilon_{12}=1$, and the following notation is used

$$T=T'-\frac{3}{2}y_3V.$$

The three pairs of Higgs doublets have the quantum numbers

$$\overline{H}_i\left(1,2,-\frac{1}{2}\right)=\begin{pmatrix}\overline{H_i^0}\\ H_i^-\end{pmatrix},\quad H_i\left(1,2,\frac{1}{2}\right)=\begin{pmatrix}H_i^+\\ H_i^0\end{pmatrix}, \tag{6}$$

while the other superfields in (5) are the usual matter superfields of the MSSM.[16]

In addition, the following soft supersymmetry breaking terms must be added to the superpotential (5):

$$\begin{aligned} W_{SSB}=&m_0^2\sum_i|\varphi_i|^2+\frac{1}{2}\left(m_{1/2}\sum_k\lambda_k\lambda_k+\text{H.c.}\right)\\ &+B\{\overline{H_i^a}[\overline{X}(RI+TD)X^T]_{ij}H_j^b+\text{c.c.}\}\epsilon_{ab}\\ &+(A_Dy_D\widetilde{q}_j^bK_{ij}\overline{H_i^a}\widetilde{d}_i+A_Ly_L\widetilde{l}_i^b\overline{H_i^a}\widetilde{e}_i\\ &+A_Uy_U\widetilde{q}_i^bH_i^a\widetilde{u}_i+\text{c.c.})\epsilon_{ab}, \end{aligned}\tag{7}$$

where φ_i denotes all scalar fields with common mass m_0^2 at the unification scale, and λ_k are the gauginos with common mass $m_{1/2}$ at the same scale. Aside from the gauginos, all other fields in W_{SSB} are the low scalar components of the corresponding superfields. The notation for the scalar Higgs doublets in (7) coincides with the corresponding superfield notation (6). These soft supersymmetry breaking terms can be reduced from the corresponding $SU(5)$ invariant terms of the unified theory after the $SU(5)$ symmetry-breaking.[11]

Having rotated the superfields H_i and $\overline{H}_i$ as

$$H_i=(XH')_i=X_{i1}H_1'+X_{i2}H_2'+X_{i3}H_3', \tag{8}$$

$$\overline{H}_i=(\overline{XH'})_i=\overline{X}_{i1}\overline{H_1'}+\overline{X}_{i2}\overline{H_2'}+\overline{X}_{i3}\overline{H_3'}, \tag{9}$$

where H_i' and $\overline{H_i'}$ are the new Higgs superfields, one can conveniently rewrite (5) as

$$\begin{aligned} W=&(y_DQ_j^bK_{ij}\overline{X}_{ik}\overline{H_k'}^aD_i+y_LL_i^b\overline{X}_{ik}\overline{H_k'}^aE_i\\ &+y_UQ_i^bX_{ik}H_k'^aU_i)\epsilon_{ab}+[\overline{H_i'}^a(RI+TD)_{ij}H_j'^b]\epsilon_{ab}, \end{aligned}\tag{10}$$

and (7) as

$$\begin{aligned} W_{SSB}=&m_0^2\sum_i|\varphi_i|^2+\frac{1}{2}\left(m_{1/2}\sum_k\lambda_k\lambda_k+\text{H.c.}\right)\\ &+B[\overline{H_i'}^a(RI+TD)_{ij}H_j'^b+\text{c.c.}]\epsilon_{ab}\\ &+(A_Dy_D\widetilde{q}_j^bK_{ij}\overline{X}_{ik}\overline{H_k'}^a\widetilde{d}_i+A_Ly_L\widetilde{l}_i^b\overline{X}_{ik}\overline{H_k'}^a\widetilde{e}_i\\ &+A_Uy_U\widetilde{q}_i^bX_{ik}H_k'^a\widetilde{u}_i+\text{c.c.})\epsilon_{ab}. \end{aligned}\tag{11}$$

To get the light Higgs doublet pair, a fine tuning procedure should be performed:

$$R+T=\mu\sim10^3\ \text{GeV}. \tag{12}$$

The fine-tuning procedure is more meaningful than in the other GUTs, since in the finite model none of the parameters is running above the GUT scale.

As can be seen from (10) and (11), the first two components in the decompositions (8) and (9) remain heavy, while the third components H_3' and $\overline{H_3'}$ become light. By the decoupling theorem,[17] only this pair need be taken into account in the effective low-energy theory, whose superpotential takes the following form at the unification scale:

$$W=(y_D\overline{n_i}Q_j^bK_{ij}\overline{H_3'}^aD_i+y_L\overline{n_i}L_i^b\overline{H_3'}^aE_i$$
$$+y_Un_iQ_i^bH_3'^aU_i)\epsilon_{ab}+B(\mu\overline{H_3'}^aH_3'^b)\epsilon_{ab}\,, \tag{13}$$

where

$$\overline{n_i}=\overline{X_{i3}}\,,\quad n_i=X_{i3}\,,\quad \sum_i \overline{n}_i^2=1,\quad \sum_i n_i^2=1.$$

The corresponding soft supersymmetry breaking terms are

$$W_{SSB}=m_0^2\sum_i |\varphi_i|^2+\frac{1}{2}\left(m_{1/2}\sum_k \lambda_k\lambda_k+\text{H.c.}\right)$$
$$+B(\mu\overline{H_3'}^aH_3'^b+\text{c.c.})\epsilon_{ab}+(A_Dy_D\overline{n_i}\widetilde{q}_j^bK_{ij}\overline{H_3'}^a\widetilde{d}_i$$
$$+A_Ly_L\overline{n_i}\widetilde{l}_i^b\overline{H_3'}^a\widetilde{e}_i+A_Uy_Un_i\widetilde{q}_i^bH_3'^a\widetilde{u}_i+\text{c.c.})\epsilon_{ab}\,. \tag{14}$$

Here φ_i denotes all light scalar fields of the effective low-energy theory. In analogy with (7), the notation for the scalar Higgs doublets in (14) coincides with the notation of the corresponding superfields.

Equations (13) and (14) are the usual superpotential and soft supersymmetry breaking potential of the MSSM,[16] respectively. As is well known, in the MSSM there is no problem with unification of the gauge coupling constants of MSSM at a single point at a very high scale.[18] Moreover, only in the supersymmetric model with two Higgs doublets is this unification possible.[19]

As can be seen from (13), all information about the quark mass hierarchy at the GUT scale is contained in the Higgs sector of the finite unified theory, namely, in the unitary Higgs mixing matrix S:

$$y_i^U=n_iy^U,\quad y_i^D=\overline{n_i}y^D,\quad y_i^L=\overline{n_i}y^L,$$

$$y^U=\frac{4}{\sqrt{15}}\,g_{\text{GUT}}\,,\quad y^D=y^L=\frac{2}{\sqrt{5}}\,g_{\text{GUT}}\,, \tag{15}$$

where g_{GUT} is a gauge coupling constant of the unified theory.

These conclusions are natural and correct, and there are no subtle points if the full scalar Higgs doublet potential has no dangerous directions (along these directions it can be unbounded below) and has absolute minimum at zero at the unification scale. Since the full scalar Higgs potential has a rather complicated structure because of the large number of Higgs fields, this is not obvious. In the next section, a condition will be written for the parameters of the unified superpotential which is necessary to guarantee this.

3. DOUBLET POTENTIAL AT THE GUT SCALE

The scalar Higgs doublet potential arises from the superpotential (13) and minimal SUSY gauge interaction, when all nondynamical components of the Higgs and gauge superfields are eliminated, and from the corresponding soft supersymmetry breaking part (14).[16] Hence, the full potential has the following form at the unification scale:

$$V(\overline{H_i'},H_i')=(m_0^2+R^2+T^2)\left(\sum_i |\overline{H_i'}|^2+\sum_i |H_i'|^2\right)$$
$$+RT(D^*+D)_{ij}(\overline{H_i'}^\dagger\overline{H_j'}+H_i'^\dagger H_j')$$
$$+B[\overline{H_i'}^a(RI+TD)_{ij}H_j'^b+\text{c.c.}]\epsilon_{ab}$$
$$+\frac{g^2+g'^2}{8}\left[\sum_i |\overline{H_i'}|^2-\sum_i |H_i'|^2\right]^2$$
$$+\frac{g^2}{4}\left[(\overline{H_i'}^\dagger\overline{H_j'})^*(\overline{H_i'}^\dagger H_j')-\left(\sum_i |\overline{H_i'}|^2\right)^2\right.$$
$$+(H_i'^\dagger H_j')^*(H_i'^\dagger H_j')-\left(\sum_i |H_i'|^2\right)^2$$
$$\left.+2(\overline{H_i'}^\dagger H_j')^*(\overline{H_i'}^\dagger H_j')\right], \tag{16}$$

where summation over the repeating Higgs generation indices is implied. Also, it is assumed for brevity that

$$|\overline{H_i}|^2=|\overline{H_i^0}|^2+|\overline{H_i^-}|^2,\quad |H_i|^2=|H_i^+|^2+|H_i^0|^2.$$

It is more convenient to introduce the new notation

$$M_1^2=m_0^2+R^2+T^2+2RT\cos\theta_1\,,$$

$$\mathscr{M}_1e^{i\gamma_1}=R+Te^{i\theta_1},\quad \mathscr{M}_1=|R+Te^{i\theta_1}|,$$

$$M_2^2=m_0^2+R^2+T^2+2RT\cos\theta_2\,,$$

$$\mathscr{M}_2e^{i\gamma_2}=R+Te^{i\theta_2},\quad \mathscr{M}_2=|R+Te^{i\theta_2}|,$$

$$m^2=m_0^2+(R+T)^2,\quad \mu=R+T,$$

and rewrite the potential (16) as

$$V=M_1^2(|\overline{H_1'}|^2+|H_1'|^2)+M_2^2(|\overline{H_2'}|^2+|H_2'|^2)+m^2(|H_3'|^2$$
$$+|H_3'|^2)+B(\mathscr{M}_1e^{i\gamma_1}\overline{H_1'}^aH_1'^b+\mathscr{M}_2e^{i\gamma_2}\overline{H_2'}^aH_2'^b$$
$$+\mu\overline{H_3'}^aH_3'^b)\epsilon_{ab}+\frac{g^2+g'^2}{8}\left[\sum_i |\overline{H_i'}|^2-\sum_i |H_i'|^2\right]^2$$
$$+\frac{g^2}{4}\left[(\overline{H_i'}^\dagger\overline{H_j'})^*(\overline{H_i'}^\dagger\overline{H_j'})-\left(\sum_i |\overline{H_i'}|^2\right)^2\right.$$
$$+(H_i'^\dagger H_j')^*(H_i'^\dagger H_j')-\left(\sum_i |H_i'|^2\right)^2$$
$$\left.+2(\overline{H_i'}^\dagger H_j')^*(\overline{H_i'}^\dagger H_j')\right]. \tag{17}$$

Having parametrized the Higgs doublets as

$$H_i=U_i\begin{pmatrix}0\\ v_i\end{pmatrix},\quad \overline{H_i}=\overline{U_i}\begin{pmatrix}\overline{v_i}\\ 0\end{pmatrix},$$

where U_i and $\overline{U_i}$ are some $SU(2)$ matrices, and v_i and $\overline{v_i}$ are positive, one can derive that when the quartic terms in (17) vanish the condition for positivity of the quadratic part of (17) is

$$M_1^2(\alpha_1^2+\beta_1^2)+M_2^2(\alpha_2^2+\beta_2^2)+m^2(\alpha_3^2+\beta_3^2) \\ -2|B\mathscr{M}_1|\alpha_1\beta_1-2|B\mathscr{M}_2|\alpha_2\beta_2-2|B\mu|\alpha_3\beta_3\geqslant 0, \tag{18}$$

where

$$\alpha_1^2+\alpha_2^2+\alpha_3^2=1, \quad \beta_1^2+\beta_2^2+\beta_3^2=1,$$

$$\overline{v_i}=\alpha_i\sqrt{\sum_i \overline{v}_i^2}, \quad v_i=\beta_i\sqrt{\sum_i v_i^2}.$$

This requirement is necessary to provide stability of (17) in these directions. The quadratic form (18) is obviously positive if the following conditions are satisfied:

$$M_1^2\geqslant|B\mathscr{M}_1|, \tag{19}$$

$$M_2^2\geqslant|B\mathscr{M}_2|, \tag{20}$$

$$m^2\geqslant|B\mu|. \tag{21}$$

The conditions (19) and (20) hold in any case due to our fine-tuning procedure (12) (it is assumed that $B\sim\mu$). Note that if (21) is violated, the quadratic form would be negative when

$$\alpha_1=\alpha_2=\beta_1=\beta_2=0.$$

Thus, the condition

$$m_0^2+\mu^2\geqslant|B\mu| \tag{22}$$

is necessary for the stability of the potential (17) for large fields. At the same time, as can be seen from (18) and (19)–(21), the restriction (22) guarantees the positivity of the scalar Higgs potential (16) on any field configurations. This means that after spontaneous breaking of the $SU(5)$ symmetry the scalar Higgs doublet potential has its only and absolute minimum at zero at the GUT scale.

4. CONCLUSIONS

In this paper, the SUSY $SU(5)$ finite theory with an R-symmetrical and $(B-L)$-conserving superpotential has been considered. The Yukawa matrices of this theory were chosen in the simplest possible way, and their values were fixed by the condition of finiteness up to some arbitrariness. This arbitrariness, and the arbitrariness in the choice of the mass matrices that are not restricted by the condition of finiteness, can be used to reduce the finite GUT to the MSSM after the $SU(5)$ symmetry breaking at the unification scale. Reduction to the MSSM is necessary because only in a SUSY theory with two Higgs doublets is the unification of the gauge couplings of the MSSM at a single point possible.[18] In the supersymmetric theories with a more extended Higgs sector such a unification is problematical.[19]

In this work, the analysis of the reduction of the finite GUT to the MSSM is simpler than that in Ref. 11. For the low-energy theory to be self-consistent, it is necessary that the scalar Higgs doublet potential have no dangerous directions and have an absolute global minimum at zero at the unification scale after the $SU(5)$ symmetry breaking. These requirements impose the restriction (22) on the parameters of the finite GUT. If this restriction holds, both the requirements are met. As for the rest, after the $SU(5)$ symmetry breaking one gets the MSSM as the low-energy theory with the boundary conditions at the GUT scale for the Yukawa couplings (15).

I am grateful to the Russian Fund for Fundamental Research for financial support of this work (Grant No. 96-02-17379a). I am also grateful to the Heisenberg–Landau program for providing a stipend. This work is also supported by the International Center for Fundamental Physics in Moscow (INTAS Grant No. 93-2492-ext).

*)e-mail: ikond@thsun1.jinr.dubna.su

---

[1] M. T. Grisaru, W. Siegel, and M. Rocek, Nucl. Phys. B **159**, 429 (1979).
[2] A. Parkes and P. West, Phys. Lett. B **138**, 99 (1984); D. R. T. Jones and L. Mezincesku, Phys. Lett. B **136**, 242 (1984).
[3] S. Hamidi, J. Patera, and J. H. Schwarz, Phys. Lett. B **141**, 349 (1984); S. Rajpoot and J. G. Taylor, Phys. Lett. B **147**, 91 (1984).
[4] D. R. T. Jones and L. Mezincesku, Phys. Lett. B **138**, 293 (1984); P. West, Phys. Lett. B **137**, 371 (1984).
[5] A. Parkes, Phys. Lett. B **156**, 73 (1985).
[6] A. V. Ermushev, D. I. Kazakov, and O. V. Tarasov, Nucl. Phys. B **281**, 72 (1987); D. I. Kazakov, Mod. Phys. Lett. A **9**, 663 (1987); D. R. T. Jones, Nucl. Phys. B **279**, 153 (1986).
[7] D. I. Kazakov and I. N. Kondrashuk, Int. J. Mod. Phys. A **7**, 3869 (1992); D. Kapetanakis, M. Mondragon, and G. Zoupanos, Z. Phys. C **60**, 181 (1993); J. Kubo, M. Mondragon, and G. Zoupanos, Nucl. Phys. B **424**, 291 (1994); Kanazawa University Preprint, KANAZAWA-95-05 (1995); M. Mondragon and G. Zoupanos, CERN Preprint CERN-TH-7098/93; M. Mondragon and G. Zoupanos, Nucl. Phys. B (Proc. Suppl.) C **37**, 98 (1995); J. Kubo, M. Mondragon, and G. Zoupanos, Preprint MPI-PhT/95-132; J. Kubo, M. Mondragon, M. Olechowski, and G. Zoupanos;, Preprint MPI-PhT/95-133.
[8] G. G. Ross and R. G. Roberts, Nucl. Phys. B **377**, 517 (1992).
[9] H. E. Haber, Preprint SCIPP 91/06; R. A. Flores and M. Sher, Ann. Phys. **148**, 95 (1983); H. E. Haber and K. Kane, Phys. Rep. **117**, 75 (1985); R. Barbieri, Rivista Nuovo Cim. **11** No. 4, 1 (1988); M. Sher, Phys. Rep. **179**, 273 (1989); I. N. Kondrashuk, JETP Lett. **62**, 472 (1995); I. N. Kondrashuk, Int. J. Mod. Phys. A **11**, 989 (1996).
[10] J. E. Björkman, D. R. T. Jones, and S. Raby, Nucl. Phys. B **259**, 503 (1985).
[11] D. I. Kazakov, M. Yu. Kalmykov, I. N. Kondrashuk, and A. V. Gladyshev, Nucl. Phys. B **471**, 389 (1996).
[12] N. Sakai, Z. Phys. C **11**, 153 (1982).
[13] N. Polonsky and A. Pomarol, Phys. Rev. Lett. **73**, 2292 (1994).
[14] H. P. Nilles, Phys. Rep. C **110**, 1 (1984); L. E. Ibáñez and D. Lust, Nucl. Phys. B **382**, 305 (1992); V. Kaplunovsky and J. Louis, Phys. Lett. B **306**, 269 (1993); R. Barbieri, J. Loius, and M. Moretti, Phys. Lett. B **312**, 451 (1993); **316**, 632(E) (1993); A. Brignole, L. E. Ibáñez, and C. Munoz, Nucl. Phys. B **422**, 125 (1994); L. Hall, J. Lykken, and S. Weinberg, Phys. Rev. D **27**, 2359 (1983).
[15] D. I. Kazakov, JINR Preprint E2-94-162.
[16] H. E. Haber, Preprint SCIPP 92/33 of the Santa Cruz Institute of Particle Physics (1992).
[17] T. Appelquist and J. Carazzone, Phys. Rev. D **11**, 2856 (1975).
[18] U. Amaldi, W. de Boer, and H. Fürstenau, Phys. Lett. B **260**, 447 (1991).
[19] S. Dimopoulos and H. Georgi, Nucl. Phys. B **193**, 150 (1981).

Published in English in the original Russian journal. Reproduced here with stylistic changes by the Translation Editor.

Resonant scattering of three-level Rydberg atoms in a microwave field

A. V. Bezverbnyĭ, I. M. Beterov, A. M. Tumaĭkin, and I. I. Ryabtsev

Novosibirsk State University, 630000 Novosibirsk, Russia

Institute of Semiconductor Physics, Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia

(Submitted 4 April 1996)
Zh. Éksp. Teor. Fiz. **111**, 796–815 (March 1997)

We examine the theory of potential scattering of Rydberg atoms in a microwave field. The model of a three-level atom is employed to calculate the radiative force emerging in the resonant coherent interaction with the microwave field for the case of a two-photon resonance and high intensities, using the method of quasienergies of the system consisting of the atom and the field. We determine the probabilities of Landau–Zener transitions in the spatial regions where under two-photon resonance conditions the quasienergies of the atoms approach one another by a small quantity. We also study the dynamics of the variation of the spatial profile of a beam of Rydberg atoms caused by resonant scattering. Finally, we give the results of the first experimental observation of the variation of the transverse beam profile when Rydberg atoms pass through a nonuniform microwave field formed in a rectangular waveguide and in resonance with the two-photon 36P–37P transition. [S1063-7761(97)00303-X]

1. INTRODUCTION

The resonant interaction of Rydberg atoms and microwave radiation is one of the areas of the physics that is being vigorously explored.[1] Lately a number of papers have appeared that study the effect of spatially nonuniform microwave fields on the motion of Rydberg atom in resonant one-photon[2] and multiphoton[3] interactions.

The present study is devoted to a theoretical analysis of the process of resonant scattering of a beam of Rydberg atom in a microwave field of a standing linearly polarized wave, based on the assumption that the three-level model of atomic energy states is valid. In Sec. 2 we describe the interaction of the atoms with the microwave field. Section 3 is devoted to a study of ''dressed'' states of Rydberg atoms in a microwave field. The analysis is done by applying the method of quasienergies to separate resonant states of the system consisting of an atom and the field. In Sec. 4 we investigate the probabilities of Landau–Zener transitions in spatial regions where the corresponding quasienergies of the atoms move closer together by a small quantity in conditions of a two-photon resonance. In Sec. 5 we calculate the radiative force emerging as a result of the resonant interaction of the atoms and the microwave field. Here we study various limiting cases depending on the initial population of the selected Rydberg states of atoms: in two-photon resonance, and at high microwave field intensities. We compare our results with those for a one-photon transition, a case thoroughly studied for optical transitions. Section 6 is devoted to an analysis of the kinetics of an atomic beam as a whole, with the effects of spontaneous decay of Rydberg states of atoms ignored. In Sec. 7 we examine the dynamics of the variations of the spatial profile of the beam caused by scattering processes in the light field. There we also give the results of experiments on the scattering of a beam of Rydberg sodium atoms in the 36P state in a spatially nonuniform microwave field in the form of a standing wave.

2. A MODEL FOR THE INTERACTION OF RYDBERG ATOMS AND A MICROWAVE FIELD

Let us take a beam of Rydberg atoms propagating along the x axis with a thermal velocity $v \sim 10^5$ cm/s and entering a region with a standing microwave field directed along the z axis:

$$\mathbf{E}(\mathbf{r},t)=\{e^{-i\omega t}E_0\mathbf{e}_0 \sin(kz)+ \text{c.c.}\}, \tag{1}$$

where E_0 is amplitude of the field, ω is the wave frequency, and $\mathbf{e}_0$ is the linear polarization vector. For instance, for sodium atoms in highly excited states with principal quantum numbers $n=20$–40, the corresponding resonance wavelength λ is of order 3–5 mm and lies in the microwave range.[4]

In analyzing the interaction of atoms with such a field we limit ourselves to a simplified energy-level structure consisting of only three states of sodium atoms, in accordance with Ref. 4: we isolate the states 36P, 37S, and 37P from the entire set of energy states existing in the given energy range (see Fig. 1a). This means we ignore multiphoton resonance processes involving more than two photons, and the other states (D and F) are not taken into consideration either because of the existing selection rules in a linearly polarized field or because of the nonresonant nature of the interaction. The simplified energy-level structure of the atoms has the form depicted in Fig. 1b. Here, following Ref. 5, we use the following notation:

E_i, with $i=1$, 2, 3, stands for the values of the energies of the given states (we ignore the fine structure of P states);

$\Delta_2=(E_2-E_1)/\hbar-\omega$ is the detuning of the microwave field from resonance for the 36P–37S levels;

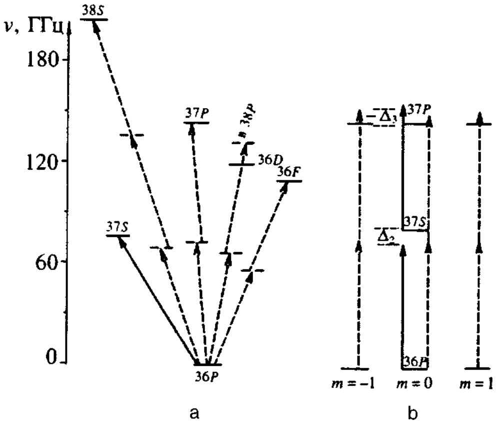


FIG. 1. The energy-level diagram of the Rydberg states of the sodium atom near $n=36$. The dashed lines indicate multiphoton transitions, and the solid lines one-photon transitions.

$\Delta_3=(E_3-E_2)/\hbar-\omega$ is the detuning of the microwave field from resonance for the $37S$–$37P$ levels;

$\Delta_{32}=(E_3-E_1)/\hbar-2\omega=\Delta_3-\Delta_2$ is the detuning of the microwave field from two-photon resonance of the $36P$–$37P$ levels.

The following aspects set our model apart from the earlier models of three-level atoms resonantly interacting with light fields.

1. The lifetimes of the Rydberg states increase with the principal quantum number in proportion to n^3–n^5. For the values of the principal quantum number considered here ($n=36$ and 37) these lifetimes reach several hundred microseconds. This means that processes of spontaneous emission into the vacuum modes of the microwave field can be treated by perturbation techniques. For Rydberg atoms the incoherent processes of interaction with the thermal-reservoir modes become most important. However, when the temperature of the surrounding medium is fairly low, this type of incoherent interaction can also be considered a perturbation. Therefore, to a high accuracy we can consider the processes of resonant interaction of Rydberg atoms and microwave fields as being coherent to lowest order. This approximation is valid even when the transverse profiles of the microwave field are fairly broad, of order $r_0\sim 10$ cm.

2. The dipole moment of the transitions increases considerably since geometrically Rydberg atoms are large. The corresponding dipole moments d are of order $10^3 ea_0$, where a_0 is the Bohr radius. Hence the effect of saturation of resonant transitions in microwave fields sets in at fairly low intensities. For example, for the transitions in sodium under discussion the saturation intensity I_{sat} is of order $5\times 10^{-7}\,\mathrm{W/cm^2}$.

3. The momentum $p_{\mathrm{ph}}=\hbar k=h/\lambda$ transferred from the field to an atom as a result of a single act of stimulated absorption is roughly 10^4–10^3 smaller than the momentum of recoil from an optical photon.

4. The distribution of some groups of levels (for example, S and P) is almost uniform, so that multiphoton resonances also play an important role (in addition to one-photon resonances). Here, however, we limit ourselves to two-photon processes.

If we limit ourselves to coherent processes of the interaction of atoms and a microwave field, it is sufficient to analyze the corresponding Schrödinger equation for the atoms and to assume that the field is fixed and classical (Eq. (1)):

$$i\hbar\frac{\partial}{\partial t}\hat{\Psi}=(\hat{H}_0+\hat{V}_{\mathrm{int}})\hat{\Psi}. \tag{2}$$

Here $\hat{H}_0=\hat{T}_{\mathrm{kin}}+\hat{V}_0$ is the operator of the energy of the atom in the absence of the interaction with the microwave field. Note that the given operator determines the dynamics of the atom in relation to both translational degrees of freedom (the kinetic-energy operator $\hat{T}_{\mathrm{kin}}=-\hbar\nabla^2/2m$) and internal degrees of freedom (the operator $\hat{V}_0$). The energy level diagram in Fig. 1 determines the eigenvalues E_i of the operator $\hat{V}_0$, while the corresponding wave functions φ_i^0 determine the probability amplitudes of the corresponding states. Note the following feature of the model presented in Fig. 1b: all the Zeeman sublevels present can be divided into three independent groups, in accordance with the stimulated transitions that emerge in the model. In the first, and most important, group the P-atoms with angular-momentum projections $m=0$ and the S-atoms ($37S$) participate in the coherent interaction with the linearly polarized field. The dynamics of the populations of these states is determined by one-photon and two-photon resonant transitions. The other two independent subgroups are formed by the $36P$ and $37P$ sublevels with projections $m=\pm 1$ and are connected by two-photon transitions. However, because of the absence of an intermediate real level (see Fig. 1b) the interaction with the field is nonresonant. There are two reasons why the dynamics of these subgroups will not be examined here: first, because of the smallness of the resulting effects due to the nonresonant nature of the interaction; second, because these processes contribute nothing to the gradient radiative force (25).$^{1)}$

In Eq. (2), $\hat{V}_{\mathrm{int}}=-\hat{\mathbf{d}}\cdot\mathbf{E}$ is the operator of the energy of the interaction of atoms with the microwave field ($\hat{\mathbf{d}}$ is the atomic dipole moment operator), and $\hat{\Psi}$ is the column vector of the atomic wave function,

$$\hat{\Psi}=\begin{pmatrix}\Psi_1\\ \Psi_2\\ \Psi_3\end{pmatrix}, \tag{3}$$

which in the representation of the self-energy states $\{\varphi_i^0\}$, of a ''bare'' atom described earlier couple the corresponding probability amplitudes Ψ_i (here $i=1$ corresponds to $36P$ at $m=0$, $i=2$ corresponds to $37S$, and $i=3$ corresponds to $37P$ at $m=0$), whose squares $|\Psi_i|^2$ determine the populations of the given states.

Since we are considering only coherent atom–field interaction process, specifying the initial conditions is also an important part of the model:

$$\hat{\Psi}(t=0)=\begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix}. \tag{4}$$

When we are dealing with optical transitions from the ground state of the atom, the initial conditions are usually chosen to be $C_1=0$ and $C_2=C_3=0$. However, specifying the initial conditions in our case is a problem in its own right since, firstly, the $36P$ state with the lowest energy is not the ground state at all (it is obtained as a result of a certain multistage process[4]) and, secondly, groups of Rydberg atoms with closely spaced energy values are actually formed in the excitation process.

Therefore, usually all three probability amplitudes C_i are finite. However, in analyzing specific results we mainly examine, for the sake of simplicity, the ideal case where in the process of formation of Rydberg atoms all the atoms accumulate in the $36P$ sublevel with $m=0$, i.e., $C_1=1$ and $C_2=C_3=0$.

3. QUASIENERGY ATOMIC STATES

We analyze the dynamics of the atoms with respect to the internal degrees of freedom in their interaction with an electromagnetic wave by employing the quasienergy method.[6] Let us write the components of the atomic wave function in the form of an expansion in the time harmonics of the microwave field:

$$\Psi_p(t)=\sum_{l=-\infty}^{\infty}\Psi_{p,l}\exp[il\omega t].$$

Next we use the resonance approximation, which means ignoring the contribution of the terms in Eq. (2) that oscillate in time. As demonstrated by Akulin and Karlov,[6] in such a model it is sufficient to analyze the approximate closed system of equations for the three following components:

$$\Psi_{1,0}=a(t)\exp\left(-i\frac{E_1}{\hbar}t\right),$$

$$\Psi_{2,-1}=b(t)\exp\left(-i\frac{E_1}{\hbar}t\right), \tag{5}$$

$$\Psi_{3,-2}=c(t)\exp\left(-i\frac{E_1}{\hbar}t\right).$$

To this we add one more approximation: at the first stage we ignore the contribution of the kinetic-energy operator. Note that in models in which the atoms interact with spatially uniform electromagnetic fields (a traveling wave) the kinetic energy can be taken into account fairly simply by allowing for the corresponding Doppler shifts $\Delta_D=\mathbf{k}\cdot\mathbf{v}$. In our model we examine the case of the spatially nonuniform field (1) (a standing wave). The fact that in the atom–field interaction we ignore the atom's kinetic energy means that the atomic beam is well-collimated: the time $t_{\text{int}}=2r_0/v_T$ of the atom–field interaction is not sufficiently long for the atom to be shifted along the field by a distance of the order of one wavelength:

$$v_z t_{\text{int}}\ll\lambda. \tag{6}$$

Due to the relatively large wavelength of the microwave fields this condition is easily met even for beams with a divergence angle

$$\Delta\alpha=\frac{v_z}{v_T}\sim\frac{\lambda}{r_0}\approx 10^{-2},$$

where the transverse size of the microwave field in the x direction is of order $r_0\sim 10$ cm.

The corresponding system of equations is

$$\frac{\partial a(\mathbf{r},t)}{\partial t}=-i\Omega_{12}(\mathbf{r})b(\mathbf{r},t),$$

$$\frac{\partial b(\mathbf{r},t)}{\partial t}=-i\Delta_2 b(\mathbf{r},t)-i\Omega_{12}^*(\mathbf{r})a(\mathbf{r},t)-i\Omega_{23}(\mathbf{r})c(t), \tag{7}$$

$$\frac{\partial c(\mathbf{r},t)}{\partial t}=-i\Delta_{32}c(\mathbf{r},t)-i\Omega_{23}^*(\mathbf{r})b(\mathbf{r},t).$$

Here we have introduced the following notation for the Rabi frequencies:

$$\Omega_{ij}(\mathbf{r})=\Omega^0\sin(kz),\quad \Omega_{ij}=\Omega_{ji}^*, \tag{8}$$

$$\Omega^0=\frac{d_{ij}E_0}{\hbar}, \tag{9}$$

where the d_{ij} are the dipole moments of the corresponding transitions.

Note that by ignoring the contribution of the kinetic-energy operator we ensure that the system of equations (7) describes the dynamics of the internal degrees of freedom of an atom. The effect of translational motion on the dynamics of the internal degrees of freedom of atoms can be large in regions where the field is low, i.e., at the nodes of a standing wave. As is well known, in this case the Landau–Zener transitions between states with closely spaced energy levels become important. In what follows this aspect is studied separately.

We examine the solution of the system of equation (7) by diagonalizing the system and finding its eigenvalues and eigenvectors. In matrix form the system (7) becomes

$$i\frac{\partial}{\partial t}\hat{X}(\mathbf{r},t)=\hat{A}(\mathbf{r})\hat{X}(\mathbf{r},t),$$

$$\hat{X}=\begin{pmatrix} a \\ b \\ c \end{pmatrix},\quad \hat{A}=\begin{pmatrix} 0 & \Omega_{12} & 0 \\ \Omega_{12}^* & \Delta_2 & \Omega_{23} \\ 0 & \Omega_{23}^* & \Delta_{32} \end{pmatrix}. \tag{10}$$

The eigenvalues of the matrix $\hat{A}$ are given by the following Cardan formulas:

$$\lambda_1(s_1+s_2)-\frac{a_2}{3},$$

$$\lambda_{2,3}=-\frac{s_1+s_2}{2}-\frac{a_2}{3}\pm i\frac{\sqrt{3}}{2}(s_1-s_2),$$

$$s_{1,2}=\sqrt[3]{r\pm\sqrt{q^3+r^2}},\qquad q=\frac{1}{3}a_1-\frac{1}{9}a_2^2,$$

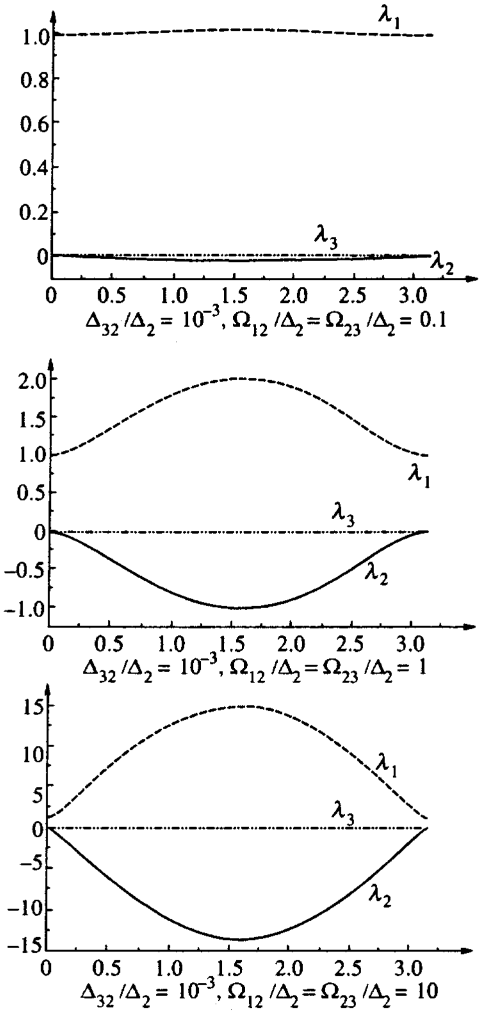


FIG. 2. Spatial dependence of the eigenvalues of the matrix A for different detunings and intensities of the microwave field.

$$r=\frac{1}{6}(a_1a_2-3a_0)-\frac{1}{27}a_2^3,\quad a_0=\Delta_2|\Omega_{12}|^2,$$

$$a_1=\Delta_2\Delta_{32}-|\Omega_{23}|^2-|\Omega_{12}|^2,\quad a_2=-\Delta_2-\Delta_{32}.$$

Figure 2 depicts the spatial dependence of the given eigenvalues at half of the wavelength of the microwave field.

The corresponding eigenvectors are defined as

$$\hat{\Lambda}_i=N_i^{-1}\begin{pmatrix}1\\ \lambda_i/\Omega_{12}\\ (\Omega_{23}^*\lambda_i)/[\Omega_{12}(\lambda_i-\Delta_{32})]\end{pmatrix}\tag{11}$$

with the normalization constants

$$N_i=\sqrt{1+\frac{\lambda_i^2}{|\Omega_{12}|^2}+\frac{\lambda_i^2|\Omega_{23}|^2}{(\lambda_i-\Delta_{32})^2|\Omega_{12}|^2}}\ .\tag{12}$$

Note the important physical essence of the given states: they describe ''dressed'' states of an atom, with the eigenvalues λ_i (Fig. 2) determining, among other things, level repulsion caused by the switch-on of the field.[6]

Analysis of Fig. 2 shows that under two-photon resonance conditions, when $\Delta_{32}\ll\Delta_2,\Omega_{23,12}$ holds, near the field's nodes the energies the states 2 and 3 are close, $|\lambda_2-\lambda_3|\approx\Delta_{32}$. Identification of the corresponding dressed states (11) shows that for Δ_2 and Δ_{32} positive the state Λ_1 corresponds in the node region $\Omega\to 0$ to the bare state 37S, Λ_2 corresponds to 36P, and Λ_3 corresponds to 37P. When Δ_{32} is negative, the situation changes in the following manner:[2)] Λ_3 corresponds to 36P, and Λ_2 to 37P.

In the approximation of small values of Δ_{32}, the energy values are given by the following formulas:

$$\lambda_{1,2}=\frac{\Delta_2\pm\sqrt{\Delta_2^2+4(|\Omega_{23}|^2+|\Omega_{12}|^2)}}{2}+\delta\lambda_{1,2},\tag{13}$$

$$\lambda_3=\delta\lambda_3,$$

$$\delta\lambda_1=\Delta_{32}\frac{|\Omega_{12}|^2}{|\Omega_{23}|^2+|\Omega_{12}|^2},$$

$$\delta\lambda_{2,3}=\Delta_{32}\frac{|\Omega_{23}|^2}{2(|\Omega_{23}|^2+|\Omega_{12}|^2)}\times\left[1\mp\frac{\Delta_2}{\sqrt{\Delta_2^2+4(|\Omega_{23}|^2+|\Omega_{12}|^2)}}\right].\tag{14}$$

As noted earlier, in the regions where λ_2 and λ_3 are close the effect of the kinetic-energy operator cannot be ignored. Here quantum hopping from state λ_2 to state λ_3 (or in the opposite direction) become possible (Landau–Zener transitions; see Ref. 5). Below we calculate the probabilities of such transitions.

The general solution of the Schrödinger equation (2) can be expressed in terms of the eigenvectors defined above in accordance with the initial conditions as follows:

$$\hat{\tilde{\Psi}}(t)=\sum_{s=1}^{3}a_s\hat{\Lambda}_s\exp(-i\lambda_s t),\tag{15}$$

$$\hat{\tilde{\Psi}}(t)=\begin{pmatrix}\Psi_1\exp\left(i\frac{E_1}{\hbar}t\right)\\ \Psi_2\exp\left(i\frac{E_1}{\hbar}t+i\omega t\right)\\ \Psi_3\exp\left(i\frac{E_1}{\hbar}t+2i\omega t\right)\end{pmatrix},$$

where the factors a_s can be found from the initial conditions at $t=0$ and $\Omega_{23,12}=0$.

For instance, at $t=0$ only the 36P sublevel is populated,

$$\begin{aligned}&C_1=a_2=1,\quad C_{2,3}=a_{1,3}\ \text{for}\ \Delta_{32}>0,\\ &C_1=a_3=1,\quad C_{2,3}=a_{1,2}\ \text{for}\ \Delta_{32}<0.\end{aligned}\tag{16}$$

Figure 3 depicts the distribution of the probability amplitudes for these initial conditions with different values of the parameters of the problem. The upper diagram which shows the case of a two-photon resonance when the population of the intermediate level proves is low, is especially noteworthy.

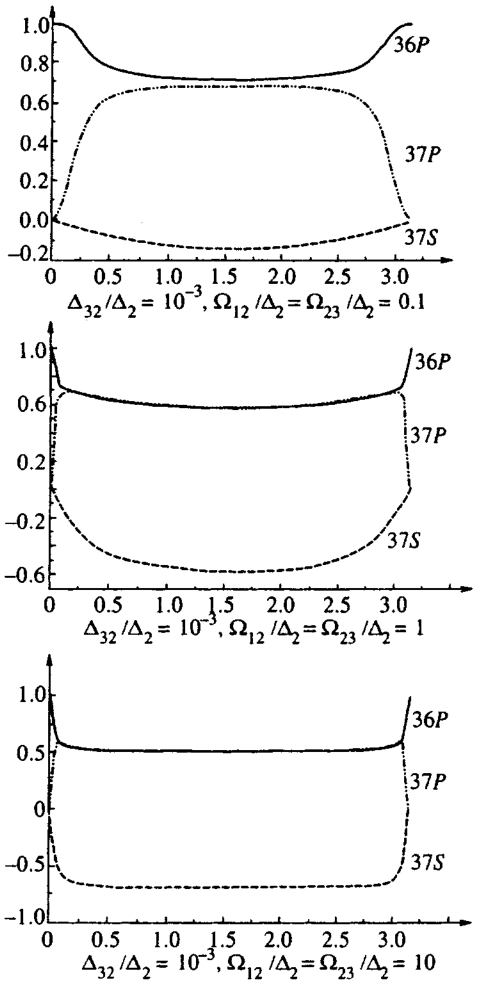


FIG. 3. The probability amplitudes of various states with the initial conditions $a_2=1$ and $a_1=a_3=0$.

4. LANDAU–ZENER TRANSITIONS UNDER TWO-PHOTON RESONANCE CONDITIONS

Let us estimate the probability of the above Landau–Zener transitions between the states $\hat{\Lambda}_{2,3}$ near the nodes of a standing wave. To this end we linearize the initial Rabi frequencies,

$$\Omega_{ij}(\mathbf{r})\approx\frac{d_{ij}E_0}{\hbar}kz=\Omega_{ij}^0kz,$$

and employ the condition that the translational motion of the center of mass of an atom is given, $\mathbf{R}(t)\approx\mathbf{R}_0+\mathbf{v}t$ (see Ref. 5). Under these simplifying assumptions the initial Schrödinger equation (2) acquires the form

$$i\frac{da}{dt}=\Omega_{12}^0kzb,$$

$$i\frac{db}{dt}=(\Omega_{12}^0)^*kza+\Delta_2b+\Omega_{23}^0kzc, \tag{17}$$

$$i\frac{dc}{dt}=(\Omega_{23}^0)^*kzb+\Delta_{32}c.$$

Under two-photon resonance conditions, the probability amplitude for the 37S state we have $|b|\ll|a|,|c|$ by virtue of the small parameter $|\Omega_{12}^3|/\Delta_2$, and we can use the following approximation:

$$b\approx-\frac{(\Omega_{12}^0)^*kza+\Omega_{23}kzc}{\Delta_2}. \tag{18}$$

As a result, the system of equations (17) is reduced to a system of two equations for the probability amplitudes a and c:

$$i\frac{da}{dt}=iv_z\frac{da}{dz}=-\frac{V_1^2(kz)^2a+V_1V_2(kz)^2c}{\Delta_2}, \tag{19}$$

$$i\frac{dc}{dt}=iv_z\frac{dc}{dz}=-\Delta_{32}c-\frac{V_2^2(kz)^2c+V_2V_1(kz)^2a}{\Delta_2}, \tag{20}$$

$$V_1=\Omega_{12}^0,\quad V_2=\Omega_{23}^0. \tag{21}$$

Note that the Landau–Zener Hamiltonian that follows from this system of equations differs considerably for the ordinary Hamiltonian which is linear in z for the case of two-level atoms.[5] This, in particular, leads to a situation in which it is impossible to derive a Landau–Zener transition matrix for such a Hamiltonian, contrary to the case of Ref. 5. But the probability W_{ac} of transitions from the 36P state to the 37P state can be estimated by employing the complex-time method.[7] We can assume with fairly high accuracy that $V_1\approx V_2=V$. In this approximation the corresponding probability is

$$W_{ab}=\exp\left[-\frac{1}{4}B(3/2,1/4)\left(\frac{\Delta_{32}}{\Delta_L}\right)^{3/2}\right], \tag{22}$$

$$B(3/2,1/4)=\frac{\Gamma(3/2)\Gamma(1/4)}{\Gamma(7/4)}\approx3.5,\quad \Delta_L=\left[\frac{V^2(kv)^2}{\Delta_2}\right]^{1/3}. \tag{23}$$

The characteristic parameter here is the frequency Δ_L: for $\Delta_{32}\sim\Delta_L$ the Landau–Zener transitions become important. We estimate Δ_L for the following orders of magnitude of the quantities involved in the problem:

- for the transition in sodium with an exact two-photon resonance the detuning Δ_2 is $2\pi\times2.6$ GHz;
- the dipole moment is estimated at $1500ea_0$;
- the velocity of the atoms along the wave vector is estimated at $v_z\sim\Delta\alpha_{\text{div}}v_T=10^{-3}\times600$ m/s, where $\Delta\alpha_{\text{div}}$ is the divergence angle of the atomic beam, and v_T is the corresponding thermal velocity of the atoms; and
- $I_{\max}\sim20$ mW/cm^2 is the maximum intensity of the microwave field from the open end of the waveguide.

For these values of the parameters Δ_L reaches its maximum value, roughly $2\pi\times80$ kHz, i.e., Δ_L is extremely small, smaller in order of magnitude than the fine structure ~114 and 124 MHz (Ref. 4) of the 36P and 37P states. Note, however, that in this model only the single Zeeman sublevel with the projection $m=0$ interacts resonantly with the field.

As we show below, the effect of Landau–Zener transitions near nodes can essentially diminish the gradient radiative force.

5. GRADIENT FORCE IN A MICROWAVE FIELD

By the gradient force acting on an atom in a microwave field we mean the corresponding vector $\mathbf{F}(\mathbf{r})$ entering into the Fokker–Planck equation for the atomic distribution function $f(\mathbf{r},\mathbf{p},t)$,

$$\left(\frac{\partial}{\partial t}+\mathbf{v}\cdot\nabla\right)f(\mathbf{r},\mathbf{p},t)=-\frac{\partial}{\partial\mathbf{p}}(\mathbf{F}(\mathbf{r})f(\mathbf{r},\mathbf{p},t)) \tag{24}$$

and defined as

$$\mathbf{F}=\sum_{\alpha,\beta}^{3}\nabla\Omega_{\alpha,\beta}\rho_{\beta\alpha}. \tag{25}$$

Here summation is over the entire set of states of the atom. The matrices $\rho_{\alpha\beta}$ in the adopted coherent atom–field interaction model can be expressed in terms of the corresponding probability amplitudes of the levels:

$$\rho_{\alpha\beta}=\Psi_{\alpha}^{*}\Psi_{\beta}. \tag{26}$$

For the solution (15) obtained earlier this means that

$$\mathbf{F}=\sum_{i,j=1}^{3}\sum_{m,n=1}^{3}\nabla\Omega_{i,j}\exp(i(\lambda_m-\lambda_n)t)\Lambda_{m,j}^{*}\Lambda_{n,i}a_m^{*}a_n, \tag{27}$$

where m and n label the dressed states defined by (11), and i and j label the matrix elements of these states.

Let us now show that the corrections oscillating in time can be ignored. For instance, in the particular case of positive detunings Δ_{32} considered here the smallest difference is

$$|\lambda_2-\lambda_3|\sim 2\frac{|\Omega|^2}{\Delta_2}. \tag{28}$$

At such intensities, for $\Omega\sim 0,1$ and $\Delta_2\approx 1$ GHz, in the time $\tau=d_0/v_T\approx 35\ \mu$s it takes the atoms to pass through the region with the microwave field, the corresponding terms in (29) perform about $N_{\text{osc}}=10^4$–10^5 oscillations. Hence in averaging over the velocities v_x these contributions can be discarded, provided that the experiment does select groups of atoms with fixed longitudinal velocities:

$$\mathbf{F}\approx\sum_{i,j=1}^{3}\sum_{n=1}^{3}\nabla\Omega_{i,j}|a_n|^2\Lambda_{n,j}^{*}\Lambda_{n,i}=\sum_{i=1}^{3}|a_i|^2F_i. \tag{29}$$

We also assume that the interaction between the microwave field and the Rydberg atoms is switched on smoothly, i.e., by allowing for an approximately Gaussian decay of the intensity of the microwave field toward the periphery we can assume that the probability of the transitions between the corresponding quasienergy states considered earlier is negligible. Due to the adiabaticity of the switch-on process the quantities a_n are determined solely by the initial conditions.3)

Let us take, for example, the boundary conditions (16). In this case the expression for the force simplifies ($n=2$ at $\Delta_{32}>0$):

$$\mathbf{F}\approx\nabla\Omega_{12}\Lambda_{2,2}^{*}\Lambda_{2,1}+\nabla\Omega_{23}\Lambda_{2,3}^{*}\Lambda_{2,2}+\text{ c.c.} \tag{30}$$

If we now take values of the problem's parameters such that Δ_{32} can be neglected in comparison to $\Omega_{12,23}$, Δ_2, and Δ_{23}, we can use the corresponding approximate value (13) for λ_2: As a result we arrive at the following expression for the gradient radiative force:

$$F_z=-\frac{\partial V_{\text{eff}}}{\partial z}, \tag{31}$$

where the effective potential is defined as

$$V_{\text{eff}}=\frac{\hbar}{2}\sqrt{(\Delta_2)^2+4(|\Omega_{12}(z)|^2+|\Omega_{23}(z)|^2)}. \tag{32}$$

Below we list the important features of this potential.

1. The potential resembles what is known as the nonresonant potential[5] in the theory of coherent interaction of two-level atoms with the field of a standing light wave. Here the inclusion of an additional, third, level into the picture of the resonant interaction with the field effectively increases the potential (32) due to the additional term Ω_{23} by a factor of approximately $\sqrt{2}$. It can be suggested that allowing for three-photon and higher-order processes may further increase the effectiveness of the potential by a factor of $\sqrt{m}$, where m is the corresponding order of the process involving photons.

2. In the case of a two-photon resonance, when

$$\frac{\Omega_{12}}{\Delta_2}\ll 1,\quad \Delta_{32}\ll\Omega,\Delta_2, \tag{33}$$

with $\Omega_{12}\approx\Omega_{23}=\Omega(z)$ for Rydberg atoms, the effective potential becomes the well-known potential for two-photon transitions,[6]

$$V_{\text{eff}}=\hbar\frac{\Omega^2(z)}{\Delta_2}. \tag{34}$$

3. Finally, we note that the results are independent of the sign of Δ_{32} and are determined mainly by the value of Δ_2: when the microwave field intensity satisfies $\Omega>\Delta_2$, the amplitude of the effective potential (32) depends on Ω almost linearly. Actually, at such microwave field intensities we cannot limit ourselves to two-photon processes in the model, and we must take into account higher-order multiphoton processes.

If, however, the initial conditions are defined in a more complicated way and there are components with a_1, $a_3\neq 0$, we must use the general formula (29). Figure 4 depicts the graphs of the corresponding components of the constituent forces in (29) at the half-wavelength of the field. In particular, we see that as the component a_3 increases due to Landau–Zener transitions, the overall force diminishes, since $|F_3|\ll|F_2|$.

6. THE KINETICS OF RYDBERG ATOMS IN A MICROWAVE FIELD

The kinetics of a beam of Rydberg atoms in microwave fields is described by Eq. (24), with (30) as the force. Let us examine the time-independent pattern of the atoms scattered by the microwave field. In this case the initial kinetic equation assumes the form

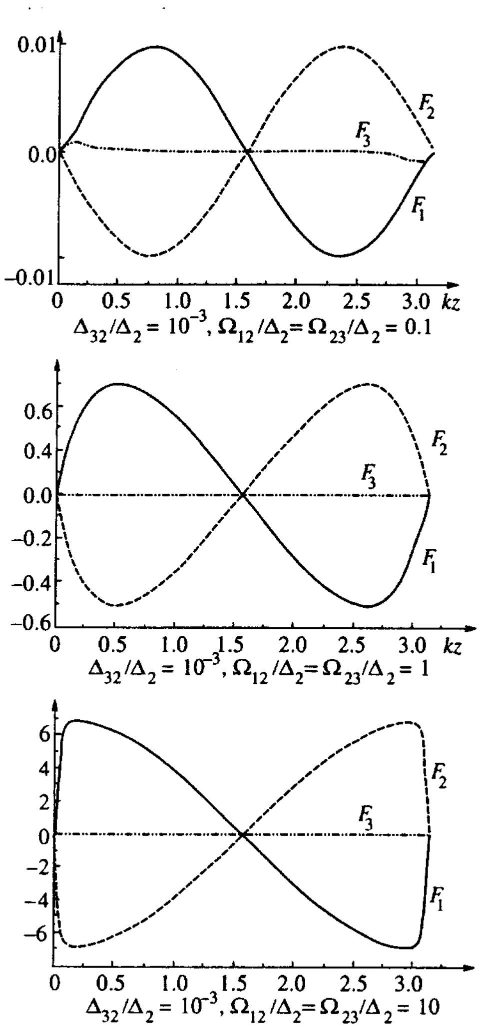


FIG. 4. The components of the constituent forces for different detunings and intensities of the microwave field in relative units.

$$\left(v_x \frac{\partial}{\partial x} + v_z \frac{\partial}{\partial z} \right) f(\mathbf{r},\mathbf{v}) = -F_z(z)\varphi(x) \frac{\partial}{\partial p_z} f(r,v), \tag{35}$$

where the function $\varphi(x)$ determines the profile of the microwave amplitude in the direction x of propagation of the atomic beam.

An important aspect of solving Eq. (35) is formulating the appropriate boundary conditions. In this case the boundary condition is fixed by the distribution function for the atoms at the ''muzzle'' of the atomic gun ($x=0$). We assume that the atoms are thermalized in the longitudinal (v_x) and transverse (v_z) projections of the velocities but that the temperatures are different. The corresponding velocity distributions are shaped like Maxwellian distributions.[8] Similarly, the spatial distribution, or profile, of the atomic beam in the cross section (along the z axis) is determined by a Gaussian law. Thus, the boundary condition can be written as

$$f(x=0, z, v_x, v_z) = \frac{1}{N} \exp\left(-\frac{(z-z_0)^2}{r_b^2} - \frac{v_x^2}{v_T^2} - \frac{v_z^2}{v_\perp^2} \right) \theta(v_x), \tag{36}$$

where N is the normalization factor, r_b is the radius of the atomic beam, z_0 specifies the position of the center of the atomic beam along the z axis, $\theta(y)$ is the Heaviside step function, $v_\perp = \Delta\alpha v_T$ is the transverse thermal velocity of the atoms in the beam, $\Delta\alpha$ is the beam divergence angle in radians (the extent to which the beam is collimated), and v_T is the longitudinal thermal velocity.

Let us solve Eq. (35) by perturbation techniques. Here for the zeroth approximation we can take the solution of the initial kinetic equation (35) with a zero right-hand side,

$$f_0(x>0,z,v_x,v_z) = \frac{1}{N} \exp\left(-\frac{(z-z_0-(v_z/v_x)x)^2}{r_b^2} - \frac{v_x^2}{v_T^2} - \frac{v_z^2}{v_\perp^2} \right) \theta(v_x). \tag{37}$$

It describes the spread of the transverse profile of the atomic beam caused by thermal motion of the atoms along the z axis. If the effect of the microwave field on the atomic-beam profile is regarded as a perturbation, the initial distribution function can be written as

$$f(\mathbf{r},\mathbf{v}) \approx f_0(\mathbf{r},\mathbf{v}) + \delta f(\mathbf{r},\mathbf{v}), \tag{38}$$

where the kinetic equation for the corresponding correction is

$$\left(v_x \frac{\partial}{\partial x} + v_z \frac{\partial}{\partial z} \right) \delta f(\mathbf{r},\mathbf{v}) = -F_z(z)\varphi(x) \frac{\partial}{\partial p_z} f_0(\mathbf{r},\mathbf{v}). \tag{39}$$

The solution of this equation can be written in the form

$$\delta f(\mathbf{r},\mathbf{p}) = -\int_0^x d\xi \frac{F_z[z+(v_z/v_x)(\xi-x)]\varphi(\xi)}{v_x} \times \left\{ \frac{\partial}{\partial p_z} - \frac{\xi-x}{p_x} \frac{\partial}{\partial z} \right\} f_0[\xi, z+(v_z/v_x) \times (\xi-x), v_x, v_z]. \tag{40}$$

In particular, if we take a simplified version of the problem, where the microwave field begins immediately after the source of the atomic beam and its profile has sharp boundaries, $\varphi(x) = \theta(2r_0 - x)$ (recall that r_0 is the traverse radius of the profile of the microwave field), the computational formula for the correction following the region with the microwave field ($x>2r_0$) assumes the form

$$\delta f(x>2r_0, z, \mathbf{p}) = -\int_0^{2r_0} d\xi \frac{F_z(z+(v_z/v_x)(\xi-x))}{v_x} \left\{ \frac{\partial}{\partial p_z} - \frac{\xi-x}{p_x} \frac{\partial}{\partial z} \right\} \times f_0(\xi, z+(v_z/v_x)(\xi-x), v_x, v_z), \tag{41}$$

or, using the explicit form of the zeroth approximation (37), we obtain

$$\delta f(z,x,v_x,v_z) = -2 \int_0^{2r_0} d\xi \frac{F(z+(v_z/v_x)(\xi-x))}{v_x}$$

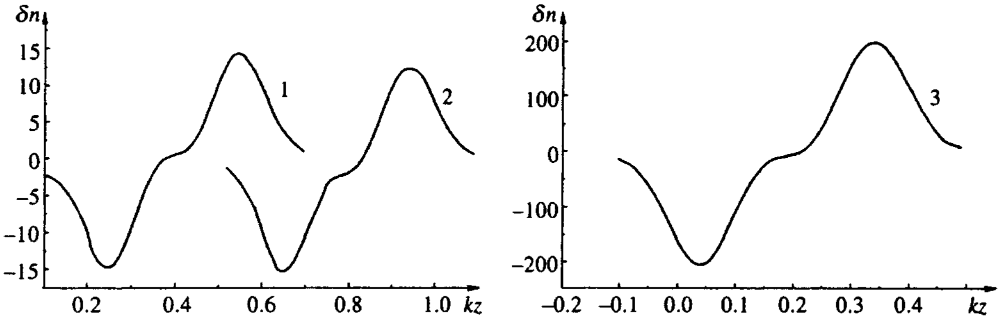


FIG. 5. The results of calculations of the spatial profile of the atomic beam for different positions of its center in relation to the nodes of a standing microwave: $kz=\pi/8$ (curve *1*), $kz=\pi/4$ (curve *2*), $kz=0.188$ (curve *3*).

$$\times\left[-\frac{p_z}{p_\perp^2}-\frac{(z-z_0-(v_z/v_x)x)\xi}{r_b^2 p_x}\right]f_0(z,x,v_z,v_x). \tag{42}$$

The perturbation theory we have just described actually amounts to an expansion in powers of the time of flight (the interaction time) of the atom through the microwave field, $t_{\text{int}}=2r_0/v_T$, and the first correction is linear in this parameter. Note that the higher−order corrections give rise to divergences, related to the large interaction times for atoms whose velocities v_x are low. As is well known, these divergences can be removed by summing the entire perturbation−theory series. In the present model, however, calculating the first correction is sufficient because of the small effect of the microwave field on the dynamics of the atoms. For example, let us estimate in order of magnitude the corresponding smallness parameter in the problem.

1. We assume that the force acting on the atom attains its maximum value, and its order of magnitude, in accordance with Eqs. (31) and (32), is estimated at $F\sim\hbar k\Omega$. For the limiting value we take the Rabi frequency at which the dynamic Stark broadening Δ_2 is of order $2\pi\times 2.6$ GHz.

2. In accordance with the coherent nature of the atom–field interaction chosen in the model, the interaction time t_{int} cannot exceed the characteristic time of radiative decay of the corresponding Rydberg states. In our case we limit this value to $t_{\text{int}}\approx 35\ \mu$s.

Then the corresponding smallness parameter in the perturbation theory described above has the form

$$\nu=\frac{Ft_{\text{int}}}{p_z}=\frac{8\Omega^2}{\sqrt{(\Delta_2)^2+8\Omega^2}}\,\frac{\hbar k t_{\text{int}}}{\Delta\alpha_{\text{div}} m v_T}, \tag{43}$$

and with the above parameters and the corresponding mass of sodium atoms we obtain $\nu\sim 0.1$–1. Note that for Rabi frequencies $\Omega\geqslant\Delta_2$ the parameter ν becomes greater than unity and our perturbation theory becomes invalid.

The above result (42) for the first correction to the distribution function can be simplified if we allow for the fact that in the microwave field the wavelength $\lambda=4$ mm, as a characteristic parameter for the spatial gradient for F_z, is large compared to the integration parameter: $(v_z/v_x)2r_0\sim\alpha_{\text{div}}2r_0\ll\lambda$. For instance, for microwave fields with transverse dimensions $2r_0\sim 10$ cm the value of the small parameter is

$$\frac{\alpha_{\text{div}}2r_0}{\lambda}\sim 10^{-1}.$$

Then, expanding the integrand in a Taylor series and keeping only the first correction, we arrive at the following approximate result for the given correction:

$$\delta f=f_0\frac{4F(z-(v_z/v_x)x)}{v_x}\times\left[\frac{p_z}{p_\perp^2}r_0-\frac{(z-z_0-(v_z/v_x)x)r_0^2}{r_b^2 p_x}\right]. \tag{44}$$

Further calculations of the profile of the atomic beam at the detector with a known distribution function are done according to the following algorithm:

$$n(z)=\int d\mathbf{v}\ f(z,x-L,v_x,v_z)n_0, \tag{45}$$

where L is the position of the detector of atoms along the x axis, and n_0 is the average atomic density in the beam in the initial stage.

In particular, if we plug the above correction (44) into this formula, the final expression for the correction to the density caused by the radiative force becomes

$$\delta n(z,L)=n_0(z,L)\left\langle\frac{4F(z-(v_z/v_x)L)}{v_x}\left[\frac{p_z}{p_\perp^2}r_0-\frac{(z-z_0-(v_z/v_x)L)r_0^2}{r_b^2 p_x}\right]\right\rangle_{v_z,v_x}, \tag{46}$$

where $n_0(z,L)$ is the distribution of the atomic beam density at the detector in the absence of a microwave field. The total density of the atomic beam is

$$n(z,L)=n_0(z,L)+\delta n(z,L). \tag{47}$$

Figure 5 depicts the results of calculations of the spatial profile of the atomic beam for different values of the parameters of the problem. The left diagram shows the distributions $\delta n(z)$ at the detector at point $x=L$ for the following values of the dimensionless parameters: the detuning of the two-photon transition is $\tilde{\Delta}_{32}=\Delta_{32}/\Delta_2=5\times 10^{-4}\sim 0$; the

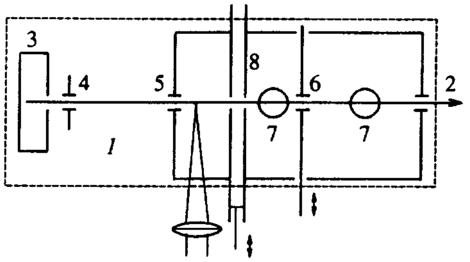


FIG. 6. Experimental setup: *1*—vacuum chamber, *2*—beam of sodium atoms, *3*—oven, *4* and *5*—diaphragms, *6*—moving slit, *7*—entrance windows of the channel multipliers, and *8*—waveguide.

Rabi frequencies are $\tilde{\Omega}_{i,j}=\Omega_{i,j}/\Delta_2=1$; the width of the microwave beam is $\tilde{r}_0=r_0/\lambda=10$; the detector position is $\tilde{L}=L/\lambda=10^3$; the divergence of the atomic beam is $\alpha=v_\perp/v_T=10^{-3}$; the width of the atomic beam is $\tilde{r}_b=r_b/\lambda=10^{-1}$; the position of the center of the atomic beam in relation to the nodes of the standing wave of the microwave field is $\tilde{z}_0=kz_0=\pi/8$ for the curve *1* and $\tilde{z}_0=\pi/4$ for the curve *2*; and the scale along the vertical axis is chosen in units of the characteristic parameter of the problem to be $\widetilde{\delta n}=\hbar\Delta_2/mv_t^2\sim10^{-4}$, which, of course, in this case is a small parameter.

The right diagram in Fig. 5 depicts the density distribution δn for the following altered parameters of the problem: Rabi frequencies $\tilde{\Omega}_{i,j}=\Omega_{i,j}/\Delta_2=10$; and the position of the center of the atomic beam is chosen by the maximum of the gradient force, i.e., at $\tilde{z}_0\approx0.188$.

An important aspect should be noted. The above correction to the density (46) becomes divergent as a result of averaging because of the contribution of atoms with low velocities $v_x\sim0$, since the time it takes such atoms to interact with the microwave field tends to infinity. The above results are given for velocities $v_x>10^{-6}v_T$.

7. EXPERIMENT

We measured the profile of a beam of Rydberg sodium atoms in two-photon resonance with a standing wave of the microwave field (at frequency 72.6 GHz) on the $36P-37P$ transition. Earlier we studied this transition in detail in two-photon microwave spectroscopy and double-Stark-resonance experiments.[4] In addition, other multiphoton resonances were also observed in the given frequency range.[9]

The experimental setup is shown schematically in Fig. 6. The $36P$ Rydberg state of a sodium atom was excited in a three-stage process $3S-3P-4S-36P$ by the radiation from three pulsed locked tunable lasers. The parameters of the lasers were close to those described in Ref. 4. The laser beams were matched on dichroic mirrors and were focussed on the vacuum chamber *1* at right angles to the atomic beam *2*. The sodium atoms were evaporated in the oven *3*, with the beam being formed by two diaphragms, *4* and *5*, which limited the initial divergence to about 3×10^{-3} rad. The Rydberg atoms were detected by the means of selective field ionization[10] in a pulsed electric field. The produced photoelectrons were gathered at the entrance windows of the channel electron multipliers *7*. The electric signals from the output of the multipliers were processed in the pulse-counting mode with an accumulation time of 10 s.

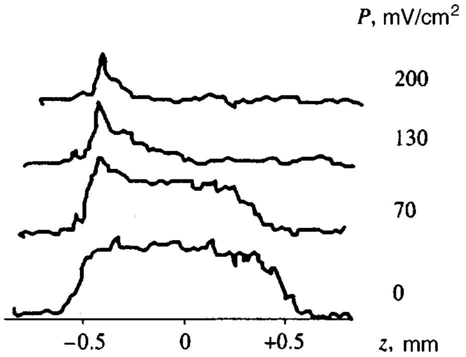


FIG. 7. Experimental traces of the profile of a beam of Rydberg sodium atoms for different intensities of microwave radiation in resonance with the $36P-37P$ two-photon transition (the atomic beam passes through the region with the maximum gradient force).

The detection device consisted of two parts. In one the Rydberg atoms were excited by laser radiation and were forced to interact with the microwave field when the beam of excited atoms was sent through the waveguide *8*. Here the signals from tuning the field frequency to the multiphoton resonances were registered. The total number of Rydberg atoms passing through the moving diaphragm *6* was measured in the second part of the device. By changing the position of the diaphragm one could obtain the distribution of the number density of the Rydberg atoms in the beam along one axis. Scattering or deflection of the beam of Rydberg atoms caused by the transfer of transverse momentum from the microwave field changes the measured beam profile. In addition we were able to vary the gradient of the microwave field in the opening of the waveguide by changing the position of the reflecting wall (piston), which ensured the formation of a quasistanding wave. The peak power flux of the microwave field could reach 0.2 W/cm^2. This was done by using a frequency-tunable oscillator with a G4-142 backward-wave valve.

Figure 7 depicts the characteristic traces of the profiles of a beam of Rydberg sodium atoms with and without a microwave field. The frequency of the microwave radiation was tuned exactly in resonance with the two-photon transition $36P-37P$. The initial profile of the beam was a nearly symmetric trapezoidal distribution with a width of about 1 mm, determined by the size of the collimating diaphragms. The contribution of the measuring device to the width was less than 0.3 mm. The signal noise was related primarily to the fluctuations in the frequency and amplitude of the laser radiation. Also, a slight constant background noise was observed, which was attributed to the charged particles from the magnetic discharge pump which produced the high vacuum, entering into the registration area. Introduction of the microwave field was found to change the profile of the atomic beam considerably, depending on how the standing wave was formed in the waveguide. Figure 7 corresponds to tuning to one of the slopes of the standing waves, where the gradient force is at its maximum. We also see that the distribution of the Rydberg atoms in the cross section after the atoms have gone through the opening in the waveguide depends on the intensity of the microwave field and on the point of interaction. When the minimum of the microwave was formed in the opening, the profile of the beam of Rydberg atoms was found to change insignificantly.

Analysis of the experimental results suggests two basic physical effects that determine the change in the profile of the beam of Rydberg atoms when the atoms pass through a nonuniform resonant microwave field.

The first, is the mechanical effect produced by the microwave field. This effect is caused by the gradient force and was discussed above. In our experiment we were able to deflect the Rydberg atoms to angles comparable to the initial divergence of the atomic beam. For this reason the atoms passing through the antinodes of the standing microwave are deflected to larger angles than when passing through the slopes of the wave. According to the above theory (see Sec. 5), the dependence of the angle of deflection of the beam by the gradient force on the velocity and on the time it takes the Rydberg atoms to interact with the microwave field complicates the pattern of the spatial distribution of the atoms considerably. Here, by allowing for the distribution of the atoms over the longitudinal and transverse velocities we arrive at a complex scattering pattern, and the microwave field may act as a positive or negative lens. In addition, the position of the detector also affects the beam profile.

Actually, the narrowing of the atomic beam observed in Fig. 7 corresponds to the focussing mode. But at the same time we see that the total number of the Rydberg atoms is not conserved. This is due either to scattering by very large angles, as a result of which the signal from the Rydberg atoms falls below the background noise, or to quenching of the Rydberg atoms in the microwave field. Such quenching may be caused by ionization coupling with the continuum of the type of diffusion ionization.[11] However, in our experiments the probability of such ionization is negligible since the ionization threshold corresponds to microwave fields whose strength must be higher by a factor of 10 to 100 for states with such values of n.

A more likely explanation is the phenomenon of radiative quenching of Rydberg states in an intense microwave field, related to the transition of the atoms to lower states. For states $36P$ the lifetime of Rydberg atoms, with allowance for the decrease in lifetime due to the interaction with thermal radiation at liquid nitrogen temperatures is 170 μs and exceeds, albeit insignificantly, the time it takes the Rydberg atoms with typical velocities to travel from the point of laser excitation to the second channel multiplier. In this case practically all Rydberg atoms are detected. Depending on the principal quantum number n, the lifetime of Rydberg states increases at least like n^3, and for lower levels it decreases according to the same law. This could lead to a situation in which for essentially lower states not all the Rydberg atoms that have interacted with microwave radiation would be detected. Calculations have shown that at the frequency of the microwave radiation used in the experiment there are no transitions, either resonant or multiphoton, from the excited $36P$ state of sodium to lower states. But in strong microwave fields the resonance defect may be balanced by the dynamic Stark effect, and then the resulting resonances[9] could drive up the probabilities of radiative transitions, including multiphoton transitions, considerably.

The detected profiles of a beam of Rydberg sodium atoms depicted in Fig. 7, with the atoms passing through a spatially nonuniform microwave field formed in the waveguide, most probably reflect a combination of kinetic effects and multiphoton radiative quenching of Rydberg states. Moreover, strong signals have been detected in the range of intensities of the microwave field in which the three-level model, for which there is a reliable theory, ceases to be valid.

To observe the kinetic effects more clearly and to compare the results with the above theory, we need experiments at minimum intensities of the microwave field, which requires increasing the sensitivity of the device considerably and retaining spatial resolution. Nevertheless, our experiments have revealed a considerable variation in the profile of a beam of Rydberg atoms in the presence of a spatially nonuniform microwave field, which qualitatively supports our picture of resonant scattering of Rydberg atoms. To ensure quantitative agreement the experimental setup must be redesigned, and other resonances, where the contribution of radiative quenching would be less important, should be found. Theoretical calculations that allow for higher-order multiphoton interaction must also be done. Estimates show that at high intensities of the microwave field the contribution of the above factor may be extremely significant.

8. CONCLUSION

By using the present method we can study, at least in principle, the scattering of atoms when the interaction in the electromagnetic fields is coherent, to any photon order. Here not only can we calculate the effect of a force on an individual atom but we can also estimate the kinetic characteristics of the atomic beam as a whole. Below we list the possible avenues of research.

1.Taking into account a large number of levels in the limit of an m-photon resonance.

2. Taking into account the perturbations of relaxation processes for Rydberg states.

3. Studying other field configurations. Here a standing wave is not a good example, since for a microwave field the spatial gradients are low. Note that the results given in this paper (e.g., (32)) are valid not only for a standing wave but also for a wave of arbitrary spatial configuration (to be sure, a linearly polarized field). Thus, the choice of the appropriate sharp spatial nonuniformity in a microwave field with a gradient extending over a region with dimensions $\ll\lambda$ could boost the scattering effect considerably.

4. Landau–Zener transitions require further study. In particular, it would be interesting to select models with corresponding field and atomic configurations, where the Landau–Zener transitions would have a strong effect on the scattering of atoms by the field.

5. Various polarization field configurations, possibly including a static field, could be used as an instrument to control scattering processes.

Partial financial support for this work was provided by the Universities of Russia Program and the Russian Fund for Fundamental Research (Grants No. 93-02-15299and No. 95-02-04752).

[1)] If one allows only for the dipole approximation in the interaction operator, the value of the gradient radiative force, as Eq. (25) implies, is determined only by the off-diagonal elements of the density matrix whose quantum numbers are fixed by the selection rules corresponding to one-photon absorption–emission processes. Hence the effect of multiphoton processes can manifest itself only indirectly. In particular, the groups of Zeeman sublevels and the corresponding transitions unimportant in the given model could become important if another type of polarization of the microwave field (circular, for instance) is added to the linear polarization.

[2)] Note that all subsequent formulas refer to the case $\Delta_2 > 0$, which is important for practical reasons since we usually have $E_2 - E_1 > E_3 - E_2$.

[3)] Nonadiabaticity also manifests itself near nodes because of Landau–Zener transitions. Here, however, we ignore such transitions.

[1] R. F. Stebbings and F. B. Danning (eds.), *Rydberg States of Atoms and Molecules*, Cambridge Univ. Press, New York (1983), p.1.

[2] D. F. Zaretskiĭ and M. M. Yudkevich, Opt. Spektrosk. **67**, 742 (1989) [Opt. Spectrosc. (USSR) **67**, 752 (1989)].

[3] I. M. Beterov and I. I. Ryabtsev, Laser Phys. **4**, 953 (1994).

[4] I. M. Beterov, A. O. Vyrodov, I. I. Ryabtsev, and N. V. Fateev, Zh. Éksp. Teor. Fiz. **101**, 1154 (1992) [Sov. Phys. JETP **74**, 616 (1992)].

[5] A. P. Kazantsev, G. I. Surdutovich, and V. P. Yakovlev, *Mechanical Action of Light on Atoms* World Scientific, Singapore (1990).

[6] V. M. Akulin and N. V. Karlov, *Intense Resonant Interactions in Quantum Electrodynamics*, Springer, Berlin (1991).

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

[8] E. M. Lifshitz and Pitaevskiĭ, *Physical Kinetics*, Pergamon Press, Oxford (1981).

[9] I. M. Beterov, G. L. Vasilenko, I. I. Ryabtsev *et al.*, in *Intense Laser Phenomena* (Series on Optics and Photonics), edited by I. Kiyan and M. Ivanov, World Scientific, Singapore (1992), p. 100.

[10] V. S. Letokhov, *Laser Photoionization Spectroscopy*, Academic, Orlando (1987).

[11] N. B. Delone, V. P. Kraınov, and D. L. Shepelyanskiĭ, Usp. Fiz. Nauk **140**, 355 (1983) [Sov. Phys. Uspekhi **26**, 551 (1983)].

Translated by Eugene Yankovsky

Exact nonrelativistic expressions for the tensor for scattering of light by atoms

M. A. Preobrazhenskiĭ

Voronezh State Academy of Architecture and Civil Engineering, 394000 Voronezh, Russia

(Submitted 7 June 1996)
Zh. Éksp. Teor. Fiz. **111**, 816–830 (March 1997)

Exact nonrelativistic analytical expressions are derived for dipole two-photon transitions between arbitrary multiplets of the hydrogen atom and positive hydrogenlike ions. The result is expressed in terms of a single Gauss hypergeometric function and polynomials whose degrees increase linearly with the number of nodes of the bound states of the quantum system. The cross sections of elastic scattering of light by K- and L-shells of the hydrogen atom are given as an example. It is demonstrated that by expanding the discrete-spectrum wave functions in ultraspherical polynomials it is also possible to obtain analytical expressions of the cross sections of two-photon transitions between states described by the Simons model potential. The basis consisting of Chebyshev polynomials is shown to be the best expansion basis, and the coefficients of such an expansion are given for a broad range of parameters of the problem. Calculation of the polarizability of the $5S$-state of the rubidium atom is chosen as an example. Finally, the results are compared with the experimental data and the theoretical results of other researchers. [S1063-7761(97)00403-4]

1. INTRODUCTION

As is known,[1] the matrix elements C_{ik} of the dipole tensor of scattering of light by an atom can be expressed in terms of the wave functions of the initial ($|1\rangle$) and final $|2\rangle$ states as follows:

$$C_{ik}=\langle 2|d_i G_{E_1+\omega_1} d_k + d_k G_{E_1-\omega_2} d_i|1\rangle, \tag{1}$$

where the d_i are the components of the dipole-interaction operator, G is the atomic Green's function, E_1 is the energy of the initial level, and $\omega_{1,2}$ are the energies of the absorbed and scattered photons, respectively. (Here and in what follows the atomic system of units is employed.) The dipole approximation is applicable in studies of the interaction between bound atomic states and an electromagnetic wave because the linear dimensions r of atoms are small compared to the wavelength λ, i.e., $r \ll \lambda$. In the optical range this is true for states whose principal quantum number n does not exceed 30 (see Ref. 2).

The C_{ik} can be used to express the probabilities of elastic and inelastic scattering of light and the shifts of energy levels in an electromagnetic field. For instance, for an isolated atomic level, the tensor C_{ik} degenerates into the polarizability tensor α_{ik}, related to the level shift ΔE as follows:

$$\Delta E = -\frac{\alpha_{ik}\mathscr{E}_i\mathscr{E}_k}{2}, \tag{2}$$

where the $\mathscr{E}_i$ are the components of the electric field vector. The atomic states with positive orbital quantum numbers l cannot be isolated and are degenerate at least in M, the projection of the total angular momentum J. The dependence of C_{ik} on both M and the indices i and k can be found in general form via the Wigner–Eckart theorem, which makes it possible to determine the dependence of ΔE on magnetic quantum numbers.[3] Here, for an isolated atom, it has proved more convenient to express ΔE not in terms of the tensor α_{ik} but in terms of the scalar and tensor polarizabilities proportional to the components of this tensor.

Hydrogenlike and Rydberg states are also degenerate in l. Generally speaking, the contribution to the shift of an atomic level is provided not only by the terms of α_{ik} proportional to the square of the field strength but also by terms linear in $\mathscr{E}$. It is the linear terms that completely determine ΔE in the static limit. However, in the range of optical frequencies ω such terms provide only a small contribution to the Stark effect.[3] In the presence of a field so strong that the fine structure of the levels can be ignored, the problem can, generally speaking, be reduced to diagonalizing the tridiagonal quasienergy matrix of rank $n-|M|$, where n is the principal quantum number, with elements proportional to (1) (see Ref. 4). In particular cases of fields that do not lift the degeneracy completely the calculation of ΔE is simpler. For instance, in linearly polarized fields the level parity is conserved and the quasienergy matrix can be expanded in a direct sum of matrices that mix states of fixed parity.[3]

The situation becomes more complicated when we wish to describe the Stark effect for an atomic multiplet. The rank of the matrix that must be diagonalized depends on the frequency and strength of the field and cannot be specified in advance. It is determined by the number of levels for which the shift of a separate level is on the order of the width of the multiplet. For the excited states of most atoms in real fields all levels belonging to a multiplet are mixed by the field, while different multiplets can be assumed isolated.

Since in all cases except that of a static field the calculation of the scattering matrix is a mandatory stage in studying level shifts, probabilities of elastic and Raman scattering of light, two-photon ionization, and other two-photon processes, the problem has attracted great interest and many methods have been employed for its solution. Time-dependent perturbation theory in the Hartree–Fock method,[5] many-particle perturbation theory,[6] and the method of random phase with exchange[7] are the methods commonly used for studying ground states of complex atoms. However, many difficulties of a technical nature are encountered when one attempts to apply these methods to excited degenerate states. In this case the use of the Coulomb approximation proves to be justified from the physical standpoint. However, even in this approximation exact analytical expressions for the elements of the scattering matrix have only been obtained for the ground and first excited states of the hydrogen atom.[3,8] Direct summation of the terms generated by the Sturm expansion of the Green's function is possible only if the one-photon ionization channel is closed.[4] A method based on solving an inhomogeneous differential equation[9] uses the spherical symmetry of the S-state explicitly. But even in this case the considerable technical difficulties involved allow only the two lowest states to be studied. The use of the spectral expansion and the integral representation[10] of the Green's function leads to cumbersome numerical calculations, whose complexity rapidly increases with n.

If n is greater than 2, calculating C_{ik} for an arbitrary frequency ω of the external field requires additional approximations even for levels of the hydrogen atom and hydrogen-like ions. Some researchers have used the expansion of C_{ik} in inverse powers of ω (see Ref. 11), but this resulted in obtaining only a few terms of the expansion. The slow convergence of such an expansion and the nonanalytic nature of the behavior of the imaginary part of C_{ik} limit the applicability of this approximation.

The semiclassical expansion in inverse powers of the principal quantum number n (see Ref. 12) is also weakly convergent. Technical difficulties limit the accuracy of this expansion to terms of order $1/n^2$, and the fact that the level width is nonanalytic in the energy makes it impossible to obtain the imaginary part of ΔE in this approximation.

Methods based on summing a finite number of oscillator strengths calculated in the Kramer approximation[13] also have limited applicability. Since in this case the contribution of the continuous spectrum to α_{ik} is completely ignored, the approximation is inapplicable for frequencies ω exceeding the ionization potential of the level being excited. Here, as shown in Ref. 3, for ground atomic states the virtual transitions to the continuous spectrum determine up to 50–90 % of the polarizability value. For excited states the contribution varies dramatically, and for values of the orbital quantum number l greater than unity it can reach considerable values even when the ionization channel is closed. On the other hand, the approximation also becomes invalid for low frequencies ω. The thing is that in the Kramer approximation the oscillator strengths do not depend on l. This makes it impossible to allow for the mixing of levels that are degenerate in l, and this mixing cannot be ignored for low frequencies ω.

The situation requires developing new methods for obtaining exact analytical expressions for C_{ik}. The present paper is devoted to solving this problem in the nonrelativistic dipole approximation.

2. EXACT ANALYTICAL EXPRESSIONS FOR THE COMPONENTS OF THE TENSOR FOR SCATTERING OF LIGHT BY HYDROGEN LEVELS

The standard methods of the theory of angular momentum[14] make it possible to integrate in (1) over the angular variables and to find the dependence of C_{ik} on the magnetic quantum numbers of the initial, virtual, and final atomic states (M, m, and M_f, respectively). By applying the Wigner–Eckart theorem to (1) we can express C_{ik} in terms of Clebsch–Gordan coefficients and the reduced perturbation-theory composite matrix element:

$$C_{ik}=(-1)^{J-M+j-m}\begin{pmatrix} J & 1 & j \\ M & q_i & m \end{pmatrix}\left\{\begin{matrix} j & 1 & J_f \\ M & q_k & M_f \end{matrix}\right)$$
$$\times(2\|dG_{E_1+\omega_1}d+dG_{E_1-\omega_1}d\|1), \tag{3}$$

where $q_{i,k}=0,\pm1$. Calculation of (3) has been done for various angular-momentum coupling schemes and presents no difficulties (see, e.g., Refs. 3 and 15). In what follows we will need the explicit expression for C_{ik} in the LS coupling scheme in terms of the radial composite matrix element T and the Racah coefficients W:

$$C_{ik}=(-1)^{J-M-m-j+2S-L-L_1}\begin{pmatrix} J & 1 & j \\ M & q_i & m \end{pmatrix}$$
$$\times\begin{pmatrix} j & 1 & J_1 \\ M & q_k & M_1 \end{pmatrix}(2j+1)[(2J+1)(2J_1+1)]^{1/2}$$
$$\times W(LSlj;S1)W(ljL_1J_1;S1)L_{\max}L_{1\max}$$
$$\times[T(nL,\nu_+l,n_1L_1)+T(nL,\nu_+l,n_1L_1)]. \tag{4}$$

Here L, l, and L_f are, respectively, the orbital quantum numbers of the initial, virtual, and final states; $L_{\max}$ is the largest of the numbers L and l; S is the spin quantum number; the effective principal quantum number of the virtual level is defined as

$$\nu_{+,-}=[2(-E\pm\omega)]^{1/2}; \tag{5}$$

and the perturbation-theory radial composite matrix element is given by

$$T(n_fL_f,\nu l,nL)=\langle nL|r^3g_l(\nu,r,r_1)r_1^3|n_fL_f\rangle, \tag{6}$$

where $\langle nL|$ and $|n_fL_f\rangle$ are the radial parts of the wave functions of the initial and final atomic states, and $g_l(\nu,r,r_1)$ is the radial part of the Green's function. Calculating T constitutes the main difficulty in the problem.

The general algorithm for calculating two-photon dipole radial matrix elements for hydrogen states can be found in Ref. 15. Reasoning along similar lines, we arrive at an expression for the elements of the scattering matrix. We substitute the Sturm representation of the radial part of the Coulomb Green's function[3] in terms of the Laguerre polynomials L_k^{2l+1} in (6) (see Ref. 16),

$$g_l(\nu,r,r_1)=\frac{4}{\nu}\sum_{k=0}^{\infty}\frac{(rr_1)^l\exp(-r/\nu-r_1/\nu)}{k!(k+2l+1)!(k+l+1-\nu)}$$
$$\times L_k^{2l+1}(2r/\nu)L_k^{2l+1}(2r_1/\nu), \tag{7}$$

and the explicit expression for the bound-state wave function,

$$|nL\rangle=\frac{2^{L+1}}{n^{L+2}}[(n+L)!(n-L-1)!]^{1/2}\times\sum_{\alpha=0}^{n-L-1}\frac{(-2r/\nu)^{\alpha}}{(n-L-1-\alpha)!(2L+1+\alpha)!\alpha!}. \tag{8}$$

The matrix element (6) can be expressed in terms of the Laplace transform for L_k^{2L+1}. Using the expression for a Laplace transform in terms of the hypergeometric function,[16] we get

$$T(n_fL_f,\nu l,n_iL_i)=C_iC_f\sum_{k=0}^{\infty}\frac{(2l+k+1)!}{k!(k+l+1-\nu)}\times\left[\frac{\nu-n_i}{\nu+n_i}\,\frac{\nu-n_f}{\nu+n_f}\right]^k\times I(n_iL_i,\nu,lk)I(n_fL_f,\nu,lk). \tag{9}$$

Here we have introduced the following notation:

$$C_i=\frac{[(n_i+L_i)!(n_i-l_i-1)!]^{1/2}}{n_i^{L_i+2}\nu^{l+1/2}(2l+1)!}\left(\frac{\nu n_i}{\nu+n_i}\right)^{l+L_i+4}, \tag{10}$$

$$I(n_iL_i,\nu,lk)=\sum_{\alpha=0}^{n_i\text{-}L_i\text{-}1}\frac{(l+L_i+\alpha+3)!}{(n_i-l_i-\alpha-1)!(2L_i+\alpha+1)!\alpha!}\times\left(\frac{-2\nu}{\nu+n_i}\right)^{\alpha}F\left[-k,l-L_i-\alpha-2,2l+2,\frac{2n_i}{\nu-n_i}\right], \tag{11}$$

where F is the Gauss hypergeometric function, which in our case is a finite polynomial in the variable $2n_i/(\nu-n_i)$ whose degree is the smaller of the numbers k and $L_i-l+\alpha+2$. At the same time, F can be considered a polynomial of degree $L_i+\alpha+2-l$ in the parameter k.

This property makes it possible to analytically continue $T(n_fL_f,\nu l,n_iL_i)$ into the region of imaginary ν, where the expansion (9) diverges. Physically this region corresponds to an open channel of one-photon ionization of the initial atomic state. Let us write $I(n_iL_i,\nu,lk)$ in the form of a polynomial $Q_{j_i}^{x_i,y_i}(k)$ of the variable k with coefficients that depend on

$$x_i=\frac{n_i}{\nu-n_i},\quad y_i=\frac{\nu}{\nu+n_i}. \tag{12}$$

The degree j_i of this polynomial is determined by the leading term in the expansion of F with the maximum value $\alpha_{\max}^i=n_i-L_i-1$ and is $j_i=n_i+1-l$, with the result that the degree j of the product $P_{if}(k)$ of the polynomials Q_i and Q_f is determined by the total number of nodes of the wave functions of the initial and final states, n_{r_i} and n_{r_f}:

$$j=n_{r_i}+n_{r_f}+2-2l; \tag{13}$$

the $P_{if}(k)$ are polynomials of the variable k with coefficients that depend on the parameters x_i, y_i, x_f, and y_f. Expanding $P_{if}(k)$ in the polynomials

$$R_m=\prod_{\gamma=0}^{m}(k+2l+2+\gamma),\quad m=1,\ldots,j,\quad R_0=1, \tag{14}$$

we find that

$$P_{if}(k)=\sum_{m=0}^{j}b_m(x_i,y_i,x_f,y_f)R_m. \tag{15}$$

Since the system of polynomials (14) is not orthogonal, we cannot obtain closed expressions for the coefficients $b_m(x_i,y_i,x_f,y_f)$ of the expansion (15). Such an expansion can be done by standard methods: either by Horner's method via successive division of $P_{if}(k)$ by R_m $(m=0,\ldots,j)$, or by solving a system of $j+1$ linear equation obtained by comparing the coefficients of like powers of the variable k on the right- and left-hand sides of Eq. (15). The fact that the leading coefficients of all the polynomials in (14) are equal to unity makes it possible to construct a more economical recursion procedure for obtaining the $b_m(x_i,y_i,x_f,y_f)$ coefficients, a generalization of the moving band technique (see Ref. 17, Chap. 1) to the case of asymmetric functions. If we subtract from $P_{if}(k)$ the polynomial R_j with a weight equal to the leading coefficient of $P_{if}(k)$, the remainder is a polynomial of degree $j-1$. By recursively repeating this procedure with the remainders $j-1$ times we arrive at the system of coefficients b_m. This algorithm consists only of subtraction of polynomials with integral coefficients, is free from numerical errors, and can easily be programmed.

Now we allow for the fact that

$$(k+2l+1)!R_m=(k+2l+m+1)!, \tag{16}$$

and reduce (9) to

$$T(n_fL_f,\nu l,n_iL_i)=C_iC_f\sum_{m=0}^{j}b_{i,l,f,m}(x_i,y_i,x_f,y_f)\times\sum_{k=0}^{\infty}\frac{(2l+k+1+m)!}{k!(k+l+1-\nu)}\times\left[\frac{\nu-n_i}{\nu+n_i}\,\frac{\nu-n_f}{\nu+n_f}\right]^k. \tag{17}$$

Since the inner sum with respect to k can be expressed by a hypergeometric series,[16]

$$\sum_{k=0}^{\infty}\frac{(2l+k+1+m)!}{k!(k+l+1-\nu)}\left[\frac{\nu-n_i}{\nu+n_i}\,\frac{\nu-n_f}{\nu+n_f}\right]^k=\frac{(2l+1+m)!}{l+1-\nu}\,{}_2F_1\left[2l+m+1,l+1-\nu,l+2-\nu,\frac{\nu-n_i}{\nu+n_i}\,\frac{\nu-n_f}{\nu+n_f}\right], \tag{18}$$

we see that Eq. (17) constitutes a finite sum of contiguous Gauss hypergeometric functions:

$$T(n_fL_f,\nu l,n_iL_i)=\sum_{m=0}^{j}d_{i,l,f,m}(x_i,y_i,x_f,y_f)F_{if}(m). \tag{19}$$

Here the following notation has been introduced:

$$
\begin{aligned}
&d_{i,l,f,m}(x_i,y_i,x_f,y_f)\\
&\quad = C_iC_f\frac{(2l+1+m)!}{l+1-\nu}b_{i,l,f,m}(x_i,y_i,x_f,y_f),
\end{aligned}
\tag{20}
$$

$$
F_{if}(m)={}_2F_1\left[2l+m+1,l+1-\nu,l+2-\nu,\frac{\nu-n_i}{\nu+n_i}\,\frac{\nu-n_f}{\nu+n_f}\right].
$$

The algorithm not only makes it possible to describe the behavior of the scattering tensor in the region of analyticity but also gives the correct position of the poles: $\nu=l+n+1$ for $n=0,\ 1,\ \dots,\ \infty$.

The above algorithm for deriving analytical expressions for the components of the scattering tensor in terms of hypergeometric functions contains only the procedures of dividing polynomials with rational coefficients and collecting like terms in the variables x and y, which means that it can easily be programmed. The program implementing the algorithm contains only integer operations and is free from numerical errors.

By way of an example we give the explicit expressions for the diagonal (in the principal quantum numbers: $x_i=x_f=x$ and $y_i=y_f=y$) transitions from the ground and first excited states of the hydrogen atom:

$$
\begin{aligned}
&\frac{T(1s,\nu 1,1s)}{2^8 3^2}\\
&\quad =F(0)(1-4x+4x^2)+xF(1)(4-9x)+5x^2F(2),
\end{aligned}
$$

$$
\begin{aligned}
&\frac{T(2s,\nu 1,2s)}{2^{11}3^3 5}\\
&\quad =F(0)\Bigg[\frac{1/5-2y+5y^2}{8}+\frac{x(-1/5+3y-10y^2)}{2}+2x^2\\
&\qquad \times\left(\frac{1}{5}-6y+300y^2\right)+16x^3y(1-10y)+160x^4y^2\Bigg]\\
&\quad +xF(1)\Bigg[\frac{1/5-3y+10y^2}{8}+x(-9/10+26y-130y^2)\\
&\quad -56x^2y(1-10y)-800x^3y^2\Bigg]+2x^2F(2)\Bigg[\Bigg(\frac{1}{4}-7y\\
&\quad +35y^2\Bigg)+32xy(1-10y)+728x^2+y^2\Bigg]-64x^3yF(3)\\
&\quad \times(1-10y+48xy)+336x^4F(4)y^2,
\end{aligned}
$$

$$
\begin{aligned}
&\frac{T(2p,\nu 1,2s)}{2^{10}3^2 5}\\
&\quad =F(0)\Bigg[\frac{1-5y}{16}+\frac{x(-3+20y)}{4}+3x^2\\
&\qquad \times(1-10y)+4x^3(-1+20y)-80x^4y\Bigg]\\
&\qquad +xF(1)\Bigg[\left(\frac{3}{4}-5y\right)+13x\\
&\qquad \times\left(-\frac{1}{2}+5y\right)+14x^2(1-20y)+400x^3y\Bigg]+x^2F(2)\\
&\qquad \times\Bigg[\left(\frac{7}{2}-5y\right)+16x(-1+20y)-728x^2+y\Bigg]+x^3F(3)\\
&\qquad \times[6(1-20y)+576xy]+168x^4F(4)(-1+y),
\end{aligned}
\tag{21}
$$

$$
\begin{aligned}
&\frac{T(2p,\nu 0,2p)}{2^7 3}\\
&\quad =F(0)\left(\frac{1}{24}-x+10x^2-\frac{160}{3x^3}+160x^4-256x^5+\frac{512}{3x^6}\right)\\
&\quad +xF(1)(1-23x-208x^2-928x^3+2048x^4-1792x^5)\\
&\quad +16x^2F(2)\left(\frac{13}{16}-16x+116x^2-368x^3+432x^4\right)\\
&\quad +16x^3F(3)\left(\frac{19}{3}-98x+496x^2-824x^3\right)\\
&\quad +160x^4F(4)(3-32x+84x^2)+640x^5F(5)(2-11x)\\
&\quad +x^6F(6)\frac{4480}{3},
\end{aligned}
$$

$$
\begin{aligned}
\frac{T(2p,\nu 1,2p)}{2^9 3\times 5^2}&=F(0)\left(\frac{1}{16}-x+6x^2-16x^3+16x^4\right)\\
&\quad +xF(1)(1-13x+56x^2-80x^3)\\
&\quad +x^2F(2)\left(7-64x+\frac{738x^2}{5}\right)\\
&\quad +24x^3F(3)\left(1-\frac{24x}{5}\right)+\frac{168x^4F(4)}{5},
\end{aligned}
$$

$$
\begin{aligned}
\frac{T(2p,\nu 2,2p)}{2^{12}3^2 5^3}&=3F(0)\left(\frac{1}{8}-x+2x^2\right)+xF(1)(3-13x)\\
&\quad +7x^2F(2).
\end{aligned}
$$

In the formal static limit $\nu=n_{i,f}+l_{i,f}+1$ the initial series (17) can have only one first-order pole at $n_{r_{i,f}}=l-l_{i,f}$, which is possible only if $\Delta l=0,-1$. In all other cases (17) tends to a finite limit. Clearly, in the finite sums in (21) all the terms $x_{i,f}^j$ containing poles of orders higher than in (17) cancel out. For instance, for $T(1s,\nu 1,1s)$ there exists a finite static limit, $T(1s,\nu 1,1s)=27/4$, which yields a correct value for the static polarizability of the ground state of the hydrogen atom: $\alpha_{1s}=-9/2$.

The above algorithm makes it possible to express the tensor for scattering of light by any hydrogenlike atomic state in terms of a finite sum of contiguous Gauss hypergeometric functions. The three-term Kummer recursion relations allow any contiguous hypergeometric functions to be reduced to two functions. The fact that the second and third arguments in $F_{if}(m)$ differ by unity makes it possible to reduce (20) to a single hypergeometric function and several elementary functions. Let us prove this assertion.

Applying the Kummer relation

$$F(a+1,b,c,z)=\frac{F(a,b,c,z)+bF(a,b+1,c,z)}{a} \tag{22}$$

to $F_{if}(m)$ and allowing for the fact that $F(a,b,b,z)$ is an elementary function,

$$F(a,b,b,z)=(1-z)^{-a}, \tag{23}$$

we find that

$$F_{if}(m)=\frac{l+m+1+\nu}{2l+2+m}F(m-1)+\frac{l+1-\nu}{2l+2+m}Q^{2l+2+m}, \tag{24}$$

with

$$Q\equiv\frac{(\nu+n_i)(\nu+n_f)}{2\nu(n_i+n_f)}. \tag{25}$$

By applying (24) successively we can express $F_{if}(m)$ with an arbitrary m and hence the sum (20) in terms of $F_{if}(0)$ and an elementary function:

$$F_{if}(m)=\frac{l+m-1+\nu}{2l+2+m}P_mF_{if}(0)+(l+1-\nu)\sum_{j=0}^{m-1}\frac{P_jQ^{2l+2+m-j}}{2l+1+m-j}, \tag{26}$$

where

$$P_j=\prod_{k=0}^{j-1}\frac{l+j+\nu-k}{2l+1+j-k}. \tag{27}$$

To illustrate the above ideas, we give two examples of radial composite matrix elements expressed in terms of the single hypergeometric function $F(0)$:

$$T(1s,\nu 1,1s)=\frac{3}{2}\left(\frac{2\nu}{1+\nu}\right)^{10}\times\frac{(7\nu^2-18\nu+12)F(0)+(2-\nu)[(3\nu-8)Q^4+4Q^5]}{\nu^3(\nu-1)^2(2-\nu)},$$

$$T(2p,\nu 2,2p)=\frac{2^7 15^3}{\nu^5(\nu-2)^2}\left(\frac{2\nu}{2+\nu}\right)^{14}\times\left\{\frac{(49\nu^2-162\nu+324)F(0)}{24}+(3-\nu)\times\left[\frac{(5\nu-24)Q^6}{3}+\frac{(3\nu-32)Q^7}{4}\right]\right\}. \tag{28}$$

The relationships between contiguous hypergeometric functions can be used to prove a more general assertion: not only can an individual component of the tensor for scattering of light by an atomic shell be expressed in terms of a single hypergeometric function but so can the entire tensor. This property simplifies both the summation and averaging over the atomic multiplet components and the diagonalization of the quasienergy matrix.

According to the selection rules in the dipole approximation, the angular momentum l of the virtual state can be L_i, L_i+1, or L_i-1. Hence $F_{if}(0)$ can be expressed in terms of the angular quantum number of the initial state as follows:

$$\begin{aligned}F^{+}_{if}(0)&=F(2L_i+4,L_i+2-\nu,L_i+3-\nu,z),\\ F_{if}(0)&=F(2L_i+2,L_i+1-\nu,L_i+2-\nu,z),\\ F^{-}_{if}(0)&=F(2L_i,L_i-\nu,L_i+1-\nu,z).\end{aligned} \tag{29}$$

Let us express $F^{+}_{if}(0)$ in terms of $F_{if}(0)$. The reduction of the first parameter in the function $F^{+}_{if}(0)$ to the value $2L_i+2$ is done by applying the transformation (22) twice. A simultaneous decrease of the second and third parameters by unity can be achieved by a single Kummer transformation:

$$F(a,b,c+1,z)=\frac{c[F(a,b-1,c,z)-(1-z)F(a,b,c,z)]}{z(a-c)}. \tag{30}$$

In view of the fact that we have $c=b$ in (29), the function $F(a,b,c,z)$ is, in accordance with (23), an elementary function, with the result that

$$F^{+}_{if}(0)=\frac{L_i+2-\nu}{2L_i+3}\times\left\{Q^{2L_i+3}+\frac{L_i+1+\nu}{(2L_i+2)z}[Q^{2L_i+2}-F_{if}(0)]\right\}. \tag{31}$$

Applying the transformation

$$F(a-1,b,c,z)=\frac{bF(a,b+1,c,z)/z-(c-b-a)F(a,b,c,z)}{c-a} \tag{32}$$

to $F^{-}_{if}(0)$ twice and allowing for the fact that $F(a,b+1,c,z)$ is an elementary function, we can reduce the first parameter in $F^{-}_{if}(0)$ to $2L_i+2$. The Kummer transformation

$$F(a,b,c-1,z)=\frac{bF(a,b+1,c,z)/z+(c-b-1)F(a,b,c,z)}{c-a} \tag{33}$$

makes it possible to express $F^{-}_{if}(0)$ in terms of $F_{if}(0)$ as follows:

$$F^{-}_{if}(0)=\frac{Q^{2L_i}}{L_i+\nu}(2L_iQ-L_i+\nu)-\frac{z2L_i(2L_i+1)}{(L_i-\nu)(L_i+\nu+1)}F_{if}(0). \tag{34}$$

In particular, for P-states we have

$$F^{+}_{if}(0)=\frac{3-\nu}{5}\left[Q^5+\frac{2+\nu}{4z}(Q^4-F_{if}(0))\right],$$
$$F^{-}_{if}(0)=\frac{Q^2}{\nu+1}(2Q-1+\nu)-\frac{6z}{(1-\nu)(2+\nu)}F_{if}(0). \tag{35}$$

3. ANALYTICAL EXPRESSIONS FOR THE ELEMENTS OF THE TENSOR FOR SCATTERING OF LIGHT BY HYDROGENLIKE STATES OF COMPLEX ATOMS

When the one-electron approximation can be used to describe the behavior of the optical electron of a complex atom, we can employ the model-potential method to construct the wave function and the Green's function. The only method that leads to analytical results and is therefore most often used in multiphoton calculations is that of the Simons model potential[3]

$$V(\mathbf{r})=-\frac{Z}{r}+\sum_l \frac{B_l(E)P_l}{r^2}. \tag{36}$$

Here $B_l(E)$ is a parameter determined from the condition that the poles of the Green's function and the experimental spectrum of the atom coincide, and P_l is the projection operator on the subspace of angular-momentum eigenfunctions with a given l. Like as any one-electron method, it incorrectly describes the atomic-core region, where correlations have a strong effect. However, in calculating the probability of processes involving photons the dipole-interaction operator d reduces the relative contribution of this region to the integral (6).

The Simons potential is characterized by an incorrect asymptotic behavior in the opposite limit of large values of r, too. However, the wave functions and the Green's function in the region of large r are exponentially small. In addition, the nonlocal character of the potential in the angular variables makes it possible to offset this deficiency at least partially by an appropriate choice of the parameters $B_l(E)$ (see Ref. 3). As numerous calculations have shown, there are a broad range of radiation frequencies and level energies within which photoprocesses involving excited states of atoms of alkali and alkali-earth metals, noble gases, and other elements are described fairly well with such a potential. In addition, the potential also describes phototransitions from the ground states of atoms with one valence electron. This makes it possible to use the approximation with the Simons model potential to calculate the tensor for scattering of light by a complex atom.[3]

The wave function $|\nu_i\lambda\rangle$ and the Green's function $g_\lambda(\nu,r,r_1)$ of an optical electron in this potential are[3]

$$|\nu_i\lambda\rangle=\frac{2^{\lambda+1}}{\nu^{\lambda+2}}[(\nu_i+\lambda)!n_r!]^{1/2}\times\sum_{\alpha=0}^{n_r}\left(\frac{-2r}{\nu_i}\right)^{\alpha}\frac{1}{n_r!(2\lambda+1+\alpha)!\alpha!} \tag{37}$$

$$g_\lambda(\nu,r,r_1)=\frac{4}{\nu}\sum_{k=0}^{\infty}\frac{(rr_1)^{\lambda}\exp(-r/\nu-r_1/\nu)}{k!(k+2\lambda+1)!(k+\lambda+1-\nu)}\times L_k^{2\lambda+1}(2r/\nu)L_k^{2\lambda+1}(2r_1/\nu),$$

when n_r is the radial quantum number. The effective principal quantum number of a bound state, $\nu_{i,f}$, is defined as $\nu_{i,f}=(-2E_{i,f})^{-1/2}$, and the parameter λ is a function of the energy E of the optical electron and the angular quantum number l. The parameter is selected in such a way that the experimentally established values of the bound-state energies coincide with the poles of the Green's function. The necessary condition for the Simons model-potential method to be applicable in calculating the photoprocess probabilities is the smoothness of the function $\lambda(E)$:

$$\left|\frac{\partial\lambda}{\partial E}\right|\ll 1.$$

Substituting (37) into (6), we arrive at the generalized expressions (9)–(11) for the case of complex atoms:

$$T(\nu_f\lambda_f,\nu\lambda,\nu_i\lambda_i)=C_iC_f\sum_{k=0}^{\infty}\frac{F(2\lambda+k+2)!}{k!(k+\lambda+1-\nu)}\times\left[\frac{\nu-\nu_i}{\nu+\nu_i}\,\frac{\nu-\nu_f}{\nu+\nu_f}\right]^{k}I(\nu_i\lambda_i,\nu,\lambda k)I(\nu_f\lambda_f,\nu,\lambda k), \tag{38}$$

where

$$C_i=\frac{[\Gamma(n_i+L_i+1)n_{ri}!]^{1/2}}{n_i^{\lambda_i+2}\nu^{\lambda_i+1/2}\Gamma(2\lambda_i+2)}\left(\frac{\nu n_i}{\nu+n_i}\right)^{\lambda+\lambda_i+4},$$

$$I(n_i\lambda_i,\nu,\lambda k)=\sum_{\alpha=0}^{n_{ri}}\frac{\Gamma(\lambda+\lambda_i+\alpha+4)}{\Gamma(n_{ri}-\alpha)\Gamma(2\lambda_i+\alpha+2)!\alpha!}\times\left(\frac{-2\nu}{\nu+n_i}\right)^{\alpha}F\left[-k,\lambda-\lambda_i-\alpha-2,2\lambda+2,\frac{2n_i}{\nu-n_i}\right]. \tag{39}$$

A particular case of (37) and (38) for transitions that are diagonal in the principal quantum number was obtained by Manakov and Ovsiannikov.[10]

The expansion (37) converges only for real ν, which from the standpoint of physics means that the one-photon ionization channel is closed. It is impossible to directly apply the above algorithm for analytic continuation of $T(n_fL_f,\nu l,n_iL_i)$ to the imaginary ν region, since $F[-k,\lambda-\lambda_i-\alpha-2,\,2\lambda+2,\,2n_i/(\nu-n_i)]$ is not a finite polynomial of variable k for noninteger values of the parameter $\lambda_i-\lambda$. For the analytic continuation algorithm to become applicable we must reduce $I(n_i\lambda_i,\nu,\lambda k)$ with noninteger λ and λ_i to a sum of polynomials $Q_{j_i}^{x_i,y_i}(k)$. To this end we expand $r^{\lambda_i+2-\lambda}$ in integer powers of the variable r:

$$r^{\lambda_i+2-\lambda}=\sum_{\delta=0}^{\delta_{\max}}h_\delta(\lambda-\lambda_i)r^{\delta}. \tag{40}$$

A function can be expanded in power series in orthogonal Jacobi polynomials in an infinite number of ways, with the various series differing by the weight function with respect to which orthogonality is established. By collecting terms with equal powers of r we arrive at the power expansions (40). For $\delta_{\max}\to\infty$ all these expansions are exact and equivalent. But if the purpose of the expansion is to attain an accuracy that is fixed in a finite interval with a minimum number $\delta_{\max}$ of expansion terms, the convergence varies. Since the error in calculating the cross section of a photoprocess involving a complex atom is in any case bounded below by the one-electron approximation of the model potential, there is no need to strive for absolute accuracy in the expansion. It is precisely this situation that we are studying here.

The expansion basis must be selected so as to ensure the most economical (for the given accuracy) description of the integrand in (6) in the interval $0>r\geqslant\nu$ of variation of variable r, the interval that provides the main contribution to the integral. Generally speaking, the local Taylor and Maclaurin expansions are extremely inefficient here, since they use only the information about the behavior of a function at a single point. Much more effective is the procedure of expanding in orthogonal polynomials, and from the overall point of view the best is the expansion in ultraspherical polynomials.[17]

The contributions of the different parts of the domain of integration in (6) differ for different values of the parameters of the problem. This fact can be taken into account by an appropriate choice of the weight factors in the orthogonal expansion and, hence, of the form of the ultraspherical polynomial.[17] The basis consisting of Chebyshev polynomials $T_k(x)$ ensures a uniform estimate for the relative maximum deviation $\Delta_{\max}$ of the approximation polynomial from the expanded function in the entire interval of values of r. It is this property of Chebyshev polynomials that determines their advantage in approximating a function over a broad interval of variation of the parameters of the problem. Such a choice of the basis makes it possible to determine the form of the power expansion irrespective of the type of atom and the field parameters. Here the relative error Δ in calculating the integral cannot exceed $\Delta_{\max}$. In the interval $-1<r<1$ the polynomials $T_k(x)$ are determined by the recurrence relations

$$T_{k+1}(x)=2xT_k(x)-T_{k-1}(x),$$

$$T_0(x)=1,\quad T_1(x)=x. \tag{41}$$

The coefficients a_k of the expansion of a function $f(x)$ in Chebyshev polynomials are given by the following expression:

$$a_k=\int_{-1}^{1} f(x)T_k(x)(1-x^2)^{1/2}dx. \tag{42}$$

Substituting (40) into (6) we get

$$I(n_i\lambda_i,\nu,\lambda k)$$

$$=\sum_{\delta=0}^{\delta_i} h_\delta(\lambda-\lambda_i)\sum_{\alpha=0}^{n_{ri}}\frac{\Gamma(\lambda+\lambda_i+\alpha+4)}{\Gamma(n_{ri}-\alpha)\Gamma(2\lambda_i+\alpha+2)\alpha!}$$

$$\times\left(\frac{-2\nu}{\nu+n_i}\right)^{\alpha}F\left[-k,-\delta-\alpha,2\lambda+2,\frac{2n_i}{\nu-n_i}\right]. \tag{43}$$

Since now the second parameter in F is a negative integer, we have an expression similar to (15):

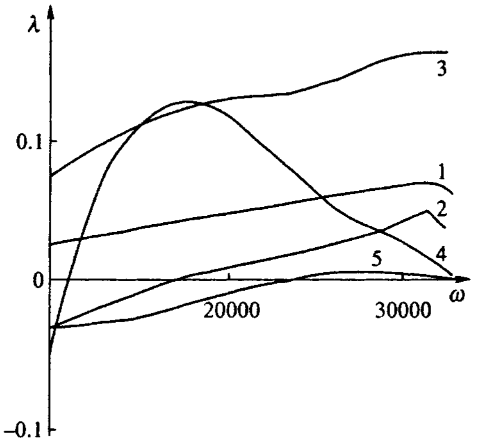


FIG. 1. The energy dependence of the parameters λ of the model potential for the rubidium atom. The curves *1–5* correspond to s-, $p_{1/2}$-, $p_{3/2}$-, d-, and f-states. Energies are measured in units of cm^{-1}. The given values of λ_l are normalized by the condition $\lambda_l=\lambda-a_l$, where $a_s=0.8$, $a_{p_{12}}=1.3$, $a_{p_{3/2}}=1.2$, $a_d=1.68$, and $a_f=2.98$.

$$P_{if}(k)=\sum_{m=0}^{j} B_m(x_i,y_i,x_f,y_f)R_m, \tag{44}$$

where $B_m(x_i,y_i,x_f,y_f)$ is simply a linear combination of the quantities $b_m(x_i,y_i,x_f,y_f)$ determined earlier, with weights defined by the coefficients of the approximation (40):

$$B_m(x_i,y_i,x_f,y_f)=\sum_{\delta=0}^{\delta_i} h_\delta(\lambda-\lambda_i)$$

$$\times\sum_{\delta_1=0}^{\delta_f} h_{\delta_1}(\lambda-\lambda_f)b_m(x_i,y_i,x_f,y_f). \tag{45}$$

Thus, the problem of calculating the elements of the tensor for scattering of light by a complex atom amounts to determining the energy dependence of the parameters λ, λ_i, and λ_f and calculating the coefficients $h_\delta(\lambda-\lambda_{i,f}+2)$. The first part of this problem for each atom can easily be solved by approximating the values of λ calculated from the experimental spectrum.[19] By way of an example, Fig. 1 depicts the energy dependence of the λ parameters for the rubidium atom, which is used below in calculations.

The solution of the second part of the problem is independent of the features of a specific atom. Table I lists the values of the coefficients $h_\delta(\rho)$ of the lowest-order terms in the expansion and Table II the values of the coefficients of the highest-order terms, which ensure an approximation accuracy $\Delta_{\max}=10^{-4}$ for $0<r<40$. These data make it possible to calculate the probabilities of photoprocesses in which ν, ν_{if}, $\nu\leqslant 40$ hold . The interval of variation of the parameter ρ describes all the transitions between the levels of alkali and alkali-earth metals and noble gases. Here the approximations that vanish to the specified accuracy are discarded, and c/n stands for $c\times 10^{-n}$. Table I and II show that, with the exception of the trivial cases of the neighborhoods of integer values of ρ, where one term with $\delta=\rho$ in the expansion (40) is sufficient, the expansion rapidly converges

TABLE I.

ρ	δ 0	1	2	3	4	5	6	7	8
0.6	1	4	−7	8.7	−8	4.541	−1.788	0.506	−0.1059
0.7	−2	3.5	−7	6.5	−4.9	2.83	−1.09	0.309	−0.06482
0.8	−1	2.1	−3.5	2.8	−2.7	1.5	−0.59	0.167	−0.03455
0.9		2.6	−1.227	1.339	−0.99	0.594	−0.233	0.06621	−0.01385
1.0		1							
1.1	0.1	0.5	0.95	−0.733	0.67	−0.3712	0.145	−0.04077	0.008526
1.2	0.5	0.2	1.4	−1.464	1.07	−0.595	0.225	−0.0629	0.01308
1.3	−0.4	1.2	1.2	−1.272	1.15	−0.64	0.253	−0.0705	0.0147
1.4	0.05	0.31	1.08	−0.6713	0.344	−0.1241	0.3142	−0.00569	0.7485/4
1.5	0.1	0.3	0.9655	−0.3927	0.14148	−0.03759	0.006787	−8.585/4	0.7659/4
1.6	−0.027	0.274	0.903	−0.2188	0.05504	−0.009698	0.001162	−0.9382/4	0.502/5
1.7	0.033	0.215	0.9106	−0.1463	0.02905	−0.004056	0.3778/3	−0.2293/4	0.8688/6
1.8	0.041	0.16	0.9234	−0.8634	0.01314	−0.001406	0.9748/4	−0.4156/5	0.9878/7
1.9	−0.038	0.12	0.9851	−0.02804	0.002309	−0.1252/3	0.3728/5	−0.4612/7	
2.0			1						
2.1	−0.0089	−0.11	1.056	0.03913	−0.00291	0.1528/3	0.445/5	0.548/7	
2.2	0.085	−0.18	1.089	0.09076	−0.00634	0.3261/3	0.947/5	0.1153/7	
2.3	0.074	−0.27	1.092	0.1561	−0.0101	0.5089/3	0.146/4	0.1772/6	
2.4	0.438	−0.36	1.063	0.2362	−0.01392	0.6851/3	0.195/4	0.2347/6	
2.5	0.01	−0.29	0.993	0.3315	−0.01736	0.8334/3	0.234/4	0.2806/6	
2.6	−0.02	−0.4	0.883	0.4421	−0.01985	0.9261/3	0.257/4	0.3061/6	
2.7	1.73	−0.5	0.746	0.5667	−0.02056	0.9294/3	0.255/4	0.3013/6	
2.8	0.95	−0.5	0.6835	0.6572	−0.01173	0.3128/3	0.391/5		
2.9	0.39	−0.1	0.3815	0.8213	−0.008	0.2049/3	0.252/5		
3.0				1					
3.1	−0.096	0.8	−0.708	1.243	0.008875	−0.9342/4			
3.2	−0.4	0.8	−0.936	1.1375	0.03685	−0.8004/3	0.9357/5		
3.3	−1	1.1	−1.483	1.547	0.07107	−0.001431	0.1638/4		
3.4	1	1.8	−1.893	1.687	0.1205	−0.002209	0.2471/4		
3.5	−10.1	5	−2.451	1.766	0.1894	−0.003086	0.3366/4		
3.6	−11	3	−2.821	1.763	0.2825	−0.003948	0.4186/4		

for $\nu \approx 10^0$, and for an arbitrary ν it also rapidly converges for $\rho > 1.5$. In this case, even when ν is large, the required accuracy is ensured by a small number of terms with $\delta \leqslant 5$. For δ small and highly excited states, the number $\delta_{\max}$ of terms in (40) that must be taken into account increases, and one is forced to allow for the leading terms in the expansion (Table II), making the use of the method in calculating the tensor of scattering of light much more complicated.

In this case the basis of ultraspherical polynomials and, in particular, of Chebyshev polynomials, which yields an error estimate that is uniform over the entire approximation interval, does not provide a correct description of the problem. For different parameter intervals the range of values of r providing the greatest contribution to (6) is different. Here the basis of Jacobi polynomials with asymmetric weight factors emphasizing such a range of values of r has proved more economical than that of ultraspherical polynomials. Naturally, with this choice of basis, the expansion (40) becomes more economical but loses in generality. There exists an important limit of a strong asymmetry of the integrand for

TABLE II.

ρ	δ 9	10	11	12	13	14	15	16	17
0.6	0.01668	−0.2/2	0.183/3	−0.128/4	0.677/6	−0.267/7	0.7567/9	−0.15/10	0.172/12
0.7	0.01016	−0.12/2	0.111/3	−0.778/5	0.412/6	−0.162/7	0.46/9	−0.89/11	0.104/12
0.8	0.00545	−0.65/3	0.6/4	−0.417/5	0.221/6	−0.869/8	0.2465/9	−0.48/11	0.56/13
0.9	0.00218	−0.26/3	0.238/4	−0.166/5	0.88/7	−0.346/8	0.982/10	−0.19/11	0.223/13
1.0									
1.1	−0.00134	0.16/3	−0.146/4	0.1017/5	−0.54/7	0.2121/8	−0.6/11	0.116/11	−0.14/13
1.2	−0.00205	0.245/3	−0.224/4	0.156/5	−0.825/7	0.3247/8	−0.92/10	0.178/11	−0.21/13
1.3	−0.0023	0.275/3	−0.251/4	0.175/5	−0.235/7	0.3648/8	−0.103/9	0.2/11	−0.23/13
1.4	−0.748/3	0.505/5	−0.256/6	0.913/5	−0.217/9	0.307/11	−0.2/13		
1.5	−0.479/5	0.205/6	−0.574/8	0.943/10	−0.7/12				
1.6	−0.17/6	0.332/8	−0.28/10						
1.7	−0.19/7	0.173/9							
1.8	−0.1/8								

which the Chebyshev-polynomial basis can be used for large values of ν and small values of ρ as well.

As shown in Ref. 10, the range of values of r contributing to the integral (6) narrows as the energy of bound atomic states and the frequency of the electromagnetic wave grow. Since the model potential provides an incorrect description of the behavior of wave functions for $r \leqslant r_c$, where r_c is the atomic-core radius, this fact imposes restrictions on the applicability of the Simons model potential (and hence of the method) to calculations of the tensor for scattering of electromagnetic radiation by Rydberg states. For highly excited states the range of applicability of the dipole approximation narrows, too. All this hinders theoretical studies of the interaction of light with such levels. At the same time, within the range of applicability of the method for states with principal quantum number $n \cong 10^1$, the number of terms in the expansion (40) required to ensure a given accuracy decreases with increasing $n_{i,f}$ and ω and with narrowing approximation range. This makes it possible to use the Chebyshev-polynomial basis in calculations of the tensor of scattering of light by complex atoms in these cases even when ρ is small.

Here is an example that illustrates how the method can be used to calculate the scattering tensor. Let us compute the shift of the ground $5S$-state of the rubidium atom at the frequency $\omega = 9434\ \mathrm{cm}^{-1}$ of a neodymium laser, which has recently been studied both theoretically and experimentally.[18] Since the detuning to the nearest level exceeds 0.1 a.u., the isolated-level approximation is valid for real laser fields, with the result that, in accordance with (4), the polarizability α_{5S} can be expressed in terms of the radial matrix elements (6) as follows:

$$\alpha_{5S} = -\frac{T(5,0,\nu_+ 1,5,0) + T(5,0,\nu_- 1,5,0)}{12}. \tag{46}$$

The energies of the virtual states are $E_1^{+,-} = \pm \omega_n$. Extrapolation of the data of Fig. 1 yields $\lambda_0 = 0.8236$, $\lambda_1^+ = 1.26124$, and $\lambda_1^- = 0.945$, in view of which $\rho^+ = 1.5624$ and $\rho^- = 1.8786$. Interpolation by the data listed in Table I yields

$$\begin{aligned} r^{1.5624} = & -0.22 + 0.3128r + 0.914r^2 - 0.27r^3 + 0.09r^4 \\ & -0.024r^5 + 0.0061r^6 - 0.00012r^7 \\ & +0.00084r^8 - 0.000045r^9, \end{aligned}$$

$$\begin{aligned} r^{1.8786} = & -0.03235 + 0.1335r + 0.9196r^2 - 0.038r^3 \\ & +0.0088r^4 - 0.00015r^5 - 0.0003r^6 \\ & +1.6\times 10^{-4}r^7 - 2.9\times 10^{-5}r^8 + 4.8\times 10^{-6}r^9. \end{aligned} \tag{47}$$

Since ν of the initial and virtual states (and hence the range of values of r contributing to the integral (6)) is of order 2, there is no need to include in (47) the highest-order terms in δ from Tables I and II. The discarded terms contribute a relative correction less than 10^{-4}. Plugging (47) into (45) and allowing for (20), we obtain $\alpha_{5S} = 707$ a.u., which is in good agreement with other theoretical ($\alpha_{5S} = 692$ a.u.; see Ref. 3) and experimental ($\alpha_{5S} = 769 \pm 61$ a.u.; see Ref. 18) values.

[1] A. I. Akhiezer and V. B. Berestetskiĭ, *Quantum Electrodynamics*, Wiley, New York (1974).

[2] M. A. Preobrazhenskiĭ, Opt. Spektrosk. **77**, 559 (1994) [Opt. Spectrosc. **77**, 494 (1994)].

[3] L. P. Rapoport, B. A. Zon, and N. L. Manakov, *The Theory of Multiphoton Processes in Atoms* [in Russian], Atomizdat, Moscow (1978).

[4] V. I. Ritus, Zh. Éksp. Teor. Fiz. **51**, 1544 (1966) [Sov. Phys. JETP **24**, 1041 (1967)].

[5] I. Epstein, J. Chem. Phys. **53**, 1881 (1970).

[6] H. P. Kelly, Phys. Rev. **182**, 84 (1969).

[7] M. Ya. Amus'ya, N. A. Cherpakov, and S. G. Shapiro, Zh. Éksp. Teor. Fiz. **63**, 889 (1972) [Sov. Phys. JETP **36**, 468 (1973)].

[8] M. Gavrila, Phys. Rev. **163**, 147 (1967).

[9] M. Marinescu, H. R. Sadeghpoure, and A. Dalgarno, Phys. Rev. A **49**, 5103 (1994); V. L. Yakhontov and K. Jungmann, in: *Europhys. Conf. Abstracts* (1966), p. 74.

[10] N. L. Manakov and V. D. Ovsiannikov, J. Phys. B **10**, 569 (1976).

[11] N. L. Manakov, V. A. Sviridov, and A. G. Faĭnshteĭn, Zh. Éksp. Teor. Fiz. **95**, 790 (1989) [Sov. Phys. JETP **68**, 451 (1989)].

[12] V. M. Vaĭnberg, V. D. Mur, V. S. Popov, and A. V. Sergeev, JETP Lett. **44**, 9 (1986).

[13] N. B. Delone and V. P. Kraĭnov, Zh. Éksp. Teor. Fiz. **83**, 2021 (1982) [Sov. Phys. JETP **56**, 1170 (1982)]; I. L. Beĭgman, Zh. Éksp. Teor. Fiz. **100**, 125 (1991) [Sov. Phys. JETP **73**, 68 (1991)]; I. L. Beigman, L. A. Bureyeva, and R. H. Pratt, Phys. Rev. A **49**, 5883 (1994).

[14] I. I. Sobelman, *Introduction to the Theory of Atomic Spectra*, Pergamon Press, Oxford (1973).

[15] M. A. Preobrazhenskii, Laser. Phys. **3**, 688 (1993).

[16] A. Erdélyi, *Higher Transcendental Functions* (Bateman Project), 3 vols., McGraw-Hill, New York (1953–1955).

[17] C. Lanczos, *Applied Analysis*, Prentice-Hall, Englewood Cliffs, N.J. (1956).

[18] K. D. Bonin and M. A. Kadan-Kelly, Phys. Rev. A **47**, 999 (1993).

[19] C. E. Moore, in *National Bureau of Standards, Circular No. 488*, Washington (1950).

Translated by Eugene Yankovsky

Quantum kinetic Boltzmann equation taking into account the resonant exchange of excitations

T. L. Andreeva and P. L. Rubin

P. N. Lebedev Physics Institute, Russian Academy of Sciences, 117924 Moscow, Russia

(Submitted 8 July 1996)
Zh. Éksp. Teor. Fiz. **111**, 831–837 (March 1997)

A derivation of the quantum Boltzmann equation is given for identical particles with internal degrees of freedom. It is shown that the off-diagonal (with respect to the internal degrees of freedom) term of the equation contains an energy pole term, which is not present in the most commonly used kinetic equation, known as the Waldmann-Snider equation. The physical conditions underlying the occurrence of the pole term in the quantum kinetic equation are analyzed. © *1997 American Institute of Physics.* [S1063-7761(97)00503-9]

1. INTRODUCTION

A necessary adjunct to line-shape calculations for gases in laser spectroscopy problems is a kinetic Boltzmann equation with the internal degrees of freedom of the particles taken into account (i.e., an equation for the density matrix). The spectral line shape is determined by the off-diagonal—with respect to internal degrees of freedom—element of the density matrix.[1] Typical of such degrees of freedom are the electronic states of an atom, both nondegenerate and degenerate with respect to the projections of the angular momentum.

Of special interest are gases whose magnetic moment is determined entirely by spin (S state) and in which the magnetic polarization, like the line shape in optics, is determined by the off-diagonal—with respect to the spin variables—element of the density matrix. So-called spin-polarized gases have come under intense scrutiny in recent years as various techniques afford increasing possibilities for the polarization of paramagnetic gases (see, e.g., Refs. 2 and 3). At issue is the fact that spin polarization significantly influences the macroscopic properties of quantum gases and, under definite conditions, even admits the existence of weakly damped spin waves in them.[4,5]

The primary tool for the theoretical investigation of spin-polarized gases is currently the Waldmann–Snider kinetic equation[6,7] generalized to the case of identical particles,[8] which has no pole term. The presence of the pole term in the Boltzmann equation has been mentioned in several papers,[9–11] and it leads to a number of interesting physical effects. In particular, the optical region of the spectrum acquires an additional line shift with an anomalous dependence on the temperature of the gas,[11] and spin waves occur in spin-polarized gases.[2–8,10]

Meyerovich, Stepaniants, and Laloë have recently[2] augmented the Waldmann–Snider equation with a pole term, but it is obtained as a third-order term with respect to the density of the gas, and they have stated that such a term does not occur in second order with respect to the density. In the present study we show that a pole term in the kinetic equation does in fact occur in second order with respect to the gas density, as is usual for the Boltzmann equation. We analyze the physical conditions underlying the origin of the pole term. We also show that the pole term in the kinetic equation appears when the scattering amplitudes depend on the internal state of the particles. In particular, the presence of two collision processes with different scattering amplitudes is sufficient for the onset of a pole term.

2. DERIVATION OF AN EQUATION FOR THE DENSITY MATRIX

The basic plan of derivation is the same as that used by Snider[6] except that our treatment is applicable both to degenerate (as in Snider's work) and to nondegenerate systems. We assume for simplicity that the particles in question are bosons (it can easily be shown that the final result does not depend on this choice). In coordinate representation the density matrix has the form

$$\rho_{\alpha\alpha'}(x,x',t)=\langle\psi^{+}_{\alpha'}(x',t)\psi_{\alpha}(x,t)\rangle. \tag{1}$$

Here $\psi_\alpha(x,t)$ is the wave function of the given particles in the second-quantization representation (α enumerates the internal states), and the angle brackets signify quantum-statistical averaging.

The second-quantized Hamiltonian of the system has the usual form

$$\hat{H}=\int\psi^{+}_{\mu}(x)H^{(0)}\psi_{\nu}(x)dx+\frac{1}{2}\int\psi^{+}_{\mu}(y)\psi^{+}_{\nu}(x)\times U_{\mu\nu\sigma\theta}(x-y)\psi_{\sigma}(x)\psi_{\theta}(y)dx\,dy. \tag{2}$$

Here $H^{(0)}$ is the single-particle Hamiltonian:

$$H^{(0)}\psi_{\nu}(x)\left(-\frac{\nabla_x^2}{2m}+E_{\nu}\right)\psi_{\nu}(x),$$

where m is the particle mass, ∇^2 is the Laplace operator, E_ν is the energy of the internal state, and U is the particle interaction potential taking their internal state into account. For simplicity, we set $\hbar=1$. The following properties of the interaction potential are deduced from the Hermitian property of $\hat{H}$ and the condition of symmetry under permutations of identical particles:

$$U^{*}_{\mu\nu\sigma\theta}(x)=U_{\theta\sigma\nu\mu}(x),\quad U_{\mu\nu\sigma\theta}(x)=U_{\nu\mu\theta\sigma}(-x). \tag{3}$$

The equations for the single-particle density matrix (1) and for the two-particle density matrix

$$\rho^{(2)}_{\alpha\beta\alpha'\beta'}(x,y,x',y')=\langle\psi^+_\alpha(x)\psi^+_\beta(y)\psi_{\alpha'}(x')\psi_{\beta'}(y')\rangle$$

have the standard form[9,10]

$$i\,\frac{\partial\rho_{\alpha\alpha'}(x,x',t)}{\partial t}=\left(-\frac{\nabla^2_x}{2m}+E_\alpha+\frac{\nabla^2_{x'}}{2m}-E_{\alpha'}\right)\rho_{\alpha\alpha'}\times(x,x',t)+I_{\alpha\alpha'}(x,x',t)-I^*_{\alpha'\alpha}(x',x,t),\tag{4}$$

$$\begin{aligned}i\,\frac{\partial\rho^{(2)}_{\tau_1\tau_2\tau_3\tau_4}(x_1,x_2,x_3,x_4)}{\partial t}&=\left[\left(-\frac{\nabla^2_{x_3}}{2m}-\frac{\nabla^2_{x_1}}{2m}\right)+E_{\tau_3}+E_{\tau_4}-\left(-\frac{\nabla^2_{x_1}}{2m}-\frac{\nabla^2_{x_2}}{2m}+E_{\tau_1}\right.\right.\\&\left.\left.+E_{\tau_2}\right)\right]\rho^{(2)}_{\tau_1\tau_2\tau_3\tau_4}(x_1,x_2,x_3,x_4)+U_{\tau_3\tau_4\gamma\beta}(x_3\\&-x_4)\rho^{(2)}_{\tau_1\tau_2\gamma\beta}(x_1,x_2,x_3,x_4)-U_{\beta\gamma\tau_1\tau_2}(x_2-x_1)\\&\times\rho^{(2)}_{\beta\gamma\tau_1\tau_2}(x_1,x_2,x_3,x_4),\end{aligned}\tag{5}$$

where $I_{\alpha\alpha'}$ is the collision integral:

$$I_{\alpha\alpha'}(x,x',t)=\int U_{\lambda\alpha\sigma\theta}(x-y)\rho^{(2)}_{\lambda\alpha\theta\sigma}(x',y,y,x)\,dy.\tag{6}$$

Summation on repeated indices is understood. Here, as usual, ternary collisions are disregarded.

To solve Eq. (6), it is necessary to specify the asymptotic form of the solution in the region where the particles essentially do not interact. The formula for this case is well known:[9]

$$\rho^{(2)}_{\tau_1\tau_2\tau_3\tau_4}(x_1,x_2,x_3,x_4)=(1+\hat{P})\rho_{\tau_3\tau_1}(x_3,x_1)\rho_{\tau_4\tau_2}(x_4,x_2).$$

Here $\hat{P}$ is the $1\leftrightarrow2$ or (equivalently) $3\leftrightarrow4$ particle permutation operator.

Assuming, as is customary in the Boltzmann approximation, that collisions are local and transforming from the density matrix in coordinate representation to the Wigner function,[12] we readily obtain an asymptotic expression for the two-particle function:

$$\begin{aligned}\rho_{\tau_3\tau_1}(x_3,x_1)\rho_{\tau_4\tau_2}(x_4,x_2)=\int dp'dp''f_{\tau_3\tau_1}(p')f_{\tau_4\tau_2}(p'')\\\times\exp i\left[(p'+p'')(R-R')\right.\\\left.+\frac{(p'-p'')(r-r')}{2}\right].\end{aligned}\tag{7}$$

Along with the initial variables, the right-hand side of this expression uses the change of variables:

$$r=x_3-x_4,\quad R=(x_3+x_4)/2,\quad r'=x_1-x_2,$$

$$R'=(x_1+x_2)/2,$$

which provides a natural separation of the motion of the center of inertia of the colliding particles from their relative motion.

The next step in the derivation of the equation in Snider's paper[6] has been done compactly and in an abstract operator form, a device that probably accounts for certain inaccuracies in the final result. We shall cover this part of the derivation in greater detail. We first invoke the obvious identity:

$$f_{\tau_3\tau_1}(p')f_{\tau_4\tau_2}(p'')=f_{\tau'_3\tau'_1}(p')f_{\tau'_4\tau'_2}(p'')\delta_{\tau_1\tau'_1}\delta_{\tau_2\tau'_2}\delta_{\tau_3\tau'_3}\delta_{\tau_4\tau'_4}.$$

Here the function

$$\varphi\left(\tau'_3,\tau'_4,\ \frac{p'-p''}{2};\tau_3,\tau_4,r\right)=\delta_{\tau_3\tau'_3}\delta_{\tau_4\tau'_4}\exp\left(i\ \frac{p'-p''}{2}\ r\right)$$

represents the asymptotic (at infinity) value of the wave function of the two-particle collision problem: $(p'-p'')/2$ is the relative momentum (the reduced mass is $m/2$), τ'_3 and τ'_4 are quantum numbers, and τ_3 and τ_4 are the corresponding internal state variables. We illustrate this operation in the example of a particle with spin, for which the wave function of the internal state represents spinors $\delta_{(1/2)\tau}$ and $\delta_{-(1/2)\tau}$, where $\tau=1/2$ or $-1/2$.

The groups of variables (r,R,τ_3,τ_4) and (r',R,τ_1,τ_2) are separated in Eq. (4), the second group describing time-reversed motion. We can assume that the time-inversion operation is tantamount to complex conjugation of the wave function, because the pair of them comprises a boson even in the case of fermions.

An exact solution is now easily obtained for the equation for the two-particle density matrix. It suffices in Eq. (7) to replace the factors representing asymptotic wave functions of the scattering problem according to (4) by the corresponding exact wave functions ψ:

$$\varphi\left(\tau'_3,\tau'_4,\ \frac{p'-p''}{2};\tau_3,\tau_4,r\right)\to\psi\left(\tau'_3,\tau'_4,\ \frac{p'-p''}{2};\tau_3,\tau_4,r\right).$$

As a result, the expression for the two-particle function acquires the form

$$\begin{aligned}\rho^{(2)}_{\tau_1\tau_2\tau_3\tau_4}(x_1,x_2,x_3,x_4)\\=2\int dp'dp''\exp[i(p'+p'')(R\\-R')]f_{\tau'_3\tau'_1}(p')f_{\tau'_4\tau'_2}(p'')\exp(i\mathscr{E}_{\tau'_1\tau'_2\tau'_3\tau'_4}t)\psi\\\times\left(\tau'_3,\tau'_4,\ \frac{p'-p''}{2};\tau_3,\tau_4,r\right)\psi^*\\\times\left(\tau'_1,\tau'_2,\ \frac{p'-p''}{2};\tau_1,\tau_2,r'\right).\end{aligned}\tag{8}$$

Here

$$\mathscr{E}_{\tau'_1\tau'_2\tau'_3\tau'_4}=E_{\tau'_1}+E_{\tau'_2}-E_{\tau'_3}-E_{\tau'_4}.$$

We note that the factor 2 on the right-hand side of the equation is a consequence of the symmetrization of the wave function (we assume everywhere that ψ has already been symmetrized).

With the expression for the two-particle function at our disposal we can readily form the collision integral $I_{\alpha\alpha'}(x,x',t)$ [see (6)]. It is convenient to write it directly in the Wigner representation:

$$
\begin{aligned}
I_{\alpha\alpha'}(x,p,t)&\\
&=\frac{2}{\pi^3}\int d\eta\, d\xi\, dp'dp''U_{\lambda\alpha\theta\sigma}(\eta)\psi\left(\tau_3',\tau_4',\frac{p'-p''}{2};\right.\\
&\quad\left.\times\theta,\sigma,\eta\right)\psi^*\left(\tau_1',\tau_2',\frac{p'-p''}{2};\alpha',\lambda,\eta-2\xi\right)\\
&\quad\times f_{\tau_3'\tau_1'}(p')f_{\tau_4'\tau_2'}(p'')\exp\{i[\mathscr{E}_{\tau_1'\tau_2'\tau_3'\tau_4'}t\\
&\quad+\xi(2p-p'-p'')]\}.
\end{aligned}
$$

The subsequent transformation of this expression requires the use of matrix elements of the T matrix [cf. Ref. 6]. For this purpose we can use the Lippmann-Schwinger equation[13]

$$
\begin{aligned}
&\psi\left(\tau_1',\tau_2',\frac{p'-p''}{2};\alpha',\lambda,x\right)\\
&\quad=\delta_{\tau_1'\alpha'}\delta_{\tau_2'\lambda'}\exp\left(i\,\frac{p'-p''}{2}\,x\right)+G^{(+)}\left(\frac{(p'-p'')^2}{4m}+E_{\tau_1'}\right.\\
&\quad\left.+E_{\tau_2'}\right)\hat{U}\psi\left(\tau_1',\tau_2',\frac{p'-p''}{2};\alpha',\lambda,x\right).
\end{aligned}
$$

Here $G^+(E)=(E-\hat{H}^{(0)})^{-1}$ is the Green's function of the noninteracting particles. This equation provides a facile means for expressing the collision integral exclusively in terms of T-matrix elements.[14] The final form of the quantum kinetic Boltzmann equation can now be written as

$$
\begin{aligned}
&\frac{\partial f_{\alpha\alpha'}(p)}{\partial t}+i(E_\alpha-E_{\alpha'})f_{\alpha\alpha'}(p)+\left(\frac{\mathbf{p}}{m}\cdot\nabla^2\right)f_{\alpha\alpha'}(p)\\
&\quad=-i[I_{\alpha\alpha'}(p)-I^*_{\alpha'\alpha}(p)].
\end{aligned}
\qquad(9)
$$

The variables x and t have been omitted from the Wigner function and the collision integral for brevity. The integral $I_{\alpha\alpha'}(p)$ is expressed in terms of T-matrix elements as follows:

$$
\begin{aligned}
I_{\alpha\alpha'}(p)&=(2\pi)^3\hbar^2\int dp'T_{\lambda\alpha\theta\sigma}\left(\frac{p'-p}{2},\frac{p'-p}{2}\right)\exp[i\mathscr{E}_{\alpha'\lambda\theta\sigma}t]f_{\sigma\lambda}(p')f_{\theta\alpha'}(p)+(2\pi)^4\hbar^2\int dp'dp''d\tilde{p}\,\delta(p+\tilde{p}-p'\\
&-p'')\exp[i\mathscr{E}_{\tau_1\tau_2\tau_3\tau_4}t]\,\frac{T_{\alpha\lambda\tau_3\tau_4}[(p-\tilde{p})/2,(p'-p'')/2]T^*_{\alpha'\lambda\tau_1\tau_2}[(\tilde{p}-p)/2,(p'-p'')/2]}{(p'-p'')^2/4m+E_{\tau_1}+E_{\tau_2}-(\tilde{p}-p)^2/4m-E_{\alpha'}-E_\lambda-i0}\,f_{\tau_3\tau_1}(p')f_{\tau_4\tau_2}(p'').
\end{aligned}
\qquad(10)
$$

A comparison of this equation with the corresponding expression in Snider's paper[6] shows that they concur only when the T matrix does not depend on the internal quantum numbers. In this case the T-quadratic part of the collision integral contains only a term with the energy δ-function, while the pole term (interpreted as a principal-value integral) disappears.

3. DISCUSSION

Consequently, it follows from Eq. (10) that the widely used Waldmann-Snider kinetic equation[6] must be augmented with a pole term. This term vanishes only when the T matrix (and, hence, the scattering amplitude) does not depend on the internal quantum numbers of the colliding particles. This is the situation treated in Ref. 13 and explains why the authors obtained a pole term only in the third order with respect to the gas density. We consider the following special cases as examples. Let the T matrix be initially diagonal:

$$
T_{\alpha\beta\alpha'\beta'}=T_{\alpha\beta}\delta_{\alpha\alpha'}\delta_{\beta\beta'},
$$

and let it be energetically nondegenerate. This approximation is customarily used in the theory of spectral line broadening,[1] even though in reality there is always an off-diagonal term, and indeed this term describes the transition probability between nondegenerate states. For the time being, in the interest of simplicity, we consider only the linear term in T in the kinetic equation for the off-diagonal element of the Wigner distribution function (9). In this case the equation assumes the very simple form

$$
\begin{aligned}
&\frac{\partial f_{\alpha\alpha'}(p)}{\partial t}+i(E_\alpha-E_{\alpha'})f_{\alpha\alpha'}(p)+\left(\frac{\mathbf{p}}{m}\cdot\nabla\right)f_{\alpha\alpha'}(p)\\
&\quad=-(\Gamma+i\nabla^2)f_{\alpha\alpha'}(p),
\end{aligned}
$$

$$
\begin{aligned}
\Gamma+i\nabla^2=i(2\pi)^3\hbar^2\int dp'\{&[T_{\alpha\alpha}(\theta=0)f_{\alpha\alpha}(p')\\
&+T_{\alpha'\alpha}(\theta=0)f_{\alpha'\alpha'}(p')]\exp(i\Delta Et)\\
&-[T^*_{\alpha'\alpha'}(\theta=0)f_{\alpha'\alpha'}(p')\\
&+T^*_{\alpha\alpha'}(\theta=0)f_{\alpha\alpha}(p')]\exp(-i\Delta Et)\}.
\end{aligned}
$$

Here $\Delta E=E_\alpha-E_\alpha'$, and θ is the scattering angle.

This equation differs from the similar type of equation normally used (see Ref. 1). Thus, the relationship of the width and shift of the spectral line to the elements of the T matrix (S matrix) differs from the one usually encountered.[1] The more general case taking into account the quadratic term in T and the off-diagonality of the T matrix, in which case the equation acquires a pole term, poses an independent problem and will be treated separately.

The second case refers to a paramagnetic gas, where the energy levels corresponding to different spin projections are nondegenerate. The simplest model for this case essentially takes only two types of collisions into account: collisions of particles with parallel and antiparallel spins, with the conservation of total spin in both cases. When these two processes are taken into account, the T matrix (nonsymmetrized) has the form

$$T_{abcd}=T_d\delta_{ac}\delta_{bd}+T_e\sigma_{ac}^{(i)}\sigma_{bd}^{(i)}$$

(σ denotes the Pauli spin matrices). Now the pole term in the kinetic equation (9) includes only the imaginary part of the interference term $T_e^*T_d$. Consequently, only in the (hypothetical) event of exact coincidence of the scattering amplitudes for particles with parallel and antiparallel spins ($T_d=T_e$) is the pole term absent.

We note that a pole term in the kinetic equation for the density matrix of a paramagnetic gas first appeared in a paper by Silin.[10] However, the term in that paper has a different form from our expression (10). Moreover, a kinetic equation containing a pole term has been derived previously by one of the present authors (Andreeva[11]) in a study of the scattering of particles by amorphous impurities. A detailed historical discussion of the origin of the pole term in the kinetic equation can be found in Ref. 2. It follows from the results of the present study that the discussion is not over.

This work has received financial support from the Russian Fund for Fundamental Research (Grant No. 96-02-17312-a).

[1] I. I. Sobelman, L. A. Vainshtein, and E. A. Yukov, *Excitation of Atoms and Broadening of Spectral lines*, Springer-Verlag, Berlin (1981) [Russ. original, Nauka, Moscow (1979)].

[2] A. E. Meyerovich, S. Stepaniants, and F. Laloë, Phys. Rev. B **52**, 6808 (1995).

[3] N. P. Bigelow, P. J. Nacher, and M. Leduc, J. Phys. (Paris) **2**, 2159 (1992).

[4] R. P. Bashkin and A. K. Meyerovich, Adv. Phys. **30**, 1 (1981).

[5] E. P. Bashkin, Usp. Fiz. Nauk **148**, 433 (1986) [Sov. Phys. Usp. **29**, 238 (1986)].

[6] R. F. Snider, J. Chem. Phys. **32**, 1051 (1960).

[7] L. Waldmann, Z. Naturforsch. Teil A **13**, 609 (1958).

[8] A. E. Meyerovich, Phys. Rev. B **39**, 9318 (1989).

[9] N. N. Bogolyubov, in *Studies in Statistical Mechanics*, Vol. 1, J. De Boer and G. E. Uhlenbeck (eds.), Wiley, New York (1961), p. 5 [Russ. original, Gostekhizdat, Moscow (1946)].

[10] V. P. Silin, *Introduction to the Kinetic Theory of Gases* [in Russian], Nauka, Moscow (1991).

[11] T. L. Andreeva, Zh. Éksp. Teor. Fiz. **54**, 641 (1968) [Sov. Phys. JETP **27**, 342 (1968)].

[12] J. E. Moyal, Cambridge Philos. Soc. **45**, 100 (1949).

[13] B. A. Lippmann and J. Schwinger, Phys. Rev. **79**, 449 (1950).

[14] A. S. Davydov, *Theory of the Atomic Nucleus* [in Russian], Fizmatgiz, Moscow (1958).

Translated by James S. Wood

Combined configuration-superposition and many-particle perturbation calculations for atoms with two valence electrons

M. G. Kozlov and S. G. Porsev*)

B. P. Konstantinov St. Petersburg Institute of Nuclear Physics, 188350 Gatchina, Leningrad Province, Russia

(Submitted 2 September 1996)
Zh. Éksp. Teor. Fiz. **111**, 838–846 (March 1997)

A new high-precision method is used to calculate the characteristics of atoms with two valence electrons. An effective Hamiltonian for the valence electrons is formulated by many-particle perturbation theory with respect to the residual interaction of the valence electrons with the core. The configuration-superposition method is then used to find the energy levels of the atom. The application of the combined method to divalent calcium, strontium, barium, and ytterbium atoms shows that the ionization potential is obtained within 0.5% error limits. The precision attained for the first few lowest levels of the energy spectra is significantly higher than is obtained by configuration-superposition calculations alone. © *1997 American Institute of Physics.* [S1063-7761(97)00603-3]

1. INTRODUCTION

The fruitfulness of atomic physics methods in the investigation of fundamental interactions is well known, especially in the study of discrete symmetries.[1–4] The nonconservation of parity has been measured to within 1% for a number of atoms.[5–8] However, only for cesium[9,10] and francium[11] has a comparable precision been attained in the calculations required for the interpretation of experiments. The need for a new method by which to improve the precision of atomic calculations is obvious in this light. This method, representing a combination of two well-known methods (the superposition of configurations and many-particle perturbation theory) was proposed earlier.[12]

The application of many-particle perturbation theory to heavy atoms with several valence electrons cannot be expected to yield high computational precision, because electrostatic interaction between valence electrons cannot be precisely accounted for within the context of this method. However, this situation can be rectified by the configuration-superposition method (or the related multiconfigurational Hartree–Fock method), which has been used on many occasions in calculations for complex atoms.[13–20] But then the precision of the configuration-superposition calculations is also limited by the impossibility of fully taking into account the correlations between core and valence electrons. These methods can therefore complement one another, and therein lies the motivation for their combining them.

2. BRIEF DESCRIPTION OF THE METHOD

The method has been proposed in two previous papers,[12] which include detailed descriptions. We shall therefore confine the present discussion to a review of its basic precepts. The complete Hilbert space, in which the many-particle Dirac equation

$$H\Psi = E\Psi \tag{1}$$

operates, is partitioned into two subspaces, and appropriate projection operators P and Q are introduced, satisfying the condition $P+Q=1$. The Hamiltonian H and the wave function Ψ can be written as follows in the P, Q formalism:

$$H = PHP + PHQ + QHP + QHQ, \tag{2}$$

$$\Psi = P\Psi + Q\Psi \equiv \Phi + \chi. \tag{3}$$

The operator P is defined as the projector onto those states of the atom for which N_c electrons are always present in the core (i.e., the core does not contain any holes, as all N_c states are filled with electrons). The operator Q projects onto the orthogonal complement of subspace P (i.e., onto those states of the atom where at least one hole is present in the core).

The following problem is solved in the configuration-superposition stage:

$$(PHP)\Phi = E^{CI}\Phi. \tag{4}$$

The solution of this equation is not a solution of Eq. (1), since the subspace Q is not taken into account here. It has been shown[12] that the inclusion of this subspace leads to the equations

$$[PHP + \Sigma(E)]\Phi = E\Phi, \tag{5}$$

$$\Sigma(E) = (PHQ)\,\frac{1}{E-QHQ}\,(QHP). \tag{6}$$

The normalization condition $\langle\Psi_i|\Psi_k\rangle = \delta_{ik}$ can be rewritten in terms of the functions Φ_i in the form

$$\langle\Phi_i|1 + (PHQ)\,\frac{1}{E_i-QHQ}\,\frac{1}{E_k-QHQ}\,(QHP)|\Phi_k\rangle = \delta_{ik}. \tag{7}$$

If we restrict the discussion to a few of the lowest levels, i.e., if we assume that $\Delta \equiv (E_i-E_k)/2$ is much smaller than $E_{\mathrm{av}} = (E_i+E_k)/2$, we see at once that the second term in (7) can be written in the series form:

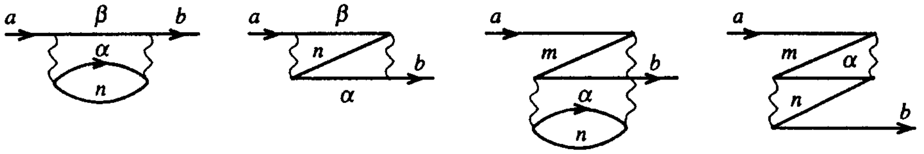


FIG. 1. Self-energy diagrams for valence electrons.

$$(PHQ)\,\frac{1}{E_i-QHQ}\,\frac{1}{E_k-QHQ}\,(QHP)$$
$$=\left(-\frac{\partial\Sigma(E)}{\partial E}-\frac{\Delta^2}{6}\,\frac{\partial^3\Sigma(E)}{\partial E^3}-\cdots\right)\Bigg|_{E=E_v}. \qquad (8)$$

The first term of the expansion (8) is already much smaller than unity. Consequently, in calculating the lowest energy levels, we can simply omit the second term in Eq. (7) and write the normalization condition in the conventional form

$$\langle\Phi_i|\Phi_k\rangle=\delta_{ik}. \qquad (9)$$

It has been shown[12] that the operator $\Sigma(E)$ can be transformed as follows for the proper choice of core orbitals:

$$\Sigma(E)=P(V-V^{N_{PT}})Q\,\frac{1}{E-QHQ}\,Q(V-V^{N_{PT}})P, \quad (10)$$

where V denotes the two-electron electrostatic interaction, and $V^{N_{PT}}$ is the interaction of N_{PT} electrons with the Hartree–Fock field. The quantity N_{PT} must satisfy the condition $N_c \leqslant N_{PT} \leqslant N$. For the calculation of the energy denominators in the indicated approximation E is replaced by E_{av} in Eq. (10).

The operator Σ is calculated by standard diagram technique. In the lowest order the corresponding diagrams can have one, two, or three outer lines.[12] Diagrams of the first type describe corrections to the one-electron part PHP (Figs. 1 and 2). Diagrams of the second type give a correction for shielding of the core by the interaction between valence electrons (Figs. 3 and 4). Diagrams corresponding to three-particle interaction, of course, do not exist for atoms with two valence electrons.

We note that the diagrams in Figs. 2 and 4 are drawn with allowance for the fact that the Hartree–Fock field includes contributions from $N_{PT}-N_c$ valence electrons. They obviously vanish for the case $N_{PT}=N_c$. Since the potential $V^{N_{PT}}$ occurs with a minus sign in Eq. (10), these diagrams are called subtractive.

3. COMPUTATIONAL PROCEDURE

The procedures used to perform the calculations can be divided into three parts. In the first stage a basis set of one-electron wave functions (orbitals) is formulated. The one-electron wave functions of the ground state of the atom and the corresponding one-electron energies are found by the Hartree–Fock–Dirac method. We use a program written by Brattsev, Deĭneka, and Tupitsyn[21] and subsequently modified by Tupitsyn (in particular, with allowance for the finite size of the nucleus, which is important in the treatment of heavy atoms). In addition to the Hartree–Fock variety we also construct virtual orbitals. Each of these can be represented by the product of one of the corresponding orbitals of the same symmetry and the exponent r with subsequent orthogonalization of the newly constructed orbitals with respect to all those preceding. A similar method for the construction of basis functions has been proposed by Bogdanovich[22,23] and has already been used by us for calculations of ytterbium[14] and bismuth.[15]

In the second stage it is required to calculate the matrix elements of the operator Σ. To calculate the self-energy diagrams and the shielding diagrams by perturbation theory, we

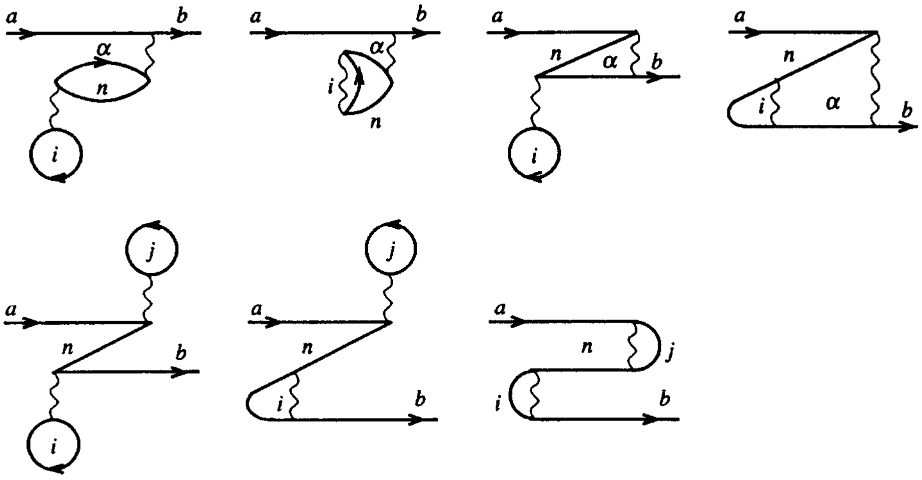


FIG. 2. Subtractive diagrams for the self-energy.

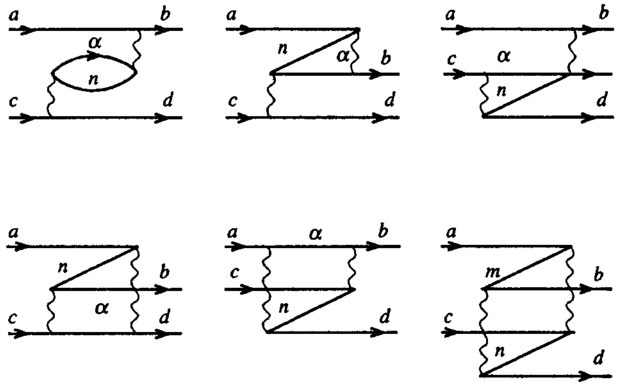


FIG. 3. Shielding diagrams.

use the constructed set of basis orbitals. Since we are concerned in this stage with the excitation of electrons having inner shells (which can be quite deeply situated), we must take into account a large number of states of the continuous spectrum as excited states. The need to include a large number of diagrams for a large number of excited states makes this part of the calculations the most time-consuming.

The third and final step is to solve Eq. (5). This can be done by the well-known configuration-superposition method. In the calculations we use a program written by Kotochigova and Tupitsyn,[24] which we have substantially modified.

4. CALCULATIONS FOR Ca, Sr, Ba, AND Yb

We have chosen the Ca, Sr, Ba, and Yb atoms to test the configuration-superposition method in combination with many-particle perturbation theory. They have in common the existence of two s-electrons in the outermost shell. It is reasonable, therefore, to assume that s-electrons are present in the valence zone and that the remaining electrons form the core.

As mentioned, the first step in calculating the energy spectrum is to obtain the one-electron functions by the Hartree–Fock–Dirac method. In this stage a self-consistent field procedure is implemented for all electrons in the atom, including electrons of the outermost shell. The resulting orbitals are then frozen in place, one electron from the last ns shell is moved to the $(n-1)d$ shell, and the Hartree–Fock–Dirac equation for this shell is solved ($n=4,5$ for Ca and Sr, and $n=6$ for Ba and Yb). The np shell is augmented analogously. The procedure culminates in the construction of Hartree–Fock orbitals for all electronic states of the main configuration of atoms and also for the np and $(n-1)d$ shells. The virtual orbitals, which together with the Hartree–Fock orbitals form the basis set, are subsequently constructed according to the procedure set forth in the preceding section. The number of orbitals that must be included in the basis will be discussed below.

The next stage is to calculate the matrix elements of the operator Σ by perturbation theory. In the case of the V^N approximation, which is the one used for the calculations, all the diagrams of Figs. 1–4 need to be taken into account (the subtractive diagrams do not vanish in this case, because $N_{PT} \neq N_c$). In the perturbation calculations the number of excited states that must be included to attain saturation is fairly high. The test of whether saturation has been reached is that the matrix elements of the operator Σ essentially cease to change as new shells are added. For example, in the calculation of the energy spectrum of Yb, the following virtual states are included: $7s-19s$, $6p-19p$, $5d-18d$, $5f-17f$, $5g-16g$. The number of orbitals included for the other atoms is approximately the same. Once the virtual orbitals have been constructed, the Hartree–Fock–Dirac operator is diagonalized on them. In this case the majority of the orbitals of each symmetry resides in the continuous spectrum, and the energy of the last orbitals is on the order of 10^2 a.u.

Next we solve Eq. (5) by the configuration-superposition method, using the calculated matrix elements of the operator Σ. In this state we take into account configurations associated with the excitation of both valence (but not core!) electrons to higher levels. The configurations are augmented until the energy of the investigated levels no longer changes. (We assume that saturation is attained when the energy of the levels does not vary by more than 30–40 cm^{-1} with the addition of new configurations.) The basis set of orbitals required to attain saturation in the configuration-superposition calculations is far smaller than in the state where many-particle perturbation theory is used. In our case it is sufficient to include 18–27 virtual orbitals. The two or three highest s-, p-, d-, f- and g-shells are added as virtual. As explained, excitations in the h-shell are insignificant for the given atoms and can be disregarded.

For comparison we also solve Eq. (4) instead of (5) to obtain a solution by the pure configuration-superposition method. The results of calculations of the lowest energy levels for both methods, the pure configuration-superposition method (SC) and its coordination with the many-particle perturbation theory (SC+MPP) for the Ca, Sr, Ba, and Yb atoms are summarized in Tables I–IV.

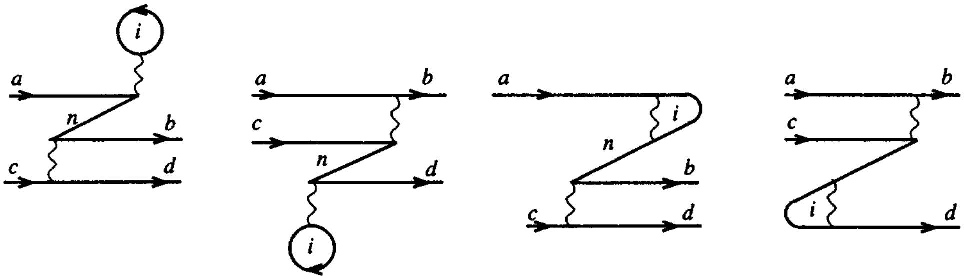


FIG. 4. Subtractive diagrams for shielding.

TABLE I. Energies of the lowest few levels of Ca (cm^{-1}) (multiplet splittings are shown in parentheses).

Configuration	Level	SC	SC+MPP	Experiment[25]
$4s^2$	$^1S_0$	0	0	0
$4s4p$	$^3P^o_0$	13720	15230	15158
$4s4p$	$^3P^o_1$	13769 (49)	15284 (54)	15210 (52)
$4s4p$	$^3P^o_2$	13870 (101)	15394 (110)	15316 (106)
$4s3d$	$^3D_1$	23661	21489	20335
$4s3d$	$^3D_2$	23664 (3)	21505 (16)	20349 (14)
$4s3d$	$^3D_3$	23664 (0)	21530 (25)	20371 (22)
$4s3d$	$^1D_2$	23642	22984	21850
$4s4p$	$^1P^o_1$	23255	23555	23652

TABLE III. Energies of several lowest levels of Ba (cm^{-1}).

Configuration	Level	SC	SC+MPP	Experiment[25]
$6s^2$	$^1S_0$	0	0	0
$6s5d$	$^3D_1$	11019	9423	9034
$6s5d$	$^3D_2$	11104 (85)	9631 (208)	9216 (182)
$6s5d$	$^3D_3$	11281 (177)	10065 (434)	9597 (381)
$6s6p$	$^3P^o_0$	10253	12221	12266
$6s6p$	$^3P^o_1$	10597 (344)	12583 (362)	12637 (371)
$6s6p$	$^3P^o_2$	11370 (773)	13448 (865)	13515 (878)
$6s6p$	$^1P^o_1$	17157	17740	18060

In addition to the energy spectra we can also calculate the ionization potentials of these atoms. Thus, in the case of two valence electrons their energy E_v is simply the sum of the first and second ionization potentials. Consequently, to find the first two ionization potentials, it is sufficient to also solve Eq. (5) for a single positive ion. If the same basis set of orbitals is used, the radial integrals needed in order to calculate the matrix elements of the operators H and Σ are the same as for the neutral atom. This fact is extremely important, since the calculation of the radial integrals for the operator Σ is the most costly operation in the computational sense. Knowing E_v for the atom and the ion, we can obtain the ionization potential for each of them without undue complication (see Tables V and Tables VI).

Analyzing the results obtained for the ionization potentials, we see that the combined method yields precision at the 0.5% level, which is roughly an order of magnitude better than obtained by pure configuration-superposition calculations. The energy spectrum obtained for the lower levels for all four atoms by the configuration-superposition method in conjunction with many-particle perturbation theory also exhibits significantly better agreement with the experimental. For odd levels the precision is close to that attained for the ionization potentials (i.e., for the most part better than 1%); for even levels (particularly the D levels) the precision is somewhat lower but still far better than for pure configuration-superposition calculations. The change in the multiplet splitting is especially dramatic when correlations are taken into account by perturbation theory. In the case of Yb, for example, it is evident from Table IV that the configuration-superposition method gives approximately one tenth of the experimental splitting of the D triplet, whereas the combined method reproduces this splitting almost perfectly. The loss of precision for even as opposed to odd levels is most likely attributable to the fact that the Hartree–Fock d-functions are a poor approximation to the true wave functions. (We note that an attempt to use virtual rather than Hartree–Fock d-functions does nothing to improve the situation.)

5. CONCLUSION

Our calculations for atoms with two valence electrons (Ca, Sr, Ba, and Yb) by the combined method of configuration superposition and many-particle perturbation theory confirms that the accuracy of computation of such atomic characteristics as the ionization potential and the energy spectrum is significantly higher than when the pure configuration-superposition method is used. In particular, the ionization potential has been reproduced within 0.5% error limits.

Work is currently in progress on the application of the combined method for transition amplitude calculations. This approach will make it possible to improve the precision of calculation of impurity P-odd amplitudes needed for the interpretation of experimental parity-nonconservation results and for testing the standard model.

The authors are grateful to Yu. G. Rakhlina for assisting with the ytterbium calculations. This work has received partial support from the Russian Fundamental Research Foundation (Grant No. 95-02-03701-a).

TABLE II. Energies of several lowest levels of Sr (cm^{-1}).

Configuration	Level	SC	SC+MPP	Experiment[25]
$5s^2$	$^1S_0$	0	0	0
$5s5p$	$^3P^o_0$	12475	14242	14318
$5s5p$	$^3P^o_1$	12648 (173)	14428 (186)	14504 (186)
$5s5p$	$^3P^o_2$	13007 (359)	14821 (393)	14899 (395)
$5s4d$	$^3D_1$	19618	18877	18159
$5s4d$	$^3D_2$	19635 (17)	18936 (59)	18219 (50)
$5s4d$	$^3D_3$	19664 (29)	19033 (97)	18319 (100)
$5s5p$	$^1P^o_1$	20863	21444	21698

TABLE IV. Energies of several lowest levels of Yb (cm^{-1}).

Configuration	Level	SC	SC+MPP	Experiment[25]
$6s^2$	$^1S_0$	0	0	0
$6s6p$	$^3P^o_0$	14357	17075	17288
$6s6p$	$^3P^o_1$	15022 (665)	17764 (689)	17992 (704)
$6s6p$	$^3P^o_2$	16527 (505)	19447 (683)	19710 (718)
$6s5d$	$^3D_1$	25216	25075	24489
$6s5d$	$^3D_2$	25238 (22)	25338 (263)	24751 (262)
$6s5d$	$^3D_3$	25299 (61)	25855 (517)	25270 (519)
$6s6p$	$^1P^o_1$	24221	25306	25068

TABLE V. Ionization potentials of Ca^+, Sr^+, Ba^+, and Yb^+ (cm^{-1}).

Ion	SC	SC+MPP	Experiment
Ca^+	91887	95537	95748
Sr^+	84635	88747	88964
Ba^+	76011	80421	80687
Yb^+	90788	97254	98269

TABLE VI. Ionization potentials of neutral atoms (cm^{-1}).

Atom	SC	SC+MPP	Experiment
Ca	47806	49142	49305
Sr	44057	45679	45926
Ba	39881	41800	42032
Yb	46759	50295	50444

*)e-mail: kozlov@lnpi.spb.su; porsev@thd.pnpi.spb.ru

---

[1] I. B. Khriplovich, *Parity Nonconservation in Atomic Phenomena* [in Russian], Nauka, Moscow (1989).

[2] A.-M. Mårtensson-Pendrill, *Methods in Computational Chemistry*, Vol. 5: *Atomic and Molecular Properties*, S. Wilson (ed.), Plenum Press, New York (1992).

[3] S. M. Barr, J. Mod. Phys **8**, 209 (1993).

[4] M. G. Kozlov and L. N. Labzowsky, J. Phys. B **28**, 1933 (1995).

[5] S. L. Gilbert, M. C. Noecker, R. N. Watts, and C. E. Wieman, Phys. Rev. Lett. **55**, 2680 (1985); M. C. Noecker, V. P. Masterson, and C. E. Wieman, Phys. Rev. Lett. **61**, 310 (1988).

[6] D. M. Meekhof, P. Vetter, P. K. Majunder, S. K. Lamoreaux, and E. N. Fortson, Phys. Rev. Lett. **71**, 3442 (1993).

[7] P. A. Vetter, D. M. Meekhof, P. K. Majumder, S. K. Lamoreaux, and E. N. Fortson, in *Proceedings of the 14th ICAP*, Boulder, Col. (1994), ID-3.

[8] M. J. D. Macpherson, K. P. Zetie, R. B. Warrington, D. N. Stacey, and J. P. Hoare, Phys. Rev. Lett. **67**, 2784 (1991).

[9] V. A. Dzuba, V. V. Flambaum, and O. P. Sushkov, Phys. Lett. A **141**, 147 (1989).

[10] S. A. Blundell, W. R. Johnson, and J. Sapirstein, Phys. Rev. Lett. **65**, 1411 (1990).

[11] V. A. Dzuba, V. V. Flambaum, and O. P. Sushkov, Phys. Rev. A **51**, 3454 (1995).

[12] V. A. Dzuba, V. V. Flambaum, and M. G. Kozlov, Phys. Rev. A **54**, 3948 (1996); V. A. Dzuba, V. V. Flambaum, and M. G. Kozlov, JETP Lett. **63**, 882 (1996).

[13] V. A. Dzuba, V. V. Flambaum, and M. G. Kozlov, Phys. Rev. A **50**, 3812 (1994).

[14] S. G. Prosev, Yu. G. Rakhlina, and M. G. Kozlov, JETP Lett. **61**, 459 (1995).

[15] M. G. Kozlov, S. G. Porsev, and V. V. Flambaum, J. Phys. B **29**, 689 (1996).

[16] I. P. Grant and H. M. Quiney, Adv. At. Mol. Phys. **23**, 37 (1988).

[17] P. Jonsson and C. Froese Fischer, Phys. Rev. A **50**, 3080 (1994).

[18] A. Aboussaid, M. R. Godefroid, P. Jonsson, and C. Froese Fischer, Phys. Rev. A **51**, 2031 (1995).

[19] C. Froese Fischer, J. E. Hansen, and H. W. van den Hart, Phys. Rev. A **51**, 1999 (1995).

[20] J. E. Wienkiewicz, S. Fritzsche, and I. P. Grant, J. Phys. B **28**, L633 (1995).

[21] V. F. Brattsev, G. B. Deǐneka, and I. I. Tupitsyn, Izv. Akad. Nauk Ser. Fiz. **41**, 2655 (1977).

[22] P. Bogdanovich, and G. Ẑukauskas, Sov. Phys. Coll. **23**, 18 (1983).

[23] P. Bogdanovich, G. Ẑukauskas, and S. Ŝadẑiuviene, Sov. Phys. Coll. **24**, 27 (1984).

[24] S. A. Kotochigova and I. I. Tupitsin, J. Phys. B **20**, 4759 (1987).

[25] C. E. Moore, Circular No. 467, U.S. Nat. Bureau Stand., Washington, D.C. (1958).

Translated by James S. Wood

Spontaneous and induced Cherenkov radiation generated by electrons in cylindrical dielectrics

N. K. Zhevago and V. I. Glebov

Kurchatov Institute Russian Scientific Center, 123182 Moscow, Russia

(Submitted 14 May 1996)
Zh. Éksp. Teor. Fiz. **111**, 847–861 (March 1997)

A theory of induced Cherenkov radiation in cylindrically symmetrical dielectrics in the case when an electron beam is moving close to the dielectric surface is presented. The spectrum of excited radiation modes has been investigated, and analytical expressions for the gain at the frequencies of various modes have been derived. *© 1997 American Institute of Physics.* [S1063-7761(97)00703-8]

1. INTRODUCTION

Induced Cherenkov radiation makes feasible free-electron lasers operating over a broad spectral range.[1–5] The major factor reducing the gain in Cherenkov free-electron lasers is multiple Coulomb scattering of electrons, which is important even in gases and relatively thin targets. A design alternative to that of a free-electron laser, in which an electron beam passes through material, is a scheme using electrons moving next to a dielectric surface.[4,6] In this case, multiple electron scattering is no longer a factor, but the intensity of Cherenkov radiation drops exponentially when the distance to the dielectric surface exceeds $\gamma\lambda\!\!\!^{-}/2\pi$, where γ is the Lorentz factor of the electrons and $\lambda\!\!\!^{-}$ is the radiation wavelength. Nonetheless, it has been possible to detect Cherenkov radiation of fairly high intensity in the microwave[7–9] and far-infrared[7] ranges in devices using waveguides coupled to dielectrics.

The paper presents a theory of Cherenkov radiation in the case when an electron beam moves near a solid dielectric cylinder or inside a hollow dielectric cylinder parallel to its axis. First we will describe the general approach to the problem of the electron energy loss, which can be also applied to other problems, such as plasmon generation, or to more complicated structures, such as an optic fiber with smoothly varying refraction index. Then we will investigate the spectrum of excited Cherenkov radiation modes in solid and hollow dielectric cylinders. The ultimate aim of this research was to study induced Cherenkov radiation at relatively short wavelengths, namely in the visible and near-infrared bands. In the limit of a cold electron beam and small gain, analytical expressions for the gain at frequencies of different modes will be given, and optimal conditions for light amplification will be presented.

2. ELECTROMAGNETIC LOSSES OF ELECTRON ENERGY IN A MEDIUM WITH CYLINDRICAL SYMMETRY

The theory of spontaneous Cherenkov radiation generated by a charge moving parallel to the axis of a cylindrical channel in a dielectric was developed by Bogdankevich and Bolotovskii.[10] But they analyzed the case of an electron moving inside a cylinder, whereas the case of an electron outside a cylinder is more interesting with a view to designing a free-electron laser. We will present a general approach to the problem in order to obtain equations for an arbitrary separation ρ_0 between the electron and the cylinder axis and more refined analytical expressions for spectra of spontaneous Cherenkov radiation, which are crucial for further analysis of induced Cherenkov radiation.

Let us first consider the case of a solid cylinder, and the more complicated case of a hollow cylinder can be analyzed similarly (Sec. 4). Let the medium around the cylinder be characterized by a dielectric permittivity $\varepsilon_1(\omega)$ (which is, generally speaking, complex), and the material of the cylinder by $\varepsilon_2(\omega)$, and let an electron move along the cylinder axis with a velocity v. In order to find solutions of the inhomogeneous Maxwell equations determined by the current and charge of a relativistic electron, we express the electromagnetic field in the form of Fourier integrals with respect to time and the distance z along the cylinder axis, and in the form of a Fourier series in the azimuthal angle ϕ, which is measured with respect to the plane passing through the cylinder axis and the electron trajectory:

$$\mathbf{E}(\mathbf{r},t)=(2\pi)^{-2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\mathbf{E}(\rho,m,k_z,\omega)\times\exp(im\phi-i\omega t+ik_zz)\,dk_z\,d\omega, \tag{1}$$

$$\mathbf{H}(\mathbf{r},t)=(2\pi)^{-2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\mathbf{H}(\rho,m,k_z,\omega)\times\exp(im\phi-i\omega t+ik_zz)\,dk_z\,d\omega.$$

The energy W lost by the electron as it moves equals the work done by the electric field on the charge:

$$W=\int_{-\infty}^{\infty}\mathbf{j}^*(\mathbf{r},t)\mathbf{E}(\mathbf{r},t)\,d^3r\,dt, \tag{2}$$

where only the longitudinal current component is nonzero:

$$j_z(\mathbf{r},t)=ev\,\delta(\mathbf{r}_\perp-\mathbf{r}_{\perp 0})\,\delta(z-vt);$$

here $\mathbf{r}_{\perp 0}$ is the electron radius-vector in the plane perpendicular to the z-axis, $\mathbf{r}=(\mathbf{r}_{\perp 0},z)$. Let us use an expansion like the one in Eq. (1) for both the electron current and electric field and perform the integration in Eq. (2). As a result, we have the spectral expansion of the energy losses in the form

$$W=\frac{ev}{\pi}\,\mathrm{Re}\sum_{m=-\infty}^{\infty}\int_0^{\infty}E_z\left(\rho_0,m,\frac{\omega}{v},\omega\right)d\omega. \tag{3}$$

As might be expected, the electron energy losses are determined by those modes whose phase velocity ω/k_z parallel to the cylinder axis coincides with the electron velocity v, and only the modes with the nonzero z-component of the electric field contribute.

Let us seek a general solution of the inhomogeneous Maxwell equations in the form of a sum of a particular solution of the inhomogeneous equations, hereinafter identified by the upper index (p), and the general solution of the homogeneous equations identified by the index (h). The first term of this sum is the electromagnetic field generated by a moving electron, and in each region of space, namely $\rho<a$ and $\rho>a$, where $\rho=|\mathbf{r}_\perp|$, it can be expressed in terms of the z-component of the vector potential A_z:

$$H_\rho^{(p)}=\frac{im}{\rho}A_z,\quad H_\varphi^{(p)}=-\frac{\partial A_z}{\partial\rho},\quad H_z^{(p)}=0,$$

$$E_\rho^{(p)}=\frac{c}{v\varepsilon}\frac{\partial A_z}{\partial\rho},\quad E_\varphi^{(p)}=-\frac{im}{v\varepsilon\rho}A_z,$$

$$E_z^{(p)}=i\left(\frac{\omega}{c}-\frac{ck_z}{v\varepsilon}\right)A_z,$$

where the subscripts denote components in the cylindrical coordinate system, and ε is the permittivity of the medium:

$$\varepsilon=\begin{cases}\varepsilon_1(\omega) & \text{for } \rho>a,\\ \varepsilon_2(\omega) & \text{for } \rho<a.\end{cases}$$

The z-component A_z of the vector-potential can be expressed in turn as

$$A_z(\rho,m,k_z,\omega)=\frac{4\pi e}{c}f_m(\rho,\rho_0)\,\delta\left(k_z-\frac{\omega}{v}\right), \tag{4}$$

$$f_m(\kappa\rho)=\begin{cases}K_m(\kappa\rho_0)I_m(\kappa\rho) & \text{for } \rho\leqslant\rho_0,\\ I_m(\kappa\rho_0)K_m(\kappa\rho) & \text{for } \rho\geqslant\rho_0.\end{cases} \tag{5}$$

Here K_m and I_m are modified Bessel functions, and $\kappa=((\omega/c)^2\varepsilon-k_z^2)^{1/2}$.

It is known[11] that the general solution of the homogeneous Maxwell equations can be expressed as a linear combination of TE-modes (in which the longitudinal electric field component $E_z^{(\mathrm{TE})}$ is identically zero) and TM-modes (in which $H_z^{(\mathrm{TM})}\equiv 0$). The nonzero components of the fields $E_z^{(\mathrm{TM})}$ and $H_z^{(\mathrm{TE})}$ satisfy the equation

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial F}{\partial\rho}\right)-(m^2-\kappa^2)F=0,$$

and the transverse components can be expressed in terms of $E_z^{(\mathrm{TM})}$ and $H_z^{(\mathrm{TE})}$. In particular, the tangential field components are expressed as

$$E_\varphi^{(h)}=\frac{i\omega}{c\kappa^2}\frac{\partial H_z^{(\mathrm{TE})}}{\partial\rho}+\frac{mk_z}{\kappa^2\rho}E_z^{(\mathrm{TM})},$$

$$H_\varphi^{(h)}=-\frac{i\omega\varepsilon}{c\kappa^2}\frac{\partial E_z^{(\mathrm{TM})}}{\partial\rho}+\frac{mk_z}{\kappa^2\rho}H_z^{(\mathrm{TE})}.$$

Let us seek the longitudinal components of free fields for $k_z=\omega/v$ in the form

$$E_z^{(\mathrm{TM})}=LA_m^{(\mathrm{TM})}K_m(\kappa_1\rho),\quad H_z^{(\mathrm{TE})}=LA_m^{(\mathrm{TE})}K_m(\kappa_1\rho),\quad \rho\geqslant a,$$

$$E_z^{(\mathrm{TM})}=LB_m^{(\mathrm{TM})}I_m(\kappa_2\rho),\quad H_z^{(\mathrm{TE})}=LB_m^{(\mathrm{TE})}I_m(\kappa_2\rho),\quad \rho\leqslant a, \tag{6}$$

where $A_m^{(\mathrm{TM})}$, $B_m^{(\mathrm{TM})}$, $A_m^{(\mathrm{TE})}$, and $B_m^{(\mathrm{TE})}$ are unknown coefficients, $\kappa=(\omega/v)\sqrt{1-\varepsilon\beta^2}$, and $\beta=v/c$. The factor $2\pi\delta(k_z-\omega/v)$ in Eq. (4) should be interpreted for $k_z=\omega/v$ as the electron free path L. The unknown coefficients are derived from continuity conditions for the tangential and longitudinal components for the electric $\mathbf{E}=\mathbf{E}^{(p)}+\mathbf{E}^{(h)}$ and magnetic $\mathbf{H}=\mathbf{H}^{(p)}+\mathbf{H}^{(h)}$ fields. These coefficients are determined by the following linear algebraic equation system:

$$A_m^{(\mathrm{TE})}K_m(\kappa_1a)=B_m^{(\mathrm{TE})}I_m(\kappa_2a),$$

$$A_m^{(\mathrm{TM})}K_m(\kappa_1a)+\frac{2ie\omega}{c^2}\left(1-\frac{1}{\varepsilon_1\beta^2}\right)f_m(\kappa_1a)=B_m^{(\mathrm{TM})}I_m(\kappa_2a)+\frac{2ie\omega}{c^2}\left(1-\frac{1}{\varepsilon_2\beta^2}\right)f_m(\kappa_2a),$$

$$\frac{i\omega}{c\kappa_1^2}\left[A_m^{(\mathrm{TE})}\kappa_1K'_m(\kappa_1a)-\frac{im}{\beta a}A_m^{(\mathrm{TM})}K_m(\kappa_1a)\right]-\frac{2iem}{v\varepsilon_1a}f_m(\kappa_1a)=\frac{i\omega}{c\kappa_2^2}\left[B_m^{(\mathrm{TE})}\kappa_2I'_m(\kappa_2a)-\frac{im}{\beta a}B_m^{(\mathrm{TM})}I_m(\kappa_2a)\right]-\frac{2iem}{v\varepsilon_2a}f_m(\kappa_2a),$$

$$\frac{i\omega}{c\kappa_1^2}\left[\frac{im}{\beta a}A_m^{(\mathrm{TE})}K_m(\kappa_1a)+\varepsilon_1\kappa_1A_m^{(\mathrm{TM})}K'_m(\kappa_1a)\right]+\frac{2e\kappa_1}{c}f'_m(\kappa_1a)=\frac{i\omega}{c\kappa_2^2}\left[\frac{im}{\beta a}B_m^{(\mathrm{TE})}I_m(\kappa_2a)+\varepsilon_2\kappa_2B_m^{(\mathrm{TM})}I'_m(\kappa_2a)\right]+\frac{2e\kappa_2}{c}f'_m(\kappa_2a),$$

where the primes denote differentiation with respect to the argument.

After solving this system of equations, we obtain the longitudinal electric field component $E_z(\rho_0,m,\omega/v,\omega)$:

$$
\begin{aligned}
&L^{-1}E_z(\rho_0\leqslant a)\\
&\quad=\frac{\Delta_m(\omega)}{D_m(\omega)}I_m(\kappa_2 a)-\frac{2ie}{\omega}\frac{\kappa_2^2}{\varepsilon_2}I_m(\kappa_2\rho_0)K_m(\kappa_2 a),\\
&L^{-1}E_z(\rho_0\geqslant a)\\
&\quad=\frac{2ie}{\omega}\left\{\frac{\kappa_1^4 I_m(\kappa_2 a)}{aD_m(\omega)}\frac{K_m^2(\kappa_1\rho_0)}{K_m(\kappa_1 a)}\right.\\
&\qquad\times\left[\frac{I_m(\kappa_2 a)K'_m(\kappa_1 a)}{\kappa_1}-\frac{I'_m(\kappa_2 a)K_m(\kappa_1 a)}{\kappa_2}\right]\\
&\qquad\left.+\frac{\kappa_1^2}{\varepsilon_1}K_m(\kappa_1\rho_0)I_m(\kappa_1 a)\left[\frac{K_m(\kappa_1\rho_0)}{K_m(\kappa_1 a)}-1\right]\right\}.
\end{aligned}
\tag{7}
$$

Here we have introduced the notation

$$
\begin{aligned}
\Delta_m=&\frac{2ie}{\omega}I_m(\kappa_2\rho_0)\left\{\varepsilon_1\kappa_1^4\kappa_2^2\right.\\
&\times\left[\frac{I_m(\kappa_2 a)K'_m(\kappa_1 a)}{\kappa_1}-\frac{I'_m(\kappa_2 a)K_m(\kappa_1 a)}{\kappa_2}\right]\\
&\times\left[\frac{K_m(\kappa_2 a)K'_m(\kappa_1 a)}{\varepsilon_2\kappa_1}-\frac{K'_m(\kappa_2 a)K_m(\kappa_1 a)}{\varepsilon_1\kappa_2}\right]\\
&\left.-\left(\frac{m}{\beta a}\right)^2\frac{(\kappa_2^2-\kappa_1^2)^2}{\varepsilon_2\kappa_2^2}I_m(\kappa_2 a)K_m(\kappa_2 a)K_m^2(\kappa_1 a)\right\},
\end{aligned}
\tag{8}
$$

$$
\begin{aligned}
D_m(\omega)=&\varepsilon_1\varepsilon_2\kappa_1^4\left[\frac{I_m(\kappa_2 a)K'_m(\kappa_1 a)}{\kappa_1}-\frac{I'_m(\kappa_2 a)K_m(\kappa_1 a)}{\kappa_2}\right]\\
&\times\left[\frac{I_m(\kappa_2 a)K'_m(\kappa_1 a)}{\varepsilon_2\kappa_1}-\frac{I'_m(\kappa_2 a)K_m(\kappa_1 a)}{\varepsilon_1\kappa_2}\right]\\
&-\left(\frac{m}{\beta a}\right)^2\left(1-\frac{\kappa_1^2}{\kappa_2^2}\right)^2 I_m^2(\kappa_2 a)K_m^2(\kappa_1 a).
\end{aligned}
$$

The expressions for the longitudinal field components in Eq. (7) are in agreement with similar results[10] derived using a different technique.

The electron energy loss is determined in the general case by the poles of the electric field component $E_z(\rho_0,m,\omega/v,\omega)$, regarded as a function of the complex frequency ω, in the upper half-plane. When the absorption of electromagnetic waves in both media can be neglected, the imaginary part of the permittivity tends to zero, and the spectrum of possible excitations resulting in energy losses is determined by the zeros of the function $D_m(\omega)$. Thus, the analysis of the spectrum of the electromagnetic energy losses in a relatively transparent material reduces to solving the dispersion equation $D_m(\omega)=0$ and calculating the sum of residues for the integrand in Eq. (3). It follows from our analysis that the dispersion equation can have solutions both in the frequency range where the real part of the dielectric permittivities ε_1 or ε_2 is negative and at frequencies where ε_1 and ε_2 are positive, and the electron velocity is larger than the threshold velocity of the Cherenkov radiation in one of the media. The solution for negative $\varepsilon(\omega)$ in metals describes generation of electron density oscillations (plasmons) and should be the subject of a separate study. We note only that the plasmon spectrum in a cylindrical sample as in a spherical[12] sample can differ from the spectrum in a continuous medium $(\varepsilon_1=\varepsilon_2=\varepsilon)$ determined by the condition $\varepsilon(\omega)=0$.

3. ENERGY LOSSES TO CHERENKOV RADIATION

Now let us investigate the case which is especially interesting from the viewpoint of generating Cherenkov radiation, when an electron travels in vacuum, $\varepsilon_1\equiv 1$, and the cylinder is made from a fairly transparent dielectric with the dielectric permittivity $\varepsilon_2(\omega)\equiv\varepsilon(\omega)$. In this case, the argument of the modified Bessel function in Eqs. (7) and (8) is purely imaginary, and it is more convenient to replace them with the Bessel functions $J_m(x)$ using the well known relations: $I_m(iz)=i^m J_m(z)$. Moreover, note that the dispersion equation $D_m=0$ is quadratic with respect to the ratio J'_m/J_m. As a result, it splits into the two equations

$$
\begin{aligned}
\frac{J'_m(x)}{J_m(x)}=&\frac{(\varepsilon\beta^2-1)^{1/2}}{2\varepsilon(1-\beta^2)^{1/2}}\left\{-(\varepsilon+1)\frac{K'_m(y)}{K_m(y)}\pm(\varepsilon-1)\right.\\
&\left.\times\left(\frac{K'_m(y)}{K_m(y)}\right)^2+\frac{4\varepsilon m^2\beta^2}{(\varepsilon\beta^2-1)^2y^2}\right\},
\end{aligned}
\tag{9}
$$

where we have introduced the notation $x=(\omega a/v)\times(\varepsilon\beta^2-1)^{1/2}$ and $y=(\omega a/v)(1-\beta^2)^{1/2}$, and both plus and minus signs on the right-hand side of this equation should be considered. Let us introduce a binary index σ, which equals 1 or -1, depending on the sign on the right-hand side of Eq. (9), and denote the zeros of the dispersion equations as $\omega_{mn\sigma}$. Further, let us calculate the sum of the residues at ordinary poles $\omega_{mn\sigma}$ at each value of σ, which is equivalent to calculating $\Delta_m(\omega)$ and the derivative $dD_m/d\omega$ at the points $\omega_{mn\sigma}$. Suppose that the frequency range of interest is far from the absorption bands and lines in the material, so that the dispersion of the permittivity $\varepsilon(\omega)$ is relatively small: $|d\varepsilon/d\omega|\ll\varepsilon/\omega$. In this case, we can assume in calculating the derivative $dD_m/d\omega$ to lowest order that ε is independent of the frequency ω. In order to simplify the resulting expressions, we use the recurrence relations between the Bessel functions and their derivatives. We replace the oscillatory terms containing Bessel functions by monotonic functions K_m using the dispersion equation (9). After a long series of transformations, which we do not describe here, we obtain the spectral distribution of the Cherenkov radiation in the form

$$
\begin{aligned}
\frac{dW}{d\omega}=&\frac{4e^2L}{a^2}\sum_{n=1}^{\infty}\left[\frac{\gamma_T^2K_0^2(y\rho_0/a)}{\gamma^2K_1^2(y)+\varepsilon K_0^2(y)}\delta(\omega-\omega_{0n1})\right.\\
&\left.+2\sum_{m=1}^{\infty}\sum_{\sigma=\pm 1}\frac{K_m^2(y\rho_0/a)}{K_m^2(y)}\frac{S_m+\sigma}{Q_m}\delta(\omega-\omega_{mn\sigma})\right].
\end{aligned}
\tag{10}
$$

We have introduced the following notation:

$$Q_m=(\varepsilon-1)(S_m+\sigma)+2\varepsilon S_m\left(R_m^2-\frac{m^2}{x^2}\right)+[\varepsilon(S_m-\sigma)+S_m+\sigma]\left(F_m^2-\frac{m^2}{y^2}\right),$$

$$F_m(y)=\frac{K_m'(y)}{K_m(y)},\quad S_m(y)=\left\{1+\left[\frac{2\beta\beta_T}{\beta^2-\beta_T^2}\frac{m}{yF_m(y)}\right]^2\right\}^{1/2},\tag{11}$$

$$R_m(y)=\left(\frac{\beta^2-\beta_T^2}{1-\beta^2}\right)^{1/2}\frac{\beta_T}{2}[(\varepsilon-1)\sigma S_m(y)-(\varepsilon+1)]F_m(y),$$

$$x=\frac{\omega a}{v}(\varepsilon\beta^2-1),\quad y=\frac{\omega a}{v}(1-\beta^2)^{1/2},$$

where $\beta_T=\varepsilon^{-1/2}$ is the electron Cherenkov-threshold velocity divided by the speed of light, and $\delta(\omega-\omega_{mn\sigma})$ is the Dirac delta-function, which should be replaced, owing to the finite cylinder length, with the function

$$s(\omega-\omega_{mn\sigma})=\frac{L}{2\pi v}\left(\frac{\sin\psi}{\psi}\right)^2,\quad \psi=\frac{L}{2v}(\omega-\omega_{mn\sigma}).\tag{12}$$

The dispersion relation (9) in this notation becomes $J_m'(x)/J_m(x)=R_m(y)$.

The first term on the right-hand side of Eq. (10) corresponds to the spectrum of axisymmetrical $(m=0)$ modes. In accordance with Eq. (9), the axisymmetrical modes are classified as TM_{0n} and TE_{0n} modes with purely transverse magnetic and electric fields, respectively, and only TM_{0n} modes corresponding to $\sigma=1$ contribute to radiation. The dispersion equation for the TM_{0n} modes can be expressed as

$$(1-\beta^2)^{1/2}J_1(x)K_0(y)+\beta_T(\beta^2-\beta_T^2)^{1/2}J_0(x)K_1(y)=0.\tag{13}$$

The other modes with $m\neq 0$ are hybrid, i.e., neither the electric nor magnetic field is purely transverse in them, which is the case in a conventional waveguide, therefore, the modes with both $\sigma=1$ and $\sigma=-1$ contribute, generally speaking, to the radiation. Numerical calculations, however, indicate that the Cherenkov radiation intensity of the $(m,n,1)$ modes is considerably higher than that of the $(m,n,-1)$ modes because the modes with $\sigma=-1$ have a nearly transverse polarization also in the case of $m\neq 0$.

In the case when the dispersion of the dielectric permittivity $\varepsilon(\omega)$ can be neglected and the electron energy is much higher than the threshold value, the approximate solution of Eq. (13) has the form

$$\omega_{0n1}=\frac{v}{a}\frac{\alpha_n}{\sqrt{\varepsilon\beta^2-1}},\tag{14}$$

where the α_n are the zeros of the Bessel function $J_0(\alpha_n)$, which for sufficiently large n are approximately located at $\pi(n-1/4)$. In this case the spectral distribution of radiation energy among the modes with $m=0$ takes the simpler form:

$$\frac{dW_0}{d\omega}=\frac{4e^2}{a^2}\frac{\gamma_T^2}{\gamma^2}\sum_{n=1}^{\infty}\frac{K_0^2(y\rho_0/a)}{K_1^2(y)}s(\omega-\omega_{0n1}),\tag{15}$$

$$\gamma=(1-\beta^2)^{-1/2},\quad \gamma_T=(1-\beta_T^2)^{-1/2},\quad \gamma\gg\gamma_T.$$

The radiation intensity versus electron energy for the most intense modes $(\sigma=1)$ calculated from Eq. (9) for quartz in the visible spectral range (ε=2.3) is shown in Fig. 1 by dashed curves (the right-hand ordinate axis), where the numbers are the indices (m,n). The wavelengths $\lambda_{mn\sigma}$ for almost all modes tend to certain limits as the electron energy increases. The only exception is the mode with $m=1$ and $n=1$, whose wavelength grows logarithmically with the electron energy.

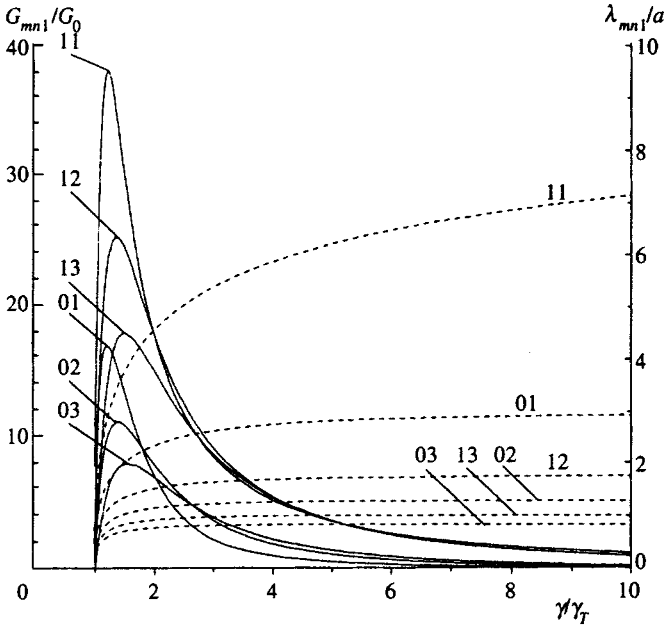


FIG. 1. Mode wavelengths λ_{mn1} (dashed curves) and respective gains G_{mn1} [in units of $G_0=10^{-3}(i/i_0)(L/a)^3$] versus the electron Lorentz factor γ for several modes in a dielectric cylinder ($\varepsilon=2.3$) of radius a. The numbers at the curves show the indices (mn). The most intense modes with $\sigma=1$ are shown.

4. CHERENKOV RADIATION IN CAPILLARY TUBES

Another case which is of interest from the viewpoint of generating induced Cherenkov radiation is an electron beam traveling inside a hollow dielectric cylinder (capillary tube). Let a be the inside radius of the capillary tube, b its outside radius, $\varepsilon_2(\omega)$ the dielectric permittivity of the capillary tube material, and $\varepsilon_1(\omega)$ the permittivity of the environment. The spectrum of spontaneous Cherenkov radiation is calculated using Eq. (3) and the technique described in Sec. 2. In the regions $\rho<a$ and $\rho>b$, the field components are calculated in a form similar to Eq. (6), and in the region $a<\rho<b$ as a linear combination of the functions $I_m(\kappa_2\rho)$ and $K_m(\kappa_2\rho)$ with unknown coefficients. The continuity conditions on the tube boundaries yield a linear equation system for eight unknown coefficients. In the case of arbitrary azimuthal indices m, the solution is rather lengthy, so we limit our analysis to the case of axisymmetrical modes with $m=0$. Note that these are the only excited modes if an electron travels along

the tube axis $\rho_0=0$. By solving the linear algebraic equation system, we obtain the following expression for the longitudinal field component $E_z(\rho_0,0,\omega/v,\omega)$ in the region $\rho_0<a$:

$$L^{-1}E_z(\rho_0<a)=\frac{2ie}{\omega a}\frac{I_0(\kappa_1\rho_0)}{I_0(\kappa_1 a)}\left[\frac{K_0(\kappa_2 a)a_{21}-I_0(\kappa_2 a)a_{22}}{a_{11}a_{22}-a_{12}a_{21}}+\frac{\kappa_1^2 a}{\varepsilon_1}K_0(\kappa_1 a)\right], \tag{16}$$

where

$$a_{11}=\frac{\varepsilon_1}{\kappa_1}I_1(\kappa_1 a)I_0(\kappa_2 a)-\frac{\varepsilon_2}{\kappa_2}I_0(\kappa_1 a)I_1(\kappa_2 a),$$

$$a_{12}=\frac{\varepsilon_1}{\kappa_1}I_1(\kappa_1 a)K_0(\kappa_2 a)+\frac{\varepsilon_2}{\kappa_2}I_0(\kappa_1 a)K_1(\kappa_2 a),$$

$$a_{21}=\frac{\varepsilon_1}{\kappa_1}K_1(\kappa_1 b)I_0(\kappa_2 b)+\frac{\varepsilon_2}{\kappa_2}K_0(\kappa_1 b)I_1(\kappa_2 b),$$

$$a_{22}=\frac{\varepsilon_1}{\kappa_1}K_1(\kappa_1 b)K_0(\kappa_2 b)-\frac{\varepsilon_2}{\kappa_2}K_0(\kappa_1 b)K_1(\kappa_2 b).$$

Let us next consider the case when the tube material is a relatively transparent dielectric ($\operatorname{Im}\varepsilon_2=0$, $\operatorname{Re}\varepsilon_2=\varepsilon>0$), and the tube is in vacuum ($\varepsilon_1\equiv 1$). We set $k=-i\kappa_2$ and use the relation

$$K_0(\kappa_2 a)=(-i\pi/2)[J_0(ka)-iY_0(ka)],$$

where Y_0 is the Neumann function. Then the integration of Eq. (3) yields the following expression for the spectral distribution of the radiated energy of the modes with $m=0$:

$$\frac{dW_0}{d\omega}=\frac{4e^2L}{a^2}\sum_{n=1}^{\infty}I_0^2(\xi\rho_0)\left\{\left[\frac{4\varepsilon}{(\pi kaf_0)^2}-(\varepsilon-1)\right]I_0^2(\xi a)-\frac{\gamma^2}{\gamma_T^2}I_1^2(\xi a)\right\}^{-1}s(\omega-\omega_{0n}). \tag{17}$$

Here

$$f_0^2=\frac{[(\gamma^2-\gamma_t^2)^{1/2}K_1(\xi b)p_0-\varepsilon^{1/2}\gamma_T K_0(\xi b)q_0]^2}{\gamma^2K_1^2(\xi b)+\varepsilon K_0^2(\xi b)},$$

$$k=\frac{\omega}{v}(\varepsilon\beta^2-1)^{1/2},\quad \xi=\frac{\omega}{v}(1-\beta^2)^{1/2},$$

$$p_0=J_0(ka)Y_0(kb)-J_0(kb)Y_0(ka),\quad q_0=\frac{\partial p_0}{\partial kb},$$

$$r_0=\frac{\partial p_0}{\partial ka},\quad s_0=\frac{\partial q_0}{\partial ka},$$

and ω_{0n} are the roots of the dispersion equation. In deriving Eq. (17), we have used the relationship[13] $p_0s_0-q_0r_0=4/(\pi^2k^2ab)$. The dispersion relation is derived using the condition that the denominator in Eq. (16) should vanish:

$$\frac{I_1(\xi a)}{I_0(\xi a)}+\frac{\varepsilon\xi}{k}\frac{kK_1(\xi b)r_0-\varepsilon\xi K_0(\xi b)s_0}{kK_1(\xi b)p_0-\varepsilon\xi K_0(\xi b)q_0}=0. \tag{18}$$

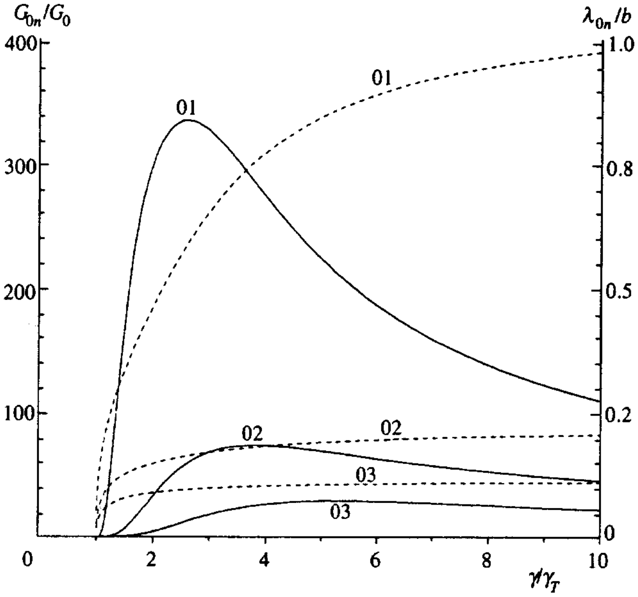


FIG. 2. Radiation wavelengths λ_{0n} (dashed curves) and gain G_{0n} in units of $G_0=10^{-3}(i/i_0)(L/b)^3$ versus electron Lorentz factor γ for several axisymmetrical modes in a capillary tube with aspect ratio $a/b=0.9$ and $\varepsilon=2.3$. The numbers at the curves are the indices (m,n) of respective modes.

When the tube radius is much larger than the wavelength ($ka\gg 1$), the resulting spectrum can be expressed in a simpler form:

$$\frac{dW_0}{d\omega}=\frac{4e^2}{a^2}\sum_{n=1}^{\infty}I_0^2(\xi\rho_0)\left\{(1+\varepsilon\delta)I_0^2(\xi a)+\left[\left(\left(\frac{\gamma}{\gamma_T}\right)^2-1\right)\delta-1\right]I_1^2(\xi a)\right\}^{-1}s(\omega-\omega_{0n}), \tag{19}$$

where $\delta=(b-a)/a$, and the dispersion relation is also simplified:

$$\tan[k(b-a)]=\frac{\varepsilon\xi}{k}\left[\frac{K_0(\xi b)}{K_1(\xi b)}+\frac{I_0(\xi a)}{I_1(\xi a)}\right].$$

Since the condition $\xi\ll k$ holds in this case, the mode frequencies for $n\neq 1$ are determined in the first approximation by the equation

$$\omega_{0n}=\pi(n-1)v/[b-a)(\varepsilon\beta^2-1)^{1/2}]. \tag{20}$$

The wavelengths of the excited modes as functions of the electron Lorentz factor γ calculated by Eq. (18) are shown by dashed lines in Fig. 2 for a capillary tube with $\varepsilon=2.3$ and the ratio $a/b=0.9$. At all electron energies much larger than the threshold ($\gamma\gg\gamma_T$), the wavelengths of all modes except (0, 1) tend to certain limits. They are determined by the ratio between the inner and outer radii, a/b, which is illustrated by the curves in Fig. 3 for $\gamma=10\gamma_T$. One can see that in thin-wall capillary tubes ($b-a\ll b$) the mode wavelength is determined by the wall thickness, rather than by the tube radius.

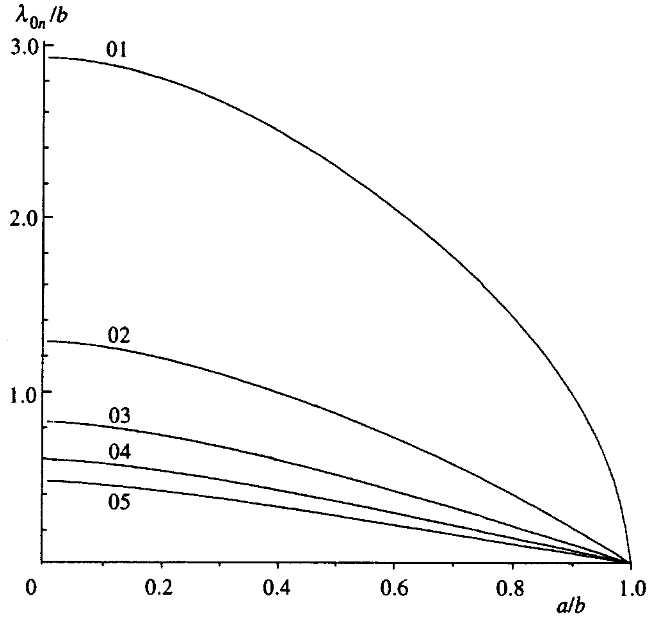


FIG. 3. Wavelengths of modes with indices $(0,n)$ in a capillary at $\gamma=10\gamma_T$ versus the ratio between capillary radii a/b.

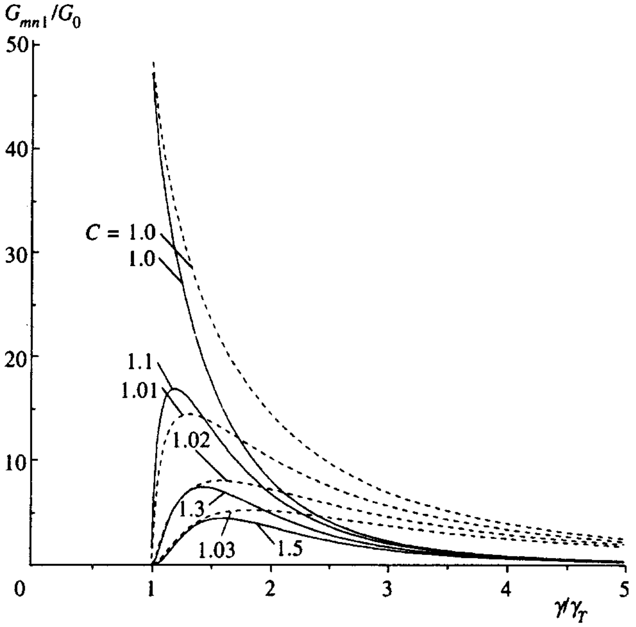


FIG. 4. The gain G_{mn1} at the frequencies of the modes with $m=0$, $n=1$, $\sigma=1$ (solid curves) and $m=0$, $n=15$, $\sigma=1$ (dashed curves) in a cylinder with $\varepsilon=2.3$ and radius a for a hollow electron beam with inner radius $r=Ca$ and outer radius $R=2a$ as a function of the Lorentz factor γ. The coefficients C are shown by numbers at the curves.

5. INDUCED CHERENKOV RADIATION

The optical gain due to induced radiation at wavelengths of various modes can easily be calculated in the limit of low optical gain and cold electron beam, when the nonlinear effects of interaction between the radiation and electrons are negligible, and the spectral width of the radiation line is determined by the length over which an electron interacts with the wave field. Let N be the initial number of photons in a certain mode (m,n,σ) between the planes $z=0$ and $z=L$. In particular, these may be photons generated by an external source. The change ΔN in the photon number is caused by both induced radiation of photons by electrons at the corresponding mode frequency and induced absorption of photons by beam electrons. The probabilities of these processes are related to the spontaneous radiation probability by the well-known relation $w_{\text{ind}}=w_{\text{cap}}=Nw_{\text{sp}}$. The parameter $\Delta N=w_{\text{ind}}-w_{\text{cap}}$ is nonzero if the recoil due to emission or absorption of a photon is taken into account. As a result, we have

$$\Delta N=N\frac{\partial w_{\text{sp}}}{\partial\psi}\Delta\psi, \tag{21}$$

where the phase ψ is determined by Eq. (12), and its change $\Delta\psi$ by the mode frequency shift due to recoil.

The probability w_{sp} of spontaneous radiation is obtained by dividing the classical radiation energy $dW_{mn\sigma}/d\omega$ of the corresponding mode calculated above by the photon energy $\hbar\omega$. As for the small shift $\Delta\psi$, it can be calculated using conservation of momentum and energy in the process of photon emission to a certain mode (m,n,σ). Let the electron have in the initial state energy E, longitudinal momentum p_z, and zero projection on the xy-plane. The corresponding parameters in the final state are E', p_z', and $p_\perp'$. The total values of these parameters for the electron–photon system in the process of radiation (absorption) are constant, i.e.,

$$E'=E\mp\hbar\omega,\quad p_z'=p_z\mp\hbar k_z,\quad p'^2_\perp=(\hbar\kappa)^2, \tag{22}$$

where $(\hbar\kappa)^2=(\hbar\omega/c)^2\varepsilon-(\hbar k_z)^2$ is the square of the photon transverse momentum in the material [see Eq. (5)]. Since the relationships $E^2-p_z^2=mc^2=E'^2-(p'^2+p'^2_\perp)$, $k_z^2+\kappa^2=\varepsilon\omega^2$, and $p_z=Ev$ also hold, we derive from Eq. (22) the condition of Cherenkov radiation (absorption) of various modes with due account of the quantum recoil effect:

$$(\omega-k_zv)\pm\frac{\hbar\omega^2}{2E}(1-\varepsilon)=0.$$

Hence

$$\Delta\psi=\hbar\omega^2(\varepsilon-1)L/(2Ev). \tag{23}$$

Then the gain G is determined using the equation

$$G=\frac{1}{e}\int_S\frac{\Delta N}{N}j(\rho_0,\varphi_0)dS, \tag{24}$$

where $j(\rho_0,\varphi_0)$ is the electron current density as a function of the distance ρ_0 from the cylinder axis and of the azimuthal angle φ_0, and integration (24) is performed over the electron beam cross section S. Using Eqs. (21) and (23), we can express the gain at the frequency of the (m,n,σ) mode as

$$G_{mn\sigma}=\frac{(\varepsilon-1)\omega L}{2eEv}\int_S j(\rho_0,\varphi_0)\frac{\partial}{\partial\psi}\frac{dW_{mn\sigma}}{d\omega}dS, \tag{25}$$

where $dW_{mn\sigma}/d\omega$ is the spectral distribution of spontaneous Cherenkov radiation in the (m,n,σ) mode. Note that the expression for the gain in Eq. (25) does not contain the Planck constant and can be derived from a purely classical theory.

In the case of a solid dielectric cylinder, we derive from Eq. (10) and (25) the gain for the mode (m,n,σ) in the form

$$G_{mn\sigma} = \frac{2i}{i_0}\,\frac{L^3(\varepsilon-1)}{a^2\lambda_{mn\sigma}\gamma\beta^2}\,\frac{\langle K_m^2(y\rho_0/a)\rangle}{K_m^2(y)}\,\frac{(S_m+\sigma)(2-\delta_{m0})}{Q_m} \times\frac{d}{d\psi}\,\frac{\sin^2\psi}{\psi^2}. \qquad (26)$$

Here we have written $i_0=mc^3/e$, i is the total electron current, $y=2\pi a/(\lambda_{mn\sigma}\gamma\beta)$, $\lambda_{mn\sigma}=2\pi c/\omega_{mn\sigma}$, δ_{m0} is the Kronecker symbol, and angular brackets denote averaging over the beam cross section. In particular, if the electrons are distributed uniformly over a ring with inner radius $r\leqslant a$ and outer radius R, we obtain, using integration formulas,[14]

$$\left\langle K_m^2\left(\frac{y\rho_0}{a}\right)\right\rangle = \frac{\rho_0^2}{R^2-r^2}\left[K_m^2\left(\frac{y\rho_0}{a}\right) - K_{m-1}\left(\frac{y\rho_0}{a}\right)K_{m+1}\left(\frac{y\rho_0}{a}\right)\right]\Bigg|_r^R. \qquad (27)$$

For axisymmetrical modes $(m=0)$, the gain determined by Eq. (26) is expressed by a simpler formula:

$$G_{0n\sigma} = \frac{2i}{i_0}\,\frac{L^3}{a^2\lambda_{0n\sigma}\gamma\beta^2}\,\frac{\varepsilon\langle K_m^2(y\rho_0/a)\rangle}{\gamma^2K_1^2(y)+\varepsilon K_0^2(y)}\,\frac{d}{d\psi}\,\frac{\sin^2\psi}{\psi^2}.$$

Since at large values of the argument the modified Bessel functions in Eq. (27) contain an exponentially decaying factor, most of the radiation at the wavelength $\lambda_{mn\sigma}$ is due to electrons moving at distances $\Delta=\rho_0-a$ from the cylinder surface within $\Delta_c=\lambda_{mn\sigma}\beta\gamma/4\pi$. Thus, the gap Δ between the electron beam and the dielectric surface determines the minimum electron Lorentz factor $\gamma_c=4\pi\Delta/\lambda_{mn\sigma}\beta$ needed for efficient amplification of radiation at the wavelength $\lambda_{mn\sigma}$. The gain described by Eq. (26) at $\psi\approx-1.3$, corresponding to the maximum of $G_{mn\sigma}$, for the most intense modes with $\sigma=1$ as a function of the ratio between γ and its threshold value γ_T is plotted in Fig. 1 for the case of $\varepsilon=2.3$ (visible band in quartz) and a beam with the parameters $r=1.1a$, $R=2a$ [see Eq. (27)]. In this case the minimum gap $\Delta=r-a$ is $0.1a$. The gain is measured in the units $G_0=10^{-3}(i/i_0)(L/a)^3$. One can see that γ corresponding to the maximum gain increases with decreasing wavelength (increasing mode index n), whereas the maximum gain itself drops. Similar effects can be observed when the mode indices (m,n) are fixed and the gap width $\Delta=r-a$ is varied (Fig. 4).

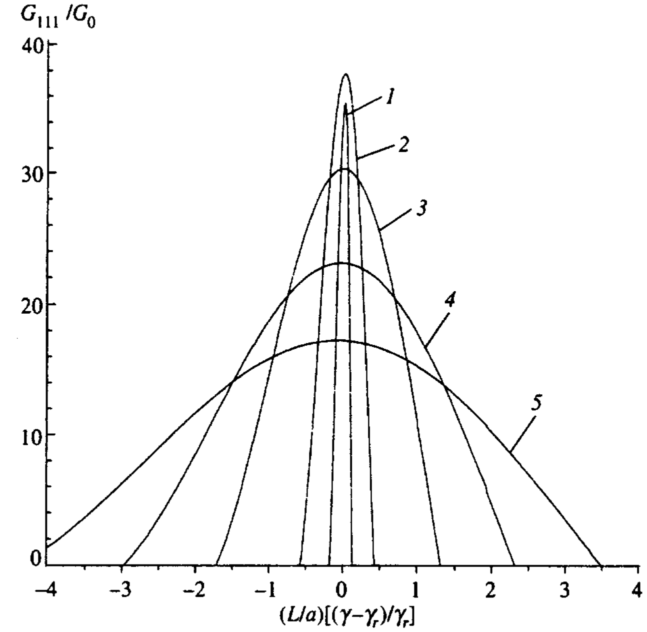


FIG. 5. The gain G_{mn1} for the mode 111 versus the deviation from the resonant electron Lorentz factor γ_r. Curve *1* corresponds to $\gamma_r/\gamma_T=1.10$; (*2*) 1.21; (*3*) 1.46; (*4*) 1.70; (*5*) 2.00.

In the case of a hollow dielectric cylinder (capillary tube), the gain G_{0n} at frequencies of harmonics with $m=0$ can be expressed as

$$G_{0n} = \frac{2i}{i_0}\,\frac{L^3\varepsilon}{a(b-a)\lambda_{0n}\gamma^3\beta^2}\,\frac{\langle I_0^2(\xi\rho_0)\rangle}{I_1^2(\xi a)}\,\frac{d}{d\psi}\,\frac{\sin^2\psi}{\psi^2}, \qquad (28)$$

where $\xi=2\pi/\lambda_{0n}\beta\gamma$. If a cylindrical electron beam has a uniform density in the region $0\leqslant\rho_0\leqslant R$, we have[14]

$$\langle I_0^2(\xi\rho_0)\rangle = I_0^2(\xi R) - I_1^2(\xi R). \qquad (29)$$

Calculations of the gain by Eq. (28) for the case $\varepsilon=2.3$ and $R=0.9a$, when the gap between the beam and the tube wall is 0.1 of its inside radius a, are given in Fig. 2. These results are similar to those for the case of a solid cylinder discussed above (Fig. 1).

6. CONCLUDING REMARKS

In deriving Eqs. (26) and (28) for the gain, we assumed that the spontaneous radiation line had a natural width due to the finite time of interaction between the dielectric and an electron. The resulting expressions are also valid in the case of inhomogeneous broadening due to the spread of the electron energy and angle distributions, if the inhomogeneous broadening is smaller than the natural line width. By varying the phase angle parameters ψ, in Eq. (12) and assuming $\delta\psi<\pi$, we obtain the following limitations:

$$\Delta\theta^2<(\varepsilon\beta^2-1)\lambda\beta/L, \qquad (30)$$

$$\Delta\gamma/\gamma^3<(\varepsilon\beta^2-1)\beta^3\lambda/L, \qquad (31)$$

where $\Delta\theta^2$ is the mean-square spread of the electron propagation angle around the dielectric axis and $\Delta\gamma mc^2$ is the spread in the electron energy. The conditions (30) and (31) were obtained assuming that the electron energy was notably different from the threshold value and the mode frequencies were determined by the approximate expressions (14) and (20).

The condition (31) can be also used in estimating the maximum efficiency of conversion of the electron energy to that of induced Cherenkov radiation if the entire spread $\Delta\gamma$ is ascribed to electron energy radiation losses. The efficiency is fairly high, especially at high electron energy $\gamma \gg \gamma_T$, but, according to curves of Figs. 1 and 2, the gain drops with the ratio γ/γ_T. In the range of electron energies close to the threshold value, where the gain is maximum, the simple estimates by Eqs. (30) and (31) do not apply. For this case, Fig. 5 shows curves of the gain as a function of the electron energy deviation from the resonance $E_r = mc^2\gamma_r$ corresponding to the maximum gain. The curves were plotted for the (111) mode in a solid cylinder and a set of gradually increasing values of $\gamma_{\rm res}$ (and corresponding wavelengths; see Fig. 1). The peak widths are proportional to the efficiency for conversion of the electron energy to induced radiation.

Induced Cherenkov radiation in a cylindrical dielectric can be used to amplify electromagnetic waves over a wide wavelength range from the microwave through the optical band. For example, according to our estimates, the gain G_{01} is about 5% for wavelength $\lambda = 190\ \mu$m, electron energy $E = 1$ MeV, and the following capillary tube parameters: $a = 100\ \mu$m, $b = 150\ \mu$m, $L = 5$ cm, and $i = 70\ \mu$A. The angular divergence of the electron beam in this case should be less than 10^{-3} rad. In this case, waves of the visible band can be amplified only using modes with large refraction indices, $n \gtrsim 10$, whereas in this case, as can be seen in Fig. 4, the gain is highly sensitive to the gap width between the beam and the dielectric wall.

The work was supported by the Russian Fund for Fundamental Research (project No. 95-2-04457a).

[1] V. M. Arutyunyan and S. G. Oganesyan, JETP Lett. **34**, 515 (1981).

[2] W. Becker and J. K. McIver, Phys. Rev. A **25**, 956 (1982).

[3] V. A. Bazylev, N. K. Zhevago, and M. A. Kumakhov, Dokl. Akad. Nauk SSSR, Fizika **263**, 855 (1982) [Sov. Phys. Doklady **27**, 306 (1982)].

[4] J. Walsh, B. Johnson, G. Dattoli, and A. Renieri, Phys. Rev. Lett. **53**, 779 (1984).

[5] N. K. Zhevago, Nucl. Instr. Meth. A **331**, 584 (1993).

[6] E. P. Garate and J. E. Walsh, IEEE Trans. Plasma Sci. **PS-13**, 524 (1985).

[7] E. P. Garate, J. E. Walsh, C. Shanghnessy, and B. Johnson, Nucl. Instr. Meth. A **259**, 125 (1987).

[8] E. P. Garate, H. Kosai, K. Evans, and A. Fisher, Appl. Phys. Lett. **56**, 1092 (1990).

[9] Guigyan Wang, Shanfu Yu, Peijue Xun *et al.*, Appl. Phys. Lett. **59**, 2378 (1991).

[10] L. S. Bogdankevich and B. M. Bolotovskiĭ, Zh. Éksp. Teor. Fiz. **32**, 1421 (1957) [Sov. Phys. JETP **5**, 1157 (1957)].

[11] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media*, Pergamon Press (1984).

[12] N. K. Zhevago and V. I. Glebov, Zh. Éksp. Teor. Fiz. **98**, 278 (1990) [Sov. Phys. JETP **71**, 154 (1990)].

[13] M. Abramowitz and I. Stigun, *Handbook of Mathematical Functions*, NBS, Applied Mathematics Series 55 (1964).

[14] I. S. Gradshteĭn and I. M. Ryzhik, *Tables of Integrals, Sums, Series, and Products*, Academic, New York (1980).

Translation provided by the Russian Editorial office.

Helium ionization with excitation of the atom accompanied by absorption and scattering of high-energy photons

M. Ya. Amus'ya and A. I. Mikhaĭlov

B. P. Konstantinov St. Petersburg Institute of Nuclear Physics, 188350 Gatchina, Leningrad Region, Russia

(Submitted 16 September 1996)
Zh. Éksp. Teor. Fiz. **111**, 862–870 (March 1997)

We calculate analytically the cross sections for ionization of the helium atom with absorption and scattering of high-energy photons. The electrons are assumed to be moving in the Coulomb field of the nucleus. The electron–electron interaction is taken into account in the first order of perturbation theory. The high-frequency limits for the ratios of these cross sections to the single-electron ionization cross sections are obtained for s-excitations in absorption and for s- and p-excitations in scattering. © *1997 American Institute of Physics.* [S1063-7761(97)00803-2]

1. Ionization with excitation of the atom accompanied by photon absorption or scattering is determined entirely by the electron–electron interaction. The helium atom is the simplest system in which such a process can take place. While two-electron ionization has been thoroughly studied both theoretically and experimentally, the process in which the removal of one atomic electron is accompanied by excitation of the other electron belonging to the same atom has been studied much less. However, this process contributes to the formation of singly charged ions and must be taken into account when the theoretical and experimental values for the ratio of the cross sections of double and single ionizations are compared.

We examine photons whose energy ω is in the interval $I \ll \omega \ll m$ (here I is the ionization energy of the atom, m is the electron mass, and we employ the relativistic system of units, in which $\hbar = c = 1$). When the photon energy is limited by the inequalities $I < \omega < \eta$, where $\eta = m\alpha Z$ is the average momentum of a bound electron ($\alpha = 1/137$, and Z is the atomic number), photoabsorption has a much higher probability than inelastic photon scattering. The probabilities of these two processes become comparable at $\omega \sim \eta$, where in the energy range $\omega \gg \eta$ photon scattering (the Compton effect) dominates. For the helium atom the threshold frequency satisfies $\omega_0 \approx \eta \approx 7$ keV.

Similarly, ionization with excitation of the atom by photons occurs primarily through photoabsorption for $I < \omega < \eta$ and through scattering for $\omega > \eta$.

2. We start with ionization with excitation of the helium atom in photoabsorption. The process has been studied in Refs. 1–3. Brown[1] and Dalgarno and Sadeghpour[2] did numerical calculations involving variational wave functions of the initial state and Coulomb wave functions of the final electrons. An analytic calculation was done in Ref. 3, but the final formula contains an error, which is corrected in the present paper. We also derive an analytical formula linking the cross sections for double ionization and single ionization with excitation.

Let us derive a formula for the cross section of the process in the first approximation in the electron–electron interaction, using Coulomb wave functions as the zeroth approximation. By Ψ_1 and Ψ_2 we denote the wave functions of the electrons in the final state. If Ψ_2 refers to the bound state, at high photon frequencies $\omega \gg I$ the energy of the ejected electron, E_1, is approximately ω and the momentum p_1 is $\sqrt{2m\omega} \gg \omega$, while the wave function Ψ_1 can be taken in the form of a plane wave, since the Coulomb parameter for this electron is small:

$$\xi_1 = \frac{\eta}{p_1} = \sqrt{\frac{I}{\omega}} \ll 1.$$

This is also true for double ionization, since, as shown in Ref. 3, the principal contribution to the cross section of this process is provided by energies ranging from $E_1 \sim \omega$ to $E_2 \sim I$.

The Feynman diagram of the process is depicted in Fig. 1. A line with a solid black dot indicates that the corresponding state is described by a Coulomb function, a wavy line depicts electron–electron interaction, and a dashed line corresponds to a photon. Estimates done in Ref. 3 show that the diagram of Fig. 1 is the leading term in the amplitude of the process. The contributions of other diagrams contain an additional small factor of order I/ω.

Writing the electron–photon interaction operator in the ordinary gradient form $\mathbf{p}\cdot\mathbf{A}/m$, where $\mathbf{A}$ is the photon vector potential and $\mathbf{p}$ is the electron momentum operator, and allowing for the fact that $\mathbf{p}_1 - \mathbf{k} \approx \mathbf{p}_1$ holds (where $\mathbf{k}$ is the photon momentum, $p_1 \gg k = \omega$), we arrive at the following expression for the amplitude of the process:

$$F = \frac{\mathbf{e}\cdot\mathbf{p}_1}{m} \int \langle \mathbf{p}_1 | G(E) | \mathbf{f}_1 \rangle \langle \mathbf{f}_1 + \mathbf{f} | \Psi_{1s} \rangle \times \frac{1}{f^2} \langle \Psi_2 | \mathbf{f}_2 \rangle \langle \mathbf{f}_2 - \mathbf{f} | \Psi_{1s} \rangle \frac{d\mathbf{f}\, d\mathbf{f}_1\, d\mathbf{f}_2}{(2\pi)^9}. \tag{1}$$

Here $E = 2E_{1s} - E_2$ is the energy of the intermediate state, $G(E)$ is the Coulomb Green's function of the electron, and $\mathbf{e}$ is the photon polarization vector.

The integrals with respect to $\mathbf{f}_1$ and $\mathbf{f}_2$ with the Coulomb functions Ψ_{1s} can easily be evaluated:

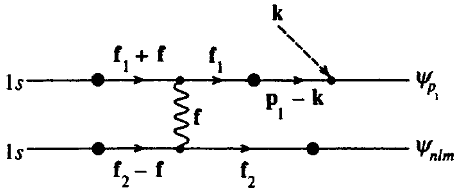


FIG. 1.

$$\int \langle \mathbf{p}_1|G(E)|\mathbf{f}_1\rangle\langle \mathbf{f}_1+\mathbf{f}|\Psi_{1s}\rangle \frac{d\mathbf{f}_1}{(2\pi)^3}=N\left(-\frac{\partial}{\partial\nu}\right) \times\langle \mathbf{p}_1|G(E)V_{i\nu}|-\mathbf{f}\rangle|_{\nu=\eta}, \tag{2}$$

$$\int \langle \Psi_2|\mathbf{f}_2\rangle\langle \mathbf{f}_2-\mathbf{f}|\Psi_{1s}\rangle \frac{d\mathbf{f}_2}{(2\pi)^3} =N\left(-\frac{\partial}{\partial\nu}\right)\langle\Psi_2|V_{i\nu}|\mathbf{f}\rangle|_{\nu=\eta}, \tag{3}$$

where $N=(\eta^3/\pi)^{1/2}$, and $\eta=m\alpha Z$.

Using the Coulomb Green's function in the momentum representation[4] and performing a series of transformations, we find that

$$\langle \mathbf{p}_1|G(E)V_{i\nu}|-\mathbf{f}\rangle=\frac{4ipm\eta}{p_1^4} \times\int_1^\infty dy\left(\frac{y+1}{y-1}\right)^{i\xi}\langle\mathbf{f}|V_{py+i\nu}|0\rangle, \tag{4}$$

$$p=\sqrt{2mE}=i\eta\lambda,\quad \lambda=\sqrt{2+\varepsilon_2},\quad \varepsilon_2=\frac{E_2}{I},$$

$$I=\frac{\eta^2}{2m},\quad i\xi=\frac{i\eta}{p}=\frac{1}{\lambda}.$$

Here $\langle\mathbf{f}|V_{i\alpha}|\mathbf{f}'\rangle$ is the matrix element of the Yukawa potential $\exp(-ar)/r$ in the momentum representation:

$$\langle\mathbf{f}|V_{ia}|\mathbf{f}'\rangle=\frac{4\pi}{(\mathbf{f}-\mathbf{f}')^2+a^2}. \tag{5}$$

After Eqs. (2)–(4) are substituted in (1) we arrive at the following expression for the amplitude of the process:

$$F=BN^2I_y\left(-\frac{\partial}{\partial\mu}\right)\frac{1}{\mu^2}\langle\Psi_2|V_{i\eta}-V_{i(\eta+\mu)}|0\rangle,$$
$$B=-\frac{4\eta^2\lambda}{p_1^3}(\mathbf{e}\cdot\mathbf{n}_1),\quad \mathbf{n}_1=\frac{\mathbf{p}_1}{p_1},\quad \mu=\eta(1+\lambda y), \tag{6}$$

where

$$I_y=\int_1^\infty dy\left(\frac{y+1}{y-1}\right)^{1/\lambda}$$

is an integral operator, and the ket vector $|0\rangle$ stands for a plane-wave state with zero momentum.

Formula (6) can be used in the cases of ionization with excitation and double ionization. In the first process, Ψ_2 is the Coulomb function $\varphi_{n\ell m}$ of a bound state. For it the matrix element in (6) is

$$\langle\varphi_{n\ell m}|V_{i\nu}|0\rangle=4\pi N_n\frac{(\nu-\eta_n)^{n-1}}{(\nu+\eta_n)^{n+1}}\delta_{\ell 0}\delta_{m0},$$
$$N_n=\left(\frac{\eta_n^3}{\pi}\right)^{1/2},\quad \eta_n=\frac{\eta}{n}. \tag{7}$$

In the event of double ionization, Ψ_2 is the Coulomb wave function of the continuous spectrum φ_{p_2}. In the coordinate representation the function $\langle\mathbf{r}|\varphi_{p_2}\rangle=\varphi_{p_2}(\mathbf{r})$ behaves asymptotically like a combination of a plane wave and a converging spherical wave. For the Coulomb function of the continuous spectrum there exists a closed nonrelativistic expression,[5] which yields[6]

$$\langle\varphi_{p_2}|V_{i\nu}|0\rangle=4\pi N_{p_2}\frac{(\nu+ip_2)^{i\xi_2-1}}{(\nu-ip_2)^{i\xi_2+1}},$$

$$N_{p_2}^2=\frac{2\pi\xi_2}{1-\exp(-2\pi\xi_2)},\quad \xi_2=\frac{\eta}{p_2}. \tag{8}$$

The same expression can be obtained if for φ_{p_2} we take a partial-wave expansion. As in (7), only the wave with $\ell=0$ contributes to the matrix element (8). Thus, in our approximation, the transition of a second electron is possible only in the s-state of the discrete and continuous spectra.

Comparing (7) and (8), we conclude that

$$\langle\varphi_{p_2}|V_{i\nu}|0\rangle=\left.\left(\frac{N_{p_2}}{N_n}\langle\varphi_{ns}|V_{i\nu}|0\rangle\right)\right|_{n\to i\xi_2}, \tag{9}$$

where $\varphi_{ns}\equiv\varphi_{n10}$.

Plugging (6) into (7), we arrive at the following expression for the amplitude of photoionization accompanied by excitation of the ion to a state with principal quantum number n:

$$F^+(n)=F^+(ns)=8\eta^{-2}N_nBI_y\chi(y)(1+\lambda y)^{-3}, \tag{10}$$

$$\chi(y)=\varphi(1,1)-\varphi(x,1)-(x-1)^2\varphi(x,2), \tag{11}$$

$$x=2+\lambda y,\quad \lambda=\sqrt{2-n^{-2}},$$

$$\varphi(q,k)=\frac{(q-n^{-1})^{n-k}}{(q+n^{-1})^{n+k}}. \tag{12}$$

Replacing N_n with N_{p_2} and n with $i\xi_2$ in (10), we arrive at an expression for the double ionization amplitude F^{++}.

The differential cross sections of the process involving photon absorption are

$$d\sigma_a^+(n)=(4\pi\alpha)^3\frac{1}{\omega}\overline{|F^+(n)|^2}\frac{d\mathbf{p}_1}{(2\pi)^2} \times\delta(\omega+2E_{1s}-E_1-E_{ns}), \tag{13}$$

$$d\sigma_a^{++}=(4\pi\alpha)^3\frac{1}{\omega}\overline{|F^{++}|^2}\frac{d\mathbf{p}_1d\mathbf{p}_2}{(2\pi)^5} \times\delta(\omega+2E_{1s}-E_1-E_2). \tag{14}$$

The bar over $|F|^2$ stands for an average over photon polarizations. Integrating with respect to $\mathbf{p}_1$ (and with respect to

TABLE I. Values of $\beta_a^+(n)$ in %.

	Ref. 2		Present work	
n	He	Li$^+$	He	Li$^+$
2	4.79	1.60	2.31	1.03
3	0.59	0.25	0.43	0.19
4	0.19	0.08	0.16	0.071
5	0.09	0.04	0.076	0.034
6	0.05	0.02	0.043	0.019

Ω_2 in (14)), we find the total cross section $\sigma_a^+(n)$ for single ionization and the energy distribution $d\sigma_a^{++}/d\varepsilon_2$ for double ionization:

$$\beta_a^+(n)=\frac{\sigma_a^+(n)}{\sigma_a^+}=\frac{Q(n)}{Z^2}, \tag{15}$$

$$\beta_a^{++}(\varepsilon_2)=\frac{1}{\sigma_a^+}\frac{d\sigma_a^{++}}{d\varepsilon_2}=\frac{R_a(\varepsilon_2)}{Z^2}\,. \tag{16}$$

Here Z is the atomic number (we are examining helium and heliumlike atoms), and σ_a^+ is the cross section of an ordinary photoeffect on a helium atom in the Born approximation:

$$\sigma_a^+=\frac{2^8\pi\alpha}{3m\omega}\left(\frac{I}{\omega}\right)^{5/2}, \tag{17}$$

$$Q(n)=\frac{2^8}{n^3}J^2(n), \quad R_a(\varepsilon_2)=2^7J^2(p_2), \tag{18}$$

$$J(n)=\lambda\int_1^\infty dy\left(\frac{y+1}{y-1}\right)^{1/\lambda}\chi(y)(1+\lambda y)^{-3}, \tag{19}$$

$$J(p_2)=J(n\to i\xi_2), \quad \xi_2=\frac{\eta}{p_2}=\frac{1}{\sqrt{\varepsilon_2}}. \tag{20}$$

The inverse is also true, i.e.,

$$J(n)=J(p_2\to i\eta_n), \quad \eta_n=\frac{\eta}{n}. \tag{21}$$

This yields a relationship linking $Q(n)$ and $R_a(\varepsilon_2)$:

$$Q(n)=\frac{2}{n^3}R_a(p_2\to i\eta_n). \tag{22}$$

By calculating $J(n)$ in the limit $n\to\infty$ we can find the asymptotic behavior of $Q(n)$:

$$Q(n\gg 1)\approx\frac{0.336}{n^3}=\frac{2R_a(0)}{n^3}. \tag{23}$$

The graph in Ref. 3 presenting the dependence of R_a on ε_2 can be used to find $R_a(0)$, which is equal to 0.168 and coincides with (23). But from the data listed in Table 1 of Ref. 3 it follows that $Q(n)=0.277/n^3$, which contradicts (23). This forced us to calculate $\beta_a^+(n)$ anew via formula (15) and to tabulate the results in Table I together with results obtained earlier by Dalgarno and Sadeghpour[2] for He and Li$^+$. As Table I clearly shows, the agreement between the numerical results of Ref. 2 and our analytical results improves as n and Z grow.

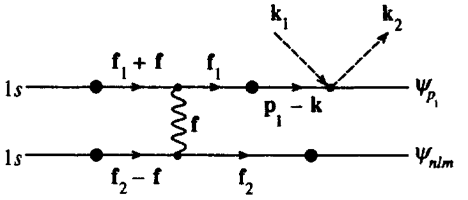


FIG. 2.

3. Now let us discuss inelastic (Compton) photon scattering in the $\eta\ll\omega\ll m$ energy range. Double ionization of the helium atom in such a process was studied in Refs. 7–12. Here we develop the ideas of Ref. 10 and derive formulas for the cross sections of ionization with excitation in first order in the electron–electron interaction (Ref. 8 contains only numerical values of such cross sections obtained in the momentum approximation).

The process is described by the Feynman diagram depicted in Fig. 2. The electron–photon interaction is represented by the operator $\mathbf{A}^2/2m$. As Refs. 9 and 10 show, this diagram plays the leading role in double ionization via Compton scattering. In double ionization the main contribution is provided by the part of the electron energy spectrum where the second electron acquires an energy $E_2\sim I$. The energy of the incident photon (minus the ionization and excitation energies) is distributed between the scattered photon and the ejected electron.

Let E_1 and $\mathbf{p}_1$ be the energy and momentum of the ejected electron, and ω_1 and ω_2 the frequencies of the incident and scattered photons. In Ref. 10 it is shown that the scattered photons predominantly carry away the energy $\omega_2\approx\omega_1-\omega_1^2/m$. Since $E_1=p_1^2/2m\approx\omega_1-\omega_2\approx\omega_1^2/m$ holds, we have $p_1\sim\omega_1$, and the Coulomb parameter satisfies $\xi_1=\eta/p_1\approx\eta/\omega_1\ll 1$. This makes it possible to use a plane wave with momentum $\mathbf{p}_1$ as the wave function of the ejected electron.

Let σ_c^+ be the cross section of ordinary Compton scattering by a helium atom, and let $\sigma_c^+(n\ell)$ be the cross section for inelastic photon scattering by helium in which one electron leaves the atom and the other is transferred to an excited state $n\ell$ of the He$^+$ ion. Equations (11), (29), and (45) of Ref. 10 make it possible to find the ratio of these cross sections:

$$\beta_c^+(n\ell)=\frac{\sigma_c^+(n\ell)}{\sigma_c^+}$$

$$=(4\pi\alpha)^2\left(-\frac{\partial}{\partial E}\right)\sum_m\langle\phi_{n\ell m}|G(E)|\phi_{n\ell m}\rangle, \tag{24}$$

$$\langle\phi_{n\ell m}|=N^2\frac{\partial^2}{\partial\nu_1\partial\nu_2}\int\frac{d\mathbf{f}}{(2\pi)^3}\langle\varphi_{n\ell m}|V_{i\nu_1}|\mathbf{f}\rangle$$

$$\times\frac{1}{f^2}\langle-\mathbf{f}|V_{i\nu_2}, \tag{25}$$

where $N^2=\eta^3/\pi$, and $E=2E_{1s}-E_{n\ell}$. After finding the derivatives we must put $\nu_1=\nu_2=\eta=m\alpha Z$.

As in Ref. 10, we approximate the Coulomb Green's function $G(E)$ by three terms of the Sturm expansion:[13]

$$G(E)\approx M\left\{\frac{|\Psi_{10}\rangle\langle\Psi_{10}|}{\lambda-1}+\frac{|\Psi_{20}\rangle\langle\Psi_{20}|}{\lambda-1/2}+\sum_{\mu}\frac{|\Psi_{21\mu}\rangle\langle\Psi_{21\mu}|}{\lambda-1/2}\right\}$$

$$=G_{10}+G_{20}+G_{21},\tag{26}$$

$$M=-\frac{\eta^2\lambda^2}{\pi\alpha Z},\quad \lambda=\sqrt{\varepsilon},\quad \varepsilon=-\frac{E}{I}=2-n^{-2}.$$

The functions ψ_{ik} in the coordinate representation have the form

$$\begin{aligned}&\langle\mathbf{r}|\Psi_{10}\rangle=e^{-\gamma r},\quad \langle\mathbf{r}|\Psi_{20}\rangle=(1-\gamma r)e^{-\gamma r},\\&\langle\mathbf{r}|\Psi_{21\mu}\rangle=\gamma(\boldsymbol{\varepsilon}_\mu\mathbf{r})e^{-\gamma r},\end{aligned}\tag{27}$$

where $\gamma=\eta\lambda$, and the $\boldsymbol{\varepsilon}_\mu$ are cyclic unit vectors, with $\mu=0,\pm1$.

As a result of inserting (26) into (24) we get

$$\beta_c^+(n\ell)$$

$$=\frac{(4\pi\alpha)^2}{2I\lambda}\frac{\partial}{\partial\lambda}M\sum_m\left(\frac{|\langle\phi_{n\ell m}|\Psi_{10}\rangle|^2}{\lambda-1}\right.$$

$$\left.+\frac{|\langle\phi_{n\ell m}|\Psi_{20}\rangle|^2}{\lambda-1/2}+\sum_{\mu}\frac{|\langle\phi_{n\ell m}|\Psi_{21\mu}\rangle|^2}{\lambda-1/2}\right).\tag{28}$$

Substituting (25) into (28) yields the following matrix elements:

$$\begin{aligned}&\langle-\mathbf{f}|V_{i\eta}|\Psi_{10}\rangle=\langle\mathbf{f}|V_{i\nu}|0\rangle,\quad \nu=\eta+\gamma,\\&\langle-\mathbf{f}|V_{i\eta}|\Psi_{20}\rangle=\left(1+\gamma\frac{\partial}{\partial\nu}\right)\langle\mathbf{f}|V_{i\nu}|0\rangle,\\&\langle-\mathbf{f}|V_{i\eta}|\Psi_{21\mu}\rangle=i\gamma(\boldsymbol{\varepsilon}_\mu\nabla_k)\langle\mathbf{f}|V_{i\nu}|\mathbf{k}\rangle|_{\mathbf{k}\to0}.\end{aligned}\tag{29}$$

Employing (29), we obtain

$$\begin{aligned}&\langle\phi_{n\ell m}|\Psi_{10}\rangle=N^2\left(-\frac{\partial}{\partial\nu}\right)\frac{1}{\nu^2}\langle\varphi_{n\ell m}|V_{i\eta}-V_{i(\eta+\nu)}|0\rangle,\\&\langle\phi_{n\ell m}|\Psi_{20}\rangle=\left(1+\gamma\frac{\partial}{\partial\nu}\right)\langle\phi_{n\ell m}|\Psi_{10}\rangle,\\&\langle\phi_{n\ell m}|\Psi_{21\mu}\rangle=i\gamma N^2\frac{\partial^2}{\partial\nu_1\partial\nu}\frac{\boldsymbol{\varepsilon}_\mu\nabla_k}{\nu}\\&\qquad\times\int_0^1 dt\langle\varphi_{n\ell m}|V_{i\Lambda}|\mathbf{k}t^2\rangle|_{\mathbf{k}\to0},\end{aligned}\tag{30}$$

where $\Lambda=\nu_1+\nu t$. Now we can use Eq. (7) and the definition (12) to obtain

$$\begin{aligned}&\langle\phi_{n\ell m}|\Psi_{i0}\rangle=\delta_{\ell0}\delta_{m0}\frac{8N_n}{\eta^2(\lambda+1)^3}\chi_i,\quad i=1,\ 2,\\&\langle\phi_{n\ell m}|\Psi_{21\mu}\rangle=\delta_{l1}\delta_{m\mu}\frac{32N_n\lambda}{\eta^2(\lambda+1)^3}\chi_3,\end{aligned}\tag{31}$$

$$\chi_1=\varphi(1,1)-\varphi(a,1)-(a-1)^2\varphi(a,2),$$

$$\chi_2=-\frac{2\lambda-1}{\lambda+1}\chi_1+2\lambda(\lambda+1)^2(2\lambda+3)\varphi(a,3),\tag{32}$$

$$\chi_3=-\frac{\chi_1}{(\lambda+1)^2}+\frac{1}{2}(\lambda+1)(2\lambda+3)\varphi(a,3),$$

where $a=\lambda+2$.

TABLE II. Values of Q_{ik} in %.

n	2	3	4	5	6
Q_{10}	6.01	1.05	0.376	0.179	9.93(−2)*
$-Q_{20}$	0.126	2.74(−2)	1.04(−2)	5.09(−3)	2.87(−3)
Q_{21}	0.479	0.115	4.51(−2)	2.24(−2)	1.27(−2)

*Note:** The number in parentheses stands for the power of 10.

As Eqs. (31) show, the second electron can be excited only to the s- and p-states of the helium ion, which is a consequence of our approximation of the Green's function. The probability of a transition to other states can be found if we leave a greater number of terms in the Sturm expansion of the Green's function. But such transitions, as Surić *et al.*[8] argue, have a small probability, and we ignore them here.

We plug (31) into (28) and find the derivative with respect to λ. For transitions to ns- and np-states we have

$$\beta_c^+(ns)=\frac{Q_{10}+Q_{20}}{Z^2},\quad \beta_c^+(np)=\frac{Q_{21}}{Z^2},\tag{33}$$

$$\begin{aligned}&Q_{10}=\frac{2^{10}}{n^3}\frac{\lambda\chi_1^2}{(\lambda+1)^6(\lambda-1)}\left(\frac{1}{\lambda-1}+\frac{6}{\lambda+1}-\frac{2}{\lambda}-\frac{2\chi_1'}{\chi_1}\right),\\&Q_{20}=\frac{2^{10}}{n^3}\frac{\lambda\chi_2^2}{(\lambda+1)^6(\lambda-1/2)}\\&\qquad\times\left(\frac{1}{\lambda-1/2}+\frac{6}{\lambda+1}-\frac{2}{\lambda}-\frac{2\chi_2'}{\chi_2}\right),\\&Q_{21}=\frac{2^{14}}{n^3}\left(1-\frac{1}{n^2}\right)\frac{\lambda^3\chi_3^2}{(\lambda+1)^6(\lambda-1/2)}\\&\qquad\times\left(\frac{1}{\lambda-1/2}+\frac{6}{\lambda+1}-\frac{4}{\lambda}-\frac{2\chi_3'}{\chi_3}\right),\end{aligned}\tag{34}$$

where $\chi_i'=\partial\chi_i/\partial\lambda$. After the derivative with respect to λ is found, we can use the definition $\lambda=\sqrt{2-n^{-2}}$ to simplify the expressions for χ_i and χ_i'.

For $n\gg1$ we have the following asymptotic formulas:

$$Q_{10}=\frac{0.196}{n^3},\quad Q_{20}=-\frac{0.00585}{n^3},\quad Q_{21}=\frac{0.0263}{n^3}.\tag{35}$$

Table II lists the values of Q_{ik} for finite n that were calculated by (34), and Table III lists the results of calculating $\beta_c^+(n\ell)$ for He by (33) and those of Surić *et al.*[8] Note that the results of numerical and analytical calculations for photon scattering (Table III) are in better agreement than those for photoabsorption.

As in the photoeffect, in Compton scattering there exists a simple relationship between the cross sections of double ionization, σ_c^{++}, and of ionization with excitation, $\sigma_c^+(n)$:

TABLE III. Values of $\beta_c^+(n\ell)$ in %.

	Ref. 8		Present work	
n	ns	np	ns	np
2	2.60	0.163	1.47	1.120
3	0.292	0.0316	0.256	0.0288
4	0.094	0.0119	0.091	0.0113
5	0.043	0.0057	0.043	0.0056
6	0.023	0.0032	0.024	0.0032

$$\sigma_c^+(n)=\sigma_c^+(ns)+\sigma_c^+(np)=\frac{2}{n^3}\frac{d\sigma_c^{++}}{d\varepsilon_2}(p_2\rightarrow i\eta_n), \quad (36)$$

or

$$Q_{ik}=\frac{2}{n^3}R_{ik}(p_2\rightarrow i\eta), \quad (37)$$

where the R_{ik} determine the ratio $\beta_c^{++}(\varepsilon_2)$ of the cross sections of double and single ionizations in Compton scattering:

$$\beta_c^{++}(\varepsilon_2)=\frac{d\sigma_c^{++}}{d\varepsilon_2}\frac{1}{\sigma_c}=\frac{R_c(\varepsilon_2)}{Z^2}=\frac{R_{10}+R_{20}+R_{21}}{Z^2}, \quad (38)$$

$$\sigma_c^+=2\sigma_T\left(1-\frac{2\omega_1}{m}\right), \quad \sigma_T=\frac{8}{3}\pi r_e^2,$$

where σ_T is the Thomson cross section, and r_e is the classical electron radius.

The R_{ik} correspond to the same terms G_{ik} of the Sturm expansion of the Coulomb Green's function (26) as the Q_{ik}. In Ref. 10 only the numerical values of $R_{ik}(\varepsilon_2)$ for different ε_2 are given. To arrive at analytical expressions we must simply multiply the Q_{ik} of (34) by $n^3/2$ and then replace n with $i\xi_2$:

$$R_{ik}(\varepsilon_2)=\left(\frac{n^3}{2}Q_{ik}\right)_{n\rightarrow i\xi_2}. \quad (39)$$

The validity of (37) can easily be established for $n\rightarrow\infty$. In this case

$$Q_{ik}(n\gg 1)=\frac{2}{n^3}R_{ik}(0). \quad (40)$$

By taking the values of $R_{ik}(0)$ from Table 2 of Ref. 10 we arrive at the expressions (35) for the Q_{ik}.

This work received partial financial support from the Section of Fundamental Sciences of the U.S. Department of Energy under the contract W-31-109-ENG-38.

[1] R. L. Brown, Phys. Rev. A **1**, 341 (1970)

[2] A. Dalgarno and H. R. Sadeghpour, Phys. Rev. A **46**, R3591 (1992).

[3] M. Ya. Amusia, E. G. Drukarev, V. G. Gorshkov, and M. P. Kazachkov, J. Phys. B **18**, 1248 (1975).

[4] V. G. Gorshkov, Zh. Éksp. Teor. Fiz. **47**, 352 (1964) [Sov. Phys. JETP **20**, 234 (1965)].

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

[6] V. G. Gorshkov, A. I. Mikhailov, and V. S. Polikanov, Nucl. Phys. **55**, 273 (1964).

[7] K. Hino, P. M. Bergstrom, Jr., and J. H. Macek, Phys. Rev. Lett. **72**, 1620 (1994).

[8] T. Surić, K. Pisk, B. A. Logan, and R. H. Pratt, Phys. Rev. Lett. **73**, 790 (1994).

[9] M. Ya. Amusia and A. I. Mikhailov, Phys. Lett. A **199**, 209 (1995).

[10] M. Ya. Amusia and A. I. Mikhailov, J. Phys. B **28**, 1723 (1995).

[11] T. Surić, K. Pisk, and R. H. Pratt, Phys. Lett. A **211**, 289 (1996).

[12] L. Spielberger, O. Jagutzki, B. Krässig, U. Meyer, Kh. Khayyat, V. Mergel, Th. Tschentscher, Th. Buslaps, H. Brauning, R. Dorner, V. Vogt, M. Achler, J. Ullrich, D. S. Ullrich, D. S. Gemmell, and H. Schmidt-Bocking, Phys. Rev. Lett. **76**, 4685 (1996).

[13] S. A. Zapryagaev, N. L. Manakov, and V. G. Pal'chikov, *Theory of One- and Two-Electron Multiply Charged Ions* [in Russian], Energoatomizdat, Moscow (1985).

Translated by Eugene Yankovsky

Comparison of methods for calculating turbulent diffusion coefficients

N. A. Silant'ev

Main Astronomical Observatory, Russian Academy of Sciences, 196140 St. Petersburg, Russia; Instituto Nacional de Astrofisica, Optica y Electronica, Puebla, Mexico

(Submitted 18 September 1996)
Zh. Éksp. Teor. Fiz. **111**, 871–881 (March 1997)

Different methods for calculating the turbulent diffusion coefficient D_T of a passive scalar impurity in an infinite homogeneous isotropic stationary turbulent medium are examined. The values of D_T calculated by these methods are compared for two limiting types of turbulence, viz., turbulence with a δ-function spectrum and turbulence with a Kolmogorov-type spectrum. The temporal dependence of the velocity correlators is assumed to be exponential. It is shown that the most accurate method is based on the use of the solution of the nonlinear equation for the averaged Green's function with consideration of the contribution from the four-point turbulent velocity correlators. A comparison with the results of other methods that are simpler from the mathematical standpoint shows that some of them also permit the calculation of D_T with relatively good accuracy. [S1063-7761(97)00903-7]

1. INTRODUCTION

The description of the transport of a passive scalar impurity (particle concentrations, temperatures) in turbulent media is one of the main problems in the theory of turbulence. In view of the complexity of the exact solution of the problem, efforts are usually confined to the use of the diffusion approximation, which is suitable for large-scale spatiotemporal averaging of an impurity field (for further details, see Refs. 1 and 2). In the diffusion approximation the main problem is the calculation of the turbulent diffusion coefficient D_T. The latter determines the mean impurity flux $[\mathbf{F}(\mathbf{r},t)=-D_T\nabla\langle n\rangle]$ at the fixed point $\mathbf{r}$ of the medium at the time t in the diffusion approximation. As will be shown below, the calculation of D_T is associated with the use of the ordinary single-particle Green's function $G(1,2)$ of the exact transport equation for the impurity field.

We note that the theory of turbulence is also concerned with the somewhat different problem of "relative diffusion," in which interest is focused on the probability $p(R,t)$ of the separation of two initially close liquid particles to a distance R during the time t. The function $p(R,t)$ approximately satisfies a diffusion equation with the diffusion coefficient $K(R,t)$. The classic work by Richardson[3] and Batchelor[4,5] was devoted mainly to the solution of this problem (see also Ref. 6 and 7). The exact determination of $K(R,t)$ requires knowledge of the two-particle Green's function $G(1,2;3,4)$. We stress that the ordinary diffusion coefficient D_T does not coincide with $K(R,t)$.

The problem of calculating D_T is also related directly to long-enduring problems of the theory of turbulence, viz., the problem of closing the hierarchy of equations for the velocity correlators and the problem of calculating the turbulent viscosity. Various methods have been developed to solve these problems.[8,9] The methods for describing turbulent transport in bounded media and for the case of free turbulence in an infinite medium differ significantly. In the present work we consider only the latter case, although the proposed methods can certainly be generalized to turbulent flows in bounded media.

Thus, the problem is formulated as follows: there is an incompressible infinite turbulent medium characterized by an ensemble of assigned velocities $\mathbf{u}(\mathbf{r},t)$. The turbulence is assumed to be homogeneous, isotropic, and stationary, the mean velocity satisfies $\langle\mathbf{u}(\mathbf{r},t)\rangle=0$, and div $\mathbf{u}=0$. The angle brackets denote averaging over the ensemble of the actual turbulent velocities $\mathbf{u}(\mathbf{r},t)$. Because of the homogeneous stationary character of the turbulence, all the averaged quantities depend on the coordinate difference $\mathbf{R}=\mathbf{r}_1-\mathbf{r}_2$ and the time difference $\tau=t_1-t_2$. To be specific, we shall discuss the diffusion of impurity particles with the concentration $n(\mathbf{r},t)$. The molecular diffusion coefficient D_m is usually much smaller than D_T, and in the final formulas we set $D_m=0$.

The numerical values of the turbulent diffusion coefficient D_T are essentially determined by the explicit form of the pair velocity correlator

$$B_{nm}(\mathbf{R},\tau)=\langle u_n(\mathbf{r}_1,t_1)u_m(\mathbf{r}_2,t_2)\rangle, \tag{1}$$

which can also be characterized by the generalized spectrum $E(p,\tau)$:

$$\langle\mathbf{u}(\mathbf{r},t)\mathbf{u}(\mathbf{r},t+\tau)\rangle=\int_0^\infty dp\; E(p,\tau). \tag{2}$$

This is because the most significant contribution to impurity transport is made by the large-scale turbulent fluctuations, which are described well by the two-point correlator (1). The description of small-scale fluctuations clearly requires knowledge of the four-point and higher velocity correlators. The generalized spectrum $E(p,\tau)$ is characterized by the lifetime τ_0, the wave number $p_0=1/R_0$, and the velocity u_0 of the turbulent motions. The turbulent diffusion coefficient D_T depends significantly on these parameters. It is convenient to represent D_T in the dimensionless form

$$D_T=\frac{u_0}{p_0}\,\overline{D_T}\left(\frac{t}{\tau_0},\xi_0,\ldots\right). \tag{3}$$

Here $\xi_0=u_0\tau_0p_0$ is a dimensionless parameter, which is very important for determining D_T. The following estimates apply:[1] $D_T\simeq u_0^2\tau_0/3\equiv(u_0/p_0)\xi_0/3$ for $\xi_0\ll 1$ and $D_T\simeq u_0/p_0$ for frozen turbulence with $\xi_0\gg 1$. Thus, the dimensionless turbulent diffusion coefficient D_T as a function of ξ_0 first increases linearly with $\xi_0/3$ and then tends monotonically to a certain limiting value $D_T(\infty)$. The monotonicity of this increase is dictated by the fact that an increase in ξ_0 corresponds to a longer lifetime for the turbulent fluctuations.

These limiting estimates can have only a qualitative character. For turbulent flows having a broad energy spectrum and a complicated dependence on the time τ, even the choice of the parameters p_0 and τ_0 is uncertain. This choice is more or less unique only for narrow-band spectra of $E(p,\tau)$. The diffusion coefficient D_T depends on the specific form of $E(p,\tau,)$ as a whole, and methods for calculating it that are independent of the formal choice of the parameters p_0 and τ_0 are needed.

In this paper we shall study mainly the stationary value of D_T corresponding to the times $t\gg\tau_0$ (or $t\gg t_0=R_0/u_0$ for frozen turbulence).

In the Lagrangian representation of the velocity field $\mathbf{v}(\mathbf{a},t)$ the exact expression for the turbulent diffusion coefficient has the form[10]

$$D_T(t)=\frac{1}{3}\int_0^t d\tau\langle\mathbf{v}(\mathbf{a},t)\mathbf{v}(\mathbf{a},\tau)\rangle, \tag{4}$$

where $\mathbf{a}$ is the initial position of a liquid or gas particle. This simple formula can be used only in cases in which the ensemble of trajectories of the liquid particles is known (as a result of measurements or a numerical experiment). As we know, the Eulerian representation of the velocity $\mathbf{u}(\mathbf{r},t)$ is used for the most part, and Eq. (4) is practically useless for calculating D_T, since passage from the Eulerian velocities $\mathbf{u}(\mathbf{r},t)$ to the Lagrangian velocities $\mathbf{v}(\mathbf{a},t)$ is a difficult problem that has not yet been solved.

In the Eulerian representation the exact expression for D_T contains averaging of the velocity components $u_n(\mathbf{r},t)$ with the Green's function $G(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)\equiv G(1,2)$:[1,2]

$$D_T(t)=\frac{1}{3}\int d\mathbf{R}\int_0^t d\tau\langle u_n(1)G(1,2)u_n(2)\rangle. \tag{5}$$

Here and below we use the convenient notation $dn\equiv d\mathbf{r}_n dt_n$, $f(n)\equiv f(\mathbf{r}_n,t_n)$, $f(1-2)\equiv f(\mathbf{r}_1-\mathbf{r}_2,t_1-t_2)$, $\mathbf{R}=\mathbf{r}_1-\mathbf{r}_2$, $\tau=t_1-t_2$, as well as the usual convention regarding the summation over repeated vector and tensor indices. The Green's function $G(1,2)$ is a random function, which depends on the velocity field $\mathbf{u}(\mathbf{r},t)$. As is usually done, we represent all the quantities in the form of a sum of the mean value and the fluctuational part:

$$n(\mathbf{r},t)=\langle n\rangle+n'(\mathbf{r},t),\quad G(1,2)=\langle G\rangle+G'(1,2).$$

The evolution of the impurity concentration $n(\mathbf{r},t)$ satisfies the equation (we recall that $\operatorname{div}\mathbf{u}=0$):

$$(\partial/\partial t-D_m\nabla^2)n(\mathbf{r},t)=-(\mathbf{u}(\mathbf{r},t)\nabla)n(\mathbf{r},t). \tag{6}$$

The Green's function $G(1,2)$ is the solution of this equation with a δ-function source $[Q(\mathbf{r},t)=\delta(\mathbf{R})\delta(\tau)]$. Using the Green's function $G_m(R,\tau)$ of this equation without the term containing the velocity $\mathbf{u}(\mathbf{r},t)$, we obtain the integral equation for $G(1,2)$:

$$G(1,2)=G_m(1-2)+\int d3\,G(1-3)(-\mathbf{u}(3)\nabla^{(3)})G(3,2), \tag{7}$$

$$G_m(R,\tau)\equiv G_m(1-2)=H(\tau)(4\pi D_m\tau)^{-3/2}\times\exp(-R^2/4D_m\tau), \tag{8}$$

where $H(\tau)=1$ for $\tau>0$ and $H(\tau)=0$ for $\tau<0$. For turbulent flows in bounded media the Green's functions G_m and $G(1,2)$ satisfy particular boundary conditions, but the general formula (5) is also suitable in these cases. The formal solution of Eq. (6) in terms of the Lagrangian velocity $\mathbf{v}(\mathbf{a},t)$ gives the expression

$$G(1,2)=\delta\left(\mathbf{r}_1-\mathbf{r}_2-\int_{t_2}^{t_1}d\tau\,\mathbf{v}(\mathbf{a},t)\right),$$

whose substitution together with the relations

$$\mathbf{r}_1=\mathbf{a}+\int_0^{t_1}d\tau\,\mathbf{v}(\mathbf{a},\tau)$$

and $\mathbf{u}(\mathbf{r},t)=\mathbf{v}(\mathbf{a},t)$ into the basic expression (5) leads at once to (4).

Thus, the different methods for calculating D_T in the Eulerian representation reduce to a particular choice of the Green's function $G(1,2)$ (or to the method for calculating it). The explicit form of the stochastic Green's function $G(1,2)$ as a functional of the random velocity $\mathbf{u}(\mathbf{r},t)$ is unknown. Therefore, ways must be sought to represent $G(1,2)$ in the form of a series of approximations, whose substitution into (5) would lead to a corresponding series of approximations for D_T. The main question here is the rapid (although only asymptotic for broad spectra) convergence of this series for different values of $\xi_0=u_0\tau_0p_0$.

2. RENORMALIZED EQUATION FOR THE GREEN'S FUNCTION

The simplest representation of the Green's function $G(1,2)$ in the form of a series is confined to the iteration of the basic equation (7). The substitution of this series of iterations into (5) leads to a series of approximations for D_T of the form

$$D_T=\frac{u_0^2\tau_0}{3}\,(a+b\xi_0^2+c\xi_0^4+\ldots). \tag{9}$$

Here the term with ξ_0^2 is specified by the four-point velocity correlator $\mathbf{u}(\mathbf{r},t)$, the term with ξ_0^4 is given by sixth-order correlators, etc. It is seen from (9) that D_T can be calculated in this manner only for $\xi_0^2\ll 1$. For turbulence with a broad Kolmogorov-type spectrum ($\propto p^{-5/3}$), only the first term in the expansion exists, and the terms with $b,c,\ldots$ diverge. The series (9) converges poorly because the expansion into a series was performed in powers of the molecular Green's function G_m, which does not describe the convective character of the transport of an impurity by turbulence.

An equation for $G(1,2)$ in a new renormalized form must clearly be obtained. In this new equation the convective mechanism of impurity transport must already be taken into account in the free term. The possibility of writing such an equation is obvious and is based on the relations[1]

$$G(1,2)=M(1-2)+\int d3\, M(1-3)\times\left[(-\mathbf{u}(3)\nabla^{(3)})G(3,2)-\int d4\, \hat{K}(3-4)G(4,2)\right], \tag{10}$$

$$M(1-2)=G_m(1-2)+\int d3\int d4\, G_m(1-3)\times\hat{K}(3-4)M(4-2). \tag{11}$$

It is easy to verify that the substitution of (11) into (10) leads to the original equation (7). The most satisfactory choice of the kernel $\hat{K}$ is probably the one for which the auxiliary function $M(R,\tau)$ would coincide with the averaged Green's function $\langle G(1,2)\rangle$. In fact, the free term $M(1-2)=\langle G(1,2)\rangle$ describes the convective transport of impurity particles in a turbulent medium. Thus, Eq. (11) for $M(R,\tau)$ should be a closed equation for the averaged Green's function $\langle G(1,2)\rangle$. The stochastic equations (6) and (7) are linear equations. It is, however, easy to see that averaging these equations does not give separate equations for $\langle n(\mathbf{r},t)\rangle$ or $\langle G(1,2)\rangle$, i.e., there will always be terms like $\langle \mathbf{u}(1)\nabla n'(1)\rangle$ or $\langle \mathbf{u}(1)\nabla G'(1,2)\rangle$. The situation is similar to the one which emerges during the derivation of the Reynolds equations for the mean velocity, i.e., the mean values depend on the contribution of the fluctuations, and the latter, in turn, are determined by the distribution of the averaged quantities. The endeavor to write a separate equation for $\langle G(1,2)\rangle$ leads to a hierarchy of equations that are nonlinear in $\langle G(1,2)\rangle$. The procedure for deriving this hierarchy was described in detail in Ref. 1. Thus, the kernel $\hat{K}$ in (11) depends on $\langle G\rangle$ and can be represented in the form of an infinite series of terms with continually increasing powers of $\langle G\rangle$. The first term in this series has the form

$$\hat{K}^{(1)}(1-2)=\langle(-\mathbf{u}(1)\nabla^{(1)})\langle G(1,2)\rangle(-\mathbf{u}(2)\nabla^{(2)})\rangle. \tag{12}$$

The substitution of this kernel into (11) leads to the simplest nonlinear equation in the hierarchy of equations for $\langle G(1,2)\rangle\equiv g(1-2)$:

$$g(1-2)=G_m(1-2)+\int d3\int d4\, G_m(1-3)\times\langle\mathbf{u}(3)\nabla^{(3)}g(3-4)\mathbf{u}(4)\nabla^{(4)}\rangle g(4-2). \tag{13}$$

This is the familiar nonlinear equation proposed by Kraichnan[11] and Roberts[12] to describe turbulent transport (the direct interaction or DIA equation). Comparing the results of the numerical simulation of the transport of impurity particles with the solution of this equation, Kraichnan subsequently showed[13,14] that they agree well with one another.

Thus, the first of the equations in the hierarchy already describes the convective transport of particles well and can be employed as the free term in the renormalized equation (10). The substitution of the iterations of Eq. (10) into (5) then leads to a series of the form

$$D_T=\frac{u_0^2\tau_0}{3}\left[\frac{a}{1+d\xi_0}+\frac{b\xi_0^2}{(1+e\xi_0)^3}+\dots\right]. \tag{14}$$

This series exhibits good asymptotic convergence for all values of the parameter $0\leqslant\xi_0<\infty$. The series (14) can be regarded as an analog of Kummer's method[15] for improving the convergence of the original series (9). The first term in (14) is obtained when the free term $M\equiv g$ of the renormalized equation (10) is substituted into (5). This term completely takes into account the contribution of all the forms and powers of the two-point correlators (1) to D_T. The second term in (14) describes the contribution of the remaining fourth-order correlators etc. Of course, the choice of the kernel $\hat{K}$ in (11) is formally arbitrary, but its apparently physically substantiated choice in the form (12) is quite fortuitous. From the physical standpoint it is clear that the contributions to D_T from the higher-order velocity correlators, which describe the details of the turbulent fluctuations, should be far less significant than the contributions of the two- and four-point correlators.

We present the explicit form of the first two terms in the series $D_T=D_T^{(0)}+D_T^{(1)}+\dots$, which is obtained as a result of the substitution of the free term $M=g$ and the first iteration of Eq. (10) into (5) (here we consider only the stationary values of D_T at $t\to\infty$):

$$D_T^{(0)}=\frac{1}{3}\int_0^\infty dp\int_0^\infty d\tau\, E(p,\tau)\tilde{g}(p,\tau), \tag{15}$$

$$D_T^{(1)}=\frac{1}{24}\int_0^\infty dp\, p\int_0^\infty dq\, q\int_{-1}^1 d\mu\, \mu(1-\mu^2)\times\int_0^\infty d\tau_1\int_0^\infty d\tau_2\int_0^\infty d\tau_3 E(p,\tau_1+\tau_2)E(q,\tau_2+\tau_3)\times\tilde{g}(p,\tau_1)\tilde{g}(q,\tau_3)\tilde{g}(|\mathbf{p}+\mathbf{q}|,\tau_2). \tag{16}$$

Here $\mathbf{p}\cdot\mathbf{q}=pq\mu$, and $\tilde{g}(p,\tau,)$ is the Fourier transform $g(R,\tau)$ in the variable $\mathbf{R}$:

$$\tilde{g}(p,\tau)=\int d\mathbf{R}\,\exp(-i\mathbf{p}\cdot\mathbf{R})g(R,\tau). \tag{17}$$

In calculating (16) we assumed that the fourth-order velocity correlators are Gaussian, i.e., they were assumed to be equal to the sum of all the possible products of the second-order correlators.

Here we present the results of the calculations of $D_T^{(0)}$ and $D_T^{(1)}$ for the two limiting types of spectra, viz., a δ-function spectrum and a very broad (Kolmogorov-type) spectrum:

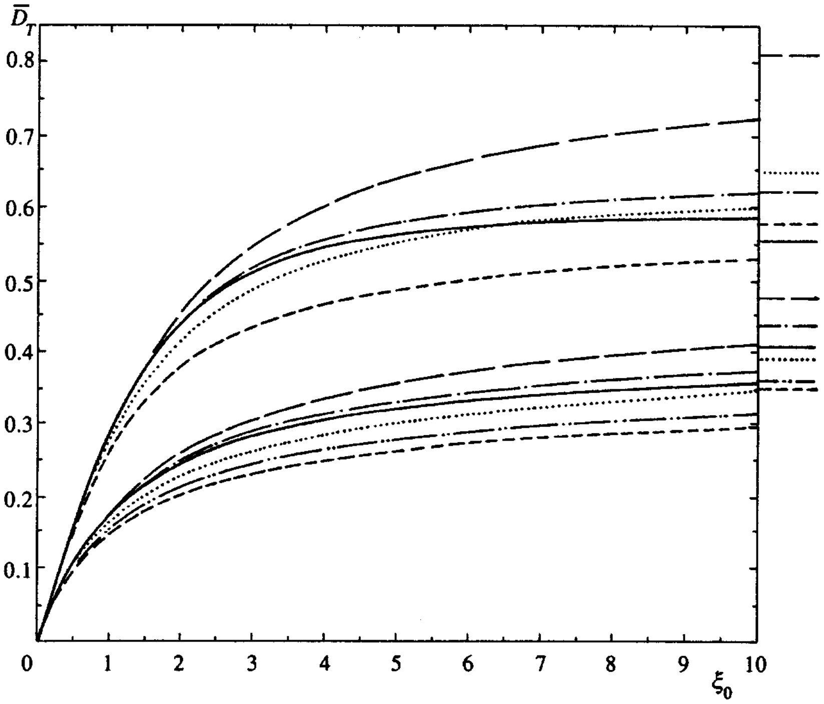


FIG. 1. Stationary values of turbulent diffusion coefficients. The upper group of curves corresponds to the δ-function spectrum $E_1(p,\tau)$, and the lower group corresponds to the Kolmogorov-type spectrum $E_2(p,\tau)$. The dot-dashed curves represent $\overline{D}_T^{(0)}(\xi_0)$, and the solid curves represent the values of $\overline{D}_T^{(0)}(\xi_0)+\overline{D}_T^{(1)}(\xi_0)$. The dotted curves give $\overline{D}_T^{(m)}(\xi_0)$. The dashed and long-dashed curves correspond to $\overline{D}_T^{(s)}(\xi_0)$ and $\overline{D}_T^{(s)}(\infty,\xi_0)$. The two-dot-dashed curves represent the combined self-consistent values of $\overline{D}_T^{(sa)}(\xi_0)$. The levels on the right correspond to the values of D_T in the limit $\xi_0\rightarrow\infty$ (frozen turbulence).

$$E_1(p,\tau)=u_0^2\delta(p-p_0)\exp(-\tau/\tau_0), \tag{18}$$

$$E_2(p,\tau)=(u_0^2/p_0)0.65159x^4/(1+x^{17/3})\exp(-\tau/\tau_0),$$

where $x=p/p_0$. In the limit $x\gg 1$, we have $E_2\propto p^{-5/3}$, i.e., the spectrum is of the Kolmogorov type. These results make it possible to evaluate the accuracy of proposed methods for intermediate types of spectra.

The values of $D_T^{(0)}(\xi_0)$ and $D_T^{(0)}(\xi_0)+D_T^{(1)}(\xi_0)$ are presented in Fig. 1. It is seen that the contribution of the fourth-order velocity correlators increases monotonically from 0% at $\xi_0=0$ to 4.1% for the E_2 spectrum and up to 5.5% for the E_1 spectrum at $\xi_0=10$. The exact values of $\overline{D}_T^{(0)}$ and $\overline{D}_T^{(1)}$ for frozen turbulence ($\xi_0\rightarrow\infty$) were obtained in Ref. 16, where they were denoted by $\underline{\underline{D}}_T^{(0)}$ and $\underline{\underline{D}}_T^{(1)}$. For the E_1 spectrum we have $\overline{D}_T^{(0)}(\infty)=0.6222$ and $\overline{D}_T^{(1)}(\infty)=-0.0691$, i.e., the correction $\overline{D}_T^{(1)}$ amounts to 11.1% of $\overline{D}_T^{(0)}$. For the E_2 spectrum we have $\overline{D}_T^{(0)}(\infty)=0.4359$ and $\overline{D}_T^{(1)}\times(\infty)=-0.0291$, i.e., the correction for this broad spectrum amounts to less than 6.7%.

These results show that the main contribution to D_T is made by the two-point velocity correlators, which describe the large-scale structure of the turbulent fluctuations. The contribution of the four-point correlators is $\leqslant 10\%$, in agreement with the comparatively small influence of small-scale turbulent motions on diffusion. It is noteworthy that the correction satisfies $D_T^{(1)}<0$. This seems physically natural, since the presence of small-scale eddies clearly means that not all the impurity particles are carried by the medium to great distances. Because the formulas are very lengthy, it is difficult to evaluate the contribution $D_T^{(2)}$ of the sixth-order correlators, which describe even smaller eddy motions. If, however, it is assumed that the relative rate of decrease of the terms in (14) is identical for at least the first terms, the maximum contribution of the sixth-order correlators will be $\leqslant 1\%$ of $D_T^{(0)}$ when $\xi_0\rightarrow\infty$. When $\xi_0\rightarrow 0$, the contribution of all the corrections $D_T^{(1)},D_T^{(2)},\ldots$ tends to zero.

The results obtained with the use of the nonlinear equation (13) and the iterations of the renormalized equation (10) are physically and mathematically sounder. The direct numerical simulation of turbulent diffusion[13] for the E_1 spectrum confirms this. Therefore, we shall assume that the values of $D_T^{(0)}+D_T^{(1)}$ are very close to the exact value of D_T and compare the results of the calculations of D_T using different approximate methods specifically with these ''standard'' values.

3. APPROXIMATE METHODS FOR CALCULATING D_T

3.1. Use of asymptotes of the nonlinear equation (13)

The numerical determination of the solutions of the nonlinear equation (13) is a fairly complicated problem. It is most convenient to use the Fourier transform in $\mathbf{R}$ and the Laplace transform in τ (we take $D_m=0$):

$$\tilde{g}(p,s)=\int d\mathbf{R}\int_0^\infty d\tau\,\exp(-s\tau)\exp\langle -i\mathbf{p}\cdot\mathbf{R})g(p,\tau), \tag{19}$$

$$\tilde{g}(p,s)=\left[s+\frac{p^2}{4}\int_0^\infty dq\int_{-1}^1 d\mu(1-\mu^2)\int_0^\infty d\tau E\right.$$
$$\left.\times(q,\tau)\tilde{g}(|\mathbf{p}-\mathbf{q}|,\tau)\exp(-s\tau)\right]^{-1}. \tag{20}$$

It is easy to see that the interations (20) comprise a continued fraction with positive terms. This means that two consecutive interations represent the exact value $\tilde{g}(p,s)$ with a deficiency

$(\tilde{g}_<(p,s)$ and an excess $(\tilde{g}_>(p,s)$. Thus, $\tilde{g}_>(p,s)=1/s$ clearly represents $\tilde{g}(p,s)$ with an excess. The substitution of this expression into (20) gives

$$\tilde{g}_<(p,s)=\left[s+\frac{u_0^2\tau_0p^2}{3(s\tau_0+1)}\right]^{-1}.$$

However, these functions are excessively rough representations of the exact value $\tilde{g}(p,s)$. A more exact representation is given by the asymptote $\tilde{g}(p,s)$ for $p\gg p_0$:

$$\tilde{g}_a(p,s)\equiv\tilde{g}_<(p,s)=2[s+(s^2+4u_0^2p^2/3)^{1/2}]^{-1}. \tag{21}$$

Substitution of (21) into (13) yields $\tilde{g}_>(p,s)$:

$$\tilde{g}_>(p,s)=\left[s+\frac{p^2}{4}\int_0^\infty dq\int_0^\infty d\tau\int_{-1}^1 d\mu(1-\mu^2)E \times(q,\tau)\tilde{g}_<(|\mathbf{p}-\mathbf{q}|,\tau)\exp(-s\tau)\right]^{-1}. \tag{22}$$

The mean value $g_m=(g_<+g_>)/2$ is a fairly good representation of the exact value of g. The substitution of g_m into (5) gives the following expression for the case of $E(p,\tau)=E(p)\exp(-\tau/\tau_0)$:

$$D_T^{(m)}(\xi_0)=\frac{\tau_0}{6}\int_0^\infty dp\,E(p)\left\{2[1+(1+4u_0^2\tau_0^2p^2/3)^{1/2}]^{-1} +\left[1+\left(\frac{p\tau_0}{2}\right)^2\int_0^\infty dq\,E(q)\int_{-1}^1 d\mu \times\frac{1-\mu^2}{1+(1+u_0^2\tau_0^2(\mathbf{p}-\mathbf{q})^2/3)^{1/2}}\right]^{-1}\right\}. \tag{23}$$

The calculation of D_T using this formula does not present any difficulties.

The results of the calculations of $D_T^{(m)}(\xi_0)$ for the spectra (18) are presented in Fig. 1 (the dotted curves). It is seen that for both spectra the plots of $D_T^{(m)}$ are close to the values of $D_T^{(0)}+D_T^{(1)}$. For $E_2(p,\tau)$ a maximum difference of $\simeq 12\%$ is observed at $\xi_0\simeq 5$, and it decreases to 4% for $\xi_0\to\infty$. In the case of the narrow-band spectrum $E_1(p,\tau)$, a maximum difference of $\simeq 13\%$ is observed for $\xi_0\to\infty$.

Thus, the comparatively simple formula (23) represents D_T quite satisfactorily, especially for frozen turbulence with a broad spectrum. More exact approximate formulas can probably be found for the Green's function $g(R,\tau)$ and used to calculate D_T.

3.2. Self-consistent method

We have already stated that the large-scale turbulent motions make the largest contribution to D_T. On the other hand, for such scales the Green's function $\langle G(1,2)\rangle$ is described well by the diffusion formula (8), in which the sum $D=D_m+D_T\simeq D_T$ must be taken instead of D_m. Therefore, it is natural to employ the expression $\tilde{g}(p,\tau)=\exp(-D_T^{(s)}p^2\tau)$ with the as yet unknown diffusion coefficient $D_T^{(s)}(\xi_0)$ as $\tilde{g}(p,\tau)$ in (15):

$$D_T^{(s)}(\xi_0)=\frac{1}{3}\int_0^\infty dp\int_0^\infty d\tau\,E(p,\tau)\exp[-D_T^{(s)}(\xi_0)p^2\tau]. \tag{24}$$

The expression (24) is a nonlinear equation for finding the self-consistent diffusion coefficient $D_T^{(s)}(\xi_0)$. This expression was apparently first proposed in Ref. 17.

The results of the calculations of $D_T^{(s)}(\xi_0)$ are presented in Fig. 1 (dashed lines). The maximum departure from $D_T^{(0)}+D_T^{(1)}$ for the δ-function spectrum $E_1(p,\tau)$ amounts to 15% at $\xi_0=3$, and then the difference decreases (9.5% at $\xi_0=10$ and 4% for $\xi_0\to\infty$). The difference for the broad spectrum $E_2(p,\tau)$ is greater: 19%, 17%, and 15%, respectively for $\xi_0=3$, 10, and ∞. Such accuracy is perfectly acceptable for many cases.

3.3. Combined self-consistent method

A more exact result is provided by the method in which the diffusion Green's function $\tilde{g}(p,\tau)=\exp(-D_T^{(sa)}p^2\tau)$ is used for $p\leqslant p_*$ and the asymptote (21) is used for $p\geqslant p_*$. For $E(p,\tau)\propto\exp(-\tau/\tau_0)$ such a combined method gives

$$D_T^{(sa)}=\frac{\tau_0}{3}\int_0^{p_*}dp\,\frac{E(p)}{1+D_T^{(sa)}p^2\tau_0} +\frac{2\tau_0}{3}\int_{p_*}^\infty dp\,\frac{E(p)}{1+(1+4p^2u_0^2\tau_0^2/3)^{1/2}}. \tag{25}$$

Continuity of the combined Green's function at the point p_* gives

$$p_*=[(u_0^2\tau_0-3D_T^{sa})/3(D_T^{(sa)})^2\tau_0]^{1/2}. \tag{26}$$

The results of the calculations based on this method for $E_2(p,\tau)$ are given in Fig. 1 (the two-dot-dashed curve). It is seen that this method gives values that are closer to $D_T^{(0)}+D_T^{(1)}$ than does the pure self-consistent method. For the δ-function spectrum $E_1(p,\tau)$ this method is as accurate as the pure self-consistent method.

3.4. Self-consistent method with a time-dependent diffusion coefficient

The stationary values of $D_T(\xi_0)$ can also be calculated as the $t\to\infty$ limit of the time-dependent turbulent diffusion coefficient $D_T(t,\xi_0)$. A self-consistent method for calculating $D_T(t,\xi_0)$ follows from (5), if a diffusion Green's function with a time-dependent diffusion coefficient is taken as the Green's function. Ultimately we obtain

$$D_T^{(s)}(t,\xi_0)=\frac{1}{3}\int_0^\infty dp\int_0^t d\tau\,E(p,\tau) \times\exp\left[-\int_0^t d\tau' D_T^{(s)}(\tau',\xi_0)p^2\right]. \tag{27}$$

If we take $D_T^{(s)}(t,\xi_0)\equiv D_T^{(s)}(\xi_0)$ here and let $t\to\infty$, Eq. (27) transforms into (24). The coefficient $D_T^{(s)}(t,\xi_0)$ takes a stationary value, in the limit $t\gg\tau_0$ for $\xi_0\ll 1$ or $t\gg t_0=1/u_0p_0$ for $\xi_0\gg 1$, i.e., the process is determined by the smaller of the characteristic times.

The results of the calculations of $D_T^{(s)}(\infty,\xi_0)$ are given in Fig. 1 (the long-dashed curve). It is seen that the curves practically coincide with $D_T^{(0)}+D_T^{(1)}$ at $\xi_0 \leqslant 1$. Then the difference increases monotonically and reaches 23% at $\xi_0=10$ for the δ-function spectrum $E_1(p,\tau)$ and 15% for the broad spectrum $E_2(p,\tau)$. In the limit $\xi_0\to\infty$ the differences are even greater, being $\simeq 48\%$ and 17%, respectively. A comparison of Eqs. (24) and (27) reveals that $D_T^{(s)}(t,\xi_0) > D_T^{(s)}(\xi_0)$. The figure confirms this. The values of $D_T^{(0)}+D_T^{(1)}$ lie between the plots of $D_T^{(s)}(\infty,\xi_0)$ and $D_T^{(s)} \times(\xi_0)$, which can be regarded as the upper and lower bounds for the real turbulent diffusion coefficient $D_T(\xi_0)$. In addition, the mean value $D_T^{(sm)}=(D_T^{(s)}(\infty,\xi_0)+D_T^{(s)}(\xi_0))/2$ of these coefficients represents $D_T^{(0)}+D_T^{(1)}$ with very good accuracy. The difference between this mean and $D_T^{(0)}+D_T^{(1)}$ for the broad spectrum $E_2(p,\tau)$ amounts to 3%, 1.5%, and 1.2% at $\xi_0=5$, 10, and ∞, respectively, i.e., in this case $D_T^{(sm)}$ is more accurate than $D_T^{(m)}$. However, the difference for the δ-function spectrum $E_1(p,\tau)$ is greater and amounts to 0.1%, 7%, and 25%, respectively, at the same values of ξ_0. For this spectrum $D_T^{(s)}(\xi_0)$ is more accurate for $\xi_0 > 10$.

4. CONCLUSIONS

A comparison of the methods considered above leads to the following conclusions.

1. The method based on the numerical solution of the nonlinear equation (13) and iteration of the renormalized equation (10) is the most accurate method for calculating turbulent diffusion coefficients.

2. Of all the approximate methods considered, the method which is based on the use of an approximate expression for the solution of the nonlinear equation (13) and gives (23) is the most accurate and preferable method.

3. For turbulence characterized by a broad spectrum, the mean value $D_T^{(sm)}$ of the stationary and nonstationary self-consistent values of D_T also represents $D_T(\xi_0)$ with good accuracy.

4. All the methods considered can be generalized for calculating the nonstationary diffusion coefficients $D_T(t,\xi_0)$. For short times $D_T(t,\xi_0)\simeq t/3t_0$ is a good approximation up to $t/t_0 \leqslant \xi_0$ for turbulence with $\xi_0 \ll 1$ and up to $t/t_0 \leqslant 1$ for $\xi_0 \gg 1$. We recall that $t_0=1/u_0p_0\simeq R_0/u_0$. For $\xi_0\ll 1$ the quantity $D_T(t,\xi_0)$ achieves a stationary value at $t/t_0\simeq 2-3$, and for $\xi_0 \gg 1$ it achieves a stationary value at $t/t_0\simeq 10-20$.

We thank the Foundation for Fundamental Research of the Russian Academy of Sciences for supporting this work.

[1] N. A. Silant'ev, Zh. Éksp. Teor. Fiz. **101**, 1216 (1992) [Sov. Phys JETP **74**, 650 (1992)].
[2] A. Z. Dolginov and N. A. Silant'ev, Astrophys. Fluid Dyn **63**, 139 (1992).
[3] F. Richardson, Proc. R. Soc. London A **110**, 209 (1926).
[4] G. K. Batchelor, Proc. Cambridge Philos. Soc. **48**, 345 (1952).
[5] *Survey in Mechanics*, G. K. Batchelor (ed.), Cambridge University Press, Cambridge (1956), p. 352.
[6] H. G. E. Hentshel and I. Procaccia, Phys. Rev. A **29**, 1461 (1984).
[7] A. S. Monin and A. M. Yaglom, *Statistical Fluid Mechanics; Mechanics of Turbulence, Vols. 1 and 2*, MIT Press, Cambridge, Mass. (1971, 1975).
[8] *Turbulence*, P. Bradshaw (ed.), Springer Verlag, Berlin (1978).
[9] E. N. Parker, *Cosmical Magnetic Fields*, Clarendon Press, Oxford (1979).
[10] G. I. Taylor, Proc. London Math. Soc. A **20**, 196 (1921).
[11] R. H. Kraichnan, J. Fluid Mech. **5**, 497 (1959).
[12] P. H. Roberts, J. Fluid Mech. **11**, 257 (1961).
[13] R. H. Kraichnan, Phys. Fluids **13**, 22 (1970).
[14] R. H. Kraichnan, J. Fluid Mech. **81**, 385 (1977).
[15] G. M. Fikhtengol'ts, *The Fundamentals of Mathematical Analysis, Vol. 2*, Pergamon Press, Oxford–New York (1965).
[16] N. A. Silant ev, Geophys. Astrophys. Fluid Dyn. **75**, 183 (1994).
[17] R. Phythian and W. D. Curtis, J. Fluid Mech. **89**, 241 (1978).

Translated by P. Shelnitz

Investigation of the turbulent mixing of thin layers of materials of different density during the laser acceleration of flat multilayer targets in the Iskra-4 facility

V. A. Andronov,[1)] S. A. Bel'kov, A. V. Bessarab, I. N. Voronich, S. G. Garanin,
A. A. Gorbunov, V. N. Derkach, G. V. Dolgoleva, A. I. Zaretskiĭ, V. M. Izgorodin,
B. N. Ilyushechkin, G. A. Kirillov, G. G. Kochemasov, Yu. V. Kuratov, V. I. Lazarchuk,
V. A. Lebedev, V. M. Murugov, L. S. Mkhitar'yan, A. V. Okutin, S. I. Petrov,
A. V. Pinegin, N. N. Rukavishnikov, A. N. Razin, A. V. Ryadov, A. V. Senik,
N. A. Suslov, S. A. Sukharev, and V. A. Tokarev

(Submitted 19 April 1996)
Zh. Éksp. Teor. Fiz. **111**, 882–888 (March 1997)

The results of the first experiments devised to investigate the mixing of thin layers of Al and Au during the laser acceleration of flat three-layer targets of Si (5 μm), Al (2 μm), and Au (0.05–0.26 μm) by radiation converted to the second harmonic from the Iskra-4 iodine laser with an intensity of $4\times10^{13}-7\times10^{13}$ W/cm^2 ($\tau_{0.5}\sim1$ ns), which acts on the Si side of the target. A method for detecting the occurrence of mixing is developed. It is established that under the experimental conditions the thickness of the mixing region is at least ~0.15 μm. The results of a theoretical analysis of the evolution of the disturbances leading to mixing are presented. © *1997 American Institute of Physics.* [S1063-7761(97)01003-2]

The turbulent mixing of thermonuclear fuel with the material of the target capsule compressing it is a serious problem in the intertial-confinement thermonuclear fusion program. This process leads to a significant increase in the energy of the laser radiation needed to effect the ignition of laser targets.[1,2]

Mathematical-physical models of the mixing process,[3] which have been experimentally confirmed in shock-tube experiments,[4] are presently known. However, the acceleration of the target capsule by laser radiation has some special features, which preclude extrapolating the existing models to the case of targets for laser-driven thermonuclear fusion without experimental verification. For this reason, specially designed experiments involving the laser acceleration of such capsules or their flat analogs would be of interest.

This paper presents the results of the first experiments performed to investigate the mixing of thin layers of Al and Au when three-layer flat Si–Al–Au targets are accelerated in the Iskra-4 facility.[5] The methodical approach to performing the experiments under consideration, whose main features were proposed in Refs. 6 and 7, was implemented in the following manner (see the schematic representation in Fig. 1). A Si–Al–Au target was irradiated from the Si side by a pulse with wavelength $\lambda=0.66$ μm, duration $\tau_p\approx0.8-1.2$ ns, and intensity $I_p\approx0.5\times10^{14}-1\times10^{14}$W/cm^2, which we shall henceforth call the power pulse. The thickness of the silicon layer was selected so that the thermal wave would not manage to reach the Al layer during the period of action of the power laser radiation, i.e., the aluminum and gold layers were accelerated in the ''cold'' state (if their heating by the x radiation of the plasma corona and the shock wave is disregarded). The conditions for the development of Rayleigh–Taylor instability, which results in turbulent mixing, obtained on the Al–Au interface because of the large difference between the densities ρ of these materials and the direction of acceleration of the materials realized $(\mathbf{g}\cdot\nabla\rho<0)$.

When the mixture of these materials formed is heated by laser radiation of suitable intensity, a plasma of Au and Al ions can be obtained, and the occurrence of mixing can be detected by observing the appearance of the x-ray line emission of those ions. For this purpose, a diagnostic pulse with duration $\tau_d\approx0.3-0.5$ nm and intensity $I_d\approx<10^{13}$ W/cm^2 is supplied coaxially with the power pulse to the rear side of the target in the experiments. The delay of the diagnostic beam relative to the power beam was selected so that the mixing process would manage to develop. In the experiments under consideration this delay amounted to $\Delta t\sim0.86$ ns.

In the absence of mixing a possible reason for the appearance of the Al line emission may be that a sufficiently thin layer of Au is ''burned through'' because of the presence of ''hot'' spots in the intensity distribution over the cross section of the diagnostic beam. However, even when mixing occurs, the presence of these ''hot'' spots must be taken into account when the experiment is compared with theoretical estimates of the time for the appearance of the Al x-ray line emission. To monitor the possibility of a manifestation of the effect just indicated on the target, an additional beam, which we shall henceforth call the control beam, was directed onto the target from the gold side. The design delay of this beam relative to the power beam was ~1.6 ns, and the spot was located at a distance of 4 mm from the diagnostic beam. Both spots were simultaneously within the field of vision of the diagnostic instrument. It was assumed that the laser flux densities and the intensity distributions in the diagnostic and control spots are similar. An identical delay of the appearance of the Al x-ray line emission relative to the beginning of the diagnostic and control laser radiation provides evidence that the Au layer is burned through at the hot points of the respective spots. The appearance of the pulse of Al x radiation from the diagnostic spot before the corresponding x-ray pulse from the control spot (or the absence of the latter pulse) can indicate the occurrence of mixing.

The target fabrication technique ensured a specular surface on the Al and Au layers with a roughness no greater than 0.05 μm.[8] The roughness of the Si surface was 0.1–0.3

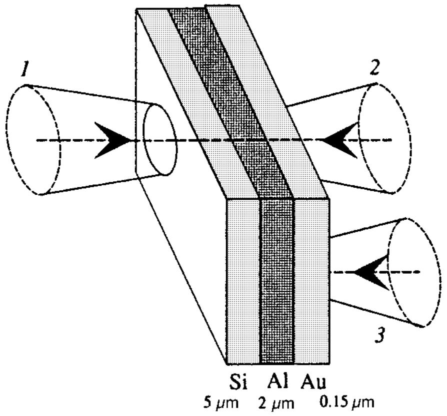


FIG. 1. Schematic representation of the mixing experiments: power (*1*), diagnostic (*2*), and control (*3*) beams.

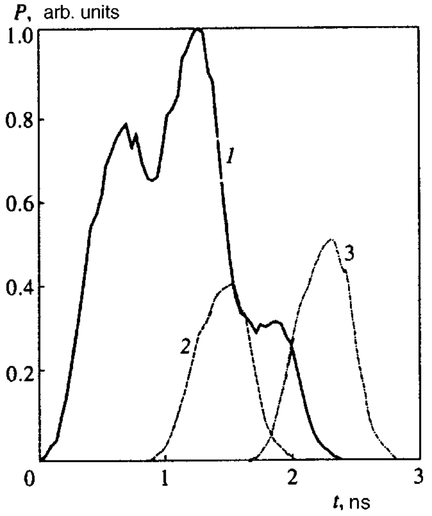


FIG. 2. Temporal irradiation diagram of targets in one of the experiments: power (*1*), diagnostic (*2*), and control (*3*) beams.

μm. In the experiments described here the thickness of the Si layer was 4.3–5 μm, the thickness of the Al layer was 1.7–2.0 μm, and the thickness of the Au layer varied in the range 0.05–0.3 μm.

The temporal parameters of the laser radiation (the shape and duration of the pulses) were determined using an SER type electro-optic device detector[9] with a time resolution of ~ 40 ps. The values of the energy in the power, diagnostic, and control beams were measured using TPI-2A calorimeters. The measurement error was $\pm 15\%$. The energy of the laser radiation absorbed by the target was also measured in the experiments. Plasma and optical calorimeters[10] placed within the vacuum interaction chamber were used for this purpose.

The dimensions of the irradiation spots of the power, diagnostic, and control spots and the nature of the distribution of the intensity of the laser radiation in them (the presence of hot points) were evaluated using x-ray pinhole cameras.

The x-ray diagnostic system provided for the recording of the following: the time-integrated x-ray line emission spectrum of the plasma in the 0.04–0.25 keV and 1.4–2.5 keV ranges using spectrographs on a diffraction grating and a KAP crystal; the x-ray line emission near the He_αAl line ($h\nu \approx 1.6$ keV) on a background of the x emission in the Au M band with a time resolution of ~ 150 ps; the x-ray emission continuum of the plasma in the 1–5 keV range with a time resolution of ~ 50 ps.

The shape and time sequence of the power, diagnostic, and control laser pulses are presented in Fig. 2.

The intensity distribution in the power beam was fairly smooth. At the same time, numerous hot microregions were observed on the pinhole images of the diagnostic and especially the control irradiation spots. The transverse dimensions of the control spots were 1.5–2 times greater than the dimensions of the diagnostic beam on the target. This caused the corresponding mean intensities in the beams to differ by a factor of 2–3. The transverse dimensions of the power and diagnostic beams on the target were similar.

The flux density of the laser radiation in the power beam varied in the range $I_p \sim 5 \times 10^{13} - 10^{14}$ W/cm^2, the duration varied in the range 0.83–1.1 ns, and the diameter of the spot varied in the range 320–520 μm. The intensity of the diagnostic beam fell within the range $2 \cdot 10^{13} - 4 \cdot 10^{13}$ W/cm^2, and the intensity of the control beam was between 0.7×10^{13} and 2×10^{13} W/cm^2 when the pulse duration was 0.3–0.5 ns. The corresponding spot diameters were 390–530 μm and 500–800 μm. The absorption factor of the energy of the power beam in the target was $K_a \approx 0.9$ in these experiments.

Let us turn to the results from recording the x-ray emission spectrum of the plasma formed on the rear side of the target. We, first of all, note that the Al line emission in the 0.04–0.25 keV range was recorded on the time-integrated spectrograms on a background of the x-ray emission of gold from both the diagnostic and control spots when the thickness of the gold layer was up to 0.15 μm. As for the control spot, this is evidence either that Au is burned through at the hot points or that a layer of Au of thickness 0.15 μm manages to vaporize within the duration of the control pulse.

The results of the temporal recording of the x radiation of the plasma in the range of quantum energies near $h\nu \sim 1.5$ keV for thicknesses of the gold layer on the target equal to 0.1 and 0.16 μm are presented in Fig. 3. In each figure the spectral photochronograms are presented on the left, and the plots of the time dependence of the intensity of the x-ray emission in the Au continuum and the He_αA1 line ($\lambda = 0.776$ nm) obtained as a result of their treatment are presented on the right. The duration of the emission of the Au continuum at half-height corresponds approximately to the duration of the diagnostic and control laser pulses (~ 0.5 ns), and the temporal distance between them is ~ 0.7 ns, which corresponds to the difference between the arrival times of the diagnostic and control beams on the rear side of the target. In this case the temporal shape of the Al emission approximately repeats the shape of the pulse of the Au x-ray continuum.

When the thickness of the gold layer is 0.1 μm, Al line emission is observed for both the diagnostic and control beams, simultaneously with the Au emission. However, for

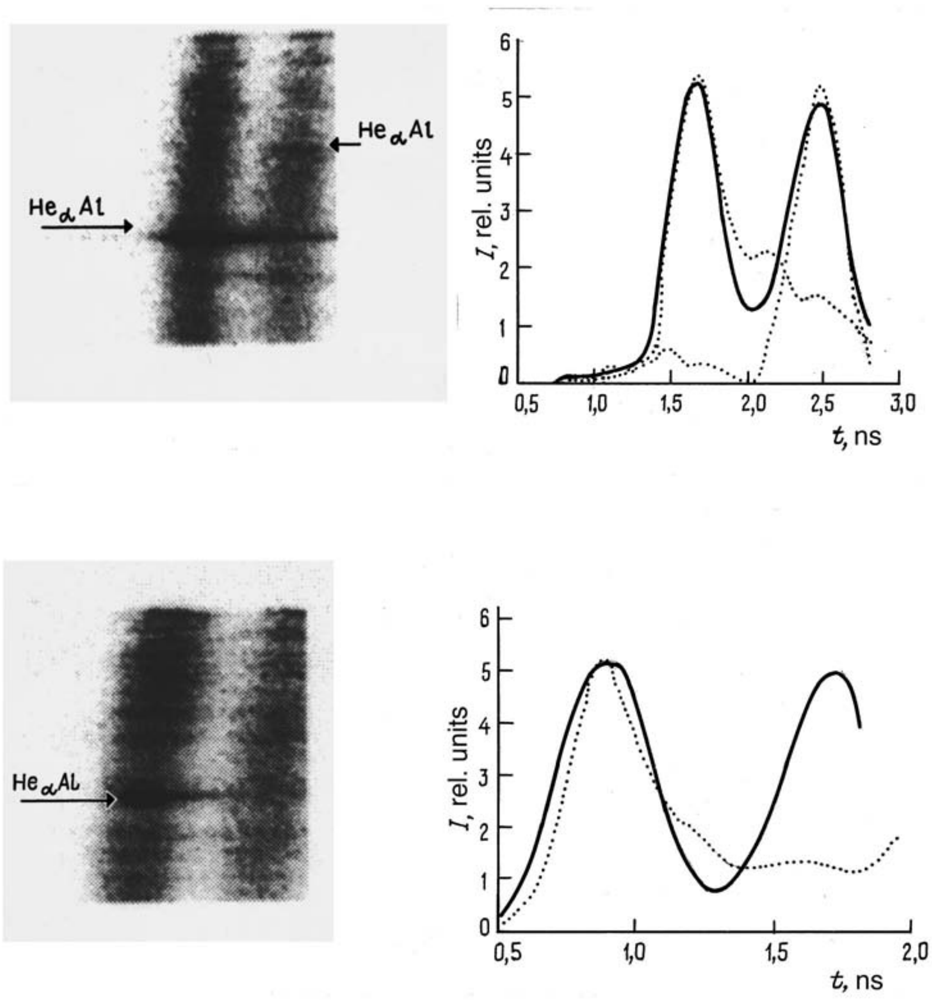


FIG. 3. Spectral photochronograms and plots of the time dependence of the intensity of the x-ray emission of gold (solid line) and aluminum (dotted line). The numbers *1* and *2* label the spectral bands corresponding to the continuous x radiation from the diagnostic and control spots; the upper part of the figure corresponds to $\delta_{\mathrm{Au}}=0.1\ \mu$m, and the lower part of the figure corresponds to $\delta_{\mathrm{Au}}=0.16\ \mu$m.

the control beam this finding is most likely attributable to the rapid burning through of the Au at the hot points, whose presence was evidenced by the pinhole image of the experiment. At the same time, the fairly good synchronism of the Al and Au x-ray emission pulses for the diagnostic beam can be attributed to the occurrence of their complete mixing as a result of acceleration of the target by the power laser radiation already at the arrival time of the diagnostic pulse.

Considerable synchronism of the Au and Al emission pulses in the diagnostic spot is also observed when the thickness of the gold layer is 0.15–0.16 μm. At the same time, there is no Al emission from the control spot. Such a picture definitely attests to the presence of a region of mixing of Al and Au of thickness 0.15–0.16 μm at the arrival time of the diagnostic pulse.

Preliminary calculations of the development of turbulent mixing when a three-layer target is accelerated were performed using physical models[3] and the SNDP program.[11] The shape of the power laser pulse was approximated in the calculations by an isosceles trapezium with a duration at the base $\tau_b=2$ ns and rise and decay times $\tau_e=0.6$ ns. The intensity at the maximum was equal to $I_p=7.5\times10^{13}$ W/cm^2. The diagnostic pulse was incident on the rear (gold) side of the three-layer target with a delay of the pulse maximum relative to the beginning of the power pulse $\Delta t=1.46$ ns. The shape of the diagnostic pulse was given by a Gaussian line profile with a duration at half-maximum $\tau_{0.5}=0.45$ ns and a peak intensity $I_d=3\times10^{13}$ W/cm^2. The thicknesses of the Si (5 μm) and Al (2 μm) layers corresponded to the experimental values. The thickness of the Au layer was varied from 0.15 to 0.3 μm.

The following physical processes were taken into account in the calculations: the gas dynamics, the electronic and ionic heat conduction, the nonequilibrium spectral diffusion of the x radiation, the relaxation of the electron and ion temperatures, and the kinetics of the ionization of the multicomponent high-Z plasma. The electronic thermal conductance was calculated with consideration of the heat flux limiting. The limiting factor was set equal to $f=0.1$. The absorption of the laser radiation was taken into account with consideration of the inverse bremsstrahlung mechanism only.

Figure 4 shows the motion of the interface between the Au and Al layers in the calculation without mixing. Up to approximately $t=0.5$ ns the motion of the boundary is governed by the removal of the thin gold layer under the action of the hard x radiation in the silicon laser corona (mainly the line emission in the silicon K_α region). The Au and Al densities are practically equalized, and no mixing appears. Then, at the moment when the shock wave appears, the interface accelerates abruptly, and a density jump forms. Rayleigh–

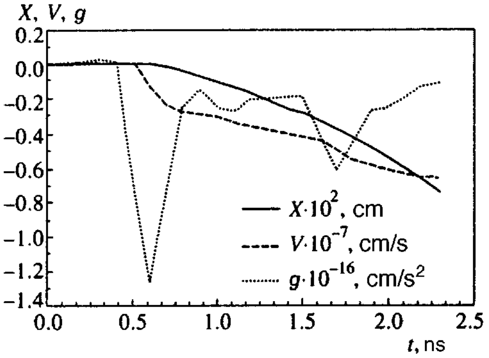


FIG. 4. $X - t$ diagram of the motion of the interface between the gold and aluminum layers and diagrams of the velocity V and acceleration g of the interface calculated without mixing.

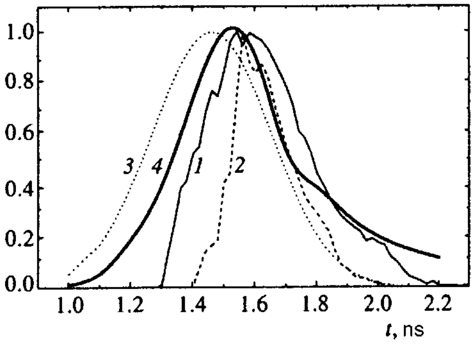


FIG. 6. Calculated time dependence of the emission in the He$_{\alpha}$Al lines from the rear side of the target for different thicknesses of the gold layer (*1* – $\delta_{\mathrm{Au}}=0.15\ \mu$m, *2* – $\delta_{\mathrm{Au}}=0.3\ \mu$m), of the diagnostic laser pulse (*3*), and of the x radiation in the continuous spectrum of the gold plasma (*4*).

Taylor instability, which results in mixing of the layers, should develop after this moment.

Figure 5 presents a density profile and the distribution of the mass concentrations of gold and aluminum calculated with consideration of mixing at $t=1.4$ ns, which corresponds approximately to the intensity maximum of the diagnostic pulse for a gold thickness equal to 0.15 μm. It is seen from the calculation that because of the rapid removal of the gold layer as a result of its being heated by the hard x radiation of the Si laser corona, the mixing zone for the experimentally realized acceleration values does not manage to cover the entire layer. At the time represented in Fig. 5 the mass of the mixed gold layer corresponds approximately to half of its initial mass. Similar behavior of the mixing zone is observed in the calculation with a thickness of the gold layer equal to 0.3 μm. The comparison of the calculated emission pulses in the He$_{\alpha}$Al lines in calculations with different thicknesses of the gold layer and of the background emission of the gold plasma is more striking.

Figure 6 presents calculated plots of the time dependence of the emission intensity of the Al lines obtained in the calculations with a thicknesses of the gold layer equal to 0.15 and 0.3 μm (curves *1* and *2*, respectively). The figure also presents the emission intensity of the gold plasma (curve *4*) in the spectral range near the quantum energy of the lines of helium-like aluminum, as well as the diagnostic laser pulse (curve *3*). The delay of the emission of the gold plasma relative to the diagnostic pulse is clearly seen (it equals approximately 100 ps). For a thin layer of gold its emission in the continuum spectrum and the emission of the aluminum lines practically coincide along the time scale (the delay amounts to ~50 ps, which is appreciably smaller than the temporal resolution of the electro-optic device detector). A significant delay (about 0.2 ns) of the beginning of the emission of the aluminum lines relative to the Au emission in the continuum spectrum is observed for the thick gold layer.

A comparison of the Au emission and the aluminum lines in the experiments with gold layers of different thickness suggests the appearance of such a delay (~50 ps) for a layer thickness $\delta_{\mathrm{Au}}\approx 0.15\ \mu$m; however, the insufficiently high temporal resolution of the detector precludes drawing an unequivocal conclusion regarding its presence and magnitude. As the calculations show, when the thickness of the gold layer is increased, reliable detection of this delay can be expected.

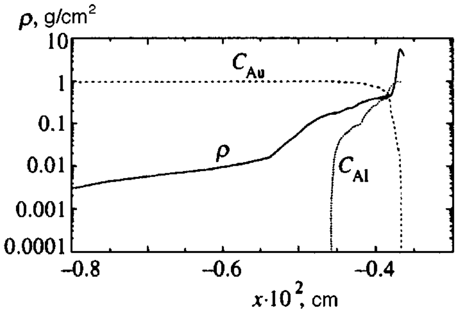


FIG. 5. Distribution of the density and the mass concentrations $C_{\mathrm{Au,Al}}$ of gold and aluminum in a calculation with mixing for a gold layer thickness equal to 0.15 μm at the time $t=1.4$ ns.

[1)] Deceased.

---

[1] G. G. Kochemasov, in *Proceedings of the 23rd European Conference on Laser Interaction with Matter. Oxford, 19–23 September 1994, Institute of Physics Conference Series No. 140*, (1995), p. 17.

[2] M. D. Cable, S. P. Hatchett, J. A. Caird *et al.*, Phys. Rev. Lett. **73**, 2316 (1994).

[3] V. A. Andronov, V. I. Kozlov, V. V. Nikiforov, A. N. Razin, and Yu. A. Yudin, Vopr. At. Nauki Tekh., Ser. Mat. Modelirovanie Fiz. Protsessov No. 2, 59 (1994).

[4] V. A. Andronov, E. E. Meshkov, V. V. Nikiforov *et al.*, Dokl. Akad. Nauk SSSR. **264**, 76 (1982) [Sov. Phys. Doklady **27**, 393 (1982)].

[5] I. N. Voronich, V. A. Eroshenko, A. I. Zaretskiĭ *et al.*, Izv. Akad. Nauk SSSR., Ser. Fiz. **54**, 2016 (1990).

[6] A. Raven, H. Azechi, T. Yamanaka, and C. Yamanaka, Phys. Rev. Lett. **47**, 1049 (1981).

[7] P. A. Holstein, B. Meyer, M. Rostaing *et al.*, C. R. Acad. Sci. **307**, 211 (1988).

[8] N. V. Pleshivtsev, *Cathode Sputtering* [in Russian], Atomizdat, Moscow (1968).

[9] V. P. Lazarchuk, V. M. Murugov, S. I. Petrov, and A. V. Senik, Fiz. Plazmy **20**, 101 (1994) [Plasma Phys. Repts. **20**, 94 (1994)].

[10] S. R. Gann and V. C. Rupert, in Laser Program Annual Report-1975, LLNL, UCRL-50021-75 (1975), p. 404.

[11] S. A. Bel'kov and G. V. Dolgoleva, Vopr. At. Nauki Tekh., Ser. Mat. Modelirovanie Fiz. Protsessov (1), 59 (1992).

Translated by P. Shelnitz

Creation of ordered structures in a classical thermal plasma containing macroparticles: experiment and computer simulation

V. E. Fortov, V. S. Filinov, A. P. Nefedov, O. F. Petrov, A. A. Samaryan, and A. M. Lipaev

Science and Research Center for the Thermal Physics of Pulsed Excitations, Russian Academy of Sciences, 127412 Moscow, Russia

(Submitted 17 June 1996; resubmitted 3 October 1996)
Zh. Éksp. Teor. Fiz. **111**, 889–902 (March 1997)

We have compared experimental measurements of ordered structures in a thermal plasma containing macroparticles of CeO_2 at atmospheric pressure and a temperature around 1700 K with the results of numerical Monte Carlo calculations for the Yukawa model. We describe several distinctive features of the way the experiments were done, including how the ordered macroparticle structures were detected. We discuss a theoretical model of the behavior of an equilibrium system of charged macroparticles in a plasma and the effective interaction potential between them. Good agreement between the experimental and numerical results is noted, and possible reasons for the observed discrepancies are discussed. [S1063-7761(97)01103-7]

1. INTRODUCTION

A characteristic of plasmas with macroparticles (often referred to as a dusty or aerosol plasma, or a plasma with a condensed dispersive phase) is the fact that the particles (whose sizes can vary from hundredths of a micron up to several tens of microns) effectively interact with the electrons and ions, and thus significantly affect the properties of the plasma. When these particles are introduced into the plasma, or just appear in it, e.g., as the result of condensation, they are charged by the electron and ion currents, and also by electron emission. The latter process can lead to a positive electric charge; in this case, particles that emit electrons can raise the electron density in the gas phase and thus its electrical conductivity. If, however, the particles trap electrons, then the opposite effect occurs.[1–4]

One of the parameters used to describe the plasma is the so-called interaction parameter γ_p, which is the ratio of the average Coulomb energy of a particle to its average thermal energy. A distinguishing feature of a plasma with macroparticles is the fact that the charge Z_p of a particle can have extremely large values (of order 10^2-10^5 electron charges). As a result, the interaction parameter γ_p, which depends on Z_p^2, can greatly exceed unity, which implies that the resulting plasma is highly nonideal. Theoretical calculations of the equilibrium properties of such a plasma show that under certain conditions strong interparticle correlations lead to the appearance of ordered structures in the distribution of macroscopic particles analogous to structures in liquids or solids.[5] Electrons and ions in this case remain an ideal gas, as in a Debye plasma.

Here we will conduct a detailed investigation of ordered structures, including the conditions for their formation, within the framework of the single-component plasma and the Yukawa models.[5–7]

For a one-component plasma the interaction potential between particles $U(r)$ is Coulombic, and the plasma thermodynamics are characterized by a parameter γ_p of the form

$$\gamma_p=(Z_pe)^2/\langle r\rangle kT_g.$$

Here T_g is the plasma temperature, $\langle r\rangle=(4\pi n_p/3)^{-1/3}$ is the mean distance between particles, and n_p is the particle density. The Yukawa model includes the effect of screening by a background charge, which leads to an interaction potential of Debye–Hückel type. The effect of screening is determined by the ratio $\kappa=\langle r\rangle/r_D$ (where r_D is the Debye radius). The model of a one-component plasma is a limiting case of the Yukawa model as $\langle r\rangle/\to r_D 0$.

Note that these models treat a classical quasineutral spatially unbounded plasma, for which critical values of the interparticle interaction parameter are obtained by numerical simulation; these critical values correspond to phase transitions.[8,9] Thus, in a one-component plasma the three-dimensional system forms a regular crystalline structure for values of γ_p larger than $\gamma_c=171$. For small values of γ_p ($\gamma_p<4$) the plasma is in a gaslike state (see Ref. 8).

One of the first experimental observations of crystal structures was in systems of micron-size charged particles of iron and aluminum confined by a certain configuration of AC and DC electric fields.[10] Later investigations revealed Coulomb crystals of atomic ions in traps of various kinds, for example in Penning traps.[11] For macroscopic particles with large negative charge ($\sim10^4e$), crystalline states are observed when the particles are introduced into the boundary space-charge layer of a high-frequency discharge[12] in which equilibrium is established between gravitational and electrostatic forces.[13–16]

In practically all the experimental papers known to us, ordered structures are observed in clouds of space charge containing from a hundred to a few thousand charged particles. The interaction potential between particles, whose form is found to have a strong effect on the phase transitions in the plasma, can in this case differ significantly from the interaction potential in a classical quasineutral plasma. For plasma structures of this kind, theory also predicts that boundary conditions have a strong effect on the phase state of the plasma.[17] Thus, for example, a cloud of particles in a spherical trap separates into spherical layers. In place of sharp phase transitions, the system undergoes a gradual evo-

1063-7761/97/030489-08$10.00

lution from a liquid state characterized by short-range order to an intermediate state characterized by the coexistence of a liquid and solid phase, and finally to a solid state.[18]

For a system of macroparticles in the boundary layer of a high-frequency discharge, a crystalline structure is observed for $2.1^{-}10^4<\gamma_p<1.6\times10^5$ and $0.6\leqslant\kappa\leqslant4.8$ (Ref. 14), or $\gamma_p\sim10^6$ for $\kappa\sim9.8$ (Ref. 16). However, numerical calculations using the Yukawa model predict that for crystallization we require $\gamma_p\approx99$ for $\kappa=0.7$, $160<\gamma_p<850$ for $1<\kappa<5$, and $\gamma_p=4.8\times10^4$ for $\kappa=10$ (Ref. 9, 19). In Ref. 20 this considerable jump in the value of γ_p (by roughly two orders of magnitude) was explained by Brownian motion of the particles. The corresponding model calculations lead to values of $\gamma_p\approx1.6\times10^4$ for $\kappa=4$. Nevertheless, a considerable difference remains (a factor of ~10) in the value of γ_p, which is assigned to perturbations of the external high-frequency field.

Thus, the study of phase transitions in systems of particles confined in traps requires both inclusion of the effect of the external field and the choice of an adequate particle interaction potential. All these factors make the problem of numerically simulating phase transitions in these systems so complicated that present-day efforts have met with only isolated successes.

In this paper we study experimentally and theoretically the appearance of ordered structures in a weakly ionized thermal plasma containing charged macroparticles undergoing practically laminar flow at atmospheric pressure and temperatures of order 2000 degrees.[21] Our theoretical analysis is based on numerical calculations using the Monte Carlo method to solve the Yukawa model. The rather large size of the region of plasma formation (a volume of ~10 cm^3, corresponding to macroparticle numbers of order 10^8 of densities $\sim10^7$ cm^{-3}), together with the uniformity of the plasma and the absence of external electric and magnetic fields, allow us to minimize the influence of boundary conditions on the phase transitions in the plasma and thereby correctly analyze the experimental results within the framework of the Yukawa model.

2. EFFECTIVE INTERACTION OF MACROPARTICLES IN A PLASMA

The problem of correctly constructing an effective pseudopotential for the interaction of macroparticles in a low-temperature plasma is extremely difficult, and up to now has not been solved. An analogous problem arises in the theory of suspensions of charged particles as well.[22,23] Many theoretical papers begin their investigations of the properties of these systems within the framework of the so-called primitive model. In this model the interaction between charged macroparticles is described by a hard-sphere Coulomb potential, for which we have $V_{mm}(r)=\infty$ for $r\leqslant2R_p$ while for $r>2R_p$

$$V_{mm}(r)=\frac{(Z_pe)^2}{r}, \tag{1}$$

where Z_p is the charge on a macroparticle and $2R_p$ is its characteristic size.

The interaction between macroparticles and microparticles (electrons and ions) is described by a potential V_{mc} such that $V_{mc}(r)=\infty$ for $r\leqslant R_p+\sigma_\alpha$, while for $r>R_p+\sigma_\alpha$

$$V_{mm}(r)=\frac{(Z_pq_\alpha)e^2}{r}, \tag{2}$$

where $q_\alpha e$ is the charge on a microparticle of type α, and σ_α is its characteristic size (for example: $\alpha=1$ and $q_\alpha=-1$ for electrons, $\alpha=2$ and $q_\alpha>0$ for ions).

The interaction between microparticles is described by a potential $V_{\alpha\gamma}$ for which $V_{\alpha\gamma}(r)=\infty$ for $r\leqslant\sigma_\alpha+\sigma_\gamma$, while for $r>\sigma_\alpha+\sigma_{\gamma\alpha}$

$$V_{\alpha\gamma}(r)=\frac{(q_\alpha q_\alpha)e^2}{r}. \tag{3}$$

The macroparticle, electron, and ion densities n_p, n_1, and n_2 should satisfy the condition of electrical neutrality:

$$Z_pn_p+q_2n_2=q_1n_1. \tag{4}$$

Because of the strong charge and density asymmetries $Z_p\gg q_1$, $Z_p\gg q_2$, $q_1\approx q_2$, $n_p\ll n_1$ and $n_p\ll n_2$, $n_2q_2\ll n_1q_1$, and also the strong interaction between particles, the primitive model is very difficult to use, both in constructing a first-principles analytic theory and in numerical simulations. At this time the most promising theoretical approaches are based on describing the interaction between macroparticles in terms of effective pseudopotentials that arise from averaging (integration) over microparticle coordinates. One of the most natural ways to introduce effective pseudopotentials for interaction between macroparticles is to generate an effective Hamiltonian by using the expression[24]

$$H_{\text{eff}}=-kT\,\ln\left\langle\exp\left(-\frac{H}{kT}\right)\right\rangle, \tag{5}$$

where H is the full Hamiltonian of the primitive model, k is Boltzmann's constant, and the brackets $\langle\ \rangle$ denote canonical averaging over coordinates and momenta of the microparticles. The effective Hamiltonian H_{eff} can have the following form:

$$H_{\text{eff}}=K_m+U_{mm}(\{R_j\})+F([\rho_c(r);\{R_j\}]), \tag{6}$$

where K_m is the macroparticle kinetic energy,

$$U_{mm}(\{R_j\})=\sum_{i<j}V_{mm}(|R_i-R_j|/2R)$$

is the potential energy of the interaction between macroparticles based on the potential (1), and $\{R_j\}$ is the set of coordinates that describe the configuration of the macroparticles. In this expression, the functional F describes the free energy of the microparticles in the external field of the macroparticles. It is known[23] that F can be represented as a functional of the local charged-particle density $\rho_c(r)$, which depends parametrically on the positions of the macroparticles $\{R_j\}$. The functional F can be written in the form of a sum of four terms as follows:

$$F=F_{id}+F_{\text{ext}}+F_{cc}+F_{\text{cor}}, \tag{7}$$

where

$$F_{id}=kT\int dr\ \rho_c(r)[\ln(\lambda_c^3\rho_c(r))-1], \tag{8}$$

$$F_{\text{ext}}=\int dr\ \rho_c(r)V_{\text{ext}}(r,\{R_j\})=\sum_j \int dr\ \rho_c(r)\times V_{mc}(|r-R_j|), \tag{9}$$

$$F_{cc}=\frac{(q_1e)^2}{2}\int dr'dr\ \frac{\rho_c(r)\rho_c(r')}{|r-r'|}; \tag{10}$$

here λ_c is the thermal wavelength of the microparticles, F_{ext} describes the interaction between micro- and macroparticles, and F_{cc} is the contribution of the average field arising from the Coulomb repulsion between electrons. Unfortunately, no explicit expression is known for the term F_{cor} that describes the correlation between macroparticles; therefore, various approximations are used in the literature. One rather promising approximation is the expression

$$F_{\text{cor}}=kT\int dr\ \rho_c(r)\Psi_{OCP}^{\text{exc}}(T,\rho_c(r)), \tag{11}$$

where $\Psi_{OCP}^{\text{exc}}(T,\rho_c(r))$ is the contribution to the free energy per unit volume for the one-component plasma model, which is known from Monte Carlo calculations. It is obvious that the functional F depends nonlinearly on the coordinates $\{R_j\}$.

In Refs. 24 and 25, the authors minimize the functional F and find the effective interaction potential between macroparticles by using the numerical Carparrinello method. Their calculations showed that over a wide range of temperatures and densities the two terms on the right side of Eq. (6), which depend on the macroparticle coordinates and which describe many-particle interactions, can be reliably approximated by a sum of pairwise additive terms, each of which in turn is well approximated by the Yukawa potential. Thus, we can use a screened Coulomb potential of the Debye–Hückel type as an effective pairwise potential to describe the interaction between macroparticles, in which the characteristic scale of the screening and the interaction parameter have a more complicated dependence on density and temperature than in the Debye case. These dependences are different for different phase states of the system.[7,24–26]

By virtue of the latter fact, a first-principles investigation of the phase diagram requires very careful and extensive calculations. The best approach is probably to use specially-designed numerical methods combined with comparisons and analysis of experimental data. In this paper, the model we adopt, which matches the conditions of the experiments, is one in which the macroparticles interact through a screened Coulomb potential (the Debye model or the Yukawa model):

$$U(r)=\frac{(Z_pe)^2\exp(-r/r_D)}{r}. \tag{12}$$

Justification for our use of the Yukawa model is also provided by analysis of the experimental data given in Ref. 21.

3. CRITERIA FOR CRYSTALLIZATION

In the literature, authors have used various phenomenological criteria to define the transition of a system to a crystalline state.[7,23,25–32] The best-known criterion is that of Lindemann,[26–28] according to which a solid phase melts when the ratio of the mean-square displacement of a particle to the mean interparticle distance begins to exceed 0.1 (although for different physical systems this number sometimes varies between 0.05 and 0.2). Other criteria are formulated in Refs. 29–32. For example, it has been shown that the first maximum of the structure factor $S(\mathbf{q})$ of a liquid on its crystallization curve, where

$$S(\mathbf{q})=1+n_p\int d\mathbf{r}[g(\mathbf{r})-1]\exp(-i\langle\mathbf{q}\cdot\mathbf{r}\rangle), \tag{13}$$

is constant and equal to 2.85. Here $g(r)$ is the binary correlation function, which is the probability for finding particles separated by a distance $r=|\mathbf{r}|$.

A simple criterion for crystallization based on the binary correlation function can be formulated by considering the ratio of the minimumm value $g_{\min}$ to the maximum value $g_{\max}$ of this function:[27–29]

$$R_g=\frac{g_{\min}}{g_{\max}}. \tag{14}$$

The transition from liquid to crystalline state corresponds to values of this parameter equal to 0.2.

4. EXPERIMENT

Our experimental setup, which includes a generator of plasma and diagnostic apparatus for determining the parameters of the particles and the gas, is described in detail in Refs. 21, 33. The source of plasma (a two-flame propane-air Mecker burner) creates a laminar jet of plasma with uniformly distributed parameters (temperatures, electron and ion concentrations) in the region interior to the jet. The pressure of the combustion products is atmospheric. The diameter of the plasma jet was 25 mm, with a velocity V_g that varied in the range 2–3 m/s. The concentrations of electrons and ions in the plasma are in the range 10^9-10^{10} cm^{-3}, with equal electron and ion temperatures: $T_i=T_e=T_g$. Spectrometer measurements of the particle temperature[34] show that the latter is close to the gas temperature $(T_p=T_g)$.

Our object of study was a weakly ionized thermal plasma with CeO_2 macroparticles. Cerium dioxide was chosen as a material for the macroparticles both because of its chemical inertness and because of the small amount of work needed to liberate thermal electrons ($\sim$2.75 eV, see Ref. 35) from the surface of heated CeO_2 particles. Particles of cerium oxide powder contain impurities of alkali metal compounds. As a result, spectral measurements in the plasma jet detected the presence of sodium atoms with a low ionization potential. Thus, the basic components of the plasma were charged particles of CeO_2, electrons, and singly-charged Na^+ ions.

In order to study Coulombic ordered structures in this plasma it is necessary to have data both on the particle charge and on the basic plasma parameters. Our experiments

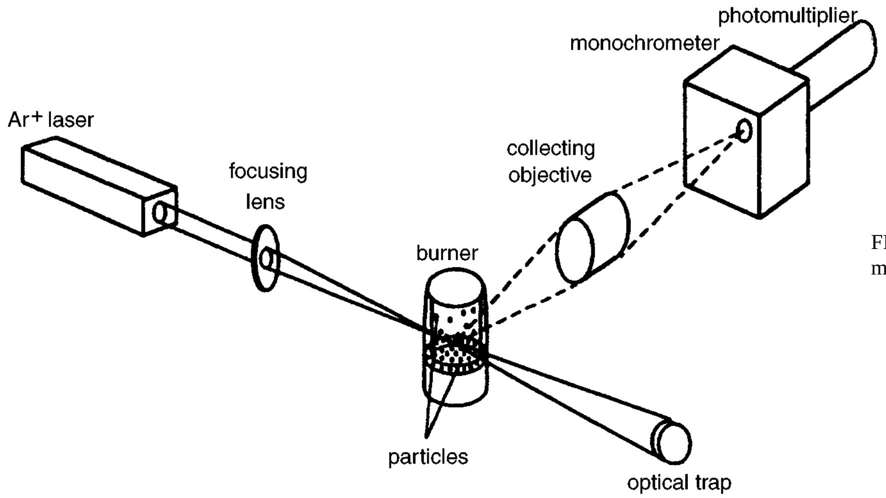


FIG. 1. Sketch of optical measurements of spatial macroparticle structures.

determined the particle sizes, their densities, the plasma temperature, the local electron density, and the density of sodium ions.[21,33,36–38]

Plasma diagnostics were carried out both by probe and optical methods. The density n_i of positive alkali metal ions was measured by the electric probe method.[36,37] The method used to determine the local electron density n_e was based on measuring the current I and the longitudinal electric field intensity E in the plasma.[36] The temperature of the gas and the density of alkali metal atoms were measured by traditional methods: the generalized reversal method and the method of total absorption.[33]

In order to determine the average (Sauter) diameter D_p and the macroparticle density n_p in the plasma jet we used a novel laser method.[38] The method is based on measurement of the extinction (attenuation of light) in a dispersive medium at small scattering angles, and is designed to determine the characteristics of particles in the size range 0.5–15 μm. Measurement errors of the extinction of about 2% give rise to errors in finding the particle sizes of about 3%, and errors in the density of about 10%.

In order to measure the spatial positions of the macroparticles we used a laser time-of-flight detector.[21] The measurement volume for the time-of-flight detector is formed by focusing the beam of an argon laser ($\lambda = 0.488\ \mu$m) in a given region of the plasma jet (Fig. 1). The pulses of scattered light from individual particles are converted by a photomultiplier into electric signals. These signals are processed to calculate the correlation function $g(r)$, where $r = V_p t$. Here V_p is the average velocity of the particles ($V_p \approx V_g$) and t is the time. In what follows, the binary correlation function $g(r)$ will be used to analyze the spatial structure of the particles.

In measurements with CeO_2 particles, the particle density n_p varied in the range $(0.2-5.0)\times 10^7\ cm^{-3}$, the temperature of the plasma T_g in the range 1700–2200 K. Consequently, the ion density n_i varied from $0.4\times 10^{10}\ cm^{-3}$ to $4.0\times 10^{10}\ cm^{-3}$, and the electron density n_e from $2.5\times 10^{10}\ cm^{-3}$ to $8\times 10^{10}\ cm^{-3}$. The average diameter D_p of a CeO_2 particle is about 0.8 μm.

The measurements of the spatial structures of macroparticles were compared with the results obtained for an aerosol jet at room temperature. In this case, the interior flame of the burner contained only air and CeO_2 particles. This system simulates a plasma with a random (chaotic) spatial distribution of macroparticles (a gaslike plasma).

Figure 2 shows typical binary correlation functions $g(r)$ for CeO_2 particles in an aerosol jet at room temperature ($T_g \cong 300$ K) and in a plasma ($T_g = 2170$ K and $T_g = 1700$ K). It is easy to see that the correlation functions

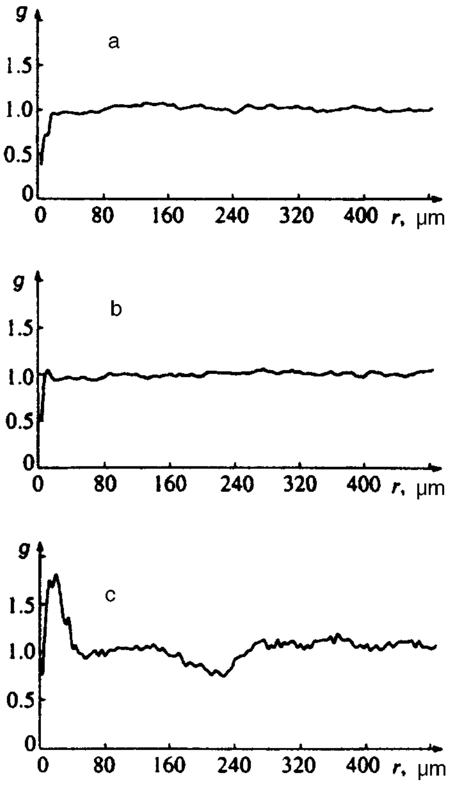


FIG. 2. Binary correlation function $g(r)$ for CeO_2 particles in an air jet at room temperature $T_g \cong 300$ K and $\gamma_p = 0$ (a), in a plasma ($Z_p = 500$) at a temperature $T_g = 2170$ K, $\gamma_p = 40$ and $\kappa = 3.5$ (b), and at a temperature $T_g = 1700$ K, $\gamma_p = 120$, and $\kappa = 1.6$ (c).

$g(r)$ for the aerosol jet and for the plasma at a temperature $T_g = 2170$ K and particle density $n_p = 2.0 \times 10^6$ cm^{-3} are practically the same (Figs. 2a, b). We therefore conclude that particles in the plasma are weakly interacting and the formation of ordered structures is impossible. This is also confirmed by independent diagnostic measurements (probe and optical) made in the plasma. Thus, the values of the interaction parameter γ_p computed from the measurement results, and the parameter $\kappa = \langle r \rangle / r_D$ that takes into account Debye screening, are respectively 40 and 3.5.

At a plasma temperature of about 1700 K and CeO_2 particle density of order 10^7 cm^{-3}, analysis of the binary correlation function $g(r)$ reveals short-range order, which is characteristic of liquids (Fig. 2c). That is, the particles form an ordered structure.[21] Under these conditions the ion density ($n_i = 4.2 \times 10^9$ cm^{-3}) is roughly an order of magnitude lower than the electron density ($n_e = 7.2 \times 10^{10}$ cm^{-3}). The charge of the particles obtained from the quasineutrality condition in the form $Z_p n_p = n_e$ ($n_i \ll n_e$) is positive, with a value of about $10^3 e$. This value is accurate to within a factor of 2, which can be explained by the thermal emission of electrons from the particle surface.[1,2,35] The parameters γ_p and κ obtained from these diagnostic measurements had values $>$ 120 and 1.6 respectively, indicating a system of weakly interacting particles. According to criteria of Refs. 2, 7, 19, this implies that a gas–liquid phase transition is occurring.

5. NUMERICAL RESULTS

The plasma with macroparticles was numerically simulated using the Monte Carlo method.[39] For the calculations the following parameters for the plasma with CeO_2 particles were chosen: macroparticle density $n_p = 5.0 \times 10^7$ cm^{-3}, electron density $n_e = 7.2 \times 10^{10}$ cm^{-3}, ion density $n_i = 0.42 \times 10^{10}$ cm^{-3}, and plasma temperature $T_g = 1700$ K. The Debye radius r_D and average distance between particles in this case were 11 μm and 17 μm respectively. The electron density and ion density in the plasma was roughly three orders of magnitude larger than the macroparticle density. Therefore, in simulating the plasma the Monte Carlo calculations can be carried out only after introducing an effective interaction potential between the macroparticles, which arises, as we mentioned above, from averaging over the positions of electrons and ions. In this context it is worth recalling that the question of the form of the effective potential is not finally resolved; however, at this time the Debye potential (12) is the most reliable choice.

In these calculations it is convenient to pick the Debye radius r_D as a unit of length, where this radius depends on the electron and ion density of the plasma. Then (12) becomes

$$U(r) = \frac{(Z_p e)^2 \exp(-r)}{r_D r}, \tag{15}$$

where we have introduced the dimensionless distance $r = r/r_D$ from the macroparticle.

The correlation parameter $\Gamma_p = Z_p^2 e^2 / k T_g r_D$ can be written in the form $\Gamma_p = \gamma_p \langle r \rangle / r_D$. Under conditions characteristic of these experiments, i.e., $\langle r \rangle = 17\ \mu$m and $r_D = 11\ \mu$m, we obtain $\gamma_p \simeq 120$ and $\Gamma_p = \gamma_p \langle r \rangle / r_D \cong 185$ for $Z_p = 500$, and $T_g = 1700$ K.

As an alternate model for analysis and comparison of numerical and experimental results, we chose the model of a single-component plasma in which the macroparticles interact via the Coulomb potential (1) superposed on a uniform compensating background of the opposite charge. As is well known from the literature,[5,8] the model of a single-component plasma has been studied in considerable detail by Monte Carlo methods. For this reason it is expedient to compare the numerical results for both models. In doing this, we must keep in mind that the model of a single-component plasma has the property of similarity, which implies that all the results depend only on the single dimensionless parameter γ_p. However, in the Debye model the results depend on two dimensionless parameters, namely, γ_p and $\kappa = \langle r \rangle / r_D$. Numerical results within the Debye model were obtained for the two values $\kappa = 1.0$ and 2.0. These values correspond to the experimental conditions, and allow us to analyze thoroughly the tendency for the computed quantities to change when we compare with the results of the single-component plasma model.

In Monte Carlo calculations we usually consider a finite number of particles N located in a cell of size L with periodic boundary conditions. Current speeds of available computers and the requirement that computation times be reasonable (one point per day) restricts the problems that can be solved to those for which values of N range up to 125. Thus, in our case the size of the cell L, which is also conveniently measured in units of the Debye radius, equaled

$$\frac{L}{r_D} = \left(\frac{N}{n_p r_D^3} \right)^{1/3} = \left(\frac{4\pi}{3} N \right)^{1/3} \frac{\langle r \rangle}{r_D}. \tag{16}$$

In the calculations with $\kappa = 1.0$ and 2.0, the quantity L/r_D was equal to 8.05 and 16.1 respectively, and the interaction parameter varied over wide limits. Let us begin our analysis of the numerical results with a discussion of the binary correlation functions $g(r)$ obtained within the one-component model. Figure 3 shows the macroparticle $g(r)$ for values of the parameter γ_p from 1 to 140.[8] Note that in this model crystallization occurs at $\gamma_p = 171$, and that the binary correlation function $g(r)$ for $\gamma_p = 171$ satisfies the crystallization criteria mentioned above.

The results of calculating $g(r)$ for the Yukawa model and analogous values of the parameter Γ_p for values of $\kappa = 1.0$ and 2.0 are shown in Figs. 4 and 5. Comparison and analysis of Figs. 3, 4, and 5 shows that within the framework of the Debye model $g(r)$ is close to the analogous $g(r)$ obtained within the framework of the one-component model; for $\kappa = 1.0$ this is quite natural, since as $r_D \to \infty$ the Debye model reduces to the one-component plasma model. Increasing the parameter Γ_p to 1000 for $\kappa = 1.0$ (which significantly exceeds values of Γ_p obtained under experimental conditions) does not lead to crystallization of macroparticles according to this criteria. Evidence of this is the ratio $R_g > 0.2$ and the behavior of the computed structure factor

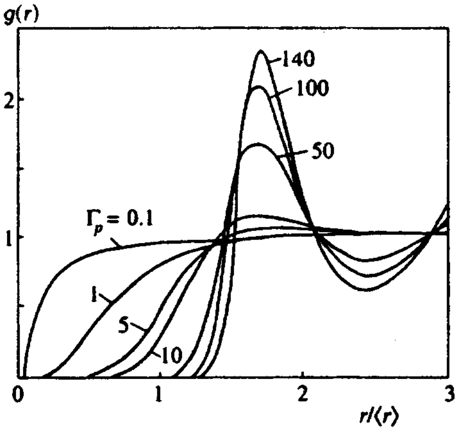


FIG. 3. Binary correlation functions $g(r)$ of the one-component plasma model for various values of Γ_p.

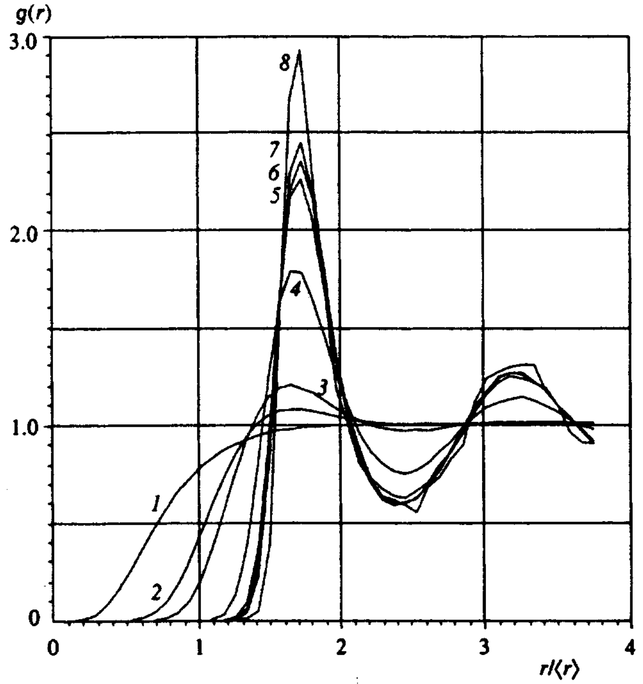


FIG. 5. Binary correlation functions $g(r)$ of the Debye plasma model for $\kappa=1$. $\Gamma_p=1.5$ (*1*), 7.5 (*2*), 15 (*3*), 75 (*4*), 150 (*5*), 185 (*6*), 210 (*7*), and 1000 (*8*).

shown in Fig. 6. The maximum value of $S(q)$ at $\Gamma_p=1000$ is smaller than 2.5, whereas according to Ref. 29 it should reach 2.85 on the crystallization curve of $S(q)$.

Comparison of these numerical results for $g(r)$ with experiment is complicated by the fact that the waist of the laser beam has a diameter several times smaller than the interparticle distance. Therefore, the diagnostic methods used for these ordered structures in the plasma give us information not about the correlation function itself, but rather about the correlation function averaged over the measurement volume formed by the focusing of the laser beam.

In order to estimate an effective size d_{eff} that takes into account the diameter of the laser constriction, we note that when the distance from such a particle is smaller than d_{eff} the probability of finding another particle reduces to zero. This is valid for noninteracting particles when $d_{\mathrm{eff}} \gg D$ holds. Therefore, in order to estimate d_{eff} we use measurements of the correlation function for CeO_2 particles in an air jet (Fig. 2b). The effective size d_{eff} lies in the interval within which the value of the correlation function equals zero. According to

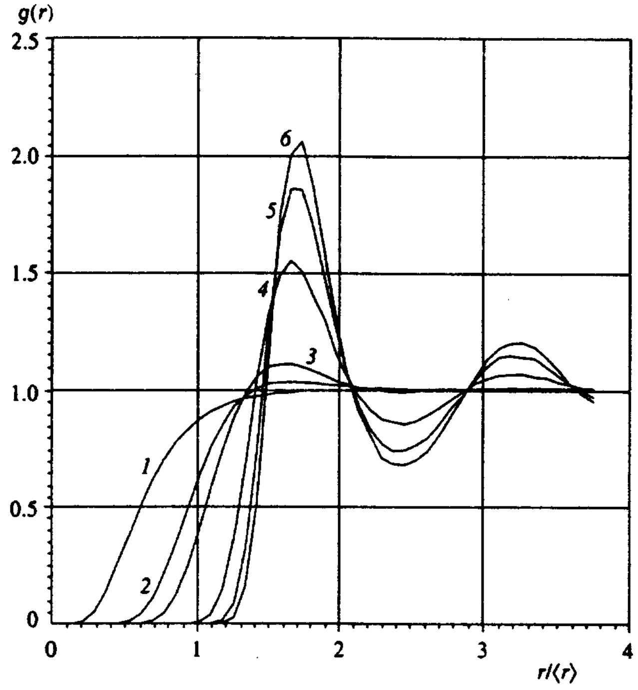


FIG. 4. Binary correlation functions $g(r)$ of the Debye plasma model for $\kappa=2$. $\Gamma_p=$(*1*), 7.5 (*2*), 15 (*3*), 75 (*4*), 150 (*5*), and 210 (*6*).

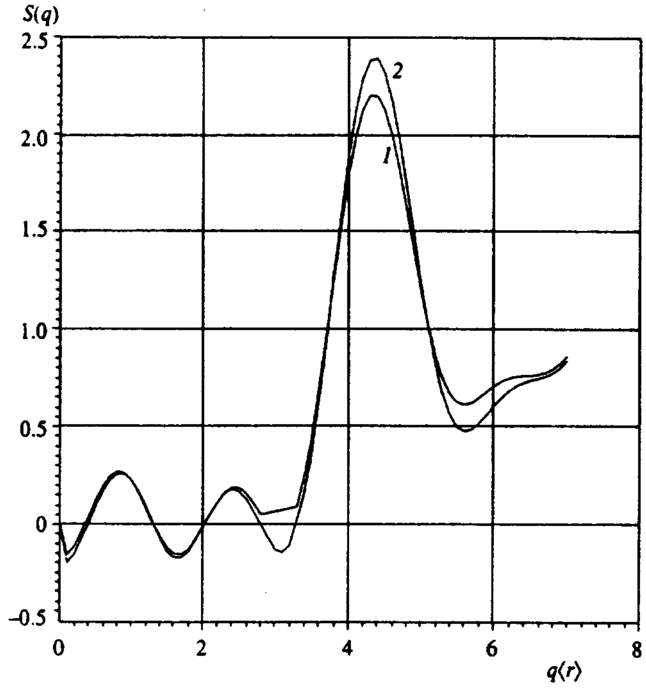


FIG. 6. Structure factor for the Debye plasma model with macroparticles for $\kappa=1$. $\Gamma_p=216$ (*1*) and 1000 (*2*).

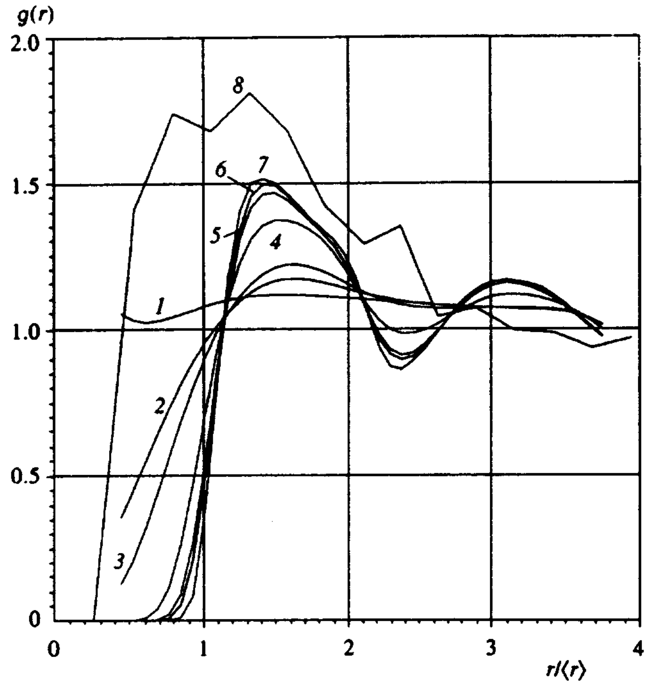


FIG. 7. Averaged theoretical and experimental binary correlation functions $g(r)$ of the Debye plasma model for $\kappa=1$. $\Gamma_p=1.5$ (*1*), 7.5 (*2*), 15 (*3*), 75 (*4*), 150 (*5*), 216 (*6*), and 1000 (*7*); (*8*)—experimental binary correlation function.

our estimates, the size d_{eff} of the region of averaging is roughly 5 μm.

In order to convert to the corresponding averaged correlation function we must go from the probability density to probability distributions obtained by the Monte-Carlo method, for which $g(r)$ must be multiplied by the volume of a spherical layer of thickness dr. By integrating this over a spherical layer of thickness d_{eff} and dividing the results by its volume, we obtain the averaged correlation functions $\langle g(r)\rangle$. The functions obtained are shown in Fig. 7. In this same figure we show the experimental binary correlation function (curve *8*) obtained under the experimental conditions listed above (Fig. 2c). Comparison of these figures shows that the agreement between numerical and experimental results is rather good. These results are also in agreement with the calculations of the binary correlation functions given in Refs. 9, 39. The discrepancies, in particular the broader experimentally obtained correlation function, could be due to the coexistence of regions with random positions of the particles and regions with an ordered structure (domains). In the latter case, the particles could be located at positions smaller than $\langle r\rangle$.[40]

6. CONCLUSION

In this work we have compared experimental measurements of spatially ordered macroparticle structures in the volume of a thermal plasma with the results of numerical Monte Carlo calculations based on the Yukawa model. Ordered structures of macroparticles are observed to form in a plasma consisting of positively charged micron-size CeO_2 particles, electrons emitted by them, and singly charged sodium ions at atmospheric pressure and temperatures of about 1700 K. Analysis of the theoretical and experimental binary correlation functions $g(r)$ show that the latter are in good agreement after averaging the theoretical $g(r)$ over the measurement volume formed by focusing of the laser beam. Numerical results also confirm that under the conditions of the experiment no crystallization of macroparticles takes place in the plasma. However, a gas–liquid phase transition is clearly observed, along with spatial ordering of the macroparticles associated with strong interparticle interactions. This assertion is based on analysis of the computed correlation functions and verification that the two criteria for crystallization mentioned above ($R_g>0.2$ and $S(q)>2.85$) are fulfilled.

This work was carried out with the partial financial support of the Russian Fund for Fundamental Research (Grant No. 95-02-04561) and of an INTAS-RFBR program (Grant No. 95-1335).

[1] M. S. Sodha and S. Guha, Adv. Plasma Phys. **4**, 219 (1971).
[2] V. E. Fortov and T. Yakubov, *Physics of Nonideal Plasma*, Hemisphere, London (1990).
[3] T. M. Sugden and B. A. Thrush, Nature **168**, 703 (1951).
[4] K. E. Shuler and J. Weber, J. Chem. Phys. **22**, 491 (1954).
[5] S. Ichimaru, Rev. Mod. Phys. **54**, 1017 (1982).
[6] H. Ikezi, Phys. Fluids **29**, 1764 (1986).
[7] M. O. Robbins, K. Kremer, and G. Grest, J. Chem. Phys. **88**, 3286 (1988).
[8] W. L. Slattery, G. D. Doolen, and H. E. DeWitt, Phys. Rev. A **21**, 2087 (1980).
[9] R. T. Farouki and S. Hamaguchi, Appl. Phys. Lett. **61**, 2973 (1992).
[10] R. F. Wuerker, H. Shelton, and R. V. Langmuir, J. Appl. Phys. **30**, 342 (1959).
[11] S. L. Gilbert, J. J. Bollinger, and D. J. Wineland, Phys. Rev. Lett. **60**, 2022 (1988).
[12] T. Trottenberg, A. Melzer, and A. Piel, Plasma Sources, Sci. and Technol. **4**, 450 (1995).
[13] J. H. Chu and L. I, Phys. Rev. Lett. **72**, 4009 (1994).
[14] H. Thomas, G. E. Morfill, V. Demmel *et al.*, Phys. Rev. Lett. **73**, 652 (1994).
[15] Y. Hayashi and K. Tachibana, J. Appl. Phys. **33**, L804 (1994).
[16] A. Melzer, T. Trottenberg, and A. Piel, Phys. Lett. A **191**, 301 (1994).
[17] A. Rahman and J. P. Schiffer, Phys. Rev. Lett. **57**, 1133 (1986).
[18] D. H. E. Dubin and T. M. O'Neil, Phys. Rev. Lett. **60**, 511 (1988).
[19] M. J. Stevens and M. O. Robbins, J. Chem. Phys. **98**, 2319 (1992).
[20] X. H. Zheng and J. C. Earnshaw, Phys. Rev. Lett. **75**, 4214 (1995).
[21] V. E. Fortov, A. P. Nefedov, O. F. Petrov *et al.*, JETP Lett. **63**, 7 (1996).
[22] P. Meaking and A. T. Skjeltorp, Adv. Phys. **42**, 1 (1993).
[23] H. Lowen, Phys. Reports **237**, 249 (1994).
[24] H. Lowen, P. A. Madden, and J. P. Hansen, Phys. Rev. Lett. **68**, 1081 (1992); J. Chem. Phys. **98**, 3275 (1993).
[25] H. Lowen and G. Kramposthuber, Europh. Lett. **23**, 673 (1993).
[26] E. J. Meijer and D. Frenkel, J. Chem. Phys. **94**, 69 (1991).
[27] F. Lindemann, Z. Phys. **11**, 609 (1910).
[28] A. R. Ubbelohnde, *The Molten State of Matter* (J. Wiley, Chichester, 1978).
[29] J. P. Hansen and L. Verlet, Phys. Rev. Lett. **184**, 151 (1969).
[30] H. Lowen, T. Palberg, and R. Simon, **70**, 1557 (1993).
[31] A. V. Indrani and S. Ramaswamy, Phys. Rev. Lett. **73**, 360 (1994).

[32] M. Fuchs, Phys. Rev. Lett. **74**, 1490 (1995).
[33] A. B. Kondrat'ev, A. P. Nefedov, O. F. Petrov, and A. A. Samaryan, Teplofiz. Vys. Temp. **32**, 452 (1994) [in Russian].
[34] A. P. Nefedov, O. F. Petrov, and O. S. Vaulina, Spect. Radiat. Transfer **54**, 453 (1995).
[35] V. S. Fomenko, *Handbook of Thermionic Properties,* Plenum, New York (1966).
[36] V. F. Kosov, V. I. Molotkov, and A. P. Nefedov, Teplofiz. Vys. Temp. **29**, 633 (1991) [in Russian].
[37] M. S. Benilov, Teplofiz. Vys. Temp. **26**, 993 (1988) [in Russian].
[38] O. S. Vaulina, A. P. Nefedov, O. F. Petrov *et al.*, Zh. Prikl. Spekt. **63**, 133 (1996) [in Russian].
[39] V. M. Zamalin, G. E. Norman, and V. S. Filinov, *The Monte Carlo Method in Statistical Thermodynamics* (Nauka, Moscow, 1977).
[40] V. E. Fortov, A. P. Nefedov, O. F. Petrov *et al.*, Phys. Lett. A **219**, 89 (1996).

Translated by Frank J. Crowne

Interaction of a high-power laser beam with low-density porous media

A. E. Bugrov, S. Yu. Gus'kov, and V. B. Rozanov

P. N. Lebedev Institute of Physics, Russian Academy of Sciences, 117924 Moscow, Russia

I. N. Burdonskiĭ, V. V. Gavrilov, A. Yu. Gol'tsov, E. V. Zhuzhukalo, N. G. Koval'skiĭ, M. I. Pergament, and V. M. Petryakov

Troitsk Insitiute for Innovative and Thermonuclear Studies, 142092 Troitsk, Moscow Region, Russia

(Submitted 26 June 1996)
Zh. Éksp. Teor. Fiz. **111**, 903–918 (March 1997)

We have experimentally investigated the processes of laser light absorption and energy transfer in porous targets made of ''agar-agar'' ($C_{14}H_{18}O_7$) with an average density of 1–4 mg/cm^3 illuminated by the focused beam of a neodymium laser with an intensity of 10^{14} W/cm^2 within a pulse of duration 2.5 ns. Many important scientific and technical problems, e.g., inertial-confinement thermonuclear fusion, the creation of lasers in the x-ray regime, and the modeling of astrophysical phenomena under laboratory conditions, can be successfully addressed by using low-density porous media as components of such targets. In our experiments with porous targets of variable density and thickness we used optical and x-ray diagnostic methods, which ensured that our measurements were made with high temporal and spatial resolution. We show that a region forms within the porous target consisting of a dense high-temperature plasma which effectively absorbs the laser radiation. Energy is transferred from the absorption region to the surrounding layer of porous material at up to 2×10^7 cm/s. Experimental data are in good agreement with the predictions of our theoretical model, which takes into account the specific features of absorption of laser radiation in a porous material and is based on representing the energy transfer within the material as a hydrothermal wave. © *1997 American Institute of Physics.* [S1063-7761(97)01203-1]

1. INTRODUCTION

The use of low-density porous media as components of various types of targets illuminated by high-power laser pulses is a very promising approach to the study of the physics of interaction between a laser beam and matter, and can lead to the successful resolution of a number of important scientific and technical problems. Among these are inertial-confinement thermonuclear fusion, the creation of sources of coherent radiation in the x-ray region of the spectrum, simulation of astrophysical phenomena under laboratory conditions, and finally fundamental studies of the properties and states of matter under dynamic loading in the megabar pressure range.

Continuing efforts to initiate thermonuclear fusion in microtargets by high-power laser beams have led to considerable refinement of the fundamental requirements on the parameters of laser radiation and target construction. The use of complex targets made with materials whose average densities are two or three orders of magnitude smaller than normal solid-state densities can in principle ensure symmetric compression of the thermonuclear fuel and implementation of thermonuclear ignition with significantly more relaxed requirements on the parameters of the laser system and conditions for illumination. In the direct-illumination scheme, it is possible to significantly alter the spatial density distribution of the plasma corona that forms at the illuminated surface of a spherical thermonuclear target by depositing on the surface a layer of foam material with suitably chosen thickness and initial density made from elements with low atomic numbers. Use of this procedure appears to lead to efficient smoothing out of nonuniformities in the illumination at high values of the ablation pressure at the target surface. In the indirect-illumination scheme, a low-density material placed at the inner surface of a shell-converter made of heavy elements and illuminated by laser beams can confine the expanding reradiating plasma, thereby ensuring the required uniformity of the distribution of x-ray intensity at the surface of the thermonuclear target. A porous medium is a fundamentally necessary element in the construction of promising targets of the ''laser greenhouse'' type as well.[1] A further attractive possibility is the option of introducing a small amount of various elements into the porous target, thereby influencing the spectral composition of the x-ray emission from the plasma.

A fundamental feature of porous low-density targets is their nonuniformity. The structure of these materials is, as a rule, a disordered aggregate of particles of various shapes and normal solid-state densities together with pores. The absorption of high-power laser radiation, the mechanism for energy transport, and the fluid dynamic processes in media with discrete structures can have a very distinctive character and has recently attracted more and more interest and attention.[2]

In contrast to the classical process of light scattering in a dispersive medium when the latter is illuminated by a laser beam with intensity exceeding 10^{12} W/cm^2, in a porous medium heating and expansion of the structural elements takes place, causing the dimensions of the structure to increase.

This in turn causes the length over which absorption is effective to decrease with time.

Consider a case where the average density of electrons in the porous target is below the critical density for the given wavelength of the laser light. The length of the interaction region, which is related to the overlap of the laser beam cross section by the expanding dense elements of the porous material, will decrease until the particles of the material have dispersed to below a critical density. After this, a region of bulk absorption of the laser radiation forms, whose longitudinal size (along the direction of the laser beam) is determined by classical collision mechanisms:

$$L_T=\frac{9.2\times10^{-8}}{Z}\left(\frac{A}{Z}\right)^2\frac{T^{3/2}}{\lambda^2\rho^2}.$$

Here A, Z are the atomic number and charge of the plasma ions respectively, λ is the wavelength of the laser light (in μm), T is the electron temperature (in keV), and ρ is the plasma density (in g/cm^3). Finally, after total internal vaporization of the porous material (at which point the density of the plasma becomes equal to the average initial density), the size of the region of bulk absorption will depend only on the plasma temperature. (We assume that the pulse persists until all these processes are able to proceed to their completion.)

What kind of mechanism should operate to transport energy from the bulk absorption element to the cooler inner layers of the target? An electron thermal conductivity wave should form only after the pores in the material are filled with plasma during the internal vaporization of the solid elements. However, under our conditions, energy can be transferred from the zone where the laser radiation is absorbed by a wave that is best referred to as hydrothermal. This wave forms as fluid dynamic plasma currents diffuse through the low-density porous material; behind the wave front, a rapid equalization of the plasma temperature occurs due to electron thermal conductivity. The velocity of a hydrothermal wave is close to the velocity of sound, which for the expected plasma temperatures of order 1 keV comes to $\sim 2\times10^7$ cm/s, whereas the velocity of an electron thermal conductivity wave at the same temperature and densities of $10^{-3}-10^{-2}$ g/cm^3 exceeds $\sim 5\times10^7$ cm/s.

This paper presents experimental and theoretical studies of the interaction of high-intensity laser light with low-density porous media. In Sec. 2 we discuss our experimental conditions, the parameters of the targets used, and also our suite of diagnostic equipment. In Sec. 3 we give the results of experiments and their discussions. Section 4 deals with the development of a theoretical model and comparison of computational results with experimental data. The main conclusions are formulated in Sec. 5.

2. EXPERIMENTAL CONDITIONS AND DIAGNOSTIC APPARATUS

In these experiments on the interaction of high-power laser light with low-density targets we used the Mishen' facility[3] with the following illumination conditions: the optical wavelength was 1.054 μm; the half-width of the laser pulse was $\sim$2.5 ns with a rise time of 0.3 ns; the light was focused onto the target by a lens with a relative aperture of 1:10 (the half-width of the intensity distribution within the focal spot was $\sim$150 μm); and the optical flux density at the surface of the illuminated target was 10^{14}W/cm^2. In these experiments we used planar targets with thicknesses of 100–1000 μm made of ''agar-agar'' ($C_{14}H_{18}O_7$) with densities in the range 1–10 mg/cm^3, and also multilayer targets in which layers of aluminum with thickness 1–6 μm were deposited on the illuminated surface, back surface, or both surfaces of the low-density material. The agar has a chaotic structure made up of fibers with solid-state density and a diameter of 1–2 μm, and a spacing between fibers of 10–50 μm. The technology for fabricating targets from agar was developed at TRINITI by S. F. Medovshchikov and S. L. Nedoseev.[4]

The diagnostic complex we used to investigate the mechanisms for interaction of the laser light with low-density targets and the energy transport in these targets included a number of methods based on measuring the x-ray emission of the plasma, and also optical diagnostic methods. The x-ray methods included time-integrated calorimetric measurements in the wavelength range 0.5–2 nm; imaging of the plasma using a pinhole camera behind various filters for observing both along the surface of the irradiated target (with a spatial resolution $\sim$15 μm) and at an angle of $\sim$30° to the direction of the laser beam (with spatial resolution $\sim$25 μm); and recording of the x-ray emission from the plasma using photodiodes with a time resolution of $\sim$0.5 ns. The optical methods included multiframe shadow photography of the expansion of the plasma as it forms (the exposure time of an individual frame was $\sim$0.3 ns, the wavelength of the probe light was $\lambda=0.53$ μm, and the spatial resolution was $\sim$30 μm); measurements using an Agat-SF photodetector of the time evolution of illumination of the back surface of the irradiated target in the wavelength range 400–700 nm with a spatial resolution $\sim$ 30 μm and time resolution of 50 ps; and time-integrated measurements of the emission spectra scattered into the aperture of a focusing lens. The positioning of the x-ray and optical diagnostics is shown in Figs. 1a, 1b.

3. EXPERIMENTAL RESULTS

Let us first discuss the results of measuring the illumination of the back surface of the target in the visible wavelength range 400–700 nm, spatially (along the direction of the discriminating slit) and temporally resolved measurements made with an Agat-SF electrooptic camera. Figures 2a–2c show measurements of the time evolution of the back-surface illumination of the target, recorded in experiments in which three different types of planar targets made of agar were illuminated. In Fig. 2a, note the first flash, which coincides in time with the start of the laser pulse illuminating the target. This flash is caused by the passage of the laser radiation through the target at the initial stage of the illumination process. We note that we observed an analogous effect in previous experiments on illuminating thin plastic films made of Mylar.[5] As follows from Fig. 2a, the duration of the phase of partial transparency is 200–300 ps. As we might expect, this flash is absent when a layer of aluminum of thickness 1.5 μm is deposited on the back surface of the target (Fig. 2b). In all the image-converter pictures shown we observe an

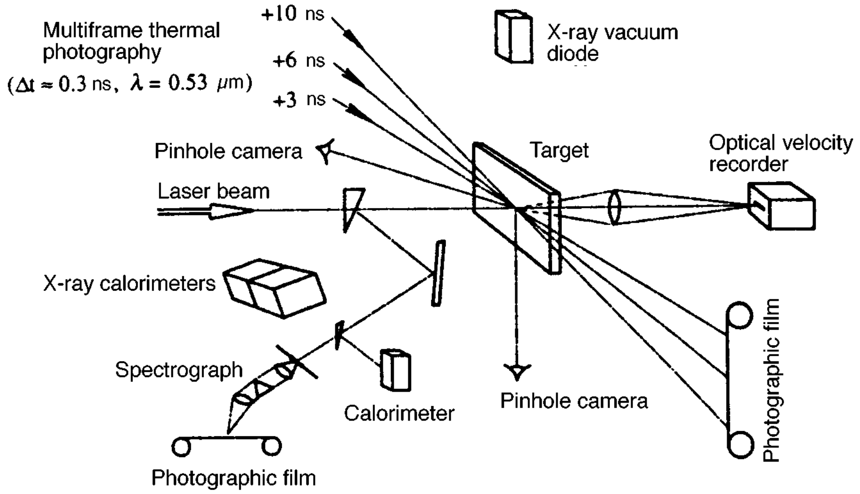


FIG. 1. Sketch of the positioning of diagnostic apparatus.

abrupt growth in the intensity of illumination at the target's back surface with a time delay relative to the start of the laser pulse. The delay in appearance of this flash relative to the beginning of the laser pulse increases with increasing target thickness, as can be seen by comparing Figs. 2a, 2b, and 2c. If we assume that the origin of the flash is associated with energy transport to the back surface of the target from the developing zone of efficient absorption of the laser light, assuming the velocity of energy transfer in the porous medium is constant, then it is easy to obtain an estimate of this velocity v_T, and also to estimate the thickness of the absorption zone Δ. Under our experimental conditions, the results shown in Fig. 2 imply that the corresponding values are $v_T \sim 2\times 10^7$ cm/s and $\Delta \sim 150$ μm (Here and in what follows, we will neglect the relatively small (~ 100 ps) time delay connected with the passage of the shock wave through the aluminum layer of thickness 1.5 μm.)

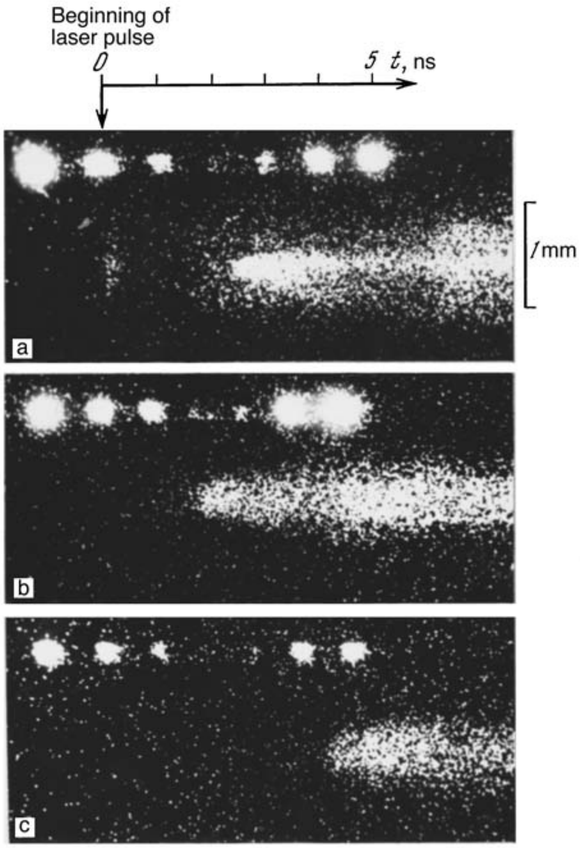


FIG. 2. Time history of the illumination of the back surface of a target made of agar with density 1 mg/cm^3 and thickness 600 μm (a), 500 μm (b), and 1000 μm (c), In cases b and c an aluminum film with thickness 1.5 μm was deposited on the back surface of the target.

The question of what mechanism can transfers energy from the zone of absorption with such a high velocity is a topic for special discussion. One possible mechanism could be the propagation of a shock wave from the high-pressure zone (the zone where the laser light is absorbed) into the interior of the target. In order to test this assumption we set up experiments in which we illuminated three-layer targets made up of a porous medium (agar with a density of ~ 1 mg/cm^3 and thicknesses from 500 μm to 1000 μm) inserted between layers of aluminum at the illuminated and back surfaces of the target (of thickness 6 μm and 1.5 μm respectively). Under these conditions, energy is transferred through the porous medium primarily by propagation of a shock wave. From our previous experiments[6] on the illumination of aluminum foils with thickness 6 μm, we know that for a laser pulse with duration ~ 2.5 ns and an optical flux density of $\sim 10^{14}$ W/cm^2 an aluminum layer with thickness ~ 4 μm accelerates in the direction of propagation of the laser beam and acquires a velocity of (6–8)$\times 10^7$ cm/s. In experiments with a three-layer target, the accelerated and relatively cold (5–10 eV) aluminum plasma plays the role of the piston, which excites a shock wave in the porous material. The experiments show that the rate of propagation of the shock wave in the agar practically coincides with the piston velocity and is a few times smaller than the energy transport velocity (v_T) recorded when a single-layer target made of agar is illuminated. Thus, the shock-wave mechanism is not responsible for the energy transport in one-layer porous targets.

Information about the dimensions and temperature of the plasma wave formed with a target made of agar was obtained by recording x-ray photons in the range of energy 1–1.5 keV with a pinhole camera. In Fig. 3 we show pinhole camera

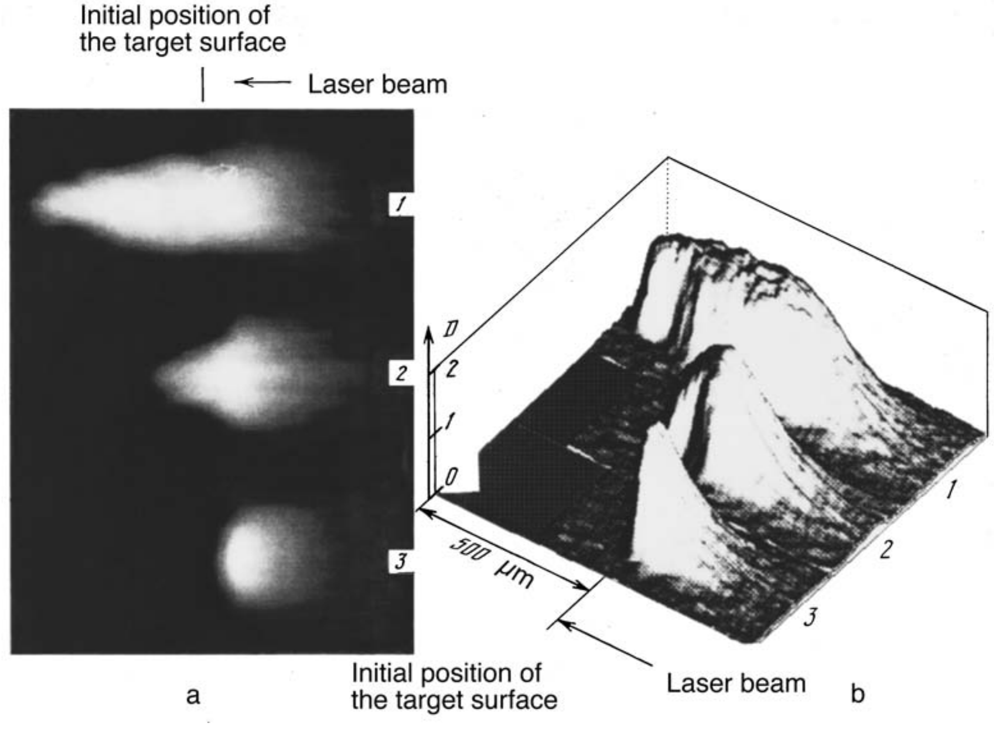


FIG. 3. Pinhole camera picture (a) and the corresponding darkness distribution (b) obtained from illuminating a target made of agar with thickness 500 μm and densities 1 mg/cm^3 (*1*) and 4 mg/cm^3 (*2*), and also during illumination of a target made of Mylar with thickness 10 μm (*3*).

pictures obtained with a beryllium filter 50 μm thick and observations perpendicular to the direction of the laser beam, in experiments where the illuminated targets were of agar with thickness of 500 μm and densities 1 mg/cm^3 (*1*) and 4 mg/cm^3 (*2*). For comparison, the same figure shows a pinhole camera picture obtained by illuminating a target made of Mylar with solid-state density and a thickness of 10 μm (*3*). The results of processing these pinhole camera pictures are shown in Fig. 4. It is clear from this figure that the longitudinal (in the direction of the laser beam) dimensions of the developing high-temperature plasma layer are 400–500 μm for a target with initial density of 1 mg/cm^3 and 150–200 μm for a target with initial density 4 mg/cm^3. The electron temperature of the plasma is determined by the filter method, and for targets of all types was 0.8–1 keV, decreasing somewhat in the direction into the depth of the porous target. For example, when the target density is 1 mg/cm^3 we have $T_e \approx 0.6-0.7$ keV at a distance of 350 μm from the illuminated surface. Note that during illumination of a target with decreased density the intensity of x-ray light in the region of photon energies 1–1.5 keV exceeds the corresponding value obtained in experiments with the Mylar target by a factor of 2–3.

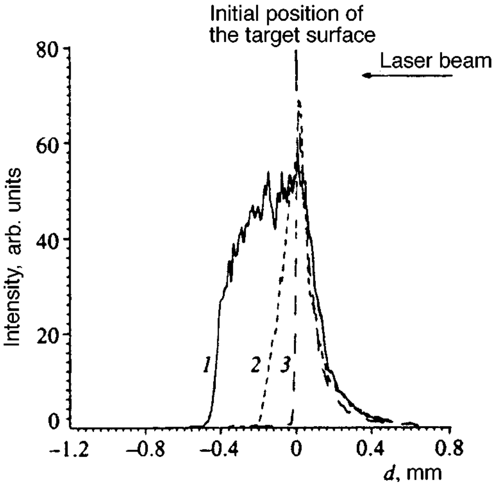


FIG. 4. Distribution of intensity of x rays from the plasma along the direction of the laser beam obtained as a result of processing the pinhole camera oscura pictures shown in Fig. 3.

Taken together, these experimental data are in good agreement with the results of multiframe shadow photography of porous targets illuminated by a high-power laser beam. Figures 5a, 5b show shadow photographs recorded in experiments on an agar target with thickness 500 μm and density 1 mg/cm^3 (a) and a Mylar film with thickness 10 μm (b). It is clear that the overall pictures of the plasma dispersal differ significantly for these cases. For the illuminated Mylar film the transverse size of the high-pressure zone at the back surface of the target practically coincides with the diameter of the focal spot of the laser beam (Fig. 5b), whereas for the low-density target this dimension, as is apparent from Fig. 5a, is several times larger than the diameter of the focal spot. We find that practically all the material of the porous target is in motion (see Fig. 5a) only three nanoseconds after the start of the illumination. Figure 6 shows the results of shadow photographs of the process of plasma expansion in experiments where agar targets with density 1 mg/cm^3 and two different thicknesses are illuminated; at the back surface of both targets aluminum layers were deposited with thicknesses of 1.5 μm. These photographs allows us to measure the rate of directional motion of the aluminum layer and estimate the pressure (energy density) at the back surface of the porous material, which for the case of targets made of agar with thickness 500 μm and

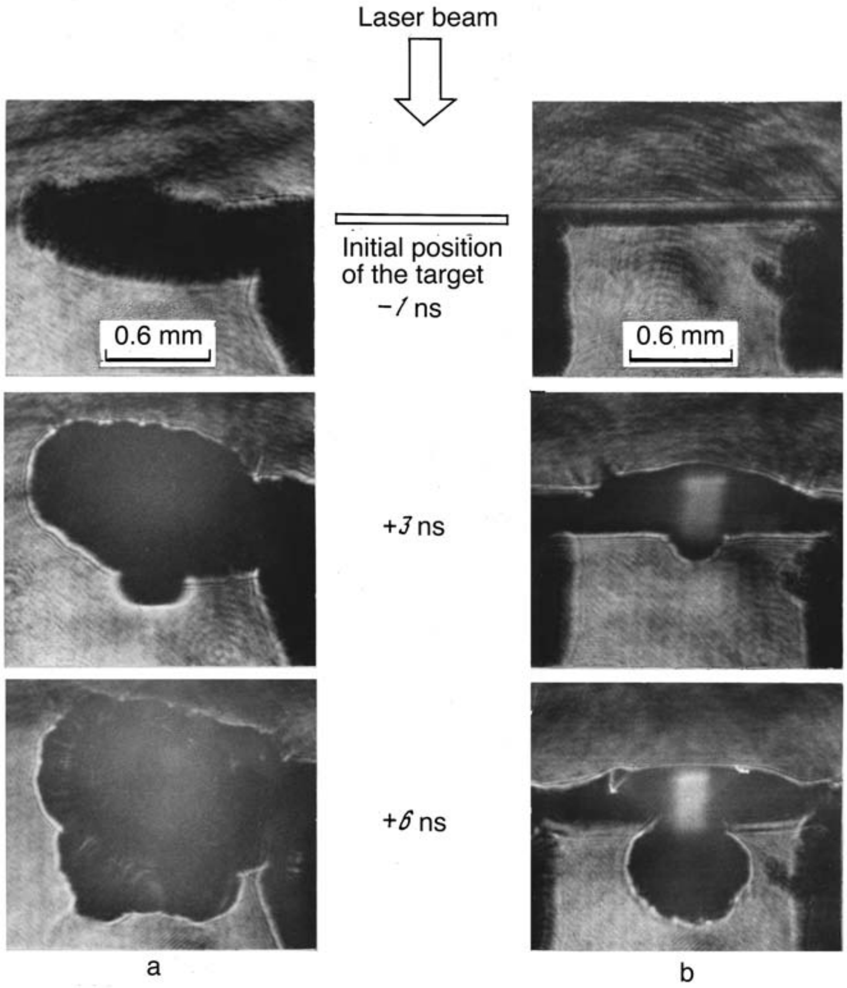


FIG. 5. Shadow photographs taken in experiments in which an agar target with density 1 mg/cm^3 and thickness 500 μm is irradiated (a) and a target made of Mylar with thickness 10 μm is irradiated (b).

density 1 mg/cm^3 (see Fig. 6a) amount to $\sim 2\times 10^7$ cm/s and $\sim$2.5 Mbar respectively. As the initial thickness of the porous material increases, the velocity of the aluminum foil decreases (see Fig. 6b). Increasing the initial density of the agar target leads to a similar effect. Aluminum foil with thickness 1.5 μm efficiently confines the expansion of the heated material of the porous target into vacuum, and motion is observed only within the zone of maximum pressure. The transverse size of this zone depends weakly on thickness and the density of the porous material, and equals $\sim$500 μm, i.e., it exceeds the size of the focal spot at the illuminated surface of the target by a factor of several.

The distinctive features of the interaction of high-power laser light with our porous targets arise from the lower density and larger extent of the plasma that forms. When the material from a target made of agar with density 4 mg/cm^3 is totally ionized, the electron density is close to the critical density for the fundamental harmonic of a neodymium laser, while the longitudinal size of the plasma that forms, as the data shown above illustrate, is several hundred times the wavelength of the laser radiation. It is known[7] that at sufficiently high optical flux densities ($\geqslant 10^{14}$ W/cm^2) these conditions are favorable for the development of various types of anomalous processes in the plasma (stimulated Brillouin scattering, stimulated Raman scattering, and parametric decay instabilities), which could lead to undesirable consequences when porous media are used as targets for laser thermonuclear fusion (decrease in the absorption coefficient of the laser light, generation of fast particles, decrease in the conversion coefficient of the laser light to x-rays, etc.). However, preliminary experiments in which we measured the energy and spectrum of the radiation scattered into the aperture of a focusing lens showed that for optical flux densities of $\sim 10^{14}$ W/cm^2 the fraction of scattered energy came to 5 to 6 percent of the pulse energy for illumination of either porous targets or targets made of Mylar with solid-state densities. It is interesting to note that radiation is observed in the scattering spectrum at the second harmonic frequency ($\lambda = 0.53$ μm) even when the illuminated agar target has an initial density of 1 mg/cm^3, although in this case the maximum electron density of the plasma for such a target at full ionization of the porous material ($n_e \approx 3\times 10^{20}$cm^{-3}) is several times lower than the critical value.

4. THEORETICAL MODEL FOR THE INTERACTION OF PULSED LASER LIGHT WITH A POROUS MATERIAL

4.1. Absorption of laser light in a porous material

Let us discuss the problem of absorption of a laser beam in a planar layer of porous material, taking into account the

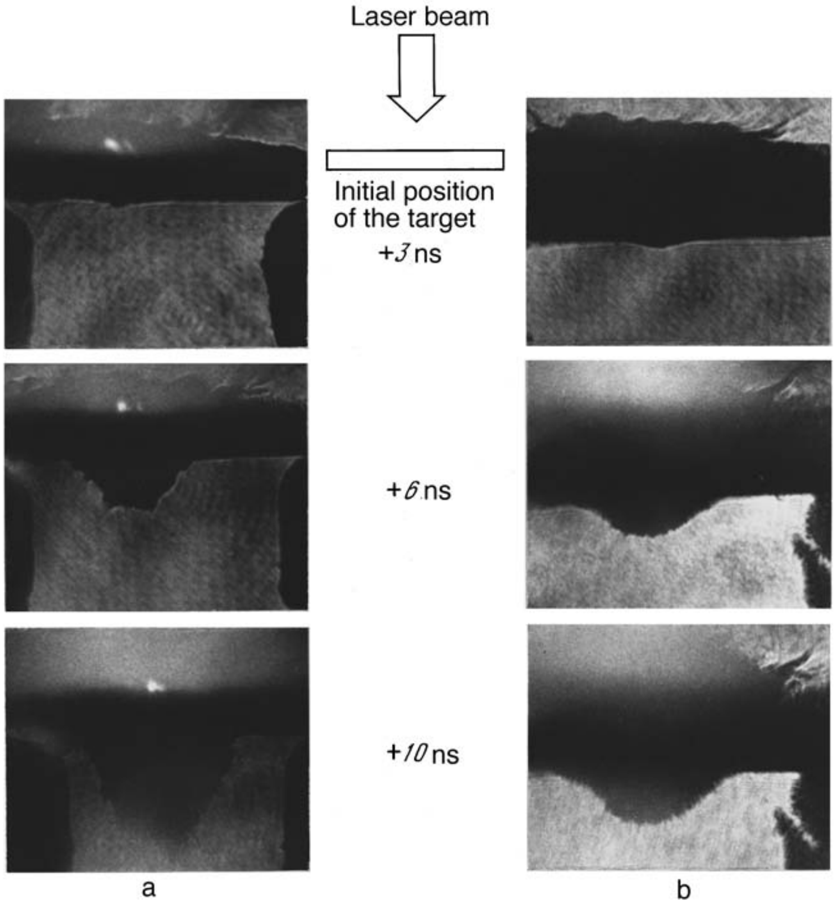


FIG. 6. Shadow photographs recorded in experiments on irradiated targets made of agar with density 1 mg/cm^3 and thicknesses 350 μm (a) and 1000 μm (b). An aluminum layer with thickness 1.5 μm was deposited on the back surface of the targets.

processes of internal vaporization of initially solid elements as they are heated by the laser light. The properties of a porous material are determined by its atomic composition, density, and the sizes of the solid particles of material, and also the dimensions of the pores. In the experiments we used targets made of agar, which consists of randomly located fibers of solid-state densities and lengths $l\sim 10$–50 μm and radii $b_0\sim 1$ μm ($l \gg b_0$). The average pore size r_p and the average number n of fibers per unit volume are expressed in terms of the dimensions and bulk density ρ_0 of the thread materials, and also the average bulk density of the porous material ρ_a, in the following way:

$$r_p \simeq \pi^{1/3} b_0^{2/3} l^{1/3} (\rho_0/\rho_a)^{1/3}, \quad n=\rho_a/(\rho_0 \pi b_0^2 l). \tag{1}$$

For porous materials consisting of randomly located fibers, the degree of geometric overlap of the laser beam per unit length along the direction of propagation of the beam is given by the expression $\sigma_* = 2/\pi b_0 l n$.

For a laser beam propagating into the porous material a distance x from the surface, the fractional overlap of the optical flux, when we neglect light scattering, is

$$\sigma = 1-\exp(-\sigma_* x) \simeq \sigma_* x. \tag{2}$$

Since the length of an agar fiber greatly exceeds its radius ($l \gg b_0$), the expansion of an individual fiber as it is heated by the laser light can be treated in the approximation of cylindrical dispersal. In this case, according to (2) the increase in the degree of opacity of the porous material in the process of internal vaporization is directly proportional to the radius of the expanding fiber. As we show below, under the conditions of our experiments the duration of the internal vaporization process (homogenization) of the porous material in the region where the laser beam acts is shorter than the rise time of the laser pulse. Doing the calculations for cylindrical isothermal expansion of a fiber of porous material for a linearly growing laser light power, we obtain the following expression for the opacity σ and the absorption length for laser radiation $L=\sigma_*^{-1}$:

$$\sigma = x/L, \quad L=L_0[1+(t/t_*)^2]^{-2}, \quad 0 \leqslant t \leqslant t_{\mathrm{cr},a}. \tag{3}$$

Here $L_0=\pi^2 \rho_0 b_0/2\rho_a$ is the absorption length corresponding to the initial state of the porous material; $t_* = [24\pi^2/(\gamma-1)(\rho_0 b_0^3/k_a I_m t_1)]^{1/2}$; k_a is the average absorption coefficient of laser light in an individual expanding thread; t_1 is the duration of the leading edge of the laser pulse; $I_m = I(t_1)$; γ is the adiabatic constant; $t_{\mathrm{cr},a} = [(\rho_0/\rho_{\mathrm{cr},a})^{1/4}-1]^{1/2} t_*$ are completion times for the processes that generate the region of absorption of the laser beam: t_{cr} is the internal vaporization time of a thread, i.e., the time it takes to reach the critical density (for a material with density $\rho_a < \rho_{\mathrm{cr}}$), while t_a is the total internal vaporization time, i.e., the time it takes to reach its average density ρ_a (for a material with density $\rho_a > \rho_{\mathrm{cr}}$).

Figure 7 shows the time dependence of the length of the absorption region for the laser beam in a porous material for

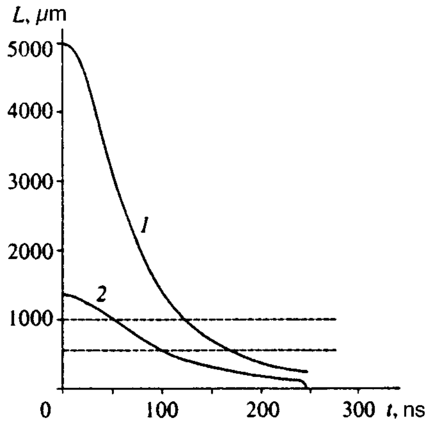


FIG. 7. Time dependence of the absorption length of a laser beam in a porous material with average density 1 mg/cm^3 (*1*) and 4 mg/cm^3 (*2*). The dashed lines show the characteristic thicknesses of the targets used.

two values of its average density (1 mg/cm^3 is a subcritical density and 4 mg/cm^3 is a supercritical density), calculating using Eq. (3). The absorption coefficient was taken to be 0.5, and the initial radius of the thread was 1 μm. The calculations show that the absorption region for laser light in the targets used in our experiments forms in a time shorter than the rise time of the laser pulse ($t_1 = 0.3$ ns). The absorption length for the laser beam in a porous material with a density close to the critical density decreases in the process of internal vaporization from its initial value L_0, which depends on the density of the material ρ_a and is a few mm, to a value $L \approx L_0(\rho_{\mathrm{cr}}/\rho_0)^{1/2}$ corresponding to the critical density of the expanding fibers, which equals ≈ 200 μm. According to (3), for a target of thickness Δ_a the transparency stage lasts a time

$$t_{\mathrm{tr}} = [(L_0/\Delta_a)^{1/2} - 1]^{1/2} t_*, \quad \Delta_a < L_0. \tag{4}$$

Then for a target with average density $\rho_a = 1$ mg/cm^3 the duration of the transparency stage is 150 ps for a target thickness of 500 μm, and 120 ps for a target thickness of 1000 μm. And, in fact, this time was less than 300 ps in experiments.

The spatial extent (along the laser beam) of the specific (per ion) energy absorption is determined from the expression

$$E(x,t) = \frac{1}{B\rho_a} \int_0^{t_f(x)} k_a I(t) \sigma(x,t)\, dt, \tag{5}$$

in which t_f is the time of arrival at the point with coordinate x of the front that marks the start of the region of cut-off absorption, and $B = [Z/(\gamma - 1)A] \times 10^{15}$ erg/g·keV is the specific heat of the plasma. For a porous material with subcritical average density the spatial extent of the specific energy in the absorption region with final length $L_{\mathrm{cr}} \approx \mathrm{L}_0(\rho_{\mathrm{cr}}/\rho_0)^{1/2}$ is quite uniform. The magnitude of the energy per unit volume absorbed in this region up to time T_{cr} according to (3) and (5) is

$$E_{\mathrm{cr}} = \frac{32}{3\pi^2(\gamma - 1)} \left(\frac{\rho_0}{\rho_{\mathrm{cr}}} \right)^{3/2} \left(\frac{\mathrm{k}_a \mathrm{I}_m b_0}{B^2 \rho_0 t} \right)^{1/2}. \tag{6}$$

In reality, formation of the region of absorption of laser light in the porous material ends with the total internal vaporization of the material. Using the expression for the characteristic length of a classical collisional absorption mechanism, and also Eq. (6), and setting $k_a \approx 0.5$, we can find the temperature and size of the absorption region at this total internal vaporization time. For a target made of agar with average density $\rho_a \approx 1$ mg/cm^3 it is easy to obtain the values $T \approx 400$ eV, $L_T \approx 300$ μm.

Thus, the effect of increasing opacity of the porous material as it internally vaporizes under the action of the laser light can explain the finite duration of the transparency stage observed in experiments on planar porous targets, and the formation within these targets of a region of bulk absorption of the laser light. For most of the duration of the laser pulse light is absorbed in just this region due to classical collision mechanisms, and the hydrothermal wave transports energy to the surrounding layers of porous materials.

4.2. Energy transfer in the porous material

The problem of propagation of a spherical hydrothermal wave is described by a system of equations which includes an equation for the velocity of the wave front:

$$\frac{dZ}{dt} = [(\gamma - 1)BT]^{1/2}, \quad t \gg t_{\mathrm{cr}}, \quad Z \gg L_t(t_{\mathrm{cr}}), \quad Z \gg r_f \tag{7}$$

(where r_f is the radius of the focal spot) and an equation for the energy, assuming that the temperature behind the wave front is uniform (the temperature behind the hydrothermal wave front is equalized by the ultrasonic electron thermal conduction wave):

$$2\pi Z^3 T B \rho_a = \begin{cases} \int_0^t E_L\, dt, & 0 \leq t \leq t_L \\ E_L, & t > t_L \end{cases}. \tag{8}$$

The solution of this system of equations for the case of constant laser light power (E_L = const) has the form

$$Z = \left[\frac{3}{2} \left(\frac{5}{3} \right)^2 \frac{(\gamma - 1)E_L}{\pi \rho_a} \right]^{1/5} t^{2/5} \varphi(t), \tag{9}$$

$$T = \left[\frac{3}{2} \left(\frac{5}{3} \right)^3 \frac{E_L}{(\gamma - 1)^{3/2} B^{5/3} \pi \rho_a} \right]^{2/5} t^{-6/5} \varphi^2(t). \tag{10}$$

Here

$$\varphi(t) = \begin{cases} (t/t_L)^{1/5}, & t_{\mathrm{cr}} \ll t \leq t_L \\ l, & t \geq t_L \end{cases}.$$

whereas the pressure behind the front of the hydrothermal wave is

$$P = \left[\left(\frac{3}{5} \right)^3 \frac{3}{2} \frac{(\gamma - 1)E_L \rho_a^{3/2}}{\pi} \right]^{2/5} t^{-6/5} \varphi^2(t). \tag{11}$$

Note that for $t > t_L$ the solution to the problem of propagation of a hydrothermal wave coincides up to constant coefficients with the solution to the problem of a strong explosion.[8] The difference in the coefficients is connected with the different velocities of the hydrothermal and shock waves. According

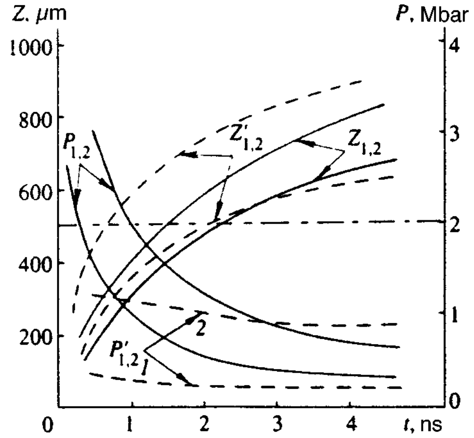


FIG. 8. Time dependence of the radius of the energy transport wave front (curves $Z_{1,2}$ and $Z'_{1,2}$) and the pressure behind the wave front (curves $P_{1,2}$ and $P'_{1,2}$). Label 1 is for an average density of 1 mg/cm^3, label 2 is for 4 mg/cm^3. Curves $Z_{1,2}$ and $P_{1,2}$ are plotted for transport of energy by a hydrothermal wave, curves $Z'_{1,2}$ and $P'_{1,2}$ for transport by an electron thermal-conduction wave.

to (9) when a laser beam strikes a planar target with thickness Δ_a the hydrothermal wave reaches the back surface of the target at the time

$$t_r=\begin{cases}\left[\left(\frac{3}{5}\right)^2\frac{2}{3}\,\frac{\pi\rho_a t_L\Delta_a^5}{(\gamma-1)E_L}\right]^{1/3}, & t_r\leqslant t_L\\ \left[\left(\frac{3}{5}\right)^2\frac{2}{3}\,\frac{\pi\rho_a\Delta_a^5}{(\gamma-1)E_L}\right]^{1/2}, & t_r\geqslant t_L\end{cases}. \tag{12}$$

At this time, according to (9)–(12) the temperature and pressure of the plasma behind the wave front equal

$$T=\begin{cases}\left[\frac{3}{5}\left(\frac{3}{2}\right)\frac{E_L}{\pi(\gamma-1)^{1/2}B^{3/2}t_L\rho_a\Delta_a^2}\right]^{2/3}, & t_r\leqslant t_L\\ \frac{3}{2}\,\frac{E_L}{\pi\rho_a B\Delta_a^3}, & t_r\geqslant t_L\end{cases}, \tag{13}$$

$$P=\begin{cases}\left[\left(\frac{3}{5}\right)\frac{(\gamma-1)E_L\rho_a^{1/2}}{\pi t_L\Delta_a^2}\right]^{2/3}, & t_r\leqslant t_L\\ \frac{3}{2}\,\frac{(\gamma-1)E_L}{\pi\Delta_a^3}, & t_r>t_L\end{cases}. \tag{14}$$

In analyzing energy transport by a hydrothermal wave and comparing it with experimental results, it is also useful to include a calculation of the propagation velocity of a spherical electron thermal conduction wave in a uniform medium of density ρ_a under the action of a constant-power laser pulse:[1]

$$Z\approx\left[\frac{19}{7}\,\frac{f\psi}{B^{7/2}}\left(\frac{3}{2\pi}\,\frac{E_L}{t_L}\right)^{5/2}\right]^{2/19}\left(\frac{t}{\rho_a}\right)^{7/19}. \tag{15}$$

Here $\psi\approx 10^{19}/Z$ erg/(keV$^{7/2}$ cm×s) is the coefficient in the expression for the thermal conductivity, and f is the factor that limits the thermal flux. Figure 8 shows the time dependence of the radii of the hydrothermal wave front and of the electron thermal conduction wave front, and also the pressure behind the front for these waves calculated using expressions (9), (11), (15). Calculations for the hydrothermal wave give results close to experiment for the arrival time of the energy transport wave at the back surface of a planar target and for the pressure behind the wave front at this time. For a target with thickness $\Delta_a=0.5$ mm these quantities come to 1.3 ns and 0.6 Mbar, respectively for a material whose mean density $\rho_a=1$ mg/cm^3, and 2.2 ns and 1.2 Mbar for a material whose density $\rho_a=4$ mg/cm^3. The propagation velocity of the hydrothermal wave equals 4×10^7 cm/s for a material of density is $\rho_a=1$ mg/cm^3 and 2.7×10^7 cm/s for a material whose density $\rho_a=4$ mg/cm^3. The classical electron thermal conduction wave (for a thermal flux bounding factor $f=1$) propagates with a higher velocity: 6×10^7 cm/s for $\rho_a=1$ mg/cm^3 and 3.8×10^7 cm/s for $\rho_a=4$ mg/cm^3, which implies an arrival time of the wave at the back surface of the target that is considerably smaller than the experimental value (especially for the target with $\rho_a=1$ mg/cm^3): $t_r=0.6$ ns for $\rho_a=1$ mg/cm^3 and $t_r=1.8$ ns for $\rho_a=4$ mg/cm^3. The electron thermal conduction wave gives results close to experiment only when we introduce significant limiting factors for the thermal flux ($f\approx0.01$ for $\rho_a=1$ mg/cm^3 and $f\approx0.1$ for $\rho_a=4$ mg/cm^3).

Thus, our model of energy transfer via hydrothermal waves when laser light is absorbed in a porous medium is adequate to describe the experimental results. We note that the energy transport wave passes through a target of thickness $\Delta_a=0.5$ mm in a time shorter than the duration of the laser pulse $t_r<t_L$. Estimates using Eq. (13) give values of the temperature $T\approx800$ eV for a target of this thickness and a density $\rho_a=1$ mg/cm^3 at the instant the hydrothermal wave reaches the back side of the surface. In contrast, the backward (classical) bremsstrahlung absorption length in a plasma with temperature 800 eV and an electron density corresponding to a target with density $\rho_a=1$ mg/cm^3 is 720 μm, i.e., larger than the target thickness. Therefore, for porous targets with these parameters we can reach the stage of secondary transparency to the laser light, connected with the formation of the high-temperature plasma over the entire thickness of the target.

In concluding this section, let us consider the question of acceleration of the planar layer of solid material at the back of the target surface. Acceleration of a planar layer of solid material after the arrival of the hydrothermal wave at the back surface of the porous material takes place, while the hydrothermal wave continues to propagate through the porous material; however, it is now in the form of a cylindrical wave. Solution to the equations of motion for a thin solid layer under these conditions gives the following estimate for the asymptotic velocity of the layer:

$$v_s\approx\left(\frac{P(t_r)}{\rho_a}\right)^{1/2}\left(\frac{\rho_a\Delta_a}{\rho_s\Delta_s}\right), \tag{16}$$

where ρ_s and Δ_s are the density and thickness of the solid material layer respectively. For a two-layer target consisting of a layer of agar with $\rho_a=1$ mg/cm^3, $\Delta_a=0.5$ mm and a layer of aluminum with $\rho_s=2.7$ g/cm^3, $\Delta_s=1.5$ μm, Eq. (16) gives $v_s\approx5\times10^6$ cm/s for the magnitude of the velocity in the solid layer, which is in good agreement with experiment.

5. CONCLUSION

We have experimentally investigated the interaction of high-power laser pulses ($\lambda = 1.054$ μm, $\tau = 2.5$ ns, $I = 10^{14}$ W/cm^2) with planar low-density porous targets (1–4 mg/cm^3).

We find that targets with thicknesses of several hundred microns are partially transparent to the laser radiation for times less than 300 ps from the start of illumination.

We have shown that even at the initial stages of illumination ($t = 0.5$ ns) a region of dense high-temperature plasma forms within the porous target at a depth 100–200 μm, in which the laser radiation is effectively absorbed.

The velocity of energy transport from the region of absorption of the laser light to the surrounding layers of porous material reaches 2×10^7 cm/s, while the plasma temperature rises to 0.8 keV within the internal regions of the target over the time of the laser pulse.

We have developed a theoretical model that takes into account the properties of the absorption of laser radiation in a porous medium with low average density and a hydrothermal wave based on the representation of energy transport in the porous medium.

The results of these experiments are found to be in reasonable agreement with the theoretical predictions. Comparison of the experimental data with the results of our theoretical analysis allows us, in particular, to estimate the pressure of the plasma at the back surface of the porous target at the time the hydrothermal wave arrives. At the back surface of a target with density 1 mg/cm^3 and thickness 500 μm the pressure of the plasma exceeds 1 Mbar within an area with dimensions that are several times larger than the diameter of the focal spot.

The authors are grateful to V. F. Vasil'ev, V. V. Kryzhko, B. N. Mironov, and V. G. Nikolaevskiĭ for help in carrying out the experiments, and also to O. L. Dedova and M. Yu. Sadkov for participating in the making of our illustrations.

This work was carried out with the financial support of the Russian Fund for Fundamental Research (project No. 96-02-19157) and the International Science and Technology Center (Project No. 029-94).

[1] S. Yu. Gus'kov, N. V. Emitrenko, and V. B. Rozanov, Zh. Éksp. Teor. Fiz. **108**, 548 (1995) [JETP **81**, 296 (1995)].

[2] J. Lindl, Phys. Plasmas **2**, 3933 (1995).

[3] V. A. Bolotin, I. N. Burdonsky, V. V. Gavrilov *et al.*, Rev. Sci. Instrum. **61**, 3259 (1991).

[4] A. V. Batjunin, A. N. Bulatov, A. V. Branitskij *et al.*, ''Beams-90,'', *8th Int. Conf. on High-Power Particle Beams*, Novosibirsk, USSR, 1990. paper V05, p. 109.

[5] V. V. Aleksandrov, M. B. Brenner, S. V. Loburev *et al.*, Kvant. Elektron. **13**, 677 (1986) [Sov. J. Quantum Electron. **16**, 443 (1986)].

[6] L. A. Bol'shov, I. N. Burdonskiĭ, A. L. Velikovich *et al.*, Zh. Éksp. Teor. Fiz. **92**, 2060 (1987) [Sov. Phys. JETP **65**, 1160 (1987)].

[7] G. Duderstadt and G. Moses, *Inertial-Confinement Fusion,* Wiley, New York (1982).

[8] L. I. Sedov, Dok. Akad. Nauk SSSR **42**, 112 (1946).

Translated by Frank J. Crowne

Nonlinear repolarization processes in ferroelectric liquid-crystal thin films

S. A. Pikin

Institute of Crystallography, Russian Academy of Sciences, 117333 Moscow, Russia

E. I. Demikhov

Institute of Solid State Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Province, Russia

E. S. Pikina

Institute for Problems of Petroleum and Gas, Russian Academy of Sciences, 117917 Moscow, Russia

M. V. Gorkunov

Moscow State Engineering–Physical Institute, 115409 Moscow, Russia

(Submitted 12 July 1996)
Zh. Éksp. Teor. Fiz. **111**, 919–937 (March 1997)

The dynamics of spontaneous polarization switching of the ferroelectric smectic C^* in a variable electric field are examined theoretically and experimentally with the help of polarized light scattering. The observed effect of quasiresonant scattering both in freely suspended smectic films and in ordinary electro-optical cells is interpreted within the framework of the nonlinear model of isolated movable kinks in the director orientation distribution. It is shown that the maximum of the scattering intensity at the characteristic frequency of the applied electric field disappears at low temperatures and for small thicknesses of the smectic film. The dependence of the ''resonant'' frequency on the electric field amplitude, the proximity to the phase transition temperature, the film thickness and thickness of the ferroelectric domains, and also various material parameters is found. Estimates are made of such important characteristics as the dielectric anisotropy, viscosity, and elasticity of the smectic films. The effect of film thickness on the density distribution of the polar anisotropy energy in the film and on the corresponding shape of the moving orientation front within the film are discussed.
[S1063-7761(97)01303-6]

1. INTRODUCTION

The dynamics of ferroelectric liquid crystals (LC) in electric fields is usually associated with the existence of collective modes,[1–12] which are damped. Specifically, this means they are associated with the Goldstone mode, which describes the relaxation of perturbations of the phase of the order parameter (spontaneous polarization **P**), and a soft mode, which describes the relaxation of perturbations of the amplitude of the order parameter.

These modes are readily observed by means of dielectric spectroscopy as a response of ferroelectric liquid crystals to the application of a relatively weak electric field E, which, as a rule, is smaller than some critical value E_c called the unwinding field of the polarization helix[5] (such a helix exists in practically all ferroelectric liquid crystals in thermodynamic equilibrium). Here the phase φ, which is the azimuthal angle of the orientation vector **P** lying in the plane of the smectic layer of the ferroelectric liquid crystal, and the amplitude $P=\mu\theta$, where μ is the piezoelectric modulus (a material parameter of the liquid crystal) and θ is the polar angle of the collective tilt of the molecules in the smectic layer, vary weakly and almost independently of each other.[5]

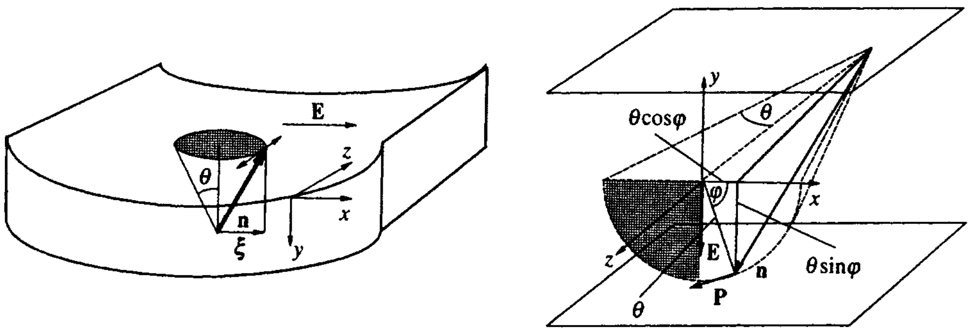
In strong fields exceeding the unwinding field, perturbations of the angles φ and θ become large and are in general interdependent. Thus, ferroelectric liquid crystals repolarize in strong fields. For the appropriate boundary conditions on the surface of a liquid-crystal film, sufficiently weak anisotropy of the liquid crystal in the plane of the smectic layer, and relative rigidity of the orientational structure of the liquid crystal along the crystal axis perpendicular to the smectic layer, variations of the tilt angle θ can be weak, but variations of the azimuthal angle φ become large.

A large number of varied applications of electro-optical cells based on ferroelectric liquid crystals possessing such properties come at once to mind. Thus, the corresponding characteristics of azimuthal switching of the spontaneous polarization are of significant interest; however, these essentially nonlinear repolarization processes have been little studied.

In the present paper we present a fundamentally nonlinear model of azimuthal repolarization of ferroelectric liquid crystals in a variable electric field together with experimental data on light scattering in repolarizing liquid-crystal films. These studies allow us to draw definite conclusions about the suitability of the proposed model. We also advance the hypothesis that the described nonlinear polarization dynamics is also applicable in the case of solid solutions of ferroelectrics with the corresponding anistropy properties.

The main idea is that repolarization, i.e., rotation of the polarization vector **P** under the action of the electric field **E** through the angle π does not happen all at once throughout the volume of the liquid crystal, but begins in the ''weak-

1063-7761/97/030506-10$10.00



a b

FIG. 1. Geometry of ferroelectric liquid-crystal films: a) geometry #1—freely suspended film, b) geometry #2—''bookshelf'' (ordinary cell).

est'' points of the film, e.g., on its surface, assorted defects, etc., via local generation of a so-called orientational kink, which is characterized by an abrupt change in the dependence $\varphi(\mathbf{r})$ over a relatively small scale $\Delta\mathbf{r}$. A kink $\Delta\varphi \sim \varphi(\mathbf{r}+\Delta\mathbf{r})-\varphi(\mathbf{r})\sim\pi$ then propagates into the interior of the film with some velocity v which depends on the applied field to distances d which are much greater than the width $|\Delta\mathbf{r}|$ of the kink, where the distances d can characterize the thickness of the film, the width of the ferroelectric domain, the distance between defects, etc. (Such a moving kink is sometimes called a soliton.) In a variable field this moving kink takes on the nature of an oscillator.

This repolarization mechanism, i.e., the dynamics of motion of the orientational kink (solitons), leads to an entirely different form of the perturbations of the dielectric tensor (permittivity tensor) $\epsilon(\mathbf{r},t)$ in comparison with the Goldstone contribution to these perturbations. In particular, light scattering off a Goldstone mode cannot explain the curves of the observed frequency dependence of the integrated scattering intensity $I(\omega)$, which have a quasiresonant character, where ω is the frequency of the applied field $E=\tilde{E}\cos(\omega t)$. We show below that the soliton mechanism of perturbations in the permittivity provides a completely adequate explanation of the experimental results.

2. THE EQUATION OF AZIMUTHAL MOTION AND ISOLATED ORIENTATIONAL KINKS

We consider two situations in which the applied electric field $\mathbf{E}$ is parallel to the smectic layers and induces azimuthal rotations of the polarization vector $\mathbf{P}$ by the angle $\varphi(\mathbf{r},t)$ in the plane of the smectic layer, that is to say, the case of a freely suspended film, in which the smectic layers are parallel to the surface of the film (Fig. 1a), the case of ''bookshelf'' geometry, i.e., the case of a surface-stabilized ferroelectric liquid crystal in which the smectic layers are perpendicular to the film surface (Fig. 1b). In these cases the equations of motion for the director and polarization vector have the form[13–17]

$$PE\sin\varphi - U\theta^2\sin\varphi\cos\varphi + K\theta^2\frac{\partial^2\varphi}{\partial y^2}=\gamma\theta^2\frac{\partial\varphi}{\partial t}, \quad (1)$$

where K is the elasticity constant, γ is the viscosity coefficient, and the quantity

$$U\theta^2=(U_0+\epsilon_a E^2)\theta^2 \quad (2)$$

is the anisotropy energy and ϵ_a is the dielectric anisotropy. We assume that the quantity $U=U_0+\epsilon_a E^2$ is positive, i.e., the easy (most favored) direction of the polarization vector $\mathbf{P}$ corresponds to the angles $\varphi=0$ and $\varphi=\pi$. The angle φ in these cases varies along the y axis which lies in the plane of the smectic layer.

The right-hand side of Eq. (1) corresponds to the existence of the invariant expression

$$\gamma\left(\xi_1\frac{\partial\xi_2}{\partial t}-\xi_2\frac{\partial\xi_1}{\partial t}\right)=\gamma(\boldsymbol{\xi}\times\dot{\boldsymbol{\xi}})_y \quad (3)$$

for the viscous torque, which retards the rotation of the two-dimensional vector $\boldsymbol{\xi}$ with components

$$\xi_1=\theta\cos\varphi, \quad \xi_2=\theta\sin\varphi, \quad (4)$$

describing the orientational order parameter of the ferroelectric liquid crystal.[5]

The anisotropy parameter $U\theta^2$ characterizes the energy of the anisotropic dipole–dipole interactions which give rise to the polarization fluctuations inducing a Coulomb interaction between the polarization charges, which in turn checks the development of such fluctuations. The preferred orientation of the spontaneous polarization vector in the liquid-crystal film may be conditioned by the boundary conditions and external factors and may be fixed in the film, thanks to the presence of free current carriers. Negative values of the parameter U correspond to the preferred orientation of the polarization vector $\varphi=\pi/2$, i.e., perpendicular to the applied electric field, while in the case of the bookshelf geometry (see Fig. 1b) the preferred direction is parallel to the film surface. In the case of a freely suspended thin film (Fig. 1a) the smectic layer can be homogeneous along the z axis and inhomogeneous in x and y due to the influence of the film surfaces and the electrodes. In particular, very thin films can have an easy direction, the z axis, with $U<0$. The temperature dependence of the parameter U_0 is completely determined by the fluctuational nature of the considered anisotropy: in the phenomenological description it satisfies $U_0 \propto -\theta^2 \propto T_{AC*}-T$, i.e., the corresponding contribution to the free energy is

$$U\theta^2=-\mathrm{const}\cdot\theta^4+\epsilon_a\theta^2E^2. \quad (5)$$

Here T_{AC*} is the temperature of the smectic A–smectic C^* phase transition.

Thus, it may be supposed that in very thin films the parameter U can take negative values, at least at sufficiently low temperatures (at sufficiently large values of the tilt angle θ). In relatively thick films in the absence of a field the parameter U is evidently small since in the limiting case of an unbounded ferroelectric liquid crystal the smectic layers should be practically isotropic. When a field $\mathbf{E}$ is applied perpendicular to the surface of a thick film the parameter U can take positive values ($\epsilon_a > 0$) since usually the electrode surfaces define the orientation of the polarization vector $\mathbf{P}$ to be normal to the film surface. These considerations are purely qualitative and motivational. Below we show that they can be backed up by light-scattering experiments in various liquid-crystal films where the parameters U_0 and ϵ_a can be estimated from the experimental data.

Equation (1) can be used to describe rotations of the director and of the polarization under the action of an applied field $E=\widetilde{E}\cos(\omega t)$. In this case we will make use of an approximate expression for the anisotropy energy which corresponds to the time-averaged value of the expression $U\theta^2$, i.e., without allowance for the second harmonic:

$$\langle U\theta^2\rangle \approx -\mathrm{const}\cdot\theta^4 + \frac{1}{2}\epsilon_a\theta^2\widetilde{E}^2, \tag{6}$$

where by θ we understand some effective value of the tilt angle of the director depending on the field amplitude $\widetilde{E}$. This approximation, as a consideration of problems of this kind shows, gives an adequate qualitative description. Below we will drop the time-averaging notation $\langle\dots\rangle$.

It is convenient to introduce the dimensionless variable $s=y/\eta$, the dimensionless parameter $b=d/\eta$, and parameters having dimensions of frequency and length, respectively

$$a=\frac{\mu\widetilde{E}}{\gamma\theta}, \quad \eta=\sqrt{\frac{K}{U}}, \tag{7}$$

where d is the characteristic length in the film. Given the assumptions we have made, Eq. (1), as shown in the Appendix, has a solution of the form

$$\varphi(s,t)=\arctan\frac{1}{\sinh[s-s_0-(a/\omega)\sin(\omega t)]}, \tag{8}$$

where s_0 is some constant, and for constant and homogeneous U, in the notation of Eq. (A1) we have

$$\widetilde{\Xi}=\frac{\mu\widetilde{E}}{U\theta}, \quad \vartheta=\frac{\omega\gamma}{U}, \quad \tau=\frac{tU}{\gamma}, \quad u(s)=s=\frac{y}{\eta}. \tag{9}$$

Solution (8) describes the motion of an isolated kink along the y axis with velocity

$$v=\eta a\cos(\omega t). \tag{10}$$

In a constant field, i.e., in the limit $\omega\to 0$, Eqs. (8) and (10) describe the motion of the kink[13,14] with constant velocity $v=\eta a$:

$$\varphi(s,t)=\arctan\frac{1}{\sinh(s-s_0-at)}. \tag{11}$$

It should be noted that we are considering large reorientation angles $\varphi\sim\pi$, i.e., the given case differs fundamentally from the usual dielectric measurements in weak electric fields. Moreover, in this situation repolarization takes place between two homogeneous orientational states ($\varphi=0$ and $\varphi=\pi$), i.e., in the absence of an orientational helix. As is well known, such a helix can be unwound between the walls of the electro-optical cell if the smectic film is thin enough ($d<0.1$–$1\ \mu$m), and also if the electric field E exceeds the unwinding field E_c. The value of E_c is a function of the temperature and vanishes at the second-order phase transition temperature T_c (Ref. 5). Therefore even in relatively thick films at temperatures close to T_c we are in the right to consider homogeneous orientational states in moderate fields.

3. DYNAMICAL INHOMOGENEITY OF THE PERMITTIVITY AND SCATTERING OF POLARIZED LIGHT

Experimentally, such repolarization processes can be studied by scattering polarized light off the dynamic inhomogeneities of the permittivity. The total integrated scattered light intensity is proportional to the integral of the square of the scattering amplitude $\alpha(y,t)$ over the variables y and t.

The function $\alpha(y,t)$ is proportional to the quantity $\mathbf{i}\cdot\Delta\hat{\varepsilon}(y,t)\cdot\mathbf{f}$, where $\mathbf{i}$ and $\mathbf{f}$ are the initial and final polarizations of the light, and $\Delta\hat{\varepsilon}$ is the permutation of the dielectric tensor (permittivity tensor) for light with the given wavelength. In the case under consideration the quantity $\mathbf{i}\cdot\Delta\hat{\epsilon}\cdot\mathbf{f}$ is proportional to the difference of the products of the components of the order parameter

$$\xi_i(y,t)\xi_f(y,t)-\xi_i(y,0)\xi_f(y,0), \tag{12}$$

and therefore the total scattered light intensity is proportional to

$$I_{if}=\int_0^{2\pi/\omega}dt\int_0^d dy[\xi_i(y,t)\xi_f(y,t)-\xi_i(y,0)\xi_f(y,0)]^2. \tag{13}$$

Two experimental geometries are of interest:

1) a freely suspended film (Fig. 1a) with the y axis aligned with the normal to the plane of the film, in an external field E directed along the x axis; light with polarization i_x or i_z is normally incident upon the film.

2) An electro-optical cell (Fig. 1b) with the y axis perpendicular to the plane of the electrodes, in which the external field E lies; the incident light has polarization i_x or i_z.

If the incident and scattered light are polarized identically (either along the x axis or along the z axis) in both cases, then the quantity I characterizing the total scattered light intensity is, according to Eqs. (4), (12), and (13), equal in both cases to

$$I_{xx}=I_{zz}=\theta^4\int_0^{2\pi/\omega}dt\int_0^d dy[\cos^2\varphi(y,t) - \cos^2\varphi(y,0)]^2. \tag{14}$$

In the case of crossed polarizers (i_x and f_z) for geometries #1 and #2 we obtain respectively the equations

$$I'_{xz}=\theta^4\int_0^{2\pi/\omega}dt\int_0^d dy[\sin\,\varphi(y,t)\cos\,\varphi(y,t)$$

$$-\sin\,\varphi(y,0)\cos\,\varphi(y,0)]^2, \tag{15}$$

$$I''_{xz}=\theta^2\int_0^{2\pi/\omega}dt\int_0^d dy[\cos\,\varphi(y,t)-\cos\,\varphi(y,0)]^2. \tag{16}$$

For geometry #2 and the light incident in the z direction in crossed polarizers (i_x and f_y) the quantity I_{xy} coincides with the value of I_{xz} for geometry #1. It is useful to note in the calculations of the intensity I that in geometry #2 (see Fig. 1b) the projection of the director $\mathbf{n}$ on the z axis is approximately equal to $\sqrt{1-\theta^2\cos^2\varphi}$.

Using the relations

$$\sin\,\varphi(y,t)=\frac{1}{\cosh(s-A)},\quad \cos\,\varphi(y,t)=\tanh(s-A),$$

$$\sin\,\varphi(y,0)=\frac{1}{\cosh\,s},\quad \cos\,\varphi(y,0)=\tanh\,s, \tag{17}$$

$$A\equiv A(t)=\frac{a}{\omega}\sin(\omega t),\quad b=\frac{d}{\eta},$$

in which we have left out the unimportant constant s_0, we can calculate the integrals over the coordinate y in expressions (14)–(16). We obtain

$$I_{zz}(\omega)=\theta^4\eta\int_0^{2\pi/\omega}dt\Bigg\{\tanh\,b-\frac{1}{3}\tanh^3\,b+\tanh\,A$$

$$-\frac{1}{3}\tanh^3\,A+\tanh(b-A)-\frac{1}{3}\tanh^3(b-A)$$

$$+\frac{2}{\sinh^2\,A}\Bigg[\tanh\,b+\frac{(1-\tanh^2\,A)\tanh\,b}{1-\tanh\,b\,\tanh\,A}$$

$$+\frac{2}{\tanh\,A}\ln(1-\tanh\,b\,\tanh\,A)\Bigg]\Bigg\}, \tag{18}$$

$$I'_{xz}(\omega)=\theta^4\eta\int_0^{2\pi/\omega}dt\Bigg\{\frac{1}{3}\tanh^3\,b+\frac{1}{3}\tanh^3(b-A)$$

$$+\frac{1}{3}\tanh^3\,A+\frac{2}{\sinh\,A}\ln(1-\tanh\,b\,\tanh\,A)$$

$$-\frac{2\cosh\,A}{\sinh^2\,A}\Bigg[\tanh\,b+\frac{(1-\tanh^2\,A)\tanh\,b}{1-\tanh\,b\,\tanh\,A}$$

$$+\frac{2}{\tanh\,A}\ln(1-\tanh\,b\,\tanh\,A)\Bigg]\Bigg\}, \tag{19}$$

$$I''_{xz}(\omega)=-\theta^4\eta\int_0^{2\pi/\omega}dt\Bigg\{\tanh\,b+\tanh(b-A)+\tanh\,A$$

$$+\frac{2}{\tanh\,A}\ln(1-\tanh\,b\,\tanh\,A)\Bigg\}. \tag{20}$$

Integrals (18)–(20) cannot be expressed in terms of elementary functions and, what is more, at very low frequencies ω of the external variable field the limits of integration in time must be changed for the following reasons. It is clear that if the period of the oscillations of the electric field is much greater than the time it takes a kink to move across the film (perpendicular or parallel to its surface), i.e., for $2\pi/\omega\gg d/v$, perturbations of the permittivity will exist only for times of order d/v. Therefore, in the low-frequency limit, when $A(t)\approx at$ holds, the integrals in Eqs. (18)–(20) should be calculated over the interval $0\leqslant t\leqslant b/a$, i.e., the quantities $I(\omega\to 0)$ tend toward some finite limit. At very high frequencies, when the spatial amplitudes of the oscillations of the kinks are of the order of $\eta a/\omega\ll d$, the intensities I should tend to zero. The asymptotic limit $I(\omega\to\infty)$ can be calculated with the help of expansions in the small quantity A. In particular, expressions (18)–(20) have the asymptotic limits

$$I_{zz}(\omega\to\infty)\approx\theta^4\eta\left(\frac{4}{3}\tanh^3\,b-\frac{4}{5}\tanh^5\,b\right)\frac{\pi a^2}{\omega^3}, \tag{21}$$

$$I'_{xz}(\omega\to\infty)\approx\theta^4\eta\Bigg(6\tanh\,b-7\tanh^3\,b$$

$$+\frac{16}{5}\tanh^5\,b\Bigg)\frac{\pi a^2}{\omega^3}, \tag{22}$$

$$I''_{xz}(\omega\to\infty)\approx\theta^2\eta\left(\tanh\,b-\frac{1}{3}\tanh^3\,b\right)\frac{\pi a^2}{\omega^3}. \tag{23}$$

The limits $I(\omega\to 0)$ can also be calculated for small values of the parameter b if we make the substitution

$$\int_0^{2\pi/\omega}A^2(t)dt=\frac{\pi a^2}{\omega^3}\to a^2\int_0^{b/a}t^2dt=\frac{b^3}{3a}, \tag{24}$$

which gives

$$I_{zz}(\omega\to 0,\ b\to 0)\sim\theta^4\eta b^6/a, \tag{25}$$

$$I'_{xz}(\omega\to 0,\ b\to 0)\sim\theta^4\eta b^4/a, \tag{26}$$

$$I''_{xz}(\omega\to 0,\ b\to 0)\sim\theta^2\eta b^4/a. \tag{27}$$

Expressions (25)–(27) can be used to estimate the effect of scattering when the kinks in fact disappear for $b\ll 1$ and the dependence $I(\omega)$ has a quasiresonant character. At intermediate values of ω the function $I(\omega)$ should have a maximum. This has the physical meaning that the permittivity experiences its largest perturbation when the kink runs across a distance d during a time of the order of the half-period π/ω.

$$d\sim\frac{v}{\omega_{\text{ext}}}\sim\frac{\eta a}{\omega_{\text{ext}}},\quad\text{i.e. }\ \omega_{\text{ext}}\sim\frac{a}{b}, \tag{28}$$

where the quantity ω_{ext} corresponds to the maximum intensity $I_{\max}=I(\omega_{\text{ext}})\sim\theta^4\eta b/a$. Consequently, we have

$$\omega_{\text{ext}}\sim\frac{a}{b}\sim\frac{\mu\eta\tilde{E}}{\gamma\theta d}\sim\frac{\mu\tilde{E}}{\gamma\theta d}\sqrt{\frac{K}{U}},$$

$$I_{\max}(\omega\sim\omega_{\text{ext}})\sim\frac{\theta^4 d}{a}\sim\frac{\gamma\theta^5 d}{\mu\tilde{E}}. \tag{29}$$

Numerical calculations confirm these qualitative estimates. Figures 2a–b reveal the role of the upper integration

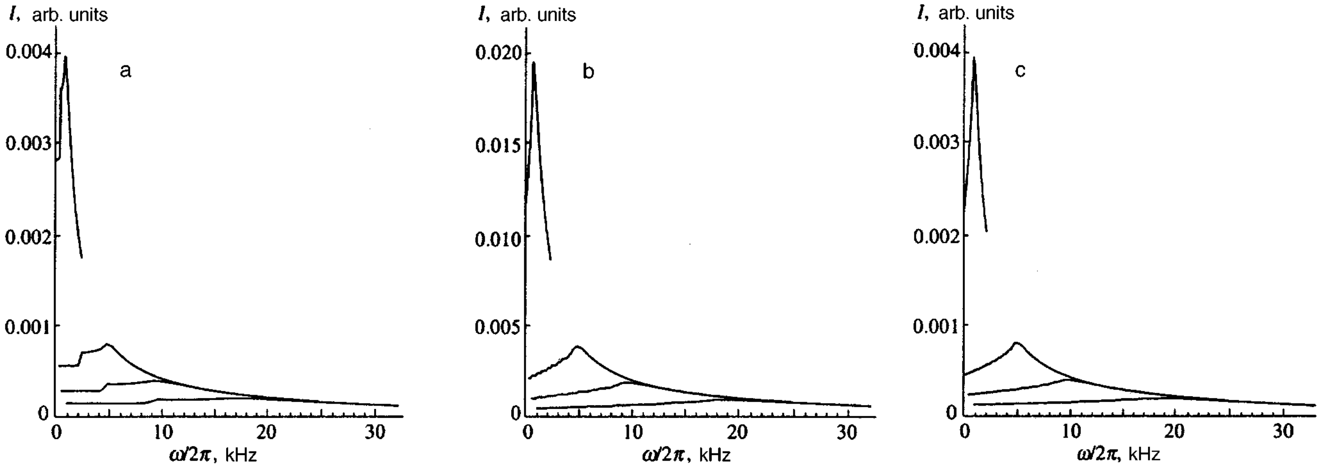


FIG. 2. Numerical calculations of the frequency dependence of the integrated light scattering intensity (crossed polarizers). The effect of the integration limit in time is shown: a) $2\pi/\omega$ for $\omega > a/b$ and $2\pi a/b$ for $\omega < a/b$; b) $10\pi/\omega$ for $\omega > a/b$ and $10\pi a/b$ for $\omega < a/b$; c) $20\pi/\omega$ for $\omega > a/b$ and $20\pi a/b$ for $\omega < a/b$. The parameters a and b in the calculations of the four curves from top to bottom took the respective values $a=10^4$, $b=10$; $a=10^5$, $b=20$; $a=10^5$, $b=10$; $a=10^5$, $b=5$.

limit in expressions (18)–(20), which was chosen to be equal to $2\pi/\omega$, $10\pi/\omega$, and $20\pi/\omega$, respectively, for $\omega > a/b$ and $2\pi a/b$, $10\pi a/b$, and $20\pi a/b$, respectively, for $\omega \leqslant a/b$. It is clear that the figures are qualitatively similar, but at low frequencies the behavior of the functions $I(\omega)$ depends on the upper limit t_0, which is determined in the experiment by the instrument parameters.

Relations (29) show that at a fixed temperature, i.e., for $\theta=$ const and $U \sim \tilde{E}^2$ the quantities $\omega_{\rm ext}$ and $I_{\max}$ as functions of the field amplitude $\tilde{E}$ behave as follows:

$$\omega_{\rm ext} = \text{const}, \quad I_{zz_{\max}} \sim I'_{xz_{\max}} \propto 1/\tilde{E}. \tag{30}$$

It would be well to underscore the main features of the behavior of the intensity maximum $I''_{xz_{\max}}$ as a function of $\tilde{E}$ for constant θ: $I''_{xz_{\max}}=$const, which follows from relations (16) and (20). Thus, the field dependence of the quasi-resonant behavior of the intensity in geometries #1 and #2 differs in the case of crossed polarizers.

The temperature dependence of $\omega_{\rm ext}$ and $I_{\max}$ is determined by the temperature dependence of the tilt angle θ. Thus, as the temperature is lowered the tilt angle increases and correspondingly the frequency $\omega_{\rm ext}$ decreases and the intensity $I_{\max}$ increases. If, in accordance with relation (5), as the temperature is lowered in a sufficiently thin film the quantity U vanishes at some value of the temperature T, then at this point the width of the kink η diverges (the parameter b vanishes) and, according to relations (30), the frequency $\omega_{\rm ext}$ grows without bound and the intensity $I_{\max}$ vanishes at this temperature.

4. DISCUSSION OF EXPERIMENTAL RESULTS

The observed frequency dependence of the light scattering intensity (Figs. 3–6) is reminiscent of resonant behavior and can be explained by the motion of kinks described above. Fits of the experimental dependences $I(\omega)$ for geometries #1 and #2 at large frequencies $\omega > \omega_{\rm ext}$ to the func-

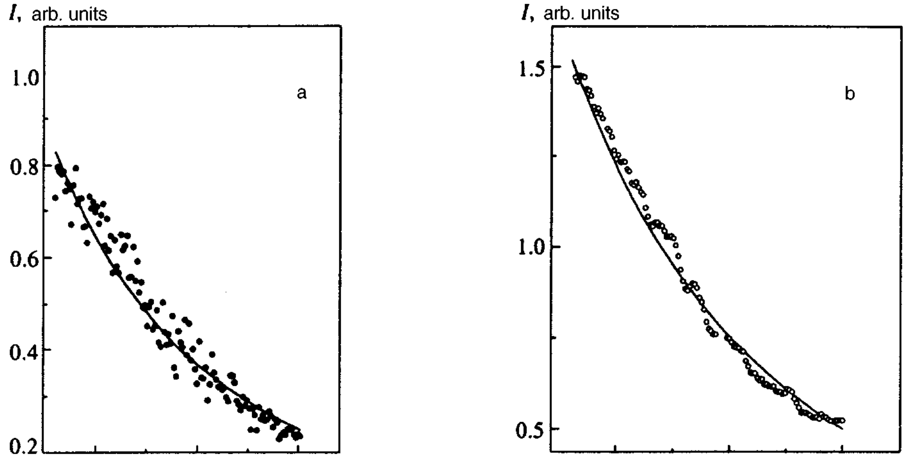


FIG. 3. Fit to the high-frequency limb of the function $I_{xz}(\omega)$ by the power-law ω^{-3} in geometry #1 (a) and #2 (b). In the experiments we used the liquid crystal 4-(2*s*,3*s*)-2-[chloro-3-methylpentanoyloxy]-4′-heptyloxybiphenyl with smectic *A*-smectic C^* phase transition temperature $T_{AC*}=43$ °C and spontaneous polarization varying within the limits 1300–290 nC/cm^2 in the C^* phase.

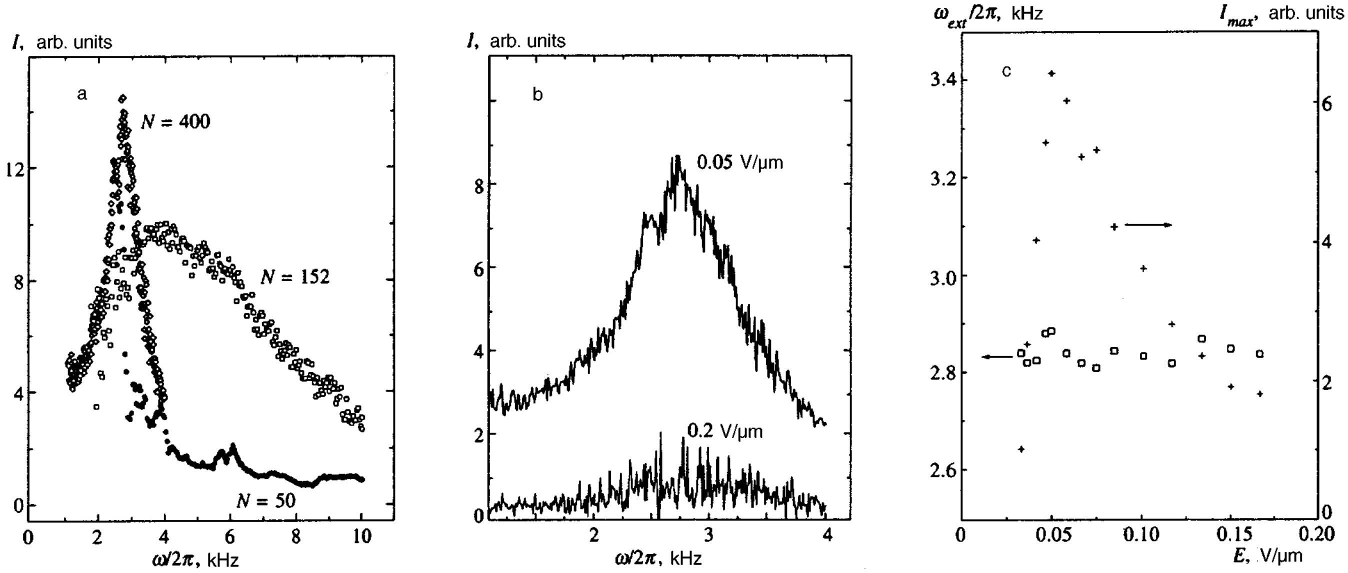


FIG. 4. Frequency characteristics of quasi-resonance light scattering in geometry #1 for a) different numbers of smectic layers N and b) amplitudes of the variable electric field. Figure 4c shows the resonance scattering characteristics as functions of the field amplitude for $N=300$, $T_{AC*}-T\approx 0.5$ °C.

tional form $G(\omega+B)^{-C}$ are shown in Fig. 3: the best fit results for the law $I(\omega)\sim\omega^{-3}$ with $G\sim a^2\sim 10^8\ \mathrm{s}^{-2}$ (the theoretical dependences were calculated for values of the parameter $a\geqslant 10^4\ \mathrm{s}^{-1}$).

Figure 4a presents data of an experiment in geometry #1 in which the thickness of the freely suspended film was decreased. The calculated curves for $a=\mathrm{const}$ and decreasing b exhibit qualitatively similar behavior: a decrease of $I_{\max}$ and an increase of ω_{ext} (cf. Figs. 4 and 2).

Figures 4b and c show that for fixed tilt angle θ in a freely suspended film an increase in the amplitude of the field leads to a decrease of the maximum intensity for essentially constant ω_{ext}, which is explained by relations (28) and (29). However, at very small values of $\tilde{E}$ the nature of the dependence $I_{\max}(\tilde{E})$ changes abruptly (Fig. 4c), which is explained by the disappearance of kinks for $b\ll 1$, when the width η of a kink becomes larger than the characteristic dimension of the film d. Indeed, as was noted above, estimates (24)–(26) for $a\propto b\propto\eta^{-1}\propto\tilde{E}$ show that for small $\tilde{E}$ the function $I_{\max}(\tilde{E})$ begins to grow steeply with increasing $\tilde{E}$.

Figure 5 illustrates the two types of temperature dependence of $\omega_{\max}$ and $I_{\max}$ for thin and thick freely suspended films. In the thin film disappearance of scattering ($I_{\max}\rightarrow 0$) and a critical increase of ω_{ext} are observed at some temperature while in the thick film no such critical temperature is

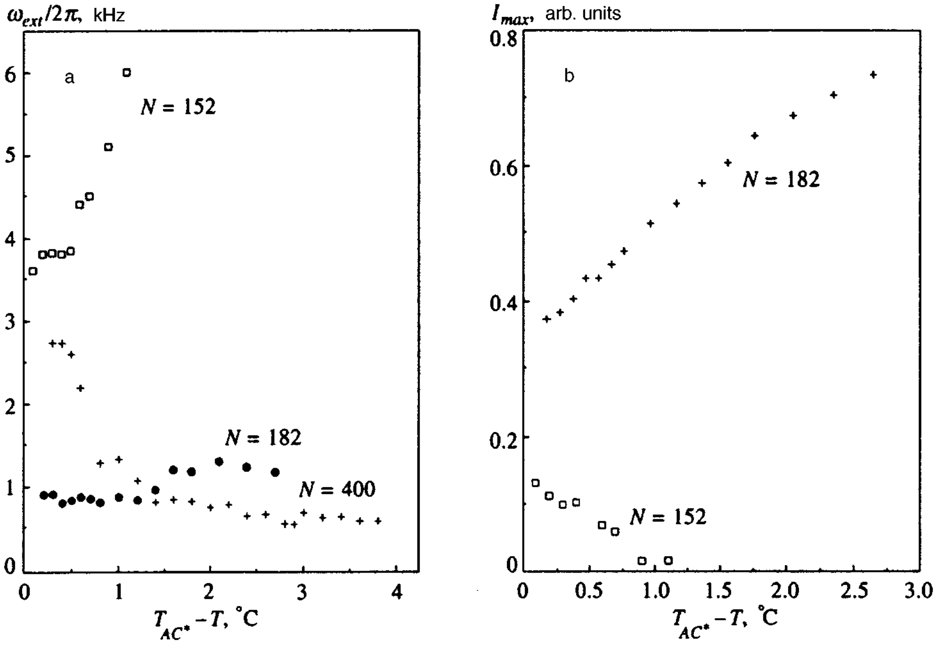


FIG. 5. Temperature dependence of the resonant frequency (a) and the scattering amplitude (b) in thin and thick freely suspended films.

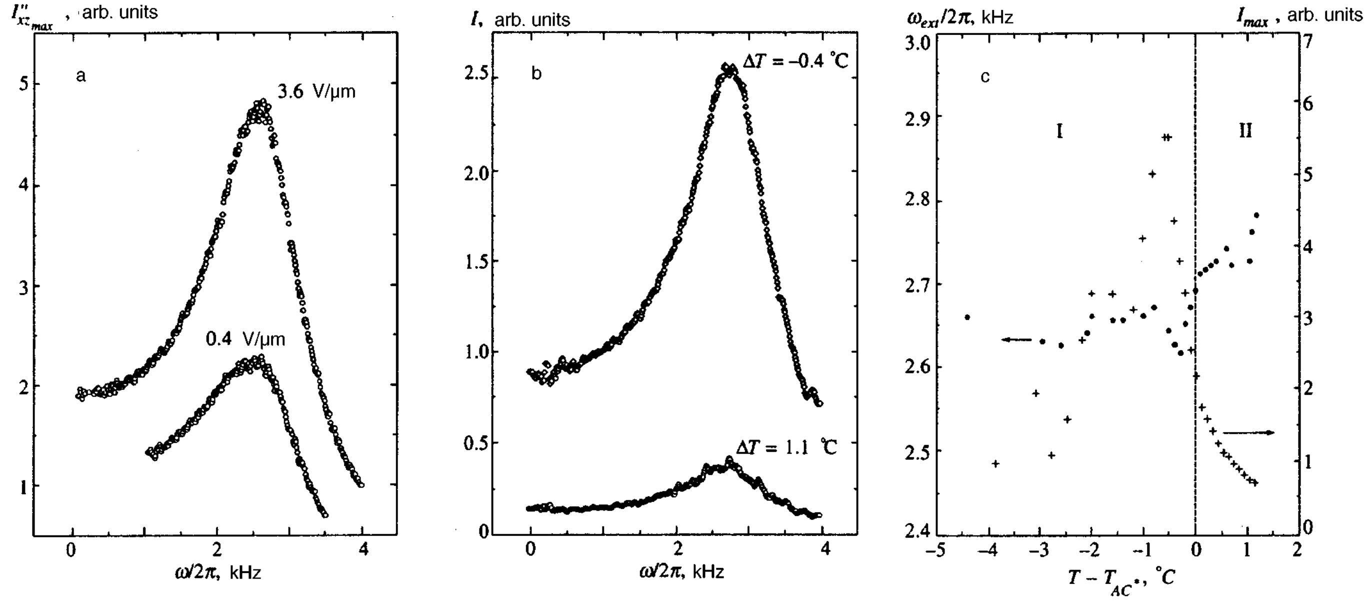


FIG. 6. Frequency characteristics of quasiresonant light scattering in geometry #2 a) for different amplitudes of the variable electric field for $\Delta = T_{AC*} - T = 0.5$ °C and b) for different temperatures. Figure 6c shows the resonance characteristics as functions of temperature; $d = 25$ μm; I—smectic C^*, II—smectic A.

observed, and $I_{\max}$ grows while $\omega_{\rm ext}$ decreases continuously with decrease of temperature. These data are easily explained with the help of relations (5)–(7) if it is assumed that in the thin film we have $U_0 < 0$ and in the thick film $U_0 \to 0$. The geometry of the investigated thin films (see Fig. 1a) can cause the appearance of the easy axis (z axis) and the negative value of U_0 since the film is practically homogeneous in z, but inhomogeneous in x and y due to the influence of the film surfaces and the electrodes. Here note should be made of a remarkable feature of such freely suspended films: banded domains exist in them spontaneously, parallel to the z axis, with characteristic width $\Delta x \sim 10^2$ μm.[18–22] Therefore for thin films (less than 1 μm), when the kink becomes thicker than the film, kinks can arise and migrate only in the xz plane of the film. At the same time, the width Δx of a domain becomes the characteristic dimension d and the parameter $b \sim \Delta x/\eta$ remains large. For very small film thicknesses (less than 0.1 μm), when the indicated domains disappear,[22] the time it takes the kinks to move across the film becomes very large, for which reason in the latter case we do not observe any quasiresonant light scattering phenomena.

A word should be said about the complexity of experimental studies of the dependence $I(\omega)$ at very low frequencies. These difficulties are connected with the proper choice of the instrument time constant $t_0 < \omega_{\rm ext}^{-1}$, which defines the upper integration limit in expressions (18)–(20). If the parameter t_0 is a random quantity, then the observed scattering intensity can have random jumps in the low-frequency region, which indeed took place in the experiment.

Another feature of the low-frequency dependence $I(\omega)$ shows up in films with small characteristic dimension d, i.e., for $b \ll 1$. As was noted above, in this case there is no maximum intensity; however, expressions of the form (22)–(24) remain valid for the frequency interval $a \ll \omega \ll b/a$, where in the integrals (18)–(20) the upper limit t_0 remains fixed, leading to the replacement of the functional dependence $I(\omega) \propto \omega^{-3}$ by the law $I(\omega) \propto \omega^{-2}$.

In thick films with geometry #2 in the case of crossed polarizers, experiment shows (see Fig. 6a) that the dependence of the maximum intensity $I''_{xz\max}$ on the field amplitude $\tilde{E}$ differs fundamentally from that for geometry #1. This maximum increases as the field increases at constant temperature near the phase transition temperature T_{AC*}, where the ratio of the maxima for temperatures below and above T_{AC*} is significantly greater than unity (see Fig. 6b). As the temperature increases through the phase transition temperature T_{AC*}, the quantity $I''_{xz_{\max}}$ falls rapidly while the quantity $\omega_{\rm ext}$ rapidly grows (see Fig. 6c). Such a qualitative difference between the experimental data for the geometries #1 and #2 is explained by several factors. First, as was noted above, at temperatures not too close to T_{AC*} and not for too large fields $\tilde{E}$, the chirality of the smectic C^* stabilizes the azimuthal helix of spontaneous polarization and hinders motion of the orientational kinks. In the immediate vicinity of the phase transition, when the effective field E is greater than the unwinding field $E_c \propto \theta$, kinks can move, and here the polar angle θ, which depends strongly on temperature and the applied field, determines the corresponding quasiresonant behavior according to relations (29). Making use of well-known thermodynamic expressions for the function $\theta(T,E)$ in these simple cases and taking relations (29) into account, we obtain qualitative dependences for the characteristics of the quasiresonance peak:

$$\begin{aligned} \omega_{\rm ext} &\propto \frac{1}{\sqrt{T_{AC*}-T}} - \frac{\text{const}\cdot \mu E}{\sqrt{(T_{AC*}-T)^3}}, \\ I_{xz_{\max}} &\propto \sqrt{(T_{AC*}-T)^3} + \text{const}\cdot \mu E \end{aligned} \tag{31}$$

in the phase C^*, when $\theta \propto \sqrt{T_{AC*}-T}$ for $E<\mathrm{const}\sqrt{(T_{AC*}-T)^3}$;

$$\omega_{\mathrm{ext}} \propto E^{-1/3}, \quad I_{xz_{\max}} \propto E \tag{32}$$

in phases C^* and A for $E>\mathrm{const}\cdot\sqrt{|T_{AC*}-T|^3}$ and $\theta \propto E^{1/3}$; and

$$\omega_{\mathrm{ext}} \propto \frac{T-T_{AC*}}{E}, \quad I_{xz_{\max}} \propto \frac{E^3}{(T-T_{AC*})^3} \tag{33}$$

in phase A for $E<\mathrm{const}\cdot\sqrt{(T_{AC*}-T)^3}$ and $\theta \propto E/(T-T_{AC*})$. Expressions (31)–(33) provide a completely adequate explanation of the experimental data graphed in Fig. 6c.

To estimate the parameters a and b in the kink model and describe the experimental data, we use the following characteristic orders of magnitude:

$$\mu \sim 10^{-3}\ \mathrm{C/m^2}, \quad K \sim 10^{-12}\ \mathrm{N}, \quad \gamma \sim 10^{-1}\ \mathrm{Pa\cdot s},$$

$$\widetilde{E} \sim 3\cdot 10^6\ \mathrm{V/m}, \quad d \sim 10^{-5}\ \mathrm{m}, \quad \varepsilon_a \sim 10^{-11}\ \frac{\mathrm{m^2N}}{\mathrm{C^2}},$$

$$\theta \sim 10^{-1} \quad \text{for } \Delta T = T_{AC*}-T \sim 1\ \mathrm{K}. \tag{34}$$

Thus we obtain

$$a=\frac{\mu\widetilde{E}}{\gamma\theta} \sim 3\times 10^5\ \mathrm{s^{-1}}, \quad \eta \sim \frac{1}{\widetilde{E}}\sqrt{\frac{K}{\epsilon_a}} \sim 10^{-7}\ \mathrm{m},$$

$$b=\frac{d}{\eta} \sim 10^2, \quad \omega_{\mathrm{ext}} \sim 3\times 10^3\ \mathrm{s^{-1}}.$$

In fact the quantities γ, K, and ϵ_a are not known exactly; however, they can be estimated from the experimental data by choosing the parameters a and b that most adequately fit the experimental data.

5. CONCLUSION

The observed frequency dependence of the light scattering intensity I_{xz} is reminiscent of resonant behavior and is explained by the motion of the orientational kinks discussed above. Here a substantial role is played by the boundaries of the homogeneous layer of the C^* smectic phase. Under the action of an external field such isolated kinks can form on these boundaries, after which they move into the interior of the ferroelectric film. Both the surface of the film and the domain walls, in the case in which a domain structure is formed in the film, can serve as such boundaries. It makes sense to speak of the motion of kinks if the distance between the characteristic boundaries exceeds the width of a kink. Therefore in thick films, whether freely suspended or found between the solid electrode surfaces, isolated kinks can move in three dimensions, whereas in the case of thin films they can move only along its surface. The helical structure of the C^* smectic is unfavorable for the appearance of kinks, and therefore the above-discussed quasi-resonant behavior of the light scattering characteristics is not observed in the C^* phase far from the transition point T_{AC*}, where the helix exists. Since the critical unwinding field of the helix is $E_c \sim q_0^2 K\theta/\mu$, where q_0 is the wave number of the unperturbed helix and $\theta^2 \sim (T_{AC*}-T)/T_{AC*}$, for $\mu \sim 10^{-3}$ C/m^2, $K \sim 10^{-12}$ N, $q_0 \sim 10^7$ m^{-1}, and $E \sim 10^4$ V/m in thick films the described effects can take place in a narrow temperature interval about T_{AC*}, specifically for $\Delta T = (T_{AC*}-T) \leqslant \Delta T_c \sim T_{AC*}(\mu E/Kq_0^2)^2 \sim 1$ K, where $\theta(\Delta T_c) \sim 10^{-1}$.

In thin films, where the orientational helix is absent thanks to the influence of the boundaries, these effects exist over a wider temperature interval.

The motion of the kinks is substantially influenced by the dielectric anisotropy in the plane of the smectic layer. This anisotropy can be due to the two-dimensional nature of the C^* smectic, the anisotropy of the dipole–dipole interactions, and electrical effects of the outer boundaries. Specific combinations of these conditions lead to various regimes of kink motion and, accordingly, to different temperature and field dependences of the characteristics of quasiresonant scattering of polarized light. Guided by the experimental data, with the help of the dynamic kink model discussed above we have arrived at estimates of the unknown material parameters, namely the orientational viscosity and elasticity, and also the dielectric anisotropy.

It is of interest to study the situation in which the ferroelectric liquid-crystal film has not one, but several different characteristic scales d, e.g., as a result of the coexistence of several domain structures. In this case, one can expect the appearance of several quasiresonances in the light scattering intensity, corresponding to different characteristic frequencies. We explored this possibility in the case of the ferroelectric liquid crystal FLC-240 ($T_{AC*}=53.5$ °C), which has a relatively high spontaneous polarization $P=40$ nC/cm^2. In Refs. 23–25 it was shown that such liquid crystals in an ordinary electro-optical cell possess a specific static domain structure which is characterized by a strong temperature dependence $d_1(T)$ of the domain widths. Along with such domains, under certain critical conditions in these liquid crystals a dynamic domain structure appears, caused, for example, by an electrohydrodynamical instability which is characterized by a strong frequency dependence of the width of the domains $d_2(\omega)$ (Ref. 5), where the inequality $d_2(\omega)<d_1(T)$ is entirely possible. Consequences of this kind require additional study. In the present paper we have only presented evidence of the appearance of two quasiresonances (see Fig. 7) for different values of the frequency of the applied electric field which adequately correspond to the two observed domain structures and depend substantially on the temperature and field amplitude.

ACKNOWLEDGMENTS

We would like to express our deep gratitude to L. A. Beresnev for providing a quantity of the liquid crystal FLC-240, to U. Hoffmann and S. Astaf'ev for technical assistance, and to the German Society of Natural Scientists (E.I.D) for support. This work was carried out with the partial support of INTAS (Grant No. 94-4078) and the Russian Foundation for Basic Research (Grant No. 96-02-16667).

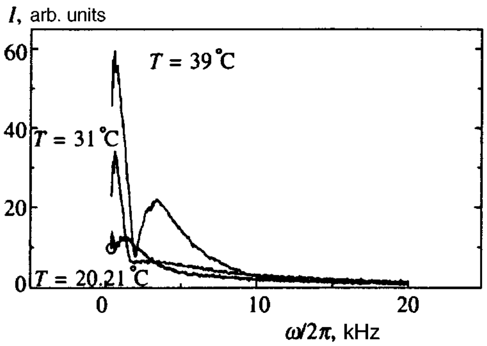


FIG. 7. Frequency dependence of scattering intensity with two maxima, observed in the smectic liquid crystal FLC-240 with coexistence of two domain structures.

APPENDIX A

We consider a class of nonlinear equations of motion which can be written formally as

$$\left(\Xi+\sum_i \frac{\partial^2 u}{\partial s_i^2}\right)\sin\varphi-\frac{1}{2}\sum_i\left(\frac{\partial u}{\partial s_i}\right)^2\sin 2\varphi+\sum_i\frac{\partial^2\varphi}{\partial s_i^2}=\frac{\partial\varphi}{\partial\tau}, \tag{A1}$$

where $u=u(\mathbf{s})$ is an arbitrary function, $\Xi=\tilde{\Xi}\cos(\nu\tau)$, $\varphi=\varphi(\mathbf{s},\tau)$, and $\mathbf{s}$ and τ are dimensionless variables. If it is assumed that at the time $\tau=0$ (the time at which the field is switched on) such a region exists inside the ferroelectric liquid-crystal film, bounded by the line $u(\mathbf{r})=u_0=\text{const}$, and that inside it we have $\varphi(\mathbf{s},\tau=0)=0$ and outside it $\varphi(\mathbf{s},\tau=0)=\pi$ or *vice versa*, where the width $|\Delta\mathbf{s}|\sim 1$ of the transition region is much smaller than the characteristic dimensionless film parameter s, then it can be easily seen that Eq. (A1) has the exact solutions

$$\varphi_1^{\pm}=\pm\arctan\frac{1}{\sinh[u(\mathbf{s})-u_0-(\tilde{\Xi}/\nu)\sin(\nu\tau)]}. \tag{A2}$$

The form of the solution (A2) is uniquely determined by the following relations:

$$\sin\varphi_1^{+}=\frac{1}{\cosh R_1},\quad \cos\varphi_1^{+}=\tanh R_1,$$

$$\sin 2\varphi_1^{+}=\frac{2\sinh R_1}{\cosh^2 R_1},$$

$$\frac{\partial\varphi_1^{+}}{\partial s_i}=-\frac{1}{\cosh R_1}\frac{\partial R_1}{\partial s_i},\quad \frac{\partial\varphi_1^{+}}{\partial\tau}=-\frac{1}{\cosh R_1}\frac{\partial R_1}{\partial\tau},$$

$$\frac{\partial^2\varphi_1^{+}}{\partial s_i^2}=-\frac{1}{\cosh R_1}\frac{\partial^2 R_1}{\partial s_i^2}+\frac{\sinh R_1}{\cosh^2 R_1}\left(\frac{\partial R_1}{\partial s_i}\right)^2, \tag{A3}$$

where

$$R_1(\mathbf{s},\tau)=u(\mathbf{s})-u_0-\frac{\Xi}{\nu}\sin(\nu,\tau). \tag{A4}$$

The solutions (A2) describe the motion of isolated orientational kinks in the function $\varphi(\mathbf{s},\tau)$. For example, for $\nu=0$ they describe the motion of a kink in a constant field along the s axis (the solutions with subscripts 1 and 2 correspond to motion in opposite directions):

$$\begin{aligned}\varphi_1^{\pm}(s,\tau)&=\pm\arctan\frac{1}{\sinh(s-s_0-\Xi\tau)},\\ \varphi_2^{\pm}(s,\tau)&=\pm\arctan\frac{1}{\sinh(s-s_0+\Xi\tau)},\end{aligned} \tag{5}$$

with constant velocity $V^{\pm}=ds/dt=\pm\Xi$, where the plane of the front is perpendicular to the direction $\mathbf{s}$.

The first term on the left-hand side of Eq. (A1) describes the linear interaction between the spontaneous polarization and the effective field containing a homogeneous part Ξ (the external field) and the contribution of some internal field due to the electrical properties of the liquid-crystal layer. The second term, which is quadratic in the polarization, describes the influence of the anisotropy energy on the motion of the spontaneous polarization vector; here the anisotropy energy density of the smectic layer can be locally inhomogeneous, thanks to the influence of the film surface, domain walls, and structural defects. The third term gives the usual contribution of the orientational elasticity energy. The right-hand side of Eq. (A1) describes the contribution of the orientational viscosity and the corresponding moment of the friction forces. The dimensionless variables $s_i\equiv x_i/\eta_i$ and $\tau\equiv t/t_0$ reflect the existence of characteristic spatial (η_i) and temporal (t_0) scales, defined by the material parameters of the liquid crystal.

Equation (A1) makes it possible to estimate, at least qualitatively, the influence of inhomogeneity of the ferroelectric liquid−crystal film on the repolarization processes. We will give a few examples.

1) Let the anisotropy energy density in the plane of the smectic layer (the s_1s_2 plane) in the vicinity of some point defect vary quadratically in the coordinates s_1 and s_2, and the effective field be homogeneous and constant, i.e.,

$$\begin{aligned}&\left(\frac{\partial u}{\partial s_1}\right)^2+\left(\frac{\partial u}{\partial s_2}\right)^2=4(s_1^2+s_2^2),\\ &\frac{\partial^2 u}{\partial s_1^2}+\frac{\partial^2 u}{\partial s_2^2}=\text{const}.\end{aligned} \tag{A6}$$

In this case the functions $u(s_1,s_2)$ may take the following forms:

$$s_1^2\pm s_2^2,\quad 2s_1s_2. \tag{A7}$$

Depending on the shape of the line $u(s_1,s_2)=u_0$ describing the surface of constant phase $\varphi=\pi/2$ at the initial time $\tau=0$, the motion of the phase front $\varphi(s_1,s_2,\tau)=\pi/2$, i.e., an orientational kink, can take various forms over the course of time τ. According to Eqs. (A5) and (A7), the corresponding fronts can have the form of circles, parabolas, and hyperbolas, i.e., the following proportional dependences are possible:

$$s_1^2\pm s_2^2\propto\tau,\quad 2s_1s_2\propto\tau. \tag{A8}$$

2) Let the anisotropy energy density fall off exponentially according to the law $\exp(-2s_1)$ in some direction s_1 in

the smectic layer, for example, with depth into the film away from a planar defect (the surface of the film), and let the effective field be homogeneous and constant, i.e.,

$$\left(\frac{\partial u}{\partial s_1}\right)^2 \pm \left(\frac{\partial u}{\partial s_2}\right)^2 = e^{-2s_1}, \quad \frac{\partial^2 u}{\partial s_1^2} + \frac{\partial^2 u}{\partial s_2^2} = 0.$$

In this case the functions $u(s_1, s_2)$ may take the forms

$$e^{\pm s_1} \sin s_2, \tag{A9}$$

which lead to wedge-shaped fronts, where the wedges are arrayed periodically along an isolated surface (along the s_2 axis) with period equal to the effective penetration depth of the front along the s_1 axis.

The above simple examples, of course, do not exhaust the shapes of kink motion in a constant field. It is clear that the shapes and velocities of such fronts depend strongly on the local properties of the ferroelectric liquid-crystal films. In the case of the action of a variable field Ξ, the solution (A2) shows that the kinks execute oscillatory motion relative to some initial position with frequency ν and amplitude proportional to $\tilde{\Xi}/\nu$.

Generally speaking, the boundary conditions on the surfaces of a film of finite thickness have the form[14,15]

$$Q \sin \varphi|_{\text{sur}} = G \left.\frac{\partial \varphi}{\partial s}\right|_{\text{sur}}, \tag{A10}$$

where Q is the effective parameter of polar adhesion to the surface and G is the effective elastic coupling constant. Noting the relations

$$\sin \varphi_1^{\pm} = \pm \frac{1}{\cosh(s - s_0 - \tilde{\Xi}\tau)} = -\frac{\partial \varphi_1^{\pm}}{\partial s},$$

$$\sin \varphi_2^{\pm} = \pm \frac{1}{\cosh(s - s_0 + \tilde{\Xi}\tau)} = \frac{\partial \varphi_2^{\pm}}{\partial s}.$$

We can rewrite Eq. (A10) in the form

$$\frac{Q \pm G}{\cosh(s - s_0 - \tilde{\Xi}\tau)|_{\text{sur}}} = 0,$$

$$\frac{Q \mp G}{\cosh(s - s_0 + \tilde{\Xi}\tau)|_{\text{sur}}} = 0.$$

Thus, condition (A10) is approximately fulfilled for $\tilde{\Xi}\tau \gg 1$ and $(s - s_0)|_{\text{sur}} \gg 1$. In this case, the solutions (A5) can have the following meaning: if an orientational kink appears near the film surface having width $\Delta s_x \sim 1$ $(\Delta x \sim \eta)$, then it can move a distance $\tilde{\Xi}\tau \gg 1$ in the s_x direction during the time τ, and in this case boundary condition (A10) is satisfied with exponential accuracy.

If the parameter Q is positive, then the derivative $(\partial\varphi/\partial s)|_{\text{sur}}$ in Eq. (A10) should also be positive. In this case the solutions (A5) describe the motion of a $-\pi$-kink along the s_x axis and a $+\pi$-kink in the opposite direction. For negative values of Q the solutions (A5) describe the motion of a $+\pi$-kink along the s_x axis and a $-\pi$-kink in the opposite direction.

[1] K. Yoshino, T. Uemoto, and Y. Inuishi, Jpn. J. Appl. Phys. **16**, 571 (1977).

[2] J. Hoffmann, W. Kuszyński, and J. Malecki, Mol. Cryst. Liq. Cryst. **44**, 287 (1978).

[3] A. Levstik, B. Žekš, I. Levstik *et al.*, J. de Phys. **40**, 303 (1979).

[4] R. Blinc and B. Žekš, Phys. Rev. A **18**, 740 (1978).

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

[6] A. Levstik, T. Carlsson, C. Filipič, I. Levstik, and B. Žekš, Phys. Rev. A **35**, 3527 (1987).

[7] S. Nonaka, K. Ito, M. Isogai, and M. Odamura, Jpn. J. Appl. Phys. **26**, 1609 (1987).

[8] M. Ozaki, T. Hatai, and K. Yoshino, Jpn. J. Appl. Phys. **27**, 1906 (1988).

[9] A. M. Biradar, Š. Wrobel, and W. Haase, Phys. Rev. A **39**, 2693 (1989).

[10] T. Carlsson, B. Žekš, C. Filipič, and A. Levstik, Phys. Rev. A **42**, 877 (1990).

[11] F. Gouda, K. Sharp, G. Anderson *et al.*, Jpn. J. Appl. Phys. **28**, 1887 (1989).

[12] F. Gouda, G. Anderson, S. T. Lagerwall *et al.*, Liq. Cryst. **6**, 219 (1989).

[13] P. Schiller, G. Pelzl, and D. Demus, Liq. Cryst. **2**, 21 (1987).

[14] S. A. Pikin, Mol. Cryst. Liq. Cryst. **179**, 201 (1990).

[15] S. A. Pikin, Ferroelectrics **117**, 197 (1991).

[16] E. I. Demikhov, S. A. Pikin, and E. S. Likina, Phys. Rev. E **52**, 6250 (1995).

[17] E. I. Demikhov, S. A. Pikin, and E. S. Likina, J. de Phys. II **6**, 753 (1996).

[18] E. I. Demikhov, Europhys. Lett. **25**, 259 (1994).

[19] E. I. Demikhov and H. Stegemeyer, Liq. Cryst. **18**, 37 (1995).

[20] E. I. Demikhov, Phys. Rev. E **51**, 12 (1995).

[21] E. I. Demikhov and S. A. Pikin, JETP Lett. **61**, 686 (1995).

[22] E. I. Demikhov, E. Hoffmann, H. Stegemeyer *et al.*, Phys. Rev. E **51**, 5954 (1995).

[23] L. A. Beresnev, M. V. Loseva, N. I. Chernova *et al.*, JETP Lett. **51**, 584 (1990).

[24] S. A. Pikin, L. A. Beresnev, S. Hiller *et al.*, Mol. Mater. **3**, 1 (1993).

[25] L. A. Beresnev, E. Schumacher, S. A. Pikin *et al.*, Jpn. J. Appl. Phys. **34**, 2404 (1995).

Translated by Paul F. Schippnick

The metal–insulator transition and the phase transition in metal–ammonia solutions

A. A. Likal'ter

Institute for High Temperatures, Russian Academy of Sciences, 127412 Moscow, Russia

(Submitted 15 July 1996)
Zh. Éksp. Teor. Fiz. **111**, 938–948 (March 1997)

The ground state of impurity metal (sodium) atoms in liquid ammonia close to the solvated state of the free electrons is considered. It is shown that the critical solubility point lying on the metal side of the metal–insulator transition is determined by the Coulomb interaction between the ions and electrons in the overlapping impurity states, classically accessible spheres of which form an infinite percolation cluster. The percolation conductivity via the impurity states is estimated. The estimate agrees with the experimental data near the critical solubility point.

[S1063-7761(97)01403-0]

1. INTRODUCTION

For a long time metal solutions in liquid ammonia have attracted attention by virtue of the unique variety of effects observable in them, yet they belong to the class of disordered electron systems which have received of yet little theoretical study. From the theoretical viewpoint, they are of interest not only because of their unique properties, but also the general properties which define the behavior of apparently completely different physical systems. In dilute solutions with metal concentration less than 1% the atoms are almost completely dissociated into ions and solvated electrons. Of especial interest is the possibility, as the concentration is increased, of observing a continuous transition from a weakly conducting electrolyte to a metal solution with conductivity greater than $10^2 \Omega^{-1}\ \mathrm{cm}^{-1}$. The strong variation in their properties links metal–ammonia solutions with other electron systems undergoing a metal-nonmetal transition. They are closest in their properties to expanded metals and heavily doped semiconductors.[1]

The analogy with semiconductors is based on the high dielectric constant of liquid ammonia, thanks to which the metal atoms can form impurity electron states whose radius is greater than the distance between the solvent molecules. In contrast to semiconductors, the impurity states in ammonia have been little studied, since the main role in dilute solutions is played by solvated electrons. Localization of solvated electrons is the reason why dilute metal–ammonia solutions are analogous in many of their properties to electrolytes.[2] However, as in heavily doped semiconductors in which the impurity states overlap,[3] concentrated solutions have metallic properties.

It is noteworthy that solutions of some metals, in particular sodium, allow one to observe the phase transition with stratification into a metallic and a dielectric phase with different concentrations. As was noted long ago, this transition, reaching completion at the critical solubility point, is analogous to the condensation of metal atoms.[4] A connection between this phase transition and the metal–insulator transition has also been noted.[5] The metallic electrical conductivity at the critical solubility point of sodium-ammonia solutions shows that near this point two metallic phases can be found in equilibrium having different concentrations of the dissolved metal.

The recently developed theory of plasma critical points of metals is based on their being found near the metal–nonmetal transition point.[6,7] Therefore, they are characterized by a metallic state with strong electron–ion coupling which is described in terms of the virtual atomic structure model. In particular, the position of the plasma critical points is determined by the interaction of virtual atoms with overlapping electron shells. It may be expected that the critical solubility points of metal solutions are also determined by the interaction of overlapping impurity states.

Despite the apparent obviousness of the above analogies, metal–ammonia solutions have their own idiosyncracies, rendering their theoretical treatment more difficult. These include, in particular, frequency dispersion and nonlinearity of the dielectric constant, caused respectively by the inertia of the dipolar molecules and saturation of their polarization near the ions. In this paper we estimate the parameters of the impurity states and extend the theory of plasma critical points to metal–ammonia solutions.

The plan of the paper is as follows. Section 2 introduces the dielectric model of an electron solution, which is used to describe the critical solubility point. Section 3 describes a model of the impurity states playing the main role in connection with the percolation metal–insulator transition. Section 4 presents a theory of conductivity in the vicinity of the transition point and compares the results of this theory with the experimental data. Section 5 is the conclusion.

2. DIELECTRIC MODEL OF THE SOLUTION

The dielectric constant of liquid ammonia associated with electron polarization of the molecules at optical frequencies is roughly equal to 2. The high dielectric constant at low frequencies, becoming as large as 22, is due mainly to orientation of the constant dipoles of the ammonia molecules. Thanks to the delayed response of the molecules, a unique form of self-interaction of the electron takes place which can be described (outside the preferred localization region) by an asymptotic potential energy $-(\varepsilon_\infty^{-1}-\varepsilon_0^{-1})e^2/r$, where e is the electron charge, and ε_∞ and ε_0 are respectively the optical and static dielectric constants.[8] The corresponding asymptotic limit for the impurity state obtained by adding the potential energy of the electron in the field of an ion $-e^2/\varepsilon_0 r$ is simply

$$u=-e^2/\varepsilon_\infty r. \tag{1}$$

The radius of the classical accessible sphere in potential (1) is equal to

$$R_a=e^2/\varepsilon_\infty I, \tag{2}$$

where I is the binding energy of the impurity states.

A characteristic parameter is the fraction of the volume occupied by the classically accessible spheres

$$\zeta=\frac{4\pi}{3}R_a^3 n=\left(\frac{R_a}{R_s}\right)^3, \tag{3}$$

where n is the density of the metal atoms and R_s is the radius of the Wigner–Seitz sphere

$$R_s=(4\pi n/3)^{-1/3}.$$

The virtual structure of the overlapping impurities in metal solutions is assumed to lie in the interval between the percolation threshold $\zeta\approx 1/3$ (identified with the metal-insulator transition point) and random dense packing of spheres $\zeta\approx 2/3$.

The interaction between the overlapping impurity states in a percolation cluster is characterized by the Coulomb coupling parameter

$$\Gamma=e^2/\varepsilon_e R_s T, \tag{4}$$

where T is the temperature. Since this interaction is governed by the region of preferred localization of the electrons, the static dielectric constant enters into the coupling parameter. The effective value of the dielectric constant ε_e, as was mentioned above, differs from its macroscopic value due to saturation of the polarization of the dipolar molecules near the ions.

Another dimensionless parameter is the ratio of the temperature to the ionization potential. It is not independent, since it can be expressed in terms of the parameters ζ and Γ already introduced. In particular, formulas (2)–(4) yield the ratio

$$\frac{T_c}{I}=\frac{\varepsilon_\infty \zeta_c^{1/3}}{\varepsilon_e \Gamma_c}, \tag{5}$$

for the critical solubility point. Here the subscript c indicates a critical parameter. Assuming similarity of plasma critical points, one might expect that the dimensionless parameters ζ_c and Γ_c are universal. The ratio T_c/I, however, is not universal. According to Eq. (5) it depends on the ratio of the dielectric constants.

Using known values for the parameters $\zeta_c\approx 0.365$ and $\Gamma_c\approx 17$ (Refs. 7 and 9), formulas (2) and (3) yield the following expression for the critical density

$$n_c=2.92\times 10^{19}(\varepsilon_\infty I)^3, \tag{6}$$

and formula (5) yields the expression

$$T_c=0.0425\frac{\varepsilon_\infty}{\varepsilon_e}I, \tag{7}$$

for the critical temperature. I and T_c are measured in electron volts, and n_c in cm^{-3}.

When the density corresponding to the critical concentration $x_c=0.0415$ and the critical temperature $T_c=230$ K (Ref. 10) is substituted into Eqs. (6) and (7), Eq. (6) yields the estimate $I=1.665$ eV for the binding energy and Eq. (7) the estimate $\varepsilon_e=6.9$ for the effective dielectric constant. The binding energy of the impurity states turns out to be near the binding energy of the solvated electron. This estimate is backed up below by a calculation for a simple model of the impurity states.

Let us consider how to explain the relatively small effective dielectric constant. First of all, the minimum scale at which this quantity can have meaning is greater than or of the order of the mean distance between the molecules

$$R_0=(4\pi N/3)^{-1/3},$$

where N is the density of the ammonia molecules. At such distances the ion field is large enough for saturation of polarization. The polarization vector of the dipolar molecules in the field of an ion is equal to

$$P=\chi\mathscr{E}=\frac{\varepsilon_e(r)-\varepsilon_\infty}{4\pi}\frac{e}{\varepsilon_e(r)r^2}, \tag{8}$$

where χ is the dielectric susceptibility of the dipoles, $\mathscr{E}$ is the intensity of the electric field of the ion, $\varepsilon_e(r)$ is the effective dielectric constant, which for polarization saturation becomes a function of the distance from the ion. The magnitude of the saturation polarization is given by

$$P=dN, \tag{9}$$

where $d=0.58ea_B$ is the dipole moment of the ammonia molecules and a_B is the Bohr radius. Equating expressions (8) and (9) and solving the resulting equation for $\varepsilon_e(r)$, we obtain

$$\varepsilon_e(r)=\frac{\varepsilon_\infty}{1-r^2/R_d^2}, \quad R_0<r<R_d, \tag{10}$$

where R_d is the polarization saturation radius,

$$R_d=\sqrt{\frac{e}{4\pi dN}}.$$

Schematically the dielectric constant is given by formula (10) for those values of r for which ε_e is less than its macroscopic value, and is equal to ε_0 at large distances (Fig. 1). However, before it can be used in the equations of electrostatics, it must be averaged over space. Since the potential of a single ion can be represented as an integral of $e/\varepsilon_e(r)r^2$, the average of the dielectric constant over a Wigner–Seitz cell is found as follows:

$$\frac{1}{\varepsilon_e}=\left(\frac{1}{R_0}-\frac{1}{R_s}\right)^{-1}\int_{R_0}^{R_s}\frac{dr}{\varepsilon_e(r)r^2}.$$

For conditions at the critical solubility point, averaging gives $\varepsilon_e\approx 9$, which is in reasonable agreement with the estimate obtained above based on the critical parameters. Thus, as a consequence of polarization saturation, the average dielectric constant, defining the interaction between the impurity states

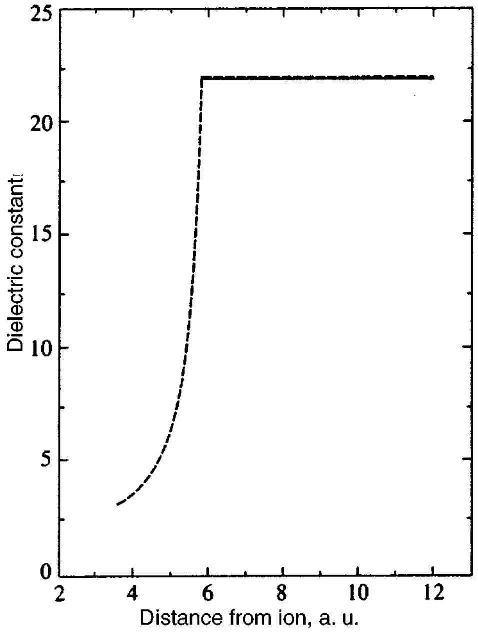


FIG. 1. Effective dielectric constant, with allowance for saturation of polarization of dipolar molecules at small distances from the ion. The radius of polarization saturation is $R_d = 5.72a_B$.

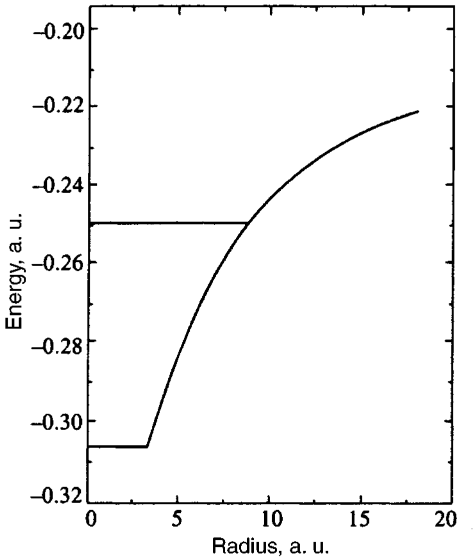


FIG. 2. Pseudopotential and impurity level $E = -0.2492$ a.u., obtained by solving the Schrödinger equation. Ionic core radius $R_c = 3.265a_B$.

in a percolation cluster, is roughly three times smaller than its macroscopic value, so the treatment of the critical solubility point based on similarity with plasma critical points of metals is self-consistent.

3. IMPURITY STATES

For a more detailed analysis of the impurity states, we choose the Heine–Abarenkov pseudopotential (constant inside the ionic core), which reproduces the energy level of the free atom.[11] In ammonia solution the ionic core of the impurity atom is smaller than the molecular cell, and the Coulomb potential outside the core falls off as $1/r$ and is inversely proportional to the dielectric constant. With allowance for the polarization field of the constant solvent dipoles induced by the impurity electron cloud, the pseudopotential takes the form

$$U = \begin{cases} -e^2/R_c, & r < R_c, \\ -[1/\varepsilon_e + (1/\varepsilon_\infty - 1/\varepsilon_e)\nu(r)]e^2/r + C, & r > R_c, \end{cases} \tag{11}$$

where the factor $\nu(r)$ describes the deviation of the electric potential of the electron cloud from a Coulomb potential, and the constant C is determined by the condition of continuity.

The potential of the electron cloud satisfies the Poisson equation

$$\frac{1}{r}\frac{d^2}{dr^2}(r\varphi) = \frac{4\pi e \psi^2(r)}{\varepsilon_e}, \tag{12}$$

where $\psi(r)$ is the wave function of the electron. Since the dependence of the potential on the charge distribution has an integrated character, we may substitute a quite simple test wave function into this equation. In particular, we use a pseudo wave function that differs from zero only outside the ionic core

$$\psi(r) = \frac{B}{\sqrt{4\pi}} e^{-kr}, \quad r > R_c, \tag{13}$$

where

$$k = \sqrt{2mI/\hbar^2},$$

m is the electron mass, $\hbar$ is Planck's constant, and B is a normalization factor,

$$B = 2k^{3/2}\exp(kR_c)/\sqrt{1 + 2kR_c + 2k^2R_c^2}.$$

The solution of the Poisson equation has the form

$$\varphi = -e\nu(r)/\varepsilon_e r, \tag{14}$$

where

$$\nu(r) = 1 - \frac{1 + kr}{1 + 2kR_c + 2k^2R_c^2}\exp[-2k(r - R_c)]. \tag{15}$$

For r near the classically accessible radius R_c, the pseudopotential defined by formulas (11) and (15) is close to a Coulomb potential (Fig. 2).

The impurity level is found from the Schrödinger equation (in atomic units)

$$\frac{d^2\chi}{dr^2} + 2(E - U)\chi = 0, \tag{16}$$

where $\chi = rR(r)$, $R(r)$ is the radial wave function, E is the energy, and the boundary conditions are

$$\chi(0) = 0, \quad z = \frac{d\ln\chi(\infty)}{dr} < 0.$$

Solving the Schrödinger equation for the core region gives the additional boundary condition

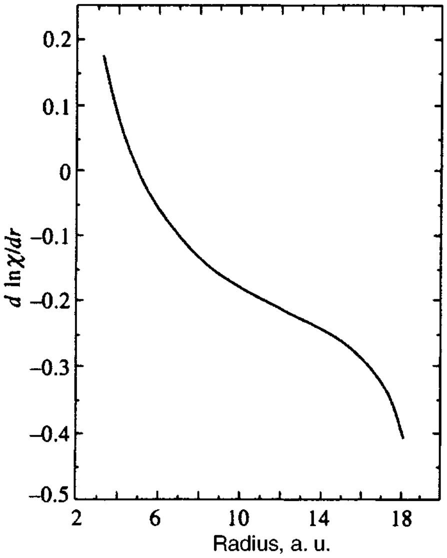


FIG. 3. Numerical solution for the logarithmic derivative of the pseudo wave function $z=d\ \ln\chi/dr$. The impurity level is subtended between two near-lying energy values for which $z(r)$ diverges at large r with negative sign (as in the figure) or positive sign (not shown).

$$z(R_c)=\kappa\ \cot(\kappa R_c),\quad \kappa=\sqrt{2(E+1/R_c)}. \tag{17}$$

In the outer region it is convenient to solve the equivalent nonlinear first-order equation

$$\frac{dz}{dr}+z^2+2(E-U)=0 \tag{18}$$

instead of Eq. (16) (Ref. 12). The impurity level is subtended between near-lying energy levels at which $z(r)$ diverges at large distances with positive and negative sign (Fig. 3). The binding energy, which is only weakly sensitive to variation of the parameters (in particular, the active dielectric constant), is equal to 1.53 eV, which agrees with the estimate based on the critical solubility point. In practice, however, much higher accuracy is needed. In particular, the binding energy of the impurity states should be greater than the binding energy of the solvated states (for which normalization to the experimental optical spectra gives 1.6 eV, Ref. 1).

4. PERCOLATION CONDUCTIVITY

In the vicinity of the metal–insulator transition point the conductivity is due to electron diffusion between the classically accessible spheres of the virtual impurity states. The solvated electrons diffuse along with them. Assuming that the classically accessible radius of the solvated states is close to that of the impurities, from the percolation viewpoint they may be considered as equivalent. With the exception of the latter point, the metal-nonmetal transition is completely analogous to the percolation transition in impurity semiconductors.[13] The transition point corresponds to the percolation threshold of classically accessible spheres of electrons. The Fermi energy of the mixed impurity (solvated) states with continuous energy spectrum depends on proximity to the transition point. As one approaches this point, the ratio of the Fermi energy to the temperature tends to zero, so that excitations of the impurities (solvated electrons) are described by Boltzmann statistics. The conductivity below this point is due to thermal excitation of the electrons to the percolation level, i.e., it has an activation character. Without repeating all the arguments,[13] we may simply review here the main relations needed to describe percolation conductivity.

The modified Drude formula has the form

$$\sigma=e^2n_e\tau\vartheta/m, \tag{19}$$

where n_e is the density of the valence electrons, $\tau=l/v_T$ is the relaxation time, $l\approx R_s$ is the minimum mean free path for scattering off the impurities, $v_T=\sqrt{8T/\pi m}$ is the mean thermal velocity, and ϑ is the localization factor.

Partial localization of the electrons is connected with the mobility gap defined by the two parameters Δ_k (the absolute and relative gaps),

$$\Delta_k=I-\frac{e^2}{\varepsilon_\infty}\left(\frac{4\pi n}{3\zeta_k}\right)^{1/3}\quad (k=1,2), \tag{20}$$

where $\zeta_1\approx 1/3$ and $\zeta_2\approx 2/3$. The first and second gaps define the excitation energies, below which the mobility is respectively equal to zero and less than the minimum gas-kinetic value.

On the nonmetal side of the transition the localization factor is an exponentially decreasing function of the first mobility gap:

$$\vartheta=\frac{T(F_1-F_2)}{\Delta_2-\Delta_1}\approx\frac{2}{\sqrt{\pi}}\frac{\sqrt{\Delta_1 T}}{\Delta_2-\Delta_1}\exp\left(-\frac{\Delta_1}{T}\right), \tag{21}$$

$$\Delta_2>\Delta_1\gg T,$$

where the function F_k is a combination of incomplete gamma functions $\Gamma(m,x)$:

$$F_k=\frac{2}{\sqrt{\pi}}\left[\Gamma\left(\frac{5}{2},\frac{\Delta_k}{T}\right)-\frac{\Delta_k}{T}\Gamma\left(\frac{3}{2},\frac{\Delta_k}{T}\right)\right].$$

On the metal side of the transition

$$\vartheta=\frac{3T/2-\Delta_1-TF_2}{\Delta_2-\Delta_1}\approx\frac{3T/2-\Delta_1}{\Delta_2-\Delta_1}, \tag{22}$$

$$\Delta_1\leqslant 0,\quad \Delta_2\gg T.$$

According to Eq. (22), in the metallic state the localization factor tends to unity for the width of the gap $\Delta_2\sim T$.

Growth of the localization factor on the metal side of the transition is linked with growth of the Fermi energy and with degeneracy of the electrons. The Fermi energy of the mixed impurity states is equal to

$$\varepsilon_F'=\hbar^2k_F^2\vartheta^2/2m,\quad k_F=(6\pi^2n_i/g_a)^{1/3}, \tag{23}$$

where $k_F=mv_F/\hbar$ is the Fermi wave vector of the delocalized valence electrons, v_F is the Fermi velocity, g_a is the statistical weight of the impurity level, and the prime denotes deviation from an electron gas. For degenerate electrons the relaxation time in the Drude formula is

$$\tau=l/v_F',\quad v_F'=v_F\vartheta. \tag{24}$$

Substituting formulas (24) into formula (19) and cancelling out the localization factor which now appears both in the numerator and the denominator, we obtain the minimum metallic conductivity

$$\sigma=\frac{e^2 n_e l}{m v_F}. \tag{25}$$

Thus, for the degenerate electrons the conductivity no longer has a percolation character. Consequently, the mean free path for impurity scattering becomes larger than the distance between the impurity atoms. The minimum mean free path in the Ioffe–Regel' sense is

$$l \sim \hbar/\Delta p, \tag{26}$$

where Δp is the quantum indeterminacy of the momentum. At high enough temperatures the mean thermal momentum mv_T may serve as a measure of the uncertainty in the momentum, and the thermal wavelength may serve as a measure of the minimum mean free path. We thus obtain from formula (19)

$$\sigma \approx \frac{e^2 n_e}{k_F m v_T} \approx \frac{e^2 n_e R_s}{m v_T}. \tag{27}$$

Thus, to within a numerical factor of order unity, the conductivity would be the same in the Boltzmann case with mean free path R_s. To this accuracy, formulas (19)–(22) for the percolation conductivity can be directly extrapolated to the electron degeneracy region on the metal side of the metal–insulator transition.

However, in the limiting case of strong degeneracy (at sufficiently low temperatures or high concentrations) the uncertainty in the momentum can be expressed only in terms of the Fermi momentum, i.e., $\Delta p=\gamma m v_F$, where γ is a coefficient. According to relation (26), the mean free path for strong degeneracy is $l \sim 1/\gamma k_F$. Schematically, we will assume that the transition between cases of moderate and strong degeneracy takes place at the point where $v_T=\gamma v_F$ holds, so that Δp varies continuously. Thus, using the condition of matching with the extrapolated Boltzmann formulas, we obtain

$$l=R_s/\gamma. \tag{28}$$

Thus, the parameter γ may be defined as the ratio of the inverse mean free path in the limit of strong degeneracy to the inverse of the distance between impurities. When using this parameter which defines the conductivity significantly above the metal-nonmetal transition point, there is no need to explicitly allow for scattering of electrons by the ammonia molecules.

The percolation conductivity of sodium–ammonia solutions with different concentrations is compared with the experimental data in Fig. 4, which shows qualitative agreement between theory and experiment.

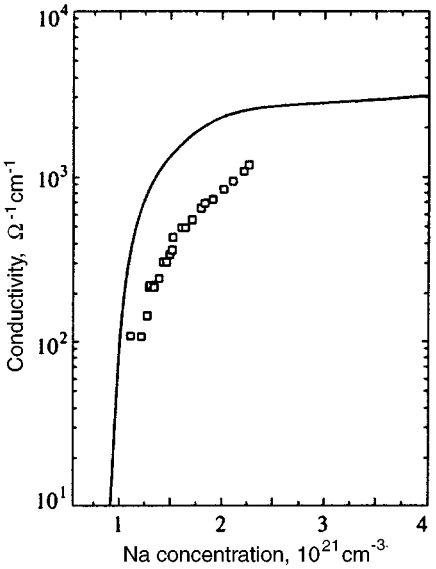


FIG. 4. Dependence of the conductivity of an ammonia solution of sodium on the sodium concentration at the critical solubility temperature $T_c=230$ K. The points correspond to experiment,[10] the curve plots the values calculated in the present work. The metal–insulator transition point lies not far below 10^{21} cm^{-3}. The binding energy of the impurity states $I=1.665$ eV. The normalization parameter for the conductivity at high concentrations $\gamma=0.25$.

5. CONCLUSION

Solutions of metals in ammonia are a unique example of electron solutions containing solvated electrons. Such solutions contain the metal atoms in impurity states whose role grows as the concentration is increased. Overlap of the classically accessible spheres of the impurity and solvated states is responsible the metal–insulator transition. Despite the qualitative differences between metal–ammonia solutions and doped semiconductors or expanded metals, the metal–insulator transitions in them have an identical percolation nature and are described by a universal theory.

The metal–insulator transition in metal–ammonia solutions has an effect on the phase transition analogous to the liquid–gas phase transition in pure metals. The parameters of the critical solubility point are determined by the Coulomb interaction of the ions and electrons in the overlapping impurity states, and the effective dielectric constant, which weakens this interaction, is substantially less than its macroscopic values due to saturation of the polarization of the dipolar molecules near the ions. An analogy with plasma critical points was used to estimate the binding energy of the impurity states. This estimate is confirmed by solving the Schrödinger equation for a simple model of the impurity states. The percolation conductivity via the impurity states near the metal–insulator transition point also agrees with experiment. Taken together, the self-consistency of the various estimates corroborates the given interpretation of the electronic structure of metal–ammonia solutions in the vicinity of the metal–insulator transition point.

[1] J. C. Thompson, *Electrons in Liquid Ammonia* (Clarendon Press, Oxford, 1976).

[2] M. F. Deĭgen, Zh. Éksp. Teor. Fiz. **26**, 293 (1954).

[3] B. I. Shklovskiĭ and A. L. Efros, *Electronic Properties of Doped Semiconductors* (Springer-Verlag, New York, 1984).

[4] K. S. Pitzer, J. Am. Chem. Soc. **80**, 5046 (1958).

[5] N. F. Mott, *Metal–Insulator Transitions* (Taylor & Francis, London, 1991).

[6] A. A. Likal'ter, Usp. Fiz. Nauk **162**, 119 (1992) [Sov. Phys. Usp. **35**, 591 (1992)].
[7] A. A. Likal'ter, Phys. Rev. B **53**, 4386 (1996).
[8] J. Jortner, J. Chem. Phys. Phys. **30**, 839 (1959).
[9] A. Likal'ter and H. Schneidenbach, JETP Lett. **63**, (1996).
[10] F. Noll, W.-C. Pilgrim, and R. Winter, Z. Phys. Chem. **156**, 303 (1988).
[11] N. F. Ashcroft and J. Lekner, Phys. Rev. **145**, 83 (1966).
[12] A. I. Baz', Ya. B. Zel'dovich, and A. M. Perelomov, *Scattering, Reactions and Decay in Nonrelativistic Quantum Mechanics* (Israel Program for Scientific Translations, Jerusalem, 1966).
[13] A. A. Likal'ter, Zh. Éksp. Teor. Fiz. **107**, 1996 (1995) [JETP **80**, 1105 (1995)].

Translated by Paul F. Schippnick

Surface ordering near the smectic-*A*–smectic-*C* transition in thin, free-standing, liquid-crystal films

V. K. Dolganov

Institute of Solid State Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

E. I. Demikhov

Institute of Physical Chemistry, University of Paderborn, 33095 Paderborn, Germany

R. Fouret and C. Gors

Laboratoire de Dynamique et Structures des Materiaux Moleculaires, Universite de Lille I, 59655 Villeneuve d'Ascq Cedex, France

(Submitted 7 August 1996)
Zh. Éksp. Teor. Fiz. **111**, 949–953 (March 1997)

Optical reflectivity measurements have been conducted near the smectic-*A*–smectic-*C* phase transition in free-standing films with thickness between two and eleven molecular layers. The temperature dependence of the reflectivity in thin film differs significantly from that in thick films. The optical thickness per layer increases in films with two to five layers as a result of cooling, in contrast with thick films. The average layer spacing was found to decrease with decreasing film thickness. © *1997 American Institute of Physics.* [S1063-7761(97)01503-5]

Phase transitions in thin, liquid-crystal, free-standing films have recently been studied extensively.[1–10] Free-standing films reveal rich variety of properties unknown for the bulk samples. The discrete layer-by-layer surface freezing has been observed near the Sm-*A*–Sm-*I*, Sm-*A*–Hex-*B*, Sm-*A*–Cry-*B* bulk transitions.[4,5,10,11] In contrast, ellipsometric measurements of the molecular tilt angle have demonstrated that the order parameter for the Sm-*A*–Sm-*C* transition is a continuous function of temperature in thin films.[12,13] The surface ordering phenomena near the Sm-*A*–Sm-*C* transition are very unusual. Boundary layers in free-standing films are tilted in the temperature interval of Sm-A phase.[12,13] This tilt causes an increase in the transition temperature on decreasing the number of layers.[12–15] The optical reflectivity measurements are an informative tool for studies of phase transitions in free-standing films. Until now, however, detailed optical reflectivity measurements have not been conducted in extremely thin films near the Sm-*A*–Sm-*C* transition.

In this paper we report the results of high-precision, optical-reflectivity measurements near Sm-*A*–Sm-*C* transition for free-standing films varying from two to eleven molecular layers. We have observed an anomalous temperature dependence of the optical thickness in ultrathin layers. We present the thickness dependence of the average interlayer spacing on the film thickness. The penetration length of the Sm-*C* surface ordering was found to be significantly larger than in the case of the layer-by-layer transitions.

The experiments were performed on free-standing films of *p*-decyloxybenzoic acid-*p*-*n*-hexyloxyphenyl ester. The bulk samples possess the following phase sequence: Sm-*C* (77 °C) Sm-*A* (83 °C) nematic (89 °C) isotropic. Using the polarizing microscope we did not observe discontinuities of the optical properties at the bulk transition temperature (T_{AC}). These observations indicate that the Sm-*A*–Sm-*C* transition is a second-order or weak first-order transition. Our experimental setup enabled simultaneous optical observations and measurements of the reflection and transmission intensities to be made. The films were illuminated with nearly normally incident light. The temperature was controlled with an accuracy of ±0.01 °C, X-ray-diffraction studies of the bulk samples were carried out using a curved linear position-sensitive multidetector and a curved quartz monochromator. The layer spacing was $d_A = 3.02$ nm at 82 °C and increased slightly in Sm-*A* temperature range with decreasing temperature (about 10^{-3} nm/°C). In the Sm-*C* phase d_C decreased significantly with decreasing temperature ($d_C = 2.92$ nm at 71 °C).

The 2 to 11-layer films were spread over a 6-mm-diam hole in a 0.2-mm-thick steel plate. Two methods of films preparation were used. Thin films could be spread by a movement of a mobile steel slip across the hole. As a rule, it took many attempts to obtain a film of the required thickness. The second method was based on the step-by-step thinning of free-standing films above the bulk Sm-*A*–isotropic[16] or Sm-*A*–nematic[17,18] phase transitions. The film thickness decreased in a stepwise manner on increasing the temperature. After preparing a film of a required thickness, the heating was stopped and the film cooled in the temperature interval of the bulky Sm-*A*. Combining these two methods, we prepared films ranging from 2 and 11 layers. The film thickness and interlayer spacings were determined by optical-reflectivity measurements using the equation[19]

$$I(\lambda)=\frac{(n^2-1)^2\,\sin^2(2\pi nL/\lambda)}{4n^2+(n^2-1)^2\,\sin^2(2\pi nL/\lambda)}, \tag{1}$$

where N is the number of layers, $n = 1.48$ is the refractive index, and d is the interlayer spacing. For thin films Eq. (1) can be simplified:[20]

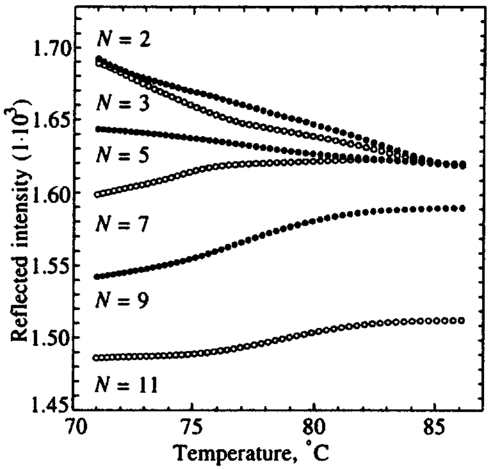


FIG. 1. Temperature dependence of the reflectivity in various free-standing films. The data from 3-, 5-, 7-, and 11-layer films were multiplied by $4/N^2$. The temperature was decreased at the rate 2 mK/s.

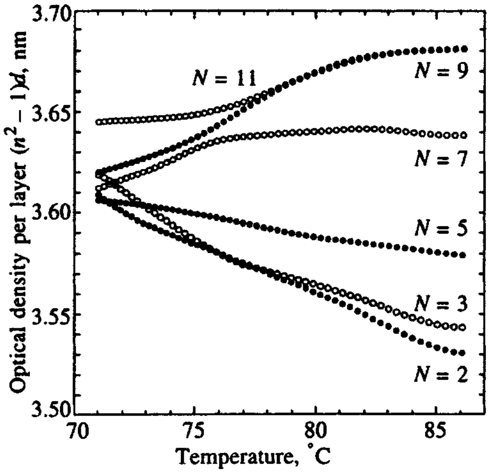


FIG. 2. Temperature dependence of the optical density per layer $(n^2 - 1)d$ for 2-, 3-, 5-, 7-, 9-, and 11-layer films.

$$I(\lambda)=\frac{N^2\pi^2 d^2(n^2-1)^2}{\lambda^2}. \tag{2}$$

Equation (2) allows the determination of the optical thickness per layer $(n^2 - 1)d$. In our case this equation gives the optical thickness with acceptable accuracy only for a very thin film. For thicker films, it was necessary to use the exact equation [Eq. (1)]. The thickness of very thick films, together with the refractive index, can be obtained from the fitting of the reflectivity spectra $I(\lambda)$ with Eq. (1).[14] However, this does not work properly for thin films $(N < 15)$, because the λ-dependence of the reflectivity is very smooth. In this case, we used the bulk value for the refraction index and reduced the number of fitting parameters. The numerical estimates show that this procedure allows us to find optical thickness without significant error (no more than 0.3%).

Figure 1 shows the temperature dependences of the reflectivities from 2- and 11-layer films ($\lambda = 550$ nm) multiplied by $4/N^2$ to compare results from films of different thicknesses. As expected for the second-order transition, we observed the continuous variation in the reflected intensities. Unexpected result is a drastic change in the temperature dependences with decreasing number of layers. The variation in the slope of the curves was found to occur for the number of layers ranging from 7 to 5.

Using the reflectivity data in Fig. 1, the temperature dependences of the optical thickness per layer $(n^2 - 1)d$ were calculated. These dependences are shown in Fig. 2. The change of the thickness is continuous, which corresponds to the second-order Sm-A–Sm-C transition. For the 7 to 11 layer films the optical thickness decreases on cooling. Qualitatively similar decrease in the optical thickness at the Sm-A–Sm-C transition was observed previously for very thick films.[14] The interlayer spacing was estimated using the optical thickness per layer and the refractive index. In thick films (9 to 11 layers) at 82 °C (Sm-A phase) d_A coincided with the value obtained from x-ray measurements with accuracy of 2%. Since the value of the optical reflectivity was measured with an uncertainty of about 1% and n with accuracy of 0.02, there was a good agreement between optical and x-ray data in the Sm-A phase. However, the relative decrease of the layer spacing $(d_A - d_C)/d_A \approx 0.03$. (x-ray measurements) in the range of temperatures from 82 to 71 °C was more than the decrease of the optical thickness in thick films. In our opinion, this difference stems from the change of the refractive index.

Data for the $N < 5$ layer films display the anomalous increase of the optical thickness and seem to indicate that the Sm-A–Sm-C transition is not present in thin films. Extremely thin films exhibit similar variation of $(n^2 - 1)d$ in the entire temperature range, in which these thicknesses are stable (up to 104 °C for a two-layer film). These results can be explained if we assume that the surface field stabilizes the Sm-C molecular ordering near the surface. The anomalous temperature dependence of the optical thickness in ultrathin films $(N < 5)$ can be interpreted in two ways: i) the increase of the refractive index as a result of cooling, ii) the change in the interlayer spacing. Our data are not sufficient to distinguish between these two cases. It should be noted that there cannot be a simple analogy with the behavior of thick films. The increased packing efficiency and the quench of the layer fluctuations in thin films,[21] as compared with thick films, can be the reason for the anomalous change of the optical thickness.

Recently, x-ray reflectivity studies of ultrathin Sm-C^* films on subtrates have shown an increase in the smectic layer spacing with decreasing number of smectic layers (about 20%).[22] Figure 3 shows the plot of the optical thickness per layer versus the number of layers. Our results show that at low temperature (71 °C, Sm-C phase) the change in the interlayer spacing is less than 1%. At high temperatures we observed significant differences in the optical interlayer spacings in thin and thick films (Fig. 3). The $(n^2 - 1)d$ changes are continuous within the measurement error. From Fig. 3 it is clear that $(n^2 - 1)d$ increases at different rates as the temperature changes. It should be noted that some surface ordering has been observed even below the bulk transition temperature (75 °C Fig. 3). The s-like shape of the thickness dependence (86 °C) provides strong evidence of

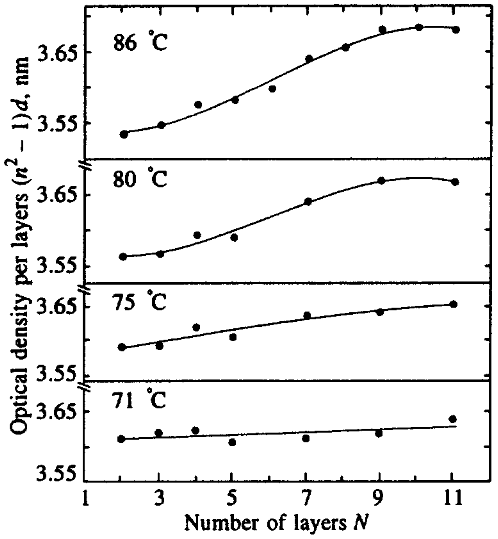


FIG. 3. Optical density per layer plotted as a function of the number of layers N.

the transition from the Sm-C surface to Sm-A ordering. In the Sm-C phase the thickness of the single smectic layer, d_C, may be assumed to be $d_C = d_A \cos\theta$, where θ is the tilt angle. By assuming that the value of the refractive index does not depend on the film thickness and that the thickness of the eleven-layer film at 86 °C is determined mainly by the Sm-A ordering, we could estimate θ at about 18° for a two-layer film. This value of θ is slightly lower than that found in the DOBAMBC two-layer film (approximately 25°, $T - T_{AC} \approx 10$ °C).[12] For the layer-by-layer ordering,[4,5,10,11] only the surface layers exhibit transition at a temperature ≈10 °C above the bulk transition. It is obvious that the penetration depth in our case is several times greater (Fig. 3). We estimate the surface order penetration depth to be about three layers. Another distinguishing feature of this transition is the behavior of the penetration depth near T_{AC}. For the layer-by-layer transition[4,5,10,11] the penetration depth diverges near the bulk transition temperature, at which there is an abrupt transition in the whole sample. In contrast, the penetration depth remains finite in the temperature range of the bulk Sm-C phase near the bulk transition temperature. These results can be attributed to the strong surface ordering and to the possibility of the continuous changes of the molecular tilt angle in the adjacent layers.

In summary, we have presented the first study of the Sm-A—Sm-C transition in extremely thin films using optical reflectivity measurements. The temperature dependence of the layer spacing in thin films differs significantly from that in thick films near the bulk Sm-A–Sm-C transition. Our observations indicate that surface field stabilizes the Sm-C phase in extremely thin films. In the high-temperature region, the layer spacing decreases with decreasing number of layers and remains approximately constant at lower temperatures. These observations indicate that the Sm-C–Sm-A transition takes place with increasing film thickness.

The research described in this publication was made possible, in part, by Grants INTAS 94-4078, the Russian Fund for Fundamental Research 95-02-05343, program ''Statistical physics and nonlinear dynamics'' and from the French National Education Minister.

[1] S. P. Pershan, *Structure of Liquid Crystal Phases*, World Scientific, Singapore (1988), p. 1.
[2] M. Cheng, J. T. Ho, S. W. Hui, and P. Pindak, Phys. Rev. Lett. **59**, 1112 (1987).
[3] E. B. Sirota, S. P. Pershan, L. B. Sorensen, and J. Collett, Phys. Rev. A **36**, 2890 (1987).
[4] B. D. Swanson, H. Stragier, D. J. Tweet, and L. B. Soresen, Phys. Rev. Lett. **62**, 909 (1989).
[5] T. Stoebe, R. Geer, C. C. Huang, and J. W. Goodby, Phys. Rev. Lett. **69**, 2090 (1992).
[6] T. Stoebe, C. C. Huang, and J. W. Goodby, Phys. Rev. Lett. **68**, 2944 (1992).
[7] P. Pieranski *et al.*, Physica A **194**, 364 (1993).
[8] T. Stoebe, C. C. Huang, and J. W. Goodby, Phys. Rev. E **48**, 404 (1993).
[9] C. Bahr D. Fliegner, C. J. Booth, and J. W. Goodby, Europhys. Lett. **26**, 539 (1994).
[10] A. J. Jin, T. Stoebe, and C. C. Huang, Phys. Rev. E **49**, R4791 (1994).
[11] V. K. Dolganov, R. Fouret, and C. Gors, JETP Lett. **63**, 285 (1996).
[12] S. Heinekamp, R. Pelkovits, E. Fontes *et al.*, Phys. Rev. Lett. **52**, 1017 (1984).
[13] C. Bahr and D. Fliegner, Phys. Rev. A **46**, 7657 (1992).
[14] I. Kraus, P. Pieranski, E. I. Demikhov, and H. Stegemeyer, Phys. Rev. E **48**, 1916 (1993).
[15] E. I. Demikhov, U. Hoffman, and H. Stegemeyer, J. Phys. France II **4**, 1865 (1994).
[16] T. Stoebe, P. Mach, and C. C. Huang, Phys. Rev. Lett. **73**, 1384 (1994).
[17] E. I. Demikhov, V. K. Dolganov, and K. P. Meletov, Phys. Rev. E **52**, R1285 (1995).
[18] V. K. Dolganov, E. I. Demikhov, R. Fouret, and C. Gors, Phys. Lett. A **220**, 242 (1996).
[19] M. Born and E. Wolf, *Principles of Optics*, Pergamon, New York (1980), p. 1.
[20] T. Stoebe and C. C. Huang, Phys. Rev. E **49**, 5238 (1994).
[21] R. Holyst, Phys. Rev. A **44**, 3692 (1991).
[22] M. Tarabia, G. Cohen, J. Gersten, and D. Davidov, Phys. Rev. E **51**, 799 (1995).

Published in English in the original Russian journal. Reproduced here with stylistic changes by the Translation Editor.

The magnetic field and energy of an Abrikosov vortex in an anisotropic London superconductor

I. M. Dubrovskiĭ

Institute of Metal Physics, Ukrainian National Academy of Sciences, 252142 Kiev, Ukraine

(Submitted 5 February 1996)
Zh. Éksp. Teor. Fiz. **111**, 954–963 (March 1997)

The behavior of a straight Abrikosov vortex in an anisotropic uniaxial London superconductor is studied. Analytical expressions are derived that approximately describe the magnetic field in three regions: the asymptotic region, where the distance r from the vortex line is greater than $\lambda\Gamma$ (λ is the London length and Γ is the anisotropy constant), the intermediate region $\lambda<r<\lambda\Gamma$, and the region $r<\lambda$. It is found that in the intermediate region with high anisotropy the component of the magnetic field along the vortex line changes sign for a certain interval of angles between the vortex line and the anisotropy axis. Because of this the interaction of parallel vortices whose plane is parallel to the anisotropy axis has a minimum and a maximum. This means that numerous metastable vortex lattices can exist. Additional terms in the vortex self-energy are obtained, and although they are smaller than the leading logarithmic term, they display a different dependence on the angle between the vortex line and the anisotropy axis. [S1063-7761(97)01603-X]

1. The magnetic field of a straight Abrikosov vortex in an anisotropic superconductor with large values of the Ginzburg–Landau parameter κ was studied in Refs. 1–4. An equation describing the magnetic field $\mathbf{h}(\mathbf{R})$ can be derived from the Ginzburg–Landau theory (see Ref. 3 and 5):

$$\mathbf{h}+\mathrm{curl}(\hat{\mu}\ \mathrm{curl}\ \mathbf{h})=\frac{\Phi_0}{2\pi\lambda^2}\Delta(\mathbf{R})\mathbf{v}. \tag{1}$$

In what follows we confine ourselves to the case of uniaxial anisotropy, when the effective mass along the axis is greater than in the perpendicular plane. Then, along the major axes the tensor $\hat{\mu}$ has the following components:

$$\mu_{11}=\mu_{22}=1,\quad \mu_{33}=\Gamma^2>1, \tag{2}$$

where Φ_0 is the quantum of magnetic flux (fluxoid), $\mathbf{v}$ is the unit vector along the Abrikosov-vortex line, $(\mathbf{h}(0)\cdot\mathbf{v}>0$, λ is the depth of penetration of the magnetic field parallel to the anisotropy axis, and $\mathbf{R}$ is the two-dimensional radius vector in the plane perpendicular to $\mathbf{v}$ and measured, in units of λ, from the point of intersection of the plane and the Abrikosov-vortex line. We denote the polar coordinates in this plane by R and φ, and the angle between the anisotropy axis and the Abrikosov-vortex line, measured from the axis to the direction of $\mathbf{v}$, by θ. As for the function $\Delta(\mathbf{R})$, it is known that for $R>R_0=\eta(\varphi,\theta)\kappa^{-1}$ it is negligible and that

$$\int_S \Delta(\mathbf{R})d\mathbf{R}=2\pi. \tag{3}$$

The values of the function $\eta(\varphi,\theta)$ differ little from unity, and S stands for the entire plane perpendicular to $\mathbf{v}$. In Refs. 1–4, $\Delta(\mathbf{R})$ was replaced by $2\pi\delta(\mathbf{R})$, where $\delta(\mathbf{R})$ is a delta function in a plane, i.e., instead of $\mathbf{h}(\mathbf{R})$ one actually finds $\mathbf{g}(\mathbf{R})$, the Green's function of Eq. (1). (More precisely, $\mathbf{g}(\mathbf{R})$ is one of the three vectors comprising the Green's tensor of Eq. (1).) The Fourier transform of $\mathbf{g}(\mathbf{R})$ was obtained in Refs. 2 and 6, while in Ref. 3 $\mathbf{g}(\mathbf{R})$ was studied via numerical integration and asymptotic expansion of the inverse Fourier transform. Here we derive analytical expressions for $\mathbf{g}(\mathbf{R})$ in the regions $R\ll 1$ and $1\ll R\ll\Gamma\sqrt{1-\sin^2\theta(1-\Gamma^{-2})}$, and refine the asymptotic expression for $\mathbf{g}(\mathbf{R})$ when $R\gg\Gamma$. In obtaining the results we often use formulas from Refs. 7 and 8. Actually,

$$\mathbf{h}(\mathbf{R}_1)=\frac{1}{2\pi}\int_S \mathbf{g}(\mathbf{R}_1-\mathbf{R}_2)\Delta(\mathbf{R}_2)d\mathbf{R}_2. \tag{4}$$

In what follows we show that for $R<1$ the averaging of (4) is important in the case of an anisotropic superconductor. Bearing Eq. (4) in mind, we also refine the expression for the Abrikosov-vortex energy derived in Ref. 6.

2. Let us decompose $\mathbf{g}(\mathbf{R})$ along the directions of the Cartesian axes of a system of coordinates with axis 3 directed along $\mathbf{v}$, axis 1 perpendicular to the anisotropy axis, and axis 2 directed in such a way that $0\leqslant\theta\leqslant\pi/2$. We assume that the components are functions of the polar coordinates R and φ in the plane perpendicular to $\mathbf{v}$. Then $\mathbf{g}(\mathbf{R})$ and $\mathbf{h}(\mathbf{R})$ can be represented by the following Fourier series:

$$g_1(\mathbf{R})=\sum_{n=1}^{\infty} g_{1n}(R)\sin 2n\varphi,$$

$$g_2(\mathbf{R})=\sum_{n=0}^{\infty} g_{2n}(R)\cos 2n\varphi, \tag{5}$$

$$g_3(\mathbf{R})=\sum_{n=0}^{\infty} g_{3n}(R)\cos 2n\varphi.$$

Next we show that the other Fourier amplitudes are zeros. Let us discuss in greater detail the calculation of $g_3(\mathbf{R})$. The integral formula for this function obtained in Refs. 2 and 6 can be derived by means of some identities. It has the following form:

$$g_2(\mathbf{R})=\frac{\Phi_0}{(2\pi\lambda)^2}\int_0^{2\pi}\int_0^{\infty}\exp[iq\ R\cos(\varphi-\psi)]$$
$$\times\Bigg[\left(\frac{\cos^2\gamma}{1-\sin^2\gamma\ \sin^2\psi}-\frac{\cos^2\theta}{1-\sin^2\theta\ \sin^2\psi}\right)$$
$$\times\frac{1}{q^2+\Gamma^{-2}(1-\sin^2\gamma\ \sin^2\psi)^{-1}}$$
$$+\frac{\cos^2\theta}{1-\sin^2\theta\ \sin^2\psi}\,\frac{1}{q^2+1}\Bigg]qdq\ d\psi, \qquad (6)$$

where q and ψ are the polar coordinates of the two-dimensional vector $\mathbf{q}$,

$$\sin^2\gamma=\sin^2\theta(1-\Gamma^{-2}),\quad 0\leqslant\gamma\leqslant\theta\leqslant\pi/2, \qquad (7)$$

with $\gamma=\theta$ only if $\theta=0$. Using formula (8.511.4) of Ref. 7,

$$\exp(iz\ \cos\ \alpha)=J_0(z)+2\sum_{n=1}^{\infty}i^nJ_n(z)\cos\ n\alpha, \qquad (8)$$

we arrive at the following expressions for the Fourier amplitudes:

$$g_{3n}=\frac{\Phi_0}{(2\pi\lambda)^2}(2-\delta_{n,0})\int_0^{2\pi}\cos\ 2n\psi\Bigg[\left(\frac{\cos^2\gamma}{1-\sin^2\gamma\ \sin^2\psi}\right.$$
$$\left.-\frac{\cos^2\theta}{1-\sin^2\theta\ \sin^2\psi}\right)G_n\left(\frac{R}{\Gamma(1-\sin^2\gamma\ \sin^2\psi)^{1/2}}\right)$$
$$+\frac{\cos^2\theta}{1-\sin^2\theta\ \sin^{\ 2}\psi}G_n(R)\Bigg]d\psi. \qquad (9)$$

Here

$$G_n(aR)=(-1)^n\int_0^{\infty}\frac{J_{2n}(qR)qdq}{q^2+a^2}$$
$$=(-1)^n\int_0^{\infty}\frac{J_{2n}(z)zdz}{z^2+a^2R^2}. \qquad (10)$$

These integrals are evaluated in the Appendix, with the result that

$$G_n(aR)=K_{2n}(aR)$$
$$-\frac{1}{2}\sum_{m=0}^{n-1}\frac{(-1)^m(2n-m-1)!2^{2n-2m}}{m!(aR)^{2n-2m}}, \qquad (11)$$
$$G_0(aR)=K_0(aR),$$

i.e., $G_n(aR)$ for $n\neq 0$ is equal to the corresponding modified Bessel function $K_{2n}(aR)$ minus the terms in the series representation of this function (see formula (8.446) in Ref. 7) that diverge at $R=0$. Equations (6)–(9) suggest that since the factor of $\cos n\psi(\sin n\psi)$ in the integrand is a function only of $\sin^2\psi$, only the terms in the Fourier series proportional to $\cos 2n\varphi$ are retained. Since

$$\sin^{2m}\psi=\sum_{k=0}^{m}S_k\ \cos[2(m-k)\psi] \qquad (12)$$

(see formula (1.320.1) of Ref. 7) and $\sin^2\psi$ enters into the integrand only with a factor $\sin^2\theta$ or $\sin^2\gamma$, the expansion of the amplitude g_{3n} in a series in $\sin^2\theta$ begins with the power $2n$. This ensures good convergence of the Fourier series.

Let us study the asymptotic behavior of $g_{3n}(R)$ in the limit $R\gg\Gamma$ $(n\neq 0)$. With the help of the identity

$$\Gamma^2(1-\sin^2\gamma\ \sin^2\psi)\equiv(\Gamma^2-1)(1-\sin^2\theta\ \sin^2\psi)+1, \qquad (13)$$

we can easily show that all the terms of the asymptotic expansion of the integrand in (9) originating from the sum in (11) are proportional to $\sin^2\psi$ with a power of at most $n-1$, with the result that they vanish when integrated with respect to ψ. The leading term of the asymptotic expansion of the modified Bessel function does not depend on the order of the function. Then Eqs. (5), (9), and (11) and the well-known definition of the coefficients of a Fourier series lead to the following result:

$$g_3(\mathbf{R})=\frac{\Phi_0}{2\pi\lambda^2}\sqrt{\frac{\pi}{2R}}\Bigg\{\left(\frac{\cos^2\gamma}{1-\sin^2\gamma\ \sin^2\varphi}\right.$$
$$\left.-\frac{\cos^2\theta}{1-\sin^2\theta\ \sin^2\varphi}\right)\Gamma^{1/2}(1-\sin^2\gamma\ \sin^2\varphi)^{1/4}$$
$$\times\exp\left[-\frac{R}{\Gamma}(1-\sin^2\gamma\ \sin^2\varphi)^{-1/2}\right]$$
$$+\frac{\cos^2\theta}{1-\sin^2\theta\ \sin^2\varphi}\exp(-R)\Bigg\}. \qquad (14)$$

Let us now compare (14) with the corresponding Eqs. (15)–(19) of Ref. 3. The dependence of $g_3(\mathbf{R})$ on R in both cases is the same: an exponential decreases as R grows, with a pre-exponential factor proportional to $R^{-1/2}$. The $\varphi-$dependent factor in the exponent (the expression in square brackets) of the first term is also the same, to first order in $\sin^2\gamma$. The factor of the second exponential function is also partly the same, but in the present paper it proves to be universal for all angles φ and θ, while in Ref. 3 it differs significantly in some sectors, even exhibiting a different dependence on R. The most important is the difference in the factor of the first exponential function. In particular, in Eq. (14) it is always nonnegative, which is understandable if we allow for (7). Generally, Eq. (14) suggests that $g_3(\mathbf{R})$ for $R\gg\Gamma$ is always positive, i.e., the inversion effect discovered by Grishin *et al.*[3] does not manifest itself in the asymptotic region. In Ref. 3 the existence of inversion in this region rests on the assumption that the second term of the asymptotic expansion is greater in absolute value than the first (see Eqs. (15) and (20) in Ref. 3). From (11) it follows that the function $G_n(aR)$ may have zeros originating from the sum, but, as we have just seen, for $R\gg\Gamma$ the contribution of the sum from (11) to $g_{3n}(R)$ vanishes and only the asymptotic behavior of the modified Bessel functions, which are always positive, is important here.

Inversion in $g_3(\mathbf{R})$ was also discovered by Grishin *et al.*[3] in their numerical calculations in the intermediate region of values of R. From (9) it follows that the effect can occur when the first and second terms in the integrand have different signs. Let us study the region

$$1 \ll R \ll \Gamma \cos\gamma. \tag{15}$$

Then, using (11), we can write the following approximate relationships:

$$G_n\left(\frac{R}{\Gamma\sqrt{1-\sin^2\gamma\,\sin^2\psi}}\right) \approx \frac{(-1)^n}{2n},$$

$$G_n(R) \approx (-1)^n \frac{2n}{R^2},$$

$$G_0\left(\frac{R}{\Gamma\sqrt{1-\sin^2\gamma\,\sin^2\psi}}\right) \approx -\ln\frac{R}{2\Gamma} - C + \frac{1}{2}\ln(1-\sin^2\gamma\,\sin^2\psi), \tag{16}$$

$$G_0(R) \approx \sqrt{\frac{\pi}{2R}}\,\exp(-R),$$

where C is Euler's constant. Plugging (16) into (9), evaluating the integrals via formulas (3.615.1) and (4.399.3) of Ref. 7, and summing the series via formulas (5.4.9.6) and (5.4.9.13) of Ref. 8, we find that

$$\begin{aligned} g_3(\mathbf{R}) = \frac{\Phi_0}{2\pi\lambda^2}\Bigg\{ &(\cos\gamma - \cos\theta)\left(-\ln\frac{R}{2\Gamma} - C\right) \\ &+ \cos\theta\sqrt{\frac{\pi}{2R}}\exp(-R) - \cos\gamma \\ &\times \ln\left[\frac{1+\cos\gamma}{2\cos\gamma}\left(1 - 2\tan^2\frac{\gamma}{2}\cos 2\varphi + \tan^4\frac{\gamma}{2}\right)^{1/2}\right] \\ &+ \cos\theta\,\ln\left[\frac{1+\cos\theta}{\cos\gamma + \cos\theta}\right. \\ &\times \left.\left(1 - 2\tan^2\frac{\theta}{2}\cos 2\varphi + \tan^4\frac{\theta}{2}\right)^{1/2}\right] + \frac{4\cos\theta}{R^2} \\ &\times \tan^2\frac{\theta}{2}\,\frac{(\tan^4(\theta/2)+1)\cos 2\varphi - 2\tan^2(\theta/2)}{(1 - 2\tan^2(\theta/2)\cos 2\varphi + \tan^4(\theta/2))^2}\Bigg\}. \end{aligned} \tag{17}$$

The expression (17) in the sectors where $\cos 2\varphi$ is close to -1 can be negative. For instance, for the values of the parameters adopted in Ref. 3 in numerical calculations, $\Gamma = 8$ and $\theta = \pi/6$ at the point with coordinates $R = 5.2$ and $\varphi = \pi/2$, which corresponds to the negative minimum of $g_3(\mathbf{R})$ in Eq. (17), we have $g_3 = -3.3 \times 10^{-3}\Phi_0/2\pi\lambda^2$. (In the notation adopted in the present paper, at $\Gamma = 8$ the value of λ is half the value of λ of Ref. 3.) The departure from the value depicted in Fig. 1c of Ref. 3 amounts to about 10%. However, in contrast to the results of Ref. 3, in the present work the regions where $g_3 < 0$ are limited.

For the values of R we are considering here the field $\mathbf{h}(\mathbf{R})$ differs little from $\mathbf{g}(\mathbf{R})$. Indeed, the characteristic distances over which $g_{3n}(R)$ changes considerably are of order unity (λ in the dimensional units of length). Then the terms in the Fourier series (5) change considerably over a distance of $\min(1, R/n)$. If this distance is much larger than κ^{-1}, integration in (4) does not change the corresponding term. Thus, over the given distance R integration changes only the terms in the Fourier series with $n \geq R\kappa$, and, as shown earlier, the smallness of these terms is of order $\sin^{2n}\theta$.

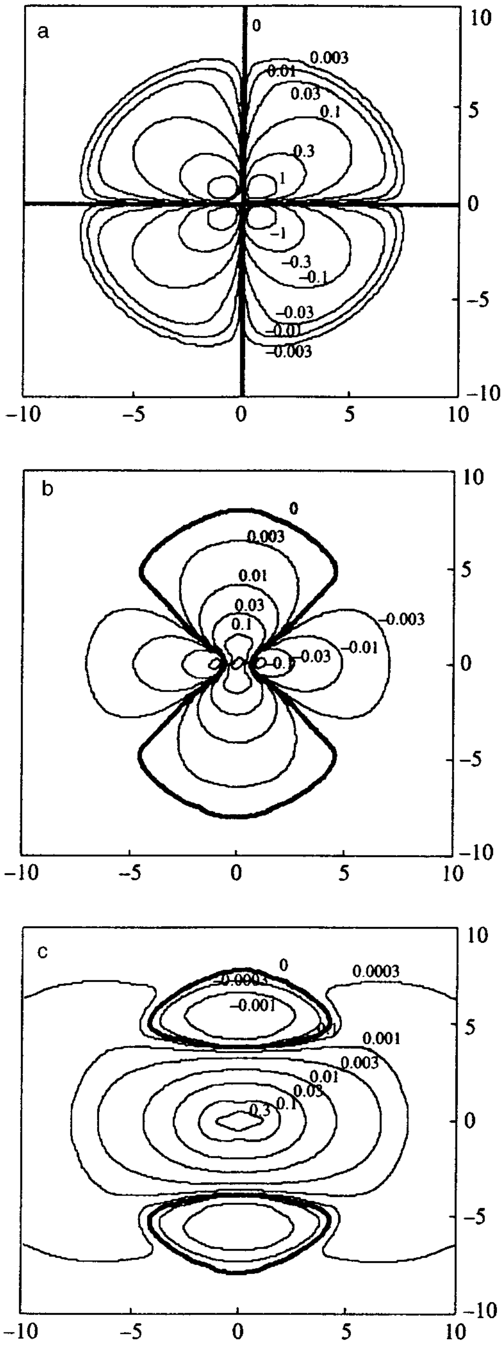


FIG. 1. Equal-value curves for the components of the function $\mathbf{g}_s(\mathbf{R})$, normalized to $\Phi_0/\pi\lambda^2$, in the plane perpendicular to the axis of an Abrikosov vortex: a—$g_{s1}(\mathbf{R})$, b—$g_{s2}(\mathbf{R})$, and c—$g_{s3}(\mathbf{R})$. The heavy curves separate the regions of positive and negative values.

For $R < 1$ we have

$$G_n\left(\frac{R}{\Gamma\sqrt{1-\sin^2\gamma\,\sin^2\psi}}\right)\approx G_n(R)\approx\frac{(-1)^n}{2n},$$

$$G_0\left(\frac{R}{\Gamma\sqrt{1-\sin^2\gamma\,\sin^2\psi}}\right)\approx -\ln\frac{R}{2\Gamma}+\frac{1}{2}\times\ln(1-\sin^2\gamma\,\sin^2\psi)-C, \qquad (18)$$

$$G_0(R)=-\ln\frac{R}{2}-C,$$

which yield

$$g_3(\mathbf{R})=\frac{\Phi_0}{2\pi\lambda^2}\left[(\cos\gamma-\cos\theta)\ln\Gamma+\cos\gamma\left(-\ln\frac{R}{2}-C\right)-\cos\gamma\,\ln\frac{1+\cos\gamma}{2\cos\gamma}+\cos\theta\,\ln\frac{1+\cos\theta}{\cos\gamma+\cos\theta}-\frac{1}{2}\cos\gamma\,\ln\left(1-2\tan^2\frac{\gamma}{2}\cos 2\varphi+\tan^4\frac{\gamma}{2}\right)\right]. \qquad (19)$$

Not only does this expression diverge, it becomes strongly dependent on φ as $R\to 0$. This singularity, which is natural for a Green′s function, vanishes in $\mathbf{h}(\mathbf{R})$ if formula (4) is employed. In (19) we put

$$R=|\mathbf{R}_1-\mathbf{R}_2|,\qquad \cos 2\varphi=\cos\left[2\arctan\left(\frac{R_1\sin\varphi_1-R_2\sin\varphi_2}{R_1\cos\varphi_1-R_2\cos\varphi_2}\right)\right]. \qquad (20)$$

Note that $\Delta(\mathbf{R}_2)R_2$ has a sharp peak at $R_2=R_0$ with a width of order κ^{-1}. Then integration in (4) with respect to R_2 can be done in general form by taking the slowly varying factors of $\Delta(\mathbf{R}_2)R_2$ outside the integral sign at $R_2=R_0$, ignoring the dependence of R_0 on φ to the same accuracy, and employing Eq. (3). As a result of integration with respect to φ_2 we find that the axisymmetric term in $h_3(\mathbf{R}_1)=h_3(R_1)+\tilde{h}_3(R_1,\varphi_1)$ for $R_1\geqslant\overline{\eta\kappa}^{-1}$ coincides with a similar term in (19), while for $R_1<\overline{\eta\kappa}^{-1}$ it remains constant (the factor $\overline{\eta}$ differs little from unity). The φ_1-dependent term in $h_3(\mathbf{R}_1)$ has the form

$$\tilde{h}_3(R_1,\varphi_1)=\frac{\Phi_0}{2(2\pi\lambda)^2}\cos\gamma\int_0^{2\pi}\ln\left\{1+\tan^4\frac{\theta}{2}-2\tan^2\frac{\theta}{2}\times\cos\left[2\arctan\left(\frac{R_1\sin\varphi_1-\overline{\eta\kappa}^{-1}\sin\varphi_2}{R_1\cos\varphi_1-\overline{\eta\kappa}^{-1}\cos\varphi_2}\right)\right]\right\}d\varphi_2. \qquad (21)$$

At $R_1=\overline{\eta\kappa}^{-1}$,

$$\cos\left[2\arctan\left(\frac{\sin\varphi_1-\sin\varphi_2}{\cos\varphi_1-\cos\varphi_2}\right)\right]=-\cos(\varphi_1+\varphi_2). \qquad (22)$$

Then, according to Eq. (2.6.36.9) of Ref. 8, we have $h_3(\overline{\eta\kappa}^{-1},\varphi_1)\equiv 0$. Reasoning along similar lines we can show that $h_3(0,\varphi_1)\equiv 0$.

Thus, Eqs. (14), (17), (19), and (21) describe the component of the magnetic field parallel to the Abrikosov-vortex line in different ranges of values of R. Here the magnetic field changes sign only in the intermediate region.

Similarly, we can write expressions for the other components of the Green's function. In the asymptotic region $R\gg\Gamma$,

$$g_1(\mathbf{R})=\frac{\Phi_0}{2\pi\lambda^2}\,\frac{\sin 2\theta\,\sin 2\varphi}{4}(\Gamma^2-1)\sqrt{\frac{\pi}{2R}}\times\left\{\left[\frac{\Gamma^{1/2}(1-\sin^2\gamma\,\sin^2\varphi)^{1/4}}{(\Gamma^2-1)(1-\sin^2\theta\,\sin^2\varphi)}-\frac{1}{\Gamma^{3/2}(1-\sin^2\gamma\,\sin^2\varphi)^{3/4}}\right]\times\exp\left[-\frac{R}{\Gamma(1-\sin^2\gamma\,\sin^2\varphi)^{1/2}}\right]-\frac{1}{(\Gamma^2-1)(1-\sin^2\theta\,\sin^2\varphi)}\exp(-R)\right\}, \qquad (23)$$

$$g_2(\mathbf{R})=-\frac{2\cos^2\varphi}{\sin 2\varphi}g_1(\mathbf{R}). \qquad (24)$$

In the intermediate region (15),

$$g_1(\mathbf{R})=\frac{\Phi_0}{2\pi\lambda^2}\cot\theta\left\{\frac{4}{R^2}\,\frac{\cos^{-4}(\theta/2)\cos\theta\,\sin 2\varphi}{[1-2\tan^2(\theta/2)\cos 2\varphi+\tan^4(\theta/2)]^2}+\arctan\left[\frac{\tan^2(\gamma/2)\sin 2\varphi}{1-\tan^2(\gamma/2)\cos 2\varphi}\right]-\arctan\left[\frac{\tan^2(\theta/2)\sin 2\varphi}{1-\tan^2(\theta/2)\cos 2\varphi}\right]\right\}, \qquad (25)$$

$$g_2(\mathbf{R})=\frac{\Phi_0}{2\pi\lambda^2}\,\frac{\sin 2\theta}{4}\left\{\frac{1}{\cos^2(\theta/2)}\sqrt{\frac{\pi}{2R}}\exp(-R)+\left[\frac{\Gamma^2-1}{\Gamma^2\cos^2(\gamma/2)}-\frac{1}{\cos^2(\theta/2)}\right]\left(\ln\frac{2\Gamma}{R}-C\right)-\frac{2\cos\theta}{R^2\cos^4(\theta/2)}\,\frac{[1+\tan^4(\theta/2)]\cos 2\varphi-2\tan^2(\theta/2)}{[1-2\tan^2(\theta/2)\cos 2\varphi+\tan^4(\theta/2)]^2}+\frac{2\cos\gamma}{\sin^2\theta}\ln\left[\frac{1+\cos\gamma}{2\cos\gamma}\left(1-2\tan^2\frac{\gamma}{2}\cos 2\varphi+\tan^4\frac{\gamma}{2}\right)^{1/2}\right]-\frac{2\cos\theta}{\sin^2\theta}\ln\left[\frac{1+\cos\theta}{\cos\gamma+\cos\theta}\times\left(1-2\tan^2\frac{\theta}{2}\cos 2\varphi+\tan^4\frac{\theta}{2}\right)^{1/2}\right]\right\}. \qquad (26)$$

It may seem that these expressions are incorrect since they do not vanish when we let Γ go to 1. In this case, however, the intermediate region determined by the inequalities (15) vanishes.

For $R\ll 1$,

$$g_1(\mathbf{R})=\frac{\Phi_0}{2\pi\lambda^2}\cot\,\theta\,\arctan\left[\frac{\tan^2(\gamma/2)\sin\,2\varphi}{1-\tan^2(\gamma/2)\cos\,2\varphi}\right]. \tag{27}$$

The dependence of $g_1(\mathbf{R})$ on R appears when the higher-order approximations in R are taken into account. The function $g_2(\mathbf{R})$, as well as $g_3(\mathbf{R})$, has a logarithmic singularity at $R=0$:

$$\begin{aligned}g_2(\mathbf{R})=&\frac{\Phi_0}{2\pi\lambda^2}\frac{(\Gamma^2-1)\sin\,2\theta}{2}\left\{\frac{\ln(2/R)-C}{2\Gamma^2\cos^2(\gamma/2)}\right.\\&+\frac{\ln\,\Gamma}{2}\left[\frac{1}{\Gamma^2\cos^2(\gamma/2)}-\frac{1}{(\Gamma^2-1)\cos^2(\theta/2)}\right]\\&+\frac{1}{(\Gamma^2-1)\sin^2\theta}\left[\cos\,\gamma\,\ln\frac{1+\cos\,\gamma}{2\cos\,\gamma}\right.\\&\left.-\cos\,\theta\,\ln\frac{1+\cos\,\theta}{\cos\,\gamma+\cos\,\theta}\right]+\frac{\cos\,\gamma}{2(\Gamma^2-1)\sin^2\theta}\\&\left.\times\ln\left[1-2\tan^2\left(\frac{\gamma}{2}\right)\cos\,2\varphi+\tan^4\frac{\gamma}{2}\right]\right\}.\end{aligned} \tag{28}$$

This means that even at the center of the vortex neither $\mathbf{g}(\mathbf{R})$ nor $\mathbf{h}(\mathbf{R})$ is parallel to $\mathbf{v}$.

Figure 1 depicts the equal-value curves for the three components of the function $\mathbf{g}_s(\mathbf{R})$ formed by matching the above asymptotic expressions for $\mathbf{g}(\mathbf{R})$ at parameter values $\Gamma=8$ and $\theta=\pi/6$. The matching was done by including one expression and excluding the other via coefficients varying from zero to unity in the intermediate regions $0.5\leqslant R\leqslant 1$ and $6\leqslant R\leqslant 8$. Hence Fig. 1 provides a qualitative representation of the function $\mathbf{g}(\mathbf{R})$, while in the regions of applicability of the asymptotic formulas this representation is even quantitative. Figure 1 can be compared with the corresponding figure in Ref. 3, which was obtained by numerically integrating the inverse Fourier transform for the same parameter values. Note that the unit of length in the present paper is half that in Ref. 3. In both papers the pattern of $g_1(\mathbf{R})$ (Fig. 1a) is the same. The function $g_2(\mathbf{R})$ in Fig. 1b corresponds even quantitatively to the g_2 of Ref. 3 in the region represented in that paper. The position of the negative minimum on the horizontal axis cannot be obtained exactly from the asymptotic formulas for the chosen values of the parameter. However, this region does not provide a full description of the function $g_2(\mathbf{R})$. From Eqs. (23) and (24) and from the corresponding asymptotic formula of Ref. 3 it follows that $g_2(\mathbf{R})$ is nonpositive for $R\gg\Gamma$ and vanishes on the vertical axis, since it is proportional to $\cos^2\varphi$. Hence the zero-level curve becomes closed on the vertical axis and then continues along this axis. Figure 1c shows that the region of negative values of $g_3(\mathbf{R})$ is limited, in contrast to the behavior depicted in the corresponding figure in Ref. 3.

3. The leading contribution to the energy of a singular vortex is provided by the magnetic field and the kinetic energy of the superconducting currents[3]

$$E=\frac{\lambda^2}{8\pi}\int_S[|\mathbf{h}|^2+(\mathrm{curl}\ \mathbf{h})\cdot\hat{\mu}\cdot(\mathrm{curl}\ \mathbf{h})]d\mathbf{R}. \tag{29}$$

Here E is the energy density per unit length of the Abrikosov vortex. Integrating the second term by parts and plugging (1) and (4), we obtain

$$E=\frac{\Phi_0}{32\pi^3}\int_S\int_S\Delta(\mathbf{R}_1)g_3(\mathbf{R}_1-\mathbf{R}_2)\Delta(\mathbf{R}_2)d\mathbf{R}_1d\mathbf{R}_2. \tag{30}$$

The cutoff procedure in $\mathbf{q}$-space carried out in Ref. 6 means replacing one of the functions $\Delta(\mathbf{R})$ by $2\pi\delta(\mathbf{R})$ and the other by

$$\begin{aligned}\Delta(R)&=\frac{1}{2\pi}\int_0^{\kappa}\int_0^{2\pi}\exp[iqR\,\cos(\varphi-\psi)]qdqd\psi\\&=\frac{\kappa J_1(\kappa R)}{R}.\end{aligned} \tag{31}$$

Instead we allow for the fact that, as noted earlier, $R\Delta(R)$ has a sharp peak at $R=R_0\approx\overline{\eta\kappa}^{-1}$. Then in $g_3(\mathbf{R}_1-\mathbf{R}_2)$ of Eq. (30) we can put $R_1=R_2=\overline{\eta\kappa}^{-1}$ and, using Eqs. (3), (19) and (20), we obtain

$$\begin{aligned}E&=\frac{\Phi_0}{32\pi^3}\int_0^{2\pi}\int_0^{2\pi}g_3(\overline{\eta\kappa}^{-1},\varphi_1;\overline{\eta\kappa}^{-1},\varphi_2)d\varphi_1d\varphi_2\\&=\left(\frac{\Phi_0}{4\pi\lambda}\right)^2\left[\cos\gamma\left(\ln\frac{2\kappa}{\bar{\eta}}-C\right)+(\cos\,\gamma-\cos\,\theta)\right.\\&\left.\times\ln\,\Gamma-\cos\,\gamma\,\ln\frac{1+\cos\,\gamma}{2\,\cos\,\gamma}+\cos\,\theta\,\ln\frac{1+\cos\,\theta}{\cos\,\gamma+\cos\,\theta}\right].\end{aligned} \tag{32}$$

Here the first term coincides with the well-known formula derived in Ref. 6. The employed method of integration makes it possible to estimate the relative accuracy of the calculations. For the first term this accuracy is $(\ln\,\kappa)^{-1}$, i.e., the term is determined to within logarithmic accuracy, as noted in Ref. 6. The other terms are determined to within a relative accuracy of κ^{-1} and are generally smaller than the first. They are characterized, however, by a different dependence on the angle between the Abrikosov vortex and the anisotropy axis: at $\theta=0$ they are zero, and their contribution monotonically increases with θ. For $\Gamma=8$, $\kappa=50$, and $\theta=\pi/2$ the correction amounts to 14% of the value of the first term.

The energy of interaction, per unit length, of two parallel Abrikosov vortices separated by a distance L so large that $h_3(\mathbf{L})\approx g_3(\mathbf{L})$ is

$$E_{\mathrm{int}}=\frac{\Phi_0}{4\pi}g_3(\mathbf{L}). \tag{33}$$

If the plane containing the Abrikosov vortices passes through the anisotropy axis, for parameter values allowing for the inversion of the sign of $g_3(\mathbf{L})$, in the intermediate region (15) the E_{int} vs L dependence acquires a negative minimum at $L=L_{\mathrm{min}}$ and a positive maximum at $L=L_{\mathrm{max}}$ (here $L_{\mathrm{min}}<L_{\mathrm{max}}\sim\Gamma$). In this case, at Abrikosov-vortex densities $N<L_{\mathrm{max}}^{-2}$ there can be two ideal vortex lattices in the system: a metastable Abrikosov lattice in which all the vortices are separated by distances greater than L_{max} and are repulsive, and a stable lattice (described in Ref. 3) consisting of repul-

sive planes that are parallel to the anisotropy axis and in which the Abrikosov vortices are separated by distances of about $L_{\min}$. There can also be metastable groups of nonideal lattices in which chains of different lengths consisting of Abrikosov vortices separated by distances of about $L_{\min}$ are surrounded on all sides by Abrikosov vortices separated by distances greater than $L_{\max}$. These metastable lattices can be one of the reasons for the hysteresis phenomena observed and discussed in Ref. 9 and the works cited therein.

APPENDIX

Using the recursion relation for the Bessel function of the first kind (see formula (8.471.1) of Ref. 7), we find that

$$J_{2n}(x)=(-1)^n\sum_{k=0}^{n}(-1)^k J_k(x) A_{nk}, \qquad \text{(A1)}$$

where the A_{nk} are positive integral coefficients. We have

$$2^k\int_0^\infty x^{-k+1}J_k(x)\frac{dx}{x^2+b^2} = \frac{2}{(k-1)!\,b^2} - 2^{k+1}\int_0^\infty x^{-(k-1)+1}J_{k-1}(x)\frac{dx}{(x^2+b^2)^2}. \qquad \text{(A2)}$$

Here integration by parts is done according to formula (5.52.2) of Ref. 7. Doing this k times and using formula (6.565.4) of Ref. 7, we get

$$2^k\int_0^\infty x^{-k+1}J_k(x)\frac{dx}{x^2+b^2} = \sum_{m=1}^{k}\frac{(-1)^{m+1}2^{2m-1}(m-1)!}{(k-m)!\,b^{2m}} + (-1)^k\frac{2^k}{b^k}K_k(b). \qquad \text{(A3)}$$

Next we plug (A1) into (10) and integrate according to (A3). The result is

$$G_n(aR)=\sum_{k=0}^{n}A_{nk}\left[(-1)^k\sum_{m=1}^{k}\frac{(-1)^{m+1}2^{2m-1}(m-1)!}{(k-m)!(aR)^{2m}} + \frac{2^k}{(aR)^k}K_k(aR)\right]. \qquad \text{(A4)}$$

Taking $(2/aR)^k$ outside the brackets, replacing the summation index m with $k-l$, and writing $K_k(aR)$ in the form of a series via formula (8.446) of Ref. 7, we arrive at a series expansion of $G_n(aR)$ in powers of the argument:

$$G_n(aR)=\sum_{m=0}^{\infty}\frac{1}{m!}\left(\frac{aR}{2}\right)^{2m}\sum_{k=0}^{n}\frac{(-1)^{k+1}}{(m+k)!}A_{nk} \times\left[\ln\frac{aR}{2}+C-\frac{1}{2}\left(\sum_{l=1}^{m}\frac{1}{l}+\sum_{l=1}^{m+k}\frac{1}{l}\right)\right]. \qquad \text{(A5)}$$

One can obtain the zeroth term in this expansion, $G_n(0)=(-1)^n/2n$, directly from Eq. (10) by using formula (6.561.14) of Ref. 7. For the modified Bessel function $K_{2n}(x)$ we can derive, by employing recurrence relations, a formula similar to (A1) with the same coefficients A_{nk} but without the factors $(-1)^{k+n}$. Using this formula, we can transform (A4) to

$$G_n(aR)=K_{2n}(aR)-\sum_{k=1}^{n}\sum_{l=1}^{k}A_{nk}(-1)^{k+l} \times\frac{2^{2l-1}(l-1)!}{(k-l)!(aR)^{2l}}. \qquad \text{(A6)}$$

Writing $K_{2n}(aR)$ in the form of a series via formula (8.446) of Ref. 7 and comparing the result with (A5), we conclude that the coefficients of the negative powers of aR vanish, which yields (11).

[1] V. G. Kogan, Phys. Rev. B **24**, 1572 (1981).

[2] A. M. Grishin, Fiz. Nizk. Temp. **9**, 277 (1983) [Sov. J. Low Temp. Phys. **9**, 138 (1983)].

[3] A. M. Grishin, A. Yu. Martynovich, and S. V. Yampol'skiĭ, Zh. Éksp. Teor. Fiz. **97**, 1930 (1990) [Sov. Phys. JETP **70**, 1089 (1990)].

[4] V. G. Kogan and J. R. Clem, Phys. Rev. B **24**, 2497 (1981).

[5] V. L. Ginzburg, Zh. Éksp. Teor. Fiz. **23**, 236 (1952).

[6] A. V. Balatskiĭ, L. I. Burlachkov, and L. P. Gor'kov, Zh. Éksp. Teor. Fiz. **90**, 1478 (1986) [Sov. Phys. JETP **63**, 866 (1986)].

[7] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Sums, Series and Products*, Academic Press, New York (1980).

[8] A. P. Prudnikov, Yu. A. Brychkov, and O. I. Marichev, *Integrals and Series*, Vol. 1: *Elementary Functions*, Gordon and Breach, New York (1983).

[9] L. G. Mamsurova, K. S. Pigal'skiĭ, V. P. Sakun, and L. G. Shcherbakova, Fiz. Tverd. Tela (St. Petersburg) **37**, 2954 (1995) [Phys. Solid State **37**, 1630 (1995)].

Translated by Eugene Yankovsky

“The energy spectrum of quadrupole impurity centers of different types in antiferromagnets

M. A. Ivanov

Institute of Metal Physics, Ukrainian National Academy of Sciences, 252142 Kiev, Ukraine

V. Ya. Mitrofanov and A. Ya. Fishman

Institute of Metallurgy, Ural Branch of the Russian Academy of Sciences, 620016 Ekaterinburg, Russia

(Submitted 27 April 1996)
Zh. Éksp. Teor. Fiz. **111**, 964–978 (March 1997)

We study the energy spectrum and some properties of various quadrupole centers (magnetic impurities or magnetic impurity complexes that are symmetric with respect to the magnetic sublattices of the antiferromagnet). We allow for the effect of spin–phonon coupling on the quadrupole splitting parameter and show that such coupling can lead to a considerable decrease in the value of this parameter and even change its sign. We investigate the behavior of quadrupole centers with an orbitally degenerate ground state and of quadrupole impurity complexes formed by mixed-valence ions. We demonstrate that such centers may greatly affect the resonant, magnetic, and thermodynamic properties of antiferromagnets. Finally, we analyze the existing experimental data and show that several new effects can be observed in systems with such centers (in particular, a magnetic analog of the Jahn–Teller effect and a strong magnetoelectric effect). © *© 1997 American Institute of Physics.* [S1063-7761(97)01703-4]

1. INTRODUCTION

Localized low-frequency excitations of an impurity spin positioned symmetrically with respect to the magnetic sublattices of an antiferromagnet have been studied by many researchers (see, e.g., Refs. 1–6). In the absence of a magnetic field the spectrum of such impurities is described by an effective Hamiltonian of quadrupole type with a constant A equal, in order of magnitude, to the ratio of the square of the constant J of the exchange impurity–matrix coupling to the width of the spin-wave band of the antiferromagnet, with $A>0$. As a result, the ground state of impurities with half-integer spin proves to be twofold generate: $S_z=\mp 1/2$. Here the magnetic, resonant, and thermodynamic properties of an antiferromagnet containing such centers exhibit many specific properties characteristic of crystals with two-level systems, including the presence of a peculiar glass phase.[5]

The present upsurge of interest in quadrupole centers is caused by the appearance of new objects of research in both traditional systems[7] and such specific quasi-two-dimensional magnetic materials as high-T_c superconducting oxides.[8–12] The reason for the appearance of quadrupole magnetic centers in high-T_c superconducting oxides is either nonstoichiometry of the composition or nonisovalent substitutional atoms. Therefore, it is natural that in this case the behavior of the impurity center strongly depends on the nature of localization of excess charge on the center. The same systems have been mentioned (see Ref. 8) in connection with the possibility of forming quadrupole centers with an orbitally degenerate ground state.

Following the description of quadrupole states belonging to impurity centers, a similar approach was adopted in describing the properties of some types of antiferromagnets, in which the average value of the exchange field generated by the atoms of one sublattice on the other is zero,[13] as well as in analyzing the special features of the electron spectrum of carriers in high-T_c superconducting oxides.[14]

In this paper we go into a deeper study of the energy spectrum of various types of quadrupole centers. Here for the traditional quadrupole centers we find the contribution of spin–phonon coupling to the quadrupole splitting parameter A and attempt to resolve the contradiction between the theory developed earlier and the new experimental data.[7] We focus on the analysis of the features of the low-frequency energy spectra and properties of unconventional quadrupole centers, such as centers in which the ground state is orbitally degenerate, and impurity complexes that occur due to charge transfer between ions of transition elements that are neighbors of an impurity atom (quadrupole centers with mixed-valence ions).

2. COMMON QUADRUPOLE CENTERS

2.1. Contribution of spin–phonon coupling to the quadrupole splitting parameter

An example of quadrupole centers is the orbitally nondegenerate magnetic impurity interstitial atoms positioned symmetrically with respect to the magnetic sublattices of an antiferromagnet. For such centers the Hamiltonian of the exchange interaction with the spin subsystem of the matrix can be written as

$$\mathscr{H}=\sum_{in}\sum_{s} J_{in,s}(\mathbf{S}_{in}\cdot\mathbf{S}_s),\quad \sum_{n} J_{1n,s}=\sum_{n} J_{2n,s}, \tag{1}$$

where $\mathbf{S}_s$ is the spin of an impurity atom at the interstitial site with number s, and the $J_{in,s}$ are the constant of exchange coupling between an impurity at the interstitial site s and the matrix spin $\mathbf{S}_{in}$ in the cell n of the sublattice i $(i=1,2)$.

When the exchange coupling of impurity and matrix is weak, the effective spin Hamiltonian $\mathscr{H}_{\rm eff}$ for noninteracting impurity centers can be found by perturbation techniques. The small parameter in the perturbation series is the ratio of the impurity–matrix exchange interaction to the maximum exchange interaction in the matrix, $J/\max(J_{ii'})$, where $J\equiv\Sigma_n J_{in,s}$ is the effective impurity–matrix exchange interaction, and $J_{ii'}\equiv\Sigma_n J_{in,i'n'}$ is the corresponding interaction of matrix atoms belonging to different sublattices. Then, to second order in this ratio, $\mathscr{H}_{\text{eff}}$ can be expressed as[2,3]

$$\mathscr{H}_{\text{eff}}=A(S_z^2-S(S+1)), \qquad (2)$$

where S is the impurity spin, and S_z is the projection of the spin on the antiferromagnetism axis. The quadrupole splitting parameter A in (2) is positive and satisfies

$$A\sim\frac{J^2}{\max(J_{ii'})}.$$

These statements concerning the magnitude and sign of the quadrupole splitting parameter may become invalid if there is a strong interaction between the magnetic moment of an impurity center and phonons. Lattice vibrations model the parameters of the impurity–matrix exchange interaction:

$$\mathscr{V}=\sum_{in}\sum_{\alpha s}\mathscr{K}^{\alpha}_{in,s}(\mathbf{S}_{in}\cdot\mathbf{S}_s)(\mathscr{U}^{\alpha}_{in}-\mathscr{U}^{\alpha}_{s}),$$
$$R^{\alpha}_{in,s}\equiv\frac{\partial J_{in,s}}{\partial(R^{\alpha}_{in}-R^{\alpha}_{s})},\quad \alpha=x,\ y,\ z, \qquad (3)$$

where R^{α}_{in}, R^{α}_{s} and $\mathscr{U}^{\alpha}_{in}$, $\mathscr{U}^{\alpha}_{s}$ are the coordinates of the equilibrium positions and the displacements of the matrix atoms and the impurity in the lattice. Allowing for this effect in second-theory perturbation theory in $\mathscr{V}$ (with respect to the phonon subsystem) adds the following term to the impurity-center effective Hamiltonian $\mathscr{H}_{\text{eff}}$ (Eq. (2));

$$\mathscr{H}^{(s-p)}=A^{(s-p)}S_z^2,\quad A^{(s-p)}\sim-\frac{S_0^2(\mathscr{K}^{\alpha}_{is+\delta,s})^2a^2}{\omega_D}<0, \qquad (4)$$

where S_0 is the spin of the matrix atoms, the label $s+\delta$ numbers the sites nearest to an impurity center s, a is the distance between the impurity and the nearest spin in the matrix, and ω_D is the Debye frequency.

We see that the contribution $A^{(s-p)}$ of the spin–phonon interaction (Eq. (4)) to the total quadrupole splitting constant $A_{\text{eff}}=A+A^{(s-p)}$ is negative. As a result, A_{eff} can become much smaller than A and even have different signs. The temperature dependence of the parameter $A^{(s-p)}$ at low temperatures can be approximately described if in (4) we replace S_0 with the temperature-dependent function $S_0(T)=S_0\sigma(T)/\sigma(0)$, which reflects the variation of the sublattice magnetization $\sigma(T)$ with temperature. Thus, the parameter $A^{(s-p)}$ is proportional to $\sigma^2(T)$. At the same, we can easily show that A exhibits no such temperature dependence, as a result of which the effective quadrupole splitting parameter A_{eff} can increase with temperature within a certain temperature range.

At present we know of two examples of direct spectral manifestation of ordinary quadrupole centers in magnetic materials.[1,7] In both cases, Fe^{3+} ions at the interstitial sites of the FeCl_2 lattice (Ref. 1) or occupying ‘‘illegal’’ positions c in the orthoferrite YFeO_3 (Ref. 7) act as a kind of probe. For the impurity center $\text{Fe}^{3+}(c)$ in YFeO_3, the presence of five lines in the spectrum has made it possible to determine the constants in the effective impurity Hamiltonian and to estimate the impurity–matrix exchange coupling parameters.[7] Here the experimentally determined quadrupole splitting constant $A\cong0.5\ \text{cm}^{-1}$ proved to be smaller by a factor of ten than the fraction $J^2/J_{12}\sim10\ \text{cm}^{-1}$ found from theoretical estimates.[7]

The only obvious reason for the experimental value of A of Ref. 7 to be so small is almost perfect balance between the exchange and spin–phonon contributions to A_{eff}. In this connection it seems natural to study analogous quadrupole centers in substances that are similar in structure, and to establish whether such a perfect balance of these contributions is a unique property or a general law.

2.2. Strongly coupled quadrupole centers

A characteristic feature of high-$T_c$superconducting oxides is that the exchange interaction between the holes formed as a result of doping and the matrix spins is much stronger than the exchange interaction in the matrix proper, i.e., $|J|>|J_{ii'}|$. In describing the energy spectra of such strongly coupled quadrupole centers it is convenient to use the approach suggested in Ref. 4. Here for the zeroth approximation we take the spectrum of an isolated complex consisting of an impurity center and the nearest spin neighbors. Then the interaction of this complex and the remaining matrix is taken into account by perturbation techniques. Here the smallness parameter is $1/z_0$, where z_0 is the number of the nearest magnetic neighbors of a spin in the matrix. For instance, the spectrum of a complex consisting of an impurity spin and the two nearest magnetic ions belonging to different sublattices of an antiferromagnet can be described in the zeroth approximation by the following expression:

$$E(\overline{S},S_{12},M_{\overline{S}})=\frac{J_{10}(\overline{S(S+1)}-S_{12}(S_{12}+1)-S(S+1))}{2}+J_{12}\left(\frac{S_{12}(S_{12}+1)}{2}-S_0(S_0+1)\right), \qquad (5)$$

where J_{10} and J_{12} are the parameters of the exchange interaction in the complex between the spins of the impurity and matrix and between the spins in the matrix, respectively, S_{12} is the total spin of two matrix atoms, and $\overline{S}$ is the total spin of the complex.

We see that the nature of the ground state of the impurity complex strongly depends on the signs of J_{10} and J_{12}, in contrast to the case of a weakly coupled quadrupole center, where the sign of the impurity–matrix exchange interaction has no effect on the structure of the spectral terms. Here for a complex where antiferromagnetic interaction J_{10} is dominant ($J_{10}>0$) the ground state corresponds to a maximum value of S_{12} and a minimum value of $\overline{S}$.

In this approximation, when the interaction between the isolated complex and the matrix is ignored, all energy levels (5) are degenerate in the magnetic quantum number $M_{\overline{S}}$ corresponding to the projection of the total spin of the complex.

The degeneracy is lifted when we allow for the effective exchange fields, described by the Hamiltonian $\mathscr{H}_{\text{eff}}(\alpha)$, through which the remaining magnetic atoms of the matrix act on the spins of the complex ($\alpha=0$, 1, 2 is the number of the atom in the complex). Second-order perturbation theory in the above parameter yields the following corrections $\Delta E(S,S_{12},M_{\overline{S}})$ to the energy $E(S,S_{12},M_{\overline{S}})$:

$$\Delta E(\overline{S},S_{12},M_{\overline{S}})=A(\overline{S},S_{12})[M_{\overline{S}}^2-\overline{S}(\overline{S}+1)], \tag{6}$$

where the constants $A(\overline{S},S_{12})$ prove to be positive (as they are in (2)), but their value is determined by the exchange parameters only for the matrix atoms.[4] Thus, the ground state of the complex in this approximation is achieved only when the projection $M_{\overline{S}}$ of the total spin S is at its minimum. However, in the present case the renormalization of the quadrupole splitting parameter caused by spin–phonon coupling is still determined by the corresponding parameters for the matrix atoms.

Note that the impurity center examined by Aristov and Maleyev[9], which incorporates two Cu^{2+} ions and a spin-1/2 hole (O^-) between the ions, fully corresponds to the model of a strongly coupled quadrupole impurity adopted in the present paper and in Ref. 4. Here, in the (generally accepted) case of antiferromagnetic interaction between the hole and the copper ions, the ground state in the spin triad Cu^{2+}–O^-–Cu^{2+} is, according to (5), the state with an intermediate spin $S_{12}=1$ and a total spin $S=$ 1/2 rather than $S=3/2$, as is the case in Ref. 9. Nevertheless, here too the ground state proves to be twofold degenerate and contributes considerably to the magnetic properties of high-T_c superconducting oxides.

3. ORBITALLY DEGENERATE QUADRUPOLE CENTERS

Let us now examine the case where a position symmetric with respect to the magnetic sublattices of the antiferromagnetic is occupied by a Jahn–Teller impurity ion in which the ground state is twofold orbitally degenerate in a cubic crystalline field (the E-state). Such a situation occurs, for instance, when Jahn–Teller centers replace nonmagnetic tetrahedral cations in the spinel lattice with antiferromagnetic ordering of the magnetic ions in octopositions.

The exchange interaction of a Jahn–Teller impurity in the nearest magnetic ions can be described by the following Hamiltonian:[15]

$$\mathscr{H}=\sum_{\delta}\left(L(\mathbf{S}\cdot\mathbf{S}_\delta)+\widetilde{L}\sum_{\mu=\theta,\varepsilon}C_{E\mu,\delta}(\mathbf{S}\cdot\mathbf{S}_\delta)\mathbf{U}_{E\mu}\right),$$
$$\mathbf{U}_{E\theta}=\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix},\quad \mathbf{U}_{E\varepsilon}=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \tag{7}$$

where the S_δ are the operators of the spins of the nearest neighbors of the impurity, L and $\widetilde{L}$ are the exchange interaction parameters for an impurity with orbital degeneracy, and the $\mathbf{U}_{E\mu}$ are the orbital operators specified in the space of the wave functions of the ground E-state. The expressions for the unitary matrices $C_{E\mu,\delta}$ for some impurity-cluster types can be found in Ref. 15.

Here we limit ourselves to the case of a weak exchange interaction between the centers and the nearest neighbors, i.e., $|L|,|\widetilde{L}|\ll|H_\delta|$, where H_δ is the molecular field acting on the matrix spins. Then the effective Hamiltonian of an isolated quadrupole center is given by the following expression, accurate to within second-order terms in the above small parameters:

$$\mathscr{H}_{\text{eff}}=\mathbf{A}(S_z^2-S(S+1))+\mathbf{B}S_z, \tag{8}$$

$$\mathbf{A}=A_0\mathbf{I},\quad A_0=\frac{(L^2+2\widetilde{L}^2)S_0z}{|2H_\delta|},\quad \mathbf{I}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix},$$

$$\mathbf{B}=\sum_{\mu}B_{E\mu}\mathbf{U}_{E\mu},$$

$$B_{E\mu}=\widetilde{L}\sum_{\delta}C_{E\mu,\delta}[\langle S_\delta^z\rangle-L(S_0(S_0+1)-\langle(S_\delta^z)^2\rangle)/H_\delta],$$

where z is the number of magnetic ions nearest to the impurity.

If when we allow for magnetic order in the matrix the local symmetry of the impurity center does not change (and coincides with the crystallographic symmetry), then only the first term, which is independent of the orbital variables ($B_{E\mu}=0$), remains in the effective Hamiltonian (8) of the impurity center. Here the splitting of the magnetic states is described by the same effective Hamiltonian as in the case of orbitally nondegenerate quadrupole centers:

$$E(M)=A_0M^2.$$

But when the symmetry of the impurity cluster with allowance for the spin orientations of the magnetic ions nearest to the Jahn–Teller center is lower than the crystallographic symmetry, an additional splitting of the spectrum of localized states may occur:

$$E_{\mp}(M)=A_0M^2\mp\sqrt{(B_{E\theta}M+h_{E\theta})^2+(B_{E\varepsilon}M+h_{E\varepsilon})^2}, \tag{9}$$

where the $h_{E\mu}$ are the components of the low-symmetry crystalline field on the Jahn–Teller ion. If the lowering of the symmetry on the impurity is caused only by exchange interactions ($h_{E\mu}=0$), the twofold degeneracy of the energy levels of the Jahn–Teller center is retained for all $M\neq 0$:

$$E_{\mp}(M)=E_{\pm}(-M).$$

Equation (9) clearly shows that the value of $|M|$ for the impurity spin in the ground state depends on the parameter $\eta=\sqrt{B_{E\theta}^2+B_{E\varepsilon}^2}/A_0$ and, in contrast to the case of an orbitally nondegenerate quadrupole center, may not assume its minimum value. In particular, for an impurity with an integer spin the ground state transforms with increasing η in the following manner: the nondegenerate nonmagnetic state with $M=0$ is successively replaced by twofold degenerate states with spin projections $\mp M$ ($M\neq 0$) when the parameter η reaches the corresponding critical values $\eta_{\text{crit}}(M)=(2|M|+1)$. Such behavior of the ground magnetic state is the magnetic analog of the Jahn–Teller effect.

Let us now analyze the effect of quadrupole Jahn–Teller centers on the magnetic anisotropy energy and the magnetoelastic properties (usually Jahn–Teller ions provide an anomalously large contribution to these characteristics of the magnetic material). The corresponding terms in the impurity Hamiltonian that describe single-ion anisotropy on a Jahn–Teller center, and the interaction of the center and uniform magnetostriction strains have the form[16]

$$\mathscr{H}_a=\mathbf{U}_{E\theta}\left(D\left[S_z^2-\frac{S(S+1)}{3}\right]+V_E e_{E\theta}\right)+\mathbf{U}_{E\varepsilon}\left(\frac{D(S_x^2-S_y^2)}{\sqrt{3}}+V_E e_{E\varepsilon}\right), \tag{10}$$

where D is the single-ion anisotropy constant, and V_E is the parameter of the interaction of the Jahn–Teller ion with the E_g-strains $e_{E\theta}=(2e_{zz}-e_{xx}-e_{yy})/2$ and $e_{E\varepsilon}=\sqrt{3}(e_{xx}-e_{yy})/2$ in a cubic magnetic material.

For the sake of brevity let us examine the case where we have $B_{E\mu}=0$ and the Hamiltonian (10) can be considered a perturbation with respect to the energy levels (9). Then the magnetic anisotropy energy E_a related to the Jahn–Teller subsystem can be written as

$$E_a=-x_{JT}\left|D\left\langle M^2-\frac{S(S+1)}{3}\right\rangle_0\right|\times\sqrt{1-3\sum_{i<j}\beta_i^2\beta_j^2},\quad k_BT\ll|D|, \tag{11}$$

where x_{JT} is the concentration of Jahn–Teller centers, $\langle\cdots\rangle_0$ stands for the quantum average in the ground state, and the β_i are the direction cosines of the sublattice magnetizations with respect to the crystallographic axes. We see that even when the ground state of quadrupole centers with integer spin is nonmagnetic ($M=0$), the contribution of these centers to the magnetic anisotropy of the crystal is of the same order of magnitude as that of a Jahn–Teller ion with an ordinary dipole spectrum. The maximum splitting of the doublet and, accordingly, the greatest absolute value of E_a are achieved when magnetization is directed along the tetragonal axes of the crystal.

According to Ref. 16, the contribution of such Jahn–Teller centers, ΔB_1, to the magnetoelastic constant B_1 at low temperatures $k_BT\ll|D|$ is given by the following expression:

$$\Delta B_1=x_{JT}|V_E|\operatorname{sgn}\left(V_E D\left\langle M^2-\frac{S(S+1)}{3}\right\rangle_0\right). \tag{12}$$

What is important is that for quadrupole (with $M=0$ or $\mp 1/2$ in the ground state) and dipole (with $M=S$ in the ground state) centers not only do the values of the corresponding contributions to B_1 differ but so do their signs. A similar effect exists for the contribution of the Jahn–Teller subsystem to the magnetostriction constant $\lambda_{[001]}$.

4. QUADRUPOLE CENTERS WITH 3*d*-IONS OF MIXED VALENCE

In this section we discuss another type of quadrupole center in an antiferromagnet: an impurity cluster containing mixed-valence $3d$-ions. Such clusters form in various compounds as a result of doping the compounds by nonisovalent substitutional atoms or in the presence of nonstoichiometry in the cation or anion sublattices. Such centers are characterized by a strong resonant coupling of the $3d$-ions in equivalent crystallographic positions nearest to the defect, a situation realized by the transfer of an excess electron or hole from one magnetic $3d$-ion of the complex to another (double Zener exchange). The tunneling splitting of the states of such a cluster is at its maximum and the ground-state energy is at its minimum when all the spins in the complex are aligned. When the cluster center occupies a position symmetric with respect to the magnetic sublattices of the antiferromagnet and the resonant exchange interaction of cluster ions dominates, a new type of quadrupole center, a mixed-valence complex appears. In the mean-field approximation, the exchange interaction of such a complex and the nearest antiferromagnetically ordered spins of the matrix is found to be fully balanced.

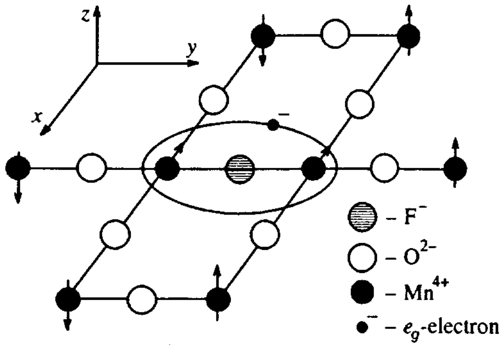


FIG. 1. A mixed-valence quadrupole center coupled with an anion substitution of F^- for O^{2-} in the Me–O plane of the antiferromagnet.

Usually one of the charge states of mixed-valence ions is orbitally degenerate. Hence describing the behavior of a mixed-valence complex requires proper generalization of double-exchange theory[17] and analysis on its basis of the structure of the ground state of the complex and the complex–matrix exchange interaction. In solving this problem we begin with a mixed-valence complex of minimum size, and then discuss more complicated quadrupole centers.

4.1. Double exchange in a pair of 3*d*-ions with orbitally degenerate resonance states

Let us take a cubic or quasi-two-dimensional antiferromagnet with anion substitution of, say, F^- for O^{2-} in the Me–O plane with antiparallel-ordered magnetic $3d$-ions Me. This gives rise to the simplest mixed-valence center: a pair of $3d$-ions of Me separated by the F^- ion and an excess e_g- or t_{2g}-electron localized at this complex (Fig. 1). For the sake of definiteness we select $Mn^{4+}(3d^3)$ ions as the main magnetic ions, as the substitution of F^- for O^{2-} leads to another state, $Mn^{3+}(3d^4)$, which corresponds to the addition of an e_g-electron to the cation subsystem. This is the situation, for instance, for the $Ca_2MnO_{4-x}F_x$ ($x\ll 1$) compound.[18]

For the basis of the lowest electron states of the cluster we take the one corresponding to the localization of the excess electron on one of the two (a,b) magnetic ions of the complex accompanied by the formation of the following configurations: $Mn^{3+}(a)$–$Mn^{4+}(b)$ and $Mn^{4+}(a)$–$Mn^{3+}(b)$. The ground state of the Mn^{4+} ion is the orbital singlet ${}^4A_2$, and the ground state of the Mn^{3+} ion in a cubic crystalline field is the orbitally degenerate ${}^4E$-state. The degeneracy is lifted by the following low-symmetry crystalline fields: the rhombic field Δ_r, caused by the presence of the F^- ion on the Mn^{3+}–Mn^{4+} line, and the tetragonal field Δ_t, present in crystals with a quasi-two-dimensional structure. Nevertheless, in analyzing the spectrum we allow for both orbital states of the Mn^{3+} ion, since the resonant interaction may prove to be comparable or even stronger than the low-symmetry fields Δ_t and Δ_r. All these interactions are assumed to be weaker than intratomic exchange.

Let us write the wave functions of the mixed-valence complex as a combination of antisymmetrized products of the wave function of the Mn^{3+} and Mn^{4+} ions:

$$\Psi_n(\bar{S}M_{\bar{S}})=\frac{\psi(a,A_2,S_0;b,E\mu,S;\bar{S}M_{\bar{S}})}{\sqrt{2}}\mp\frac{\psi(a,E\mu,S;b,A_2,S_0;\bar{S}M_{\bar{S}})}{\sqrt{2}},\tag{13}$$

where S and S_0 are the spins of the Mn^{3+} and Mn^{4+} ions, $\bar{S}$ and $M_{\bar{S}}$ are the total spin of the pair and its projection on the quantization axis, and the index μ numbers the orthogonal states of the orbital doublet on the Mn^{3+} ion. The corresponding energy levels $E_n(E_n\equiv E_n(\bar{S},M_{\bar{S}}),\ n=1–4)$ of the mixed-valence complex are given by the following expression (which does not allow for Heisenberg exchange):

$$E_{1,2}=\frac{1}{2}\left[t\mp\sqrt{(t+2\Delta_r)^2+4\Delta_t^2+2\Delta_t(t+2\Delta_r)}\right],\tag{14}$$

$$E_{3,4}=\frac{1}{2}\left[-t\mp\sqrt{(-t+2\Delta_r)^2+4\Delta_t^2+2\Delta_t(-t+2\Delta_r)}\right],$$

$$t\equiv t(\bar{S})=\frac{b(\bar{S}+1/2)}{2S_0+1},$$

$$b\cong 4b_{E\theta,E\theta}\cong\frac{4}{3}b_{E\varepsilon,E\varepsilon}\cong\frac{4}{\sqrt{3}}b_{E\theta,E\varepsilon},$$

where b is the transfer integral for the e_g-electron in the pair we are considering here,[17] and the energies $E_{1,2}$ and $E_{3,4}$ corresponds to the respective wave functions with + and − signs in (13). Figures 2a and b depicts the curves representing the dependence of the resonance state energies E_n on the double-exchange parameter $t(\bar{S})$.

An analysis of Eqs. (14) (see also Fig. 2) shows that the energy of the lowest state, E_n, decreases with increasing $|t|$ (unless $\Delta_r>|b|/2$ and $\Delta_t=0$). Since the parameter t describing resonant splitting is at its maximum at $\bar{S}_{\max}=S+S_0=2S_0+1/2$, the ground state of the mixed-valence complex correspond to parallel orientation of the spins. Here the energy of exchange splitting, corresponding to the change of $\bar{S}$ from 1/2 to $\bar{S}_{\max}$, is equal in order of magnitude to the transfer integral b if the rhombic crystalline field Δ_r is smaller than $|b|/2$. But if $\Delta_r>|b|/2$, the exchange energy is reduced considerably for $|\Delta_t/\Delta_r|\ll 1$ and amounts to a quantity of order $(\Delta_t/\Delta_r)^2 b$.

The case $\Delta_t=0$ is of special interest. It corresponds to mixed-valence complexes in cubic crystals (in the absence of random crystalline fields of tetragonal symmetry). In this case, for $\Delta_r>|b|/2$ there is no charge transfer in the ground state of the complex (with an energy $E_1=E_3=-\Delta_r$), so that the degenerate levels of the ground state correspond to total localization of an e_g-electron on the ions of the a- or b-pair. The absence of effects associated with the transfer of an e_g-electron is due in this case to the nature of the orbital wave function of the Mn^{3+} ion, which in the present conditions is oriented perpendicularly to the y axis of the complex. Note that random crystalline fields with $\Delta_t\neq 0$ may have a strong effect on the nature of the exchange interaction in such mixed-valence complexes.

The Heisenberg exchange interaction J_{ab} in the (a,b) pair provides a certain contribution to the energy levels E_n $(n=1–4)$ of the mixed-valence complex, tending to align the spins of the Mn^{3+}–Mn^{4+} ions in an antiparallel manner. These effects are important only when $\Delta_r>|b|/2$ and $|\Delta_t/\Delta_r|\ll 1$ hold, i.e., when the resonant splitting of the two lowest states is small (in absolute value comparable to, or smaller than, the Heisenberg exchange in the pair).

Below for the sake of simplicity we examine only the limiting cases in the behavior of the Mn^{3+}–Mn^{4+} pair corresponding to the Heisenberg exchange or double exchange being dominant. The degeneracy in $M_{\bar{S}}$ of the ground state

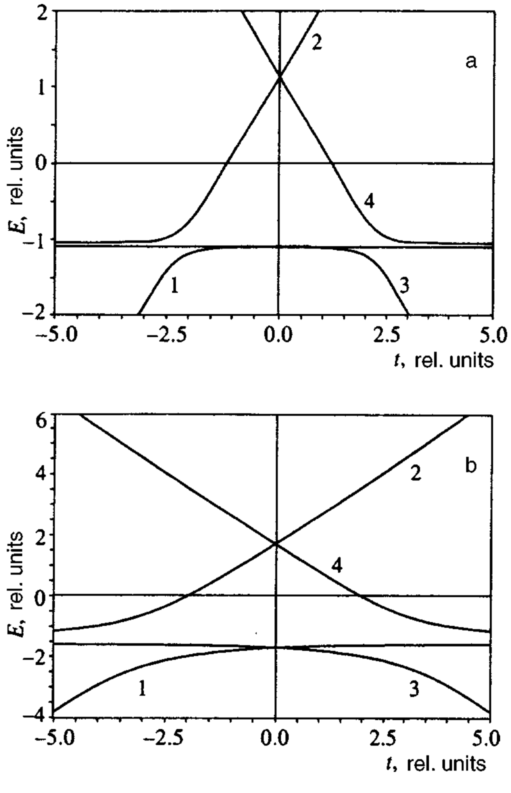


FIG. 2. The energy spectrum of a mixed-valence quadrupole center as a function of the transfer parameter t (both E and t are measured in units of Δ_r): $\Delta_t/\Delta_r=0.2$ (a), and $\Delta_t/\Delta_r=1$ (b).

($\bar{S}$ in the first case and $\bar{S}=\bar{S}_{\max}=2S_0+1/2$ in the second) of such complexes is lifted if we allow for the exchange interaction between the total spin $\bar{S}$ of the complex and the nearest spins in the matrix. We start by examining the qualitative nature of the behavior of mixed-valence centers in an antiferromagnet, ignoring the possible effects of renormalization of the parameters of the exchange interaction between the spins of the mixed-valence complex and the spins of the remaining matrix caused by the transfer of an e_g-electron.

4.2. The effect of the interaction with magnetic sublattices on the behavior of mixed-valence centers

When $(\Delta_t/\Delta_r)^2|b| \ll |J_{ab}|$ holds and we can ignore the effects of electron transfer on a mixed-valence complex, the ground state of this complex in the molecular field of the matrix is characterized by the projections of the total spin, $M_{\bar{S}}=1/2$ or $M_{\bar{S}}=-1/2$, depending at what ion (a or b) the excess charge is localized. As a result, the degeneracy of the ground state of the mixed-valence complex is of a combined nature, where one projection $M_{\bar{S}}$ of the total spin on the antiferromagnetism axis corresponds to a well-defined projection of the electric dipole moment $\mathbf{p}_0$ on the y axis of the complex. Thus, we arrive at a peculiar two-level system, in which direct tunneling between its states with $\Delta M_{\bar{S}}=\mp 1$ is impossible, so that the corresponding reorientation of such a mixed-valence complex may be only activational. Stabilization of these states of the two-level system may occur in external fields (electric, magnetic, or stress field) and in random crystalline fields.

A characteristic feature of the two-level systems considered here is the strong magnetoelectric effect that they produce. The interaction of such two-level system with an electric field E_y and a magnetic field $\mathbf{H}\|\mathbf{M}$ has the form

$$\mathscr{H}=-(p_0E_y+\mu_B gH/2)\tau_z, \tag{15}$$

$$\tau_z=\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad p_0=|e|r,$$

where e is the electron charge, r is the distance between the Mn and F ions, μ_B is the Bohr magneton, and g is the gyromagnetic ratio of the mixed-valence center. As a result, in the absence of random crystalline fields, the free energy F of the subsystem consisting of N mixed-valence centers is

$$F=-Nk_BT\,\ln\left(2\cosh\frac{p_0E_y+\mu_BH}{k_BT}\right), \quad g\cong 2. \tag{16}$$

This expression describes a magnetoelectric effect of an unusual type. At temperatures exceeding the splitting energy of the two-level system, the effect is transformed into an ordinary linear magnetoelectric effect, when the corresponding contribution F_{EH} to the free energy assumes the form $F_{EH}\sim -NE_yH/T$. In random crystalline fields with dispersion Γ the effect is reduced in proportion to T/Γ for $T<\Gamma$.

Let us now analyze the behavior of an impurity center in an antiferromagnetic matrix when double exchange is dominant in the mixed-valence complex. In this case, using second-order perturbation theory, we arrive at the following expression for the energy of quadrupole splitting of the main spectral term (E_1 or E_3):

$$\Delta E(\bar{S},M_{\bar{S}})=A(M_{\bar{S}}^2-\overline{S(S+1)}),$$
$$A=\sum_\delta \frac{J_\delta^2}{2(z-1)J_0}, \quad \bar{S}=S+S_0, \tag{17}$$

where J_0 is the value of the exchange interaction ($J_0>0$) between the Mn^{4+} ions in the matrix, and the J_δ are the parameters of exchange between the mixed-valence center and the spins of the Mn^{4+} ions nearest to the center. The convergence of the perturbation-theory series is ensured in this case by the parameter $1/z$, as it is in the case of strongly coupled ordinary quadrupole centers.

Note that in contrast to the earlier situation with $(\Delta_t/\Delta_r)^2|b| \ll |J_{ab}|$, here the dipole moment at the mixed-valence center is zero in the ground state, in view of the symmetric nature of its wave function.

Obviously, the type of quadrupole centers with an excess electron or hole considered here is always characterized by a half-integer total span $\bar{S}$. As a result, irrespective of the sign of the quadrupole splitting parameter A, the ground state of the center proves to be twofold degenerate. At present a number of basic features of the behavior of crystals with such two-level systems are known, but they are still a topic of great interest in research, including research that involves high-T_c superconducting oxides.

The specific properties of quadrupole mixed-valence centers are due to the presence of orbital degeneracy in one of the $3d$-electron configurations, $3d^n$ or $3d^{n\mp 1}$. As already noted in Sec. 3, this manifests itself in a considerable contribution of such centers to the constants of magnetic crystallographic anisotropy and magnetostriction at low temperatures. The energy of magnetic anisotropy, $E_a(M_{\bar{S}})$, of the quadrupole center (Fig. 1) can be obtained by averaging the Hamiltonian (10) over the ground-state wave function Ψ_0 of the complex:

$$E_a(M_{\bar{S}})=\left[D_\|\left(\beta_z^2-\frac{1}{3}\right)+D_\perp(\beta_x^2-\beta_y^2)\right]$$
$$\times(M_{\bar{S}}^2-\overline{S(S+1)}/3), \tag{18}$$

$$D_\|=\frac{3}{2}\kappa(\bar{S})D\cos\varphi, \quad D_\perp=\frac{\sqrt{3}}{2}\kappa(\bar{S})D\sin\varphi,$$

where D is the single-ion anisotropy constant of the Mn^{3+} ion, $\kappa(S)$ is a reduction factor ($\kappa(S)<1$), and the angle φ is related to the wave function Ψ_0 as follows:

$$\Psi_0\sim\left|\frac{3z^2-r^2}{2}\right\rangle\cos\frac{\varphi}{2}+\left|\frac{\sqrt{3}(x^2-y^2)}{2}\right\rangle\sin\frac{\varphi}{2}.$$

We see that the anisotropy energy (18) is determined by the large constant D typical of Jahn–Teller ions and strongly depends on the effects of charge transfer at the impurity center. As a result, the value and sign of the corresponding effects for the impurity contribution to the magnetic-anisotropy and magnetostriction may differ from those observed in

Jahn–Teller ions, which have a fixed charge state and exist in a nonzero exchange field. Note in this connection that the anisotropy energy (18) is also a source of competing anisotropy in systems where impurity mixed-valence complexes have different orientation axes.

4.3. Renormalization of the exchange parameters; enlargement of mixed-valence complexes

Now let us examine the renormalization of the parameters of exchange between the impurity center and the matrix spins. Here we allow for the transfer of an e_g-electron between the ions of the mixed-valence complex and the magnetic ions of the matrix. We also assume that the difference u in the potential energies of the localization of an excess change on the first and second (third) magnetic neighbors of the F^- ion is much larger than the transfer integral b_1 between these ions through the oxygen ion O^{2-} (the deep impurity model). Then the corresponding effective exchange interactions can be described by the following Hamiltonian:

$$\mathscr{H}=\sum_{\delta=1}^{z'} J_\delta(\overline{\mathbf{S}}\mathbf{S}_\delta),$$

$$J_\delta=\gamma J_0+\beta\left(J_1-\frac{2b_\delta^2}{(2S_0+1)^2u_\delta}\right), \tag{19}$$

$$\gamma=\frac{S_0}{2(S+S_0)},\quad \beta=\frac{S}{2(S+S_0)},$$

where J_1 is the superexchange interaction integral $(J_1>0)$ in the Mn^{3+}–O^{2-}–Mn^{4+} configuration, the b_δ are the transfer parameters of an e_g-electron between the ions of the complex and matrix through the intermediate oxygen ion O^{2-}, and z' is the number of the nearest magnetic neighbors of the mixed-valence complex.

We see that the ferromagnetic interactions $(\sim b_\delta^2/u)$ related to the transfer of an e_g-electron may change not only the magnitude but also the sign of the parameters J_δ of the exchange interaction. Here the ferromagnetic contributions to J_δ for the bonds oriented along the x and y axes of the mixed-valence complex differ (see Fig. 1), thanks to the anisotropy of the transfer integrals b_δ:

$$\begin{aligned} b_\delta&=\frac{b'}{2}\left[1+\cos\left(\varphi+\frac{2\pi}{3}\right)\right] \text{ along the } x \text{ axis,}\\ b_\delta&=\frac{b'}{2}\left[1+\cos\left(\varphi-\frac{2\pi}{3}\right)\right] \text{ along the } y \text{ axis,} \end{aligned} \tag{20}$$

where the prime on the transfer integral (b') implies that, in contrast to b, the transfer of an e_g-electron occurs through oxygen rather than through fluorine. For instance, for $|t|\gg|\Delta_r|,|\Delta_t|$ (and, correspondingly, $\varphi\cong 2\pi/3$), the integral of transfer along the y axis is almost four times greater than that along the x axis.

The ferromagnetic (resonant) contribution may dominate in J_δ if

$$|b'|\gg|b_0|\sqrt{\frac{u}{U}}, \tag{21}$$

where U is the energy of the repulsion of electrons at a single center, and b_0 is the transfer integral of a t_{2g}-electron, responsible for the superexchange interactions between the Mn^{4+} ions in the matrix and the Mn^{3+}–O^{2-}–Mn^{4+} pair (here $J,J_0\sim b_0^2/U$). What is important is that the condition (21) is met over a broad range of values of the transfer integral b', since the potential energy difference u is usually much smaller than the repulsion energy U.

Note that for large negative values of J_δ $(|J_\delta|\gg J_0,J)$ it becomes necessary to increase the dimensions of the magnetic impurity cluster without changing the radius of electron localization of the e_g-electron. Here, in the limit of $(\Delta_t/\Delta_r)^2|b|\ll|J_{ab}|$, the effective value of the total spin $\overline{S}$ of mixed-valence centers must increase considerably, and so does the height of the potential barrier between the states of the two-level system with projections $M_{\overline{S}}=\mp\overline{S}$. As a result, the reorientation of the two-level system and the related properties may turn out to be strongly dependent on the presence of agents inducing transitions (photoinduced effects).

In the case of mixed-valence centers of the quadrupole type, an increase in the size of the impurity complex substantially increases the contribution of these complexes to the magnetic susceptibility, specific heat, etc. Similar effects of increased contribution of the impurity centers to the various observables linked, however, to different causes are known to exist in other systems, such as magnetic impurities in strongly correlated paramagnets[19] and orthogonal impurity centers in magnetic materials.[6]

4.4. More complicated types of mixed valence in antiferromagnets

Let us examine the more complicated centers with charge transfer in antiferromagnetic crystals of a cubic or quasi-two-dimensional structure. In such substances one can expect the emergence of mixed-valence centers with four $3d$-ions nearest to the corresponding defect, and it is between these ions that electron transfer occurs. Here the defect and the cations may belong to different crystallographic planes, say, when the defect is at the vertex of a pyramid (e.g., the $Ca_{2-x}Y_xMnO_4$ and $Nd_{2-x}Ce_xCuO_4$ compounds).

Let us discuss the $Ca_{2-x}Y_xMnO_4$ system in greater detail. Here there is transfer of an e_g-electron between Mn^{4+} ions, just as there is in the case of a complex with two $3d$-ions discussed above. When the transfer integral b is small (in comparison to the parameters of antiferromagnetic exchange for Mn^{4+}–Mn^{3+} ions), we can assume that the ground state of the mixed-valence complex is fourfold degenerate, in accordance with the number of possible positions for localizing an e_g-electron. Each of these four states is characterized by a specific direction of the dipole moment of the complex and a projection of the total complex spin ($M_{\overline{S}}=1/2$ or $-1/2$, depending to which antiferromagnetic sublattice the position belongs). As a result, a kind of magnetoelectric effect may manifest itself in the impurity subsystem. This effect amounts to the following. If the e_g-electron is stabilized in one of the four positions by directing the electric field along the diagonal of the pyramid's base, we have fixed the magnetic moment of the mixed-valence complex and of the entire system as a whole. The corresponding free energy of the impurity subsystem in an electric field **E** and a magnetic field **H**‖**M** is

$$F=-Nk_BT\ \ln\left\{2\left[\exp\left(\frac{\mu_BH}{k_BT}\right)\cosh\frac{p_0(E_x+E_y)}{k_BT}\right.\right.$$
$$\left.\left.+\exp\left(-\frac{\mu_BH}{k_BT}\right)\cosh\frac{p_0(E_x-E_y)}{k_BT}\right]\right\}. \tag{22}$$

We see that in contrast to the situation examined earlier and involving the Mn^{4+}–Mn^{3+}pair, the present magnetoelectric effect is even in the electric field.

When strong resonant coupling is present, i.e., when the electron is moving along the plaquette of four Mn^{4+} ions, the energy levels of the mixed-valence complex in a state with spin $S=4S_0+1/2$ are

$$E(A_1)=-\Delta_t+\frac{b}{2},\quad E(B_1)=\Delta_t+3b/2\,,$$
$$E(B_2)=-\Delta_t-b/2\,, \tag{23}$$
$$E(A_2)=\Delta_t-3b/2\,,\quad E_{\mp}(E)=\mp\sqrt{\Delta_t^2+3b^2/4}.$$

These states of the mixed-valence center are classified according to the Γ representations of the C_{4v} group ($\Gamma=A_1$, A_2, B_1, B_2, and $2E$) and are degenerate in M_S in first-order perturbation theory in the molecular field of the matrix. In the same way as for ordinary quadrupole centers, the degeneracy is lifted in the second order, and the corresponding energy levels are described by an effective Hamiltonian of type (2). The states in question have a zero dipole moment in the Mn–O plane.

Usually the ground state of a mixed-valence center is orbitally nondegenerate. But in the range of values $b\cong\mp2\Delta_t$ the spectral terms, one of which is orbitally degenerate, move closer to each other. Here the Jahn–Teller state $E_-(E)$ may turn out to be the lowest if we allow for the orbitally dependent part of the exchange interaction between the mixed-valence complex and the matrix. In this case the quadrupole center possesses the specific features discussed in Sec. 3.

Thus, the behavior of mixed-valence centers that occupy positions symmetric with respect to the magnetic sublattices of an antiferromagnet is determined by the mechanism for stabilizing the ground state, i.e., the relative role of the transfer effects. The type of mixed-valence complex (the number of ions in the complex and their spatial configuration) manifests itself in the specific values of the total spin and its projections, in the geometry of the magnetoelectric effect, in values of the parameters of the effective Hamiltonian, etc.

5. CONCLUSION

Our study of the different types of quadrupole centers in antiferromagnets suggests the following.

1. There exists an additional spin-phonon coupling mechanism for quadrupole splitting of the spectrum of impurity centers in a zero exchange field of the magnetic sublattices of the antiferromagnet. It can reduce the value of the quadrupole splitting parameter and change the sign of that parameter.

2. A special type of quadrupole center is the one in which the ground state is orbitally degenerate. The specific features of the properties of such centers are related to the individual properties of the exchange and spin–orbit couplings and to the presence of strong vibronic effects in the orbitally degenerate states. We have predicted the existence of a magnetic analog of the Jahn–Teller effect for such centers. We have also found that the contribution of the Jahn–Teller impurities to the magnetic-anisotropy and magnetostriction constants is exceptionally large, but the value and sign of the corresponding effects differ from the typical ones for ordinary Jahn–Teller ions in magnetic crystals.

3. We have studied a new class of magnetic impurities that we believe to be quite broad: mixed-valence centers positioned symmetrically with respect to the sublattices of antiferromagnets of different types. We have shown that the properties of these mixed-valence centers strongly depend on the nature of localization of excess charge. For small transfer integrals, twofold degeneracy of the ground states in the projections $\mp|M_S|$ is accompanied by dipole degeneracy with a multiplicity equal to the number of mixed-valence ions in the impurity cluster. As a result, the mixed-valence centers may give rise to strong magnetoelectric effects of an extraordinary nature. On the other hand, in systems where resonant interactions dominate, the mixed-valence centers are described by an effective Hamiltonian of the quadrupole type, whose ground state is twofold degenerate in $\mp|M_S|$. Delocalization of the excess charge in this case leads to a situation in which the dipole moment of the center in the ground state vanishes. In both cases such centers with a degenerate ground state can have an anomalously strong effect on the resonant, magnetic, and thermodynamic properties of various substances. The behavior of the magnetic susceptibility characteristic of the systems discussed above has, probably, been observed in the $Ca_2MnO_{4-x}F_x$ compound.[18]

The authors are grateful to V. M. Loktev for useful discussions of problems pertaining to high-$T_c$superconducting oxides.

The work was supported financially by the Russian Fund for Fundamental Research (Grant No. 94-03-08035) and the International Revival Fund (Grant No. ISSEP SPUO42026).

[1] M. Motokawa and M. Date, J. Phys. Soc. Jpn. **23**, 1216 (1967).
[2] M. A. Ivanov, Fiz. Tverd. Tela (Leningrad) **14**, 562 (1972).
[3] M. A. Ivanov and A. Ya. Fishman, Fiz. Tverd. Tela (Leningrad) **15**, 2378 (1973) [Sov. Phys. Solid State **15**, 1583 (1974)].
[4] M. A. Ivanov, V. Ya. Mitrofanov, and A. Ya. Fishman, Phys. Status Solidi B **61**, 403 (1974).
[5] M. A. Ivanov and E. F. Shender, Zh. Éksp. Teor. Fiz. **69**, 350 (1975) [Sov. Phys. JETP **42**, 179 (1976)].
[6] M. Ivanov, V. M. Loktev, and Yu. G. Pogorelov, Phys. Rep. **153**, 209 (1987).
[7] A. M. Balbashov, A. G. Berezin, Yu. V. Bobryshev, P. Yu. Marchukov, I. V. Nikolaev, Ya. Paches. L. Pust, E. G. Rudashevskii, and V. V. Shushpanov, Zh. Éksp. Teor. Fiz. **102**, 1397 (1992) [Sov. Phys. JETP **75**, 757 (1992)].
[8] A. O. Gogolin and A. S. Ioselevich, JETP Lett. **50**, 502 (1989).
[9] D. N. Aristov and S. V. Maleyev, Preprint 1587-LNPI (1990).

[10] M. A. Ivanov, V. M. Loktev, and Yu. G. Pogorelov, Zh. Éksp. Teor. Fiz. **101**, 596 (1992) [Sov. Phys. JETP **74**, 317 (1992)].

[11] M. A. Ivanov, V. M. Loktev, and Yu. G. Pogorelov, JETP Lett. **55**, 27 (1992).

[12] M. A. Ivanov, V. M. Loktev, and Yu. G. Pogorelov, Physica C **198**, 171 (1992).

[13] E. F. Shender, Zh. Éksp. Teor. Fiz. **83**, 326 (1982) [Sov. Phys. JETP **56**, 178 (1982)].

[14] V. M. Loktev, Fiz. Nizk. Temp. **19**, 375 (1993) [Low Temp. Phys. **19**, 263 (1993)].

[15] L. D. Fal'kovskaya, A. Ya. Fishman, M. A. Ivanov *et al.*, J. Magn. Magn. Mater. **36**, 26 (1983).

[16] L. D. Fal'kovskaya, A. Ya. Fishman, M. A. Ivanov *et al.*, J. Magn. Magn. Mater. **71**, 337 (1988).

[17] P. W. Anderson, in *Magnetism*, Vol. 1, edited by G. T. Rado and H. Suhl, Academic Press, New York (1963), p. 25.

[18] G. Le Flem, R. Colmet, J. Claverie *et al.*, J. Phys. Chem. Solids **41**, 55 (1980).

[19] I. Ya. Korenblit and E. F. Shender, Usp. Fiz. Nauk **126**, 233 (1978).

Translated by Eugene Yankovsky

Cyclotron resonance in the organic conductor (BEDO-TTF)$_2$ReO$_4$(H$_2$O) in the millimeter wavelength band

S. V. Demishev, A. V. Semeno, N. E. Sluchanko, N. A. Samarin, and I. B. Voskoboĭnikov

Institute of General Physics, Russian Academy of Sciences, 117942 Moscow, Russia

M. V. Kartsovnik and A. E. Kovalev

Institute of Solid State Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

N. D. Kushch

Institute of Chemical Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

(Submitted 3 June 1996)
Zh. Éksp. Teor. Fiz. **111**, 979–987 (March 1997)

Resonant microwave absorption in a (BEDO-TTF)$_2$ReO$_4$(H$_2$O) organic conductor single crystal at a temperature of 1.9 K, a magnetic field of up to 70 kOe, and in the frequency band between 30 and 120 GHz has been studied. A spectral component due to the cyclotron resonance (CR) of two-dimensional charge carriers has been identified for $\nu \geqslant 80$ GHz and $H \geqslant 10$ kOe. The effective mass $m(\omega)$ increases with the frequency from $m \approx 0.8m_0$ at $\nu = 80$ GHz to $m \approx 0.95m_0$ at $\nu = 120$ GHz. Measurements of the CR line position and FWHM as functions of frequency yield an independently determined imaginary part of the memory function $M(\omega)$, which controls the dynamic magnetoconductivity, and the relaxation time $\tau(\nu = 100\ \text{GHz}) \approx 2.9 \times 10^{-11}$ s, which is more than thirty times the value of this parameter in the low-frequency limit $\tau(\nu \rightarrow 0)$. The anomalous behavior of the CR line position and FWHM is caused by the dispersion of both real and imaginary parts of $M(\omega)$, which are probably due to strong Fermi-liquid effects. © *1997 American Institute of Physics.* [S1063-7761(97)01803-9]

1. INTRODUCTION

The cyclotron resonance (CR) recently detected in several quasi-two-dimensional organic metals[1–5] has generated special interest in studies of microwave absorption in these materials. The first publications on this topic (see Refs. 1 and 2, and references therein) reported the differences between the effective masses derived from quantum oscillations and CR curves. Cyclotron resonance yielded effective masses $m \sim (0.5-1)m_0$, where m_0 is the free electron mass, which were a factor of 2.5–3 smaller than the effective masses derived from Shubnikov–de Haas and de Haas–van Alphen measurements. This behavior led to a conclusion about strong Fermi-liquid effects resulting in a giant frequency-dependent effective mass renormalization.[1–5] The dispersion in the band of $\nu = \omega/2\pi = 30$–120 GHz was detected in our previous work,[4,5] where we found that the effective mass decreased with the frequency. It is remarkable that, in addition to light masses, CR modes corresponding to masses heavier than those derived from quantum oscillation measurements were also detected.[3–5]

From a technical standpoint, detection of magnetoabsorption in organic metals is a very difficult problem because of the small dimensions of single crystals of these materials. The first experiments[1,2] were performed with mosaic samples, and a technique that can be used in detecting signals due to individual organic single crystals was developed by us[3] and described in detail in subsequent publications.[4–6] As a result, the general structure of the resonant magneto-absorption spectrum in organic metals in strong magnetic fields was determined.[3,4] This structure is composed of broad absorption lines due to the cyclotron resonance on which narrower lines of various magnetic resonances are superposed, namely, the electron spin resonance and antiferromagnetic resonance, with amplitudes smaller by a factor of five to ten.[3–5] Thus, distinguishing CR lines from absorption spectra of organic metals with specific magnetic structures may be quite a difficult problem.

Moreover, it is noteworthy that most CR experiments were conducted using organic metals of the (BEDT-TTF)$_2$ MHg(SCN)$_4$ family, where M=K, Tl, NH$_4$, and BEDT-TTF is bis(ethylenedithio)tetrathiafulvalene,[1–6] the most thoroughly investigated cases being M=K, Tl,[3–5] i.e., two-dimensional conductors characterized at helium temperatures by antiferromagnetic ordering and consequently complicated magnetoabsorption spectra.

The aim of the present work directly derives from the above description of the present status of the problem. First, it is not clear to what extent the description of the cyclotron resonance in (BEDT-TTF)$_2$MHg(SCN)$_4$ applies to other two-dimensional organic conductors. Hence, the interest in studying an organic metal based on a cross-linking molecule different from BEDT-TTF and with a nonmagnetic ground state is obvious.

Second, the interactions responsible for the renormalization and dispersion of the effective mass should also give rise to the behavior of the relaxation time $\tau(\omega)$. This follows from both the theoretical analysis of cyclotron resonance in interacting systems[7] and experimental data on heavy-fermion systems.[8] The shape of CR lines in organic metals, however, has not been analyzed, and the issue of the correspondence between relaxation times in the high- and low-frequency experiments has not been addressed.

2. EXPERIMENTAL TECHNIQUES AND SAMPLES

Experiments were carried out using (BEDO-TTF)$_2$ ReO$_4$(H$_2$O) single crystals, where BEDO-TTF is bis(ethylendioxy)tetrithiafulvalene. Compared to (BEDT-TTF)$_2$MHg(SCN)$_4$, both the basic molecule forming the conducting plane and the anion have different chemical compositions. The physical and structural properties of (BEDO-TTF)$_2$ReO$_4$(H$_2$O) have been studied in detail.[9–12] It is known that the structure of this organic metal is similar to those of compounds based on BEDT-TTF, and its conductivity has a pronounced anisotropy: $\rho_c:\rho_a:\rho_b \approx 1:3:1000$ at $T=300$ K,[9–12] where a and c are the crystal axes in the conducting plane, and the b axis is perpendicular to the ac-plane. The quasi-two-dimensional nature of its conductivity at helium temperatures is confirmed by transport measurements, and studies of Shubnikov–de Haas oscillations yield the effective masses $m=0.9m_0$ and $m=1.15m_0$ (Ref. 10), corresponding to different sections of the Fermi surface. Temperatures $T \leqslant 3$ K correspond to the onset of the superconducting transition in (BEDO-TTF)$_2$ReO$_4$(H$_2$O), which ends at $T=2.4$ K (Ref. 10). Therefore, in contrast to (BEDT-TTF)$_2$ MHg(SCN)$_4$ with M=K, Tl, it seems that the ground state of (BEDO-TTF)$_2$ReO$_4$(H$_2$O) is not magnetically ordered for $T<2$–3 K, although a possible antiferromagnetic structure like a spin-density wave at higher temperatures (3 K$\leqslant T \leqslant$35 K) has been discussed in literature.[9–12]

The (BEDO-TTF)$_2$ReO$_4$(H$_2$O) single crystals studied in our experiments had typical dimensions of $2\times1\times0.05$ mm^3. The technique of measuring microwave absorption in small metallic samples was described in detail in our previous publications.[4–6] The underlying idea of the method is to compare the microwave power $P(H)$ absorbed at a fixed microwave frequency $\nu=$const when the cell in the experimental device is empty and when it contains a sample.

As in previous experiments,[4–6] radiation was detected by a small carbon bolometer placed close to a sample. Microwaves of the 30–120-GHz band were generated by a set of backward-wave oscillators with a frequency stability $\Delta\nu/\nu \sim 10^{-4}$ and were fed to the experimental cell via a waveguide.

The theoretical analysis[4–6] has demonstrated that the power absorption due to the sample is determined by the sample impedance Z_σ and the function $f(Z_c, Z_0)$ which describes the change in the reflectivity caused by introducing the sample into the cell:

$$\delta P \sim f(Z_c, Z_0)/\ \mathrm{Re}\{Z_\sigma\},$$

where Z_c and Z_0 are the impedances of the waveguide and empty cell, respectively. Since Z_0 depends on the bolometer resistance, which is a function of magnetic field, the absorption resonances due to the sample are superposed on a monotonic signal $P(H)$.

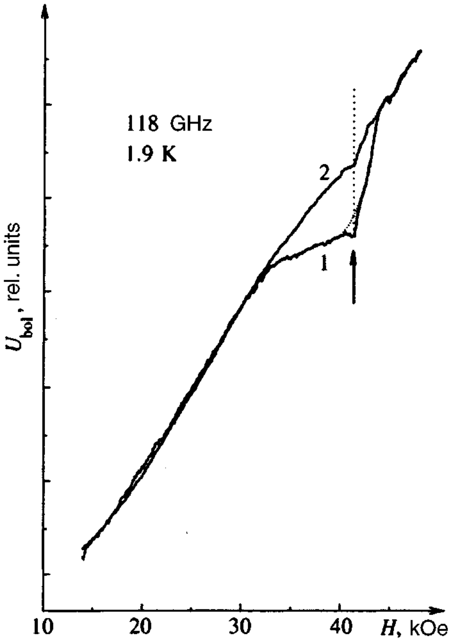


FIG. 1. Bolometer signal as a function of magnetic field: (*1*) experimental cell with the (BEDO-TTF)$_2$ReO$_4$(H$_2$O) sample; (*2*) empty cell. The arrow indicates the ESR peak corresponding to $g=2$.

At first glance, the factor $f(Z_c, Z_0)$ makes the analysis of the absorption line shape more difficult. The cell configuration, however, can be selected so that $f(Z_c, Z_0)$ is a weak function of the magnetic field, and we can take $f(Z_c, Z_0) \approx$const near a resonance if its line is sufficiently narrow. This was the case in our experimental device.

The experiments were performed in a magnetic field $H \leqslant 70$ kOe at a temperature $T=1.9$ K. The applied magnetic field and the Poynting vector of the microwave field were normal to the conducting plane in which two-dimensional current carriers moved. Assuming that at this temperature for $H=0$ the sample was in the superconducting state, we analyze in this paper the absorption for $H>10$ kOe, when the superconducting state was surely destroyed[9–12] and the sample was a normal metal. The circuit stabilizing the bolometer temperature[5] allowed us to record features in the absorbed microwave power $P(H)$ down to the level $\delta P(H)/P(0) \sim 5\times10^{-5}$. In order to rule out effects of sample overheating, we performed calibration experiments at several microwave powers.

Typical experimental curves of the signal detected by the bolometer as a function of magnetic field, $U_{\mathrm{bol}}(H)$, for the case of the empty cell and the cell with the sample are shown in Fig. 1 (curves *1* and *2*, respectively). They demonstrate that introduction of the sample into the cell leads to an additional broad absorption maximum connected with the organic metal [here we take into account that, owing to the negative slope of the bolometer characteristic, $P(H) \sim -U_{\mathrm{bol}}(H)$]. The difference between curves *2* and *1* yields a function proportional to $\delta P(H)$.

The shape of the $\delta P(H)$ curve in an organic metal is determined by the superposition of broad maxima[4–6] due to the CR and narrower resonances due to the spin degrees of freedom. The curves in Fig. 1 conform to this assumption. Specifically, a narrow ESR line of a smaller amplitude corresponding to the g-factor $g \approx 2$ (the arrow and dotted line in Fig. 1) can be seen superposed on a broad absorption maximum. Note that, unlike the case of (BEDT-TTF)$_2$MHg(SCN)$_4$,[4–6] no additional magnetic resonances have been detected in (BEDO-TTF)$_2$ReO$_4$(H$_2$O). This result is in agreement with the data[9–12] indicating that (BEDO-TTF)$_2$ReO$_4$(H$_2$O) is not magnetically ordered at liquid helium temperatures. In view of this, the question of magnetic spin-density-wave ordering in this material[9–12] seems controversial and calls for further investigation.

Note that, because of the thinness ($\sim 50\,\mu$m) of the (BEDO-TTF)$_2$ReO$_4$(H$_2$O) crystal (in earlier experiments 0.1–0.3 mm samples were used), not all the radiation was absorbed in the sample, and the fraction of microwave transmitted through the sample excited ESR in the bolometer (Fig. 1, curve *2*). Although the amplitude of this signal is lower than in the case of the cell with the metal (Fig. 1, curve *1*), the contributions due to the sample and experimental cell to the ESR are comparable, which makes the quantitative analysis of the absorption around the ESR difficult. Therefore this study is limited to the main broad maximum on the magnetoabsorption curve $P(H)$ (Fig. 1).

In order to test whether the maximum on the $P(H)$ curve is really due to the CR of two-dimensional carriers, we performed measurements in the configuration where the field was directed at an angle α with respect to the sample normal. In this case the absorption maximum (Fig. 1) was shifted to higher magnetic fields by a factor $\cos^{-1}\alpha$, which confirmed our interpretation of the absorption maximum as a CR peak due to two-dimensional carriers in (BEDO-TTF)$_2$ReO$_4$(H$_2$O), because in this case the effective component of the dc magnetic field is $H_{\perp} = H \cos \alpha$.

3. CYCLOTRON RESONANCE IN (BEDO-TTF)$_2$REO$_4$(H$_2$O)

The technique described in the previous section was used to distinguish the CR signal at different microwave frequencies. The result is given in Fig. 2. The absorption maximum was detected at frequencies $\nu > 80$ GHz. The experimental curves of $\delta P(H)$ can be approximated by Lorentzians,

$$\delta P(H) = \frac{P_{\max}(\Delta H)^2}{4(H - H_{\mathrm{res}})^2 + (\Delta H)^2}, \tag{1}$$

where H_{res} is the position of the absorption peak with the amplitude $P_{\max}$, ΔH is the FWHM of the resonant curve [open circles in Fig. 2 show experimental data, and the solid curves are best fits to Eq. (1)]. The parameters H_{res} and ΔH corresponding to the curves in Fig. 2 are plotted in Fig. 3 as functions of microwave frequency. The uncertainty of the nonlinear approximation procedure using Eq. (1) is about 5% for H_{res} and about 15% for ΔH. It is remarkable that the function $H_{\mathrm{res}}(\nu)$ is nonlinear, and the nonlinear function $\Delta H(\nu)$ has a minimum around $\nu \sim 100$ GHz (Fig. 3).

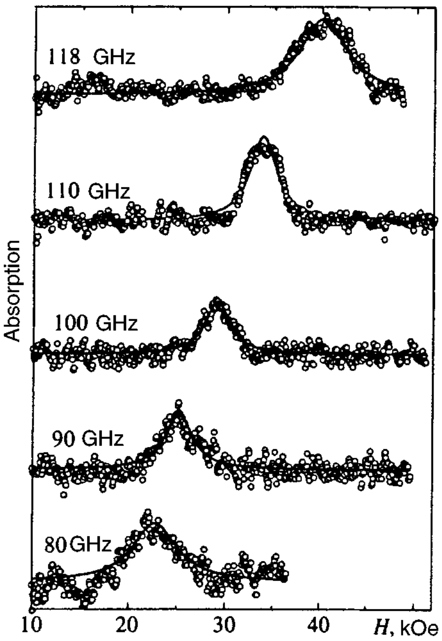


FIG. 2. Magnetoabsorption in (BEDO-TTF)$_2$ReO$_4$(H$_2$O) due to the cyclotron resonance at $T = 1.9$ K for various microwave frequencies. Experimental data are plotted by circles, and the best Lorentzian approximations by solid lines.

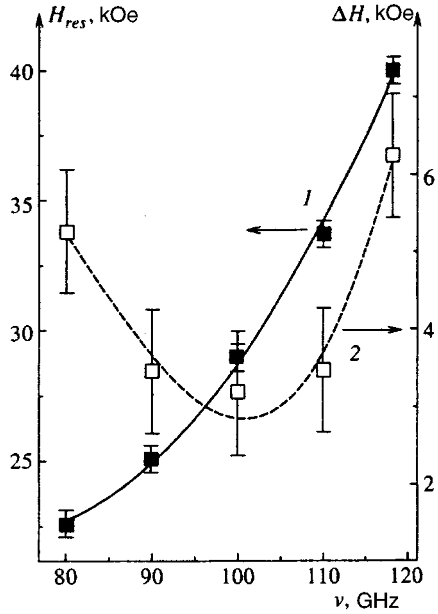


FIG. 3. (*1*) Peak positions and (*2*) FWHM of the CR line in (BEDO-TTF)$_2$ReO$_4$(H$_2$O) as functions of frequency.

In analyzing the CR in (BEDO-TTF)$_2$ReO$_4$(H$_2$O), one should bear in mind that the case of an organic conductor is essentially different from that of a normal metal.[4–6] Specifically, under the conditions for cyclotron resonance in a normal metal, carriers move in a highly nonuniform microwave field concentrated in the skin layer, so the magnetoabsorption curve $p(H)$ is a superposition of a set of peaks corresponding to multiple frequencies.[13] In an organic metal studied using this experimental geometry, in which the magnetic field H and the Poynting vector are directed perpendicular to the conducting plane, the microwave electric field also drops with increasing depth, but since the electron motion is two-dimensional, the microwave amplitude in each plane is constant, whereas the motion along the field amplitude gradient is practically nonexistent because the transport in the direction perpendicular to the layers is negligible. In this case[4–6] we have $Z_\sigma \propto 1/\sigma(\omega,H)$ and $\delta P(H) \propto \mathrm{Re}\{\sigma(\omega,H)\}$, where $\sigma(\omega,H)$ is the sample conductivity. Thus, the process of microwave absorption in an organic conductor resembles the case of a semiconductor, in which $\delta P(H) \propto \mathrm{Re}\{\sigma,(\omega H)\}$ also holds.[7]

Another essential feature of the process is that the Drude–Lorentz model for $\sigma(\omega,H)$, generally speaking, does not apply to the case of CR in a system with strong interaction, and the following function should be used:[7]

$$\sigma(\omega,H)=\frac{e^2n}{m}\,\frac{i[\omega-\omega_c+M'(\omega)]+M''(\omega)}{[\omega-\omega_c+M'(\omega)]^2+[M''(\omega)]^2}, \tag{2}$$

$$M'=\mathrm{Re}\{M(\omega)\},\quad M''=\mathrm{Im}\{M(\omega)\}, \tag{2a}$$

where $\omega_c=eH/mc$ is the cyclotron frequency, n is the carrier density, and $M(\omega)$ is the memory function, which takes into account interactions in the system.[7]

Assuming that $M'(\omega)$ and $M''(\omega)$ vary sllowly near the resonance, one can transform Eq. (2) to a quasi-Drude form[7] and obtain

$$\delta P \propto \mathrm{Re}\ \sigma(\omega,H)$$

$$=\frac{e^2n}{m}\,\frac{M''(\omega_c^*)}{[1+\partial M'(\omega_c^*)/\partial\omega]^2(\omega-\omega_c^*)^2+M^{*2}(\omega_c^*)^2}, \tag{3}$$

whereas the resonance condition takes the form

$$\omega=\omega_c^*=\frac{eH_{\mathrm{res}}(\omega)}{cm(\omega)}, \tag{4}$$

$$m(\omega)=m\left[1+\frac{M'(\omega)}{\omega}\right], \tag{5}$$

and

$$M''(\omega)=\frac{e}{2cm}\,\frac{\Delta H(\omega)\omega}{H_{\mathrm{res}}(\omega)}\,\frac{\partial H_{\mathrm{res}}(\omega)}{\partial\omega}. \tag{6}$$

Using Eqs. (4) and (6), we have derived $m(\nu)$ and $M''(\nu)$ from the experimental data plotted in Fig. 3 (see Fig. 4). In order to calculate $\partial H_{\mathrm{res}}/\partial\omega$, we approximated the experimental data by a polynomial (the solid line in Fig. 3), and in subsequent calculations used the derivative of the approximating function. In calculating $M''(\nu)$ we used the value $m=0.9m_0$.

Since the function $H_{\mathrm{res}}(\nu)$ is nonlinear, the effective mass depends on the frequency and, unlike the earlier results,[4–6] $m(\nu)$ increases with the frequency (Fig. 4), but rather then decreasing as in the case of (BEDT-TTF)$_2$MHg(SCN)$_4$. Note that the observed values of the effective mass range between $0.8m_0$ and $0.95m_0$, i.e., they are close to the value ($m=0.9m_0$ derived from Shubnikov–de Haas oscillations. But for $\nu\sim 80$ GHz the experimental curve of $m(\nu)$ shows a tendency to saturation at a level $m(\nu)<m$. Summarizing the results of the present work and other publications,[10–12] we therefore conclude that at microwave frequencies lower than 80 GHz the effective mass may decrease as a function of the frequency, i.e., in the case of (BEDO-TTF)$_2$ReO$_4$(H$_2$O) the function $m(\omega)$ may be nonmonotonic. It is also remarkable that, although the function $\Delta H(\nu)$ has a minimum (Fig. 3), the function $M''(\nu)=\mathrm{Im}\{M(\nu)\}$ increases monotonically with frequency and has no singularities in the frequency band studied.

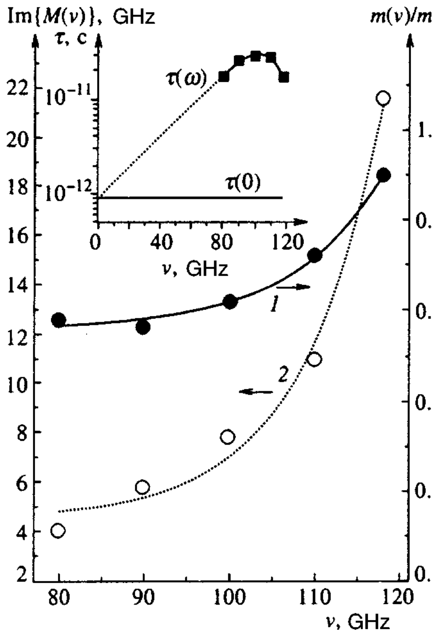


FIG. 4. (*1*) Effective mass and (*2*) imaginary part of the memory function as functions of the microwave frequency for (BEDO-TTF)$_2$ReO$_4$(H$_2$O). The insert shows the relaxation time $\tau(\nu)$ derived from the experimental data.

Our analysis indicates that Eqs. (2)–(6) can be applied to the cyclotron resonance in (BEDO-TTF)$_2$ReO$_4$(H$_2$O) at frequencies ranging between 80 and 120 GHz (Figs. 2–4). One can easily derive an expression for the relaxation time in the generalized Drude–Lorentz model with frequency-dependent effective mass and relaxation time:[5–8]

$$\tau(\nu)=\frac{[1+\partial M'(\omega)/\partial\omega]}{M''(\omega)}=\frac{H_{\mathrm{res}}}{\pi\nu\Delta H}. \tag{7}$$

It follows from Eq. (7) that, in the generalized Drude–Lorentz model, both the real and imaginary parts of $M(\omega)$ contribute to the dispersion of τ. As a result, even if $M'(\omega)$ and $M''(\omega)$ are smooth functions (Fig. 4), the curve of $\tau(\nu)$ may be nonmonotonic. The calculations according to Eq. (7) based on the data plotted in Fig. 3 are given in the insert to Fig. 4. The experimental values of $\tau(\nu)$ are about 10^{-11} s, and the maximum relaxation time $\tau\approx 2.9\times 10^{-11}$ s occurs at $\nu\sim 100$ GHz.

Now let us compare the results of this work to the values corresponding to $\omega=0$ derived from Shubnikov–de Haas oscillations.[10–12] Earlier we have noted the difference between the effective masses. As for the relaxation times, the Dingle temperature yields $\tau(0)=0.9\times 10^{-12}$ s.[10] Thus the difference between the relaxation times at low and high frequencies is also considerable (see the insert to Fig. 4), and in the case of the relaxation time, the renormalization is much larger than that of the effective mass [in the frequency range studied the maximum difference between $\tau(\omega)$ and $\tau(0)$ is a factor of 30]. Hence we assume that Fermi-liquid effects in $(\text{BEDO-TTF})_2\text{ReO}_4(\text{H}_2\text{O})$ are quite strong, and the relaxation time is a strong function of frequency in the frequency range below 80 GHz [the suggested shape of $\tau(\omega)$ is shown by the dashed line in the insert to Fig. 4]. Unfortunately, experimental investigations of CR in this band are rather difficult, and further investigation beyond the scope of the reported study is required.

4. CONCLUSION

Measurements of the resonant microwave absorption in $(\text{BEDO-TTF})_2\text{ReO}_4(\text{H}_2\text{O})$ in the 80–120-GHz band have allowed us to distinguish the contribution due to the CR of two-dimensional carriers. This is the first time when not only the peak positions, but also CR line widths have been measured. We have found out that the effective mass increases with the frequency from $m \approx 0.8m_0$ at $\nu \sim 80$ GHz to $m \approx 0.95m_0$ at $\nu \sim 120$ GHz. Alongside the effective mass dispersion, we have detected a strong renormalization of the relaxation time: at $\nu \sim 100$ GHz the ratio $\tau(\nu)/\tau(0)$ is more than 30. It seems that the most adequate description of CR in $(\text{BEDO-TTF})_2\text{ReO}_4(\text{H}_2\text{O})$ can be achieved using the magnetoconductance model[7] based on the memory function $M(\omega)$ whose real and imaginary parts are rapidly varying functions of the microwave frequency.

We are deeply indebted to B. A. Volkov, L. A. Falkovskii, and S. V. Sharov for helpful discussions. This work was part of the INTAS 93-2400 and INTAS 94-1788 international projects, and was supported by the *Fullerenes and Atomic Clusters* and *Microwave Physics* programs sponsored by the Ministry of Science of Russia, and also by the Royal Society (Great Britain).

[1] J. Singleton, F. L. Pratt, M. Doporto, T. J. B. M. Janssen, M. Kurmoo, J. A. A. J. Perenboom, W. Hayes, and P. Day, Phys. Rev. Lett. **68**, 2500 (1992).

[2] S. Hill, A. Wittlin, J. van Bentum, J. Singleton, W. Hayes, J. A. A. J. Perenboom, M. Kurmoo, and P. Day, Synth. Met. **70**, 821 (1995).

[3] S. V. Demishev, N. E. Sluchanko, A. V. Semeno, and N. A. Samarin, JETP Lett. **61**, 313 (1995).

[4] S. V. Demishev, A. V. Semeno, N. E. Sluchanko, N. A. Samarin, I. B. Voskoboĭnikov, V. V. Glushkov, A. E. Kovalev, and N. D. Kushch, JETP Lett. **62**, 228 (1995).

[5] S. V. Demishev, A. V. Semeno, N. E. Sluchanko, N. A. Samarin, I. B. Voskoboinikov, V. V. Glushkov, J. Singleton, S. J. Blundell, S. O. Hill, W. Hayes, M. V. Kartsovnik, A. E. Kovalev, M. Kurmoo, P. Day, and N. D. Kushch, Phys. Rev. B **53**, 12794 (1996).

[6] S. V. Demishev, N. E. Sluchanko, and A. V. Semeno, in *Microwave Physics* [in Russian], ed. by E. V. Suvorov, Nizhniĭ Novgorod, (1996), p. 223.

[7] J. Hajdu and G. Landwehr, in *Strong and Ultrastrong Magnetic Fields and Their Applications*, ed. by F. Herlach, Topics in Applied Physics, Vol. 57, Springer-Verlag, Berlin (1985), p. 120.

[8] L. Degiorgi, H. R. Ott, and F. Hulliger, Phys. Rev. B **211**, 220 (1995).

[9] L. I. Buravov, A. G. Khomenko, N. D. Kushch, V. N. Laukhin, A. I. Shchegolev, E. B. Yagubskii, L. P. Rosenberg, and R. P. Shibaeva, J. de Phys. I **2**, 529 (1992).

[10] S. Kahlich, D. Schweitzer, C. Rovira, J. A. Paradis, M.-H. Whangbo, I. Heinen, H. J. Keller, B. Nuber, P. Bele, H. Brunner, and R. P. Shibaeva, Z. Phys. B **94**, 39 (1994).

[11] D. Schweitzer, E. Balthes, S. Kahlich, I. Heinen, H. J. Keller, W. Strunz, W. Biberacher, A. G. M. Jansen, and E. Steep, Synth. Metals **70**, 857 (1995).

[12] S. Kahlich, W. Sommer, D. Schweitzer, I. Heinen, H. J. Keller, C. Rovira, J. A. Paradis, and M.-H. Whangbo, Synth. Metals **70**, 865 (1995).

[13] I. M. Lifshits, M. Ya. Azbel', and M. I. Kaganov, *Electron Theory of Metals*, Consultants Bureau, New York, (1973).

Translation provided by the Russian Editorial office.

Dynamic properties of an incommensurate charge density wave in monoclinic TaS_3 at low temperatures

F. Ya. Nad'

Institute of Radio-Engineering and Electronics, Russian Academy of Sciences, 103907 Moscow, Russia
Centre de Recherches sur les Très Basses Températures, Centre National de Recherches Scientifiques, 38042 Grenoble, Cédex 9, France

P. Monceau

Centre de Recherches sur les Très Basses Températures, Centre National de Recherches Scientifiques, 38042 Grenoble, Cédex 9, France

(Submitted 6 June 1996)
Zh. Éksp. Teor. Fiz. **111**, 988–1000 (March 1997)

The permittivity of monoclinic TaS_3, a quasi-one-dimensional conductor with an incommensurate charge density wave (CDW), as a function of frequency and temperature has been studied. At low temperatures and at frequencies below 1 MHz, the temperature dependence of the real part of the permittivity shows a maximum shift to lower temperatures with decreasing frequency. The temperature dependence of the relaxation time consists of two branches corresponding to macroscopic regions of CDWs with long and short relaxation times τ on the microscopic scale. With decreasing temperature, the growth of τ for large CDW regions is faster than thermal activation and shows a tendency to diverge at a finite temperature while the growth of τ due to the relaxation on the microscopic scale is slower than the activation rate. Our results show that with decreasing temperature the m-TaS_3 quasi-one-dimensional conductor goes over to a glasslike state due to the strong pinning of CDWs by randomly distributed impurities and the formation of mutually interacting solitary CDW collective excitations. © *1997 American Institute of Physics.* [S1063-7761(97)01903-3]

1. INTRODUCTION

Recently several papers have been published[1–4] stating the evidence that a maximum (singularity) exists on the temperature dependence of the real part $\varepsilon'(T)$ of the dielectric function of quasi-one-dimensional conductors measured at low frequencies between 10^{-2} and 10^7 Hz in the low temperature range. A model ascribing this maximum to an increase of the coherence length and relaxation times of CDWs in this temperature range and a transition to a ''frozen'' glasslike state has been proposed.[3,4] These effects have been found in two materials with CDWs, namely in the blue bronze $K_{0.3}MoO_3$[1,2] and orthorhombic TaS_3,[3,4] where the CDW of one type is incommensurate with the initial crystal lattice near the Peierls transition temperature T_P.

The CDW wave vector in these materials is temperature-dependent, and its value approaches four-fold commensurability with decreasing temperature.[5,6] In this connection, the question arises as to whether the observed features of $\varepsilon'(T)$ are typical of only such quasi-one-dimensional conductors where, in particular, the maximum of $\varepsilon'(T)$ can be attributed to a transition to the commensurate CDW. In order to test whether the divergence in $\varepsilon'(T)$ at low temperatures is a fundamental phenomenon deriving from basic physical principles, we have measured at low temperatures the permittivity of monoclinic TaS_3, a quasi-one-dimensional conductor in which two CDWs occur with wave vectors $\mathbf{q}_1$ and $\mathbf{q}_2$, independent of temperature and never approaching commensurability as the temperature decreases.[5]

2. EXPERIMENTAL RESULTS

We have investigated four samples of monoclinic TaS_3 with cross sections of 10^{-7}–10^{-6} cm^2 and a length of 1–2 mm. The samples were placed on a sapphire substrate. Their electrical contacts were made of silver paste or vacuum deposition of indium strips on the samples after thermal treating.[7] The contacts fabricated in vacuum had lower resistances, and were stable during thermocycling. In this paper, we present data for the $m4$ sample, which had dimensions $1200\times10\times2.5\ \mu$m, a room-temperature resistance of 144 Ω, and contacts fabricated in vacuum. Similar results have been obtained with other samples.

We have measured the dc conductivity G_{dc} and current–voltage curves (IVC) of m-TaS_3 in the temperature range from 30 to 300 K. The real part Re G and the imaginary part Im G of the ac conductivity at frequencies between 5×10^2 and 10^6 Hz were measured over a wide temperature range between room and liquid helium temperature. The measurements of the conductivity were made using an HP 4192A vector impedance analyzer under computer control. At low temperatures, when the sample resistance was greater than 1 MΩ, we used a special preamplifier in order to enhance the sensitivity at frequencies below 10^4 Hz.[4] The ac voltage applied to the sample was 5–10 mV, which corresponds to less than one-tenth of the minimum threshold field E_T throughout the studied temperature range. As follows from our data and other publications,[4,8,9] the conductivity was independent of such field amplitudes, i.e., our experiments were performed in the linear regime, and we measured the conductivity of the CDW ground state.

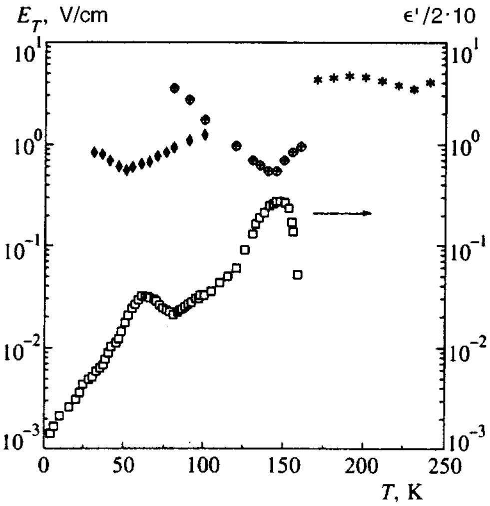


FIG. 1. Temperature dependence of threshold field and permittivity: stars—threshold field E_T for the first CDW; circles and diamonds—threshold fields E_T and E_T' for the second CDW; squares—real part of the permittivity $\varepsilon'/2\times10^8$ at a frequency of 10 kHz.

It follows from our measurements that the temperature dependence of the dc conductivity G_{dc} in our samples has a form typical of m-TaS$_3$ with two sharp drops in the conductivity at $T_1=240$ K and $T_2=160$ K, corresponding to formation of two CDWs with wave vectors $\mathbf{q}_1=(0a^*,\ 0.253b^*,\ 0c^*)$ for $T<T_1$ and $\mathbf{q}_2=((1/2)a^*,\ 0.247b^*,\ (1/2)c^*)$ for $T<T_2$.[10] In the temperature range between 100 and 160 K, the function $\log G_{\mathrm{dc}}(1/T)$ is almost linear and characterized by the activation energy $\Delta=900$ K.[11,12] The IVCs of m-TaS$_3$ samples had forms typical of quasi-one-dimensional conductors with CDWs.[13] At temperatures above the first Peierls transition T_1, the IVCs of our samples were linear in electric fields of up to about 5 V/cm, with subsequent smooth growth of the conductivity proportional to the second power of the current due to a heating effect. At temperatures below T_1, a large nonlinearity develops on the IVC at values as low as $\simeq 3$ V/cm. The pulsed measurements eliminated sample heating, showing that the nonlinearity is not due to the thermal effects, but to the beginning of CDW motion as a whole.[13] Figure 1 shows the threshold field E_T as a function of temperature. In the range of the first CDW (160–240 K), E_T has a minimum at $T=200$ K, then grows with decreasing temperature and remains relatively high, which is typical of m-TaS$_3$.[13] It is known that after formation of the second CDW ($T<160$ K), the bulk of the nonlinear conductivity is associated with it. The threshold field decreases in several times, although its temperature dependence is qualitatively similar: it drops after the transition through T_2, has a minimum, and at lower temperatures increases exponentially as $E_T(T)=E_T(0)\times\exp(-T/T_0)$, where $T_0=25$ K.[11]

In the temperature range between 100 and 160 K, the threshold field E_T is clearly defined because the CDW at $E<E_T$ is immobile, and the IVC is linear, whereas at $E>E_T$ the CDW is driven by the electric field, the IVC is quite nonlinear, which is seen most clearly in m-TaS$_3$, and narrow-band generation begins.[12] But at temperatures below 100 K, simultaneously with the continuing growth of E_T (Fig. 1), a small nonlinear contribution emerges on the IVC linear section beginning with the threshold field E_T' below E_T.[11,14] Figure 1 shows E_T' as a function of temperature. One can see that below 100 K the threshold voltage of this nonlinearity decreases with decreasing temperature, approaches a minimum, its value at the minimum being close to the minimum E_T at $T=149$ K, and after that shows a nonlinear growth.

Simultaneously with G_{dc} and IVC, we measured the complex ac conductivity $G(\omega)$ at fixed temperatures between 4.2 and 300 K. From these data we derived the permittivity using standard relations: the real part of the permittivity is $\varepsilon'(\omega)=\mathrm{Im}\sigma(\omega)/\omega$, and the imaginary part is $\varepsilon''(\omega)=[\,\mathrm{Re}\sigma(\omega)-\sigma_{\mathrm{dc}}]/\omega$, where σ denotes conductivity. Figure 1 shows a plot of $\varepsilon'(T)$ measured at the frequency $f=\omega/2\pi=10$ kHz in the temperature range of 4.2 to 160 K. One can see that it has two clearly defined maxima. One of these is near the Peierls transition temperature T_2, where the second CDW emerges. In the temperature range below this maximum, ε' drops monotonically and, according to previously reported measurements at frequencies $f>10$ MHz, this drop should extend to very low temperatures.[13] Our measurements at low frequencies, however, show that for $f<1$ MHz a clear maximum can be seen at $T=80$ K superposed on this monotonic decrease in $\varepsilon'(T)$. Comparing the curves $E_T(T)$, $E_T'(T)$, and $\varepsilon'(T)$ in Fig. 1, one can see a certain correlation among them. The maxima of $\varepsilon'(T)$ occur near the temperatures where E_T and E_T' have minima. Similar effects were observed in other quasi-one-dimensional conductors.[15] The data plotted in Fig. 1 were taken from the m-TaS$_3$ sample with a relatively low room-temperature resistance ($R_0=144\,\Omega$); therefore our device could not accurately measure Im G near T_1 at such low frequencies because of inadequate resolution. Measurements of the $m3$ sample with $R_0\sim10^3\Omega$ at a frequency of 1 MHz, however, have demonstrated that there is also a maximum near $T_1=240$ K on the $\varepsilon'(T)$ curve recorded at a frequency of 1 MHz.

Figure 2 shows the $\varepsilon'(T)$ curves near T_2 measured at several frequencies between 1 kHz and 1 MHz. As can be seen from Fig. 2, the amplitudes of these peaks increase as the frequency decreases, but their position remains constant to within the experimental accuracy. This result is in qualitative agreement with the behavior of the $\varepsilon'(T)$ curve near the Peierls transition in orthorhombic TaS$_3$,[16] NbSe$_3$,[17] and (NbSe$_4$)$_{10/3}$I.[18] Figure 3 shows the $\varepsilon'(T)$ curves near the low-temperature maximum at frequencies between 700 Hz and 1 MHz. The growth of the peak with decreasing frequency is considerably larger than near T_2. But it is more important that the peak position depends strongly on the frequency, in contrast to the region near T_2. The location of the peak shifts to lower temperatures as the frequency decreases. That is, in this quasi-one-dimensional conductor, as in other materials of this kind, such as K$_{0.3}$MoO$_3$ (Ref. 12) and orthorhombic TaS$_3$,[3,4] the function $\varepsilon'(T)$ shows a tendency to diverge with a specific frequency dependence, which is typi-

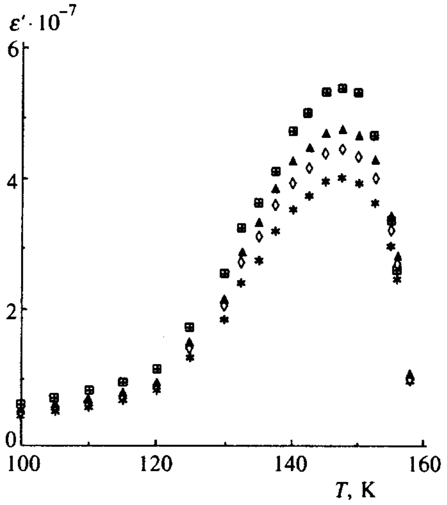


FIG. 2. Real part ε' of the permittivity versus temperature around the second Peierls transition measured at different frequencies: squares—10; triangles—100; diamonds—300; stars—1000 kHz.

cal of many disordered systems which go over to a glasslike state at low temperatures.[19–21]

The real part of the conductivity, Re G, measured at fixed temperatures as a function of frequency (Fig. 4) is similar to those measured in other quasi-one-dimensional conductors with CDWs.[2,4] In the frequency range under investigation, the low-frequency region is achieved at temperatures above 50 K in which Re G becomes independent of the frequency, i.e., it equals G_{dc}. But at lower temperatures, Re G varies with the frequency in this region, and it should approach G_{dc} at even lower frequencies.[4] At low temperatures between 4.2 and 40 K, the observed behavior of $G(\omega)$ is well approximated by the function $G(\omega) \propto \omega^{\alpha}$, where $\alpha \sim 0.8$, which is typical of low-temperature states of quasi-one-dimensional conductors,[4,22] and in the more general case, of many systems with a certain degree of disorder in which the hopping conductivity dominates.[23]

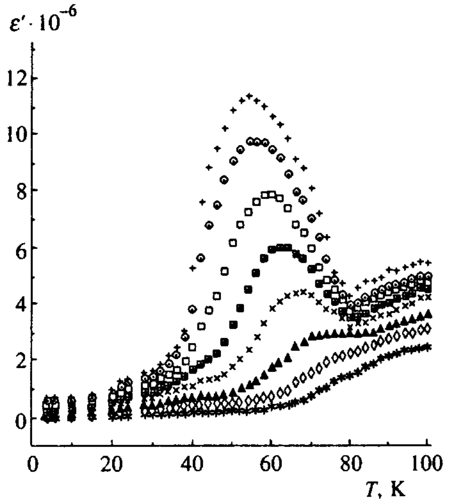


FIG. 3. Real part of the dielectric permittivity ε' versus temperature in the low-temperature range measured at different frequencies: crosses—0.7; circles—1; squares—3; squares with crosses—10; diagonal crosses—30; triangles—100; diamonds—300; stars—1000 kHz.

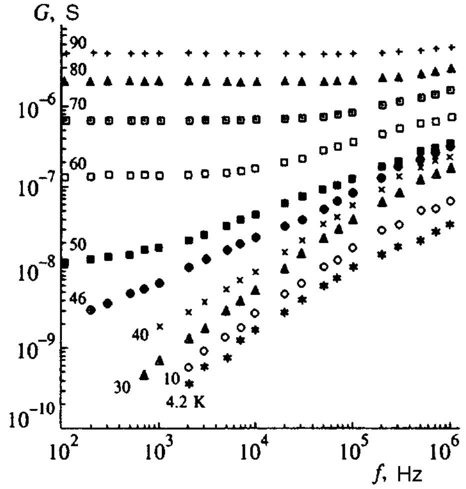


FIG. 4. Real part G of the conductivity versus frequency measured at the different temperatures indicated at the curves.

The parameter G_{dc} is necessary for calculating the dissipative part of the dielectric permittivity ε'' using the relation given above. If the transition of $G(\omega)$ to the constant value G is observed within the measured frequency range, the maximum on the $\varepsilon''(\omega)$ curve should also be located within this frequency range. This can be seen from Fig. 5, which shows the $\varepsilon''(\omega)$ curves measured at fixed temperatures between 50 and 100 K. For temperatures $T<50$ K, the maxima on the $\varepsilon''(\omega)$ curves are in the frequency range below 10^3 Hz, and we can see only the tails of these curves. For $T>50$ K, most of the $\varepsilon''(\omega)$ curve (which is also called the loss function) are in the measured frequency band. The maximum on the $\varepsilon''(\omega)$ curve shifts to higher frequencies with increasing temperature, and at $T \simeq 100$ K it approaches the edge of the frequency range in question.

Figure 5 shows the $\varepsilon''(\omega)$ curves at different temperatures and, as will be shown below, the values of $\varepsilon'(\omega)$ are considerably different from theoretical predictions obtained

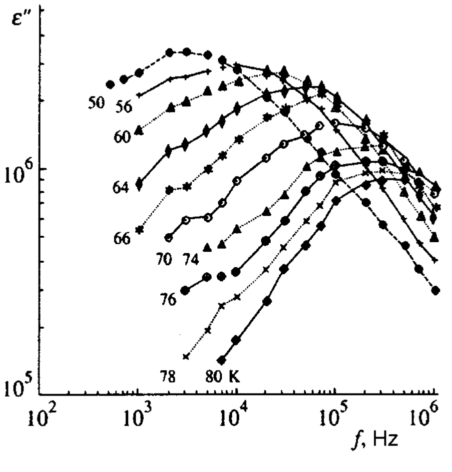


FIG. 5. Imaginary part of the dielectric permittivity ε'' versus frequency measured at the different temperatures indicated at the curves.

under the assumption of simple Debye relaxation with a single relaxation time τ_0.[24] The $\varepsilon''(\omega)$ and $\varepsilon'(\omega)$ curves obtained at low temperatures are similar to those observed in a wide range of disordered materials, including various glasses.[24] In these materials the relaxation of excitations is found to be nonexponential with an wide distribution of relaxation times. This relaxation is described by various phenomenological equations, which were generalized in the theory of dielectric relaxation in polymers:[25]

$$\varepsilon(\omega)=\varepsilon_{\mathrm{HF}}+\frac{\varepsilon_0-\varepsilon_{\mathrm{HF}}}{[1+(i\omega\tau)^{1-\alpha}]^{\beta}}, \tag{1}$$

where $\varepsilon_{\mathrm{HF}}$ corresponds to the high-frequency limit $\varepsilon'(\omega)$ at $\omega \gg \omega_0$ ($\omega_0 \equiv \tau_0^{-1}$), ε_0 is the low-frequency limit of $\varepsilon'(\omega)$ at $\omega \ll \omega_0$, and α and β characterize the width and asymmetry of the relaxation time distribution. According to Eq. (1), the relationship between ε'' and ε' is similar to the familiar Cole–Cole relation described by an asymmetrical arc of a circle.[24] The point where the high-frequency part of the $\varepsilon''(\varepsilon')$ curve crosses the $\varepsilon''=0$ axis defines the high-frequency parameter $\varepsilon_{\mathrm{HF}}$, and the intersection with the low-frequency part defines ε_0. Our plots of the function ε'' (ε') have demonstrated that $\varepsilon_{\mathrm{HF}}$ can be determined using measurements in this frequency range with sufficient accuracy. By determining this parameter at different temperatures and subtracting it from measurements of $\varepsilon'(\omega)$, we can obtain the low-frequency component of the dielectric response more accurately because the dielectric response, as is well known, can include contributions from different subsystems, such as the lattice, free electrons, etc. Figure 6 shows a plot of $\varepsilon' - \varepsilon'_{\mathrm{HF}}$ as a function of frequency (in Fig. 3 and hereinafter this difference is denoted as ε'). Throughout the temperature range studied, the measured $\varepsilon'_{\mathrm{HF}}$ was significantly smaller than ε'.[4] As can be seen from Fig. 6, log–log plots of $\varepsilon'(\omega)$ at different temperatures are qualitatively similar and can be brought to coincidence by shifting them along the log ω and log ε' axes.[4,24] In the temperature range between 4.2 and 50 K, these curves shift upwards with the temperature, and in the range between 50 and 100 K they move downwards. These curves can be roughly approximated by two parts with different slopes: a smaller slope at lower frequencies, $\varepsilon'(\omega) \propto \omega^{-m}$, where m increases from 0.1 to 0.25 as the temperature decreases, and a larger slope at higher frequencies, $\varepsilon'(\omega) \propto \omega^{-n}$, where n changes between 0.5 and 0.7. The region of the crossover between these two parts shifts to higher frequencies with increasing temperature. Figure 5 shows that the curves of $\varepsilon''(\omega)$ are also similar to each other in the log–log plots and can be brought to coincidence by shifting them along both axes.[4,24] Moreover, Figs. 5 and 6 indicate that the high-frequency parts of the $\varepsilon'(\omega)$ and $\varepsilon''(\omega)$ curves are similar and can be described by the formula

$$\varepsilon' \sim \varepsilon'' \propto \omega^{-n}.$$

3. DISCUSSION

Measurement of the permittivity as a function of temperature and frequency is a widely used technique in studies of fundamental properties of condensed matter. This method is equally applicable to electronic condensed states such as a CDW in quasi-one-dimensional conductors.

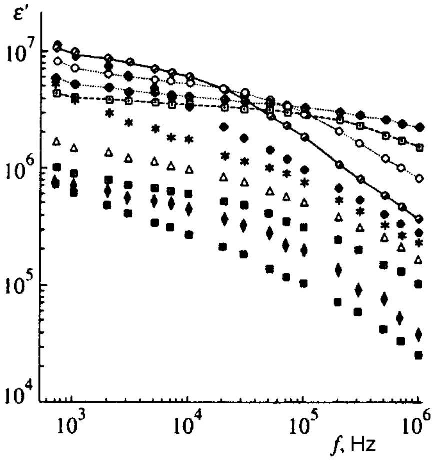


FIG. 6. Real part of the permittivity ε' as a function of frequency at different temperatures: filled squares—4; diamonds—10; squares with crosses—20; triangles—30; stars—40; filled circles—50; circles with slash—60; empty circles—70; squares with dots—80; circles with crosses—90 K.

The results described above, in particular those plotted in Fig. 1, demonstrate that the temperature dependence of the conductivity, threshold field E_T, and permittivity are interrelated. The decreases in the conductivity near T_1 and T_2, where the appropriate CDWs develope, correspond to decreases in the threshold field and, as follows from our measurements, maxima on the curves of $\varepsilon'(T)$. The amplitudes of the peaks on the curves of the permittivity $\varepsilon'(T)$ depend weakly on the frequency, but their positions on the temperature axis do not change with the frequency. This may be caused by the small relaxation time τ_0 of CDWs ($\simeq 10^{-10}$ s) near the Peierls transitions,[15] i.e., the response to the ac electric field with a frequency $>10^6$ Hz weakly depends on the frequency. The temperatures below T_1, which correspond to the $\varepsilon'(T)$ maximum and $E_T(T)$ minimum, are close to each other. However, ε' in this range is a steeper function of temperature than E_T, and as a result, the product $\varepsilon' E_T$ depends on the temperature. In the temperature range below T_1 (80–120 K), the correlation between the increase in $E_T(T)$ and decrease in $\varepsilon'(T)$ is stronger, and their product changes little. It is known that the product $\varepsilon' E_T$ is also a weak function of temperature below the Peierls transition temperature in other quasi-one-dimensional conductors.[15] The interpretation of this fact is based on the assumption that the force needed to shift the CDW as a whole, i.e., to overcome the threshold, and the field needed to polarize the CDW are proportional to one another, which is in agreement with the concept of an elastic CDW with a fairly large coherence length L_0, whose motion is blocked by a collective action of impurities within its coherence length. This case corresponds to pinning of weak CDWs and has been treated in some theoretical papers.[26,27] This means that an elastic, weakly pinned CDW is a typical state in the region near the Peierls transition.

The number of electron–hole pairs excited across the Peierls gap Δ drops exponentially with decreasing temperature, and the screening length increases accordingly. This leads to an increase in the local interaction between the CDW and strong pinning centers, hence a gradual transition from the weak, collective pinning to strong pinning.[3,4,28,29] In the case of strong pinning, the coherence of CDWs extends only over a distance $L_i \ll L_0$ between impurity centers. This is probably the reason for the drop in ε' with decreasing temperature in m-TaS$_3$ and other quasi-one-dimensional conductors.[3] Additional evidence in favor of the increase in pinning force and decrease in coherence length at low temperatures is the exponential growth of the threshold field E_T (Fig. 1), the decrease in the nonlinear contribution to the conductivity, etc.[13]

In our m-TaS$_3$ samples below 80 K, the CDW interacts, in effect, locally with simple impurity centres. The onset of its coherent motion as a whole demands high electric fields. In a weak field it is practically immobile, and its response to a weak ac field can probably be accounted for in terms of deformations of local CDW regions localized near impurities. As follows from some theoretical[29–31] and experimental[3,4,14] studies, it is most probable that deformed regions of a CDW near impurities correspond either to topological solitons with phase shifts of $\pm 2\pi$ and charges of $\pm 2e$, or dipole solitons with zero charge. It seems that most of the change in the CDW phase takes place over a length of about ten lattice constants near an impurity. Soliton tails, however, may extend over considerably longer distances. Such collective excitations, involving several electrons and similar to polarons, have effective masses larger than the free electron mass.[29] In our opinion, these collective CDW excitations determine the kinetic properties of the CDW ground state at low temperatures. The dc conductivity G_{dc} is controlled by the hopping of such collective excitations among randomly distributed impurities. In accordance with the hopping conductivity mechanism, its temperature dependence is described as follows:[11,14]

$$G_{\mathrm{dc}} \propto \exp(-T/T_0)^{1/2}.$$

As can be seen from Fig. 4, in m-TaS$_3$, as in other quasi-one-dimensional conductors, we have $G(\omega) \propto \omega^{\alpha}$, where $\alpha \sim 0.8$, which is typical of the hopping conductivity.[4,14] Measurements of the optical absorption due to states within the Peierls gap also indicate the presence of collective excitations resulting in localized electron states.[32,33]

It is natural to assume that the response of a pinned CDW to a weak electric field, described in terms of the permittivity, is also determined mainly by these collective excitations.[3,4] The number of collective excitations at low temperatures should probably be determined largely by the concentration of pinning centers for CDWs and be a flat function of the temperature. But because the CDW becomes more rigid and cooperative interaction among collective excitations increases with decreasing temperature, there is a tendency to ordering in their distributions and to extension of regions with a certain alignment of microdipoles, i.e., to an enhancement of coherence regions of collective excitations.[3,34] The permittivity should also increase, which is observed in m-TaS$_3$ (Figs. 1 and 3) and in other similar materials. Figure 1 shows that, as near the Peierls transition temperature $T_2 = 160$ K in this region of temperatures ($\simeq 50$ K), there is a local minimum of the threshold field E_T, which also indicates the possibility of extension of regions with coherent states of collective excitations, because their motion leads to the initial, weak departure of IVC from a linear curve.[11,12,14] In this temperature range, the change in ε' is larger than in $E_T'(T)$, and the product $\varepsilon' E_T'$ is no longer constant, and it is considerably smaller than near T_2 (at approximately the same minimum values of E_T and E_T'). The amplitudes of the peaks on the $\varepsilon'(T)$ curve and their positions are steeper functions of the frequency than near T_2 (Fig. 2). This difference indicates that the physical causes for the maxima on $\varepsilon'(T)$ curves in the low-temperature range may be different. This type of ground state and its behavior at a variable temperature resemble the ground state and domain growth near transitions to a glassy state in many materials with a certain degree of disorder, such as spin glasses, orientational glasses, etc.[19–21]

Simultaneously with the growth of the size L of the coherence region, their relaxation time also increases as $\tau \propto \exp\ L$.[35] We measured ε' at low but finite frequencies; for $2\pi f > \tau^{-1}$, the CDW did not have enough time to respond to the field at this frequency, so that delay effects and a decrease in ε' were detected. This dynamic effect, also named as ''slowing-down behavior,'' is responsible for the maximum on the $\varepsilon'(T)$ curve. At temperatures below the maximum of $\varepsilon'(T)$, the relaxation time is so large that the CDW is frozen in a glasslike state. Assuming that the maximum point satisfies the condition $\tau = 1/2\pi f$, we can derive from the function $\varepsilon'(T,\omega)$ the characteristic relaxation time τ^* of a CDW at low temperatures.[3,4]

Figure 7 shows the characteristic relaxation time τ^* as a function of the reciprocal temperature, which, probably, characterizes relaxation of the largest and slowest CDW regions, i.e., it describes so-called α-relaxation.[4,24] The curve of $\log \tau^* = \tau_0(1/T)$ plotted in Fig. 7 cannot be described by a pure activation mechanism (which suggests a linear function of $1/T$), but curves downward, i.e., demonstrates a tendency to divergence. According to the dynamic scaling hypothesis, which provides an adequate description of many disordered materials near a transition to a glassy state,[19–21] such a dependence should be described by the equation

$$\tau^* = \tau_0 (1 - T_c/T)^{-z\nu}.$$

Our experimental data can be approximated by this equation with $\tau_0 \sim 10^{-11}$ s, $T_c = 36$ K, and $z\nu = 13$ (Fig. 7).

Figure 7 also shows the dependence $\tau_p(T) = (\omega_p)^{-1}$ derived from the frequency ω_p corresponding to the maximum of the dissipative part of the permittivity versus frequency, $\varepsilon''(\omega)$ (Fig. 5). In the temperature range above 70 K the curves of $\tau^*(1/T)$ and $\tau_p(1/T)$ are close, which should be true when the process is characterized by a single relaxation time, or the distribution of relaxation times is narrow. In this temperature range, both curves tend asymptotically to an activation curve defined by $\tau \sim \tau_0 \exp(-\Delta/T)$ with $\Delta \simeq 900$ K, which corresponds to the gap in m-TaS$_3$ obtained from measurements of its conductivity versus temperature. This, apparently, means that in this temperature range the CDW relaxation is largely determined by single electron–hole pairs excited across the Peierls gap. One of these relaxation

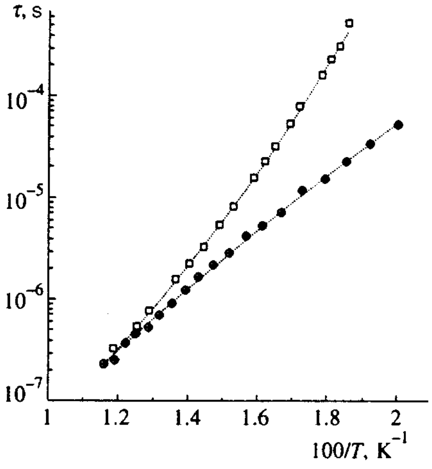


FIG. 7. Characteristic relaxation time as a function of temperature. The parameter τ_p (circles with crosses) is derived from the peak positions on the $\varepsilon''(\omega,T)$ curves. The time τ^* (squares) is derived from the positions of the peaks on the $\varepsilon'(\omega,T)$ curves.

mechanisms of local CDW deformation can be a thermally activated phase slip with dephasing between the two parts of the wave in a small region which go over to the normal state for a short time and through which the energy is dissipated.[36]

As can be seen from Fig. 7, the $\tau^*(1/T)$ and $\tau_p(1/T)$ curves differ at temperatures below $\simeq$70 K. The $\tau^*(1/T)$ curve deviates upward from the thermal activation line and, as was mentioned above, resembles the typical divergence of $\varepsilon'(T)$ near the transition to a glasslike state with a transition temperature $T_c \simeq 36$ K. This divergence corresponds to the case of barrier heights increasing with decreasing temperature, which is in accordance with the above model suggesting growth in the regions of uniform polarization of collective excitations in the CDW with decreasing temperature and the consequent increase in the height of the energy barrier to be overcome in order to change polarizations of these regions.[4,29] The curve of $\tau_p(1/T)$ deviates downward from the activation line, which corresponds to the case of energy barriers of a constant height. This means that for $\tau_p(1/T)$ the effective barrier height decreases with decreasing temperature. Probably, this shape corresponds to the case when the relaxation is due to tunneling of solitons (collective excitations) between metastable states of CDWs near pinning centers. These transitions can be the basis of a macroscopic mechanism, which determines both the hopping conductivity and the CDW polarization, hence ε', at low temperatures. According to theory, if a correlation between separate tunneling events exists, the $\tau_p(1/T)$ curve should be described by an equation similar to the one describing the upper branch:

$$\tau_p = \tau_p^0 (1 - T_c/T)^{-z\nu},$$

but with a negative T_c.[37] The experimental curve of $\tau_p(1/T)$ can in fact be approximated by this equation with $\tau_0 \sim 10^{-11}$ s, $T_c = -40$ K, and $z\nu = 24$. This seems to indicate that there can also be a correlated motion of collective excitations on this branch of the $\tau(1/T)$ curve, similarly to the case of orthorhombic TaS$_3$.[4] The existence of two branches of the temperature dependence of the relaxation time deviating upwards and downwards from the activation curve is a sign of broadening of the relaxation time distribution at low temperatures and, accordingly, of the energy distributions of barrier heights to be overcome in the process of relaxation. It is also possible, naturally, to determine an average relaxation time, whose temperature dependence would be close to that described by the Arrhenius law and would lie between the two branches shown in Fig. 7.

Thus, as can be seen from the above data, in monoclinic TaS$_3$, as in other quasi-one-dimensional conductors, several effects are observed with decreasing temperature which are characteristic of a transition to a glasslike state. The experimental data are satisfactorily described in terms of the dynamic scaling hypothesis[37] and the approximate model of the ground state in the form of a CDW and its collective soliton-like excitations at low temperatures.[3,4,29,30,38] As was noted above, there are two CDWs in m-TaS$_3$ and, according to earlier results,[5] their wave vectors are independent of temperature and do not approach values corresponding to a commensurate CDW. From this it follows that the transition to a glasslike state described in the paper is directly related neither to the presence nor absence of a temperature dependence of the wave vector, nor to a transition of the CDW to a commensurate state. The transition of a quasi-one-dimensional conductor with CDWs to a glasslike state seems to be determined mainly by processes of strong pinning of CDWs at randomly distributed impurities with formation of interacting soliton-like CDW collective excitations.

The work was supported by INTAS (Grant No. 1010-CT93-0051) and the Russian Fund for Fundamental Research (Grant No. 95-03-05811).

[1] G. Kriza, Y. Kim, A. Belezny, and G. Mihaly, Solid State Commun. **78**, 811 (1991).
[2] J. Yang and N. P. Ong, Phys. Rev. B **44**, 7912 (1991).
[3] F. Ya. Nad' and P. Monceau, Solid State Commun. **87**, 13 (1993).
[4] F. Ya. Nad' and P. Monceau, Phys. Rev. B **51**, 2052 (1995).
[5] Z. Z. Wang, H. Salva, P. Monceau *et al.*, J. Phys. Lett. **44**, L-311 (1983).
[6] J. P. Pouget *et al.*, Physica Scripta **25**, 58 (1989).
[7] D. V. Borodin, C. B. Zaitsev-Zotov, and F. Ya. Nad', Zh. Éksp. Teor. Fiz. **90**, 318 (1986) [Sov. Phys. JETP **63**, 184 (1986)].
[8] R. J. Cava, R. M. FIeming, R. G. Dunn, and E. A. Rietman, Phys. Rev. B **31**, 8325 (1985).
[9] R. J. Cava, R. M. Fleming, P. Littlewood, E. A. Rietman, L. F. Schneemyer, and R. G. Dunn, Phys. Rev. B **30**, 3228 (1984).
[10] C. Rouceau, R. Ayroles, P. Monceau, L. Guemas, A. Meerchaut, and J. Rouxel, Phys. Stat. Sol. (a) **62**, 483 (1980).
[11] M. E. Itkis, F. Ya. Nad', and P. Monceau, J. Phys. C: Condensed Matter **2**, 8327 (1990).
[12] F. Ya. Nad' and P. Monceau, Phys. Rev. B **46**, 7413 (1992).
[13] *Electronic Properties of Inorganic Quasi-One-Dimensional Compounds*, ed. by P. Monceau, Reidel, Dordrecht (1985).
[14] S. K. Zhilinskii, M. E. Itkis, I. Yu. Kal'nova, F. Ya. Nad', and V. B. Preobrazhenskii, Zh. Éksp. Teor. Fiz. **85**, 362 (1983) [Sov. Phys. JETP **58**, 211 (1983)].
[15] G. Grüner, Rev. Mod. Phys. **60**, 1129 (1988).
[16] A. H. Thompson, A. Zettl, and G. Grüner, Phys. Rev. Lett. **47**, 64 (1981).
[17] G. Grüner, L. C. Tipie, J. Sanny, W. G. Clark, and N. P. Ong, Phys. Rev. Lett. **45**, 935 (1980).
[18] T. Sekine, T. Tsuchya, and E. Matsuura, Physica B **143**, 158 (1986).
[19] U. T. Hochli, K. Knorr, and A. Loidl, Adv. Phys. **39**, 405 (1990).
[20] K. Binder and A. P. Young, Rev. Mod. Phys. **58**, 801 (1986).
[21] T. N. Nattermann and J. Villian, Phase Trans. **11**, 5 (1988).
[22] S. K. Zhilinskii, M. E. Itkis, and F. Ya. Nad', Phys. Stat. Sol. (a) **81**, 367 (1984).

[23] N. E. Mott and E. A. Davis, *Electron Processes in Non-Crystalline Materials*, Clarendon, Oxford (1989).
[24] A. K. Jonscher, *Dielectric Relaxation in Solids*, Chelsea, London (1983).
[25] S. Havriliak and S. Negami, J. Polym. Sci. C **14**, 99 (1966).
[26] H. Fukuyama and P. A. Lee, Phys. Rev. B **17**, 535 (1978).
[27] P. A. Lee and T. M. Rice, Phys. Rev. B **19**, 3970 (1979).
[28] J. R. Tucker, W. G. Lyons, and G. Gammie, Phys. Rev. B **38**, 1148 (1988).
[29] A. I. Larkin, Zh. Éksp. Teor. Fiz. **105**, 1793 (1994) [JETP **78**, 971 (1994)].
[30] S. A. Brazovskii, Zh. Éksp. Teor. Fiz. **78**, 677 (1980) [Sov. Phys. JETP **51**, 342 (1980)].
[31] C. H. Artemenko, JETP Lett. **63**, 56 (1996).
[32] M. E. Itkis and F. Ya. Nad', Synth. Metals **29**, F421 (1989).
[33] F. Ya. Nad' and M. E. Itkis, JETP Lett. **63**, 262 (1996).
[34] B. Hennion, J. P. Pouget, and M. Sato, Phys. Rev. Lett. **68**, 2374 (1992).
[35] Y. Imry and S. K. Ma, Phys. Rev. Lett. **35**, 1399 (1975).
[36] F. Ya. Nad', in *Charge-Density Waves in Solids, Modern Problems in Condensed Matter Science*, ed. by L. P. Gor'kov and G. Grüner, Elsevier, Amsterdam (1989), Vol. 25, Ch. 5.
[37] J. Souletie, J. de Phys. **51**, 883 (1990); J. de Phys. I **49**, 1211 (1988).
[38] A. F. Volkov, Phys. Lett. A **182**, 433 (1993).

Translation provided by the Russian Editorial office.

On the theory of quantum interference between inelastic and elastic electron scattering events

V. V. Rumyantsev, E. V. Orlenko and B. N. Libenson

Technical University of St. Petersburg, 195251 St. Petersburg, Russia

(Submitted 11 June 1996)
Zh. Éksp. Teor. Fiz. **111**, 1001–1015 (March 1997)

The mechanism of weak localization of relatively fast electrons scattered with a fixed energy loss from disordered media is examined. The main focus of this paper is to put forward an explanation why coherent enhancement of electron scattering in the inelastic-scattering channel takes place at angles which differ from π. A simplified kinematic model is proposed to determine the basic properties of the weak localization of electrons in the inelastic scattering channel. The model reproduces easily the range of scattering angles typical of the weak localization of electrons with a fixed energy loss. The procedure does not require calculation of the contribution from the crossed diagrams. The results agree with those based on the dynamical theory associated with the calculation of the crossed and ladder diagrams. It is possible to follow the transition from the new type of weak localization to the ordinary weak localization with decreasing energy loss. The new-type weak localization is in agreement with the regular weak localization if the energy loss is approximately equal to the energy of an optical phonon.
[S1063-7761(97)02003-9]

1. INTRODUCTION

The weak localization of conduction electrons and backscattering enhancement of classical waves in disordered media have been studied extensively during the last few decades (see, for example, Refs. 1–9). The two phenomena, which are connected with the constructive interference of random wave fields, are closely related to each other. In the case of conduction electrons, coherent quantum–mechanical backscattering can be regarded as a precursor of the exponential localization. It gives rise to a variety of quantum transport phenomena, particularly to the logarithmic increase in the resistance of metallic films with decreasing temperature which approaches absolute zero. In the case of electromagnetic waves and other classical fields, weak localization manifests itself in the enhancement of scattering in a narrow angular cone of width on the order of $1/(kl) \ll 1$ in the backward direction (k is the electron wave vector or the wave vector of a classical wave, and l is the mean free path).

Now we observe a partial shift of interest in the studies of weak localization from the problems of electron conductivity or elastic backscattering of light to new domains associated with the electron motion in disordered media. The discovery of universal conduction fluctuations has shifted the interest from average values of physical quantities to their variance and to the behavior of separate groups of electrons with fixed energies. Another point of interest is associated with the dissipation effects, because the inelastic scattering leads to the loss of phase memory of the wave function and suppresses the weak localization and the resistance fluctuations.

Finally, coherent phenomena are also of interest in the scattering of external particles (such as electrons) with a fixed energy incident on disordered samples. In contrast to electronic measurements which can only measure the conductance of a system, experiments with the beams of intermediate-energy electrons have the advantage of measuring the angular and energy spectra of electrons for an experimental realization. In Refs. 10–12, the weak localization of external electrons (with energies from tens to thousands of electronvolts) has been studied. Neutrons have also been the subject of such a consideration.[13] According to those studies, coherent phenomena can be observed in the elastic backscattering of electrons, in spite of sufficiently high energies of external electrons.

In contrast with the scattering of electromagnetic waves, the interaction of an external energetic electron with a disordered medium leads, with high probability, to inelastic scattering. The effects of inelastic processes on the conductivity under weak localization have been studied extensively (for example, see Refs. 1 and 2). We now see a new wave of this activity. One method of treating inelastic processes when they occur only in an electron reservoir (coupled to a device without inelastic processes) was developed in Refs. 14 and 15, and elsewhere subsequently. The quantum kinetic equation, which can be employed for describing quantum transport, has been derived under the assumption that the inelastic scattering is caused by noncorrelated point scatterers.[16] Much attention has been given to the effect of inelastic scattering on the observed coherent phenomena like the Aharonov–Bohm oscillations,[17] conductance fluctuations,[18,19] persistent current,[20] resonance behavior of the conductance,[21] conductance of a disordered linear chain,[22] and destruction of weak localization in inelastic scattering of particles.[23] We also point out an elegant experimental study,[24] in which the authors tried to use the weak localization as a thermometer. Some studies are devoted to the effect of dissipation on the localization of classical fields. The reflection and transmission coefficients in the presence of absorption under the localization of classical fields have been considered in Ref. 25. The effect of absorption on the wave transport has been studied in Refs. 26–35. In many cases the absorption has been introduced as a uniform imaginary energy part. The common feature of those studies is that inelastic scattering destroys the phase memory and forbids the quantum interference effect.

In some cases, however, the inelastic processes do not lead to a phase memory loss. A very simple example was considered recently.[36] In that study the authors demonstrated the effect in which the electron-photon interaction in a ballistic microstructure plays the same role as the impurity scattering in disordered media. In the presence of an external electromagnetic field all relevant photons are coherent, and spatial interference in electron-photon scattering becomes allowed, despite the inelastic nature of the collisions. The electrons do not couple to a large number of degrees of freedom, and their phase memory is preserved. The interference effects are, therefore, certainly possible in the system, even though the electron scattering is inelastic.

The quantum interference can occur even if an electron undergoes a single inelastic scattering while interacting with an incoherent electromagnetic field. Because of a single inelastic collision, the electron loses a fixed energy, $\hbar\omega$, and finds itself in the so-called inelastic-scattering channel. The energy of this electron is different from the energy of the incident particles. It can escape the medium and then be detected. In addition to the single inelastic collision, the electron should undergo at least one elastic scattering before it leaves the medium through the same surface through which it penetrates the medium. There are two ways to realize this process, since it can either start or end with an inelastic collision. The interference of electron waves associated with these complementary processes has been proved[37,38] to be constructive. It manifests itself in the enhancement of electron scattering through an angle which differs from π. The difference of this angle from π may be considerable. The new coherent phenomenon is called a new (or different) type of weak localization.

The ratio $\gamma/\omega < 1$ (where γ is the particle collision frequency, and $\hbar\omega$ is the particle energy loss) has been considered in Refs. 37–39. Recently, the opposite limiting case $\gamma/\omega > 1$, which is closer to the usual weak localization (i.e., weak localization in the elastic-scattering channel), has been considered in Refs. 40 and 41. In both cases the new type of weak localization appears to be clearly observable.

The main difference between the ordinary and the new weak localization is the typical electron scattering angle. The angular distribution of particles and radiation undergoing weak localization in a disordered medium can be found by calculating the contribution of the crossed (or so-called ''fan'') diagrams into an electron (radiation) cross section or density matrix. However, it is also useful to have a simple physical model explaining why coherent phenomena are particularly pronounced in particle scattering at certain angles. This approach is clear and enables us to evaluate the scattering angles without calculating the crossed diagrams. The main goal of the present article is to find the physical interpretation of the fact that constructive interference of the new weak localization is pronounced at scattering angles different from π.

In the case of the ordinary weak localization, there is a particular simple graphic method,[2,4] which provides insight into the phenomenon and which explains why the angle π is specific for the regular weak localization. This method takes into account that an electron with a momentum $\mathbf{k}$ is scattered via two complementary series of intermediate scattering states $\mathbf{k}\rightarrow\mathbf{k}'_1\rightarrow\mathbf{k}'_2\rightarrow\ldots\rightarrow\mathbf{k}'_{n-1}\rightarrow\mathbf{k}'_n=-\mathbf{k}$ and $\mathbf{k}\rightarrow\mathbf{k}''_1\rightarrow\mathbf{k}''_2\rightarrow\ldots\rightarrow\mathbf{k}''_{n-1}\rightarrow\mathbf{k}''_n=-\mathbf{k}$ into the $-\mathbf{k}$ state. The momentum changes are $\mathbf{q}_1,\mathbf{q}_2,\ldots,\mathbf{q}_{n-1},\mathbf{q}_n$ for the first series, and $\mathbf{q}_n,\mathbf{q}_{n-1},\ldots,\mathbf{q}_2,\mathbf{q}_1$ for the second one. The amplitudes in the final state $-\mathbf{k}$ are identical, $A'=A''=A$, and interfere constructively. This is because the complementary scattering series have the same momentum changes in opposite sequences.

The weak localization of electrons in the inelastic-scattering channel increases the electron scattering cross section at scattering angles different from π. Moreover, the effect is pronounced in a considerably wider range of angles than for the localization in the elastic-scattering channel. In this article we show that there exists a simple kinematic method which reproduces the range of scattering angles typical of the new type of weak localization. We explain the mechanism of particle localization with a fixed energy loss. The results obtained in the framework of our kinematic approach are compared with those based on the exact dynamical theory. The scattering angles typical of coherent scattering and calculated in the kinematic and dynamical approaches are in good agreement.

We shall also show that the localization of the new type turns into an ordinary localization in the limit of vanishing fixed energy loss.

2. KINEMATIC APPROACH TO DESCRIBE THE FEATURES OF WEAK LOCALIZATION

Let us consider a process in which the electron moving in a disordered medium undergoes elastic collisions and a single inelastic collision. The electron energy is assumed to be higher than the energies of the conduction electrons. The fixed energy loss $\hbar\omega$ of the electron occurs due to the inelastic collision. There are many sources of inelastic scattering, which provide a clearly distinguishable energy loss, for instance, plasmons and a number of electron atomic transitions.

In the regular weak localizations and in the new type of weak localizations the interference of the electron waves is described by crossed diagrams. In contrast with the ordinary weak localization, one of the crossed lines in the new type of weak localization corresponds to the inelastic interaction, while the others correspond to elastic interaction with randomly distributed force centers. In accordance with Eq. (18) of Ref. 37, the crossed diagrams, together with the corresponding ladder diagrams, contribute to the scattering probability factor

$$\mathscr{S}(\omega,\chi)=(2\pi)^{-3}\int_0^\infty dq_i\, q_i\, w_i(q_i,\omega)\mathscr{G}(q_i,\omega,\chi), \quad (1)$$

where χ is the electron scattering angle, $w_i(q_i,\omega)$ is the rate of inelastic scattering accompanied by excitation of the medium with a momentum q_i and energy $\hbar\omega$. The function $\mathscr{G}$ is given by

$$\mathscr{G}(q_i,\omega,\chi)=\hbar^2\int d\Omega_{\mathbf{q}_i}|G(\mathbf{k}-\mathbf{q}_i,E_k-\hbar\omega)+G(\mathbf{k}-\mathbf{Q}+\mathbf{q}_i,E_k)|^2. \tag{2}$$

Here $G(\mathbf{k}-\mathbf{q}_i,E_k-\hbar\omega)$ and $G(\mathbf{k}-\mathbf{Q}+\mathbf{q}_i,E_k)$ are the electron Green's functions. The former describes electron motion between the inelastic and elastic scattering events and the latter refers to the process with the opposite sequence of events. E_k and $\mathbf{k}$ are the initial energy and momentum of an incident electron, respectively; $\mathbf{Q}=\mathbf{q}_i+\mathbf{q}_e$ is the total momentum transfer to the medium, and $\mathbf{q}_e$ corresponds to the elastic scattering. The integration in Eq. (2) is performed over possible orientations of $\mathbf{q}_i$. Equation (2) can be rewritten in the form

$$\mathscr{G}(q_i,\omega,\chi)=\int d\Omega_{\mathbf{q}_i}\left|\frac{1}{\mathbf{v}\mathbf{q}_i-\omega-\hbar q_i^2/2m-i\gamma}+\frac{1}{\omega-\mathbf{v}'\mathbf{q}_i-\hbar q_i^2/2m-i\gamma}\right|^2. \tag{3}$$

Here we assume that the electron energy is $E_k=\hbar^2k^2/2m$, $\mathbf{v}$ and $\mathbf{v}'$ are the electron velocities in the initial and final states, respectively, and γ is the electron collision frequency.

The function (3) is convenient for studying the weak localization since it describes the propagation of electron waves between collisions. Equation (3) contains three terms. The squared absolute value of the first Green's function gives the contribution of a ladder-type diagram and describes the process in which the first collision is inelastic. The squared absolute value of the second Green's function corresponds to a ladder-type diagram for the case in which the inelastic collision is the last collision. In both cases the expressions are independent of the electron-scattering angle, $\chi=\cos^{-1}(\mathbf{v}\mathbf{v}'/vv')$. There is also the third term, which contains the product of the first and second Green's functions. It corresponds to the crossed diagrams and describes the interference of two electron waves which propagate along the same path in opposite directions. The integration over $\Omega_{\mathbf{q}_i}$ in this term does not ''kill'' the χ-dependence that describes the weak localization. The function $\mathscr{G}$ describes it at a fixed length of $\mathbf{q}_i$, while Eq. (1) is appropriate if the length of $\mathbf{q}_i$ is not fixed.

If an inelastic collision occurs between two elastic collisions, the weak localization does not exist.[38] It was also shown[38] that elastic multiple scattering at arbitrary angles does not change the angular dependence determined by the functions $\mathscr{G}$ and $\mathscr{S}$.

The squared absolute value of the first Green's function in Eq. (3) at small γ can be represented in the form

$$\left|\frac{1}{E_k-E_{\mathbf{k}-\mathbf{q}_i}-\hbar\omega-i\gamma}\right|^2\simeq\frac{\pi}{\hbar\gamma}\,\delta(E_k-E_{\mathbf{k}-\mathbf{q}_i}-\hbar\omega). \tag{4}$$

The squared absolute value of the second Green's function in Eq. (3) can be written similarly. The term in the integrand of Eq. (3), which describes the interference, has the form

$$\frac{1}{(E_k-E_{\mathbf{k}-\mathbf{q}_i}-\hbar\omega-i\gamma)(E_k-E_{\mathbf{k}-\mathbf{Q}+\mathbf{q}_i}+i\gamma)}+\frac{1}{(E_k-E_{\mathbf{k}-\mathbf{Q}+\mathbf{q}_i}-i\gamma)(E_k-E_{\mathbf{k}-\mathbf{q}_i}-\hbar\omega+i\gamma)}. \tag{5a}$$

It is evident that the ratio of Eq. (5a) to Eq. (4) is proportional to the small quantity γ. We can assume, therefore, that the frequency γ in the denominators of Eq. (5a) is (as a first approximation) an infinitesimal quantity. This enables us to rewrite Eq. (5a) in the form

$$\begin{aligned}&\left[\mathscr{P}\frac{1}{E_k-E_{\mathbf{k}-\mathbf{q}_i}-\hbar\omega}+i\pi\delta(E_k-E_{\mathbf{k}-\mathbf{q}_i}-\hbar\omega)\right]\\&\times\left[\mathscr{P}\frac{1}{E_k-E_{\mathbf{k}-\mathbf{Q}+\mathbf{q}_i}}-i\pi\delta(E_k-E_{\mathbf{k}-\mathbf{Q}+\mathbf{q}_i})\right]\\&+\left[\mathscr{P}\frac{1}{E_k-E_{\mathbf{k}-\mathbf{Q}+\mathbf{q}_i}}+i\pi\delta(E_k-E_{\mathbf{k}-\mathbf{Q}+\mathbf{q}_i})\right]\\&\times\left[\mathscr{P}\frac{1}{E_k-E_{\mathbf{k}-\mathbf{q}_i}-\hbar\omega}-i\pi\delta(E_k-E_{\mathbf{k}-\mathbf{q}_i}-\hbar\omega)\right].\end{aligned} \tag{5b}$$

An estimate reveals that in Eq. (5a) the ratio of the terms containing the product of two delta functions to the terms which contain the principal values is

$$\frac{\hbar\omega}{vq_c}\,\frac{k}{q_c}. \tag{6}$$

Here q_c is the maximum momentum q_i (for example, the cutoff plasmon momentum). As long as $k\gg q_c$, the contribution of the product of the delta functions is dominant. This means that although the quantum transport generally (and the weak localization specifically) occurs due to such electron collisions at which every next scattering begins before the end of the previous one, weak localization permits (to the first approximation) a physical interpretation, which starts from the analysis of the consequences of the simultaneous satisfaction of the two conditions,

$$E_k-E_{\mathbf{k}-\mathbf{Q}+\mathbf{q}_i}=0 \quad\text{and}\quad E_k-E_{\mathbf{k}-\mathbf{q}_i}-\hbar\omega=0. \tag{7}$$

These conditions come from the two delta functions and have the meaning of energy momentum conservation. An analysis based on Eqs. (7) is called the kinematic method in the theory of weak localization.

Below we compare the results of the kinematic analysis with the so-called dynamic results of the exact theory. We shall prove that the kinematic approach reproduces the angular properties of the weak localization with fairly good accuracy.

Let us rewrite Eqs. (7) in the form

$$E_{k'}-E_{\mathbf{k}-\mathbf{q}_i}=0, \tag{8}$$

$$E_k-E_{\mathbf{k}-\mathbf{q}_e}=0, \tag{9}$$

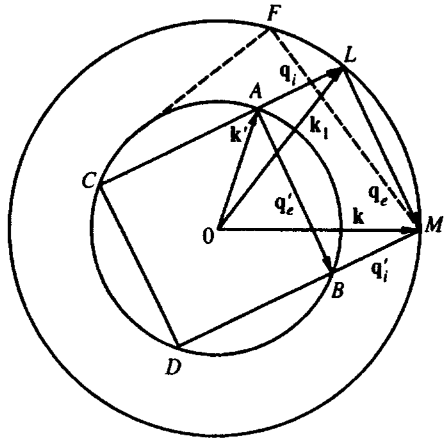


FIG. 1. Kinematic diagram of the sequence of events which give rise to a weak localization.

and show that the vectors $\mathbf{q}_i$ and $\mathbf{q}_e$, the momentum transfers during the inelastic and elastic collisions, are reciprocal orthogonal vectors.

Subtracting Eq. (8) from Eq. (9), we have

$$(\mathbf{k}-\mathbf{k}')(\mathbf{k}+\mathbf{k}')=(\mathbf{q}_i-\mathbf{q}_e)(2\mathbf{k}-\mathbf{q}_i-\mathbf{q}_e). \tag{10}$$

From Eq. (9) we obtain $q_e^2=2\mathbf{k}\mathbf{q}_e$. Since $\mathbf{Q}=\mathbf{k}-\mathbf{k}'=\mathbf{q}_i+\mathbf{q}_e$, we obtain

$$2\mathbf{k}(\mathbf{Q}-\mathbf{q}_i-\mathbf{q}_e)=Q^2-q_i^2-q_e^2. \tag{11}$$

Since $\mathbf{Q}=\mathbf{q}_i+\mathbf{q}_e$, we obtain $Q^2=q_i^2+q_e^2$. Therefore,

$$\mathbf{q}_i\mathbf{q}_e=0. \tag{12}$$

Let us now represent graphically the sequence of events which give rise to a new type of weak localization. The circles in Fig. 1 have the radii $R=k$ and $R'=k'=\sqrt{k^2-2m\hbar^{-1}\omega}$, respectively. The radii coincide with the lengths of the electron wave vectors in the initial and final states. Let us first consider the case in which the first collision is elastic. For brevity, we denote this process as $\langle m,\mathbf{k}'|H_iGH_e|n,\mathbf{k}\rangle$. Here m and n correspond to the initial and final states of the medium. The end of the vector $\mathbf{q_e}$, the momentum transfer during the elastic scattering (which may involve multiple elastic scattering), touches the circle R, so that the electron wave vector becomes equal to $\mathbf{k}_1$ as a result of elastic scattering. The following event of the scattering is an inelastic collision with a momentum transfer $\mathbf{q}_i$, and the condition $\mathbf{q}_i\mathbf{q}_e=0$ is satisfied. The vector $\mathbf{q_i}$ connects the end of the vector $\mathbf{k}_1$ with the point A on the circle $R'=k'$ (where the end of the vector $\mathbf{k}'$ rests). The energy of the final state electron is lower by $\hbar\omega$ than the energy of the initial state.

In the complementary scattering process $\langle m,\mathbf{k}'|H_eGH_i|n,\mathbf{k}\rangle$ the particle first loses its energy and only then undergoes the elastic scattering. The interference between two realizations of the scattering process is effective if the wave vectors transferred during inelastic scattering in each realization are parallel and if they have the same lengths. Therefore, the momentum transfer $\mathbf{q}_i'$ in the event of inelastic scattering in the process $\langle m,\mathbf{k}'|H_eGH_i|n,\mathbf{k}\rangle$ is shown with the segment that connects the end of the vector $\mathbf{k}$ and the point B on the circle $R'=k'$, where the vector $\mathbf{q}_i'$ is perpendicular to the vector $\mathbf{q}_e$. The vector $\mathbf{q}_e'$ in the process under consideration is shown with the segment AB. The vectors $\mathbf{q}_e$ and $\mathbf{q}_i$ in the process $\langle m,\mathbf{k}'|H_iGH_e|n,\mathbf{k}\rangle$ and the vectors $\mathbf{q}_i'$ and $\mathbf{q}_e'$ in the process $\langle m,\mathbf{k}'|H_eGH_i|n,\mathbf{k}\rangle$ convert the vector $\mathbf{k}$ to $\mathbf{k}'$.

As can be seen from Fig. 1, there is the other set of vectors $\mathbf{q}_i,\mathbf{q}_e$ and $\mathbf{q}_i',\mathbf{q}_e'$. In this case $\mathbf{q}_i$ connects the point C (at the end of $\mathbf{k}'$) with the end of $\mathbf{k_1}$. The segment CD and the segment joining the point D and the end of $\mathbf{k}$ define $\mathbf{q}_e'$ and $\mathbf{q}_i'$.

As a rule, the two sets of solutions, which correspond to the segments AB and CD, cannot be realized simultaneously. It is obvious that a solution concerned with the existence of the segment CD corresponds to the length $\mathbf{q}_i$,

$$q_i=\sqrt{k^2-(q_e/2)^2}+\sqrt{k'^2-(q_e/2)^2},$$

and to the electron-scattering angle,

$$\chi=\frac{\pi}{2}+\sin^{-1}\left(\frac{q_e}{2k}\right)+\sin^{-1}\left(\frac{q_e}{2k'}\right),$$

so that $\pi/2\leqslant\chi\leqslant\pi$.

Therefore,

$$q_i^2=k^2+k'^2+2kk'\sin\chi, \tag{13}$$

$$q_e^2=\frac{(2kk'\cos\chi)^2}{k^2+k'^2+2kk'\sin\chi}. \tag{14}$$

It is evident that $q_i/q_e>1$, an inequality that is difficult to realize. For example, in the case of plasmon excitations there is a cutoff vector $q_i=q_{ic}=\omega v_F^{-1}$ (where v_F is the electron Fermi velocity), and the probability of plasmon excitation with $q_i>q_{ic}$ is zero. In other cases of inelastic scattering the processes with small momentum transfers are also probable. However, when taking the limit of the ordinary weak localization $(k'\rightarrow k)$, the solution of Eqs. (13) and (14) corresponds to the correct description of the process. When the ratio $\hbar\omega/E$ is small, we have

$$\cos\chi=-1+\frac{1}{8[(2k/q_e)^2-1]}\left(\frac{\hbar\omega}{E}\right)^2, \tag{15}$$

and $\chi\rightarrow\pi$ as $\hbar\omega\rightarrow 0$.

When we consider the new type of weak localization, the solution associated with small momentum transfers $\mathbf{q_i}$ is realized. In this case the interference conforms to the rectangle with the vortices $ABLM$. According to Fig. 1, the dependence of the length of the vector $\mathbf{q_i}$ on χ has the form

$$q_i=\sqrt{k^2-(q_e/2)^2}-\sqrt{k'^2-(q_e/2)^2}, \tag{16}$$

and

$$\chi=\sin^{-1}\left(\frac{q_e}{2k}\right)+\sin^{-1}\left(\frac{q_e}{2k'}\right). \tag{17}$$

From Eqs. (16) and (17) we obtain

$$q_e^2=\frac{(2kk'\ \sin\chi)^2}{k^2+k'^2+2kk'\ \cos\chi},\tag{18}$$

$$q_i=\frac{2m\omega}{\hbar\sqrt{k^2+k'^2+2kk'\ \cos\chi}}.\tag{19}$$

We see from Eq. (19) that

$$\cos\chi=-\frac{k^2+k'^2}{2kk'}+\frac{2\omega^2}{vv'q_i^2}.\tag{20}$$

This expression determines the electron-scattering angle under weak localization at different values of $\mathbf{q}_i$, the inelastic momentum transfer.

If $k-k'\ll k$, then

$$\cos\chi=-1+\frac{2\omega^2}{v^2q_i^2}\ \text{ or }\ \cos\frac{\chi}{2}=\frac{\omega}{vq_i}.$$

It is convenient to rewrite them as

$$\chi=2\cos^{-1}\left(\frac{\omega}{vq_i}\right).\tag{21}$$

If the electron scattering excites the electromagnetic waves (longitudinal or transverse), Eq. (21) can be interpreted as the condition of the electron scattering through the double Čerenkov angle. This circumstance was pointed out in Ref. 37.

There is a top value of the vector $\mathbf{q}_e$ (this vector is represented by the dashed line MF) at which the vector $\mathbf{q}_i$, drawn from the point F perpendicularly to the segment MF, ceases to cross the inner circle. Accordingly, it is impossible to expect the new type of weak localization to be at electron-scattering angles close to π. We see from the geometric consideration that the top scattering angle is

$$\chi_0=\pi-\cos^{-1}\left(1-\frac{\hbar\omega}{2E}\right).\tag{22}$$

Therefore, we obtain

$$q_{i0}=k\ \sin\chi_0=\sqrt{2m\omega\hbar^{-1}}.\tag{23}$$

The new type of weak localization takes place only if $q_i<q_{i0}$.

Now we can determine the range of scattering angles of the electrons that undergo weak localization for different types of medium excitations. When a bulk plasmon is excited and its wave vector is in the range

$$\frac{\omega_p}{v}<q_i<q_c=\frac{\omega_p}{v_F},$$

the coherent phenomena at electron scattering occur at

$$0<\chi<2\cos^{-1}(v_F/v).\tag{24}$$

The range of wave vectors of the transverse electromagnetic waves (at the Čerenkov excitation) is

$$\frac{\omega}{v}<q_i<\frac{\omega}{c}\sqrt{\epsilon}.$$

The most intense scattering will then occur in the range of scattering angles

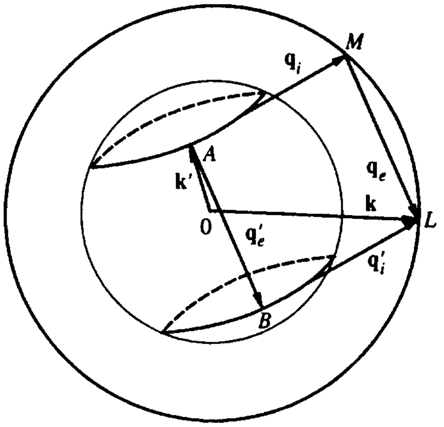


FIG. 2. Three-dimensional diagram of the realization of constructive interference.

$$0<\chi<2\cos^{-1}\left(\frac{c}{v\sqrt{\epsilon}}\right).\tag{25}$$

In the case of excitation and ionization of atoms[37] we have

$$q_i\simeq\frac{e\sqrt{Z}}{|d_{x\ mn}|},$$

where Z is the atomic number, and $d_{x\ mn}$ is the matrix element of the atomic dipole for excitation from the ground state n to the upper state m. Therefore, the coherent phenomena at electron scattering will be pronounced at

$$\chi\simeq2\cos^{-1}\left(\frac{\omega|d_{x\ mn}|}{ev\sqrt{Z}}\right).\tag{26}$$

For an optical-phonon excitation, the weak localization is effective when

$$\frac{\omega_{\text{ph}}}{v}<q_i<\frac{\sqrt{2m\omega_{\text{ph}}}}{k}\ll|\mathbf{G}|$$

($\mathbf{G}$ is the reciprocal lattice vector). The appropriate range of the electron-scattering angles is

$$0<\chi<2\cos^{-1}(\sqrt{\hbar\omega/4E}).\tag{27}$$

Until now we have assumed that every vector in Fig. 1 lies in the same plane. However, the vector $\mathbf{q}_i$, which is perpendicular to $\mathbf{q}_e$ does not necessarily lie in the plane of the vectors $\mathbf{k}$ and $\mathbf{q}_e$. If the vectors $\mathbf{q}_i$, $\mathbf{q}_e$, $\mathbf{q}_i'$, $\mathbf{q}_e'$, $\mathbf{k}$, and $\mathbf{k}'$ are situated as shown in Fig. 2, the complementary scattering processes will be accompanied by a constructive interference. In this case the ends of the vectors $\mathbf{q}_i$ and $\mathbf{q}_i'$ are situated on the circles formed by intersection of the inner sphere $R'=k'$ and the planes which are perpendicular to $\mathbf{q}_e$ and pass through the points L and M. The vectors $\mathbf{q}_i$ and $\mathbf{q}_i'$ form the ‘‘fan’’

$$\frac{\omega\sqrt{2}}{v\sqrt{1-\cos\chi}}<q_i<\min\left[\sqrt{\frac{2m\omega}{\hbar^2}},q_{i\ \max}\right].\tag{28}$$

Instead of Eq. (15) in this case we have

$$\cos\chi=-1+\frac{1}{8[(2k/q_e)^2-1]\cos\phi}\left(\frac{\hbar\omega}{E}\right)^2, \tag{29}$$

where ϕ is the azimuthal angle (in the plane perpendicular to $\mathbf{k}$) between the vectors $\mathbf{k}'$ and $\mathbf{q}_e$. Although the kinematic diagram of the realization of effective interference is now three-dimensional, our conclusions about the interference mechanism and about the features of electron-scattering angles remain valid.

3. DYNAMICAL APPROACH FOR THE DESCRIPTION OF THE NEW TYPE OF WEAK LOCALIZATION

Although the description of weak localization requires, first of all, the contribution from the poles of the Green's functions in Eq. (3), the contribution from the principal value might modify the results of the previous section. In addition, we shall consider the uncertainty of the radii of the spheres in Figs. 1 and 2 due to the image potential of the medium. The uncertainty is $\kappa\simeq\sqrt{2mU'}[2\hbar\sqrt{E}]^{-1}$, where U' is the image potential mentioned above.

Moreover, the momentum transfer $\mathbf{q}_i$ usually is not fixed in experiments. Hence we should perform integration over $\mathbf{q}_i$. It is convenient to introduce the function

$$M(\chi)=\frac{\int_0^{k+k'}dq_i\,q_i^2w_i(q_i,\omega)\mathscr{G}_c(q_i,\omega,\chi)}{\int_0^{k+k'}dq_i\,q_i^2w_i(q_i,\omega)\mathscr{G}_L(q_i,\omega)}, \tag{30}$$

which is called the degree of coherency. Equations (12), (18), and (19) in Ref. 37 clarify the definition of $M(\chi)$. In Eq. (30) $w_i(q_i,\omega)$ is the rate of excitation of a state with energy $\hbar\omega$ and momentum $\mathbf{q}_i$.

The functions $\mathscr{G}_c$ and $\mathscr{G}_L$ occur due to the crossed and ladder diagrams; these functions are defined as

$$\mathscr{G}_c(q_i,\omega,\chi)=2\,\mathrm{Re}\int d\Omega_{\mathbf{q}}\times\frac{1}{(\mathbf{v}\mathbf{q}_i-\omega-\hbar q_i^2/2m+i\gamma)(\omega-\mathbf{v}'\mathbf{q}_i-\hbar q_i^2/2m-i\gamma)}, \tag{31}$$

$$\mathscr{G}_L(q_i,\omega)=\int d\Omega_{\mathbf{q}}\left\{\frac{1}{(\mathbf{v}\mathbf{q}_i-\omega-\hbar q_i^2/2m)^2+\gamma^2}+\frac{1}{(\omega-\mathbf{v}\mathbf{q}_i-\hbar q_i^2/2m)^2+\gamma^2}\right\}. \tag{32}$$

Here the damping is $\gamma=\kappa v$.

To analyze the weak localization during excitation of the long-wavelength medium states, e.g., of plasmons or optical phonons, it is convenient to write $w_i(q_i,\omega)$ as follows:

$$w_i(q_i,\omega)=\mathscr{F}(\omega)q_i^{-2}\theta(q_{i\,\max}-q_i).$$

The function $\mathscr{F}(\omega)$ is the long-wavelength limit of the imaginary part of the reverse dielectric function of the medium accurate within a constant. The degree of coherency will then be

$$M(\chi)=\frac{\int_0^{\min(k+k',q_{i\,\max})}dq_i\,\mathscr{G}_c(q_i,\omega,\chi)}{\int_0^{\min(k+k',q_{i\,\max})}dq_i\,\mathscr{G}_L(q_i,\omega)}. \tag{33}$$

The function $\mathscr{G}_L$ determines the incoherent part of the electron cross section and takes the form

$$\mathscr{G}_L(q_i,\omega)=\frac{2\pi}{\gamma q_i}\left\{\frac{1}{v}\left[\tan^{-1}\left(\frac{vq_i+\omega+\hbar q_i^2/2m}{\gamma}\right)+\tan^{-1}\left(\frac{vq_i-\omega-\hbar q_i^2/2m}{\gamma}\right)\right]+\frac{1}{v'}\left[\tan^{-1}\left(\frac{v'q_i+\omega-\hbar q_i^2/2m}{\gamma}\right)+\tan^{-1}\left(\frac{v'q_i-\omega+\hbar q_i^2/2m}{\gamma}\right)\right]\right\}. \tag{34}$$

The coherent part is determined by

$$\mathscr{G}_c(q_i,\omega,\chi)=\frac{4\pi}{q_i}\mathrm{Re}\left\{\mathscr{Y}^{-1}(v,v',\omega,\omega_c,\chi,q_i)\times\log\left[\frac{q_i^2(\mathbf{v}\mathbf{v}')+(\hbar q_i^2/2m)^2-\omega_c^2+q_i\mathscr{Y}}{q_i^2(\mathbf{v}\mathbf{v}')+(\hbar q_i^2/2m)^2-\omega_c^2-q_i\mathscr{Y}}\right]\right\}, \tag{35}$$

where $\mathscr{Y}(v,v',\omega,\omega_c,\chi,q_i)$ has the form

$$\mathscr{Y}=\left\{\omega_c^2(\mathbf{v}-\mathbf{v}')^2-\left[(q_i[\mathbf{v}\mathbf{v}'])^2+\frac{2\hbar^2\omega\omega_cq_i^2}{m^2}-\left(\frac{\hbar q_i^2}{2m}\right)^2(\mathbf{v}+\mathbf{v}')^2\right]\right\}^{1/2}. \tag{36}$$

Here $\omega_c=\omega-i\gamma$.

The position of the angular features of the interference part in the electron cross section is mainly determined by zeros of the function $\mathscr{Y}$. The equation $\mathscr{Y}=0$ has two roots:

$$(q_i^2)_1=\frac{8\omega_c^2[[\mathbf{v}\mathbf{v}']^2+(\hbar\omega/m)^2](\mathbf{v}+\mathbf{v}')^{-2}}{[\mathbf{v}\mathbf{v}']^2+2\hbar^2\omega\omega_cm^{-2}+\sqrt{[\mathbf{v}\mathbf{v}']^4+4\hbar\omega_c(\omega-\omega_c)m^{-2}[\mathbf{v}\mathbf{v}']^2}}, \tag{37}$$

$$(q_i^2)_2=\frac{2m}{\hbar^2}\,\frac{[\mathbf{v}\mathbf{v}']^2+2\hbar^2\omega\omega_cm^{-2}+\sqrt{[\mathbf{v}\mathbf{v}']^4+4\hbar\omega_c(\omega-\omega_c)m^{-2}[\mathbf{v}\mathbf{v}']^2}}{(\mathbf{v}+\mathbf{v}')^2}. \tag{38}$$

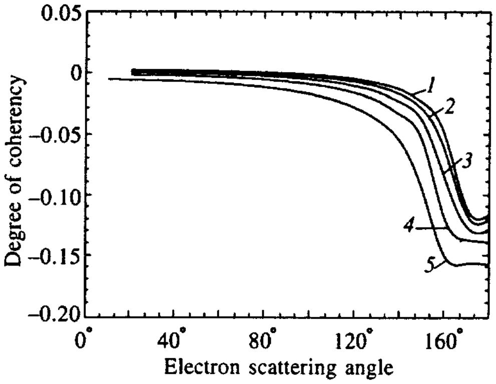


FIG. 3. Degree of coherency versus electron-scattering angle for plasmon excitation. The sharp decrease of the curves is due to the absence of the Čerenkov generation of plasmons. The difference in the degree of coherency from zero at angles which lie in the range from 2 $\cos^{-1}(v_F/v)$ up to π is due to the finiteness of γ and the contribution from the principal value in Eq. (5b). $\hbar\omega = 10$ eV, $U' = 1$ eV. Curve *1* corresponds to $E = 200$ eV, *2*—300 eV, *3*—400 eV, *4*—500 eV, and *5*—600 eV.

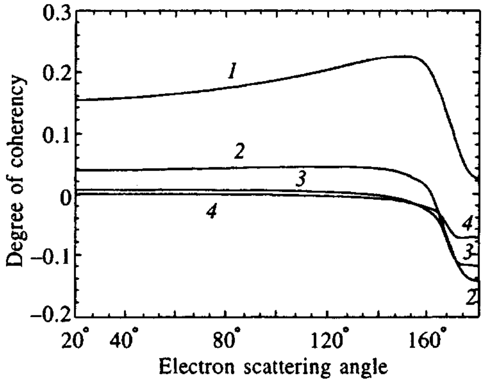


FIG. 4. Degree of coherency at different values of the energy loss. $E = 500$ eV, $U' = 1$ eV. Cureve *1* corresponds to $\hbar\omega = 2$ eV, *2*—4 eV, *3*—8 eV, and *4*—12 eV.

We can associate these roots with the value q_i defined by Eq. (19). The association is clearly seen in the limit of $\gamma \to 0$. In this case $\omega_c \to \omega$, and from Eq. (37) we obtain

$$q_{i_1}=\frac{2\omega}{\sqrt{|\mathbf{v}+\mathbf{v}'|^2}}. \tag{39}$$

Equation (39) coincides with Eq. (19). This means that the kinematic approach yields a reasonably good accuracy.

In the same approximation Eq. (38) yields

$$q_{i2}=\sqrt{\frac{k^4+k'^4-2k^2k'^2\cos 2\chi}{k^2+k'^2+2kk'\cos\chi}}. \tag{40}$$

From Eqs. (40), (18), and (19) we see the sum of Eq. (18) and the squared Eq. (19) is equal to the squared right-hand side of Eq. (40). Therefore, Eq. (40) corresponds to the case where $\mathbf{q}_i = \mathbf{Q}$. This solution is not significant for the new type of weak localization.

Since the equations are rather complicated and do not give a transparent insight into the dependence of the electron cross section on the various parameters, we present the theoretical features in Figs. 3–6 for a few typical cases.

Figure 3 shows $M(\chi)$ for the plasmon excitation in metals. Every curve shows a clearly defined slope. The nature of the sharp decrease has been explained in Sec. 2 on the basis of Eq. (22). In the case of a particular excitation the condition (22) might undergo some change. The dynamical approach takes into account these changes. In the particular case of the bulk plasmon excitation the change can be described in terms of the Čerenkov and bremsstrahlung generation of plasmons. The weak localization takes place when the plasmon generation mechanism differs only slightly from the Čerenkov mechanism. The absence of the Čerenkov generation of plasmons at

$$2\cos^{-1}(v_F/v)<\chi<\pi$$

implies that at these angles the new type of weak localization is suppressed. The difference in the cross sections from zero at the angles which lie in the range from 2 $\cos^{-1}(v_F/v)$ up to π is due to the finiteness of γ and the contribution from the principal value in Eq. (5b).

Figure 4 shows the plot of $M(\chi)$ versus the energy loss due to the plasmon excitation. With an increase in $\hbar\omega_p$, the peak position shifts to smaller χ. If $\hbar\omega = 2$ eV, the constructive interference takes place at $\chi_m = 150°$. If $\hbar\omega = 4$ eV, we have $\chi_m = 120°$. These results agree with Eq. (21) which was obtained from a kinematic analysis.

The curves $M(\chi)$ depend also on the image potential. As shown in Fig. 5, a considerable decrease in the degree of coherency at $\chi = \pi$ (i.e., in the range of bremsstrahlung plasmon generation) takes place up to $\gamma/\omega = 0.5$. Only at $\gamma/\omega \simeq 1$ the angular dependence begins to smooth out when $\chi = 2\cos^{-1}(\omega/vq_i)$. In that case it looks like a peak on the curve $M(\chi)$. With an increase in γ, it shifts to $\chi = \pi$.

Figure 6 shows a theoretical curve $M(\chi)$ for excitation of polar optical phonons. Weak localization associated with this excitation corresponds to curve *5*. We assume that $\hbar\omega_{\rm ph} = 0.05$ eV. Other curves correspond to $\hbar\omega_{\rm ph}$ larger than for ordinary optical phonons. These curves show a variation of $M(\chi)$ when the energy loss increases from phonon to plasmon losses. At $\hbar\omega < 0.5$ eV, the width of the coherent peak at $\chi \simeq \pi$ is equal to 15° and it does not get narrower

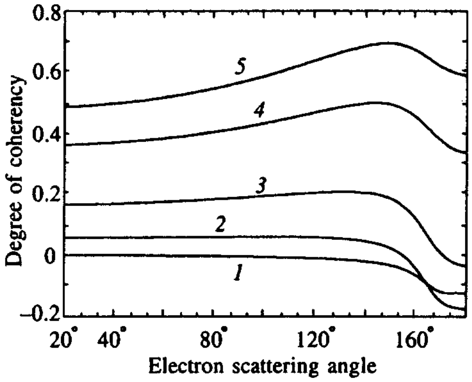


FIG. 5. Dependence of the degree of coherency on the image potential. $E = 500$ eV, $\hbar\omega = 10$ eV. Curve *1* corresponds to $U' = 1$ eV, *2*—3 eV, *3*—5 eV, *4*—8 eV, and *5*—10 eV.

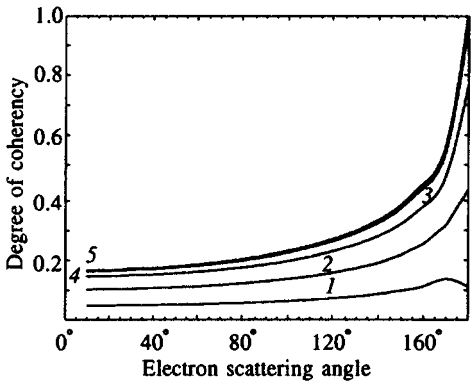


FIG. 6. Degree of coherency for quasielastic scattering. $E = 500$ eV, $U' = 1$ eV. Curve *1* corresponds to $\hbar\omega = 2$ eV, *2*—1.0 eV, *3*—0.5 eV, *4*—0.2 eV, and *5*—0.05 eV.

with a further decrease of $\hbar\omega$. This means that the weak localization of an electron during elastic scattering does not differ noticeably from the new type of weak localization during quasielastic scattering.

4. CONCLUSIONS

With a decrease in the energy loss, the new type of weak localization transforms into an ordinary weak localization. The transition is described by taking into account that the term $q_i^2/2m$ in the denominator of the Green's functions in Eq. (3) is more significant than $\hbar\omega$ in the limit $\omega \rightarrow 0$.

We have mentioned this fact at the end of Sec. 3. One can study the transition further by analyzing the variation of the ratio $\mathcal{G}_c/\mathcal{G}_L$ with $2m\omega/q_i^2$ at different scattering angles. The transition of the ordinary weak localization (i.e., at $\hbar\omega = 0$) to the new localization ($\hbar\omega \neq 0$) is shown in Fig. 7. For example, at $2m\omega q_i^{-2} = 0.8$ an announced maximum occurs for $\chi \simeq 175^\circ$. The transition seems to take place for those $\hbar\omega$ and $q_i^2/2m$ which are $\geqslant \gamma$.

It can now be stated with assurance that we have explained why the angles typical of the new type of weak localization differ from π. The simple kinematic approach enables us to estimate these angles very accurately. The dependence of the angles on the electron energy loss and on other parameters can be predicted. Finally, we have shown that there is no wall between new and ordinary weak localizations. The two phenomena are two different manifestations of the constructive quantum interference of electron waves.

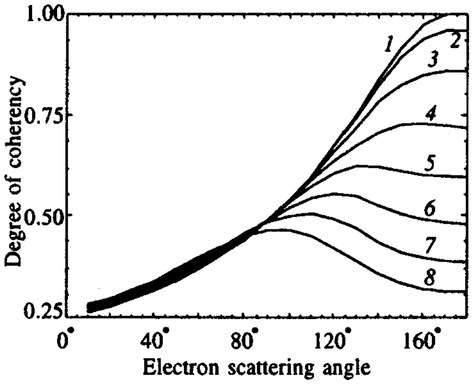


FIG. 7. Transition rate from ordinary weak localization to the new type of weak localization. $q_i^2/2mE = 0.01$, $q_c v/E = 0.2$, $\gamma/E = 0.01$. Curve *1* refers to $\hbar\omega = 0$, curve *2* corresponds to $\hbar\omega/E = 0.002$, *3*—0.004, *4*—0.006, *5*—0.008, *6*—0.010, *7*—0.012, and *8*—0.014.

We thank Prof. D. G. Yakovlev for critical reading of our paper and for helpful advice.

[1] P. A. Lee and T. V. Ramakrishnan, Rev. Mod. Phys. **57**, 287 (1985).
[2] G. Bergmann, Phys. Rev. **107**, 1 (1984).
[3] J. Rammer, Rev. Mod. Phys. **63**, 781 (1991).
[4] B. Kramer and A. MacKinnon, Rep. Prog. Phys. **56**, 1469 (1993).
[5] A. Aronov, Phys. Scr. **49**, 28 (1993).
[6] M. B. van der Mark, M. P. van Albada, and A. Lapendijk, Phys. Rev. **37**, 3575 (1988).
[7] Y. A. Kravtsov, Rep. Prog. Phys. **55**, 39 (1992).
[8] M. Kaveh, Physica B **175**, 1 (1991).
[9] *Scattering and Localization of Classical Waves in Random Media*, ed. by P. Sheng, World Scientific, Singapore (1990).
[10] R. Berkovits and M. Kaveh, Phys. Rev. B **37**, 584 (1988).
[11] R. Berkovits, D. Eliyahu, and M. Kaveh, Phys. Rev. B **41**, 407 (1990).
[12] E. E. Gorodnichev, S. L. Dudarev, and D. B. Rogozkin, Zh. Éksp. Teor. Fiz. **97**, 1511 (1990) [Sov. Phys. JETP **70**, 645 (1990)].
[13] Jun-ichi Igarashi, Phys. Rev. B **35**, 8894 (1987).
[14] M. Büttiker, Phys. Rev. B **32**, 1846 (1985).
[15] M. Büttiker, Phys. Rev. B **33**, 3020 (1986).
[16] Supriyo Datta, Phys. Rev. B **40**, 5830 (1989).
[17] S. Washburn and R. A. Webb, Adv. Phys. **35**, 375 (1986).
[18] L. P. Kanwenhoven, F. W. J. Hekking, B. J. van Wees, C. J. P. M. Harmans, C. E. Timmerling, and C. T. Foxon, Phys. Rev. Lett. **65**, 361 (1990).
[19] F. Beltram, F. Capasso, D. L. Sivco, A. L. Hutchinson, S. N. G. Chu, and A. Y. Chao, Phys. Rev. Lett. **64**, 3167 (1990).
[20] V. Ambegoakar and U. Eckern, Phys. Rev. Lett. **65**, 381 (1990).
[21] K. Maschke and M. Schreiber, Phys. Rev. B **44**, 3835 (1991).
[22] J. L. D. D'Amato and H. M. Pastawski, Phys. Rev. B **41**, 7411 (1990).
[23] E. A. Kanzieper, Zh. Éksp. Teor. Fiz. **103**, 1800 (1993) [JETP **76**, 887 (1993)]; Phys. Scr. **47**, 823 (1993).
[24] G. Bergmann, Wei Wei, Yao Zou, and R. M. Muller, Phys. Rev. B **41**, 7387 (1990).
[25] P. W. Anderson, Philos. Mag. B **52**, 505 (1985).
[26] V. Freilikher, M. Pustilnik, and I. Yurkevich, Phys. Rev. Lett. **73**, 810 (1994).
[27] K. Maschke and M. Schreiber, Phys. Rev. B **49**, 2498 (1994).
[28] R. L. Weaver, Phys. Rev. B **47**, 1077 (1993).
[29] L. F. Rigister and K. Hess, Phys. Rev. B **49**, 1900 (1994).
[30] S. John, Phys. Rev. Lett. **53**, 2169 (1984).
[31] T. R. Kirkpatrick, Phys. Rev. B **31**, 5746 (1985).
[32] C. A. Condat, J. Acoust. Soc. Am. **83**, 441 (1988).
[33] A. M. Jayannavar, Phys. Rev. B **49**, 14718 (1994).
[34] M. Yosefin, Europhys. Lett. **25**, 675 (1994).
[35] V. Freilikher, M. Pustilnik, and I. Yurkevich, Phys. Rev. B **50**, 6017 (1994).
[36] L. Y. Gorelik, A. Grincwojg, V. Z. Kleiner, R. I. Shekhter, and M. Jonson, Phys. Rev. Lett. **73**, 2260 (1994).
[37] B. N. Libenson, K. Yu. Platonov, and V. V. Rumyantsev, Zh. Éksp. Teor. Fiz. **101**, 614 (1992) [Sov. Phys. JETP **74**, 326 1992].
[38] V. V. Rumyantsev and V. V. Doubov, Phys. Rev. B **49**, 8643 (1994).
[39] E. V. Orlenko and V. V. Rumyantsev, J. Phys., Condens. Matter. **7**, 3557 (1995).
[40] E. A. Kanzieper and V. Freilikher, Phys. Rev. B **51**, 2759 (1995).
[41] E. A. Kanzieper, Phys. Rev. B **51**, 15563 (1995).
[42] G. Bergmann, Phys. Rev. B **28**, 2914 (1983).

Published in English in the original Russian journal. Reproduced here with stylistic changes by the Translation Editor.

Propagation of magnetostatic waves in unsaturated ferrite films with a strip domain structure

A. V. Vashkovskiĭ, É. G. Lokk, and V. I. Shcheglov

Institute of Radio Engineering and Electronics, Russian Academy of Sciences,
141120 Fryayzino, Moscow Region, Russia

(Submitted 24 June 1996)
Zh. Éksp. Teor. Fiz. **111**, 1016–1031 (March 1997)

The propagation of surface and volume magnetostatic waves in unsaturated films of yttrium iron garnet is studied experimentally for the case when the wavelength greatly exceeds the domain width, while the domain width is comparable to the film thickness. The characteristics of these waves are examined for symmetric linear, asymmetric linear, and symmetric zigzag strip domain structures in the films. These characteristics cannot be explained by a theory based on averaging the magnetization over all the domains. © *1997 American Institute of Physics.* [S1063-7761(97)02103-3]

1. INTRODUCTION

Studies of the propagation of magnetostatic waves in ferrite films with a domain structure have been of interest for many years because of saturation owing to various physical phenomena and effects. Most of the previous work has been devoted to the propagation of magnetostatic wave in ferrite films and wafers with regular strip domain structures (referred to as plane-parallel or laminar by some writers),[1–11] but some papers report observations of magnetostatic waves in films with an irregular domain structure.[12] The characteristics of magnetostatic waves in films with a strip domain structure have most often been examined theoretically assuming that the wavelength λ is much greater than the domain width d, while the ferrite layer thickness satisfies $s \gg d \gg \delta$ (δ is the thickness of the domain boundaries, which are assumed to be infinitely thin 180-degree walls).[1,3–6] The case $\lambda \ll d$ has been studied in Ref. 2 and $\lambda \simeq d$, in Ref. 8. Here the medium was described by magnetic susceptibility $\hat{\chi}$ and permeability $\hat{\mu}$ tensors averaged over all the domains[3,7] or the medium was treated as a ''macroscopic antiferromagnet'' in which neighboring domains play the role of antiparallel spins.[1] Both these models have made it possible to use conventional techniques from electrodynamics to obtain dispersion relations. Depending on the additional assumptions that have been made, such as the absence of an external magnetic field,[1,4,8] the absence of anisotropy fields in ferrite,[3] or the presence of an easy axis[4,8] or cubic symmetry with an anisotropy constant $K_1 < 0$[5–7] in the ferrite, the calculated $\hat{\chi}$ and $\hat{\mu}$ tensors were either diagonal[3,4] or nondiagonal[1,2,5–7], while the resulting dispersion relations described different types of surface and volume waves. The theoretical models obtained in these papers, however, only apply to a rather limited or specific circle of actual situations. In many experiments, the domain width and the film thickness do not obey the relationship $s \gg d$, so it becomes necessary to include the demagnetizing fields of the static effective magnetic surface charges created by the domain walls and the calculations cannot be carried out using the earlier theoretical models.[1,3–6]

Besides wave processes associated with magnetostatic waves, domain walls in unsaturated ferrite films can oscillate at high frequencies[11] and spin-wave excitation, such as waves or modes localized on domain boundaries,[13,14] can occur. Collective oscillations of the domain structures as an entire ensemble can exist, whose dynamic properties were described by Fillipov and Tankeev.[15]

Here we describe an experimental study of the characteristics of magnetostatic waves with wave numbers $0 < k < 1000$ cm^{-1} in the case of $\lambda \gg d$ and $s \simeq d$ for tangentially magnetized films of yttrium iron garnet (one of the most widely studied materials in research on the physics of magnetism) grown in the (111) plane. No detailed studies of this kind have been carried out previously and only some preliminary results[10,16] have been reported.

2. EXPERIMENTAL APPARATUS AND THE PARAMETERS OF THE MAGNETIC FILMS

The experiments were conducted on a device which permitted simultaneous microwave and optical measurements. The ferrite films to be studied were magnetized by a tangential uniform magnetic field $\mathbf{H}_0 \simeq 0-100$ Oe. Magnetostatic waves were excited and detected by means of moveable antennas with transducers made of gold-plated tungsten wires of length 3.5 mm and thickness 12 μm. A mechanical system was used to displace the antennas in two mutually perpendicular directions in the plane of the film and rotate them about an axis perpendicular to the surface of the film. The detector antenna was moved along a line joining the centers of the transducers to determine the wave numbers by measuring the phase shift from the phase-frequency characteristic at a fixed frequency (sliding probe technique). The orientation of the crystallographic axes of the films relative to the direction of $\mathbf{H}_0$ could be varied by rotating the film about an axis perpendicular to its plane. The domain structure was observed with the aid of a measurement microscope from the Faraday effect when a beam of light was incident on the film perpendicular to its surface.

Films of yttrium iron garnet grown in the (111) plane with a well resolved domain structure as seen by the microscope were chosen for study. This domain structure was ap-

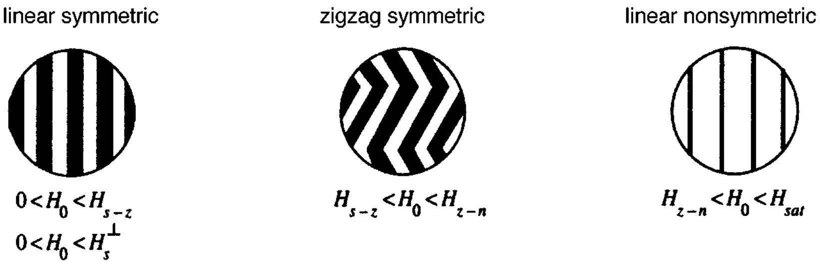


FIG. 1. Forms of strip domain structures in yttrium iron garnet films.

parently a consequence of the presence of an easy-axis magnetic uniaxial anisotropy perpendicular to the film surface, so that the local magnetization of the domains had a large component along this axis over a rather wide range of fields. The parameters of the three films that were most studied are shown in Table I, where the following notation is used: $4\pi M_0$ is the saturation magnetization, ΔH is the half width of the resonance line, s is the film thickness, H_c is the cubic anisotropy field, θ is the deviation of the uniaxial anisotropy axis from the normal to the film plane, and φ_a is the angle between the projection of the uniaxial anisotropy axis on the film plane and the [110] axis. (The parameters H_{s-z}, H_{z-n}, H_{sat}, $H_{s-n}^{\perp}$, and $H_{\mathrm{sat}}^{\perp}$ will be defined later.) The parameters s, H_c, θ, and φ_a were measured by a method that has been described elsewhere.[17] This method yielded large errors in the uniaxial anisotropy H_a and in $4\pi M_0$, so that the listed values of $4\pi M_0$ were determined from microwave measurements on films that had been magnetized to saturation.

The film size was at least 30×30 mm^2, so that edge effects were eliminated.

In order to avoid nonlinear effects (three-magnon decay and four-magnon scattering), for which the threshold level is generally higher in unsaturated yttrium iron garnet films than in saturated films[18] (in the latter this level is $\simeq 5-15$ μW, at best), the power of the cw sinusoidal microwave signal incident on the input converter was kept below 1 μW over the entire frequency range.

3. DOMAIN STRUCTURES IN THE TEST FILMS

The type of domain structure that occurs in the films depends on the magnitude of the applied magnetic field $\mathbf{H}_0$ and on the previous domain state of the film, i.e., on the type of domain structure which existed in the film prior to this, as well as the orientation of its domain boundaries relative to the applied field $\mathbf{H}_0$. In order to eliminate the effect of the previous domain structure of a film on the domain structure which develops in the film, before the experiments the film was magnetized to saturation along a crystallographic direction, after which the magnitude of the field was no longer allowed to go below 3–5 Oe (so as to preclude random domain structures). This made it possible to create just regular types of domain structures and to reproduce each type of structure the required number of times in the course of the experiments. Films in which a single regular domain structure developed over the entire sample area were studied. This made it possible to eliminate effects related to the development of block domain structures.[12] The propagation of magnetostatic waves in reproducible domain structures of this type will be examined in this paper.

When the yttrium iron garnet films were magnetized along the [110] axis by a uniform magnetic field $\mathbf{H}_0$, a regular domain structure developed in them. Depending on the magnitude of the field, the nature and parameters of this domain structure varied as follows (see Figs. 1 and 2):

(1) for $\simeq 0<H_0<H_{s-z}$ the films had a linear strip sym-

TABLE I.

	Value			
Parameter	No. 1	No. 2	No. 3	Error
$4\pi M_0$, G	1781	1787	1890	±20
$2\Delta H$, Oe	0.6	0.75	0.5	
s, μm	7.9	11.5	8.9	±0.1
H_c, Oe	−84	−83	−76	±4
$\theta°$	1.5	3.8	4.9	±1
$\varphi_a°$	−92.5	−107.6	−17.0	±2.5
H_{s-z}, Oe	28	7	4	±0.5
H_{z-n} Oe	35	13.5	6.5	±0.5
H_{sat} Oe	62	35.5	24	±1
$H_s^{\perp}$, Oe	39.3	14	7	±0.5
$H_{\mathrm{sat}}^{\perp}$, Oe	69.5	55.5	29	±1

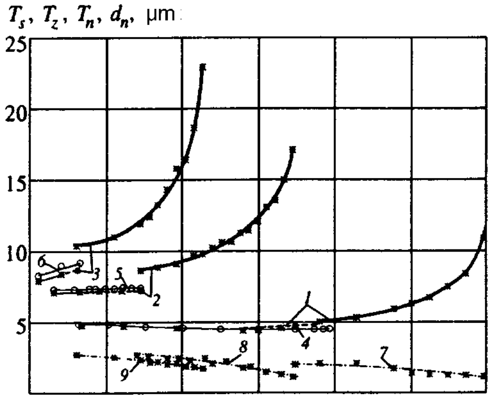


FIG. 2. Dependence on H_0 of the periods T_s of a linear strip symmetric domain structure (smooth curves), T_z of a zigzag strip symmetric structure (dashed curves), and T_n of a linear strip nonsymmetric domain structure (thick curves) and of the width d_n of the narrow domains of a nonsymmetric strip structure (dot-dashed curve) for films No. 1–No. 3 (* field H_0 applied along the [110] axis of the film, ○ field H_0 applied perpendicular to the [110] axis of the film). Curves *1*, *4*, and *7* are for film No. 1; curves *2*, *5*, and *8* for film No. 2; and, curves *3*, *6*, and *9* for film No. 3.

metric structure; neighboring domains had the same width and were oriented along the field direction;

(2) for $H_{s-z}<H_0<H_{z-n}$ the films had a zigzag strip symmetric structure; neighboring domains had the same width and were oriented at angles of $\pm 30°$ to the field, i.e., along directions specified by the projections of [111] axes on the plane of the film; here the length of the straight segment of the domains (from the point where the orientation of the domain segments relative to $\mathbf{H}_0$ changes from $-30°$ to $+30°$ to the point where it changes from $+30°$ to $-30°$) exceeded the domain width by a factor of 10–20;

(3) for $H_{z-n}<H_0<H_{\mathrm{sat}}$ the films had a linear strip nonsymmetric structure in which neighboring domains had significantly different widths and were oriented along the field; and,

(4) for $H_0>H_{\mathrm{sat}}$ there was no domain structure in the films and the films were in a saturated state.

The fields H_{s-z}, H_{z-n}, and H_{sat} for the transitions from one state of magnetic ordering to another for yttrium iron garnet films Nos. 1–3 are listed in Table I, while the periods of the symmetric T_s, zigzag T_z, and nonsymmetric T_n domain structures and the domain widths d_n of the nonsymmetric domain structure in these films can be judged from Fig. 2, which shows plots of $T_s(H_0)$, $T_z(H_0)$, $T_n(H_0)$, and $d_n(H_0)$. We note that the contrast of the nonsymmetric domain structure observed with the microscope in the interval $H_{z-n}<H_0<H_{\mathrm{sat}}$ (especially near H_{sat}) decreased significantly as the field was raised; this is evidence of an increase in the component of the domain magnetization parallel to the field.

When the films were magnetized perpendicular to the $[1\bar{1}0]$ axis by a uniform magnetic field $\mathbf{H}_0$, the nature and parameters of the domain structure varied with the magnitude of the field in the following way (see Figs. 1 and 2):

(1) for $\simeq 0<H_0<H_s^{\perp}$ the films had a linear strip symmetric structure; neighboring domains had the same width and were oriented along the field;

(2) For $H_s^{\perp}<H_0<H_{\mathrm{sat}}^{\perp}$ no domain structure was observed in the films, although the films were not in a saturated state over the entire range of fields;

(3) for $H_0>H_{\mathrm{sat}}^{\perp}$ there was no domain structure in the films and they were in a saturated state.

When characterizing the magnetic state of the films in the above field intervals, one should note that for $\simeq 0<H_0<H_s^{\perp}$ the period of the regular strip symmetric domain structure $T_s^{\perp}$ differs little from the period T_s when the $[1\bar{1}0]$ axis is oriented parallel to H_0 (see Fig. 2). Here $H_s^{\perp}$ was always slightly greater than H_{z-n} for the same film. As H_0 was increased, the contrast in the symmetric domain structure observed with the microscope gradually fell and for $H_0 \geqslant H_s^{\perp}$ no domain structure could be seen.

The field $H_{\mathrm{sat}}^{\perp}$ at which the films reached saturation magnetization could not be determined using the Faraday effect. This field could be evaluated only indirectly from measurements of the equiphase curves for the magnetostatic waves. This will be described below.

4. PROPAGATION OF SURFACE MAGNETOSTATIC WAVES IN YTTRIUM IRON GARNET FILMS WITH A STRIP DOMAIN STRUCTURE

In the following we describe the parameters and characteristics of magnetostatic waves in yttrium iron garnet film No. 1 for wave propagation perpendicular to $\mathbf{H}_0$ (transducers oriented parallel to the field $\mathbf{H}_0$). Then surface magnetostatic waves are excited in the film in the saturated state. The distinctive features of and differences in the propagation of magnetostatic waves in the other films will be noted in the course of the discussion.

First we consider the propagation of magnetostatic waves when the films were magnetized by a field $\mathbf{H}_0$ along the $[1\bar{1}0]$ axis.

For fields $H_0<H_{s-z}$, with a linear strip symmetric domain structure in film No. 1 ($H_{s-z}=28$ Oe) three excitation regions for magnetostatic waves were observed: S_{s1}, S_{s2}, and S_{s3} at frequencies of 150–500 MHz. These regions, together with plots of the low frequency limits f_{s1}, f_{s2}, and f_{s3} for the magnetostatic waves in each region as a function of the field H_0 are shown in Fig. 3a. The high frequency limits of the regions where magnetostatic waves exist could generally be determined only very approximately, since their location is substantially determined by the ratio of the amplitudes of the signal and electromagnetic pickup. (These limits are not indicated in the figures.) The $f_{s1}(H_0)$, $f_{s2}(H_0)$, and $f_{s3}(H_0)$ curves were measured from the phase–frequency characteristics by following the change with frequency in the constant value of the phase corresponding to wave numbers $k \simeq 0$. In a similar fashion, by varying the field and maintaining a fixed distance between the transducers, it is possible to follow the frequency variation of any other fixed value of the phase corresponding to another constant value of k from the phase–frequency characteristics. The curves $f_k(H_0)$ (in what follows we shall refer to them as the ''equiphase'' curves) obtained in this way for the regions S_{s1}, S_{s2}, and S_{s3} of film No. 1 are shown in Fig. 3a. Here the values of k measured by moving the detector transducer are written near the corresponding curves. (For clarity some values of k have not been written down.)

The minimum attenuation $L_{s1}^{\min}$ for the transmission coefficient in region S_{s1} for $H_0=6$ Oe was -36 dB, while the electromagnetic pickup level L_{ew} was $\simeq -55$ dB. When H_0 was raised to 18 Oe, $L_{s1}^{\min}$ increased to -29 dB and then, as the field was increased to H_{s-z}, decreased to -40 dB. In the region S_{s2}, which was observed for 24.3 Oe$<H_0<$26.4 Oe, the minimum attenuation in the amplitude–frequency characteristic was at a level of -46 dB, while in region S_{s3} when the field was changed from 6 Oe to 13.4 Oe it varied from -41 dB to roughly L_{ew}.

Dispersion curves for magnetostatic waves in regions S_{s1}, S_{s2}, and S_{s3} are shown in Fig. 4 (curves *1–4*). It is evident from Fig. 3a that the dispersion curve for region S_{s1} will have discontinuities for fields $H_0<17.5$ Oe (curve *3*, Fig. 4) and will not have them for 17.5 Oe$<H_0<$28 Oe (curve *4*, Fig. 4). The shapes of the observed curves correspond to the dispersion curves for surface magnetostatic waves. The surface character of the excited waves is also confirmed by the large difference in the damping of the

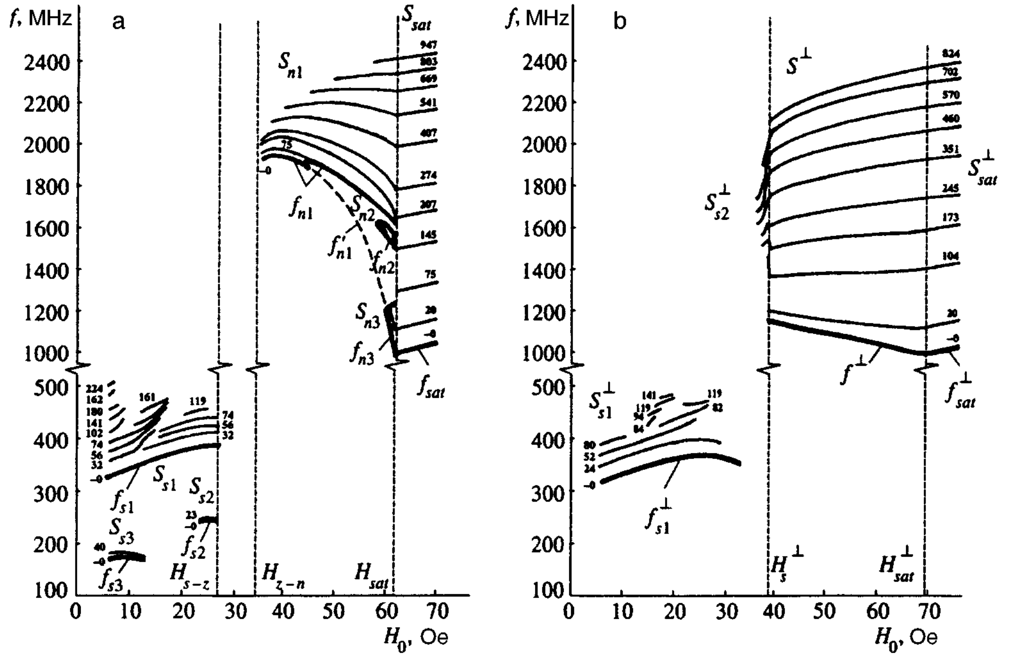


FIG. 3. The regions where surface magnetostatic waves exist and their equiphase curves for film No. 1 (the values of k in cm^{-1} are written next to each equiphase curve; the equiphase curve for $k \approx 182\ \mathrm{cm}^{-1}$ is shown within the region $S_{\underline{n2}}$); (a) field H_0 applied along the $(1\overline{1}0)$ axis of the film, (b) field $\underline{H_0}$ applied perpendicular to the $(1\overline{1}0)$ axis of the film. The frequency limits of the regions where the waves exist are denoted by heavy curves.

waves when their propagation direction is reversed. (Because of the nonreciprocity of surface magnetostatic waves, when the maximum amplitude of the wave is localized at the ferrite–substrate interface, the losses increase noticeably.)

An analysis of the amplitude- and phase–frequency characteristics of the excitations observed in regions S_{s1}, S_{s2}, and S_{s3} leaves no doubt that we have observed travelling waves. The nature of the magnetostatic waves in these regions may be related to collective oscillations of domain boundaries, as noticed previously.[6,7,10] In accordance with theoretical calculations[11] the domain boundaries have a resonant frequency in the range $\simeq 50-150$ MHz.

We note that in some yttrium iron garnet films the efficiency of exciting magnetostatic waves at fields $H_0 < H_{s-z}$ was lower than in film No. 1. For example, in film No. 2, we observed waves only with small wave numbers up to $\simeq 15\ \mathrm{cm}^{-1}$ in the $240-260$ MHz band.

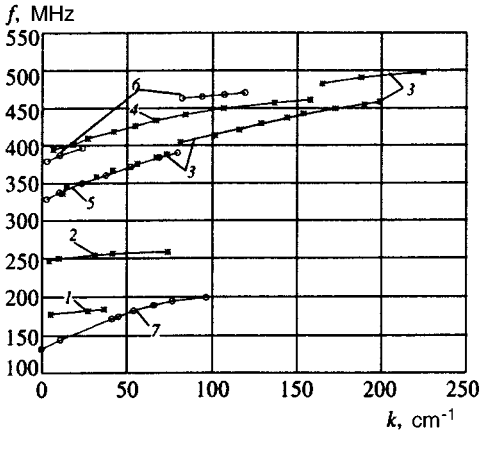


FIG. 4. Dispersion curves $f(k)$ for surface (curves *1–6*) and volume (curve *7*) magnetostatic waves with a linear strip symmetric domain structure in film No. 1 (* field H_0 applied along the $(1\overline{1}0)$ axis of the film, ○ field H_0 applied perpendicular to the $(1\overline{1}0)$ axis of the film): (*1*) for region S_{s3} with $H_0 = 6$ Oe; (*2*) for region S_{s2} with $H_0 = 26.5$ Oe; (*3*) and (*4*) for region S_{s1} with $H_0 = 6$ Oe and $H_0 = 26.5$ Oe; (*5*) and (*6*) for region $S^{\perp}_{s1}$ with $H_0 = 11$ Oe and $H_0 = 27.5$ Oe; (*7*) for region $V^{\perp}_{s2}$ with $H_0 = 39.3$ Oe.

When film No. 1 had a zigzag domain structure, in fields $H_{s-z} < H_0 < H_{z-n}$ (28 Oe$< H_0 <$35 Oe), no magnetostatic waves were observed, and this was typical of the other films that were studied.

When film No. 1 had a linear strip nonsymmetric domain structure, for fields $H_{z-n} < H_0 < H_{\mathrm{sat}}$ (35 Oe$< H_0 <$62 Oe), surface magnetostatic waves were also excited in the film. The amplitude–frequency characteristics of the transmission coefficient for these waves at different magnetization fields are shown in Fig. 5. For fields close to H_{z-n} the surface magnetostatic waves with a low frequency boundary f_{n1} in the spectrum are only slightly higher than the average electromagnetic pickup level L_{ew} (Fig. 5). As H_0 was raised the frequency bands and excitation efficiency of the surface magnetostatic waves gradually increased (Fig. 4a–d), while a narrow band in which surface magnetostatic waves were not excited appeared between the initial portion and the rest of the spectrum (Fig. 5c) and beginning at fields $H_0 \simeq 46$ Oe the initial portion of the spectrum was no longer observed. For 59 Oe$< H_0 <$62 Oe, there are two other regions in which surface magnetostatic waves exist: at frequencies of 1550–1620 MHz with an initial frequency f_{n2} and 1000–1200 MHz with an initial frequency f_{n3} (Fig. 5e). Thus, for $H_{z-n} < H_0 < H_{\mathrm{sat}}$ three regions in which magnetostatic waves exist were observed: S_{n1}, S_{n2}, and S_{n3}. The equiphase curves for these regions are shown in Fig. 3a. For $H_0 \geqslant H_{\mathrm{sat}} = 62$ Oe the film was magnetized to saturation and instead of the regions S_{n1}, S_{n2}, and S_{n3}, there was a spectrum of surface magnetostatic waves characteristic of a saturated film (Fig. 5f; the region S_{sat} of Fig. 3a) occupying the band 990–2400 MHz, in agreement with theory.[19] Note that, despite a break at the point $H_0 = H_{\mathrm{sat}}$, the equiphase curves for regions S_{n1} and S_{n3} (see Fig. 3a) organically join the equiphase curves

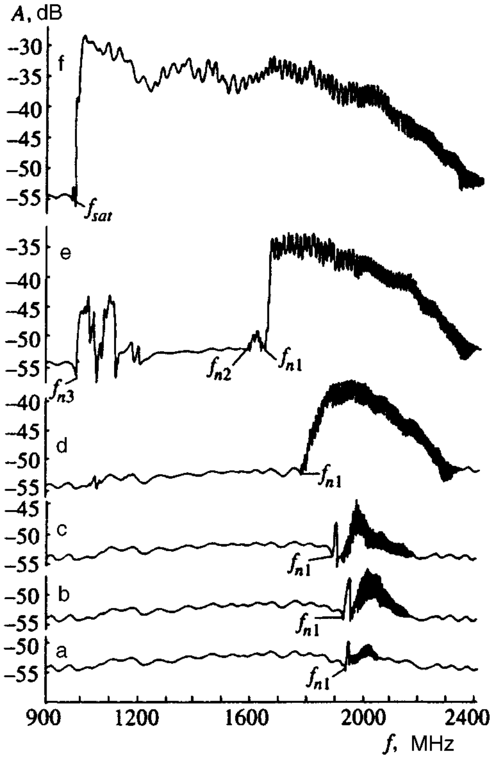


FIG. 5. Amplitude–frequency characteristics of the transmission coefficient for a microwave setup with a distance between transducers of 15 mm and the following values of the applied field H_0 (Oe): (a) 35.7, (b) 40.5, (c) 45.2, (d) 54.9, (e) 61.0, and (f) 62.2.

for the region S_{sat}, while the equiphase curves for region S_{n2} do not merge at the point $H_0 = H_{\mathrm{sat}}$ with the analogous curves for the region S_{sat}. Thus, the magnetostatic waves observed in regions S_{n1} and S_{n3} transform organically into waves that are characteristic of a saturated film, while the magnetostatic waves observed in region S_{n2} simply cease to be excited for $H_0 \geqslant H_{\mathrm{sat}}$.

Figure 6 shows the dispersion curves $f(k)$ for region S_{n1} (curves *15*), region S_{n2} (curve *6*), and region S_{sat} (curve *7*). The dispersion curves for region S_{n1} become steeper as H_0 increases; this is apparently a consequence of the increase in the projection of the average magnetization of the film parallel to the field. Unlike the dispersion curves $f(k)$ for a saturated film, in region S_{n1} at fields 45 Oe$< H_0 <$62 Oe (curves *3–5*, Fig. 6) and in region S_{n2} (curve *6*, Fig. 6) the $f(k)$ curves for films with a nonsymmetric domain structure begin not at $k \simeq 0$, but at some k_{n1} and k_{n2} which depend on H_0. In other words, the low frequency limits of the surface magnetostatic wave spectrum for regions S_{n1} and S_{n2} (the f_{n1} and f_{n2} curves in Fig. 3a) correspond to waves with $k \neq 0$. (The f_{n1} curve coincides with the equiphase curve for $k \simeq 0$ only for 35 Oe$< H_0 <$45 Oe.) Approximating the $f(k)$ curve for region S_{n1} (curves *3–5*, Fig. 6) to its intersection with the ordinate f, we obtain the frequencies f'_{n13}, f'_{n14}, f'_{n15}, ... in the surface magnetostatic wave spectrum which would correspond to waves with $k \simeq 0$. The dashed curve f'_{n1} in Fig. 3a has been constructed from these values for the frequencies. The curves f'_{n1} and f_{n1} bound the region that corresponds to the experimentally unobserved waves with $k < k_{n1}$. k_{n1} varied over $\simeq$50–150 cm^{-1} for fields in the range $\simeq$45–60 Oe.

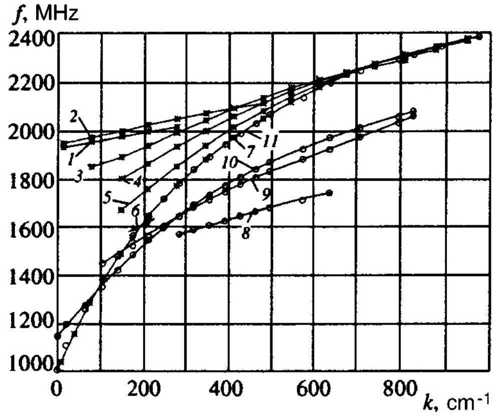


FIG. 6. Dispersion curves $f(k)$ for surface magnetostatic waves in film No. 1 for the regions S_{n1} (curves *1–5*), S_{n2} (curve *6*), S_{sat} (curve *7*, *), $S^{\perp}_{s2}$ (curves *8* and *9*), $S^{\perp}$ (curve *10*), and $S^{\perp}_{\mathrm{sat}}$ (curve *11*, ○) with the following values of the magnetic field H_0 (* field H_0 applied along the (110) axis of the film, ○ field H_0 applied perpendicular to the (110) axis of the film) (Oe): (*1*) 35.7, (*2*) 38, (*3*) 50.1, (*4*) 54.9, (*5*) 59.7, (*6*) 59.7, (*7*) (*) 62.2, (*7*) (○) 69.5, (*8*) 36.9, (*9*) 39.1, (*10*) 39.5.

In the other studies of yttrium iron garnet films the propagation of surface magnetostatic waves mostly retained the same character, but there were also differences. In many films, waves were not observed in the main region (type S_{n1}) immediately for $H_0 > H_{z-n}$, but with the appearance of a nonsymmetric domain structure in the film; thus, in film No. 2 magnetostatic waves appeared only for $H_0 > \simeq 19$ Oe and in film No. 3, for $H_0 > \cong 14$ Oe. In these films the equiphase curves for the S_{n1} type region looked the same as those for film No. 1, but without the initial segment of the curves lying between the straight lines $H_0 = H_{z-n}$ and $H_0 = 45$ Oe (i.e., in these films waves with small $k < k_{n1}$ were not excited in the S_{n1} type regions). In some films, often not just a single existence region of the type S_{n2} (which vanishes when the film is saturated) was observed to the left of the main region for the existence of magnetostatic waves, but two or three such regions distributed side by side, separated by frequency intervals in which there were no waves. The range of fields in which these regions were observed was, as a rule, wider for the other films than for the one described here (for No. 2 regions of this type were observed for 24 Oe $< H_0 <$35.5 Oe and for No. 3, for 14.5 Oe$< H_0 <$24 Oe), the minimum values of k (type k_{n2}) at which the dispersion curves for these regions originated always depended on H_0 and varied over 40–200 cm^{-1} for the different films, and the width of the range of wave numbers for these dispersion curves could be as high as 150 cm^{-1} (e.g., for No. 3). In exactly the same way, the range of fields within which magnetostatic waves were observed in regions of type S_{n3} was wider for some of the other films than for film No. 1. Thus, in film No. 3 waves were observed in a type S_{n3} region for 18 Oe$< H_0 <$24 Oe, with waves having 101 cm$^{-1} < k <$208 cm^{-1} being observed at $H_0 = 18$ Oe, waves having 65

$\text{cm}^{-1} < k < 179\ \text{cm}^{-1}$ being observed at $H_0 = 20.5$ Oe, and waves having $k < 120\ \text{cm}^{-1}$ being observed at $H_0 = 23.5$ Oe.

We now consider the propagation of magnetostatic waves when the films were magnetized by a field $\mathbf{H}_0$ perpendicular to the $[1\bar{1}\bar{0}]$ axis.

In fields $H_0 < H_s^{\perp}$ with a linear strip symmetric domain structure in film No. 1 ($H_s^{\perp} = 39.3$ Oe), two regions for excitation of surface magnetostatic waves were observed: $S_{s1}^{\perp}$ and $S_{s2}^{\perp}$ lying at frequencies of 300–500 MHz and 1500–2100 MHz, respectively. These regions, the low frequency limit $f_{s1}^{\perp}$ for the region $S_{s1}^{\perp}$, and the equiphase curves for the magnetostatic wave spectra in both regions are shown in Fig. 3b and the dispersion curves $f(k)$ are shown in Figs. 4 (for region $S_{s1}^{\perp}$) and 6 (for region $S_{s2}^{\perp}$. As can be seen from these figures, $f(k)$ for the region $S_{s1}^{\perp}$ has no discontinuities for fields of 5–15 Oe and 28–35 Oe (Fig. 4, curve *5*), while it does for fields 15 Oe$< H_0 <$28 Oe (Fig. 4, curve *6*). Surface magnetostatic waves cease to be excited in the region $S_{s1}^{\perp}$ for $H_0 \approx 35$ Oe. In fields of 35 Oe$< H_0 <$36.9 Oe magnetostatic waves are not excited at any frequency in the film, but for $H_0 = 36.9$ Oe they appear in region $S_{s2}^{\perp}$ at frequencies of 1570–1750 MHz. This new frequency range for the existence of surface magnetostatic waves expands quite rapidly and for $H_0 = H_s^{\perp} = 39.3$ Oe it extends over 1130–2120 MHz. At the same time there is a substantial reduction in the damping as these waves propagate. Thus, at a frequency $f = 1650$ MHz for $H_0 = 36.9$ Oe, the losses were -52 dB, while at $H_0 = 39.3$ Oe they were already -33 dB. It is clear from the dispersion curves $f(k)$ for the region $S_{s2}^{\perp}$ (curves *8* and *9* in Fig. 6) that magnetostatic waves with wave numbers k below a certain value $k_{s2}^{\perp}$ are not excited and that $k_{s2}^{\perp}$ tends to zero as the field approaches $H_s^{\perp}$ (see also the equiphase curves in Fig. 3b).

In film No. 1 a spectrum of surface magnetostatic waves is excited in the region $S^{\perp}$ in fields $H_s^{\perp} < H_0 < H_{\text{sat}}^{\perp}$ ($39.3 < H_0 < 69.5$ Oe) with a low frequency boundary $f^{\perp}$ (see Fig. 3b) which always corresponds to waves with $k \simeq 0$. The equiphase curves of the magnetostatic waves for this region merge smoothly (without discontinuities or abrupt changes in slope) with the equiphase curves of the waves in region $S_{s2}^{\perp}$ at $H_0 = H_s^{\perp}$ and they merge just as smoothly with the equiphase curves for the region $S_{\text{sat}}^{\perp}$ at $H_0 = H_{\text{sat}}^{\perp}$ (see Fig. 3b), which correspond to the ordinary surface magnetostatic waves excited, according to theory,[19] in a film that has been magnetized to saturation. The magnitude of the field $H_{\text{sat}}^{\perp}$ was determined using the equiphase curves for surface magnetostatic waves when the sign of the derivative of the equiphase curve for $k \simeq 0$ changes (curve $f^{\perp}$ in Fig. 3b). It was possible to determine $H_{\text{sat}}^{\perp}$ only within ± 1 Oe, since the sign change in the derivative takes place very smoothly. The dispersion relations $f(k)$ for the regions $S^{\perp}$ and $S_{\text{sat}}^{\perp}$ (for $H_0 = H_{\text{sat}}^{\perp}$) are shown in Fig. 6 (curves *10* and *11*). In accordance with the change in the character of the dispersion in the region $S^{\perp}$, as in the region $S_{s2}^{\perp}$, when H_0 increases the damping of the magnetostatic wave continues to decrease: thus, at a frequency of $f = 1650$ MHz for $H_0 = 39.3$ Oe the overall losses in the conversion and propagation of the magnetostatic waves were -33 dB and for $H_0 = 69.5$ Oe, — -26 dB.

In the other yttrium iron garnet films that were studied, the propagation of surface magnetostatic waves mostly retains the character described above. Thus, in some films with a symmetric domain wall ($H_0 < H_s^{\perp}$) the efficiency of wave excitation in the low frequency $S_{s1}^{\perp}$ type region was much lower: for example, in film No. 2 magnetostatic waves with $k < 15\ \text{cm}^{-1}$ were observed in the 235–255 MHz band with overall losses of $\simeq -53$ dB. In the high frequency $S_{s2}^{\perp}$ type region, magnetostatic waves were not excited at all in most of the test samples (for example, in sample No. 2). For $H_s^{\perp} < H_0 < H_{\text{sat}}^{\perp}$, in all the samples of yttrium iron garnet that were studied, the excitation of magnetostatic waves in the $S^{\perp}$ region was similar in character to that described above for film No. 1. However, we note that the equiphase curves for those films in which waves were not observed in the $S_{s2}^{\perp}$ type region did not have an inflection point (as in Fig. 3b) in the neighborhood $H_0 \simeq H_s^{\perp}$.

5. PROPAGATION OF VOLUME MAGNETOSTATIC WAVES IN YTTRIUM IRON GARNET FILMS WITH A STRIP DOMAIN STRUCTURE

We now estimate the parameters and characteristics of magnetostatic waves in yttrium iron garnet film No. 1 for the case in which the waves propagate parallel to $\mathbf{H}_0$ (transducers oriented perpendicular to the field $\mathbf{H}_0$). In the saturated state a backward volume magnetostatic wave is excited in the film. We note at once that, of all the films that were studied, magnetostatic waves were excited efficiently in this geometry only in film No. 1. In another two films the waves were hardly noticeable above the electromagnetic pickup level (-55 dB), so that measurements could not be made. In the remaining films magnetostatic waves were simply not excited.

First we consider the propagation of magnetostatic waves for the case in which the film is magnetized by a field $\mathbf{H}_0$ along the $[1\bar{1}\bar{0}]$ axis.

At fields $H_0 < H_{s-z}$, with a linear strip symmetric domain structure in film No. 1 magnetostatic waves were observed at frequencies of 1300–2500 MHz. The region V_s where these waves exist, its low frequency limit f_s, and the equiphase curves are shown in Fig. 7a and the dispersion curves $f(k)$ are shown in Fig. 8 (curves *1* and *2*). The dispersion curves indicate that forward volume magnetostatic waves are excited, and this is also confirmed by the fact that when the direction of propagation of the wave is reversed the damping does not change. The excitation of forward volume magnetostatic waves when there is a symmetric domain structure in the film is apparently caused by the quite large projection of the magnetization vector along the normal to the plane of the film in domains with both signs. As can be seen from Fig. 8, the dispersion curve for $H_0 = 6.3$ Oe (curve *1*) is interrupted; the spectrum of the excited waves looks like frequency bands with fairly efficient wave excitation alternating with bands in which waves are not excited. As the field H_0 is increased, the magnetostatic waves with small k gradually ceased to be excited, while the excitation bands merged. For $H_0 = 27$ Oe the dispersion curve was already continuous.

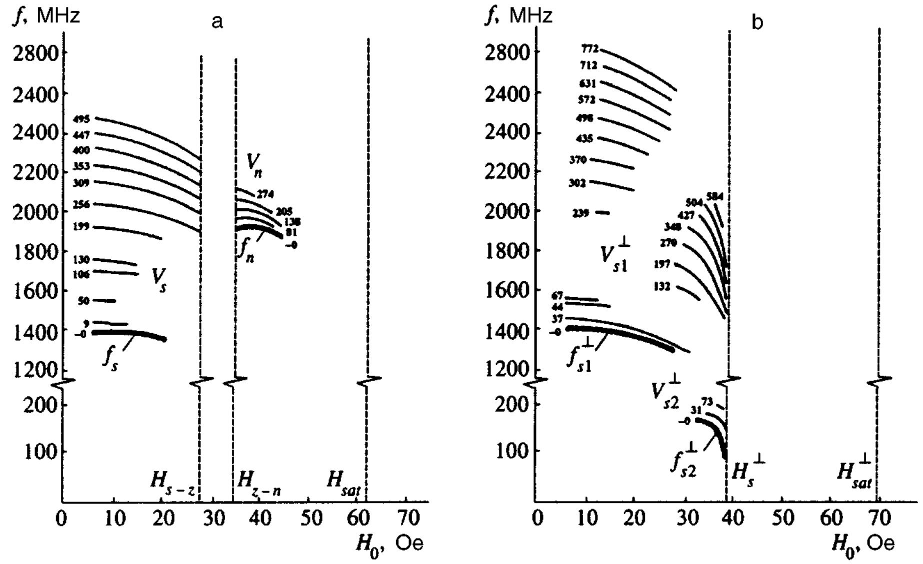


FIG. 7. The regions where forward volume magnetostatic waves exist and their equiphase curves for film No. 1 (values of k in cm^{-1} are written next to each equiphase curve) for (a) the field H_0 applied along the $(1\bar{1}0)$ axis of the film and (b) the field H_0 applied perpendicular to the $(1\bar{1}0)$ axis of the film. The frequency limits of the existence regions for the magnetostatic waves are given as thick curves.

When a zigzag domain structure was present in the film, magnetostatic waves were not excited.

When the film contained a nonsymmetric structure, excitation of forward volume magnetostatic waves was again observed at 1900–2200 MHz, but only for fields 35 Oe $<H_0<46$ Oe (saturation took place at $H_0 \simeq 62$ Oe). The region V_n where the waves exist, its low frequency limit f_n, and the equiphase curves are shown in Fig. 7a and the dispersion curves $f(k)$, in Fig. 8 (curves *3* and *4*). The dispersion curves in the frequency range occupied by magnetostatic waves were much less steep than the analogous parameters for forward volume magnetostatic waves excited in the film with a symmetric domain structure. Beginning at fields $H_0 \simeq 40$ Oe, the efficiency of exciting forward volume magnetostatic waves with wave numbers $\simeq 50\ \mathrm{cm}^{-1} < k < \simeq 100\ \mathrm{cm}^{-1}$ deteriorates significantly (for the same k and the same fields as for the propagation of surface magnetostatic waves with magnetization of the film along the $[1\bar{1}0]$ axis). Thus, the dispersion curves $f(k)$ for $H_0 > 40$ Oe (curve *4*, Fig. 8) have a discontinuity. The convergence of the equiphase curves with increasing field for the region V_n in Fig. 7a is apparently caused by the gradual disappearance of the component of the average magnetization perpendicular to the film surface.

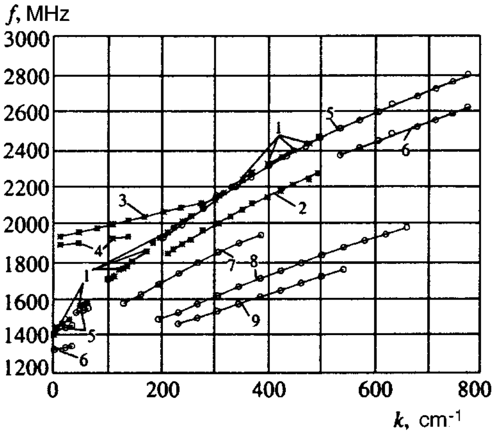


FIG. 8. The dispersion relations $f(k)$ for forward volume magnetostatic waves in film No. 1 for regions V_s (curves *1* and *2*), V_n (curves *3* and *4*), and $V_{s1}^{\perp}$ (curves *5*–*9*) with the following applied fields H_0 (Oe) (* field H_0 applied along the $(1\bar{1}0)$ axis of the film, ○ field H_0 applied perpendicular to the $(1\bar{1}0)$ axis of the film): (*1*) 6.2, (*2*) 27.3, (*3*) 35.7, (*4*) 44, (*5*) 13.2, (*6*) 26.3, (*7*) 32.2, (*8*) 37.4, and (*9*) 38.3.

At saturation ($H_0 > H_{\mathrm{sat}}$) backward volume magnetostatic waves were excited in the region V_{sat} in accordance with theory[19] (not shown in the figure).

We now consider the propagation of volume magnetostatic waves when the films are magnetized by a field $\mathbf{H}_0$ perpendicular to the $[1\bar{1}0]$ axis.

In fields $H_0 < H_s^{\perp}$ when film No. 1 had a linear strip symmetric domain structure two regions for excitation of magnetostatic waves, $V_{s1}^{\perp}$ and $V_{s2}^{\perp}$ were observed: one in the band 1300–2850 MHz and the other in the band 100–200 MHz. The equiphase curves and low frequency limits $f_{s1}^{\perp}$ and $f_{s2}^{\perp}$ for both regions where these wave exist are shown in Fig. 7b and the dispersion curves $f(k)$ are shown in Fig. 8 (for the region $V_{s1}^{\perp}$) and Fig. 4 (for the region $V_{s2}^{\perp}$). These dispersion curves indicate the excitation of forward volume magnetostatic waves in both regions. For fields $H_0 < 32$ Oe the dispersion curves $f(k)$ have discontinuities (curves *5* and *6* of Fig. 8), while for $H_0 > 32$ Oe and for the region $V_{s2}^{\perp}$ there are no discontinuities in the $f(k)$ curves (Fig. 8, curves *7*–*9*, and Fig. 4, curve *7*). With increasing H_0 the steepness of the $f(k)$ curves always decreases. Near $H_0 \simeq 27$ Oe the excitation of magnetostatic waves in the film ceases completely (except for a narrow band segment near 1350 MHz with $k < 50\ \mathrm{cm}^{-1}$). In fields $H_0 > 32$ Oe in region $V_{s1}^{\perp}$ magnetostatic waves with $k < k_{s1}^{\perp}$ are not excited with $k_{s1}^{\perp}$ rising from 130 to 240 cm^{-1} as the field is increased. In region $V_{s2}^{\perp}$, on the other hand, magnetostatic waves with $k > k_{s2}^{\perp}$ are not excited with $k_{s2}^{\perp}$ rising from 20 to 100 cm^{-1} as the field is increased.

In the second field interval, $H_s^{\perp} < H_0 < H_{\mathrm{sat}}^{\perp}$, magnetostatic waves were not excited in the film.

6. ANALYSIS OF EXPERIMENTAL DATA

In comparing the experimental data obtained for magnetization of yttrium iron garnet film No. 1 parallel and perpendicular to the [110] axis, it should be noted that the equiphase and dispersion curves for surface magnetostatic waves in regions S_{s1} and $S_{s1}^{\perp}$ are substantially the same in character for $H_0 < H_{s-z} = 28$ Oe: the equiphase curves for the same k (when $k < 80\ \text{cm}^{-1}$) with the same fields H_0 have essentially the same steepness and lie at the same frequencies (see Figs. 3a and 3b). The same can also be said about the characteristics of the forward volume magnetostatic waves in regions V_s and $V_{s1}^{\perp}$ when $H_0 < H_{s-z} = 28$ Oe (see Figs. 7a and 7b). Since the dependences of the periods T_s and $T_s^{\perp}$ of the domain structures also are essentially the same for $H_0 < H_{s-z}$ (see Fig. 2), we may conclude that the local magnetizations $\mathbf{M}_1$ and $\mathbf{M}_2$ of both kinds of domains in a symmetric domain structure have essentially the same projections on the normal to the film surface and on the direction of $\mathbf{H}_0$, for both parallel and perpendicular orientations of the [110] axis of the film relative to H_0 and, apparently, that the vectors themselves are identical: $\mathbf{M}_1^{\parallel} = \mathbf{M}_1^{\perp}$ and $\mathbf{M}_2^{\parallel} = \mathbf{M}_2^{\perp}$ for $H_0 < H_{s-z}$.

In order to explain the differences which are nevertheless observed in the characteristics of the magnetostatic waves for perpendicular and parallel orientations of the [110] axis of the film relative to $\mathbf{H}_0$ (such as a slight difference in the behavior of the equiphase curves of the surface magnetostatic waves for $k > 80\ \text{cm}^{-1}$, the existence of additional regions S_{s2} and S_{s3} when the [110] axis is parallel to $\mathbf{H}_0$, etc.), we shall examine how the magnetic ordering of the film changes as it is magnetized. The magnetization of the film, which has an easy axis anisotropy normal to the plane, can be viewed simply as a change in the orientation of the magnetizations of the domains, $\mathbf{M}_1$ and $\mathbf{M}_2$, from a position parallel to the easy axis ($\mathbf{M}_1$ directed along $\mathbf{n}$ and $\mathbf{M}_2$ in the opposite direction from $\mathbf{n}$) $\mathbf{H}_0$ to a position parallel to the applied field ($\mathbf{M}_1 \| \mathbf{H}_0$ and $\mathbf{M}_2 \| \mathbf{H}_0$). In the first case, when the [110] axis is perpendicular to $\mathbf{H}_0$, the vectors $\mathbf{n}$ and $\mathbf{H}_0$ and the [111] axis (the orientation of the vectors $\mathbf{M}_1$ and $\mathbf{M}_2$, along which the magnetic energy of a yttrium iron garnet crystal is reduced) lie in a single plane, so that all the changes in $\mathbf{M}_1$ and $\mathbf{M}_2$ as $\mathbf{H}_0$ is increased will also take place in a single plane. In the second case, when the [110] axis is parallel to $\mathbf{H}_0$, the vectors $\mathbf{n}$ and $\mathbf{H}_0$ and the type [111] axis do not lie in a single plane. Within a certain range of fields $0 < H_0 < H_{s-z}$, however, the change in the vectors $\mathbf{M}_1$ and $\mathbf{M}_2$ as $\mathbf{H}_0$ is increased nevertheless takes place in a plane passing through $\mathbf{n}$ and $\mathbf{H}_0$. This also explains the coincidence of the characteristics of the magnetostatic waves in this range of fields. Nevertheless, the closeness of the [111] axes to a plane passing through $\mathbf{n}$ and $\mathbf{H}_0$ may lead to the formation of domains of closure that cannot be seen in the microscope and have magnetizations $\mathbf{M}_3$ and $\mathbf{M}_4$ oriented along the [111] axes nearest to the direction of $\mathbf{H}_0$ and, apparently, are the reason for the differences observed in the characteristics of the magnetostatic waves in the first and second cases. As H_0 is increased, the vectors $\mathbf{M}_1$ and $\mathbf{M}_2$, which lie in a plane passing through $\mathbf{n}$ and $\mathbf{H}_0$, come ever closer to the directions specified by the [111] axes nearest to the direction of $\mathbf{H}_0$. When $H_0 = H_{s-z}$ it becomes favorable for them to orient themselves along these axes; the domain boundaries also orient themselves along these directions and a zigzag domain structure develops in the film for $H_{s-z} < H_0 < H_{z-n}$. For $H_{z-n} < H_0 < H_{\text{sat}}$ the vectors $\mathbf{M}_1$ and $\mathbf{M}_2$ appear to rotate again in the plane passing through $\mathbf{n}$ and $\mathbf{H}_0$, but because of the higher H_0 a nonsymmetric domain structure develops in the film.

The above remarks implies that only for fields $H_0 < H_{s-z}$ can the characteristics of the magnetostatic waves for the cases when the [110] axis is perpendicular and parallel to $\mathbf{H}_0$ coincide to a substantial degree.

The above discussion of the change in the character of the magnetic ordering in yttrium iron garnet films is basically hypothetical and does not explain the many experimental results described earlier. In particular, for example, it is unclear why the initial frequencies of the spectra of surface magnetostatic waves in a film with a nonsymmetric domain structure are much higher than in a film that has been magnetized to saturation.

It is easy to show that these data cannot be explained in terms of a theory which employs averaging of the magnetic permeability tensor over all the domains. (See Refs. 1, 3, and 5–7, for example.) With this approach, operating with the average effective magnetization and average effective magnetizing field, we can use the standard dispersion relations for magnetostatic waves. With the aid of these, using approximation methods, we can obtain good agreement between the calculated and experimental curves. It is quite simple to guess the shortcomings of this approach: estimates for surface magnetostatic waves show that an initial frequency for the spectrum on the order of 1500–2000 MHz (typical of many experimental dependences given in this paper) at an average magnetization of less than 1750 G (averaging over all the domains cannot yield a value exceeding $4\pi M_0$ for pure saturated yttrium iron garnet) can be obtained only with the aid of fields H_0 of magnitude 300–500 Oe and not the 5–70 Oe used in the experiments, while in order to make the width of the spectrum of the volume magnetostatic waves as large as in curves *1* and *5* of Fig. 8, the ''average effective magnetization'' should be on the order of 6000 G. Thus, although models employing ''averaging'' apparently can yield fair agreement with experiment, they cannot provide a reasonable physical interpretation of the resulting average effective magnetization and field or explain the impossibility of exciting magnetostatic waves when a zigzag domain structure exists in the films, the absence of waves in certain ranges of wave vectors k, the discontinuous character of the spectra, and the simultaneous excitation of magnetostatic waves at different frequencies. (In Fig. 7b magnetostatic waves are excited at frequencies of 100–200 MHz and 1500–2100 MHz for 33 Oe $< H_0 <$ 39 Oe.)

7. CONCLUSIONS

An experimental study has been made of wave processes in tangentially magnetized unsaturated films of yttrium iron garnet with regular symmetric, nonsymmetric, and zigzag strip domain structures. The waves observed at 800–3000 MHz are similar to volume and surface magnetostatic waves in saturated films, propagate with damping characteristic of the latter, and have wavelengths much greater than the period of the domain structure.

The experimentally studied characteristics of the magnetic waves are not explained either by a theory employing averaging of the magnetization over all the domains or with the aid of the information on the film parameters and on the nature and parameters of the domain structure given in Figs. 1 and 2 and in Table I. (For example, it is not possible to find any correlation between the period of the structure and the magnitudes of the wave vectors of these waves, which are not excited.)

In addition, low-frequency (100–500 MHz) branches of the wave oscillations are observed which appear to be related to collective resonant oscillations of the domain boundaries.

We thank V. I. Zubkov and Yu. A. Filimonov for useful discussions of this work.

This work was supported by the Russian Fund for FundamentalResearch (Grant No. 96-02-17283a).

[1] I. A. Gilinskiĭ and R. G. Mints, Zh. Éksp. Teor. Fiz. **59**, 1230 (1970) [Sov. Phys. JETP **32**, 673 (1971)].

[2] I. A. Gilinskiĭ and K. A. Ryazantsev, Fiz. Tverd. Tela **16**, 3008 (1974) [Sov. Phys. Solid State **16**, 1944 (1974)].

[3] D. D. Stancil, J. Appl. Phys. **56**, 1775 (1984).

[4] I. V. Zavislyak and V. V. Danilov, Pis'ma Zh. Tekh. Fiz. **8**, 72 (1982) [Sov. Tech. Phys. Lett. **8**, 31 (1982)].

[5] S. A. Vyzulin, S. A. Kirov, and N. E. Syr'ev, Vestnik MGU, ser. fizika i astronomiya **24**, 92 (1983).

[6] S. A. Vyzulin, S. A. Kirov, and N. E. Syr'ev, Vestnik MGU, ser. fizika i astronomiya **25**, 70 (1984).

[7] S. A. Kirov, A. I. Pil'shchikov, and N. E. Syr'ev, Fiz. Tverd. Tela **16**, 3051 (1974) [Sov. Phys. Solid State **16**, 1970 (1974)].

[8] L. V. Mikhaĭlovskaya and I. V. Bogomaz, Fiz. Tverd. Tela **19**, 1245 (1977) [Sov. Phys. Solid State **19**, 725 (1977)].

[9] V. I. Kostenko and M. A. Sigal, Phys. Stat. Sol. B **170**, 569 (1992).

[10] G. T. Kazakov, A. G. Sukharev, and Yu. A. Filimonov, in *Abstracts of Talks at the V-th All-union School on Spin-wave Microwave Electronics [in Russian]*, Zvenigorod (1991), p. 83.

[11] A. G. Gurevich and G. A. Melkov, *Magnetization Oscillations and Waves*, CRC Press, Boca Raton, FL (1996).

[12] Yu. V. Gulyaev, P. E. Zil'berman, G. T. Kazakov *et al.*, Pis'ma Zh. Tekh. Fiz. **11**, 97 (1985) [Sov. Tech. Phys. Lett. **11**, 38 (1985)].

[13] G. E. Khodenkov, Fiz. Met. Metalloved. **75** (5), 5 (1993).

[14] A. V. Mikhailov and I. A. Shimokhin, Phys. Rev. B **48**, 9569 (1993).

[15] B. N. Fillipov and A. P. Tankeev, *Dynamic Effects in Ferromagnetics with a Domain Structure* [in Russian], Nauka, Moscow (1987).

[16] A. V. Vashkovskiĭ, É. G. Lokk, and V. I. Shcheglov, JETP Lett. **63**, 572 (1996).

[17] A. V. Voronenko, S. V. Gerus, and L. A. Krasnozhen, Mikroélektronika **18**, 667 (1977).

[18] E. V. Lebedeva, A. I. Pil'shchikov, and N. S. Sedletskaya, Zh. Éksp. Teor. Fiz. **72**, 667 (1977) [Sov. Phys. JETP **45**, 350 (1977)].

[19] R. W. Damon and J. R. Eshbach, J. Phys. Chem. Sol. **19**, 308 (1961).

Translated by D. H. McNeill

Dynamic and stimulated rf echo in high-T_c superconducting powders

A. V. Drobinin and Yu. M. Tsipenyuk

P. L. Kapitza Institute of Physics Problems, Russian Academy of Sciences, 117973 Moscow, Russia

(Submitted 10 July 1996)
Zh. Éksp. Teor. Fiz. **111**, 1032–1046 (March 1997)

Two-pulse and three-pulse echoes in powdered yttrium and bismuth high-T_c superconductors are investigated to determine the dependence of the signal amplitude on the magnetic field, the temperature, and the gas pressure. The temperature is measured as a function of the relaxation time of the echo signal. The properties of the long-lived rf echo are studied in detail; it exhibits a persistent (lasting more than several hours) memory of a time series of write pulses and a cumulative storage effect. The experimental results can be explained qualitatively within the framework of the theory proposed by Asadullin [Sverkhprovodimost' **6**, 545 (1993)] to account for the nonlinear motion of vortices associated with sample defects. © *1997 American Institute of Physics.* [S1063-7761(97)02203-8]

1. INTRODUCTION

Radio-frequency echo in powders, also known as polarization, powder, or phonon echo, was discovered in the sixties and continues to intrigue both experimental and theoretical scientists to the present day. The echo was observed in powdered piezoelectrics, ferroelectrics, ferromagnets, and normal and superconducting metals (see the surveys in Refs. 1 and 2). By the nature of the dynamics of these objects, the rf echo effect is classified as an oscillator echo,[2,3] because it is a result of the dynamics of a system of nonlinear oscillators.

An interesting topic in its own right, polarization echo is also used as a method for the investigation of physical phenomena. It has far-reaching implications in radiospectroscopy, specifically in the coherent pulsed radiospectroscopy of defects in crystals and their powders. As one example, Romanov and Solovarov[4] have used rf echo to investigate the anomalous generation and conversion of acoustic modes in the vicinity of phase transitions in ferroelectric single crystals.

The sum total of experimental data indicates that in every case the phenomenon is attributable to acoustic vibrations of powder grains driven in resonance by rf pulses. Like all echo phenomena, powder echo is a nonlinear effect; however, the mechanism of the nonlinearity has its own character specific to each situation.

Immediately following the discovery of high-temperature superconductivity, rf echo was observed in high-T_c superconducting powders.[5–7] It is observed only in the superconducting phase, where the amplitude of the rf echo signal is three orders of magnitude times the amplitude of nuclear magnetic resonance (NMR) signals in these compounds (the echo is scarcely observable in the normal phase, since high-$\mathbf{T}_c$ superconductors are poor metals). Fundamental to the rf echo phenomenon is the question of what kind of mechanism underlies such strong enhancement of the effect in superconductors and what is the relationship between the superconducting and acoustical properties. In 1993 Asadullin proposed a theoretical description of rf echo in high-T_c superconductors, based on the nonlinear character of the dynamics of crystal structure defects.[8] This work lends qualitative insight into the principal attributes of rf echo observed in high-T_c superconducting powders.

Two-pulse and three-pulse rf echoes have been classified as dynamic echoes. In addition, a long-lived stimulated echo has been observed in powders of certain materials, where it occurs under the influence of a read pulse even after a pair of write pulses has terminated. This memory effect of a time series of exciting pulses is retained in the sample for a very long time (hours and even days).

Here we give the results of a detailed experimental study of dynamic and stimulated echo in yttrium and bismuth high-$\mathbf{T}_c$ superconductors; this work has been briefly reported in an earlier publication.[5]

2. EXPERIMENTAL APPARATUS

The observation of rf echo is methodologically similar to pulsed NMR experiments. The process is well known: The nuclear spin system is excited by a train of rf pulses with specially selected durations and delays between pulses. This event creates rotating magnetization in a coil, inducing a weak emf in it. To maximize the emf, the axis of the coil must be perpendicular to the external field. The emf has two components: a time-decaying free induction after each pulse and a subsequently emerging echo signal after the pulse train.

In the case of the rf powder echo signal as well, the application of essentially the same sequence of two pulses separated by a time interval τ causes an echo signal to appear in the sample at a time τ after the second pulse. The amplitude of this signal decays with time with a characteristic relaxation time of the order of a few tens of microseconds (by analogy with NMR, it is denoted by T_2, i.e., the longitudinal relaxation time), and when a third sensing (read) pulse is transmitted after a time much greater than T_2, a stimulated echo signal appears, again after a time τ. The fundamental difference of rf echo in powders from the NMR signal is its nonresonant character: For exciting pulses of a given frequency it occurs in any magnetic field.

1063-7761/97/030569-09$10.00

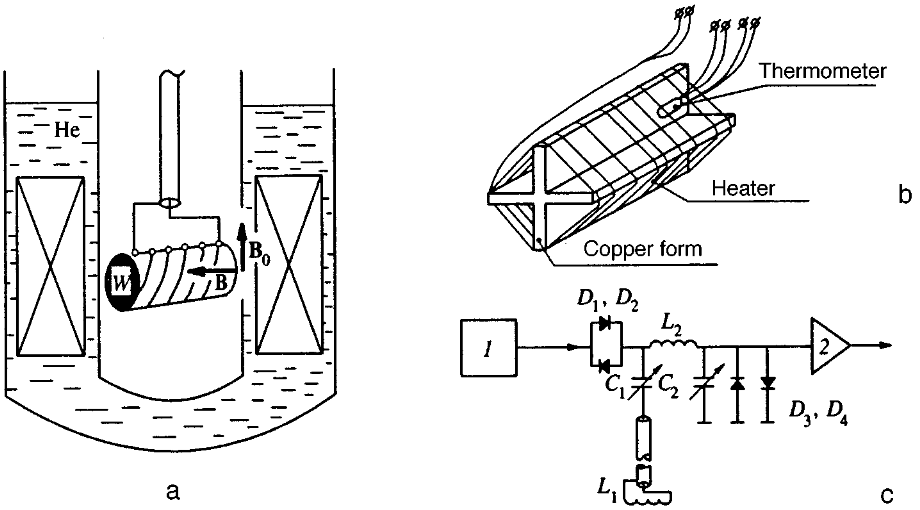


FIG. 1. Apparatus and circuit for the measurement of rf echo in high-T_c superconductors. a) Placement of the rf circuit in the cryostat; b) schematic view of the heater and attached thermometer; c) circuit schematic: 1) exciting pulse generator; 2) preamplifier.

The measurement arrangement is shown schematically in Fig. 1. A magnet with superconducting coils is used to generate a magnetic field. The rf coil is a hollow Teflon cylinder of length 22 mm and diameter 7 mm with 25 turns of 0.3-mm PEV wire wound around it. The axis of the coil is perpendicular to the external magnetic field and has an inductance $L_1 = 3\ \mu\text{H}$. The investigated powder is packed inside the coil. In addition to the powder, a Nichrome heater wound on a copper form made from two crossed copper plates of thickness 0.2 mm is also placed inside the coil to create a uniform temperature in the powder (see Fig. 1). The heater winding is bifilar. Also contained inside the coil is a semiconductor thermometer. The resistance of the thermometer is measured by a four-point technique with current reversal to compensate the thermal emf in the measurement wires. The coil with the sample enclosed is placed in a vessel filled with helium gas as a heat-transfer medium. The vessel wall and liquid helium are separated by a vacuum jacket to prevent the evaporation of helium and the formation of a temperature gradient in the vessel.

A sinusoidal signal from an rf oscillator is sent to an rf pulse shaper, whose second input receives square-wave pulses from a programmable pulse generator. The pulse shaper output signal consists of pulses of the same duration as the square-wave pulses but with an rf sine wave carrier. The pulses are sent from the shaper to an rf power amplifier, which energizes the excitation circuit. The divider diodes $D_{1,2}$ are open at the instant of transmission of the rf pulse, acquiring a voltage drop ~ 0.7 V. Since the capacitance C_2 of the sensing L_2C_2 circuit is shunted by the diodes $D_{3,4}$, the current in this circuit is determined by the reactance of the coil L_2. Consequently, most of the power from the power amplifier is transferred to the resonance-tuned exciting L_1C_1 circuit with $Q \simeq 25$. When the electric length of the connecting cable is set at $l = \lambda/4$, optimal conditions are created for transferring energy into the exciting circuit. The

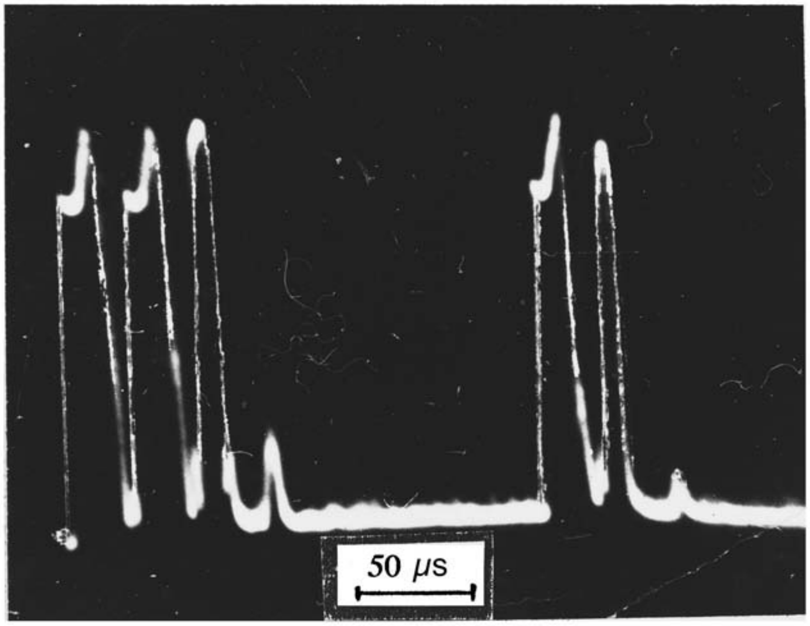


FIG. 2. Oscillogram of dynamic two-pulse and three-pulse echoes; not only the second, but also the third echo is visible.

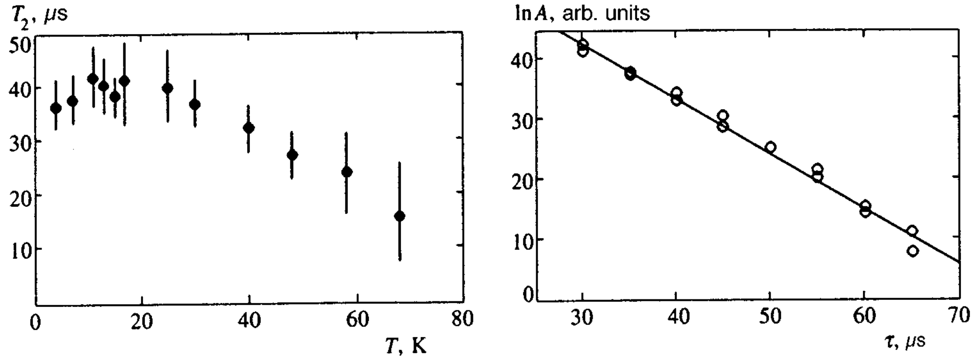


FIG. 3. Relaxation time T_2 of dynamic echo in YBaCuO powder vs temperature T ($H=25$ kOe) (left graph) and echo amplitude A vs time delay τ between exciting pulses for BiCaSrCuO ($T=7$ K, $H=25$ kOe) (right graph).

voltage across the coil L_1 attains 1 kV in this case, and the rf field is $H_{\sim}\simeq 25$ G.

The preamplifier and the receiver operate at saturation during transmission of the rf pulse. The decay of the rf voltage to the 0.7-V level requires a time of about 3 μs, after which the diodes $D_{1,2}$ close and disconnect the power amplifier. The preamplifier has a high-resistance input and a gain equal to 10. A YaS-54 spectrum analyzer is used as the rf oscillator and receiver. The amplifier and the receiver leave the saturation regime 10 μs after the pulse. Consequently, although not normally recommended, the time delay between exciting pulses is made shorter than this value. Without exciting pulses the elements L_1, C_1, and C_2 form a single series resonant circuit. The preamplifier and the receiver amplify the voltage formed across the capacitor C_2 under the influence of the sample-induced emf in the coil L_2.

The signal from the receiver is sent to a V9-5 gated voltmeter, whose gate is controlled by a programmable pulse generator. The gate is visually adjusted to the maximum of the echo signal by means of a dual-beam oscilloscope, whose two beams provide sweeps for the receiver signal and the gate pulse.

On the order of 10^6 ceramic grains of diameter ~ 100 μm were used to observe the echo in YBaCuO and BiCaSrCuO powders. The powder was prepared by crushing a polycrystalline sample and then screening fine grains from coarse grains by means of two metal screens with different mesh sizes.

No echo of any kind was observed for a powder sample containing $\sim 10^2-10^3$ grains of diameter 1 mm or smaller. Instead, a highly erratic "ringing" was visible after each rf pulse as a result of beats generated by different powder-grain oscillators.

3. EXPERIMENTAL RESULTS

When two successive rf pulses of frequency ~ 20 MHz, power ~ 100 W, and duration ~ 5 μs are fed into the resonant circuit containing the powder, a strong dynamic echo signal is observed. The rf frequency f is very close to the estimated acoustic resonance of the experimental powders. Indeed, if the characteristic grain diameter l spans one half-wavelength, then it satisfies $l=\lambda/2=s/2f$, where $s=(3-4)\cdot 10^3$ m/s is the speed of sound. At a frequency $f=2\cdot 10^7$ Hz the resonance grain diameter must in fact be of the order of 100 μm.

The echo amplitude is a maximum when the rf pulse lengths are equal, and its width increases as the duration of the sensing pulses increases. A secondary echo is observed for short time delays, as seen in Fig. 2.

As the time delay between pulses is increased, the echo amplitude decays with a time constant $T_2\simeq 35$ μs, which is approximately the same at temperatures from 4.2 K to ~ 30 K, and then T_2 begins to decrease (Fig. 3). It is evident from Fig. 3 that the decay of the echo amplitude with increasing time delay is exponential.

When a third pulse is sent to the input, a stimulated echo is observed, as seen in Fig. 2. Here the time delay between the first pair of write pulses and the third, read pulse can be much longer than the decay time constant T_2 of the two-pulse echo. The long-lived stimulated echo does not vanish even after the pair of write pulses has terminated.

Both the echo intensity and the relaxation time depend significantly on the pressure of the gas (helium in our experiments) used to fill the capsule containing the powder; the higher the gas pressure, the lower is the echo amplitude (Figs. 4 and 5). These measurements offer further evidence to the effect that the phenomenon is associated with acoustic vibrations of the sample powders.

Both the dynamic and the stimulated echo die out with increasing temperature and exist only in the superconducting

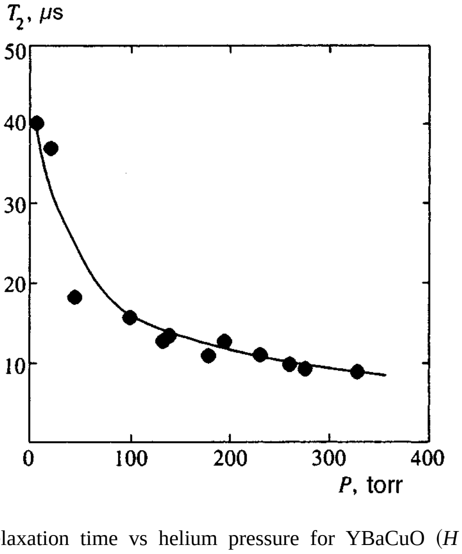


FIG. 4. Relaxation time vs helium pressure for YBaCuO ($H=25$ kOe, $T=4.2$ K).

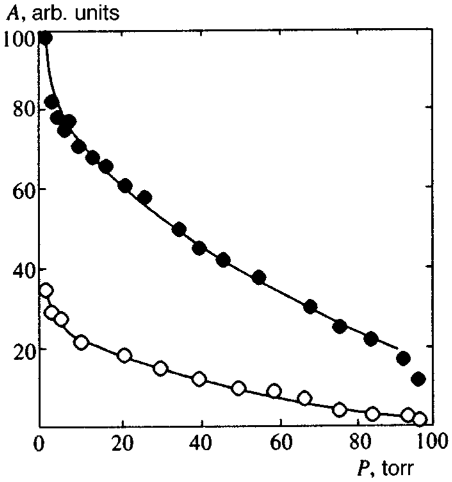


FIG. 5. Intensities of two-pulse (●) and three-pulse (○) dynamic echoes vs ambient pressure of the gas (helium).

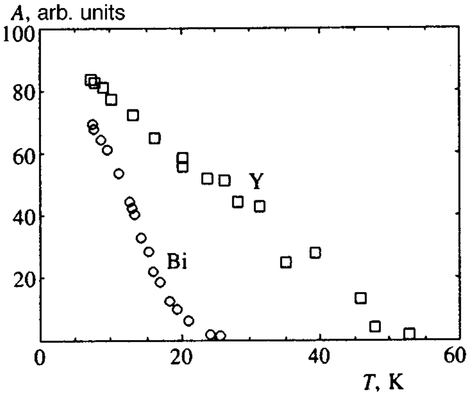


FIG. 6. Echo amplitude vs temperature for yttrium (□) and bismuth (○) samples.

phase. However, the temperature behavior of the echo amplitude differs for YBaCuO and BiSrCaCuO (Fig. 6), even though the superconducting transition temperatures of both samples have very nearly the same value, 90–92 K, a fact that has been verified both for the initial polycrystalline samples and for their powders.

Graphs of the echo amplitude as a function of the external magnetic field for the yttrium and bismuth high-T_c superconducting powders are shown in Fig. 7. For both systems the dependence is essentially linear, although for YBaCuO the amplitude begins to increase with increasing magnetic field at $\sim$3 kOe. Note that Petrov *et al.*[7] have obtained a quadratic field dependence of the echo amplitude, but their measurements were performed in a narrow range of magnetic fields, only up to 7 kOe.

A characteristic feature of the stimulated echo is the cumulative storage effect of repeated inputs of coherent pairs of rf write pulses. After a pair of write pulses is cut off, the amplitude of the long-lived echo decreases by one half in 10–20 s, then after a few minutes it settles into a constant level U_2 approximately to 0.4 times the initial value U_1, and it remains unchanged for at least several hours. When a sequence of two rf write pulses separated by a time interval $\tau < T_2$ is applied, the stimulated echo grows with the same time scales as shown in Fig. 8. It must be emphasized that the characteristic times of this process are not tens of microseconds like the typical relaxation time of the dynamic echo, but are measured in minutes, i.e., the time scale of this effect is 10^6 times greater.

More than one, in fact several time series can be written, as shown in Fig. 9, where the powder is initially exposed to a pair of rf pulses with a 30-μs time delay between them, and then the spacing of the pulses is approximately doubled. Two echo signals are observed after the read pulse. In this situation the write pulses become read pulses, as is patently evident in the given oscillogram. However, slight ($\sim$5%) changes in either the temperature or the external field almost completely erase the memory of the time series of write pulses, and the long-lived echo signal disappears.

The intensity of the stimulated echo, like that of the dynamic echo, decreases as the ambient gas pressure is increased, but at low pressures the quantities U_1 and U_2 exhibit altogether opposite behavior: Whereas U_1 exactly repeats the pressure dependence of the dynamic echo amplitude, U_2 increases at first but then begins to diminish as the gas pressure is increased (Fig. 10).

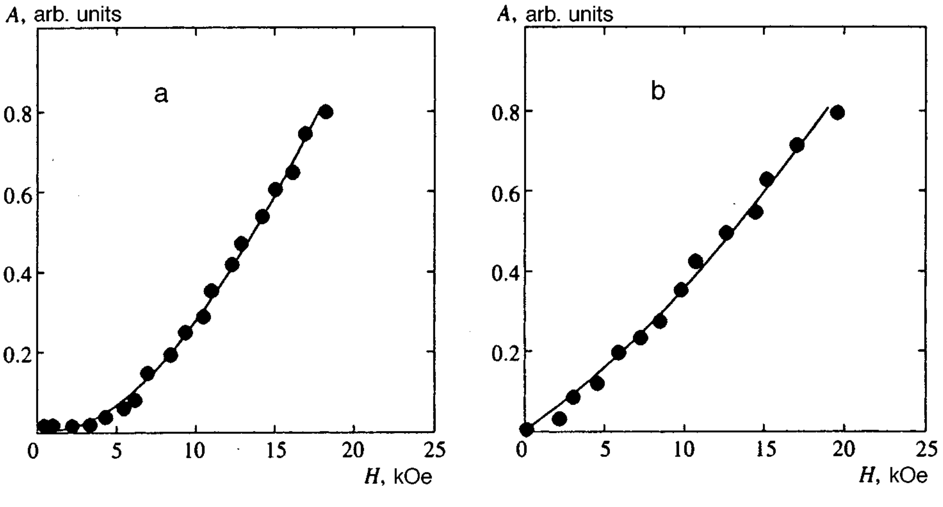


FIG. 7. Echo amplitude vs external magnetic field for yttrium (a) and bismuth (b) samples ($T = 4.2$ K).

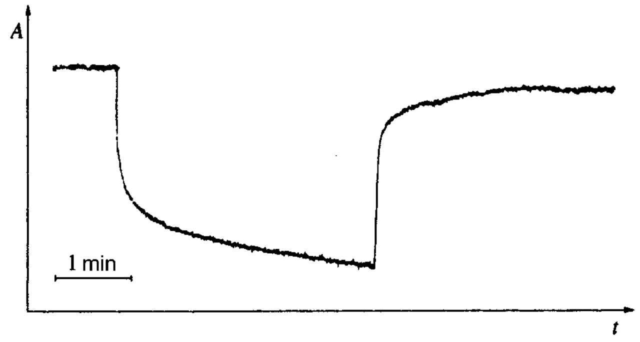


FIG. 8. Stimulated echo amplitude vs time after the input and cutoff of a pair of rf write pulses (cumulative storage effect).

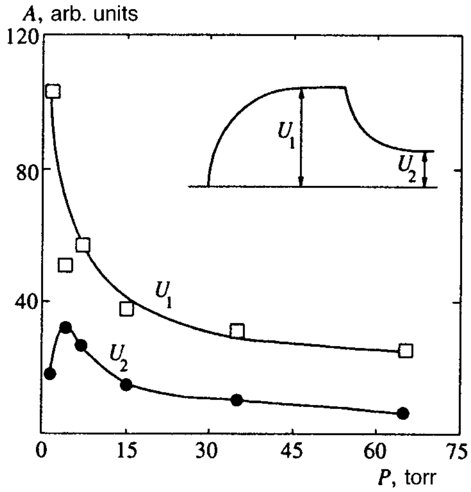


FIG. 10. Amplitudes of stimulated echo (with the constant input of two write pulses) and long-lived (with the input of a read pulse only) echoes vs ambient helium pressure for YBaCuO ($T=4.2$ K, $H=20$ kOe).

4. DISCUSSION OF THE EXPERIMENTAL RESULTS

The nonresonant dependence of the echo signal amplitude on the static external magnetic field leads to the conclusion that it is not related to the rotation of nuclear spins, but is attributable to oscillations imparted to the diamagnetic moment of the powder grains by the emf in the rf circuit. As mentioned, the acoustical nature of the echo is confirmed by the correlation of the frequency of the exciting pulses with the grain diameter and by the dependence of the echo intensity on the ambient gas pressure.

When rf pulses are applied to a powder, acoustic vibrations are excited in its grains. The electromagnetic generation of sound in normal metals has been studied at considerable length.[9,10] The feasibility of generating sound electromagnetically in high-T_c superconductor samples in an external magnetic field has already been demonstrated.[11]

Acoustic waves are generated with the same frequency as the alternating current in the exciting inductance coil by virtue of the Lorentz force acting on electrons down to the field penetration depth. The magnetic field in which the sample is immersed penetrates it as a system of vortices separated by a distance shorter than their characteristic dimensions. This means that spatial modulation of the magnetic field in the sample interior is extremely weak, so that the external field can be assumed to act directly on the electrons.

The vibrations of the powder grains induce in the receiving circuit an emf that decays within a time $\sim 1/\Delta\omega$ after each pulse, where $\Delta\omega$ is the spread of the oscillator eigenfrequencies. If a second exciting pulse is incident on the powders after a time τ, before the grain vibrations have had time to decay but after they have fallen out of phase, a second, similar decay of the induced emf sets in; after a time 2τ the grain vibration enter back into phase, producing an echo.

Gould[12] has shown that echo generation in a system of vibrating oscillators can be attributed only to their anharmonicity. The occurrence of echo in a system of anharmonic oscillators is clearly illustrated by the following simple

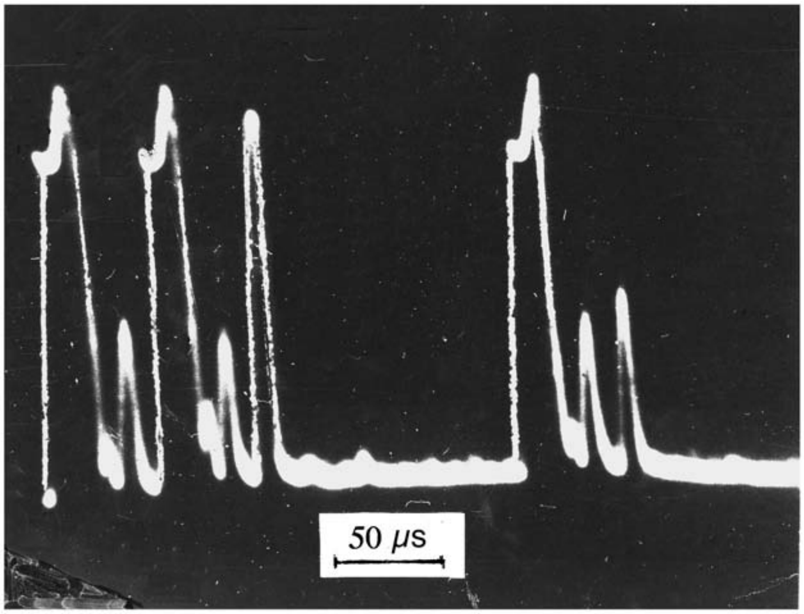


FIG. 9. Oscillogram of stimulated echo produced by two different time series of write pulses: after one series is written ($\tau=30$ μs) and the time delay between them is changed to $\tau=50$ μs, the write pulses resume the role of read pulses and are therefore followed by a stimulated echo signal.

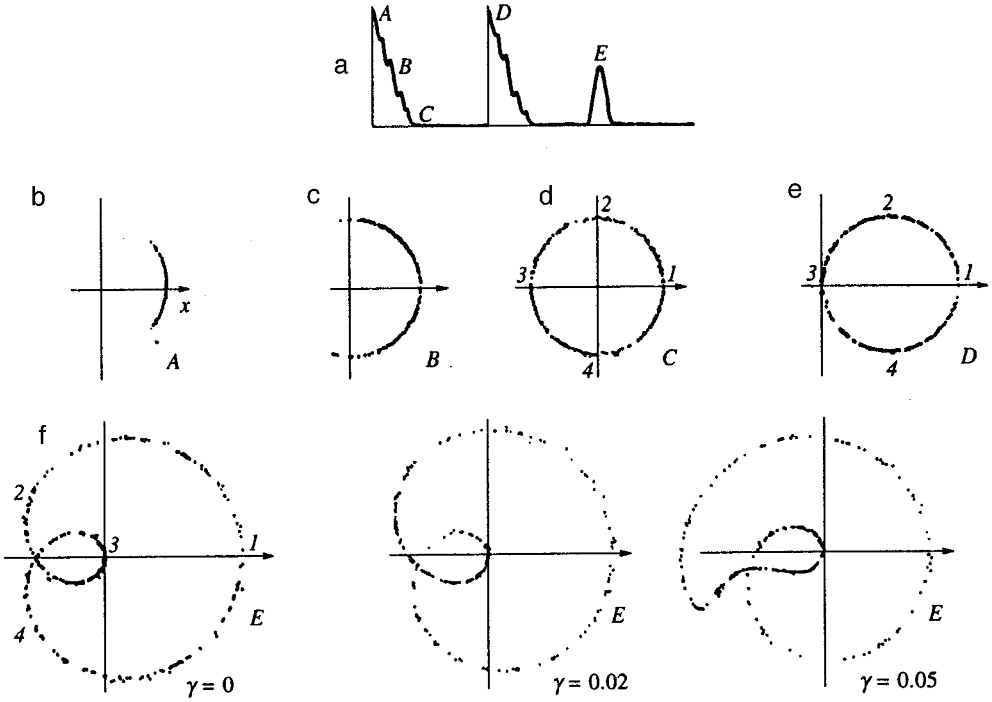


FIG. 11. Phase diagram of a system of nonlinear oscillators (results of a computer simulation of the behavior of a system of anharmonic oscillators with distinctly discernible resonance frequencies for various degrees of nonlinearity). The letters *A, B, C, D, E* denote different times in the system of exciting and sensing pulses.

model. Let us suppose that we have many nonlinear oscillators with distinctly discernible resonance frequencies and that their motion is described by the equation

$$\ddot{x}+\omega^2 x+\gamma x^3=0. \qquad (1)$$

The behavior of this system under the influence of an external force in the form of a sequential pair of pulses separated by a time interval τ is easily exhibited on a computer. The results of such a simulation are conveniently displayed on the $x\varphi$ phase plane in a rotating coordinate frame (where the x axis corresponds to vibration of an oscillator with the mean frequency) and are shown in Fig. 11.

When exposed to the first pulse, all the oscillators vibrate with the same amplitude and phase regardless of their frequency (Fig. 11b), and a macroscopic response appears in the system of oscillators. But then as time passes, owing to the different resonance frequencies, the vibration amplitudes gradually diverge in phase until the distribution of the directions of the vibration vectors becomes completely random, so that the phase diagram simply takes the form of a circle (Fig. 11d).

Under the influence of the second pulse all the oscillators acquire the same amplitude increment as in the excitation of the first pulse, and the phase diagram now represents the same circle shifted along the x axis (Fig. 11e). If the nonlinear term [the coefficient γ in Eq. (1)] is small (or equal to zero), in the subsequent evolution of the phase diagram the average projection of all the vectors describing the oscillator motions onto the x axis becomes equal to zero, and the induction signal in the receiving coil, being proportional to the vibration amplitude of the system at the mean frequency, becomes equal to zero (Fig. 11f). In the presence of nonlinearity, however, the phase trajectory is asymmetrical, and an echo signal appears after the time τ.

The theory of polarization echo in powders is based on a shift of the particle resonant frequency due to the defect structure of the sample material, in particular to dislocation deformation.[2] Asadullin[8] has developed the same approach to account for the echo in high-T_c superconductors. According to the Granato-Lücke string theory, the dislocation lines in a crystal form a three-dimensional grid. The lines are rigidly attached at the nodes of the grid and cannot be broken away by an ultrasonic field. A dislocation loop is loosely pinned in the interstitial space by all possible point defects. In the field of elastic vibrations (after pulsed excitation) small loops of length L vibrate as an elastic string without separating from the pinning sites if the elastic stress is small. The losses and the modulus step (ΔE effect) due to these loop vibrations are amplitude-independent. The response of the system is essentially nonlinear in this case, and echo does not occur.

As the amplitude of the stresses increases, the loops are catastrophically broken away from their pinning sites, but the process is reversible: After each half-period of the elastic vibrations a loop of length L breaks away and is repinned by defects. The attendant losses depend on the stress amplitude. As a dislocation moves under the influence of an rf pulse, it transfers from one local potential minimum to another. In its new position, initially at least, it is not necessarily pinned by point defects and, in that sense, can be characterized as excited. It is reasonable to assume that the density of excited dislocations is a function of the elastic stress amplitude and the pulse duration.

Under the influence of one pair of pulses only some of the mobile dislocation loops make an irreversible jump. The application of several pairs causes the pileup of such dislocations. This pileup of irreversibly excited dislocations is the basis of the cumulative formation of the long-lived stimulated echo.

The nonlinearity of the excitation process lies in the fact that the dislocation density N, which depends on the amplitude of the elastic strain $|s(x,t)|$ excited by this field, can vary during the lifetime of the rf pulse. The variation of $N(|s(x,t)|)$ influences the losses and dispersion of the elastic vibrations as well as the magnetostrictive contribution to their excitation.

We now consider the way in which, according to Asadullin,[8] an echo signal is generated in the case of superconductor in the mixed state. We assume for simplicity that the exciting pulses are identical and that acoustic vibrations $s(t)$ of amplitude s_0 are generated in the powder grains under their influence. The first exciting pulse is followed by vibrations in time at a frequency ω_1 with damping γ_1:

$$s_1(t)=s_0\ \exp[(i\omega_1-\gamma_1)t]. \tag{2}$$

The Lorentz force exerted by the rf current on the vortices is transferred to a dislocation associated with them. We assume for definiteness that the echo is associated with nonlinear properties in the dislocation vibrations. If the dislocations vibrate without breaking away from the weak pinning sites, the only consequence after the input of a second pulse is an amplitude-independent resonance shift and a change of the damping factor proportional to the dislocation density N:

$$\begin{gathered} s_2(t)=s_0\ \exp[(i\omega_2-\gamma_2)(t-\tau)], \\ \Delta\gamma=\gamma_2-\gamma_1=c_1N, \quad \Delta\omega=\omega_2-\omega_1=c_2N, \end{gathered} \tag{3}$$

where c_1 and c_2 are constants. After the second pulse the vibration amplitude is the sum of the response to the first and second pulses and at a time $t>\tau$ has the form

$$\begin{aligned} s(t)=s_1(t)+s_2(t)=s_0\{&\exp[(i\omega_1-\gamma_1)\tau]+1\} \\ &\times\exp[(i\omega_2-\gamma_2)(t-\tau)]. \end{aligned} \tag{4}$$

If $\Delta\omega$ and $\Delta\gamma$ do not depend on the amplitude, it is evident from Eq. (4) that the ith oscillator with frequency $\omega_i\neq\omega_1$ merely experiences a phase shift, but the system of oscillators with a random distribution of ω_i about the mean frequency ω_1 does not acquire a macroscopic moment, and an echo is not produced.

We now inquire what happens when the system acquires excited dislocations, the number of which depends on the vibration amplitude. During the second pulse the force acting on a unit length of the system {dislocation+vortex} is

$$\mathbf{f}=\mathbf{b}G\{s_0\ \exp[(i\omega_1-\gamma_1)\tau+s_0\}+n\mathbf{j}_2\Phi_0. \tag{5}$$

The first term in this expression for the force describes the direct effect of the acoustic vibrations (G is the shear modulus, and $\mathbf{b}$ is the Burgers vector), and the second term describes the interaction between the rf current in the second pulse and the vortices (n is the number of vortices associated with one dislocation, $\mathbf{j}_2$ is the convolution of the second rf pulse with time, and Φ_0 is the magnetic flux quantum).

The average number of dislocations shifted from their sites (excited dislocations) is proportional to the effective force, and the substitution of the force expression (5) yields

$$\begin{aligned} \overline{N(f\ )}=\alpha\mathbf{f}^2=\alpha[bG\{&s_0\ \exp[(i\omega_1-\gamma_1)\tau]+s_0\} \\ &+nj_2\Phi_0]^2=\beta\ \exp(i\omega_1\tau). \end{aligned} \tag{6}$$

Here α and β are constants.

According to (3), the frequency shift and the variation of the damping are proportional to the dislocation density, but according to (6), the number of excited dislocations depends on the time interval between pulses, so that the variations of the frequency and the damping factor acquire a factor that depends on the spacing τ of the write pulses:

$$\begin{gathered} \Delta\omega(t)=\Delta\omega_0-\Delta\omega_1\ \cos(\omega_1\tau), \\ \Delta\gamma(t)=\Delta\gamma_0+\Delta\gamma_1\ \cos(\omega_1\tau) \end{gathered} \tag{7}$$

The new frequency and the new damping factor of the vibrations contain phase information about the rf pulse in the form of the phase of the elastic vibrations $\omega_1\tau$. Now the expression (4) for the vibration amplitude of the powder gains assumes the form

$$\begin{aligned} s(t)=(s_1+s_2)&\exp[-i\Delta\omega\ \cos(\omega_1\tau)] \\ &\times\exp[-\Delta\gamma\ \cos(\omega_1\tau)]. \end{aligned} \tag{8}$$

It has been shown[8,13] that this result implies the formation of an echo signal a short time τ after the second pulse in a system consisting of a large number of particles with distinctly discernible resonant frequencies.

A remarkable characteristic of rf echo in powders is the long-term memory effect. We now elucidate the process by which a long-lived stimulated echo is produced under the influence of nonlinear mechanisms. We consider a time when the vibrations excited by the first and second pulses have decayed, but the memory of these pulses is retained in the variations of the damping $\Delta\gamma(p_1,p_2,\tau)$ and the frequency $\Delta\omega(p_1,p_2,\tau)$ of the particle eigenfrequencies (p_1 and p_2 are the dipole moments of the powder grains). These variations are attributable to the irreversible motions of defects in the grains (dislocations, etc.) under the influence of the elastic vibrations.

The third (read) pulse imparts a dipole moment p_3 to each grain. As in the case of two-pulse exposure, nonlinearity leads to echo generation. The phase memory of the preceding pulses is retained in the density of excited dislocations. As time passes, the state of the excited dislocations changes, and in the new position they are gradually pinned by point defects diffusing toward them and continue the process of diffusion to other potential minima, thereby breaking up the coherence of the state.

By this time, of course, all the elastic vibrations generated by the first two pulses have died out. The memory of them is retained only in the damping factor γ_2 and the frequency ω_2. The third pulse imparts a dipole moment to each grain, and once again an echo signal appears after a time τ if the damping in the system is amplitude-dependent.

The lifetime of the echo in this case is determined by the relaxation time of the indicated excited state. The relaxation process could be associated, for example, with the gradual pinning of dislocations by point defects diffusing toward them or with the reverse process of truncation of extended dislocation loops. The cumulative storage effect in the repeated input of coherent pairs of rf pulses is associated with the pileup of excited dislocations (or twinning boundaries). The erasure of the memory as the temperature and the external magnetic field vary is readily explained in the following model: When the temperature varies, the pinning force changes and, accordingly, the vortices transfer to other local minima of the potential relief of the lattice; when the field varies, the number of vortices changes, and they can also transfer to other pinning centers, i.e., become rearranged.

In a superconductor the breaking away of a dislocation from its pinning site inevitably sets in motion the vortex associated with it, and magnetic flux jumps take place. In exactly the same way, any motion of the vortices imparts motion to defects associated with them. This can happen only so long as the unpinning of vortices does not set in at a some temperature where the coupling between the motions of dislocations and vortices is broken.[14] The (B,T) diagram therefore acquires two regions separated by a line of irreversibility, above which the diamagnetic moment quickly relaxes, and the critical current drops abruptly with increasing temperature, both of these effects being associated with a variation of the vortex mobility in the high-T$_c$ superconductor. The energy of pinning of a solitary vortex by a point defect can be described by the expression

$$U_p \approx \xi^d B_{c2}^2, \tag{9}$$

where ξ is the coherence length, which is very small in high-T$_c$ superconductors, $1<d<3$, and B_{c2} is the second critical field. This energy must be compared with the energy of thermal vibrations kT, which lead to thermally activated vortex hops from one pinning center to another. The probability W of such hops is

$$W \propto \exp(-U_p/kT), \tag{10}$$

and the ratio U_p/kT in the high-T$_c$ superconductor can be an order of magnitude lower than in ordinary superconductors. Consequently, thermally activated vortex motion has a significant influence on the magnetic and transport properties of high-T$_c$ superconductors. In conventional superconductors this phenomenon is observed only near the transition temperature as flux creep.

The unpinning effect appears as a maximum on the temperature curve of the losses in the superconducting sample when the vortex lattice is set in motion by the external force. Data of this kind have been obtained so far in experiments to measure the real and imaginary parts of the magnetic susceptibility as well as the speed and attenuation of ultrasound, and also from the results of measurements of the current-voltage characteristics in experiments with a high-T_c superconducting single crystal vibrating in an external field (see, e.g., Ref. 15).

The mobility of the vortices can also be regarded as melting of the vortex lattice.[16] Thus, even in a perfect crystal a solitary vortex cannot move freely if $B \gg B_{c1}$, because any slight movement of the vortex relative to the stationary vortex lattice causes it to be deformed. The vibration amplitude of the vortex lattice increases as the temperature rises, and when the vibration amplitude $\langle x \rangle$ becomes sufficiently large, the vortex lattice can melt (by the Lindemann criterion $\langle x \rangle / a \approx 0.1-0.2$, where a is the spacing of the vortices), transforming into a vortex liquid.[16] The mobility of the vortices increases dramatically when the shear modulus of the vortex lattice becomes equal to zero at the melting point.

It follows from the experimental results that not only do the echo signals vanish long before the transition temperature of our high-T$_c$ superconductors is reached, but also their temperature dependences differ for the yttrium and bismuth systems. These characteristic attributes of the generation of echo signals are readily understood through the results of recent studies of (B,T) diagrams: They differ significantly for anisotropic three-dimensional (YBaCuO type) and layered (BiCaSrCuO type) high-T_c superconductors in that the boundary of the phase diagram for the yttrium system exists at much higher temperatures and fields than for the BiCaSrCuO system.

An estimate of the line of irreversibility for grains of diameter 100 μm at a frequency of 20 MHz shows that its position coincides with the curve representing the temperature dependence of the echo intensity, and the difference in the temperature behavior of the yttrium and bismuth powders is qualitatively consistent with the difference in the positions of their lines of irreversibility.

5. CONCLUSION

We have carried out a detailed experimental study of dynamic and stimulated rf echo in powdered yttrium and bismuth high-T$_c$ superconductors as functions of the magnetic field, the ambient gas pressure, and the temperature. We have measured the temperature dependence of the echo relaxation time, and we observed a long-lived stimulated echo and its cumulative storage.

Most of the results can be explained in terms of the model proposed by Asadullin, based on the nonlinear character of the motion of defects and their associated vortices in the sample. It follows from the temperature dependence of the echo amplitude that this phenomenon is intimately related to the vortex lattice dynamics and can therefore serve as an effective method of investigation of the latter.

The authors are deeply indebted to V. V. Zhuchkov for assistance with the experiments and to Ya. Ya. Asadullin for rewarding discussions.

This work has been performed in part under the auspices of Project 93210 of the State Scientific Program ''Superconductors.''

[1] W. P. Mason and R. N. Thurston (eds.), *Physical Acoustics*, Vol. 16, Academic Press, New York–London (1982), Chaps. 4–6.
[2] U. Kh. Kopvillem and S. V. Prants, *Polarization Echo* [in Russian], Nauka, Moscow (1985).
[3] É. A. Manykin and V. V. Samartsev, *Optical Echo Spectroscopy* [in Russian], Nauka, Moscow (1984).
[4] V. S. Romanov and N. K. Solovarov, JETP Lett. **61**, 231 (1995).
[5] A. V. Drobinin and Yu. M. Tsipenyuk, JETP Lett. **49**, 651 (1989).
[6] H. Nishinara, K. Hayashi, Y. Okuda, and K. Kajimura, Phys. Rev. B **39**, 7351 (1989).
[7] M. P. Petrov, D. V. Dyadyushkin, A. V. Ivanov *et al.*, Sverkhprovodimost' **3**, 2363 (1990).
[8] Ya. Ya. Asadullin, Sverkhprovodimost' **6**, 545 (1993).

[9] V. F. Gantmakher and V. T. Dolgopolov, Zh. Éksp. Teor. Fiz. **57**, 132 (1969) [Sov. Phys. JETP **30**, 78 (1969)].

[10] A. N. Vasil'ev and Yu. P. Gaĭdukov, Usp. Fiz. Nauk **150**, 161 (1986) [Sov. Phys. Usp. **29**, 889 (1986)].

[11] V. F. Gantmakher, V. T. Dolgopolov, S. S. Murzin *et al.*, Fiz. Nizk. Temp. **15**, 217 (1989) [Sov. J. Low Temp. Phys **15**, 123 (1989)].

[12] R. W. Gould, Phys. Lett. **19**, 477 (1965).

[13] R. W. Gould, Am. J. Phys. **37**, 585 (1969).

[14] E. H. Brandt, Phys. Rev. Lett. **68**, 3769 (1992).

[15] A. V. Drobinin, K. E. Kamentsev, A. A. Bush, and Yu. M. Tsipenyuk, Sverkhprovodimost' (KIAE) **6**, 1487 (1993).

[16] G. Blatter, M. V. Feigelman, V. B. Geshkenbein, A. I. Larkin, and V. M. Vinocur, Rev. Mod. Phys. **66**, 1125 (1996).

Translated by James S. Wood

Effect of an alternating field on the relaxation of Effect of an alternating field on the relaxation of magnetic flux in a Josephson medium

É. V. Matizen, P. P. Bezverkhiĭ, and V. G. Martynets

Institute of Inorganic Chemistry, Siberian Branch of the Russian Academy of Sciences, 690090 Novosibirsk, Russia

(Submitted 19 July 1996)
Zh. Éksp. Teor. Fiz. **111**, 1047–1056 (March 1997)

We establish how trapped magnetic flux depends on the frequency and amplitude of an alternating field and how such a field affects the relaxation rate of the flux. We find that the nature of the flux creep changes in the process and that relaxation of the flux stops after the external field is switched off. We examine the dynamics of flux relaxation in a ring in the approximation in which the current density is assumed homogeneous, for various density dependences of the effective vortex activation energy. The critical current density and the vortex activation energy are obtained as functions of the external field strength. Finally, we explain the observed behavior in terms of the different field profiles emerging in the rings. © *1997 American Institute of Physics.* [S1063-7761(97)02303-2]

1. INTRODUCTION

Much research has been done in studying the effect of an alternating magnetic field on the properties of granulated high-T_c superconductors (see, e.g., the review in Ref. 1). The research involved studying the complex magnetic susceptibility, the penetration of samples by weak alternating magnetic fields,[2] the formation of various current layers by such fields,[3] the effect of an alternating field on the current–voltage characteristics of the samples,[4] and other research. However, the effect of an alternating field on relaxation phenomena has been studied less.

It is known that the presence of weak fields gives rise to a random network of Josephson junctions in granulated high-T_c superconductors.[5–7] When this problem is studied phenomenologically, the system is interpreted as an effective medium with a Josephson penetration depth λ_j and an effective permeability $\mu_{\rm eff}$ that allows for Meissner screening of the granules. In fields lower than $H_{c1}^{\rm gr}$, the first critical field of the granules, the electrodynamics of the medium is determined by intergranular currents. Such an approach is valid if $\lambda_j \gg a(\sqrt{H_{c1}} \ll \sqrt{H_{c2}})$, where a is the size of the granules, and H_{c1} and H_{c2} are the first and second critical fields for the medium. In such a medium the condition for the existence of a critical state, $|dH/dx| = j_c(H)$, is valid, with the field penetrating the medium in the form of Josephson vortices, which can move and get pinned at various pinning centers, leading to various effects in flux creep.

In this work we study the effect of an alternating low-frequency magnetic field $(\widetilde{H})$ on the magnetic flux trapped in ceramic high-T_c superconducting rings when a Josephson medium is formed in rings and the amplitude of the alternating field is lower than the field capable of completely penetrating the sample (ring).

2. EXPERIMENT

The measuring cell consisted of a solenoid with two auxiliary coils forming a Helmholtz system, the coils being coaxial with the solenoid. The sample, a high-T_c superconducting ring, was placed between the two auxiliary coils. A Hall-effect detector and the junction of a thermocouple were placed at the center of the ring. The solenoid generated a constant magnetic field up to 200 Oe, while the low-inductance coils were used to generate an additional alternating magnetic field with an amplitude of up to 20 Oe. The cell was placed in a bath of liquid nitrogen or helium vapor. Measurements were done automatically and were computer-controlled. The error in measuring the magnetic field at the ring's center amounted to 0.02Oe, and the sample temperature was kept constant to within 5×10^{-2}K. The rings we studied were made of Bi-based ceramic and of granulated YBCO films on a sapphire substrate. The dimensions of the ceramic rings were: external radius $R_2 \approx 4.8$ mm, internal radius $R_1 \approx 3.2$ mm, height ~ 1.4–2 mm, and width $W = R_2 - R_1 \sim$ 1.2–2 mm; the current density was ≈ 1 kA/cm^{-2} at 4.2 K; and $T_c \approx 105$–107 K. The film rings had an external diameter of approximately 9 mm, internal diameter of approximately 6.5 mm, their height $\approx 0.5\,\mu$m; the current density was ≈ 11 kA/cm^{-2} at 78 K; and $T_c \approx 98$ K.

First we studied the magnetization of the rings, i.e., the dependence of the field B at the center of the ring on the strength H^e of the external field generated by the solenoid. Figure 1 depicts one such dependence (curve *1*), with cell-configuration corrections introduced in accordance with Ref. 8. Obviously, up to H_{cj}, corresponding to the critical states of the ring, the transport current encompasses an ever increasing cross-sectional area of the ring, starting from the outer layer. When the critical state is reached, the field expands into the hole of the ring. A further increase in the strength of the external field causes a decrease in current density. This is evident from the fact that the B vs H^e curve approaches the straight line $B = H^e$. In higher fields, $H^e > 30$ Oe, the dependence is affected by magnetization of the granules, which, on the one hand, additionally lowers the transport current density and, on the other, generates a field acting on the Hall-effect detector in the opposite direction. In high external fields and at high temperatures the field gener-

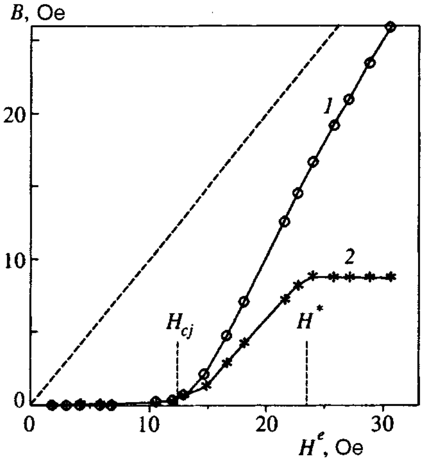


FIG. 1. Magnetization of the ring by a constant magnetic field H^e: curve *1* represents the behavior of the field B at the center of the ring as H^e increases; curve *2* the behavior of B at the center of the ring after H^e is switched off; the straight line $B=H^e$ represents the readings of a Hall-effect detector in the absence of the ring; H_{cj} stands for the penetration field strength corresponding to the critical state of the ring; H^* stands for the value of H^e at which maximum flux is trapped; and $T=77.3\mathrm{K}$.

ated by the granules exceeds the field generated by the transport current and creates several effects, including reverse hysteresis loops.[8] The curve *2* in Fig. 1 represents the trapped field after several seconds (~ 4 s) following the switch-off of the external field. We see that in fields $H^e \leqslant H_{cj}$ the flux is not trapped, in the range $H_{cj} < H^e < H^*$ the flux is trapped only partially, and at $H^e = H^*$ the trapped flux reaches its maximum value, remaining almost constant in high fields (the observed small decrease in flux is due to granule magnetization[9]). Note that for all the samples (both bismuth and yttrium) the curves exhibit similar behavior, differing only in the value of the characteristics.

The time dependence of the field B trapped at the center of the ring when the field is turned off (a relaxation curve) is generally described by a function characteristic of collective creep:[1,10]

$$B(t)=\frac{B_0}{[1+S'\ln(1+t/\tau)]^{1/\beta}}, \tag{1}$$

where B_0 is the value of the field at time zero, S' is the relaxation rate, and τ is the characteristic relaxation time.

When the external field H^e is switched on, the time dependence of $B(t)$, i.e., the introduction of the field into the ring, is described by the following formula:

$$H^e-B(t)=(H^e-B_0)\left[1+S'\ln\left(1+\frac{t}{\tau}\right)\right]. \tag{2}$$

To study the effect of an alternating field on the relaxation mode, we first found the dependence of the current density in the ring at time zero (j_0), calculated from the relationship $j_0 \propto H^e - B_0$, on the external field strength. This dependence on H^e, experimentally found from the parameters of the relaxation curve during the introduction of the field into the ring, is described satisfactorily by the function[11]

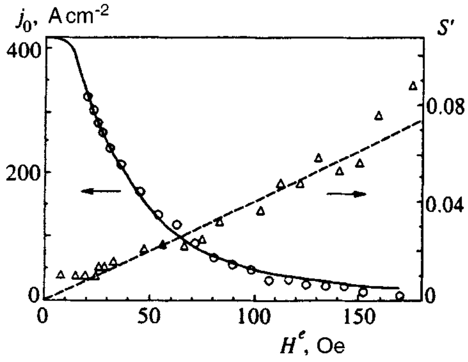


FIG. 2. The transport current density j_0 (○) and the relaxation rate S' (△) as functions of the external field strength H^e, found from the data on the relaxation of the field B at the center of the ring when H^e is introduced into the ring at 37 K: the solid curve represents the results of calculations by formula (2), and the dashed curve the results of calculations by the formula $S'=AH^e$, with $A=4\times10^{-4}\ \mathrm{Oe}^{-1}$.

$$j_0(H^e)=\frac{j_{00}H_0^2}{H^{e2}+H_0^2} \tag{3}$$

with the fitting parameters $j_{00}=420.4\ \mathrm{A/cm^2}$ and $H_0=36.5\ \mathrm{Oe}$ ($T=37$ K); see Fig. 2. Note that there is a range of weak fields in which j_0 changes little. Figure 2 also depicts the dependence of the relaxation rate S' on the external field strength.

Figure 3 depicts the curves representing the relaxation of the constant component of the trapped field $B(t)$ after H^e is switched off, in the absence of an alternating field and when such a field is switched on at the moment of trapping ($T=78$ K). Without the alternating field the relaxation curve is described by the following relationship:

$$B(t)=B_0\left[1-S'\ln\left(1+\frac{t}{\tau}\right)\right], \tag{4}$$

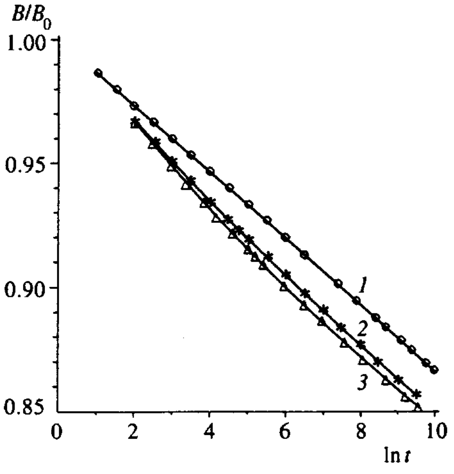


FIG. 3. Relaxation of the trapped field B in the presence of an alternating magnetic field with an amplitude H^0 after H^e has been switched off ($T=77.3$ K): curve *1* represents the behavior of B without the alternating field; curve *2* represents the behavior of B at $H^0=2.9$ Oe, $\nu=1100$ Hz, and the same initial value of H^e as for curve *1*; and curve *3* represents the behavior of B at $H^0=2.9$ Oe, $\nu=100$ Hz, and the value of H^e corresponding to the case of curve *1*.

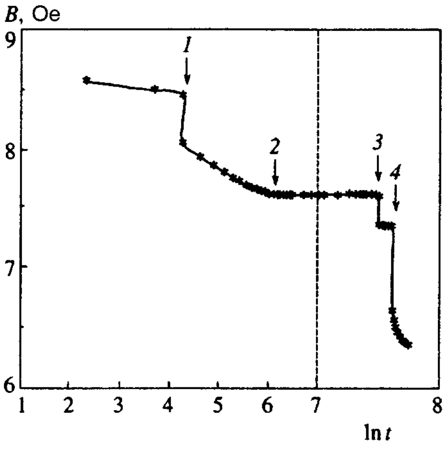


FIG. 4. The effect of introducing an alternating field on the relaxation of the constant component of the trapped field B at 77.3 K. The field H^e is switched off, and the arrows indicate the various times: at t_1 the amplitude is $H^0=1$ Oe and the frequency is $\nu=7$ kHz; at t_2 the alternating field is is switched off; at t_3 the amplitude is $H^0=2.5$ Oe and the frequency is $\nu=7$ kHz; and at t_4 the amplitude increases to $H^0=3.11$ Oe which the frequency is $\nu=7$ kHz. The dashed straight line separates the regions with different scales along the horizontal axis.

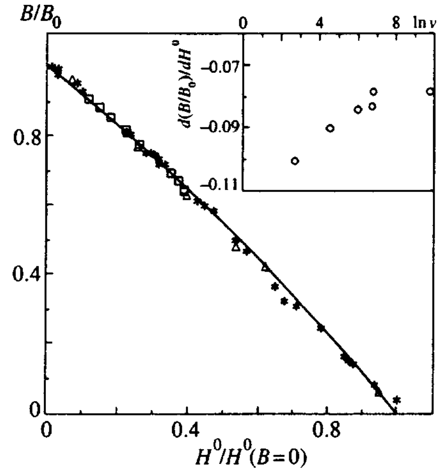


FIG. 5. The field B trapped by a ring after H^e is switched off as a function of the amplitude H^0 of the alternating field. Here B_0 is the field trapped at $H^0=0$, and $H^0(B=0)$ is the amplitude of the alternating field at which the field trapped by the ring vanishes. The solid curve corresponds to the results of calculations by formula (8). The experimental points for six different frequencies in the interval from 10 Hz to 20 kHz fit the calculated curve if for each frequency we take the corresponding value of $H^0(B=0)$. The inset shows the dependence of the average derivative $d(B/B_0)/dH^0$ on $\ln\nu$ in the interval $0.5<H^0/H^0(B=0)<1$.

with the parameters $S'=0.0133$, $B_0=8.77$ Oe, and $\tau=1$ (curve *1* in Fig. 3). Note that this dependence is the limit of the function (1) as $1/\beta\rightarrow 1$ for small values of S'. In an alternating field the relaxation curve is described by the function (1) with $S'=9.48\times10^{-3}$, $B_0=7.18$ Oe, and $1/\beta=1.83$ for an alternating field amplitude equal to 2.9 Oe and a frequency $\nu=1100$ Hz (curve *2*), and with $S'=1.48\times10^{-2}$, $B_0=7.31$ Oe, and $1/\beta=1.26$ for the same alternating field amplitude and a frequency $\nu=100$ Hz (curve *3*). The parameter calculated from the formula $\tau=S'j_cW(dH^e/dt)^{-1}$ (see Ref. 12) was equal to 0.06 s. As Fig. 3 shows, the introduction of an alternating field transforms the logarithmic dependence (4) into the nonlinear dependence (1).

Figure 4 depicts the effect of an alternating magnetic field on the relaxation of the constant component of the field B trapped by the ring. Up to time t_1 the field B relaxes with a rate $S'=0.015$. At t_1 an alternating field with an amplitude $H^0=1$ Oe and a frequency $\nu=7$ kHz is switched on, which leads to a sharp drop in B, after which the field relaxes until time t_2 with a rate $S'=0.017$, provided that this relaxation process starts at t_1. At time t_2 the alternating field is switched off over ≈0.5–1 s, after which the trapped field ceases to change, within the limits of sensitivity of the measuring device ($\delta B\approx0.004$ Oe). As a result, one can observe an interesting phenomenon—the flux in the ring becomes trapped. In the course of several hours no change in this state was observed. At time t_3 an alternating field with $H^0=2.5$ Oe and $\nu=7$ kHz is switched on, there is again a sharp increase in B, and relaxation sets in once more. At time t_4 the amplitude increases to $H^0=3.11$ Oe, B drops still further, and relaxation continues with a rate $S'=0.017$.

Figure 5 depicts, in relative units, the dependence (characteristic of all rings) of the constant component of the trapped field B (the value to which the field B rapidly drops after an alternating field $\widetilde{H}$ with an amplitude H^0 is switched on; see Fig. 4) on the amplitude and frequency of the applied alternating field. Clearly, the experimental points for all amplitudes and frequencies in these coordinates fit on a single curve. Note, however, that the effect of the frequency on the derivative dB/dH^0 is weak (see the inset in Fig. 5).

3. DISCUSSION

Before we consider the effects of an alternating field on a trapped magnetic flux, let us discuss the characteristics of the Josephson medium realized in the samples. As an example we take the results for one ceramic ring (see Fig. 1). For $H^e\geqslant H^*$ the curve *2* reaches a plateau, which corresponds to the maximum value of the trapped flux. We find $H^*\approx1.8H_{cj}$, a result that is in approximate agreement with the calculated one. Indeed, reasoning on the basis of the critical-state equation and the usual assumptions concerning the j_c vs H^e dependence, one can show that $H^*=2H_{cj}$ (see Ref. 5). When the strength of the field is further increased, the value of trapped flux (curve *2*) slowly decreases because of suppression of the critical current (the depinning current) by the external field due to degradation of the Josephson medium and granule magnetization.[9] We observed a similar behavior of the curves for the YBCO film rings.

The variation of the relaxation rate of a magnetic flux in a ceramic ring in the absence of an alternating field, which we studied in our description of the Josephson medium, is depicted in Fig. 2. For $H^e\leqslant150$ Oe the S' vs H^e dependence can be described by an expression of the form $S'=AH^e$ ($A=4\times10^{-4}$ Oe^{-1}). For $H^e<H_{cj}$ the method does not allow the dependence of S' on H^e to be detected, since in this range of field strengths the contriimportant of the field generated by the relaxing current itself is important. There are, however, certain experimental indications, obtained from single-crystal studies,[13] that for $H^e<H_{c1}$ the S' vs fH^e de-

pendence degenerates into a constant. In a field $H^e \simeq H_0$ at 37 K, the weak Josephson junctions in the medium are destroyed. This value of H_0 agrees with the value of H_{c2} obtained in measurements of the magnetization of rectangular rods fabricated from the same ceramic material.[8] This value of H_{c2} has a corresponding size $\zeta=\sqrt{\Phi_0/H_{c2}}\approx 0.7\ \mu$m, which in order of magnitude coincides with the average size of a grain (granule), with Φ_0 the quantum of magnetic flux.

As is well known,[12] if the Kim–Anderson model is applicable, i.e., if creep in a thin film generates an electromotive force whose value per unit length is

$$E=\frac{E_0}{2}\exp\left[-\frac{U_0}{k_BT}\left(1-\frac{\alpha j}{U_0}\right)\right],$$

then for $k_BT/U_0 \ll 1$ and relatively long times the relaxation rate is given by the following expression:

$$S'=\frac{k_BT}{\alpha j_0}.$$

Here U_0 is the vortex activation energy, α is a coefficient that accounts for the lowering of the barrier for vortex hopping due to the Lorentz force, and j_0 is the current density corresponding to B_0 in (1) or to H^e-B_0 in (2). For $H^e>H_{cj}$, according to (3), this yields

$$\alpha=\frac{k_BT}{Aj_{00}(t)H}\left(1+\frac{H^{e2}}{H_0^2}\right).$$

By definition, the critical current density (the depinning current) j_c is equal to U_0/α. In our case we can assume that $j_0 \sim j_c$. Then $U_0 \approx j_0\alpha = k_BT/S' = k_BT/AH^e$ when $H^e>H_{cj}$. Note that in Ref. 8 it was established that in magnetic fields induced by the critical current itself, we have $S' \propto T$, with the result that A is proportional to T and U_0 is temperature independent. An estimate of E from the data on flux relaxation at 37 K yields $E\simeq(1/2)R\mu_0B_0S'/(t+\tau)$. At $t=0$ and $\tau=1$ we have $E_0\simeq(1/2)\times R\mu_0B_0S'\approx 3\times 10^{-8}$ V/m, where R is the mean radius of the ring, $S'=0.015$, $B_0=10$ Oe$=786$ A/m, and $\mu_0=4\pi\times{}^{-7}$H/m.

To explain the effect of an alternating field on the behavior of the magnetic flux trapped by a ring we attempted to describe our experiment using the Kin–Anderson model with $U_{\text{eff}}=U_0(1-j/j_c)$. This model allows for an analytical solution for a thin ring in an approximation in which the current density is homogeneous and $k_BT/U_0 \ll 1$. Suppose that the amplitude of the alternating field is sufficiently high for the field to extend into the hole of the ring. Then the electrodynamics of the ring is described by the following equation:

$$2\pi RL\frac{dj}{dt}+\frac{E_0}{2}\exp\left(-\frac{U_0}{k_BT}\right)\exp\left(\frac{\alpha j}{k_BT}\right)$$
$$+\mu_0\pi R^2\omega H^0\sin\omega t=0. \tag{5}$$

Here the dimensions of the ring were taken into account only by the ring inductance $L=\mu_0RW/2$, where R is the mean radius of the ring, and W is the width of the ring.

The solution of Eq. (5) with the initial condition $j(t=0)=j_0$ has the form

$$j(t)=j_0\left\{1-\frac{H^0}{Wj_0}\sin\omega t-\frac{k_BT}{\alpha j_0}\right.$$
$$\left.\times\ln\left[1+C\int_0^{\omega t}\exp(-C\sin\omega t)d(\omega t)\right]\right\}. \tag{6}$$

Here we write $C=(U_0/k_BT)H^0/j_0W$ with allowance for the fact that $j_0\sim j_c$. Averaging this solution over the period yields

$$\langle j(t)\rangle=j_0\left\{1-\frac{k_BT}{U_0}\ln[1+G(H^0)\omega t]\right\}, \tag{7}$$

where

$$G(H^0)=\frac{C}{\omega t}\int_0^{\omega t}\exp(-C\sin\omega t)d(\omega t).$$

The solution implies that with the passage of time the current generating the flux trapped by the ring decreases logarithmically, as is the case with ordinary creep (see Eq. (4)), but we have $\tau=[G(H^0)\omega]^{-1}$ rather than unity. In addition, this solution provides a qualitative description of bursts of the trapped flux escaping when the alternating field is switched on (see Fig. 4). Choosing the characteristic values of the parameters of, say, one bismuth ring at 77 K, $k_BT/U_0=1.6\times10^{-2}$, $W=2\times10^{-3}$m, and $j_0=3.4\times10^6$A/m^2, we found that an alternating field of frequency 100 Hz with an amplitude $H^0\approx H_{cj}=11.4$ Oe in the experiment leads to complete destrapping flux (the field is pushed out of the hole of the ring). On the other hand, if we use Eq. (7) and determine $G(11.4$ Oe$)$ by numerical integration, we find that it takes roughly one second for the field trapped by the ring (the current in the ring) to drop from 8.63 Oe to only 6.86 Oe, while it takes a field of 76 Oe to achieve complete detrapping of the flux in the course of several seconds, in view of the almost exponential rise in the function $G(H^0)$. The discharge of the trapped field caused by introduction of an alternating field of amplitude H^0 occurs in a relatively short interval, the time it takes to introduce the field (i.e., $\sim$1–2 s); see Fig. 4. Calculation of the discharge by Eq. (7) at time t_1 in Fig. 4 has shown that such a step on the relaxation curve appears when an alternating field with $H^0=0.04$ Oe and $\nu=7$ kHz is introduced, while in the experiment the amplitude H^0 amounted to 1 Oe. We also note that the solution (6) for an infinitely thin ring contains a term with sinωt, as a result of which the alternating field reaches the center of the ring at any amplitude H^0. This, however, is not the case for a real ring of width W, and we can examine the effect of this solution on the field at the center only from that outer part of the ring that is penetrated by the alternating field.

Thus, the model provides a qualitative description of the behavior of the samples in an alternating field, but for better agreement with the experimental data it is advisable to take into account the finite dimensions of the ring and the resulting current and field distributions. Similar results concerning the effect of an alternating field are achieved by using the collective creep model

$$U_{\mathrm{eff}}=\frac{U_0}{k_BT}\left[\left(\frac{j_0}{j}\right)^{\beta}-1\right]$$

within the same approximations. Combining this model and the method of Gurevich and Brand,[10] we arrive at the following expression for the time dependence of the current density

$$j(t)=j_0\left\{1+\frac{\beta H^0}{Wj_0}\sin\,\omega t+\frac{k_BT}{U_0}\times\ln\left[1+\beta C\int_0^{\omega t}\exp(-\beta C\,\sin\,\omega t)d(\omega t)\right]\right\}^{-1/\beta},$$

which in the limit of small amplitudes H^0 and $k_BT/U_0\ll 1$ coincides with (6).

Let us now explain the results depicted in Fig. 5. Here it is natural to assume that if the conditions for a critical state are met, in rings of finite width a weak alternating field $\widetilde{H}$ penetrates the ring only up to a certain depth along the radius from the external lateral surface. This field evens out the average gradient of the field trapped in the ring. Thus, if H^0 is lower than the field in which the current is destroyed in the entire ring, then the trapped (average) transport current is destroyed only within a certain layer ΔR and, hence, the trapped flux proves to be smaller than when the current flows through the entire cross section of the ring. In this case the dependence of the constant component of the field B at the center of the ring on H^0 is given by the following expression:

$$\frac{B(H^0)}{B_0}=\frac{\ln\{1+(W/R)[1-H^0/H^0(B=0)]\}}{\ln(1+W/R)}. \tag{8}$$

Here R is the external radius of the ring, and $H^0(B=0)$ is the amplitude at which the trapped flux is destroyed by the alternating field in the entire ring. In Fig. 5 the dependence (8) is represented by the solid line, which approximates the behavior of the experimental values fairly well. The experiments in the frequency range $2<\ln\,\nu<7$ revealed a weak dependence of $B(H^0)$ on ν (see the inset in Fig. 5), with $H^0(B=0)$ increasing with frequency, with the result that qualitatively the suggested dependence (8) correctly reflects the behavior of the derivative $(1/B_0)dB/dH^0$ in the experiment. The weak frequency dependence of this derivative also follows from the solution (7), since the derivative

$$\frac{1}{j_0}\,\frac{dj}{dH^0}\sim\frac{1}{B_0}\,\frac{dB}{dH^0},$$

found from (7) is almost frequency-independent.

The observed sudden change in B (Fig. 4) corresponds to the dependence of $B(H^0)$ given by (8), while cessation of flux relaxation resembles the same phenomenon of flux trapping in the region from H_{cj} to H^* without an alternating field, which in the previous paper (Ref. 8) was explained by the formation of two field distributions in a ring with gradients opposite in sign, with the field gradient in the region adjacent to the hole directed in such a way that a force directed toward the center of the ring acts on the vortices. This explains the absence of flux relaxation in the hole. Cessation of relaxation when the field $\widetilde{H}$ is switched off has a different explanation. It is known[6,14] that relaxation of the trapped flux in a Josephson medium occurs because of the motion and emergence of Josephson vortices with a certain diffusion coefficient D as a result of the existence of a field gradient. Strong variation of D requires a frequency of the alternating field that is comparable with the frequency of vortex hopping between pinning centers, $\Omega\sim 10^6$–$10^{10}\mathrm{s}^{-1}$. Since in our experiments the frequency of the alternating field varied from 20 Hz to 20 kHz, which is much lower than hopping frequency, the sharp decrease in the rate of flux relaxation must be attributed to field gradients. Consequently, we assumed that the practical termination of relaxation is caused by a field structure modulated along the ring radius, a structure produced by a low-frequency alternating field. Differently directed varying gradients may appear in such structures when the field is applied, while a constant structure emerges when the field is switched off. Here it is assumed that the average gradient of the field in a modulated structure is zero, and so is the flux trapped by such a structure (see solution (7)).

Thus, our interpretation is as follows. When there is relaxation of the trapped flux in an alternating field whose amplitude is lower than $H^0(B=0)$, the section of the ring near the hole contains a constant field gradient that pushes the vortices out, while the section of the ring near its outer surface, penetrated by the alternating field $\widetilde{H}$, contains a modulated structure with a zero average gradient. When the field is applied, the modulated structure oscillates and does not prevent the vortices from escaping, only slightly distorting the purely logarithmic pattern of relaxation (see Fig. 3). After $\widetilde{H}$ is switched off, the modulated structure becomes frozen. The frozen structure contains field gradients directed so as to prevent the vortices from escaping, the flux in the ring freezes, and relaxation stops (see Fig. 4).

The modulated structure emerges primarily because the conditions for the critical state to emerge are met. Its stability is determined by the phenomenon of self-organized criticality,[15] amounting to the fact that when the field gradient spontaneously decreases, the current density in a section of the ring also decreases, which in turn results in an increase in the effective pinning energy driving the field gradient up, and vice versa. Such a modulated structure was observed by Batkin and Savchenko,[3] which suggests that the proposed physical picture correctly reflects the experimental situation.

4. CONCLUSION

We have found that the magnetic behavior of the rings corresponds to the low-field dynamics of a Josephson medium.[2,5,15] We have established the main characteristics of such a medium: the dependence of the critical current and the reduced logarithmic-relaxation rate on the constant field, and the dependence of the field at the center of the ring on the external field and on the trapped total current induced by switch-off of the external fields.

We have shown that the introduction of an alternating magnetic field changes the nature of relaxation of the current in the ring induced by the trapped constant external field.

We have experimentally established the dependence of the trapped flux in rings on the amplitude of the alternating field. The experimental data are found to be in good agreement with the results of calculations done on the assumption that the depth of penetration by the alternating field depends on the field amplitude.

We have discovered a change in the nature of the flux creep and the freezing of the trapped flux in the ring. This can be explained by the formation of a modulated field structure in the bulk of the ring when an external alternating field is applied.

Finally, we have used the Kim–Anderson and collective-creep models to calculate the effect of an alternating field on the flux trapped in a thin ring.

The authors are grateful to N. A. Bogolyubov for his useful remarks. The work received financial support from the Russian Ministry of Science and Technological Policies under the Program of Current Aspects of the Physics of Condensed State (Project No. 93 037) and from the Russian Fund for Fundamental Research (Project No. 96-02-19249a).

[1] G. Blatter, M. V. Feigel'man, V. B. Geshkenbein, A. I. Larkin, and V. M. Vinokur, Rev. Mod. Phys. **66**, 1125 (1994).

[2] V. P. Khavronin, S. L. Ginzburg, I. D. Luzyanin, and G. Yu. Logvinova, Sverkhprovodimost': Fiz. Khim. Tekhnol. **5**, 809 (1992) [Superconductivity **5**, 811 (1992)].

[3] V. N. Batkin and O. Ya. Savchenko, Pis'ma Zh. Tekh. Fiz. **19**, No. 7, 29 (1993) [Tech. Phys. Lett. **19**, 409 (1993)].

[4] S. N. Smirnov, S. M. Vatnik, A. N. Smal', and Yu. A. Grishin, Physica C **182**, 261 (1991).

[5] S. L. Ginzburg, V. P. Khavronin, I. D. Luzianin *et al.*, Physica C **174**, 109 (1991).

[6] E. Z. Meılikhov, Sverkhprovodimost': Fiz. Khim. Tekhnol. **2**, 5 (1989) [Superconductivity **2**, 7 (1989)].

[7] V. F. Gantmakher, V. N. Zverev, and V. M. Teplinskiĭ, JETP Lett. **59**, 874 (1994).

[8] É. V. Matizen, P. P. Bezverkhiĭ, V. G. Martynets *et al.*, Sverkhprovodimost': Fiz. Khim. Tekhnol. **7**, 1398 (1994).

[9] P. P. Bezverkhiĭ, N. A. Bogolyubov, and V. G. Martynets, Sverkhprovodimost': Fiz. Khim. Tekhnol. **3**, 469 (1990) [Superconductivity **3**, 444 (1990)].

[10] A. Gurevich and E. H. Brandt, Phys. Rev. Lett. **73**, 178 (1994).

[11] S. L. Ginzburg, Preprint No. 1615, Leningrad Nuclear Physics Institute, Leningrad (1990).

[12] A. A. Zhukov, Sverkhprovodimost': Fiz. Khim. Tekhnol. **5**, 260 (1992) [Superconductivity **5**, 257 (1992)].

[13] A. D. Caplin, G. K. Perkins, and L. F. Cohen, Supercond. Sci. Technol. **8**, 366 (1995).

[14] E. H. Brandt, Int. J. Mod. Phys. **5**, 751 (1991).

[15] S. L. Ginzburg, Zh. Éksp. Teor. Fiz. **106**, 607 (1994) [JETP **79**, 334 (1994)].

Translated by Eugene Yankovsky

Magnetization of a diffusive ring: Beyond the perturbation theory

V. V. Afonin

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, St. Petersburg, Russia

Yu. M. Galperin

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, St. Petersburg, Russia; Department of Physics, University of Oslo, P.O. Box 1048 Blindern, 0316 Oslo, Norway

(Submitted 10 June 1996)
Zh. Éksp. Teor. Fiz. **111**, 1057–1070 (March 1997)

Average persistent current over a set of diffusive metallic rings with fixed number of electrons is considered. We study the case in which the phase breaking time is much greater than an inverse average interlevel distance. In such a case, many return events for an electron must be taken into account. As a result, one arrives at a nonperturbative problem for a cooperon mode fixed by an external magnetic field. This multi-cooperon problem has been considered previously by Altland *et al.*, [Europhys. Lett. **20**, 155 (1992)] and in several following papers within the framework of supersymmetric approach. Such an approach involves very tedious calculations which were performed using a computer algebraic package. Here we solve the problem in question with the help of a replica trick. It is demonstrated that the replica trick in combination with a proper analytical continuation in the replica space allows one to obtain the result in much more explicit way.
[S1063-7761(97)02403-7]

1. INTRODUCTION

Magnetic properties of small conductors were studied extensively during the last several years (see Refs. 1 and 2 and the bibliography cited there). It has been understood that the magnetic moment (and the associated persistent current) induced by an external magnetic flux is a very specific manifestation of mesoscopic behavior. While originally predicted to appear in clean, one-dimensional, metallic rings,[3] most of the recent discussions about persistent currents have been focused on metallic rings that contain impurities.[4] Static magnetic properties of small rings and dots were studied by several authors.[5–11] An important step in the understanding of magnetization of mesoscopic quantum rings took into account the difference between canonical and grand canonical ensembles.[7,10,12–14] It was shown that the magnetization of isolated rings with a fixed number of particles is much larger than that of the ensemble of rings kept under fixed chemical potential. As a result, the main contribution to the magnetic moment was expressed in terms of the fluctuation of the number of particles at fixed chemical potential, $\langle(\delta N)^2\rangle$. The latter quantity was analyzed in Refs. 10 and 14 under the condition $\hbar/\Delta\tau_\phi \gg 1$. Here Δ is an average interlevel distance at the Fermi level, $\Delta^{-1}=\nu V$ (ν is the density of states at the Fermi level, and V is the volume) and τ_ϕ is the phase-breaking time.

Let us discuss the physical meaning of the parameter $\hbar/\Delta\tau_\phi$. As is well known,[15] in the absence of an external magnetic field the quantum correction to the conductivity is proportional to the classical probability W for an electron with a velocity v and momentum p to return to the vicinity of the starting point (more exactly, into the volume of the order of $v\,dt(\hbar/p)^2$, which is important for quantum interference). The probability W is given by the expression[15]

$$W\propto\frac{v\hbar^2}{p^2}\int^{\tau_\phi}dt\,P(\mathbf{r},t)\Big|_{\mathbf{r}=0}, \tag{1}$$

where $P(\mathbf{r},t)$ is the probability density. Here we employ the fact that in a diffusive regime it is a smooth function of coordinates at the scale of the mean free path ℓ. To estimate $P(0,t)$ we take into account that the electron diffusion is restricted by a finite volume of the sample. In such a case, we have

$$P(0,t)\propto\frac{1}{V}\sum_{n,\mathbf{n}_\perp}\exp\left[-D\left(\frac{n^2}{R^2}+\frac{\mathbf{n}_\perp^2}{d_\perp^2}\right)t\right]. \tag{2}$$

Here D is the diffusion constant, R is the radius of the ring, and $d_\perp$ is its transverse dimension. The numbers n, $\mathbf{n}_\perp$ have the meaning of quantum numbers for longitudinal and transverse diffusive modes, respectively. For a thin ring, $d_\perp \ll R$, only $\mathbf{n}_\perp=0$ is important. We can see that at $D\tau_\phi/R^2\gg 1$ the sum over discrete n, $\mathbf{n}_\perp$ in (2) cannot be replaced by an integral. Otherwise, only $n=0$ is important, and $W\sim\tau_\phi\Delta/\hbar$. If this quantity is small, we can restrict the analysis to a single return event.

Let us now concentrate on the case of external magnetic field. In a magnetic field, the number n in the expression (2) must be replaced by $n-\Phi/\Phi_0$, where Φ is the magnetic flux embedded in the ring, and $\Phi_0=\pi\hbar c/e$.[16] It is clear that the quantum contribution is maximal if Φ/Φ_0 is close to an integer number n_0. If the difference $\tilde{n}\equiv n_0-\Phi/\Phi_0=0$, we have the same situation as that in the absence of a magnetic field—only the mode with $n=n_0$ is important. One can expect that this property is also the case at finite $|\tilde{n}|\ll 1$. Indeed, for $n\neq n_0$

$$\delta W\propto\sum_{n\neq n_0}\frac{\Delta}{D\tilde{n}^2/R^2+1/\tau_\phi}\sim\frac{\Delta R^2}{D}$$

(for the last estimate we have assumed $D\tau_\phi/R^2 \gg 1$). Consequently, if $\delta W \ll 1$ one can ignore the contributions of all the modes with $n \neq n_0$ to the probability for the return. However, the corresponding contribution of the mode with $n = n_0$ is not small at $\Delta\tau_\phi/\hbar \gtrsim 1$. Hence, we arrive at the problem of calculating the localization contribution in the case

$$D/R^2 \gg \Delta/\hbar \gtrsim 1/\tau_\phi .$$

In this region we can still use a single-mode approximation, but the perturbation theory involving a single return event fails.

The problem in question was addressed by Altland *et al.*[17,18] (see also Refs. 19–21). The authors used the so-called Q-Hamiltonian approach within the framework of the supersymmetric method. An intrinsic feature of this method is that one has to cancel out specific nonphysical contributions. Therefore, the supersymmetric approach involves tedious algebraic calculations. Consequently, the authors of Refs. 17 and 18 extensively used a computer algebraic package. As a result, the intermediate equations have not been published, because, as it was stated, the computer printout had many pages.

On the other hand, another approach—the so-called replica method—exists.[22] According to this method, one has to replace the system under consideration by N systems which are identical to the original one and at the end tend $N\to 0$. Usually, after such a procedure one obtains relatively simple expressions. The limiting transition $N\to 0$ (if done properly) automatically cancels out the nonphysical contributions, which has to be done explicitly within the supersymmetric approach.

To take the full advantage of this property, one needs a regular procedure to calculate the limit $N\to 0$. The aim of the present paper is to suggest a procedure of analytical continuation of a nonperturbative, two-particle Green's function from integer N to the whole complex plane which includes the point $N = 0$. Such a procedure allows one to calculate the limit rather automatically, without the need of direct cancellation of nonphysical contributions. We obtain an analytical nonperturbative expression for the persistent current in a mesoscopic diffusive ring and compare it with the results of Refs. 17 and 18.

The paper is organized as follows. In Sec. 2 the basic equations for the fluctuation of the number of particles, as well as for the persistent current are analyzed. The effective action in the single-mode approximation is considered in Sec. 3. In Sec. 4 the particle number autocorrelation function and persistent current are calculated in the nonperturbative region, and results are summarized. In the following calculations we set $\hbar = 1$. Then $\hbar$ will be restored in the estimates and final results.

2. BASIC EQUATIONS

According to Ref. 10, the main contribution to the persistent current I can be expressed in terms of the magnetic flux Φ embedded in the ring as follows:

$$I = \frac{c\Delta}{2}\frac{\partial}{\partial\Phi}\langle(\delta N)^2\rangle_{\mu=\langle\mu\rangle}, \tag{3}$$

where $\langle(\delta N)^2\rangle_{\mu=\langle\mu\rangle}$ is the particle number autocorrelation function, calculated at a given value of chemical potential. The latter can be expressed in terms of single-electron Green's functions as[23]

$$\langle(\delta N)^2\rangle = \int_{-\mu}^{0} d\epsilon_1 d\epsilon_2 K(\epsilon_1,\epsilon_2), \tag{4}$$

where

$$K(\epsilon_1,\epsilon_2) = \frac{1}{\pi^2}\int d\mathbf{r}_1 d\mathbf{r}_2 \{\langle \mathrm{Im}\, G^R_{\epsilon_1}(\mathbf{r}_1,\mathbf{r}_1) \times \mathrm{Im}\, G^R_{\epsilon_2}(\mathbf{r}_2,\mathbf{r}_2)\rangle - \langle \mathrm{Im}\, G^R_{\epsilon_1}(\mathbf{r}_1,\mathbf{r}_1)\rangle \times \langle \mathrm{Im}\, G^R_{\epsilon_2}(\mathbf{r}_2,\mathbf{r}_2)\rangle\}. \tag{5}$$

Here $\langle ...\rangle$ means the usual impurity average. The quantity (5) has been calculated in Ref. 10 in the limiting case $\Delta\tau_\phi \ll \hbar$. Our aim is to go beyond this limiting case, i.e. to calculate the correlation function for arbitrary $\Delta\tau_\phi/\hbar$, keeping $p_F\ell \gg \hbar$. For this purpose we employ the method used by Efetov, Larkin, and Khmelnitskii[22] with minor modifications. Namely, we will use the so-called Q-Hamiltonian approach within the framework of the replica trick. The confined expression for the correlation function $K(\omega)(\omega = \epsilon_1 - \epsilon_2)$ can be written in the form (cf. Ref. 22)

$$K(\omega) = \left[\frac{\nu^2}{N^2\int DQ\,\exp(-F)}\int d\mathbf{r}_1 d\mathbf{r}_2 \int DQ e^{-F} \times \mathrm{Tr}[\Lambda Q(\mathbf{r}_1)]\mathrm{Tr}[\Lambda Q(\mathbf{r}_2)]\right]_{N\to 0}, \tag{6}$$

where

$$F = \frac{\pi\nu}{4}\int d\mathbf{z}\,\mathrm{Tr}\left[D\left(\nabla Q + \frac{ie}{c}\mathbf{A}[Q,\Lambda]_-\right)^2 + 2\left(i\omega - \frac{1}{\tau_\phi}\right)\Lambda Q\right]. \tag{7}$$

Here $\mathbf{A}$ is the vector-potential, and $[A,B]_- \equiv AB - BA$. Taking into account only the elastic scattering by short-range, nonmagnetic impurities, we can specify Q as $2N \times 2N$ Hermitian matrices, $Q^2 = 1$, $\mathrm{Tr}\, Q = 0$, N is the number of replicas, while

$$\Lambda = \begin{pmatrix} \hat{1} & 0 \\ 0 & -\hat{1} \end{pmatrix},$$

where $\hat{1}$ is the $N \times N$ unit matrix. The parameter τ_ϕ^{-1} is introduced phenomenologically. We assume that the phase breaking is due to the inelastic processes. Following Ref. 22, we use the parametrization

$$Q = \Lambda\,\exp(W), \quad W = \begin{pmatrix} 0 & B \\ -B^+ & 0 \end{pmatrix},$$

where B is an arbitrary $N \times N$ matrix.

3. EFFECTIVE ACTION IN SINGLE-MODE APPROXIMATION

Let us consider a ring with the radius R and the width $d_\perp \ll R$. We can therefore take into account only the dependence of the matrices B on the angular coordinate φ. Expanding this dependence into the discrete Fourier series, $B = \Sigma_n B_n \exp(in\varphi)$, we introduce the mode number n. As was explained in Sec. 1, only one mode with $n = n_0$ corresponding to $\min(n - \Phi/\Phi_0)$ is important (this assumption will be justified at the end of Sec. 4A). Retaining only this mode and assuming $\nabla_\varphi = (1/R)\partial/\partial\varphi$, we obtain $W\nabla_\varphi W + \nabla_\varphi W W = 0$. Hence,

$$\nabla_\varphi Q = \frac{1}{R}\Lambda\frac{\partial}{\partial\varphi}e^{W} = (\nabla_\varphi W)W^{-1}\sinh W$$
$$= \frac{in_0}{R}\sinh W.$$

We can then expand $\sinh^2 W = (1/2)(\cosh 2W - 1)$ as a series in

$$W^{2k} = (-1)^k \begin{pmatrix} \sqrt{B_{n_0}B^+_{n_0}} & 0 \\ 0 & \sqrt{B^+_{n_0}B_{n_0}} \end{pmatrix}^{2k}.$$

The item $\mathrm{Tr}(\Lambda Q)$ can be treated in a similar way. As a result, we obtain the following expression for F:

$$F = \frac{\pi}{2\Delta}\left[\frac{D}{R^2}\left(n_0 - \frac{\Phi}{\Phi_0}\right)^2 \mathrm{Tr}\sin^2(\sqrt{B_{n_0}B^+_{n_0}}) + 2\left(i\omega - \frac{1}{\tau_\phi}\right)\mathrm{Tr}\cos(\sqrt{B_{n_0}B^+_{n_0}})\right]. \tag{8}$$

We must now consider an important point. An arbitrary complex $N \times N$ matrix B can be described by two Hermitian $N \times N$ matrices. These matrices are defined as

$$B = \rho\exp(i\varphi), \quad B^+ = \exp(-i\varphi)\rho. \tag{9}$$

The quantity F depends only on the matrix ρ. On the other hand, an arbitrary Hermitian matrix could be diagonalized, the eigenvalues being real. One can immediately see that the integral over ρ in the expression (6), with F taken from Eq. (8), for the correlation function diverges. This divergence, in fact, does not occur, because the eigenvalues of ρ must be defined in a finite interval. Indeed, one has to define the variables ρ in a way to obtain a one-to-one correspondence between ρ and Q. On the other hand, one can explicitly show that

$$Q = \Lambda e^{W} = \begin{pmatrix} \cos\sqrt{BB^+} & \dfrac{\sin\sqrt{BB^+}}{\sqrt{BB^+}}B \\ B^+\dfrac{\sin\sqrt{BB^+}}{\sqrt{BB^+}} & -\cos\sqrt{B^+B} \end{pmatrix}$$
$$= \begin{pmatrix} \cos\rho & \sin\rho\, e^{i\varphi} \\ e^{-i\varphi}\sin\rho & -\cos\rho^T \end{pmatrix}, \tag{10}$$

where ρ^T denotes the transposed matrix ρ. Here we have employed the relation

$$\rho^T = \sqrt{e^{-i\varphi}\rho^2 e^{i\varphi}}, \tag{11}$$

which is a consequence of the symmetry properties of the initial replica Hamiltonian (see Appendix A). It is clear that the matrix Q is a periodic function of ρ, and that one has to specify a region at least not larger than one period in order to obtain a one-to-one correspondence. Moreover, to obtain proper analytical properties (damping is the lower semiplane of the ω-variable) of the action F (8), we must define the integration limits as $(-\pi/2, \pi/2)$. Finally, the action F reads as

$$F = \frac{\pi}{2\Delta}\left[\frac{D}{R^2}\left(n_0 - \frac{\Phi}{\Phi_0}\right)^2 \mathrm{Tr}\sin^2\rho + 2\left(i\omega - \frac{1}{\tau_\phi}\right)\mathrm{Tr}\cos\rho\right]. \tag{12}$$

Now let us transform the variables from B, B^+ to ρ, $u \equiv \exp(i\varphi)$, where the Jacobian is (see Appendix B)

$$\frac{D(B,B^+)}{D(\rho,u)} = 2(\det u^{-1}\rho)^N. \tag{13}$$

We see that the variables u can be integrated out and cancelled with the denominator in Eq. (6).

4. PARTICLE NUMBER AUTO-CORRELATION FUNCTION

A. Eigenvalue representation

Let us now return to Eq. (6). Since $\mathrm{Tr}\,\Lambda Q = 2\,\mathrm{Tr}\cos\rho$, we see that the integrand depends only on the eigenvalues of ρ. Hence, we should transform the variables to the eigenvalues and some other ones which could be integrated out in the numerator and the denominator. This transform is outlined in Appendix 5. As a result, we obtain

$$K_N = \frac{(V\nu)^2\int_0^1\{d\lambda\}\theta(\lambda^{(i)})\left[\sum_{j=1}^{N}\cos\dfrac{\pi\lambda^{(j)}}{2}\right]^2}{N^2\int_0^1\{d\lambda\}\theta(\lambda^{(i)})}, \tag{14}$$

where $\{d\lambda\} \equiv \Pi_{i=0}^{N-1}d\lambda^{(i)}|\lambda^{(i)}|^{2i+N}$, and

$$\theta(x) = \exp\left[-\frac{\pi D}{2\Delta R^2}\left(n_0 - \frac{\Phi}{\Phi_0}\right)^2\sin^2\left(\frac{\pi x}{2}\right) - \frac{\pi}{\Delta}\times\left(i(\epsilon_1 - \epsilon_2) - \frac{1}{\tau_\phi}\right)\cos\left(\frac{\pi x}{2}\right)\right]. \tag{15}$$

The expression (15) contains three dimensionless parameters:

$$\gamma \equiv \frac{\pi\hbar}{\Delta\tau_\phi}, \quad \Omega \equiv \frac{\pi(\epsilon_1 - \epsilon_2)}{\Delta}, \quad \mathscr{E} \equiv \frac{\hbar\pi D\tilde{n}^2}{2R^2\Delta}, \tag{16}$$

where $\tilde{n} \equiv (n_0 - \Phi/\Phi_0)$. It is important to keep in mind the following. If $\max(\gamma,\Omega) \gg 1$, only the small λ's are important. Hence, one obtains at the result which is found in the framework of the perturbation theory.[10,14] However, if both γ and Ω are small one must sum the multi-cooperon contributions, which cannot be done in the framework of the perturbation theory. There is a substantial simplification in the case

$$\tilde{n} \ll 1, \quad \text{but} \quad \mathscr{E}_c \equiv \frac{\hbar\pi D}{2R^2\Delta} \gg 1. \tag{17}$$

In this case, only one mode with $|n_0 - \Phi/\Phi_0| \ll 1$ is important; this is the case in which Eq. (15) is valid. Consequently, we consider the case in which the inequalities (17) hold, but the quantities γ and Ω can be arbitrary. In fact, the mode n_0 must be considered in a nonperturbative manner, while the other modes can be treated in the framework of the perturbation theory.

B. Analytical continuation

We are not able to analytically calculate the expression (14) for an arbitrary N. Instead, we will perform analytical continuation of this expression to arbitrary N, and then calculate its limit as $N \to 0$.

We introduce the quantity

$$Z_N = \prod_{k=0}^{N-1} \int_0^1 dx_k x_k^{2k+N} \theta(x_k), \tag{18}$$

where

$$\theta(x) = \exp[-\mathscr{E} \sin^2(\pi x/2) - (i\Omega - \gamma)\cos(\pi x/2)]. \tag{19}$$

It is convenient to define

$$\zeta_N \equiv \ln Z_N \equiv \zeta^R + \zeta^A,$$

where

$$\zeta^R \equiv \frac{1}{2} \sum_{k=\frac{N+1}{2}}^{\frac{3N-1}{2}} \ln \int_0^1 dx\, x^{2k-1+\delta}\theta(x),$$
$$\zeta^A \equiv \frac{1}{2} \sum_{k=-\frac{N+1}{2}}^{-\frac{3N-1}{2}} \ln \int_0^1 dx\, x^{-2k-1+\delta}\theta(x), \tag{20}$$

δ is a small positive number which later will vanish. We introduce this parameter to keep the important integrals convergent in the limit $N \to 0$. The first step is to express the sum over k in terms of the contour integral over complex k. For this purpose we should keep in mind that the derivative $\partial f[2\pi i(k+1/2)]\partial k$ [where $f(z) = (e^z + 1)^{-1}$] has second-order poles at integer numbers. Consequently, one can express $\zeta_{R(A)}$ as

$$\zeta_{R(A)} = \int_{C^{\pm}} dk \left(\frac{\partial f(2\pi i k)}{\partial k} \right) F^{\pm}(k), \tag{21}$$

$$F^{\pm}(k) = \frac{1}{2} \int^k dk' \ln\left[\int_0^1 dx\, x^{\pm 2k' - 1 + \delta} \theta(x) \right]. \tag{22}$$

The contours $C^{\pm}$ are shown in Fig. 1. These expressions are correct only if other singularities, except for the poles of f, are not important. We can show that the function $F^{+}(k)$ has singularities only in the left-hand semi-plane of the complex variable k, while the function F^{-} has singularities only in the right-hand semi-plane of k. To prove this statement, one must expand the function $\theta(x)$ in a Taylor series. For the following, it is convenient to rotate the k-plane through $\pi/2$ by introducing a new variable, $k_1 \equiv 2\pi i k$. The transformed contours $\overline{C}^{\pm}$ are shown in Fig. 2.

Making use of exponential convergence of the integral due to the properties of $\partial f/\partial k$, we transform the contour integrals to the integrals along the real axis. For simplicity let us assume that N is even. As a result, we have

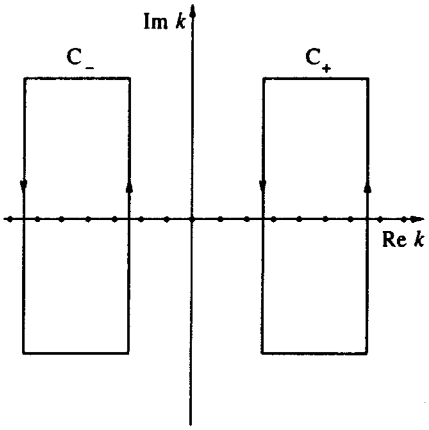


FIG. 1. The integration contours $C_{\pm}$ for $N = 4$.

$$\zeta_{R(A)} = \mp \int_{-\infty}^{\infty} dk \left(\frac{\partial f(k)}{\partial k} \right) \left[F^{\pm}\left(\frac{k \pm i\pi N}{2\pi i} \right) - F^{\pm}\left(\frac{k \pm 3i\pi N}{2\pi i} \right) \right]. \tag{23}$$

We can now perform an analytical continuation over N. For the functions $F^{\pm}$ the continuation must be done in a different way for the reason to be discussed later. For this purpose we replace iN by $\pm N_0$ in the functions $F^{\pm}$, respectively. Here N_0 is a real, positive quantity which will eventually tend to zero. Finally, we have

$$\zeta_{N_0} = \int_{-\infty}^{\infty} dk \left[\left(\frac{\partial f(k - \pi N_0)}{\partial k} \right) - \left(\frac{\partial f(k - 3\pi N_0)}{\partial k} \right) \right] \times \left[F^{-}\left(\frac{k}{2\pi i} \right) - F^{+}\left(\frac{k}{2\pi i} \right) \right]. \tag{24}$$

As a result, the lowest-order term in the N_0-expansion of the function ζ_{N_0} is $\propto N_0^2$. Finally, we obtain

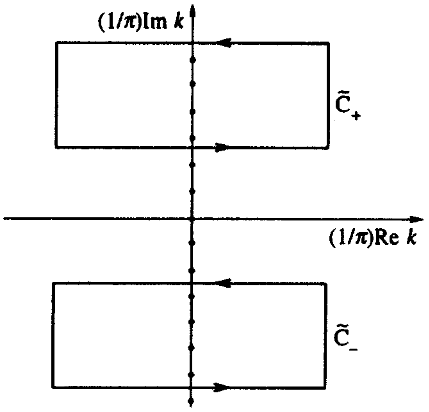


FIG. 2. The integration contours $\widetilde{C}_{\pm}$ for $N = 4$.

$$\zeta = 2\pi i N_0^2 \int_{-\infty}^{\infty} dk \left(\frac{\partial^2 f(k)}{\partial k^2} \right) \ln \left[\int_0^1 dx \; x^{-ik/\pi - 1 + \delta} \theta(x) \right]. \tag{25}$$

The reason for splitting the function ζ_N into ζ^R and ζ^A with the replacements $N \rightarrow \pm iN_0$ is as follows. As N tends to zero, the integration contour comes infinitely close to the cut of the logarithm functions in the expressions for $F^{\pm}$. Such a situation is not the case for any finite N, and it leads to a nonphysical pinch which should be subtracted. Within the above-mentioned procedure such a contribution is purely imaginary, while the one of interest to us is real. The imaginary contributions to F^+ and F^- have opposite signs. Thus, the nonphysical contribution is automatically cancelled in the sum $\zeta^R + \zeta^A$. We note that these terms are of the first order in N_0, and that they must vanish; otherwise, the two-particle Green's function would be divergent.

In fact, a similar trick has been used by Matsubara to formulate the thermal Green's function technique (see, e.g., Ref. 24). Let us compare our analytical continuation of the function ζ to the analytical continuation of the two-particle Matsubara Green's function $K(\Omega_m)$, where Ω_m is the external Matsubara frequency. In each case one must use two functions, which are regular in the upper (retarded) and the lower (advanced) semi-plane, respectively. The two above-mentioned functions can then be combined into one with a cut in the complex plane. The physical reason for such a splitting into R and A parts is to cancel the nonphysical contributions. In the Matsubara case the nonphysical contributions to $K(\Omega_m)$ arise at the point $\Omega_m \rightarrow 0$ and cancel out after a similar continuation $\Omega_m \rightarrow i\Omega$ of the sum over Ω.

C. Persistent current

Following Ref. 10, we express the current according to Eqs. (3) and (4). On the other hand,

$$K(\epsilon_1, \epsilon_2) \propto e^{-\zeta} \frac{\partial^2}{\partial \epsilon_1 \partial \epsilon_2} e^{\zeta}.$$

Finally, we obtain

$$I = -\frac{c\Delta}{2} \frac{\partial}{\partial \Phi} \lim_{N_0 \rightarrow 0} \frac{\zeta_{N_0}(\epsilon_1 = \epsilon_2 = 0)}{N_0^2} = -J_0 \tilde{n} \mathscr{G}, \tag{26}$$

where $J_0 = \hbar c D / R^2 \Phi_0 = eD/\pi R^2$, and

$$\mathscr{G} = i \int_{-\infty}^{\infty} dk$$
$$\times \frac{\partial^2 f(k)}{\partial k^2} \frac{\int_0^1 dx \; \sin^2(\pi x/2) x^{-ik/\pi - 1 + \delta} \theta(x, \Omega = 0)}{\int_0^1 dx \; x^{-ik/\pi - 1 + \delta} \theta(x, \Omega = 0)}. \tag{27}$$

One can check directly that at

$$\Delta \ll \frac{\hbar}{\tau_\phi} \ll \frac{\hbar D}{R^2}, \quad |\tilde{n}| \ll 1$$

the above expressions lead to the expressions obtained in Refs. 10 and 14. One can calculate the integrals with the help of the steepest-descent method to find[10,14]

$$I = -J_1 \tilde{n}, \quad J_1 = \frac{e\Delta}{\pi^3} \frac{D\tau_\phi}{R^2}. \tag{28}$$

In the region

$$\gamma \ll 1, \quad \mathscr{E}_c \gg \mathscr{E} \gg 1$$

one can also develop a perturbation theory. Indeed, only small values of x in the integrals in Eq. (27) are important. The physical reason for this is the magnetic-field-induced phase breaking. In this region we obtain the result

$$\mathscr{G} = \frac{1}{\pi^2 \tilde{n}^2} \frac{R^2 \Delta}{\hbar D}, \quad I = -\frac{1}{\pi^3 \tilde{n}} \frac{e\Delta}{\hbar}. \tag{29}$$

This result agrees with the asymptotic result of Ref. 14 for $\gamma \gg 1$, $\sqrt{R^2/D\tau_\phi} \ll \tilde{n} \ll 1$. Note that the result (29) obtained for $\gamma \ll 1$ is valid in the region

$$\sqrt{R^2 \Delta / \hbar D} \ll \tilde{n} \ll 1.$$

For the case $\gamma \ll 1$, $\mathscr{E} \ll 1$, where the perturbation theory is not applicable, we set $\theta(x, \Omega = 0) = 1$. As a result, we obtain $\mathscr{G} = 0.21$, and the current is

$$I = -0.21 \frac{eD}{\pi^3 R^2} \tilde{n}. \tag{30}$$

We observe a maximum at $\tilde{n} \sim \sqrt{R^2 \Delta / \hbar D}$, where the maximal current is

$$I_{\max} \sim e \sqrt{\Delta D / \hbar R^2}. \tag{31}$$

Expressions (29) and (30) are fully consistent with the curve calculated in Refs. 17 and 18 with the help of a computer algebraic package. Let us consider the dependence of the maximal current on $\gamma \approx \hbar / \Delta \tau_\phi$. At $\gamma \gg 1$ the perturbation theory[10,14] predicts the maximum of the current at $\tilde{n} \sim \sqrt{R^2 / D \tau_\phi}$; the maximum value is

$$I_{\max} \sim e \frac{\Delta}{\hbar} \sqrt{\frac{\tau_\phi D}{R^2}}. \tag{32}$$

Consequently, $I_{\max} \propto \gamma^{-1/2}$ at $\gamma \gg 1$, and it is γ-independent at $\gamma \ll 1$. In this region we estimate the persistent current to be

$$J_{\max} \sim \frac{ev_F}{R} \sqrt{\frac{\Delta \tau_{\mathrm{el}}}{\hbar}},$$

where τ_{el} is the elastic relaxation time. The quantity $\Delta \tau_{\mathrm{el}} / \hbar$ for a typical metal can be estimated as $(\ell / R)(a^2 / A)$, where ℓ is the mean free path, a is a typical interatomic distance, and A is the cross section of the ring. Equation (30) shows that at $\Delta \tau_\phi / \hbar \gg 1$ the phase-breaking time τ_ϕ is not contained in the expression for the persistent current. Returning to Eq. (1), we must conclude that at $t \gtrsim \hbar / \Delta$ the electronic wave packet is not smeared in space. This means that the n_0-mode of the cooperon is localized in a sense; the localization length is of the order of $\sqrt{\hbar D / \Delta}$. Of course, this does not mean localized, because other modes are still under weak localization conditions.

The range of parameters in which the theory is applicable and where it leads to nontrivial results can be expressed as follows:

$$1 \ll R / \ell \ll K, (1/K)(\tau_\phi / \tau_{\mathrm{el}}),$$

where $K \sim (pd_\perp/\hbar)^2$ is the number of transverse channels. The left inequality is the criterion for a diffusive motion, the first right inequality is the Thouless criterion $\mathscr{E}_c \gg 1$, and the last right inequality is the condition $\Delta\tau_\phi/\hbar \gg 1$. We see that one needs low temperatures to satisfy the inequality $\tau_\phi/\tau_{\text{el}} \gg K \gg 1$, and also samples of very small size. No previous experiments, to the best of our knowledge, satisfy this set of conditions.

5. DISCUSSION

As one can see from the preceding sections, the results of the replica procedure which are complicated for arbitrary integer N, are rather simple in the limit $N\to 0$. In this limit the nonphysical contributions are cancelled automatically, while in the supersymmetric method this has been done explicitly. An important feature which leads to such a simplification is the procedure of the analytical continuation which is done before direct calculations. Specifically, one has two functions which are analytical in the upper (lower) semiplane of the complex plane of N, respectively. The proper analytical continuation is a combination of these two functions. Consequently, it has a cut at $\operatorname{Im} N = 0$. The procedure used above allows one to cancel automatically the nonphysical pinch in the two-particle Green's function, which otherwise would exist at $N_0 = 0$. We believe that such a construction is important, in general, for the calculations involving the replica trick. In such a way we reproduce analytically and rather simply the results obtained in Refs. 17 and 18 by a computer algebraic package.

We wish to thank V. L. Gurevich, V. Yu. Petrov, and A. D. Mirlin for valuable comments. One of the author (V.V.A.) is grateful to the Research Council of Norway for a financial support within the Cultural Exchange Program (KAS).

APPENDIX A: EFFECTIVE ACTION-DERIVATION

Here we rederive the expression (6) following Ref. 22, in order to clarify important symmetry properties. Following Ref. 22 we use the replica trick and introduce field operators:

$$\boldsymbol{\Psi} = \{\psi_1,\dots,\psi_N,\psi_1^+,\dots,\psi_N^+\}, \qquad \boldsymbol{\Psi}^+ = \begin{pmatrix} \psi_1^+ \\ \cdots \\ \psi_N^+ \\ -\psi_1 \\ \cdots \\ -\psi_N \end{pmatrix}.$$

Here $\psi_i\psi_i^+ + \psi_i^+\psi_i = 0$. The action can be written as follows:

$$F = i\int (d\mathbf{r})\boldsymbol{\Psi}^+(\mathbf{r})(\hat{E}-\hat{\mathscr{H}})\boldsymbol{\Psi}^+(\mathbf{r}),$$

$$\hat{E} = E\hat{I}, \quad \hat{\mathscr{H}} = [H_0 + U_{\text{el}}(\mathbf{r})]\hat{I} - \left(\frac{\omega}{2} + i\delta\right)\Lambda. \tag{A1}$$

Here H_0 is the free-electron Hamiltonian, $\hat{I}$ is the $2N \times 2N$ unit matrix, and $U_{\text{el}}(\mathbf{r}) = U_0\Sigma_i^M \delta(\mathbf{r} - \mathbf{r}_i)$, where M is the total number of impurities. The first N rows of $\hat{\mathscr{H}}$ describe the evolution of the retarded Green's functions, while the last N rows describe an evolution of the advanced Green's functions. The following step is the averaging over the positions of the impurities. We have

$$\Sigma \equiv \prod_{i=1}^{M} \int \frac{d\mathbf{r}_i}{V} \exp\left[iU_0 \sum_{i=1}^{M} \boldsymbol{\Psi}^+(\mathbf{r}_i)\boldsymbol{\Psi}(\mathbf{r}_i)\right]$$
$$= \left[\int \frac{d\mathbf{r}}{V} \prod_{f=1}^{2N} (1 + iU_0\psi_f^+(\mathbf{r})\psi_f(\mathbf{r}))\right]^M. \tag{A2}$$

Here we have taken into account that only the linear terms in ψ_f^+ and ψ_f can enter the continual integral for the correlation function (Grassman algebra). For the same reason, one must allow for only the terms with different f when calculating the product. For a weak scattering and in the thermodynamic limit $M,V\to\infty$, $M/V = \text{const}$,

$$\Sigma \approx \exp(\delta\mu + i\Gamma),$$

$$\delta\mu = \frac{MU_0}{V}\int d\mathbf{r}\boldsymbol{\Psi}^+(\mathbf{r})\boldsymbol{\Psi}(\mathbf{r}),$$

$$\Gamma = \frac{g_0^2}{2}\int d\mathbf{r} \sum_{f\neq g} \psi_f^+(\mathbf{r})\psi_f(\mathbf{r})\psi_g^+(\mathbf{r})\psi_g(\mathbf{r}). \tag{A3}$$

Here $\delta\mu$ is a shift in the chemical potential, and $g_0^2 = 2MU_0^2/V$ (g_0 is the coupling constant). By analogy with Ref. 22, we introduce an auxiliary scalar field which is represented by Hermitian matrices Q. As a result, the effective ψ^4 interaction can be decoupled as follows:

$$\exp\left[-\frac{g_0^2}{2}\int d\mathbf{r}\sum_{f\neq g}\psi_f^+(\mathbf{r})\psi_f(\mathbf{r})\psi_g^+(\mathbf{r})\psi_g(\mathbf{r})\right] = \frac{\int \mathscr{D}Q\, \exp[-\operatorname{Tr}\int d\mathbf{r}(\pi\nu/4\tau_{\text{el}})Q^2(\mathbf{r}) - (1/2\tau_{\text{el}})\boldsymbol{\Psi}^+(\mathbf{r})Q\boldsymbol{\Psi}(\mathbf{r})]}{\int \mathscr{D}Q\, \exp[-\operatorname{Tr}\int d\mathbf{r}(\pi\nu/4\tau_{\text{el}})Q^2(\mathbf{r})]}. \tag{A4}$$

Here we used the definition $2\pi\nu g_0^2\tau_{\text{el}} = 1$. This expression is the same as Eq. (17) from Ref. 22. To analyze the symmetry properties of the impurity-averaged Hamiltonian, we take into account that the initial Hamiltonian possesses the property $\mathscr{H}_{ij}|_{i,j<N} = \mathscr{H}^*_{i+N,j+N}$. This property must be kept after impurity averaging and after introducing the field Q. In terms of Q it reads as $iQ_{ij} = -iQ^*_{i+N,j+N}$. Taking into account Eq. (10), we obtain the relation (11). The following steps are exactly the same as those in Ref. 22.

APPENDIX B: CALCULATION OF THE JACOBIAN

Let us arrange the columns of the $2N^2 \times 2N^2$ matrix $\partial(B,B^+)/\partial(\rho,\mu)$ as

$$\{B_{11},\dots,B_{N1};B_{12},\dots,B_{NN};B^+_{11},\dots,B^+_{NN}\}$$

and the rows as

$$\{\rho_{11},\dots,\rho_{N1};\rho_{12},\dots,\rho_{N2},\dots,\rho_{NN};$$
$$u_{11},\dots,u_{N1};u_{12},\dots,u_{N2},\dots,u_{NN}\}.$$

Taking into account the matrix identities

$$dB = d\rho u + \rho du, \quad dB^+ = u^{-1}d\rho - u^{-1}du\ u^{-1}\rho, \tag{B1}$$

we express $\partial(B,B^+)/\partial(\rho,u)$ as follows:

$$\frac{\partial(B,B^+)}{\partial(\rho,u)} = \begin{pmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22}, \end{pmatrix}. \tag{B2}$$

where $\hat{A}_{ik}$ are the $N^2 \times N^2$ matrices. One can show that

$$\hat{A}_{11} = u \times \hat{1}, \quad \hat{A}_{22} = -(u^{-1}\rho) \times u^{-1},$$

$$\hat{A}_{21} = \underbrace{\rho \otimes \rho \otimes \dots \rho,}_{N} \quad \hat{A}_{12} = \underbrace{u^{-1} \otimes u^{-1} \otimes \dots u^{-1},}_{N} \tag{B3}$$

Here × means the Kronecker product, and ⊗ means the direct product.[25] Making use of the identity[25]

$$\det(A \times B) \equiv (\det A)^p (\det B)^q,$$

(where q and p are the ranges of the matrices A and B, respectively), and of the Laplace expansion of the determinant,[25] we obtain Eq. (13).

APPENDIX C: VARIABLE TRANSFORMATION

Let us consider the set of variables which includes the eigenstates $\lambda_{(b)}$ and N eigenvectors $\mathbf{X}^{(b)}$. One can see from the definition $\rho_{ik}X_k^{(b)} = \lambda^{(b)}\delta_{ik}X_i^{(b)}$ that any vector of the type $e^{i\chi^{(b)}}\mathbf{X}^{(b)}$ (where $\chi^{(b)}$ is an arbitrary phase) satisfies the equation with the same $\lambda^{(b)}$ and ρ_{ik}. Consequently, one must exclude N extra variables $\chi^{(b)}$. We therefore require the diagonal elements $X_i^{(i)}$ to be real. Consequently, the matrix $X_k^{(i)}$ can be constructed according to the following procedure. The first column, $X_i^{(1)}$, contains $N-1$ variables $X_i^{(1)}$, $i \neq 1$, while the last (real) column, $X_1^{(1)}$, is calculated from the requirement of normalization. In the next column, $X_i^{(2)}$, the last $N-2$ variables are chosen independently. The element $X_1^{(2)}$ is determined by the orthogonality of the vectors $\mathbf{X}^{(2)}$ and $\mathbf{X}^{(1)}$, while the last element, $X_2^{(2)}$ is determined by the normalization of $|\mathbf{X}^{(2)}|$. The following elements are determined by continuation of this procedure. Note that since all the off-diagonal elements are complex, we can consider the real ($U_k^{(i)}$) and the imaginary ($V_k^{(i)}$) parts. In this way we can present N^2 independent elements of the matrix ρ in terms of N eigenvalues $\lambda^{(i)}$, and $N^2 - N$ independent variables $U_k^{(i)}$ and $V_k^{(i)}$. From the definition, $\rho_{ij} = \Sigma_k X_i^{(k)}\lambda^{(k)}X_j^{(k)*}$, we can express ρ in terms of $\{U,V,\lambda\}$ as

$$\frac{\partial\rho_{ij}}{\partial U_i^{(k)}} = \lambda^{(k)}[U_j^{(k)}(1+\delta_{ij}) - iV_j^{(k)}(1-\delta_{ij})],$$

$$\frac{\partial\rho_{ij}}{\partial V_i^{(k)}} = \lambda^{(k)}[iU_j^{(k)}(1-\delta_{ij}) + V_j^{(k)}(1+\delta_{ij})],$$

$$\frac{\partial\rho_{ij}}{\partial\lambda^{(k)}} = U_i^{(k)}U_j^{(k)} + V_i^{(k)}V_j^{(k)} + i(V_i^{(k)}U_j^{(k)} - U_i^{(k)}V_j^{(k)}). \tag{C1}$$

Note that the above formulas do not contain summation over repeated superscripts. To calculate the Jacobian, we arrange the corresponding $N^2 \times N^2$ transformation matrix in the following way. The columns are labeled by N^2 ⟨⟨old⟩⟩ variables

$$\{\rho_{i1},\rho_{i2},\dots,\rho_{iN}\}.$$

The rows are labeled by N^2 ‘‘new’’ variables

$$\underbrace{\{U^{(1)}_{k>1}\}}_{N-1},\underbrace{\{U^{(2)}_{k>2}\}}_{N-2},\dots,\underbrace{\{U^{(N-1)}_{N}\}}_{1}; \quad \underbrace{\{V^{(1)}_{k>1}\}}_{N-1},\underbrace{\{V^{(2)}_{k>2}\}}_{N-2},\dots,\underbrace{\{V^{(N-1)}_{N}\}}_{1};\underbrace{\{\lambda^{(k)}\}}_{N}.$$

Consequently, as follows from Eq. (33), the first $(N-1)$ rows contain the common factor $\lambda^{(1)}$ times the quantities which depend only on $\{U,V\}$. The next $(N-2)$ lines contain the factor $\lambda^{(2)}$, etc. The last N lines are $\{\lambda\}$-independent. As a result, the Jacobian can be expressed as $\Pi_{i=1}^N[\lambda^{(i)}]^{2(N-i)}$ × (some function of $\{U,V\}$). This expression must be multiplied by $(\det\rho)^N = (\Pi_{i=1}^N\lambda^{(i)})^N$, and we obtain Eq. (14).

[1] Y. Imry, in *Quantum Coherence in Mesoscopic Systems,* ed. by B. Kramer, Plenum, New York (1991).

[2] T. Chakraborty, Comments Cond. Mat. Phys **16**, 35 (1992).

[3] I. O. Kulik, JETP Lett. **11**, 275 (1970).

[4] L. P. Levy, G. Dolan, J. Dunsmuir, and H. Bouchiat, Phys. Rev. Lett. **64**, 2074 (1990).

[5] V. Chandrasekhar, R. A. Webb, U. J. Brady, M. B. Keta, W. J. Gallagher, and A. Kleinsasser, Phys. Rev. Lett. **67**, 3578 (1991).

[6] M. Büttiker, Y. Imry, and R. Landauer, Phys. Lett. A **96**, 365 (1983).

[7] H. Bouchiat and G. Montambaux, J. Phys. (Paris) **50**, 2695 (1989).

[8] A. Schmid, Phys. Rev. Lett. **66**, 80 (1991).

[9] F. Oppen and E. K. Riedel, Phys. Rev. Lett. **66**, 84 (1991).

[10] B. L. Altshuler, Y. Gefen, and Y. Imry, Phys. Rev. Lett. **66**, 88 (1991).

[11] V. Ambegaokar and U. Eckern, Phys. Rev. Lett. **65**, 381 (1990).

[12] R. Landauer and M. Büttiker, Phys. Rev. Lett. **54**, 2049 (1985); M. Büttiker, Ann. N.Y. Acad. Sci. **480**, 194 (1986).

[13] H. F. Cheung, Y. Gefen, E. K. Riedel, and W.-H. Shih, Phys. Rev. B **37**, 6050 (1988).

[14] A. Kamenev and Y. Gefen, Phys. Rev. Lett. **70**, 1976 (1993).
[15] A. I. Larkin and D. E. Khmel'nitskii, Usp. Fix. Nauk **136**, 536 (1982) [Sov. Phys. Usp. **25**, 185 (1982)].
[16] A. G. Aronov and Yu V. Sharvin, Rev. Mod. Phys **59**, 755 (1987).
[17] A. Altland, S. Iida, A. Müller-Groeling, and H. A. Weidenmüller, Europhys. Lett. **20**, 155 (1992).
[18] A. Altland, S. Iida, A. Müller-Groeling, and H. A. Weidenmüller, Annals of Physics, **219**, 148 (1992).
[19] R. A. Serota and A. Zyuzin, Phys. Rev. B **47**, 6399 (1993).
[20] A. Pandey and M. L. Mehta, J. Phys. A **16**, 2655 (1983).
[21] A. Altland, S. Iida and K. B. Efetov, J. Phys. A **26**, 3545 (1993).
[22] K. B. Efetov, A. I. Larkin, and D. E. Khmelnitskii, Zh. Éksp. Teor. Fiz. **79**, 1120 (1980) [Sov. Phys. JETP **52**, 568 (1980)].
[23] B. L. Altshuler and B. I. Shklovskii, Zh. Éksp. Teor Fiz. **91**, 220 (1986) [Sov. Phys. JETP **64**, 127 (1986)].
[24] A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinskii, *Methods of Quantum Field Theory in Statistical Physics*, Dover Publications, Inc., New York (1975).
[25] *Handbook of Applicable Mathematics, Volume I: Algebra*, ed. by Ledermann and S. Vajda, University of Sussex (A. Wiley-Intersciences Publication, New York 1980).

Published in English in the original Russian journal. Reproduced here with stylistic changes by the Translation Editor.

Electrodynamics of hard superconductors in crossed magnetic fields

I. F. Voloshin, A. V. Kalinov, S. E. Savel'ev, and L. M. Fisher

All-Russian Electrical Engineering Institute, 111250 Moscow, Russia

V. A. Yampolskiĭ

Institute for Radiophysics and Electronics, 310085 Khar'kov, Ukraine

F. Perez Rodriguez

Institute of Physics of Puebla Autonomous University, Mexico

(Submitted 18 July 1996)
Zh. Éksp. Teor. Fiz. **111**, 1071–1084 (March 1997)

A model systematically accounting for the cutting of Abrikosov flux lines has been developed for the critical state of a hard superconductor in crossed dc and ac magnetic fields. The electrodynamic equations have been derived by minimizing the Gibbs free energy calculated using the proposed two-velocity hydrodynamic model. One velocity describes the motion of the vortex lattice as a whole, and the other describes the relative motion of the two intersecting sublattices. The resulting equations yield as special cases the previously known electrodynamic equations for hard superconductors. The model provides a natural explanation for the suppression of dc magnetization by a transverse ac magnetic field observed in our experiments. © *1997 American Institute of Physics.* [S1063-7761(97)02503-1]

1. INTRODUCTION

The problem of describing the electrodynamic properties of hard superconductors has noat been completely solved to this day. The static and quasi-static properties of these materials are usually described using an equation proposed by Bean[1] for the distribution of the magnetic field H:

$$\frac{dH}{dx}=\pm\frac{4\pi}{c}J_c, \qquad (1)$$

where J_c is the critical current density and c is the speed of light. This equation has allowed the dc magnetization of hard superconductors to be calculated, and determines the shapes of hysteresis loops and the energy loss due to remagnetization. This equation has been used in describing low-frequency properties of superconductors. It was employed in developing a technique for remote measurements of the critical current density J_c, and determination of its dependence on the magnetic field, temperature and other parameters using measurements of the surface impedance, ac magnetic susceptibility, third harmonic in the superconductor response to an ac magnetic field, etc.[2–5]

The Bean equation was verified by analyzing the balance of forces acting on magnetic flux lines. According to this model, called a critical-state model, the magnetic force acting on a vortex is balanced by the pinning forces due to the interaction of vortices with crystal lattice irregularities, precipitates of other phases, sample boundaries, etc.[6] Owing to pinning, the critical current density J_c is nonzero, and the magnetic field distribution is fairly accurately described by Eq. (1).

Presently the critical state equation is usually defined as the Maxwell equation for magnetic field combined with the material equation $\mathbf{J}=\mathbf{J}_c$.[7] In most cases when the configuration is simple, the direction of the vector $\mathbf{J}_c$ is determined easily. The problem is more complicated when the applied magnetic field consists of several components or its orientation is variable. Some authors (see, for example, Ref. 8) suggested for calculations of electromagnetic parameters a generalized critical-state equation in the form

$$\text{curl }\mathbf{B}=\frac{4\pi}{c}J_c\frac{\mathbf{E}}{E}, \qquad (2)$$

where **E** is the electric field and **B** is the magnetic induction. This equation can be understood by analyzing the current-voltage characteristic of a hard superconductor. A change in the magnetic flux within the superconductor generates an electric field and converts the sample to a resistive state. One can assume that, as in the case of a normal metal, the current should be aligned with the electric field. By taking the limit $E\rightarrow 0$, one obtains Eq. (2). This equation has been used in interpreting several effects associated with the nonlinear interaction between electromagnetic waves.[9,10]

Description of superconductors in crossed magnetic fields is of special interest.[8,11] Recently[12] we have detected the collapse of the dc magnetic moment M of a hard superconductor subjected to an ac magnetic field $\mathbf{h}=\mathbf{h}_0\cos\omega t$ orthogonal to the dc magnetic field **H**. Both fields were aligned with the superconducting plate surface. We found that the dc magnetic moment decreases markedly under the ac magnetic field $h_0 \ll H$, and the hysteresis on the $M(H)$ curve disappears.

There are two alternative approaches to the electrodynamics of superconductors in crossed magnetic fields. One of them is based on Eq. (2). In accordance with this equation, shielding currents, generating the dc magnetization, are driven by the ac magnetic field inwards, which leads to a decrease in the magnetic moment. A more consistent approach can be based on the features of the vortex system in crossed magnetic fields due to crossing of magnetic flux lines, which is also called *flux cutting*.[13,14] It would be natural to account for the collapse of the dc magnetization in terms of electrodynamic equations based on the model of flux cutting,[15,16] which has been successfully used in interpreting several experiments in crossed magnetic fields.[17,18] These equations describe the magnitude of the magnetic flux density B and its tilt angle ϑ with respect to a selected z-axis as functions of coordinates and are a generalized version of the Bean model by Eq. (1). In the simplest case, they can be expressed in the form

$$\frac{\partial B}{\partial x}=\frac{4\pi}{c}J_c,\quad \frac{\partial \vartheta}{\partial x}=k_c, \tag{3}$$

where J_c is the critical current density similar to that in Eq. (1), and k_c is a phenomenological parameter determined by the current component aligned with the vector **B**. Unfortunately, these equations cannot account for the collapse of the dc magnetization.

Thus, we have a paradoxical situation, when Eq. (2), which has no proper theoretical justification and ignores important features of the vortex system, describes the collapse of the dc magnetization, whereas the equations based on the physical properties of the vortex system in crossed magnetic fields yield a description in which this magnetic moment does not strongly depend on the ac field. One can ascribe this paradox to the fact that the electromagnetic equations have not been consistently derived from the model of crossing magnetic vortex filaments and are purely phenomenological.

The aim of the present study was to develop an electromagnetic model adequately taking into account the features of the vortex system in crossed magnetic fields and describing the above effect. The next section will present a detailed description of the experiment,[12] which was a crucial test of existing models of hard superconductor electrodynamics in crossed magnetic fields. Then the equation system which successfully describes the collapse of the dc magnetic moment and also describes other existing models as limiting cases will be derived using a variational technique.

2. EXPERIMENT

An efficient method for studying the behavior of hard superconductors in crossed magnetic fields is measuring the effect of ac magnetic field on dc magnetization curves. Measurements were performed with plates of textured high-T_c superconductor of the Y–Ba–Cu–O (YBCO) family cut from a bulk piece ($4\times7\times60$ mm^3). The plate surface coincided with the **ab** crystal plane. The sample microstructure had a brick-wall shape, which is typical of fused YBCO materials. The temperatures of the onset and completion of the superconducting transition were 92 and 91.3 K, respectively, and the critical current density at $T=77$ K in a magnetic field of 10 kOe was about 10^4 A/cm^2.

The magnetization curves $M(H)$ were recorded using a vibrating-sample magnetometer. A dc magnetic field of up to 10 kOe was generated by an electromagnet and directed parallel to the sample surface. An ac magnetic field $h(t)=h_0\cos\omega t$ with the amplitude h_0 of up to 600 Oe and a frequency $\omega/2\pi$ ranging between 100 and 2000 Hz was produced by a solenoid. The ac magnetic field $\mathbf{h}(t)$ was directed perpendicular to the dc field **H** in the sample plane.

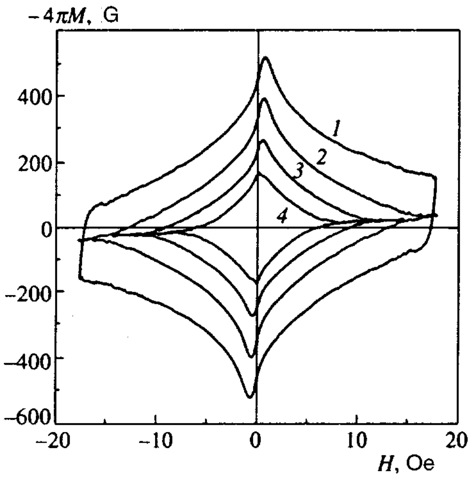


FIG. 1. Evolution of magnetization loops with increasing ac field amplitude: (*1*) $h_0=0$; (*2*) 100 Oe; (*3*) 200 Oe; (*4*) 300 Oe; $\omega/2\pi=130$ Hz, $T=77$ K.

The effect of the ac magnetic field on the dc magnetic moment is illustrated by Fig. 1. The curves of $M(H)$ were recorded in a sample with dimensions of $8\times4\times0.4$ mm and correspond to different amplitudes of the ac field. One can see that the hysteresis of the magnetization is smaller at higher h_0. Moreover, for large H the magnetization hysteresis is totally suppressed and cannot be seen on some parts of the magnetization curve. The share of these parts increases markedly with h_0. Some estimates have demonstrated that these parts of the magnetization curve correspond to H and h_0 for which the ac magnetic field penetrates throughout the sample volume. The effect has been shown to be insensitive to the ac field frequency. It is noteworthy that the suppression of the dc magnetization caused by the ac field is irreversible, i.e., when the ac magnetic field is turned off at an arbitrary dc field $H=H_1$, the magnetic moment is frozen and does not return to its initial value corresponding to $M(H_1)$ on curve *1* (Fig. 1).

These experimental results provide indirect evidence in favor of the following description of direct and alternating shielding currents in the sample. Alternating currents induced by the applied ac magnetic field $h(t)$ push the direct shielding current from the surface layer into the sample interior. As a result, the regions containing the ac magnetic field do not contain currents generating dc magnetization, i.e., the dc magnetic field in these regions is uniform. If the ac magnetic field penetrates throughout the sample volume, the magnetization curve has no hysteresis, and this state persists even when the applied ac magnetic field is turned off. This decoupling of direct and alternating shielding currents leads to other effects detected in experiments, such as the collapse of the transport current due to a collinear ac magnetic field, the suppression of induced circular currents in superconducting loops, etc.[8,11] Direct measurements of the spatial current distribution in a sample using a Hall probe[4] have produced experimental evidence in favor of this model, describing the decoupling of shielding currents. The theory described in the next section and based on the kinetics of the vortex system provides an interpretation of all the experimental facts.

3. THEORY

3.1. Two-velocity hydrodynamic model of vortex lattice

In order to describe the electrodynamics of hard superconductors in crossed magnetic fields, let us consider the behavior of the vortex system in the macroscopic approximation. The state of the system is characterized not by the positions and velocities of isolated flux lines, but by some averaged parameters describing the state of vortex ensembles, i.e., we use the quasi-hydrodynamic approximation. In this model, we introduce two continuous functions, namely the vortex density n and the mean velocity V of a sufficiently large group of vortices. The velocity V is nonzero when the applied magnetic field is varied, and we assume that this variation is slow (quasi-stationary). If the field on the superconductor surface is constant, the vortex lattice remains in its current state for an infinite time. In the state of equilibrium, the magnetic force acting on the vortices is fully balanced by the pinning force. This means that our consideration is limited to the critical state model, and both the magnetic flux creep and viscous flow of vortices are ignored. The electrodynamic equations for such a system are derived by varying the Gibbs energy with respect to small perturbations of the vortex coordinates. The Gibbs energy includes both the magnetic energy of the system and the work against pinning forces.

In what follows, we will assume that the superconductor occupies the half-space $x>0$. Suppose that both the absolute value and direction of the applied magnetic field are variable, but the field is always parallel to the sample surface. In this case, the vortex lattice state can change for several reasons. First, flux lines driven into the sample from the surface and tilted with respect to those in the bulk can change their alignment owing to their magnetic interaction. A simple analysis, however, indicates that this process is energetically unfavorable. Specifically, the gain in the magnetic energy due to vortex realignment along the magnetic field is proportional to its length, whereas the work done by the pinning force is proportional to its length squared. Second, nonparallel vortices can cross one another during their motion. When two vortices cross, there are two options. They may pass through one another without changing their alignments, or two new vortices parallel to one another may be generated as a result of their crossing.[13,14] In this process, the vortices change their directions, which become closer to that of the magnetic field, so they gain some energy. In what follows, for convenience, we will refer to the cutting as the interaction of two vortices in which initial vortices disappear and two new ones are generated. The probability of this process is characterized by the parameter p^*, which depends on their mutual alignment and may also be a function of the vortex density n. The parameter p^* can be calculated, in principle, only using the microscopic theory. Our task is construction of a model of the critical state for the vortex lattice taking into account the process of flux-line cutting.

First of all, let us define the required macroscopic variables. In this situation, there is a spatial distribution of vortices with different orientations. A simple analysis (see Appendix A) demonstrates that the flux-line cutting is impossible if we introduce a single velocity $V(x,t)$ as a continuous function of time and coordinate characterizing motion of vortices at a given point. We should introduce, at least, one additional velocity $U(x,t)$ characterizing the relative motion of vortices. Introducing the second velocity means that we assume the existence of two vortex groups with different velocities at a given point (a two-velocity quasi-hydrodynamic approximation was discussed in the earlier publication[14]):

$$V_A=V+\frac{U}{2},\quad V_B=V-\frac{U}{2}. \tag{4}$$

In fact, we have separated all the vortices into two groups, A and B. Let us assume, for simplicity, that the vortex densities in both groups are equal, i.e., $n_A=n_B=(1/2)n(x,t)$. It is obvious that the vortices of groups A and B, having the mean velocities V_A and V_B, should have different vortex angles $\vartheta_A(x,t)$ and $\vartheta_B(x,t)$ relative to a certain axis z. Specifically, vortices that have not undergone cutting arrive at the point x from its neighborhood of the size $2l$. The length l can be derived from the parameter of the vortex lattice $a=1/\sqrt{n}$ and the dimensionless parameter p, which can be calculated by averaging the microscopic flux-line cutting probability p^* over the angles. As a result, we have

$$l=1/p\sqrt{n}. \tag{5}$$

This length is the vortex free path. In what follows, we will characterize the system by the average angle $\vartheta(x,t)=(1/2)\times(\vartheta_A+\vartheta_B)$ and the difference $\Delta\vartheta$ between the mean vortex angles in the groups A and B. This difference can be estimated through the gradient of the average vortex angle $\vartheta(x,t)$:

$$\Delta\vartheta=\vartheta_A-\vartheta_B=-l\frac{\partial\vartheta}{\partial x}. \tag{6}$$

Generally speaking, there are vortices of different orientations with an angular spread $\Delta\vartheta$ around each point x, and the theoretical description should include a distribution function $f(x,t,\vartheta)$. In this study, however, we have limited our consideration, for simplicity, to two vortex groups with the average angles ϑ_A and ϑ_B.

3.2. Transport equations for the vortex density and angle

Given the average translational V and relative U velocities in the system, both the vortex density n and the average vortex angle ϑ are transported across the sample volume. It is obvious that the velocity V gives rise to changes in both $n(x)$ and $\vartheta(x)$, whereas the relative velocity U generates only an additional flow of the vortex angle. The vortex density $n(x)$ should satisfy the continuity condition

$$\frac{\partial n}{\partial t}=-\frac{\partial(nV)}{\partial x}. \tag{7}$$

The hydrodynamic component of the angle flow $g_\vartheta^{\mathrm{hydr}}$, determined by the velocity V, can be expressed as

$$g_\vartheta^{\mathrm{hydr}} = nV\vartheta. \tag{8}$$

The other component of the angular flow is due to the relative motion of the vortices. In order to obtain the expression for this flow, it is convenient to introduce the plane $x = x_0$ moving at the velocity $V(x_0,t)$ with respect to the laboratory coordinate system. Vortices of type B on the right of the plane and those of type A on the left cross the plane and transport a certain average angle. It is clear that the numbers of the vortices crossing the plane from the left and right are equal because the total flow of vortices across the plane is zero. This means, in fact, that $(n/2)U/2$ vortices of type A cross the plane per unit time, and the equal number of vortices of type B cross the plane in the opposite direction. The factor 1/2 multiplying n is introduced because the concentration of vortices of each type is $n/2$, and the factor 1/2 at U is due to the fact that the average vortex velocity with respect to the plane is $U/2$. The angle transported across the plane when a vortex of type A is replaced by one of type B equals $\Delta\vartheta$ [Eq. (6)]. Thus, the angle flow due to the relative motion of vortices is determined by the expression

$$g_\vartheta^{\mathrm{rel}} = -\frac{1}{2}n \cdot \frac{1}{2}Ul\frac{\partial\vartheta}{\partial x}. \tag{9}$$

By adding Eqs. (8) and (9) for both angle flows and using Eq. (6), which yields the difference between the angles, we obtain the following equation for the angle transport:

$$\frac{\partial(n\vartheta)}{\partial t} = -\frac{\partial}{\partial x}(nV\vartheta) + \frac{\partial}{\partial x}\left(\frac{1}{4}nUl\frac{\partial\vartheta}{\partial x}\right). \tag{10}$$

3.3. Gibbs energy and basic electrodynamic equations

This section describes derivation of electrodynamic equations for a hard superconductor taking into account interactions among vortices, between a vortex and the sample surface, pinning centers, and flux-line cutting. This can be done by varying the Gibbs energy G, which includes the magnetic energy, the work done by pinning forces when the vortex lattice is translated, and the work by pinning forces in unbending a vortex after crossing, with respect to the velocities, $\delta V(x,t)$ and $\delta U(x,t)$. To begin with, note the fundamental difference between the Gibbs energy variation taking into account flux-line cutting and the case when this effect is ignored,[19–21] i.e., when the magnetic field does not change its direction. In the latter case, the relative vortex velocity is zero, so, the variation should be performed only with respect to one variable, namely V. Note that the variation of the Gibbs energy with respect to the velocities V and U is equivalent to the variation of the vortex lattice as a whole with respect to the hydrodynamic displacement, δu^{hydr}, and of sublattices A and B with respect to the relative displacement, δu^{rel}:

$$\delta u^{\mathrm{hydr}} = \tau\delta V, \quad \delta u^{\mathrm{rel}} = \tau\delta U, \tag{11}$$

where τ is an arbitrary time interval smaller than the characteristic time over which fields in the sample change.

It is convenient to derive the electromagnetic component of the Gibbs energy in the nonlocal case, when the magnetic flux density and vortex density are related not by a simple linear formula $B = n\Phi_0$, where Φ_0 is the magnetic flux quantum, but by the integral formula[19,21]

$$\mathbf{B} = \mathbf{B}_m + \int dx' h_1(x,x')\mathbf{e}(x')n(x'). \tag{12}$$

Here the kernel $h_1(x,x')$ and the Meissner component $\mathbf{B}_m$ of the magnetic induction are defined by the equations

$$h_1 = \frac{\Phi_0}{2\lambda}\left(e^{-|x-x'|/\lambda} - e^{-(x+x')/\lambda}\right), \tag{13}$$

$$\mathbf{B}_m = \mathbf{H}_0 e^{-x/\lambda}, \tag{14}$$

where λ is the London penetration depth, $\mathbf{H}_0$ is the applied magnetic field, the vector $\mathbf{e}(x) = (\mathbf{e}_A + \mathbf{e}_B)/2$ defines the average direction of vortices at a given point in space, and the unit vectors $\mathbf{e}_A$ and $\mathbf{e}_B$ define the directions of the vortices of types A and B. By using the technique described in Ref. 20 and taking into account Eq. (12), we derive from the well-known expression for the Gibbs energy of a vortex lattice in the London approximation[21] the following formula:

$$G_{\mathrm{em}} = \frac{1}{8\pi}\int dx\, dx' [\Phi_0 h_1(x,x')\mathbf{e}(x)\mathbf{e}(x')n(x)n(x') - 2\mathbf{H}_0\mathbf{e}(x')n(x')h_1(x,x')]. \tag{15}$$

The variation of the Gibbs energy with respect to vortex displacements can be expressed in the form of two terms, $\delta G_{\mathrm{em}}^{\mathrm{hydr}}$ and $\delta G_{\mathrm{em}}^{\mathrm{rel}}$, each of which is a function of one independent variable, δu^{hydr} and δu^{rel}, respectively. From Eq. (15), one can derive the following expressions for variations of both components (see Appendix B):

$$\delta G_{\mathrm{em}}^{\mathrm{hydr}} = \frac{\Phi_0}{4\pi}\int dx\; n\delta u^{\mathrm{hydr}}\mathbf{e}(x)\frac{\partial\mathbf{B}}{\partial x}, \tag{16}$$

$$\delta G_{\mathrm{em}}^{\mathrm{rel}} = \frac{\Phi_0}{4\pi}\int dx \frac{1}{4}nl\frac{\partial\vartheta}{\partial x}\delta u^{\mathrm{rel}}\mathbf{e}_x\frac{\partial}{\partial x}[\mathbf{e}(x),\mathbf{B}(x) - \mathbf{H}_0], \tag{17}$$

where $\mathbf{e}_x$ is the unit vector aligned with the x-axis.

Now consider the Gibbs energy component due to the work done by pinning forces on the vortex lattice. The work done in a time τ per pair of vortices can be expressed as

$$A = -f_p|V_A|\tau - f_p|V_B|\tau, \tag{18}$$

where f_p is the magnitude of the maximum pinning force, and V_A and V_B are the sublattice velocities [Eq. (4)]. It follows from Eqs. (4), (11), and (18) that the variation of the work

$$\delta A = -f_p\left(\delta u^{\mathrm{hydr}} - \frac{\delta u^{\mathrm{rel}}}{2}\right)\operatorname{sign}\left(V - \frac{U}{2}\right) - f_p\left(\delta u^{\mathrm{hydr}} + \frac{\delta u^{\mathrm{rel}}}{2}\right)\operatorname{sign}\left(V + \frac{U}{2}\right) \tag{19}$$

depends sensitively on which of two motions (translational motion of the lattice as a whole or the relative displacement of the two sublattices) prevails. One can see that two fundamentally different cases may occur. If $|V|>|U|/2$ holds, then the vortices of both A and B types travel in one direction in the laboratory coordinate system, and the energy loss defined by Eq. (19) is determined by the hydrodynamic displacement:

$$\delta A=-2f_p\,\mathrm{sign}(V)\,\delta u^{\mathrm{hydr}}. \tag{20}$$

Otherwise ($|V|<|U|/2$) we derive from Eq. (19) a different formula:

$$\delta A=-f_p\,\mathrm{sign}(U)\,\delta u^{\mathrm{rel}}. \tag{21}$$

The fundamental difference between these two cases is that the work done by pinning forces, and hence the corresponding component of the Gibbs energy, depends on two different independent variables. In the first case, this is the lattice average displacement, and in the second case, the relative displacement of the sublattices.

By combining the two latter equations, we can express the Gibbs energy variation as follows:

$$\delta G_p=\frac{1}{2}\int dx\; n(x)f_p\left[2\,\delta u^{\mathrm{hydr}}\,\mathrm{sign}(V)\,\theta\left(|V|-\frac{|U|}{2}\right)\right.$$
$$\left.+\,\delta u^{\mathrm{rel}}\,\mathrm{sign}(U)\,\theta\left(\frac{|U|}{2}-|V|\right)\right], \tag{22}$$

where $\theta(x)$ is the Heaviside step function.

Finally, let us consider the Gibbs energy component related to the work done by pinning forces in unbending flux lines after their cutting. After the cutting of two flux lines of the A and B sublattices, new flux lines are generated. Some of their sections are aligned with the type A lines, and others with the type B vortices. Then the new lines are unbent so that their sections of the average length L and aligned with vortices of types A and B should turn towards one another through an angle of $\Delta\vartheta/2$. The pinning forces act against such a turn, and the energy loss per unit length of the vortex is determined by the expression

$$\delta A=\frac{1}{L}\left[\frac{1}{2}f_pL^2\frac{|\Delta\vartheta|}{2}\right]=\frac{1}{4}f_pLl\left|\frac{\partial\vartheta}{\partial x}\right|. \tag{23}$$

The length L is easily expressed in terms of the vortex lattice constant a and the angle $|\Delta\vartheta|$ between the vortices. With due account of Eqs. (5) and (6), we obtain

$$L=\frac{a}{|\Delta\vartheta|}=\frac{1}{\sqrt{n}|\partial\vartheta/\partial x|l}=\frac{p}{|\partial\vartheta/\partial x|}. \tag{24}$$

In order to derive the expression for the Gibbs energy component due to unbending of flux lines, one should calculate the number of lines cut during the relative displacement δu^{rel} of the sublattices A and B in unit time. This simple estimate yields

$$\delta N_{\mathrm{cut}}=n\,\mathrm{sign}(U)\,\delta u^{\mathrm{rel}}dx/l. \tag{25}$$

The unknown Gibbs energy variation is derived from Eqs. (23)–(25):

$$\delta G_{\mathrm{cut}}=\frac{1}{4}\int dx\; npf_p\,\mathrm{sign}(U)\,\delta u^{\mathrm{rel}}. \tag{26}$$

Thus, Eqs. (16), (17), (22), and (26) determine the Gibbs energy variations due to the mean and relative displacements of flux lines. By equating the coefficients at δu^{hydr} and δu^{rel} to zero, we obtain the sought electrodynamic equations:

$$\mathbf{e}(x)\frac{\partial\mathbf{B}}{\partial x}=-\frac{4\pi}{c}J_c\,\mathrm{sign}(V)\,\theta\left(|V|-\frac{|U|}{2}\right), \tag{27}$$

$$\mathbf{e}_x\frac{\partial}{\partial x}[\mathbf{e}(x),\mathbf{B}(x)-\mathbf{H}_0]=-\frac{4\pi}{c}J_c\frac{p^2\sqrt{n}}{\partial\vartheta/\partial x}\,\mathrm{sign}(U)$$
$$\times\left[1+\frac{2}{p}\theta\left(\frac{|U|}{2}-|V|\right)\right], \tag{28}$$

where $J_c=cf_p/\Phi_0$.

In the local limit of interest to us, $\lambda\to 0$, the direction of the vector $\mathbf{e}$ coincides with that of the magnetic induction vector, and the equation system takes a simpler form. It can be different, depending on the relation between the velocities of hydrodynamic and relative displacements. For example, for $|V|>|U|/2$ we have

$$\frac{\partial B}{\partial x}=\pm\frac{4\pi}{c}J_c,$$
$$\frac{\partial\vartheta}{\partial x}=\pm\sqrt{\frac{4\pi}{c}J_c\frac{p^2\sqrt{n}}{H_0\cos(\vartheta-\vartheta_0)}}, \tag{29}$$

while for $|V|<|U|/2$ we obtain

$$\frac{\partial B}{\partial x}=0,$$
$$\frac{\partial\vartheta}{\partial x}=\pm\sqrt{\frac{4\pi}{c}J_c\frac{p^2\sqrt{n}}{H_0\cos(\vartheta-\vartheta_0)}\left(1+\frac{2}{p}\right)}, \tag{30}$$

where ϑ_0 is the tilt angle of the magnetic field H_0 with respect to the z-axis. Finally, in the degenerate case, when $|V|=|U|/2$ holds, the number of variables in the problem is reduced. After transformations similar to those performed above, we obtain the following equation:

$$\frac{\partial B}{\partial x}(\pm)\frac{H_0}{2p\sqrt{n}}\frac{\partial\vartheta}{\partial x}\frac{\partial\sin(\vartheta-\vartheta_0)}{\partial x}=\pm\frac{4\pi}{c}J_c[1(\pm)p]. \tag{31}$$

A more detailed analysis demonstrates that the plus or minus signs in all equations can be selected independently, but those in parentheses should be consistent with each other. Note that we prefer to express our results in terms of the parameter p, not the related parameter l. The dimensionless parameter p seems to be more convenient because it is closely related to the real microscopic probability of flux-line cutting.

4. DISCUSSION

The electrodynamic equations derived from the simple two-component model allow us to analyze the behavior of hard superconductors in crossed magnetic fields. Equations (29) and (30) supplemented with the transport equations (7) and (10), and continuity conditions for the velocities V and U and the magnetic induction B form a complete system of equations for the variables $B=n\Phi_0$, ϑ, V, and U. Note that the first lines in Eqs. (29) and (30) do not contain p, i.e., they are independent of the selection of the flux-line cutting model. Depending on the values of the velocities V and U, the system of equations B and ϑ takes one of the possible forms [with different selections of plus and minus signs in Eqs. (29)–(31)]. In this respect, our model is similar to the traditional one-component model of the critical state, where the sign of the current density is also determined by the direction of vortex motion. As in the traditional model of the critical state, the selection of electrodynamic equations (29), (30), or (31) is determined by the magnetic history of the system. In principle, starting with an arbitrary initial state, one can reconstruct the system evolution using the electrodynamic equations.

Note that system of equations in the form of Eq. (29) is fully equivalent to the model often applied to the electrodynamic properties of hard superconductors in crossed magnetic fields.[15,16] Specifically, the second line of Eq. (29) can be easily transformed to $\partial\vartheta/\partial x=\pm k_c$. Our analysis has demonstrated that, generally speaking, the parameter k_c depends not only on B, but also on the angle ϑ.

Let us apply the resulting equations to our experimental results. Direct calculations of the magnetic induction by Eq. (29), which holds for $|U|<2|V|$, have shown that at all amplitudes of the ac magnetic field the dc magnetic field component is nonuniform, so the dc magnetization of the sample is not suppressed.

Therefore, let us try to interpret our experimental results in terms of the second pair of Eqs. (30). According to the first of these equations, the magnitude of the magnetic induction $B(x,t)$ is independent of the coordinate x. Since h_0 in the experiment was much smaller than the typical strength of the dc field H, the uniformity of B is equivalent to the uniformity of the z-component of magnetic induction throughout the volume where the ac field penetrates. The increase in h_0 leads to an increase in its penetration depth and suppression of the dc sample magnetization, which is in agreement with our experimental data. We should, however, test whether this solution is consistent with the complete systems of equations. In other words, we should test whether the condition $|U|>2|V|$ for the applicability of Eq. (30) is satisfied. This condition is surely satisfied, in particular, if $V=0$ holds throughout the region where the ac field penetrates. This can be proved using Eqs. (7) and (10). Given the solution $B=\text{const}$ and the smallness of the ac field in comparison with the dc field, we find from Eq. (7) that the hydrodynamic velocity is also independent of position. The velocity $V(x,t)$ is a continuous function. Since $V(x,t)=0$ holds in the region where the magnetic field is constant, i.e., where the vortex lattice is ‘‘frozen,’’ it is zero throughout the sample. At the same time, it follows from the second line of Eq. (30) and the transport equation (10) for the vortex angle that the relative velocity satisfies $U\neq 0$. Hence the necessary condition for the applicability of Eq. (30) is met.

Thus, Eq. (30) allows us to give a qualitative interpretation of our experimental results. The results obtained by Hasanain *et al.*,[22] where the dc magnetization was suppressed by a single scan, the transverse magnetic field, can probably be also interpreted in these terms.

The analysis of Eq. (30) combined with Faraday's law demonstrates that, in the limit $h_0\ll H$, the direction of the alternating shielding current coincides with the direction of the induced electric field. As a result, Eq. (30) is transformed to Eq. (2). Therefore Eq. (2) can be used in interpreting the observed effect.[12] For this reason, this equation yields correct results in many other cases.[8,9,11] In the general case, one should test in solving a specific problem whether Eq. (2) applies, or a more sophisticated model must be used.[15,16] Moreover, it is plausible than in some regions of a superconductor Eq. (30) applies, whereas in other regions Eq. (29) should be used. In other words, a careful analysis of the complete equation system described in this paper is needed for a theoretical description of the electrodynamic properties of hard superconductors in crossed magnetic fields.

The work is a part of national programs on high-T_c superconductivity (the ‘‘Collapse’’ projects sponsored by the governments of Russia and Ukraine), and was supported by the Russian Fund for Fundamental Research (Project 96-02-17730) and the Mexican Committee on Science and Technology (CONACyT), Grant 3004E306.

APPENDIX A

Let us prove that the flux-line cutting cannot be described in terms of one continuous function, namely the velocity $V(x,t)$ of flux lines. Specificaly, the equation of motion for an arbitrary vortex has the form

$$dx_1/dt=V(x_1,t). \tag{A1}$$

For a neighboring vortex, we have the equation

$$\frac{dx_2}{dt}=V(x_1+(x_2-x_1),t)\approx V(x_1,t)+(x_2-x_1)\left.\frac{\partial V}{\partial x}\right|_{x=x_1}. \tag{A2}$$

The separation between them is described by the equation

$$\frac{d(x_2-x_1)}{dt}=(x_2-x_1)\left.\frac{\partial V}{\partial x}\right|_{x=x_1}. \tag{A3}$$

Its solution has the form

$$(x_2-x_1)=(x_2-x_1)|_{t=0}\exp\left(\int_0^t dt\left.\frac{\partial V}{\partial x}\right|_{x=x_1}\right)\neq 0. \tag{A4}$$

The latter equation indicates that the separation between the vortices only asymptotically approaches zero, and the flux-line cutting is impossible.

APPENDIX B

Let us derive expressions for the variation of the electromagnetic component of the Gibbs energy due to the average and relative displacements of vortices. To this end, we express G_{em}[15] as

$$G_{\text{em}}=\frac{1}{8\pi}\int dxdx'\{h_1(x,x')\Phi_0 n(x)n(x')$$

$$\times\cos[\vartheta(x)-\vartheta(x')]-2H_0n(x')h_1(x,x')$$
$$\times\cos[\vartheta(x')-\vartheta_0]\}. \tag{B1}$$

It follows from the transport equations (7) and (10) that small displacements of the vortices lead to the following variations in the vortex density and vortex angle:

$$\delta n=-\frac{\partial}{\partial x}(n\,\delta u^{\mathrm{hydr}}), \tag{B2}$$

$$\delta\vartheta=-\delta u^{\mathrm{hydr}}\frac{\partial\vartheta}{\partial x}+\frac{1}{n}\frac{\partial}{\partial x}\left(\frac{1}{4}n\,\delta u^{\mathrm{rel}}l\frac{\partial\vartheta}{\partial x}\right). \tag{B3}$$

By varying G_{em} with respect to the flux-line density and its tilt angle with respect to the z-axis and using the symmetry property of the function $h_1(x,x')=h_1(x',x)$, (Eq. 13) we obtain

$$\delta G_{\mathrm{em}}=\frac{\Phi_0}{4\pi}\int dx\,dx'$$
$$\times\left\{\delta n(x)\left[h_1(x,x')n(x')\cos(\vartheta(x)-\vartheta(x'))\right.\right.$$
$$\left.-\frac{H_0}{\Phi_0}h_1(x,x')\cos(\vartheta(x)-\vartheta_0)\right]-\delta\vartheta(x)$$
$$\times\left[h_1(x,x')n(x)n(x')\sin(\vartheta(x)-\vartheta(x'))\right.$$
$$\left.\left.-\frac{H_0}{\Phi_0}n(x')h_1(x,x')\sin(\vartheta(x')-\vartheta_0)\right]\right\}. \tag{B4}$$

By substituting into the latter equation Eqs. (B2) and (B3), integrating the term with $\delta n(x)$ by parts with due account of the symmetry of $h_1(x,x')$, and replacing $\cos[\vartheta(x)-\vartheta(x')]$ and $\cos[\vartheta(x')-\vartheta_0]$ with the respective scalar products, we derive the equation

$$\delta G_{\mathrm{em}}=\frac{\Phi_0}{4\pi}\int dx\,dx'n(x)$$
$$\times\left\{\delta u^{\mathrm{hydr}}\mathbf{e}(x)\frac{\partial}{\partial x}\left[h_1(x,x')n(x')\mathbf{e}(x')\right.\right.$$
$$\left.-\frac{\mathbf{H}_0}{\Phi_0}h_1(x,x')\right]+\frac{1}{4}\delta u^{\mathrm{rel}}l\frac{\partial\vartheta}{\partial x}\frac{\partial}{\partial x}$$
$$\times\left[h_1(x,x')n(x')\sin(\vartheta(x)-\vartheta(x'))\right.$$
$$\left.\left.-\frac{H_0}{\Phi_0}h_1(x,x')\sin(\vartheta(x)-\vartheta_0)\right]\right\}. \tag{B5}$$

After expressing the sines in terms of vector products and using the relation

$$H_0\int dx'h_1(x,x')=\Phi_0[H_0-B_m(x)], \tag{B6}$$

which can be easily verified, we obtain Eqs. (16) and (17).

[1] C. P. Bean, Phys. Rev. Lett. **8**, 250 (1962).
[2] C. P. Bean, Rev. Mod. Phys. **36**, 31 (1964).
[3] A. M. Campbell, J. Phys. C **2**, 1492 (1969); **4**, 3186 (1971).
[4] I. F. Voloshin, N. M. Makarov, V. A. Yampol'skii, and L. M. Fisher, JETP Lett. **53**, 115 (1991).
[5] L. M. Fisher, V. S. Gorbachev, I. F. Voloshin *et al.*, Phys. Rev. B **46**, 10986 (1992).
[6] A. M. Campbell and Evetts, *Critical Current in Superconductors*, Taylor and Francis Ltd., London (1972).
[7] J. R. Clem, J. Appl. Phys. **50**, 3518 (1979).
[8] I. V. Baltaga, L. M. Fisher, N. M. Makarov *et al.*, Phys. Lett. A **148**, 213 (1990).
[9] L. M. Fisher, N. V. Il'in, I. F. Voloshin *et al.*, Physica C **197**, 161 (1992).
[10] I. V. Baltaga, K. V. Il'enko, N. M. Makarov *et al.*, Sol. State Commun. **93**, 697 (1995).
[11] I. V. Baltaga, L. M. Fisher, N. M. Makarov *et al.*, Fiz. Nizkikh Temp. **21**, 411 (1995) [Low Temp. Phys. **21**, 320 (1995)].
[12] L. M. Fisher, A. V. Kalinov, I. F. Voloshin *et al.*, Sol. State Commun. **97**, 833 (1996).
[13] E. H. Brandt, J. R. Clem, and D. G. Walmsley, J. Low Temp. Phys. **37**, 43 (1979).
[14] E. H. Brandt, J. Low Temp. Phys. **39**, 41 (1980).
[15] J. R. Clem, Phys. Rev. B **26**, 2463 (1982).
[16] J. R. Clem and A. Perez-Gonzalez, Phys. Rev. B **30**, 5041 (1984).
[17] R. Boyer, G. Fillion, and M. A. R. LeBlanc, J. Appl. Phys. **51**, 1692 (1980).
[18] R. Cave and M. A. R. LeBlanc, J. Appl. Phys. **53**, 1631 (1982).
[19] I. F. Voloshin, V. S. Gorbachev, S. E. Savel'ev *et al.*, JETP Lett. **59**, 55 (1994).
[20] L. M. Fisher, I. F. Voloshin, V. S. Gorbachev *et al.*, Physica C **245**, 231 (1995).
[21] V. S. Gorbachev and S. E. Savel'ev, Zh. Éksp. Teor. Fiz. **109**, 1247 (1995) [JETP **81**, 694 (1995)].
[22] S. K. Hasanain, Sadia Manzoor, and A. Amirabadizadeh, Supercond. Sci. Technol. **8**, 519 (1995).

Translated provided by the Russian Editorial office.

Magnetic fields at ^{181}Ta nuclei in Laves phases of RFe_2 (R=Nd, Pr, Sm, Gd, Dy, Yb, Lu)

A. A. Sorokin, B. A. Komissarova, G. K. Ryasnyĭ, L. G. Shpin'kova, Z. Z. Aksel'rod

L. V. Skobel'tsin Science and Research Institute of Nuclear Physics, Moscow State University

A. V. Tsvyashchenko, E. N. Shirani, and L. N. Fomicheva

L. F. Vereshchagin Institute of High-Pressure Physics, 142092 Troitsk, Moscow Region, Russia

(Submitted 15 August 1997)
Zh. Éksp. Teor. Fiz. **111**, 1085–1091 (March 1997)

By studying the perturbed angular correlations of γ-rays emitted during the decay of ^{181}Hf impurities in the Laves phases of $PrFe_2$, $DyFe_2$, and $YbFe_2$, we have investigated the magnetic hyperfine interaction between these compounds and its daughter nucleus ^{181}Ta, and have determined the temperature dependence of the magnetic hyperfine fields. At room temperature we obtained the following values of these magnetic hyperfine fields B_{hf}: $B_{\text{hf}}(PrFe_2) = 7.6(1)$ T, $B_{\text{hf}}(DyFe_2) = 15.5(5)$ T, and $B_{\text{hf}}(YbFe_2) = 18.8(3)$ T. When taken together with data obtained previously, the results of our experiments show that for Ta nuclei in the RFe_2 Laves phases the values of B_{hf} depend strongly on whether R is a light or a heavy rare- earth element, which allows us to conclude that in these phases the value of the magnetic moment induced at the impurity Ta nuclei depends on the interatomic distance. © *1997 American Institute of Physics.* [S1063-7761(97)02603-6]

1. INTRODUCTION

The Laves phases form a general class of intermetallic compounds with the chemical formula AB_2, which crystallize either in the $MgCu_2$ cubic structure (C15) or the $MgZn_2$ (C14) or $MgNi_2$ (C36) hexagonal structures. A detailed classification of the Laves phases was given in the monograph Ref. 1. The members of this class of compounds with transition metal and rare-earth elements are especially interesting, since they exhibit great variety in their electric and magnetic properties, depending on their constituent atoms. The recent development of powerful computational methods makes it possible to arrive at a fairly realistic description of the electronic structure of these intermetallics, from which predictions can be made regarding their macroscopic properties. Verification of these predictions calls for experimental techniques that make it possible to determine the parameters of the hyperfine interaction between nuclei of ions incorporated into their composition, since these parameters are very sensitive to the details of the imtermetallic electronic structure. The need for these techniques has led to wide application of investigative methods such as nuclear magnetic resonance, the Mossbauer effect (or nuclear gamma resonance), and the method of perturbed angular correlations of cascade gamma quanta.

All the rare-earth elements except Eu combine with iron to form the intermetallic compounds RFe_2 with the structure of the cubic Laves phases. The magnetism of these intermetallics is determined by the $3d$-bands of Fe and by the well-localized $4f$- moments of the rare-earth ions. In the majority of cases, the magnetic moments of the rare-earth elements are oriented antiparallel to the Fe moments, i.e., the corresponding compounds are not ferromagnets but ferrimagnets. The exceptions are $PrFe_2$ and $NdFe_2$ (see Ref. 2): according to the data of Ref. 3, $PrFe_2$ is a ferromagnet. In $NdFe_2$ the mutual orientation of the Fe and Nd moments is much more complex; in Ref. 2 it was shown that the moments of Nd are oriented antiparallel to one another and are almost perpendicular to the iron moments. In the Laves phases of RFe_2 the $4f$-moments of the rare-earth elements interact with the iron $3d$-bands through the rare-earth valence electrons, which are polarized by their interaction with the $3d$-band electrons. Because this interaction also determines the peculiar magnetic properties of these intermetallics, the experimental study of these compounds is of great interest.

We attempted to study several aspects of this interaction by using data on hyperfine magnetic fields at substitutional impurity Ta ions located at the sites of rare-earth ions. We previously studied compounds based on the light rare-earth elements (R=Nd, Sm, Gd) and Lu with a fully occupied $4f$-shell. The results of our measurements of hyperfine fields at Ta in $NdFe_2$, $SmFe_2$, $GdFe_2$, and $LuFe_2$ were published previously in separate short communications.[4–7] In the present paper we continue our investigation of hyperfine magnetic fields at Ta nuclei in the Laves phases with light rare earths ($PrFe_2$) and extend it to compounds with heavy rare earths ($DyFe_2$ and $YbFe_2$).

2. EXPERIMENTAL METHOD

The compounds $PrFe_2$ and $YbFe_2$ can be synthesized only under special conditions, specifically by sintering or smelting the original materials under high pressures.[2] We prepared our samples (as we did the compound $GdFe_2$ in Ref. 4) by smelting at a pressure up to 80 kbar in a special high-pressure chamber. Construction of the chamber and the method for smelting are described in Ref. 8. The samples of $DyFe_2$ were prepared by smelting in an induction furnace containing an argon atmosphere at ordinary pressure. In the course of smelting these samples, in all cases we added a rather small amount (usually less than 1 percent by weight of the Laves phase of $HfFe_2$, previously irradiated in a reactor

1063-7761/97/030599-04$10.00

and containing the isotope ^{181}Hf ($T_{1/2}=43$ days) with high specific activity. This procedure ensured that our final product was a solid solution of $HfFe_2$ in RFe_2, with impurity atoms localized in the RFe_2 lattice at substitutional sites of the rare-earth ions. The quality of the samples was checked by x-ray structure analysis, and also measurement of the Mossbauer spectra for Fe, and by comparing them with published data.

During the β decay of ^{181}Hf into ^{181}Ta, a series of 133–142 keV γ transitions is excited, in the course of which the nucleus passes through an isomer state of 482 keV with a lifetime $\tau=15.6$ ns. The spin of this state is $I = 5/2$ and the magnetic moment is $\mu=3.25\mu_N$ (the g factor is $g=\mu/I=1.30$, where μ_N is the nuclear magneton). The magnetic hyperfine field acting on a Ta nucleus was determined from the method of perturbed angular $\gamma\gamma$ correlations in the cascade. Measurements were made using an automatic three-detector scintillation coincidence spectrometer,[9] which allowed us to simultaneously record coincidence spectra $N(t,\theta)$ between cascade γ rays as a function of the time for angles $\theta=90°$ and $\theta=180°$ between the detectors. According to the widely used procedure based on the theory of perturbed angular correlations,[10] information about the magnetic hyperfine fields is obtained from a frequency analysis of the anisotropy spectra of the angular correlations $R(t)$ determined from $N(t,\theta)$ by the expression

$$R(t)=2[N(t,180°)-N(t,90°)]/[N(t,180°)+2N(t,90°)]. \tag{1}$$

For static magnetic hyperfine interactions in an unmagnetized ferromagnetic sample (i.e., with chaotic orientation of the domains) the anisotropy of the perturbed angular correlations is expressed in the form[10]

$$R(t)=A\{0.2+0.4[\exp(-\Lambda\omega_L t)\cos\omega_L t+\exp(-2\Lambda\omega_L t)\cos 2\omega_L t]\}. \tag{2}$$

Here $\omega_L=2\pi g\mu_N B_{\mathrm{hf}}/h$ is the Larmor precession frequency and B_{hf} is the magnetic hyperfine field acting on the nucleus. The coefficient A is determined by the known nuclear parameters for a given cascade; in our case, including the corrections for the angular resolution of the detectors, we had $A=-0.24$. The exponent Λ characterizes the scatter in the hyperfine interaction frequencies due to crystal imperfections in the region around the probe nucleus (assuming that the distribution of frequencies around the average value is Lorentzian).

3. MEASUREMENT RESULTS AND DISCUSSION

Figures 1 and 2 show experimental anisotropy spectra of perturbed angular correlations for Ta in $PrFe_2$ and $YbFe_2$ measured at temperatures from $T=80$ K to $T>T_C$. In these figures the modulation at the Larmor spin precession frequency is easy to see. The relatively weak attenuation of the precession is evidence that the probe Ta nuclei are localized at substitutional sites of the rare-earth ions with point cubic symmetry. This attenuation is caused by the lattice defects randomly distributed around the probe nucleus, which create electric field gradients with a characteristic random distribution in frequency and direction. By processing these spectra, we find that the gradients are characterized by a mean-square value of the quadrupole frequency $\langle\omega_Q\rangle=25$ MHz, which is considerably smaller than the Larmor frequency ($\omega_L=450$ MHz for $PrFe_2$ and 1100 MHz for $YbFe_2$ at room temperature).

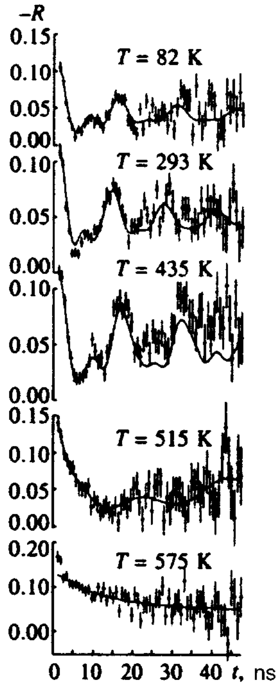


FIG. 1. Spectra of angular correlation anisotropy for ^{181}Ta in $PrFe_2$ measured at different temperatures.

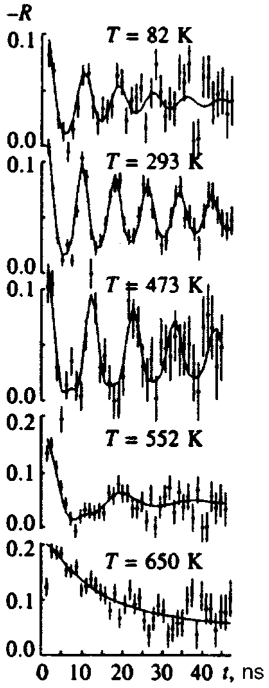


FIG. 2. Spectra of angular correlation anisotropy for ^{181}Ta in $YbFe_2$ measured at different temperatures.

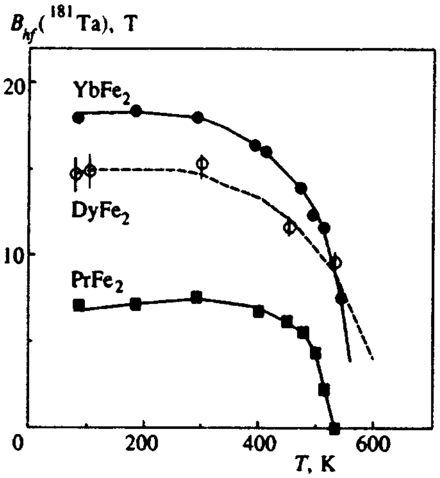


FIG. 3. Temperature dependence of the magnetic hyperfine fields for ^{181}Ta nuclei in $YbFe_2$, $DyFe_2$, and $PrFe_2$.

The values of the hyperfine field obtained from analysis of these spectra are shown in Fig. 3 as functions of temperature for these compounds, and also for $DyFe_2$. The observed falloff in $B_{hf}(Ta)$ at high temperatures is in good agreement with known values of T_C (see Refs. 2, 11)—except for $YbFe_2$, for which a value 543(5) K is obtained instead of the only available value $T_C \simeq 560$ K in the literature. In the temperature range below $\sim 0.5 T_C$, in contrast to the ordinarily observed increase in the hyperfine field, we observed a slight decrease with temperature. We have observed analogous but even more marked behavior in studies of $NdFe_2$, $SmFe_2$, and $GdFe_2$. In Table I we show values of the hyperfine field for Ta extrapolated to $T=0$ K. The signs of the hyperfine field, when given, were determined directly from measurements in an external magnetic field. In the table we also list the fundamental properties of these systems as published in Refs. 2 and 11.

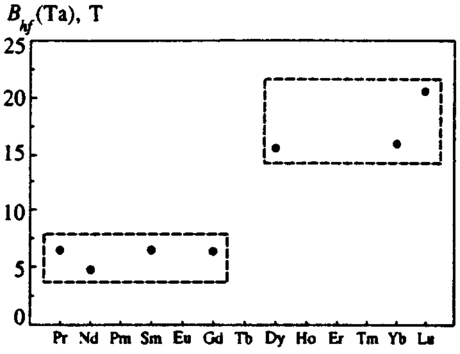


FIG. 4. Magnetic fields at ^{181}Ta nuclei extrapolated to $T=0$ K in the compounds RFe_2.

Figure 4 shows the dependence of the magnitude of $B_{hf}(Ta)$ on the atomic number of the rare-earth elements. The RFe_2 phases clearly separate into two groups, one for light R and one for heavy R. For the light group the values of B_{hf} are small and close to one another (5–6 T), despite the different types of magnetic ordering (ferromagnetic for $PrFe_2$ and $NdFe_2$ and ferrimagnetic for $SmFe_2$ and $GdFe_2$) and the different values of the magnetic moments of the rare-earth elements (e.g., $< 0.5\mu_B$ for Sm and $7\mu_B$ for Gd). In the heavy group, B_{hf} is 10–15 T larger in absolute value, and also is insensitive to increases in the magnetic moment of the rare-earth ions ($9\mu_B$ for Dy, $4\mu_B$ for Yb, while for Lu $\mu(4f)=0$). Note that the lattice constant of RFe_2 decreases monotonically from 7.47 Å to 7.22 Å with increasing atomic number of the rare-earth element (see Table I).

Studies of the magnetic hyperfine interactions of impurity ^{181}Ta ions in the ferromagnetic Laves phase of AFe_2 (where A is a $3d$- or $4d$-transition element)[12–16] show that these ions possess rather small intrinsic magnetic moments (a few tenths of a Bohr magneton) oriented antiparallel to the macroscopic magnetization of the host, i.e., the magnetic moments of the Fe ions. The presence of this moment gives rise to an additional contribution to the magnetic hyperfine field at the nucleus of the Ta ion, essentially due to polarization of the s-electrons of the host.

Calculations of the electronic structure of the $4(5)d$-$3d$ Laves phase published after 1984 (see Refs. 17–19 and the citations in the review article Ref. 19), and also the results of the experimental papers Refs. 20–25, show that both host and impurity $4d$- and $5d$-ions (e.g., Nb in $ZrFe_2$; see Ref. 21) possess intrinsic magnetic moments of order $0.3-0.7\mu_B$ antiparallel to the moments of the Fe ions in the Laves phases. Consequently, the hyperfine magnetic fields at nuclei of ions of the nd elements should be the sum of two contributions: one from the interaction between the ion va-

TABLE I. Properties of the Laves phase of RFe_2 and values of $B_{hf}(Ta)$ reduced to $T = 0$ K.

Phase	a, Å	T_C, K	Magnetic ordering	μ(Fe), μ_B	$\mu(4f)$, μ_B	B_{hf}(Ta), T
$PrFe_2$	7.47	543	ferro	1.6	3.2	6.2
$NdFe_2$	7.46	578	ferro	1.6	3.27	$-$ 5.0
$SmFe_2$	7.42	676	ferri	1.6	<0.7	6.0
$GdFe_2$	7.39	796	ferri	1.62	7.0	$+$ 6.0
$TbFe_2$	7.35	697	ferri	1.7	8.4	-
$DyFe_2$	7.32	630	ferri	1.7	9.4	15.5
$HoFe_2$	7.30	608	ferri	1.7	9.4	-
$ErFe_2$	7.28	587	ferri	1.7	8.4	-
$TmFe_2$	7.23	600	ferri	1.7	6.1	-
$YbFe_2$	7.24	543	ferri	1.64	4.0	17.5
$LuFe_2$	7.22	596	ferri	1.67	0	-20.5

lence electrons and the polarized electrons of the $3d$-band (B_v), and one from the polarization of the core (B_{host}):

$$B_{\text{hf}}=B_v+B_{\text{host}}. \tag{3}$$

These contributions have opposite signs: $B_v<0$, whereas $B_{\text{host}}>0$. It is known[11] that in rare-earth Laves phases the magnetic moment of iron, which determines the value of the negative contribution to B_{hf}(Ta) remains practically unchanged ($1.6-1.7\mu_B$) as R in RFe_2 varies over the entire rare-earth series; therefore, the observed change in B_{hf}(Ta) must be due to a sudden change in the positive contribution B_{host}, which is determined by the intrinsically localized $5d$-moment of the Ta ion. According to the estimates of Ref. 26 the magnetic moment of the $5d$-shell $\mu_B(5d) = 1\mu_B$ due to polarization of the host creates a field at the nucleus of $B_{\text{host}}\simeq 50$ T. Thus, the observed change of 10–15 T in the positive contribution corresponds to a change in the intrinsic moment of the Ta ion of $0.2-0.3\mu_B$, or an equivalent change in the degree of localization of this moment, i.e., its diffuseness.

At this point it is worth discussing an analogy between the behavior of the intrinsic moment of Ta ions in the series RFe_2 and the magnetic moment of Mn ions in the series RMn_2. In Ref. 27, which summarizes data on measurements of magnetic hyperfine fields at Mn nuclei using nuclear magnetic resonance, the authors showed that in the light-R phases from Pr to Tb the Mn ions possess large magnetic moments, from ~ 2.5 to $3.5\mu_B$, but that $\mu_B(\text{Mn})=0$ holds in the heavy-R phases R=Ho, Er, Tm. Yoshimura *et al.*[27] conclude that the magnitude of the magnetic moment of the Mn ions is determined by the interatomic distance: large values of μ_B(Mn) are observed in the range of lattice constants above roughly 7.5 Å; however, as the lattice constant decreases with increasing atomic number of the rare-earth element to values smaller than this value due to lanthanide contraction, the Mn ion moment goes to zero.

Although our situation is different—the ions we are studying are isolated impurities in the R sublattice rather than host ions of the lattice—the strength of the analogy leads us to conclude that the behavior of the intrinsic $5d$-moment of Ta ions in the series of RFe_2 phases is similarly determined by interatomic distances, and that its abrupt decrease correlates with the decrease in lattice constant from 7.38 Å for GdFe_2 to 7.31 Å for DyFe_2. It is possible that this is caused by a decrease in the splitting of the $5d$-band as we successively decrease the $4f$-spins in the Gd–Yb series while simultaneously shifting the Fermi energy due to the lanthanide contraction. The real cause can be identified by calculating the local $5d$-density at the Ta ions as a function of these parameters. The high sensitivity of the hyperfine field at the Ta to details of the electronic structure of the $4f-3d$-intermetallics can be used to test these hypotheses experimentally.

The authors are grateful to S. I. Reĭbiman for measuring and analyzing the nuclear gamma resonance spectra of our samples, and to N. N. Delyagin for discussing the results.

[1] M. Yu. Tesyluk, *Intermetallic Compounds with the Laves-Phase Structure* (Nauka, Moscow, 1981) [in Russian].
[2] V. Meyer, F. Hartmann-Boutron, Y. Gros, and Y. Berthier, J. Physique **42**, 605 (1981).
[3] M. Shimotomai, H. Miyake, and M. Doyama, J. Phys. F **10**, 707 (1980).
[4] A. A. Sorokin, G. K. Ryasny, B. A. Komissarova, L. G. Shpinkova, A. V. Tsvyashchenko, L. N. Fomichova, and E. N. Shirani, Sol. State Commun. **88**, 529 (1993).
[5] G. K. Ryasny, A. A. Sorokin, Z. Z. Akselrod, E. N. Shirani, B. A. Komissarova, L. N. Kryukova, and L. G. Shpinkova, Sol. State Commun. **92**, 741 (1994).
[6] A. A. Sorokin, G. K. Ryasny, B. A. Komissarova, L. G. Shpinkova, E. N. Shirani, V. I. Krylov, A. V. Tsvyashchenko, and L. N. Fomichova, Sol. State Commun. **81**, 66 (1992).
[7] O. I. Kochetov, B. A. Komissarova, V. I. Krylov, A. I. Muminov, Ya. Sazhinski, A. A. Sorokin, and V. M. Tsupko-Sitnikov, Abst. Repts from the *27th Conf. on Nuclear Spectroscopy and Structure of Atomic Nuclei* (Nauka, Moscow, 1987), p. 543 [in Russian].
[8] A. V. Tsvyashchenko, J. Less-Common Metals **99**, L9 (1984).
[9] Z. Z. Akselrod, B. A. Komissarova, L. N. Kryukova, G. K. Ryasnyĭ, S. A. Sergeev, and A. A. Sorokin, Instrum. Experim. Techniques **36**, 28 (1983) [in Russian].
[10] *Alpha-, Beta-, and Gamma-Ray Spectroscopy*, Vol. 3, K. Siegbahn, ed. (North-Holland, Amsterdam (1968).
[11] W. E. Wallace, *Rare Earth Intermetallics* (Acad. Press, New York, 1973).
[12] B. A. Komissarova, L. N. Kryukova, G. K. Ryasny, and A. A. Sorokin, *5th Int. Conf. on Hyperfine Interactions* (West Berlin, 1980) [book of abstracts, C59-1].
[13] Z. Z. Akselrod, M. Budzynski, T. Khazratov, B. A. Komissarova, L. N. Kryukova, S. I. Reiman, G. K. Ryasny, and A. A. Sorokin, Hyperfine Interactions **14**, 7 (1983).
[14] Z. Z. Akselrod, M. Budzynski, B. A. Komissarova, L. N. Kryukova, G. K. Ryasny, and A. A. Sorokin, Phys. State. Solidi A **119**, 667 (1983).
[15] Z. Z. Akselrod, B. A. Komissarova, L. N. Kryukova, G. K. Ryasny, A. A. Opalenko, and A. A. Sorokin, Phys. State. Solidi A **73**, K299 (1983)
[16] A. A. Sorokin, B. A. Komissarova, L. G. Shpinkova, G. K. Ryasny, and E. N. Shirani, Ser. Fiz. **58**, 10 (1994).
[17] P. Mohn and K. Schwarz, Physica B **130**, 26 (1985).
[18] H. Yamada and M. Shimizu, J. Phys. F: Metal Phys. **16**, 1039 (1986).
[19] H. Yamada, Physica B **149**, 390 (1988).
[20] T. Dumelo, P. C. Riedi, P. Mohn, K. Schwarz, and Y. Yamada, J. Mag. Mag. Mater. **54–57**, 1081 (1986).
[21] T. Dumelo and P. C. Riedi, Hyperfine Interactions **34**, 1061 (1987).
[22] C. Ritter, J. Phys. Condens. Matter **1**, 2765 (1989).
[23] P. Warren, J. B. Forsyth, G. J. McIntyre, and N. Bernoeft, J. Phys. Condens. Matter **4**, 5795 (1992).
[24] F. Baudelet, Ch. Giorgetti, S. Pizzini, Ch. Brouder, E. Dartige, A. Fontain, J. P. Kappler, and G. Krill, J. Electron Spectroscopy and Related Phenomena **62**, 153 (1993).
[25] V. S. Pokatilov, Zh. Éksp. Teor. Fiz. **93**, 944 (1987) [Sov. Phys. JETP **66**, 531 (1987)].
[26] A. J. Freeman, J. V. Mallow, and P. S. Baqus, J. Appl. Phys. **41**, 1321 (1970).
[27] K. Yoshimura, M. Shiga, and Y. Nakamura, J. Phys. Soc. Jpn. **55**, 3585 (1986).

Translated by Frank J. Crowne

A hybrid-phonon resonance in a quasi-two-dimensional nanostructure

V. A. Margulis

N. P. Ogaryov Mordovian State University, 430000 Saransk, Russia

(Submitted 29 August 1996)
Zh. Éksp. Teor. Fiz. **111**, 1092–1106 (March 1997)

The coefficient of absorption of electromagnetic radiation by a quasi-two-dimensional electron gas placed in an oblique magnetic field is found. The scattering of electrons by optical phonons is shown to lead to resonant absorption. The shape of the resonance peaks on the absorption curve is studied, and their doublet nature is demonstrated. Finally, the dependence of the resonance peaks on the angle between the magnetic field vector and the confinement plane is investigated. [S1063-7761(97)02703-0]

1. INTRODUCTION

The study of intraband optical transitions of quasi-two-dimensional electrons in an oblique magnetic field yields important information about the energy spectrum of charge carriers and their interaction with the scatterers.[1–3] A review of the early work in this area can be found in the article by Ando *et al.*,[4] while the more recent results can be found in Refs. 5 and 6. Hybridization of electric and magnetic quantization in a quasi-two-dimensional nanostructure leads to a purely discrete energy spectrum of the electrons. In view of this, resonant absorption of electromagnetic radiation caused by electron transitions between two hybrid energy levels becomes possible. To study these transitions we must select an appropriate model for the confinement potential of the quasi-two-dimensional electron gas in the nanostructure. The model employed in Ref. 7–9 in studying the effects caused by an oblique magnetic field **B** uses a parabolic well with a confinement potential $U=m^*\Omega^2z^2/2$, where m^* is the effective mass of the charge carriers, and Ω is the frequency of the potential, related to the characteristic confinement dimension l by $l=\sqrt{\hbar/m^*\Omega}$. The explanation for such a choice lies in two important facts. First, the Hamiltonian of the one-electron states in this model is quadratic, so that by applying a certain canonical transformation of the phase space the spectrum of the system can be reduced to a sum of the spectra of two harmonic oscillators, while the eigenfunctions factorize, i.e., become a product of oscillator functions.[10] The filling factor of these states is determined by the degeneracy of the hybrid-oscillator levels and is equal to $eB_z/c\hbar$, where B_z is the component of the magnetic field $\mathbf{B}=(B_x,0,B_z)$ perpendicular to the confinement plane. Second, in accordance with the generalized Kohn theorem,[11] electron–electron interactions in this case have no effect on the electron transitions.

Since the spectrum and wave functions of one-electron states have a simple analytical form, it is possible (and this will be shown shortly) to derive explicit analytical expressions for the coefficient of absorption of a high-frequency electromagnetic field.

Scattering by phonons in a two-dimensional nanostructure can lead to a process in which the transition between electronic states occurs under the simultaneous action of two factors, i.e., when the absorption of a quantum $\hbar\omega$ of the high-frequency field is accompanied by absorption or emission of an optical phonon. We call a resonance in absorption due to processes of this type a hybrid-phonon resonance.

2. THE ENERGY SPECTRUM AND THE WAVE FUNCTIONS

The energy spectrum of a electron in a parabolic well placed in an oblique magnetic field has the form[7]

$$\varepsilon_\beta=\hbar\omega_1\left(n+\frac{1}{2}\right)+\hbar\omega_2\left(m+\frac{1}{2}\right),\quad n,\ m=0,\ 1,\ \dots, \tag{1}$$

where the hybrid frequencies are[8]

$$\omega_{1,2}^2=\frac{\omega_c^2+\Omega^2}{2}\pm\sqrt{\frac{(\omega_c^2+\Omega^2)^2}{4}-\omega_c^2\Omega^2\cos^2\theta_0},$$

and θ_0 is the angle between field **B** and the confinement plane. The respective wave functions in the mixed coordinate–momentum representation can be described as follows:

$$|\beta\rangle=\frac{1}{\sqrt{L_x}}\exp\left(\frac{ip_xx}{\hbar}\right)\varphi_n\left(\frac{u}{l_1}\right)\varphi_m\left(\frac{v}{l_2}\right), \tag{2}$$

where $\varphi_n(x)$ are the oscillator functions, L_x is the normalization size, $\beta=(p_x,\ n,\ m)$, and the hybrid-magnetic lengths are $l_{1,2}=\sqrt{\hbar/m^*\omega_{1,2}}$. The variables u and v are related to the electron coordinates and momenta as follows:

$$\begin{aligned}u&=(y-y_0)\cos\alpha-\frac{1}{m^*\Omega}(p_z-p_z^0),\\ v&=(y-y_0)\sin\alpha+\frac{1}{m^*\Omega}(p_z-p_z^0),\end{aligned} \tag{3}$$

where $p_z^0=-m^*\omega_cy_0\sin\theta_0$, $y_0=-cp_x/eB\cos\alpha$, and the angle α is defined by the relation[7]

$$\tan 2\alpha=\frac{2\omega_c\Omega\sin\theta_0}{\omega_c^2-\Omega^2}. \tag{4}$$

The hybrid-quantization levels (1) are smeared by thermal motion by an amount of order T and by collisions by an amount of order $\hbar/\tau$, the latter being determined by the electron relaxation time τ. Clearly, a hybrid-phonon resonance can be observed only if all levels are well-resolved and the

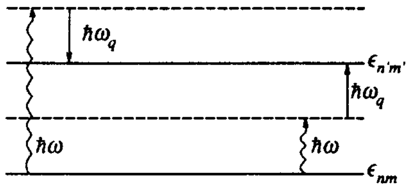


FIG. 1. Transitions leading to resonant absorption in second-order perturbation theory.

photon and phonon frequencies are sufficiently monochromatic. Hence in what follows we assume that the photon frequency is high ($\omega\tau \gg 1$) and the hybrid confinement is sufficiently strong ($\omega_{1,2}\tau \gg 1$) and quantizing ($\hbar\omega_{1,2} \gg T$).

The phonon frequency is quite monochromatic if the interaction involves long-wave optical phonons. The characteristic distance over which the electron wave function varies is of the same order as the hybrid-magnetic lengths $l_{1,2}$. For realistic situations in nanostructures these lengths are large compared to the lattice constant, with the result that the interacting phonons are long-wave.

Now let us examine the physical nature of a hybrid-phonon resonance. In the absence of size confinement along the z axis, motion along the magnetic field is free and, hence, the electron spectrum is quasidiscrete (Landau levels). This fact is known to lead to root singularities in the density of the initial and final states, which in turn lead to logarithmic singularities in the cyclotron-phonon absorption of high-frequency radiation.[12–14] The singularities become smeared if we allow for the collisional width of the level, $\sim\hbar/\tau$, or for optical phonon dispersion.

For a quasi-two-dimensional nanostructure the physical nature of the singularity is different. In our case the probability of an electron transition occurring is proportional to $\delta(\epsilon_\beta - \epsilon_{\beta'} + \hbar\omega \pm \hbar\omega_q)$, where $\epsilon_{\beta,\beta'}$ are electron energies, and $\hbar\omega_q$ is the optical phonon energy. Because the spectrum of one-electron states is discrete, if we ignore optical phonon dispersion, the absorption coefficient $\Gamma(\omega)$ containing this factor has, delta-function singularities at points where the photon frequency satisfies the condition

$$\hbar\omega_1(n-n') + \hbar\omega_2(m-m') + \hbar\omega \pm \hbar\omega_q = 0. \tag{5}$$

If phonon dispersion is taken into account, the singularity becomes smeared (as shown below), and at the singularity the absorption vanishes.

The possible types of electron transition are depicted in Fig. 1.

3. A GENERAL EXPRESSION FOR THE ABSORPTION COEFFICIENT

Using an approach based on the method suggested by Fröhlich[15] and examined in the case of transitions in a magnetic field by Bass and Levinson,[12] we can find the absorption coefficient by applying ordinary perturbation-theory techniques for the interactions of electrons with the high-frequency field, H_R, and the lattice, H_L, which are switched on simultaneously. The transitions depicted in Fig. 1 emerge in second-order perturbation theory in $H_R + H_L$.

The energy of noninteracting electrons, photons, and phonons is described by the following Hamiltonian:

$$H_0 = \sum_\beta \epsilon_\beta a_\beta^+ a_\beta + \sum_{\mathbf{f}} \hbar\omega_f b_{\mathbf{f}}^+ b_{\mathbf{f}} + \sum_{\mathbf{q}} \hbar\omega_q c_{\mathbf{q}}^+ c_{\mathbf{q}}, \tag{6}$$

where a_β^+ (a_β), $b_{\mathbf{f}}^+$ ($b_{\mathbf{f}}$), and $c_{\mathbf{q}}^+$ ($c_{\mathbf{q}}$) are the creation (annihilation) operators for electrons, photons, and phonons, respectively. Following the ideas of Ref. 12, we can write the matrix elements of the transitions depicted in Fig. 1 as follows:

$$\langle\alpha|H|\alpha'\rangle = \sum_{\beta''} \frac{\langle\beta,0|H_R|\beta'',-\mathbf{f}\rangle\langle\beta'',0|H_L|\beta',\pm\mathbf{q}\rangle}{\epsilon_\beta - \epsilon_{\beta''} + \hbar\omega} + \sum_{\beta''} \frac{\langle\beta,0|H_L|\beta'',\pm\mathbf{q}\rangle\langle\beta'',0|H_R|\beta',-\mathbf{f}\rangle}{\epsilon_{\beta'} - \epsilon_{\beta''} - \hbar\omega}. \tag{7}$$

In Eq. (7) the initial state $|\alpha\rangle$ is simply $|\beta,\dots,N_{\mathbf{f}},\dots,N_{\mathbf{q}},\dots\rangle \equiv |\beta,0,0\rangle$, and the final state $|\alpha'\rangle$ is $|\beta',\dots,N_{\mathbf{f}}-1,\dots,N_{\mathbf{q}}\pm 1,\dots\rangle \equiv |\beta',-\mathbf{f},\pm\mathbf{q}\rangle$. It is now convenient to write the absorption coefficient in a form similar to that used in Ref. 12:

$$\Gamma(\omega) = \overline{\Gamma^+}(\omega) + \overline{\Gamma^-}(\omega),$$

where $\overline{\Gamma}$ stands for thermal averaging over the initial phonon states, and

$$\Gamma^{(\pm)}(\omega) = \frac{2\pi\sqrt{\varepsilon(\omega)}}{c\hbar N_{\mathbf{f}}}(1 - e^{-\hbar\omega/T}) \sum_{\beta,\beta',\mathbf{q}} f_0(\epsilon_\beta) |\langle -\mathbf{f}, \pm\mathbf{q},\beta'|V|0,0,\beta\rangle|^2 \delta(\epsilon_\beta - \epsilon_{\beta'} \mp \hbar\omega_q + \hbar\omega), \tag{8}$$

where $\varepsilon(\omega)$ is the real part of the dielectric constant (for which in the frequency range considered here there is no dispersion), c is the speed of light in vacuum, $N_{\mathbf{f}}$ is the number of initial-state photons with a frequency $\omega = cf/\sqrt{\varepsilon(\omega)}$, and

$$f_0(\varepsilon_\beta) = \frac{8\pi\hbar n_0 L_z}{m^*\omega_c \cos\theta_0} \sinh\frac{\hbar\omega_1}{2T} \sinh\frac{\hbar\omega_2}{2T} \exp\left(-\frac{\epsilon_\beta}{T}\right),$$

is the electron distribution function for a nondegenerate gas, where n_0 is the electron concentration and L_z is the normalization length. In Eq. (8) we introduced the matrix element

$$\langle V\rangle = \sum_{\beta''} \frac{\langle\beta',0|H_R|\beta'',-\mathbf{f}\rangle\langle\beta'',0|h_L|\beta,\pm\mathbf{q}\rangle}{\epsilon_\beta - \epsilon_{\beta''} + \hbar\omega} + \sum_{\beta''} \frac{\langle\beta,0|h_L|\beta'',\pm\mathbf{q}\rangle\langle\beta'',0|H_R|\beta',-\mathbf{f}\rangle}{\epsilon_{\beta'} - \epsilon_{\beta''} - \hbar\omega}. \tag{9}$$

The interaction with the lattice is represented by the Hamiltonian

$$H_L = \sum_{\mathbf{q}} h_L(\mathbf{q}),$$

where

$$h_L(\mathbf{q}) = D_{\mathbf{q}} c_{\mathbf{q}} e^{i\mathbf{qr}} + \text{c.c.},$$

and $D_{\mathbf{q}}$ is the electron–phonon coupling constant.

We write the Hamiltonian H_R representing the interaction with the high-frequency field in the form[12]

$$H_R=\frac{e}{m^*}\sqrt{\frac{2\pi\hbar}{\varepsilon(\omega)\omega V}}b_{\mathbf{f}}\,\mathbf{e}_{\mathbf{f}}\,\mathbf{P}. \tag{10}$$

Here V is the normalization volume, and $\mathbf{P}$ is the generalized momentum in a constant magnetic field:

$$\mathbf{P}=\mathbf{p}+\frac{e}{c}\mathbf{A},\quad \mathbf{A}=\left(\frac{eB_z}{c}y,\,0,-\frac{eB_x}{c}y\right).$$

Below we calculate the absorption for linear polarization and choose the polarization vector $\mathbf{e}_{\mathbf{f}}$ in the y direction. Then

$$H_R=\frac{e}{m^*}\sqrt{\frac{2\pi\hbar}{\varepsilon(\omega)\omega V}}b_{\mathbf{f}}\,p_y. \tag{11}$$

In the calculations of the matrix elements of H_R that follow, the high-frequency field is assumed uniform. For this the photon wavelength λ must be much larger than $l_{1,2}$, which imposes certain restrictions on the hybrid-oscillation frequencies $\omega_{1,2}$. If this condition is met, the electron–photon transitions are dipole transitions, with the result that

$$\langle\beta,0|H_R|\beta',-\mathbf{f}\rangle=\frac{e}{m^*}\sqrt{\frac{2\pi\hbar N_{\mathbf{f}}}{\varepsilon(\omega)\omega V}}\langle\beta|p_y|\beta'\rangle. \tag{12}$$

4. FORMULAS FOR THE MATRIX ELEMENTS

Since the one-electron states $|\beta\rangle$ are given in the mixed coordinate–momentum representation, to calculate the matrix elements (12) we must write the operator p_y in the same representation. Using (3), we get

$$p_y=\cos\alpha\frac{\partial}{\partial u}+\sin\alpha\frac{\partial}{\partial v}. \tag{13}$$

Plugging (13) into (12) and performing simple transformations, we get

$$\begin{aligned}\langle\beta,0|H_R|\beta',-\mathbf{f}\rangle={}&\frac{ie\hbar}{m^*l_1l_2}\sqrt{\frac{\pi\hbar N_{\mathbf{f}}}{\varepsilon(\omega)\omega V}}\delta(k_x',k_x)\\&\times\{l_2\delta(m',m)\cos\alpha[\sqrt{n}\,\delta(n',n-1)\\&-\sqrt{n+1}\,\delta(n',n+1)]\\&+l_1\delta(n',n)\sin\alpha[\sqrt{m}\,\delta(m',m-1)\\&-\sqrt{m+1}\,\delta(m',m+1)]\}.\end{aligned} \tag{14}$$

After integration with respect to the variable z the matrix elements of the electron–phonon interaction become

$$\begin{aligned}\langle\pm\mathbf{q},\beta'|h_L|0,\beta\rangle={}&D_{\mathbf{q}}\sqrt{N_{\mathbf{q}}+\frac{1}{2}\pm\frac{1}{2}}\delta(k_x',k_x\pm q_x)\\&\times\langle n'm'|\exp[i(q_yy+q_zz)]|n,m\rangle.\end{aligned} \tag{15}$$

To find the matrix element on the right-hand side of Eq. (15) we write the exponential function in the mixed coordinate–momentum representation:

$$\begin{aligned}\exp[i(q_yy+q_zz)]={}&\exp[iq_y(u\cos\alpha+v\sin\alpha+y_0)]\\&\times\exp\left[q_zl^2\left(\sin\alpha\frac{\partial}{\partial u}-\cos\alpha\frac{\partial}{\partial v}\right)\right].\end{aligned} \tag{16}$$

We introduce the generalized momenta p_u and p_c that are the canonical conjugates of u and v. Then

$$p_u=\frac{\hbar}{i}\frac{\partial}{\partial u},\quad p_v=\frac{\hbar}{i}\frac{\partial}{\partial v},$$

and the second exponential factor on the right-hand side of Eq. (16) becomes $\exp[iq_zl^2(p_u\sin\alpha-p_v\cos\alpha)]$. Next we use the relationship

$$\exp\left(\frac{ia}{\hbar}p_u\right)\varphi_n\left(\frac{u}{l_1}\right)=\varphi_n\left(\frac{u+a}{l_1}\right), \tag{17}$$

which follows from the fact that the exponential operator is the generator of the translation group.

Combining (16) and (17), we can easily obtain an expression for the matrix elements:

$$\begin{aligned}\langle n',m'|&\exp[i(q_yy+q_zz)]|n,m\rangle=\exp(iq_yy_0)\\&\times\left\langle\varphi_{n'}\left(\frac{u}{l_1}\right)\middle|\exp(iq_yu\cos\alpha)\middle|\varphi_n\left(\frac{u+l^2q_z\sin\alpha}{l_1}\right)\right\rangle\\&\times\left\langle\varphi_{m'}\left(\frac{u}{l_2}\right)\middle|\exp(iq_yv\sin\alpha)\middle|\varphi_m\left(\frac{v-l^2q_z\cos\alpha}{l_2}\right)\right\rangle.\end{aligned} \tag{18}$$

Instead of the projections of the phonon wave vector we introduce the following dimensionless quantities:

$$\begin{aligned}g_y&=\frac{l_1q_y}{\sqrt{2}}\cos\alpha,\quad g_z=\frac{l^2q_z}{\sqrt{2}l_1}\sin\alpha,\\d_y&=\frac{l_2q_y}{\sqrt{2}}\cos\alpha,\quad d_z=\frac{l^2q_z}{\sqrt{2}l_2}\sin\alpha.\end{aligned} \tag{19}$$

After simple but lengthy transformations, evaluation of the integrals with oscillator functions in (18) yields

$$\begin{aligned}\langle n',m'|&\exp[i(q_yy+q_zz)]|n,m\rangle\\&=\sqrt{\frac{m'!n'!}{m!n!}}(-1)^{n+m-n'-m'}g^{n-n'}d^{m-m'}\\&\times\exp\left(iq_yy_0-\frac{d^2+g^2}{2}\right)\exp(i[\varphi(n-n')+\psi(m-m')])\\&\times L_{n'}^{n-n'}(g^2)L_{m'}^{m-m'}(d^2),\end{aligned} \tag{20}$$

where $\tan\varphi=g_y/g_z$, $\tan\psi=d_y/d_z$, $d^2=d_y^2+d_z^2$, $g^2=g_y^2+g_z^2$, and $L_N^{N'}$ are generalized Laguerre polynomials.

We introduce the notation

$$\begin{aligned}\langle n',m'|&\exp[i(q_yy+q_zz)]|n,m\rangle\\&=\exp(iq_yy_0)J(n',m',n,m).\end{aligned} \tag{21}$$

Then, using (9), (4), (15), and (21), we obtain

$$\begin{aligned}
&\langle -\mathbf{f},\pm\mathbf{q},\beta'|V|0,0,\beta\rangle\\
&=\frac{ie\hbar}{m^*l_1l_2}\sqrt{\frac{\pi\hbar N_{\mathbf{f}}(N_{\mathbf{q}}+1/2\pm 1/2)}{\omega\varepsilon(\omega)V}}D_{\mathbf{q}}\delta(k_x,k'\pm q_x)\\
&\times\left\{\frac{l_2\sqrt{n+1}\,\cos\alpha}{\hbar(\omega-\omega_1)}J(n+1,m;n',m')\right.\\
&-\frac{l_2\sqrt{n}\,\cos\alpha}{\hbar(\omega+\omega_1)}J(n-1,m;n',m')+\frac{l_1\sqrt{m+1}\,\sin\alpha}{\hbar(\omega-\omega_2)}\\
&\times J(n,m+1;n',m')-\frac{l_1\sqrt{m}\,\sin\alpha}{\hbar(\omega+\omega_2)}J(n,m-1;n',m')\\
&-\frac{l_2\sqrt{n'}\,\cos\alpha}{\hbar(\omega-\omega_1)}J(n,m;n'-1,m')+\frac{l_2\sqrt{n'+1}\,\cos\alpha}{\hbar(\omega+\omega_1)}\\
&\times J(n,m;n'+1,m')-\frac{l_1\sqrt{m'}\,\sin\alpha}{\hbar(\omega-\omega_2)}J(n,m;n',m'-1)\\
&\left.+\frac{l_1\sqrt{m'+1}\,\cos\alpha}{\hbar(\omega+\omega_2)}J(n,m;n',m'+1)\right\}\exp(iq_yy_0).
\end{aligned}\tag{22}$$

Using (20) and (21) we can transform (22) into

$$\begin{aligned}
&\langle -\mathbf{f},\pm\mathbf{q},\beta'|V|0,0,\beta\rangle\\
&=-\frac{ie\hbar}{m^*l_1l_2}\sqrt{\frac{\pi\hbar N_{\mathbf{f}}(N_{\mathbf{q}}+1/2\pm 1/2)}{\omega\varepsilon(\omega)V}}D_{\mathbf{q}}A(\omega)\\
&\times\exp(iq_yy_0)\,\delta(k_x,k'_x\pm q_x)J_{nn'}(g^2)J_{mm'}(d^2),
\end{aligned}\tag{23}$$

where

$$\begin{aligned}
&J_{nn'}(x^2)\\
&=\frac{n'!}{n!}x^{n-n'}\exp\left(-\frac{x^2}{2}\right)\exp\left(-\frac{1}{2}iq_yq_zl^2\,\sin\alpha\,\cos\alpha\right)\\
&\times(-1)^{n-n'}\exp[i\varphi(n'-n)]L_n^{n-n'}(x^2),
\end{aligned}\tag{24}$$

and $J_{mm'}(x)$ can be found from (24) by replacing φ with ψ. The frequency factor

$$\begin{aligned}
A(\omega)=&gl_2\,\cos\alpha\left(\frac{e^{-i\varphi}}{\omega_1-\omega}+\frac{e^{i\varphi}}{\omega_1+\omega}\right)\\
&+dl_1\,\sin\alpha\left(\frac{e^{i\psi}}{\omega_2-\omega}-\frac{e^{-i\psi}}{\omega_2+\omega}\right).
\end{aligned}\tag{25}$$

This can also be used to write

$$|A(\omega)|^2=2[a(\omega)q_z^2+b(\omega)q_y^2],\tag{26}$$

where

$$\begin{aligned}
a(\omega)=&\frac{l^4}{l_1^2l_2^2}\,\frac{\omega_2^2-\Omega^2}{\omega_2^2-\omega_1^2}\left(\frac{\omega_1l_2^2}{\omega_1^2-\omega^2}\sqrt{\frac{\omega_1^2-\Omega^2}{\omega_1^2-\omega_2^2}}\right.\\
&\left.+\frac{\omega_2l_1^2}{\omega_2^2-\omega^2}\sqrt{\frac{\omega_2^2-\Omega^2}{\omega_2^2-\omega_1^2}}\right)^2,
\end{aligned}\tag{27}$$

$$\begin{aligned}
b(\omega)=&\omega^2l_1^2l_2^2\,\frac{\omega_1^2-\Omega^2}{\omega_1^2-\omega_2^2}\left(\frac{1}{\omega_1^2-\omega^2}\sqrt{\frac{\omega_1^2-\Omega^2}{\omega_1^2-\omega_2^2}}\right.\\
&\left.+\frac{1}{\omega_2^2-\omega^2}\sqrt{\frac{\omega_2^2-\Omega^2}{\omega_2^2-\omega_1^2}}\right)^2.
\end{aligned}$$

In deriving (27) we employed the following relationships:

$$\cos^2\alpha=\frac{\omega_1^2-\Omega^2}{\omega_1^2-\omega_2^2},\quad \sin^2\alpha=\frac{\omega_2^2-\Omega^2}{\omega_2^2-\omega_1^2}.$$

5. THE CHARACTERISTICS OF RESONANT ABSORPTION PEAKS

We introduce partial absorption coefficients by the following formula:

$$\Gamma^{(\pm)}(\omega)=\sum_{n,m,n',m'=0}^{\infty}\overline{\Gamma^{(\pm)}}(n,m;n',m').\tag{28}$$

We also introduce the quantity

$$\Delta\omega=\omega_1(n-n')+\omega_2(m-m')+\omega\mp\omega_q.\tag{29}$$

Using the results of Sec. 4 and summing over k_x, we get

$$\begin{aligned}
\overline{\Gamma}^{(\pm)}(n,m;n'm')=&\frac{e^2\omega_c\cos\theta_0}{m^*c\hbar^2l_1^2l_2^2\omega L_z\sqrt{\varepsilon(\omega)}}\\
&\times(1-e^{-\hbar\omega/T})\sum_{\mathbf{q}}|D_{\mathbf{q}}|^2\left(\frac{n'!m'!}{n!m!}\right)\\
&\times|A(\omega)|^2d^{2(m-m')}g^{2(n-n')}e^{-d^2-g^2}[L_{n'}^{n-n'}(g^2)\\
&\times L_{m'}^{m-m'}(d^2)]^2(N_T+1/2\pm 1/2)\,\delta(\Delta\omega)f_0(\varepsilon_\beta),
\end{aligned}\tag{30}$$

where N_T is the Planck distribution function, which results from thermal averaging over phonon states.

Equation (30) clearly shows that if one ignores optical phonon dispersion ($\omega_q=\omega_0$), the partial coefficients $\Gamma^{(\pm)}(n,m;n',m')$ have delta-function singularities at points where the detuning from resonance, $\Delta\omega$, is zero. Let us now allow for weak phonon dispersion. For long-wave phonons we assume a parabolic dispersion law: $\omega_q=\omega_0(1-\omega_0^{-2}v_s^2q^2)$, where ω_0 is the optical−phonon threshold frequency, and v_s is the speed of sound.[13] In this case the integral with respect to $|\mathbf{q}|$ in (30) can easily be evaluated thanks to the presence of a delta function of the form $\delta(\Delta\omega\pm\omega_0^{-1}v_s^2q^2)$. The electron–phonon coupling constant for PO- and DO-phonons is[12]

$$|D_{\mathbf{q}}|^2=\frac{2\pi\hbar^2\alpha_L\omega_0}{m^*V}\begin{cases}\dfrac{\sqrt{2m^*\hbar\omega_0}}{q^2} & \text{for } PO\text{-phonons,}\\[2ex] \dfrac{4\hbar^2}{\sqrt{2m^*\hbar\omega_0}} & \text{for } DO\text{-phonons.}\end{cases}\tag{31}$$

Then for $\Gamma_{DO}^{(\pm)}(n,m;n',m')$ we obtain

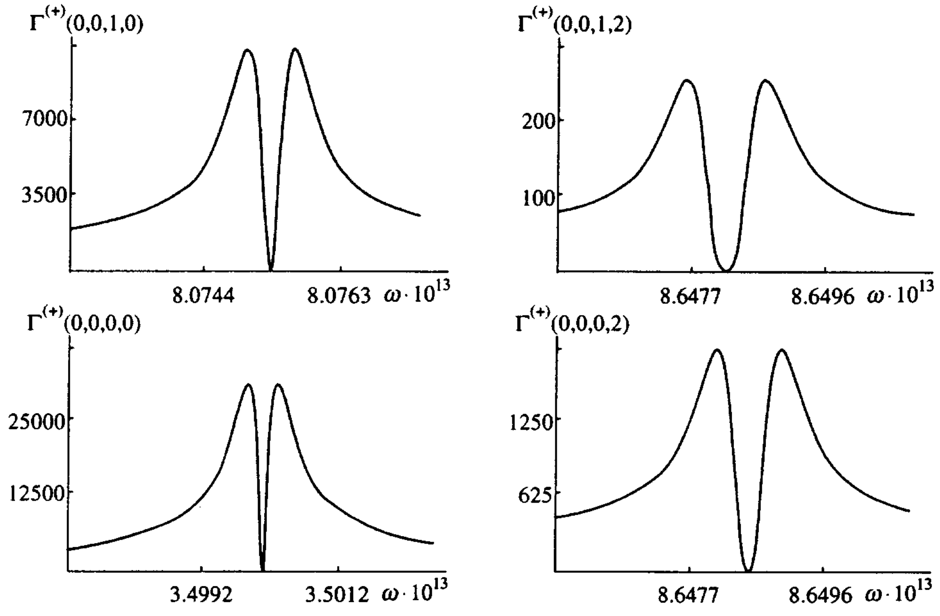


FIG. 2. Partial absorption coefficients in the case of emission of DO-phonons; $\theta_0=30°$.

$$\Gamma^{(\pm)}_{DO}(n,m;n',m') = \frac{2e^2 n_0(N_0+1/2\pm 1/2)(1-e^{-\hbar\omega/T})\hbar\omega_0^3(\Delta\omega)^{3/2}\omega_c\Omega}{\pi c\omega m^* v_s^5\sqrt{2m^*\hbar\varepsilon(\omega)}} \times\sinh\frac{\hbar\omega_1}{2T}\sinh\frac{\hbar\omega_2}{2T} \times\exp\left[-\frac{\omega_1(n+1/2)+\hbar\omega_2(m+1/2)}{T}\right] \times\Phi(n,m;n',m'). \qquad (32)$$

Here α_L is the dimensionless electron–lattice coupling constant, $N_0=(e^{\hbar\omega_0/T}-1)^{-1}$, and the integral over the angles $\Phi(n,m;n',m')$ is given by

$$\Phi(n,m;n',m') = \frac{n'!m'!}{n!m!}\int_0^{2\pi}d\varphi\int_0^{\pi}\sin\vartheta d\vartheta(a\cos^2\vartheta + b\sin^2\vartheta\sin^2\varphi)x_1^{n-n'}x_2^{m-m'}e^{-x_1-x_2} \times[L_{n'}^{n-n'}(x_1)L_{m'}^{m-m'}(x_2)]^2, \qquad (33)$$

where

$$x_{1,2}=\frac{1}{2}\frac{\omega_0|\Delta\omega|}{l_{1,2}^2 v_s^2}\left(\frac{\omega_1^2-\Omega^2}{\omega_1^2-\omega_2^2}l_{1,2}^4\sin^2\vartheta\sin^2\varphi + \frac{\omega_2^2-\Omega^2}{\omega_2^2-\omega_1^2}l^4\cos^2\vartheta\right), \qquad (34)$$

For the partial absorption coefficient $\Gamma^{(\pm)}_{PO}(n,m;n',m')$ we have

$$\Gamma^{(\pm)}_{PO}(n,m;n',m')=\frac{(mv_s)^2}{2\hbar|\Delta\omega|}\Gamma_{DO}(n,m;n',m').$$

The maxima in $\Gamma(n,m;n',m')$ are obviously caused by the maxima in $\Phi(n,m;n',m')$. Let us examine transitions that originate from the ground state $(n=m=0)$. As Eq. (33) implies, in this case the integrand in the expression for $\Phi(0,0;n',m')$ has the form $(\Delta\omega)^{n'+m'}\exp[-f(\vartheta,\varphi)\Delta\omega]$. For small values of $\Delta\omega$, this function increases according to a power law, and then falls off exponentially.

If we allow for the smearing of the hybrid-oscillation levels caused by collisions, then $\Gamma(\Delta\omega)$ must be replaced by $\mathrm{Re}\,\Gamma(\Delta\omega+i\gamma)$, with the collisional spread $\gamma=\tau^{-1}$. In Eq. (8) we must replace the delta function by the Lorentzian $\pi\tau/[1+\tau^2(\Delta\omega)^2]$. In this case $\Gamma(\Delta\omega)$ has delta-function-like spikes with a halfwidth equal to τ^{-1}.

6. RESULTS OF NUMERICAL ANALYSIS

Let us study the absorption coefficient more closely. We confine ourselves to the ultraquantum case, where transitions originate from the ground state of the system. The above analytical expressions for the partial absorption coefficients $\Gamma^{(\pm)}(0,0;n,m)$ imply that at the point where the frequency of the electromagnetic radiation satisfies $\Delta\omega=0$ these coefficients vanish. Within a small neighborhood of this point, $\Phi(0,0;n,m)$ as a function of $\Delta\omega$ has two symmetrically positioned sharp peaks (to the left and right of the point), and hence $\Gamma^{(\pm)}(0,0;n,m)$ has the same peaks near this point, i.e., the partial absorption peaks have a doublet structure. We also note that the exponential nature of the dependence of the partial absorption coefficients on the numbers n and m of the hybrid-oscillator levels guarantees a sharp decrease in partial absorption as the level numbers grow.

In view of the complexity of the analytical formulas (32)–(34), a detailed description of the absorption peaks requires numerical studies of the dependence of $\Gamma^{(\pm)}(0,0;n,m)$ on the radiation frequency, the magnetic field strength, and the angle between the magnetic field and the confinement plane.

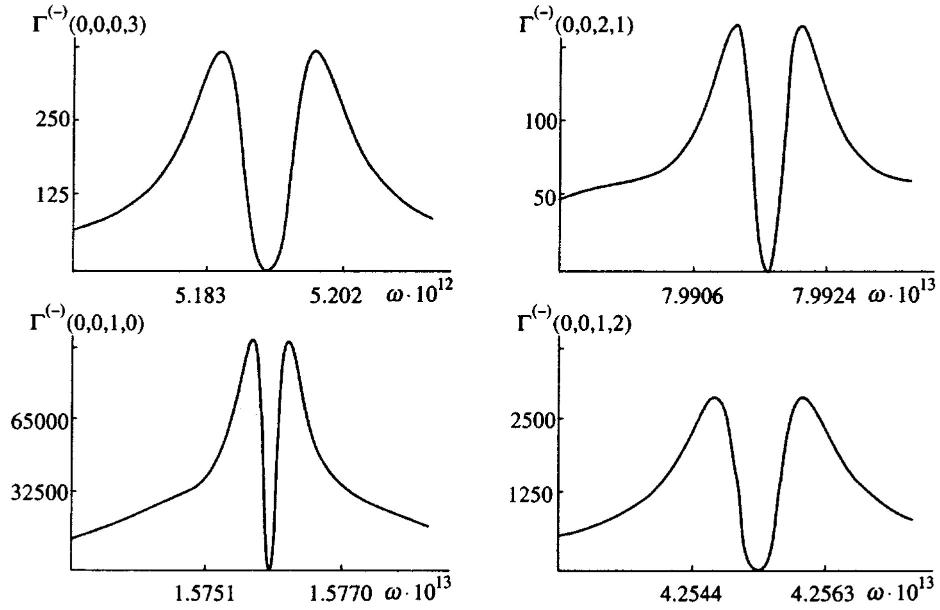


FIG. 3. Partial absorption coefficients in the case of absorption of DO-phonons; $\theta_0=60°$.

Figures 2–5 depict several partial absorption coefficients as functions of the radiation frequency ω for DO- and PO-phonons. The curves were found by numerically integrating Eqs. (32)–(34). In constructing the graphs we used the numerical values of the parameters for InSb:[12] $\varepsilon_0=17.5$, $\varepsilon_\infty=16$, and $\omega_0=3.7\times10^{13}\text{s}^{-1}$ at an electron density $n=10^{16}$ cm^{-3}, temperature $T=100$ K, and $\omega_c=3.4\times10^{13}$ s^{-1}. The graphs in Figs. 2–5 correspond to different values of the angle between the magnetic field and the confinement plane, and the doublet structure in the absorption peak is clearly visible. The peaks comprising the doublet in all figures have a similar asymmetric shape: in relation to the point at which $\Delta\omega=0$ the left peak gradually rises but then suddenly drops, while for the right peak the situation is the opposite.

Note that the distance between the doublet components in frequency ($\sim10^9\text{s}^{-1}$) is small compared to the characteristic frequencies ω_1, ω_2, Ω, ω_c, and ω_0 in the system.

The dependence of the partial absorption coefficients on the magnetic field is roughly the same as depicted in Figs. 2–5. For the values of B at which $\Delta\omega=0$ the absorption is nil. The peaks on the $\Gamma^{(\pm)}$ vs B curves exhibit the same doublet structure as those in $\Gamma^{(\pm)}(\omega)$. Besides being dependent on the magnitude B of the magnetic field, the absorption coefficients $\Gamma^{(\pm)}$ depend on the angle θ_0 between the plane with field **B** and the confinement plane. Figures 2–5 clearly show that the amplitude and absorption of the peaks strongly depend on θ_0, but the general shape of the curves and, in

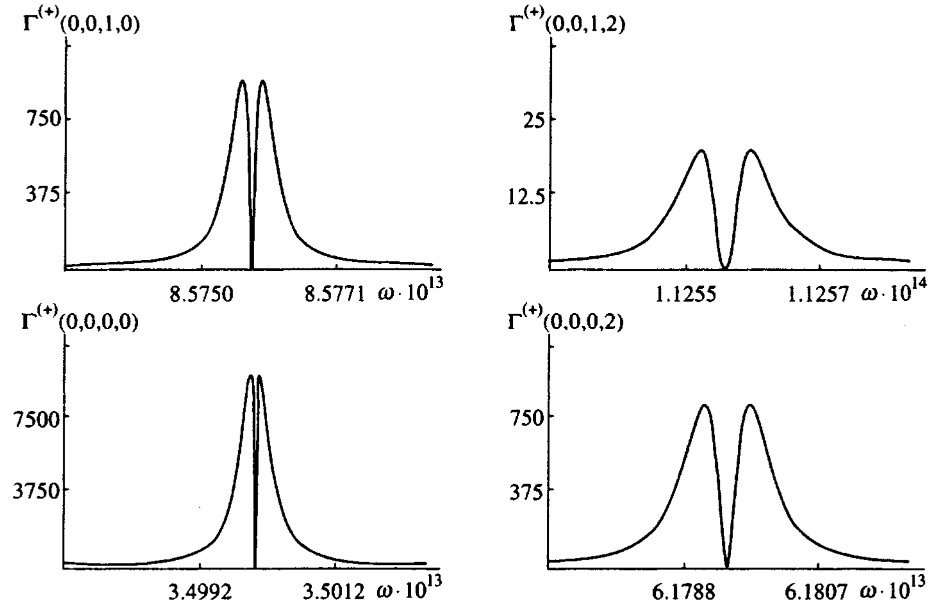


FIG. 4. Partial absorption coefficients in the case of emission of PO-phonons; $\theta_0=60°$.

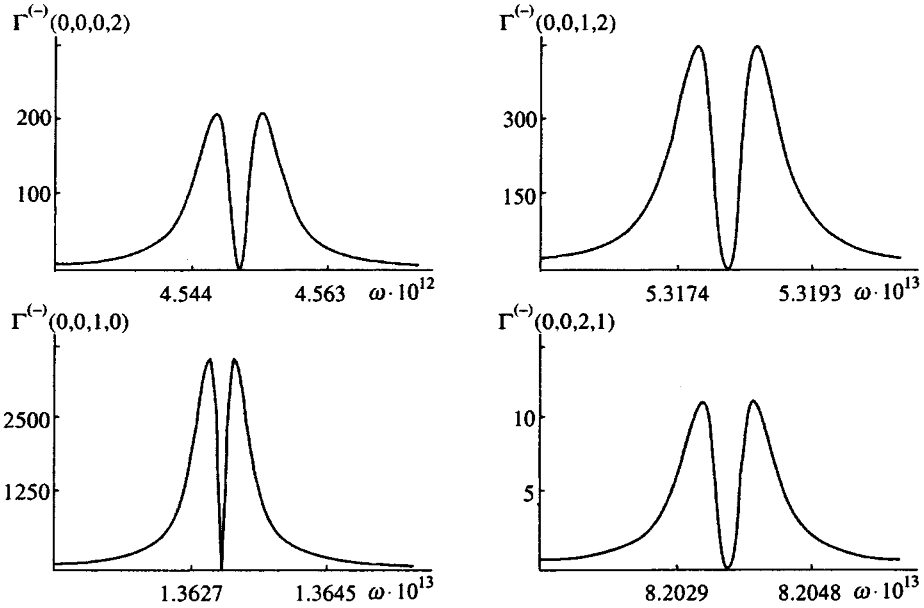


FIG. 5. Partial absorption coefficients in the case of absorption of PO-phonons; $\theta_0 = 45°$.

particular, their doublet nature remain the same. Figure 6 depicts the dependence of the partial absorption coefficients on the angle between **B** and the confinement plane. Clearly, the dependence of $\Gamma^{(\pm)}(0,0;n,m)$ on θ_0 is represented by a resonance curve. The resonance peaks are symmetric about the value of θ_0 at which $\Delta\omega = 0$. On the whole the general nature of the $\Gamma^{(\pm)}$ vs θ_0 dependence is the same as for $\Gamma^{(\pm)}(\omega)$ and $\Gamma^{(\pm)}(B)$.

7. DISCUSSION

The absorption peaks in the hybrid-phonon resonance are due primarily to the selection rules for the transitions in second-order perturbation theory and the law of energy conservation in such transitions. If we ignore optical phonon dispersion, the partial absorption peaks in $\Gamma^{(\pm)} \times (n,m;n',m')$ have delta-function singularities at the points where $\Delta\omega = 0$. Allowing for dispersion generally broadens these peaks, but it also leads to an interesting change in the shape of the absorption curve. The above results suggest that in this case the partial absorption coefficients in $\Gamma(\omega)$ have two narrow and high peaks to the left and right of the point at which $\Delta\omega = 0$ holds, while at the point itself absorption is nil. The separation in frequency between the peaks is extremely small. Thus, the doublet structure reveals itself because phonon dispersion is taken into account. We also note that the results of Sec. 6 imply that the doublet components are asymmetric. The widths of the maxima in the Γ vs ω curve are much smaller than the characteristic frequencies of the system. Figures 2–5 suggest that the ratio of the peak value of $\Gamma(\omega)$ to the value in the tail of the absorption line is

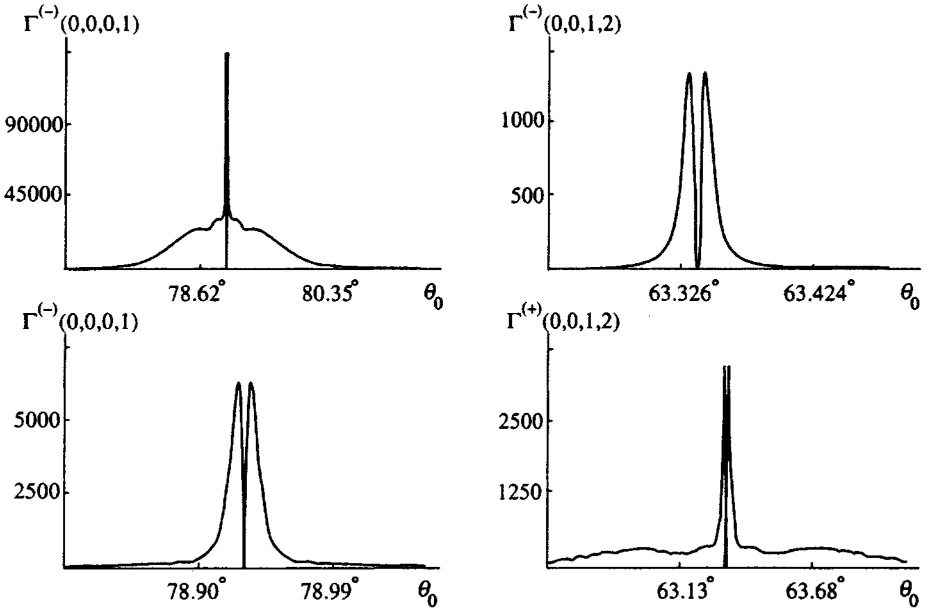


FIG. 6. Dependence of partial absorption coefficients on the angle θ_0 at $\omega_c = 3.4\times10^{13}\,\mathrm{s}^{-1}$, $\omega = 4\times10^{13}\,\mathrm{s}^{-1}$, and $T = 100$ K.

much larger than unity, with the result that near resonance points only the corresponding partial absorption coefficient contributes.

Note that this method of calculation is restricted by the condition $\omega\tau \gg 1$ (Ref. 16), with τ the relaxation time in the system, so that the passage to the limit $\omega \to 0$ in order to obtain DC conductivity in an oblique magnetic field is invalid.

The frequency factors $a(\omega)$ and $b(\omega)$ in the absorption coefficient have singularities at the points $\omega=\omega_1$ and $\omega=\omega_2$. These singularities, however, are unimportant in studies of the hybrid-phonon resonance, since they are shifted in relation to the points where $\Delta\omega=0$ by the optical phonon frequency, with the result that they lie in the distant region of the wings of the hybrid-phonon resonance lines.

There is also an analog of a cyclotron resonance in the system, which is due to purely electromagnetic transitions between hybrid-oscillator levels broadened by scattering.

To find the absorption we employ the same approach as we did in calculating the hybrid-phonon resonance. In first-order perturbation theory the absorption coefficient is

$$\Gamma(\omega)=\frac{2\pi\sqrt{\varepsilon(\omega)}}{c\hbar N_{\mathbf{f}}}(1-e^{-\hbar\omega/T})\sum_{\beta\beta'} f_0(\epsilon_\beta)|\langle\beta,0|H_R|\beta', -\mathbf{f}\rangle|^2\delta(\epsilon_\beta-\epsilon_{\beta'}+\hbar\omega). \tag{35}$$

Using (14) and summing, we get

$$\Gamma(\omega)=\frac{2\pi^2e^2n_0}{cm^*\sqrt{\varepsilon(\omega)}}\frac{\omega^2-\Omega^2}{\omega_1^2-\omega_2^2}[\delta(\omega_1-\omega)-\delta(\omega_2-\omega)]. \tag{36}$$

We allow for the time dependence in the hybrid-oscillator levels by introducing a phenomenological parameter $1/\tau$, the halfwidth of the absorption peak, where τ is the phenomenological relaxation time. Then Eq. (36) yields

$$\Gamma(\omega)=\frac{2\pi^3e^2n_0\tau}{cm^*\sqrt{\varepsilon(\omega)}}\left[\frac{\omega_1^2-\Omega^2}{\omega_1^2-\omega_2^2}\frac{1}{1+\tau^2(\omega-\omega_1)^2}+\frac{\omega_2^2-\Omega^2}{\omega_2^2-\omega_1^2}\frac{1}{1+\tau^2(\omega-\omega_2)^2}\right]. \tag{37}$$

We call this resonance at points $\omega=\omega_{1,2}$ a hybrid resonance. Since the hybrid-oscillator frequencies $\omega_{1,2}$ depend on the magnitude of the magnetic field and the angle between the magnetic field and the confinement plane, the position of the peaks on the Γ vs ω curve depends on these parameters. These peaks are shifted at least by the optical phonon frequency in relation to hybrid-phonon resonance peaks, with the result that the latter lie in the distant region of the hybrid resonance wings. Elementary estimates based on (37) and the numerical results of Sec. 6 yield $\Gamma^{(\pm)}(0,0;n,m) \gg \Gamma(\omega)$ in the wings of the hybrid resonance.

The physical nature of the hybrid-phonon resonance, discussed in the Introduction, suggests that the statistics of the electrons has no effect on the emergence of this resonance. Describing the shape of the absorption curve in a degenerate gas requires allowing for the Pauli exclusion principle for the final states. Moreover, at low temperatures the phonon number N_T is exponentially small, $\sim\exp(-\hbar\omega_0/T)$, and in the expression for $\Gamma(\omega)$ one can retain only the terms corresponding to processes with phonon emission. In view of this, for a degenerate gas Eq. (30) acquires the form

$$\begin{aligned}\Gamma(n,m;n',m')\\ &=\frac{e^2\omega_c\cos\theta_0(1-e^{-\hbar\omega/T})}{m^*c\hbar^2\omega\sqrt{\varepsilon(\omega)}l_1l_2V}\frac{n!m!}{n'!m'!}\sum_{\mathbf{q}}|D_{\mathbf{q}}|^2|A(\omega)|^2\\ &\times d^{2(m-m')}g^{2(n-n')}e^{-d^2-g^2}[L_{n'}^{n-n'}(g^2)L_{m'}^{m-m'}(d^2)]^2\\ &\times(N_{\mathbf{q}}^T+1)\delta(\Delta\omega)f_0(\epsilon_\beta)[1-f_0(\epsilon_{\beta'})],\end{aligned} \tag{38}$$

where $f_0(\epsilon_\beta)$ is the Fermi function, V is the normalization volume, and $\Delta\omega=\omega_1(n-n')+\omega_2(m-m')+\omega-\omega_q$. In the limit of $T=0$ the distribution is a step function, with the initial–state electrons ''occupying'' the energy interval $[\mu-\hbar\omega+\hbar\omega_q,\mu]$ and the final−state electrons the interval $[\mu,\mu+\hbar\omega-\hbar\omega_q]$. This suggests that the partial contributions to absorption are provided only by electrons with oscillator quantum numbers lying within the respective intervals.

A comparison of (38) and the results of Sec. 4 shows that the hybrid−phonon resonance peaks are at the same points and have the same structure as the peaks for a nondegenerate gas.

The resonance under discussion is observed against the background of lattice absorption by *TO*-phonons. These, however, provide only a monotonic contribution in the neighborhood of a hybrid−phonon resonance, which can be excluded by using the expression for the lattice absorption coefficient discussed in Ref. 12.

Note, finally, that the condition for hybrid quantization, $\hbar\omega_2 > T$ (since $\omega_1>\omega_2$), imposes a restriction on large values of the angle between the field and the confinement plane:[8]

$$\cos^2\theta_0 \gg \left(\frac{T}{\hbar\omega_2}\right)^2\left(1+\frac{\omega_c^2}{\Omega^2}\right).$$

Thus, for a field **B** that is almost parallel to the confinement plane the above results cease to be valid.

This work was supported by the Russian Fund for Fundamental Research and the Universities of Russia Program.

[1] T. Ando, J. Phys. Soc. Jpn. **44**, 765 (1978).
[2] T. Ando, Phys. Rev. B **19**, 2106 (1979).
[3] W. Beinvogl and J. F. Koch, Phys. Rev. Lett. **40**, 1736 (1978).
[4] T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys **54**, 437 (1982).
[5] D. Huang, G. Gumbs, and N. Horing, Phys. Rev. B **49**, 11 463 (1994).
[6] C. T. Liu, K. Nakamura, and D. C. Tsui, Appl. Phys. Lett. **55**, 168 (1989).
[7] R. Merlin, Solid State Commun. **64**, 99 (1987).
[8] V. A. Geĭler, V. A. Margulis, A. G. Nesmelov, and I. I. Chuchuev, Fiz. Tverd. Tela (Leningrad) **36**, 1994 (1994) [Phys. Solid State **36**, 1090 (1994)].
[9] J. H. Oh, K. J. Chang, G. Jhm, and D. J. Lee, Phys. Rev. B **50**, 15 397 (1994).

[10] V. A. Geĭler, V. A. Margulis, and I. V. Chudaev, Zh. Éksp. Teor. Fiz. **109**, 762 (1996) [JETP **82**, 409 (1996)].
[11] L. Brey, N. E. Johnson, and B. J. Halperin, Phys. Rev. B **40**, 10 647 (1989).
[12] F. G. Bass and I. B. Levinson, Zh. Éksp. Teor. Fiz. **49**, 914 (1965) [Sov. Phys. JETP **22**, 635 (1966)].
[13] R. K. Bakanas, F. G. Bass, and I. B. Levinson, Fiz. Tekh. Poluprovodn. **12**, 1457 (1978) [Sov. Phys. Semicond. **12**, 863 (1978)].
[14] R. K. Bakanas, Fiz. Tverd. Tela (Leningrad) **12**, 3408 (1970) [Sov. Phys. Solid State **12**, 2769 (1970)].
[15] H. Fröhlich, Adv. Phys. **3**, 3325 (1960).
[16] S. Visvanathan, Phys. Rev. **120**, 376 (1960).

Translated by Eugene Yankovsky

Drag effects in a two-layer system of spatially separated electrons and excitons

Yu. E. Lozovik*) and M. V. Nikitkov

Institute of Spectroscopy, Russian Academy of Sciences, 142092 Troitsk, Moscow Region, Russia

(Submitted 3 September 1996)
Zh. Éksp. Teor. Fiz. **111**, 1107–1119 (March 1997)

We discuss drag effects in a two-layer system of spatially separated electrons and excitons: the entrainment of excitons by moving electrons, and the entrainment of electrons by moving excitons. For the case of excitons entrained by electrons we find the drag velocity v_{drag}, and for electrons entrained by excitons we compute the induced electric field E_2. These drag effects can be sensitive indicators of the phase state of the excitons and of phase transitions in the exciton system (to a liquid phase, superfluid state, etc.). © *1997 American Institute of Physics.* [S1063-7761(97)02803-5]

1. INTRODUCTION

In Ref. 1, drag effects in a two-layer system of spatially separated electrons and holes were predicted theoretically, and their influence on phase transitions to a superfluid excitonic phase investigated (see also Ref. 2 and the papers cited therein). Later, Pogrebinskiĭ[3] discussed the entrainment of electrons by electrons in a semiconductor–insulator–semiconductor structure.[4] Price proposed a practical method for observing the drag effect in heterostructures. Subsequently, the drag effect was explored in a number of theoretical and experimental papers,[5–13] where various physical realizations of the drag effect were discussed in one-dimensional, two-dimensional, and three-dimensional systems.

In this paper we discuss new effects: drag phenomena in a two-layer system of spatially separated excitons and electrons. The electrical neutrality of excitons is a source of great difficulty in the investigation of their transport properties. Up to now, information about their transport properties has been obtained from local investigations of their recombination emission.[14] Therefore, it would interesting to investigate the drag exerted on electrons by excitons that are set into motion by a concentration gradient. The drag in this case is caused by the interaction of an exciton with an electron that polarizes it. In principle, the effects we will discuss could make it possible to learn things about the transport properties of excitons, and how they change as a result of phase transitions in the exciton system, by measuring the current or voltage associated with the entrained electrons. We will also discuss the inverse effect in a system of spatially separated electrons and excitons, namely the entrainment of excitons by moving electrons. In principle, this effect could make it possible to control the motion of excitons by varying the field in the electron layer. Naturally, all these effects will also occur for electrons and excitons in a single layer. The results we obtain are qualitatively correct for this case as well.

The interaction between an electron and an exciton is weaker than the interaction between two charged particles. Consequently, drag in an electron–exciton system will be harder to detect than drag in a purely electronic system. This fact leads to complications in seeing it experimentally in an exciton–electron system. However, we hope that the effect will nevertheless be observable.

In this paper we solve a system of two Boltzmann equations for the electron and exciton distribution functions. We find the velocity v_{drag}, which excitons acquire through interacting with the moving electrons. We obtain an expression for the electric field E_2 that arises in the electron system due to the entrainment of electrons by excitons.

2. DRAG COEFFICIENTS

Consider a double quantum-well structure. Assume that excitons are created by a laser in one of the wells (which we will denote with the number 1) with nonuniform density (due to, e.g., the use of a mask that is opaque to the laser radiation, focusing, etc.). Asssume that the other well (which we label with the number 2) contains a gas of electrons with density n_2. The width of the barrier between the wells we denote by d. In this problem we will not consider tunneling effects. Our goal is to compute the response of the exciton system to an external electric field applied to the electron system, and also the response of the electron system to forces exerted by the exciton system.

In the two-layer system of electrons and excitons, the exciton mass flux $\mathbf{i}=m_1 n_1 v_1$ and the electron charge flux $\mathbf{j}=-en_2\mathbf{v}_2$ can be expressed in terms of the exciton concentration gradient ∇n_1 and the external electric field $\mathbf{E}_2$ applied to the electron subsystem by

$$\mathbf{J}=\hat{K}\mathbf{S}, \tag{1}$$

where

$$\mathbf{J}=\begin{pmatrix}\mathbf{i}_1\\ \mathbf{j}_2\end{pmatrix},\quad \hat{K}=\begin{pmatrix}-m_1D_{11} & -n_1m_1\mu_{12}\\ em_1D_{21}/m_2 & en_2\mu_{22}\end{pmatrix},$$

$$\mathbf{S}=\begin{pmatrix}\nabla n_1\\ \mathbf{E}_2\end{pmatrix}.$$

Here D_{11} is the exciton diffusion coefficient, D_{21} is the coefficient of mutual diffusion of electrons and excitons, μ_{12} is the coefficient of mutual mobility of the excitons and electrons, and μ_{22} is the mobility coefficient of the electrons.

Note that in our two-layer system the coefficients D_{11} and μ_{22} include the interaction between electrons and excitons (as is true for D_{21} and μ_{12}).

When no external electric field is applied to the electron layer, but it is not connected into a closed circuit, creation of a concentration gradient in the exciton system gives rise to an induced electric field $\mathbf{E}_2^{\text{ind}}$ in the electron layer equal to

$$\mathbf{E}_2^{\text{ind}} = -\frac{m_1 D_{21}}{m_2 n_2 \mu_{22}} \nabla n_1 = -K_{21} \nabla n_1 . \tag{2}$$

In this case, we obtain the following expression for the exciton flux from (1):

$$\mathbf{i}_1 = -(m_1 D_{11} - n_1 m_1 \mu_{12} K_{21}) \nabla n_1 . \tag{3}$$

If, however, $\nabla n_1 = 0$, then we have for the exciton flux

$$\mathbf{i}_1 = -n_1 m_1 \mu_{12} \mathbf{E}_2 , \tag{4}$$

in this case, it follows from (4) that the velocity of the excitons is

$$\mathbf{v}_1 = -\mu_{12} \mathbf{E}_2 . \tag{5}$$

3. ENTRAINMENT OF ELECTRONS BY EXCITONS

When their density is nonuniform, excitons will diffuse in the direction of its decrease. As they move, the excitons interact with electrons, thereby imparting momentum to them. As a result of this, an induced electric field $\mathbf{E}_2$ appears in the electron layer.

Consider the case where the circuit to which the electron layer is connected is open. The kinetic equations for the exciton and electron distribution functions have the form

$$\frac{\partial f_1}{\partial \mathbf{r}_1} \mathbf{v}_1 = I_1 + I_{12} , \tag{6}$$

$$\frac{\partial f_2}{\partial \mathbf{r}_2} \mathbf{v}_2 + \frac{\partial f_2}{\partial \mathbf{p}_2} \dot{\mathbf{p}}_2 = I_{21} , \tag{7}$$

here I_1 is a collision integral that includes all exciton scattering processes except scattering of excitons by electrons, while I_{12} is the collision integral that takes the latter into account. In Eq. (6) there is no term with a momentum derivative $(\partial f_1 / \partial \mathbf{p}_1) \dot{\mathbf{p}}_1$. This term vanishes because no macroscopic force acts on the excitons, and $\dot{\mathbf{p}}_1 = 0$. When the circuit that includes the electron layer is open, the collision integral I_2 for the electrons in layer 2 also vanishes and in equilibrium an induced electric field $\mathbf{E}_2$ appears which compensates the change in momentum of the electrons as a result of electron–exciton collisions. The system of Eqs. (6) and (7) can be simplified if we assume that the electron–exciton interaction has only a weak effect on the process of exciton diffusion, so that $|I_1| \gg |I_{12}|$ and the term I_{12} can be omitted. If the nonuniformity in the distribution of electrons is small, i.e.,

$$\left| \frac{\partial f_2}{\partial \mathbf{r}_2} \mathbf{v}_2 \right| \ll \left| \frac{\partial f_2}{\partial \mathbf{p}_2} \dot{\mathbf{p}}_2 \right| ,$$

then we also can neglect the spatial derivative in (7). Then we obtain

$$\frac{\partial f_1}{\partial \mathbf{r}_1} \mathbf{v}_1 = I_1 , \tag{8}$$

$$\frac{\partial f_2}{\partial \mathbf{p}_2} \dot{\mathbf{p}}_2 = I_{21} . \tag{9}$$

Let us rewrite the left side of Eq. (8) in the following form (assuming $n_1 = n_1(x)$):

$$\frac{\partial f_1}{\partial \mathbf{r}_1} \mathbf{v}_1 = \frac{\partial f_1}{\partial \mu} \frac{\partial \mu}{\partial n_1} \frac{\partial n_1}{\partial x} v_{1x} + \frac{\partial f_1}{\partial T} \frac{\partial T}{\partial \mathbf{r}_1} \mathbf{v}_1 , \tag{10}$$

where $\mu = \mu(x)$ is the chemical potential of the exciton gas. We will neglect the temperature gradient and write the collision integral I_1 in the τ-approximation:

$$I_1 = -\frac{f_1 - f_1^0}{\tau_1} , \tag{11}$$

where τ_1 is the relaxation time of the excitons. Here f_1^0 is a Bose function with $\mu = \mu(n_{10}) = \mu_0$, normalized by

$$n_{10} = \int f_1^0 \frac{d\mathbf{p}_1}{(2\pi\hbar)^2} . \tag{12}$$

As a result, Eq. (8) acquires the form

$$\frac{\partial f_1}{\partial \mu} \frac{\partial \mu}{\partial n_1} \frac{\partial n_1}{\partial x} v_{1x} = -\frac{f_1 - f_1^0}{\tau_1} . \tag{13}$$

We linearize Eq. (13); for this we let $n_1(x) = n_{10} + \delta n_1(x)$, $\mu(n_1) = \mu(n_{10}) + \delta\mu(x)$, and write f_1 in the form

$$f_1 = f_1^0 + f_1^0 (1 + f_1^0) \psi_1 . \tag{14}$$

As a result, we obtain the following expression for ψ_1:

$$\psi_1 = -\frac{\tau_1}{k_B T} \frac{\partial \mu_0}{\partial n_{10}} \frac{\partial n_1}{\partial x} v_{1x} . \tag{15}$$

Let us assume that collisions between excitons and electrons are elastic. Then the collision integral in Eq. (9) has the form

$$\begin{aligned} I_{21} = \sum_{\sigma_{2'}} \int & w(\mathbf{p}_1 \mathbf{p}_2 ; \mathbf{p}_{1'} \mathbf{p}_{2'}) \{ f_{1'} f_{2'} (1 + f_1)(1 - f_2) \\ & - f_1 f_2 (1 + f_{1'})(1 - f_{2'}) \} \delta(\varepsilon_1 + \varepsilon_2 - \varepsilon_{1'} - \varepsilon_{2'}) \\ & \times \frac{d\mathbf{p}_1}{(2\pi\hbar)^2} \frac{d\mathbf{p}_{2'}}{(2\pi\hbar)^2} , \end{aligned} \tag{16}$$

where w is the probability of scattering of an exciton by an electron.

We linearize Eq. (9) by writing f_2 in a form analogous to (14), i.e.,

$$f_2 = f_2^0 + f_2^0 (1 - f_2^0) \psi_2 , \tag{17}$$

where f_2^0 is a Fermi function that satisfies the normalization condition

$$n_2 = 2 \int f_2^0 \frac{d\mathbf{p}_2}{(2\pi\hbar)^2} . \tag{18}$$

Since $\dot{\mathbf{p}}_2 = -e\mathbf{E}_2$, we obtain

$$-e\mathbf{E}_2\frac{\partial f_2^0}{\partial \mathbf{p}_2}=\sum_{\sigma_{2'}}\int w\{f_1^0f_2^0(1+f_{1'}^0)(1-f_{2'}^0)\}(\psi_{1'}+\psi_{2'}-\psi_1-\psi_2)\,\delta(\varepsilon_1+\varepsilon_2-\varepsilon_{1'}-\varepsilon_{2'})\times\frac{d\mathbf{p}_1}{(2\pi\hbar)^2}\frac{d\mathbf{p}_{2'}}{(2\pi\hbar)^2}. \qquad (19)$$

When there is no current in the electron layer, ψ_2 and $\psi_{2'}$ vanish. We multiply both parts of Eq. (19) by $p_{2x}/(2\pi\hbar)^2$, integrate over all the momenta $\mathbf{p}_2$, and sum over projections of the spin σ_2. Then Eq. (19) can be written in the form

$$E_2=-K_{21}\nabla_x n_1, \qquad (20)$$

where the coefficient K_{21} (see Eq. (2)) equals

$$K_{21}=\frac{2\tau_1}{em_1n_2k_BT}\frac{\partial\mu_0}{\partial n_{10}}\int wf_1^0f_2^0(1+f_{1'}^0)(1-f_{2'}^0)\times p_{2x}(p_{1'x}-p_{1x})\,\delta(\varepsilon_1+\varepsilon_2-\varepsilon_{1'}-\varepsilon_{2'})\times\frac{d\mathbf{p}_1}{(2\pi\hbar)^2}\frac{d\mathbf{p}_2}{(2\pi\hbar)^2}\frac{d\mathbf{p}_{2'}}{(2\pi\hbar)^2}. \qquad (21)$$

Taking into account the conservation of momentum during collisions: $p_{1x}-p_{1'x}=p_{2'x}-p_{2x}$, and also the symmetry of the integral in Eq. (21), we obtain

$$K_{21}=\frac{\tau_1}{2em_1n_2k_BT}\frac{\partial\mu_0}{\partial n_{10}}\int wf_1^0f_2^0(1+f_{1'}^0)(1-f_{2'}^0)\times q^2\delta(\varepsilon_1+\varepsilon_2-\varepsilon_{1'}-\varepsilon_{2'})\times\frac{d\mathbf{p}_1}{(2\pi\hbar)^2}\frac{d\mathbf{p}_2}{(2\pi\hbar)^2}\frac{d\mathbf{p}_{2'}}{(2\pi\hbar)^2}, \qquad (22)$$

where $\mathbf{q}=\mathbf{p}_{2'}-\mathbf{p}_2$.

For the probability of scattering of an exciton by an electron we will use the Born approximation:

$$w(\mathbf{p}_1\mathbf{p}_2;\mathbf{p}_{1'}\mathbf{p}_{2'})=\frac{2\pi}{\hbar}|U(q)|^2, \qquad (23)$$

where $U(q)$ is the Fourier transform of the effective exciton–electron interaction energy (see Sec. 5).

Use of the Born approximation in this case is correct because the condition for its applicability has the form

$$\gamma n^{3/2}\ll\hbar v, \quad \text{if } d<n^{-1/2},$$

$$\gamma d^{-3}\ll\hbar v, \quad \text{if } d>n^{-1/2},$$

where $n=\max\{n_1,n_2\}$; the quantity γ is defined below in Eq. (71), and v is the electron–exciton relative velocity. This condition is fulfilled over a wide range of densities and temperatures. For example, for an exciton radius $a=20$ Å, a dielectric constant $\epsilon = 10$, in the medium and $d<n^{-1/2}$, we obtain the condition $n^{3/2}/v\ll10^{13}\text{cm}^{-4}$.

Using the identity

$$\delta(\varepsilon_1+\varepsilon_2-\varepsilon_{1'}-\varepsilon_{2'})=\int d(\hbar\omega)\,\delta(\varepsilon_1-\varepsilon_{1'}-\hbar\omega)\times\delta(\varepsilon_2-\varepsilon_{2'}+\hbar\omega), \qquad (24)$$

we obtain a different expression for the coefficient K_{21}:

$$K_{21}=\frac{1}{2}\frac{\tau_1}{e\hbar^2k_BTm_1n_2}\frac{\partial\mu_0}{\partial n_{10}}\int_0^\infty U^2(q)q^3dq\int_{-\infty}^\infty d\omega\times\int_{-\infty}^\infty f_1^0(1+f_{1'}^0)\,\delta(\varepsilon_1-\varepsilon_{1'}-\hbar\omega)\frac{d\mathbf{p}_1}{(2\pi\hbar)^2}\times\int_{-\infty}^\infty f_2^0(1-f_{2'}^0)\,\delta(\varepsilon_2-\varepsilon_{2'}+\hbar\omega)\frac{d\mathbf{p}_2}{(2\pi\hbar)^2}. \qquad (25)$$

We rewrite Eq. (25) in the form

$$K_{21}=\frac{1}{8\pi^2}\frac{\tau_1}{e\hbar^2k_BTm_1n_2}\frac{\partial\mu_0}{\partial n_{10}}\int_0^\infty U^2(q)q^3dq\times\int_{-\infty}^\infty\frac{\text{Im }\chi^B(q,\omega)\text{Im }\chi^F(q,\omega)}{\sinh^2(\hbar\omega/2k_BT)}d\omega, \qquad (26)$$

where

$$\chi^B(q,\omega)=-\int\frac{f^0(\varepsilon_1)-f^0(\varepsilon_{1'})}{(\varepsilon_1-\varepsilon_{1'}+\hbar\omega+i\delta)}\frac{d\mathbf{p}_1}{(2\pi\hbar)^2}, \qquad (27)$$

$$\chi^F(q,\omega)=-\int\frac{f^0(\varepsilon_2)-f^0(\varepsilon_{2'})}{(\varepsilon_2-\varepsilon_{2'}+\hbar\omega+i\delta)}\frac{d\mathbf{p}_2}{(2\pi\hbar)^2}. \qquad (28)$$

Let the system parameters be such that the excitons and electrons obey Boltzmann distributions. In this case we find a simpler expression for K_{21} from (25):

$$K_{21}=\frac{1}{4\sqrt{2\pi}}\frac{\tau_1n_{10}}{em_1}\sqrt{\frac{M}{(k_BT)^3}}\frac{\partial\mu_0}{\partial n_{10}}\int_0^\infty U^2(k)k^2\times\exp\left(-\frac{k^2\hbar^2}{8Mk_BT}\right)dk, \qquad (29)$$

where $M=m_1m_2/(m_1+m_2)$, $k=q/\hbar$ and U is the potential energy of the interaction between electrons and excitons.

We now wrte E_2 in terms of the exciton diffusion velocity v_{diff}. Doing so, we find

$$v_{\text{diff}}=\frac{1}{n_{10}}\int v_{1x}f_1^0\psi_1\frac{d\mathbf{p}_1}{(2\pi\hbar)^2}. \qquad (30)$$

Taking into account Eq. (15), we obtain from (30)

$$v_{\text{diff}}=-\frac{\tau_1}{m_1}\frac{\partial\mu_0}{\partial n_{10}}\frac{\partial n_1}{\partial x}. \qquad (31)$$

Let us write the equation for E_2 in the form

$$E_2=\lambda_{21}v_{\text{diff}}. \qquad (32)$$

Using (29), we find the following expression for λ_{21}:

$$\lambda_{21}=\frac{1}{4\sqrt{2\pi}}\frac{n_{10}}{e}\sqrt{\frac{M}{(k_BT)^3}}\times\int_0^\infty U^2(k)k^2\exp\left(-\frac{k^2\hbar^2}{8Mk_BT}\right)dk. \qquad (33)$$

4. ENTRAINMENT OF EXCITONS BY ELECTRONS

Consider the reverse situation, i.e., where the electrons drag the excitons. Let us calculate the velocity imparted to the excitons when they interact with electrons.

In this case, the kinetic equations have the form

$$\frac{\partial f_1}{\partial \mathbf{r}_1}\mathbf{v}_1 = I_1 + I_{12}, \tag{34}$$

$$\frac{\partial f_2}{\partial \mathbf{r}_2}\mathbf{v}_2 + \frac{\partial f_2}{\partial \mathbf{p}_2}\dot{\mathbf{p}}_2 = I_2 + I_{21}. \tag{35}$$

In order to simplify Eqs. (34) and (35) we will assume that the electron current is uniform, and also that the collision integral I_{21} is a small correction to I_2 and thus can be neglected in (35). Furthermore, if we are interested only in the drag velocity v_{drag}, assuming that it dominates over the diffusion velocity v_{diff}, we can omit the term containing the spatial derivative of f_1 in Eq. (34). Then we obtain

$$I_1 + I_{12} = 0, \tag{36}$$

$$\frac{\partial f_2}{\partial \mathbf{p}_2}\dot{\mathbf{p}}_2 = I_2. \tag{37}$$

The usual procedure of substituting f_2 into Eq. (37) in Eq. (17) and using the τ approximation for I_2 gives

$$\psi_2 = -\frac{\tau_2}{m_2 k_B T} e\mathbf{E}_2\mathbf{p}_2. \tag{38}$$

Here $\mathbf{E}_2 = \{E_2, 0\}$ is the intensity of the external electric field and τ_2 is the relaxation time of the electrons.

Let us write Eq. (36) in more detail:

$$I_1 = -\sum_{\sigma_2,\sigma_{2'}} \int w\{f_{1'} f_{2'}(1+f_1)(1-f_2) - f_1 f_2 (1+f_{1'}) \times (1-f_{2'})\}\,\delta(\varepsilon_1+\varepsilon_2-\varepsilon_{1'}-\varepsilon_{2'})\,\frac{d\mathbf{p}_{1'}}{(2\pi\hbar)^2}\,\frac{d\mathbf{p}_2}{(2\pi\hbar)^2}. \tag{39}$$

By substituting f_1 in the form (14) and f_2 in the form (17) into (39), and also setting $I_1 = -(f_1 - f_1^0)/\tau_1$, we obtain the linearized equation

$$f_1^0(1+f_1^0)\psi_1 = 2\tau_1 \int w f_1^0 f_2^0 (1+f_{1'}^0)(1-f_{2'}^0)(\psi_{1'} + \psi_{2'} - \psi_1 - \psi_2)\,\delta(\varepsilon_1+\varepsilon_2-\varepsilon_{1'}-\varepsilon_{2'}) \times \frac{d\mathbf{p}_{1'}}{(2\pi\hbar)^2}\,\frac{d\mathbf{p}_2}{(2\pi\hbar)^2}. \tag{40}$$

The condition $v_{\text{diff}} \ll v_{\text{drag}}$ permits us to drop the terms in ψ_1 and $\psi_{1'}$ under the integral sign in Eq. (40).

The expression for the drag velocity has the form

$$v_{\text{drag}} = \frac{1}{m_1 n_{10}} \int p_{1x} f_1^0 (1+f_1^0)\psi_1 \frac{d\mathbf{p}_1}{(2\pi\hbar)^2}. \tag{41}$$

We write the equation for v_{drag} (see (5)) as

$$v_{\text{drag}} = -\mu_{12}E_2. \tag{42}$$

In light of the relations (40) and (41), we have for μ_{12}

$$\mu_{12} = \frac{2e\tau_1\tau_2}{m_1 m_2 n_{10} k_B T} \int w f_1^0 f_2^0 (1+f_{1'}^0)(1-f_{2'}^0) p_{1x} \times (p_{2'x} - p_{2x})\,\delta(\varepsilon_1+\varepsilon_2-\varepsilon_{1'}-\varepsilon_{2'}) \times \frac{d\mathbf{p}_1}{(2\pi\hbar)^2}\,\frac{d\mathbf{p}_{1'}}{(2\pi\hbar)^2}\,\frac{d\mathbf{p}_2}{(2\pi\hbar)^2}. \tag{43}$$

Omitting calculations like those described above, we finally find an expression for μ_{12}:

$$\mu_{12} = \frac{1}{8\pi^2}\,\frac{\tau_1\tau_2}{m_1 m_2 n_{10}\hbar^2}\,\frac{e}{k_B T}\int_0^\infty dq \int_{-\infty}^\infty d\omega U^2(q) q^3 \times \frac{\operatorname{Im}\chi^B(q,\omega)\operatorname{Im}\chi^F(q,\omega)}{\sinh^2(\hbar\omega/2k_B T)}, \tag{44}$$

or in the classical case

$$\mu_{12} = \frac{\tau_1\tau_2}{4\sqrt{2\pi}}\,\frac{e n_2}{m_1 m_2}\sqrt{\frac{M}{(k_B T)^3}}\int_0^\infty U^2(k) k^2 \times \exp\left(-\frac{k^2\hbar^2}{8Mk_B T}\right) dk. \tag{45}$$

5. EFFECTIVE INTERACTION IN THE ELECTRON–EXCITON SYSTEM

In order to compute the effective interaction energy in a two-layer system of electrons and excitons we will use the self-consistent approximation. Quantities that refer to excitons we label with the number 1, while those that refer to electrons we label with 2. If the exciton radius is much smaller than the distance between electrons and excitons, then the interaction energy between an isolated electron and an exciton has the form

$$V_{e-\text{ex}}(\mathbf{r}_1 - \mathbf{r}_2, d) = -\frac{\gamma}{[(\mathbf{r}_1-\mathbf{r}_2)^2 + d^2]^2},$$

where $\gamma = \alpha e^2/2\epsilon$, α is the polarizability of the two-dimensional exciton in its ground state, d is the distance between layers, $|\mathbf{r}_1 - \mathbf{r}_2|$ the distance between excitons and electrons along the layers, and ϵ is the dielectric permittivity of the medium. In deriving expressions for the effective interaction energy in a multiparticle electron–exciton system we will assume that the exciton–exciton interaction is negligibly small compared to the electron–exciton interaction and omit it. Let us place a test charge $-e$ in the electron subsystem at the origin of the coordinates. The linearized kinetic equations for the distribution functions of excitons and electrons have the form

$$\frac{\partial f_1}{\partial \mathbf{r}_1}\mathbf{v}_1 + \frac{\partial f_1^0}{\partial \mathbf{p}_1}\dot{\mathbf{p}}_1 = 0, \tag{46}$$

$$\frac{\partial f_2}{\partial \mathbf{r}_2}\mathbf{v}_2 + \frac{\partial f_2^0}{\partial \mathbf{p}_2}\dot{\mathbf{p}}_2 = 0, \tag{47}$$

where

$$\dot{\mathbf{p}}_1 = -\frac{\partial}{\partial \mathbf{r}_1}U(\mathbf{r}_1, d), \quad \dot{\mathbf{p}}_2 = -\frac{\partial}{\partial \mathbf{r}_2}U(\mathbf{r}_2, 0).$$

The interaction energies $U(\mathbf{r},0)$ and $U(\mathbf{r},d)$ satisfy the equations

$$U(\mathbf{r},0)=\int \rho_2(\mathbf{r}')\,\frac{e^2}{\epsilon|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}' -\int \frac{\gamma\rho_1(\mathbf{r}')d\mathbf{r}'}{[(\mathbf{r}-\mathbf{r}')^2+d^2]^2}+\frac{e^2}{er}, \tag{48}$$

$$U(\mathbf{r},d)=-\int \frac{\gamma\rho_2(\mathbf{r}')d\mathbf{r}'}{[(\mathbf{r}-\mathbf{r}')^2+d^2]^2}-\frac{\gamma}{(r^2+d^2)^2}, \tag{49}$$

where

$$\rho_1(\mathbf{r})=\int f_1(\mathbf{r},\mathbf{p})\,\frac{d\mathbf{p}}{(2\pi\hbar)^2},$$

$$\rho_2(\mathbf{r})=2\int f_2(\mathbf{r},\mathbf{p})\,\frac{d\mathbf{p}}{(2\pi\hbar)^2}.$$

Fourier-transforming Eqs. (46), (49), we obtain (assuming that $\mathbf{k}=\{k,0\}$)

$$f_1(\mathbf{k},\mathbf{p})=\frac{1}{v_{1x}}\frac{\partial f_1^0}{\partial p_{1x}}\,U(\mathbf{k},d), \tag{50}$$

$$f_2(\mathbf{k},\mathbf{p})=\frac{1}{v_{2x}}\frac{\partial f_2^0}{\partial p_{2x}}\,U(\mathbf{k},0), \tag{51}$$

where f_1^0 is a Bose function and f_2^0 is a Fermi function, and

$$U(\mathbf{k},0)=\frac{4\pi e^2}{\epsilon k}\int f_2(\mathbf{k},\mathbf{p})\,\frac{d\mathbf{p}}{(2\pi\hbar)^2} -\gamma F(\mathbf{k},d)\int f_1(\mathbf{k},\mathbf{p})\,\frac{d\mathbf{p}}{(2\pi\hbar)^2}+\frac{2\pi e^2}{\epsilon k}, \tag{52}$$

$$U(\mathbf{k},d)=-2\gamma F(\mathbf{k},d)\int f_2(\mathbf{k},\mathbf{p})\,\frac{d\mathbf{p}}{(2\pi\hbar)^2}-\gamma F(\mathbf{k},d), \tag{53}$$

where the function $F(\mathbf{k},d)=(\pi k/d)K_1(kd)$, where $K_1(z)$ is a modified Bessel function.

From Eqs. (50)–(53) we obtain a system of two algebraic equations for determining $U(k,0)$ and $U(k,d)$:

$$U(k,0)=\frac{2\pi e^2}{\epsilon k}\,\beta_2U(k,0)-\gamma\beta_1F(k,d)U(k,d) +\frac{2\pi e^2}{\epsilon k}, \tag{54}$$

$$U(k,d)=-\gamma\beta_2F(k,d)U(k,0)-\gamma F(k,d), \tag{55}$$

where

$$\beta_1=\int \frac{1}{v_x}\frac{\partial f_1^0}{\partial p_x}\frac{d\mathbf{p}}{(2\pi\hbar)^2},\quad \beta_2=2\int \frac{1}{v_x}\frac{\partial f_2^0}{\partial p_x}\frac{d\mathbf{p}}{(2\pi\hbar)^2}. \tag{56}$$

If f_1^0 and f_2^0 are Boltzmann distributions, then

$$\beta_1=-\frac{n_{10}}{k_BT},\quad \beta_2=-\frac{n_2}{k_BT}, \tag{57}$$

If f_2^0 is a Fermi step function, then

$$\beta_2=-\frac{m_2}{\pi\hbar^2}. \tag{58}$$

As a result, the expression for the effective interaction energy in a multiparticle electron–exciton system has the form

$$U(k,d)=-\frac{\gamma F(k,d)}{1-2\pi e^2\beta_2/\epsilon k-\gamma^2\beta_1\beta_2F^2}. \tag{59}$$

The expression for the effective interaction energy has the same form in a one-layer system of electrons and excitons, except that the function $F(k,d)$ is replaced by $F(k,a)$, defined as follows:

$$F(k,a)=\int \frac{e^{-i\mathbf{kr}}d\mathbf{r}}{r^4}=2\pi\int_a^\infty \frac{J_0(kr)dr}{r^3},$$

where a is the exciton radius and J_0 is a Bessel function.

By defining the function $F(k,a)$ in this way, we assert that the interaction between an electron and an exciton is dipole–charge at distances down to the exciton size and equals zero at smaller distances.

We now derive the value of the parameter $\gamma=\alpha e^2/2\epsilon$ entering into Eq. (59).

To second order in perturbation theory with respect to the electron–exciton interaction operator $\hat{V}=-\mathbf{dE}=-e\mathbf{dR}/\epsilon R^3$ (where $\mathbf{d}=e\mathbf{r}$ is the exciton dipole moment and R the distance between the electron and the exciton), the energy of interaction between an exciton in its ground state and an electron has the form

$$W=\sum_k{}'\,\frac{|V_{0k}|^2}{E_0-E_k}=\frac{e^4}{\epsilon^2R^4}\sum_k{}'\,\frac{|x_{0k}|^2}{E_0-E_k}=-\frac{\alpha e^2}{2\epsilon R^4}. \tag{60}$$

We will use this expression to calculate the effective interaction in the many-particle electron–exciton system. First we calculate the polarizability α entering into the parameter γ for a two-dimensional exciton:

$$\alpha=-\frac{2e^2}{\epsilon}\sum_k{}'\,\frac{|x_{0k}|^2}{E_0-E_k}. \tag{61}$$

We introduce the auxiliary operator $\hat{b}$ by[15]

$$x=\frac{m_1}{\hbar}\frac{d\hat{b}}{dt}. \tag{62}$$

Then we obtain for α the expression

$$\alpha=\frac{2im_1e^2}{\epsilon\hbar^2}\,(x\hat{b})_{00}. \tag{63}$$

Consider the action of the operator $\hat{b}$ on the wave function ψ_0 for the exciton ground state:

$$x\psi_0=\frac{m_1}{\hbar}\frac{d\hat{b}}{dt}\,\psi_0=\frac{im_1}{\hbar^2}\,(\hat{H}\hat{b}-\hat{b}\hat{H})\psi_0, \tag{64}$$

where $\hat{H}$ is the system Hamiltonian.

Let $\hat{b}\psi_0=b(\mathbf{r})\psi_0$. The equation $[-\hbar^2\nabla^2/2m_1 + U(r)]\psi_0=E_0\psi_0$ implies that Eq. (64) has the form

$$ix\psi_0=\frac{1}{2}\,\nabla^2 b(\mathbf{r})\psi_0+\nabla b(\mathbf{r})\nabla\psi_0 . \tag{65}$$

Substituting $b(\mathbf{r})=f(r)\cos\phi$ into Eq. (65) reduces it to the form

$$ir=\frac{1}{2}\,f''+\frac{1}{2}\,\frac{f'}{r}-\frac{1}{2}\,\frac{f}{r^2}+f'\,\frac{\psi_0'}{\psi_0}. \tag{66}$$

The ground-state wave function of the two-dimensional exciton is

$$\psi_0=\sqrt{\frac{2}{\pi a^2}}\,e^{-r/a}, \tag{67}$$

where $a=\epsilon\hbar^2/2m_1e^2$ is the radius of the two-dimensional exciton in its ground state, $m_1=m_em_h/(m_e+m_h)$, m_e and m_h are the electron and hole masses.

Solving Eq. (66), taking (67) into account, we obtain

$$f=A\left(1+\frac{a}{2r}\right)+B\,\frac{e^{2r/a}}{r}-\frac{ia}{2}\times\left(r^2+\frac{3ar}{2}+\frac{3a^2}{2}+\frac{3a^3}{4r}\right). \tag{68}$$

The coefficients A and B are chosen from the condition of finiteness of $f\psi_0$ as $\tau\rightarrow 0$ and $\tau\rightarrow\infty$. As a result, the solution takes the form

$$f=-\frac{ia}{2}\left(r^2+\frac{3ar}{2}\right). \tag{69}$$

For the polarizability α, taking (63) and (67) into account we obtain

$$\alpha=\frac{21}{8}\,\frac{m_1e^2}{\epsilon\hbar^2}\,a^4=\frac{21}{16}\,a^3. \tag{70}$$

Thus, the parameter γ entering into Eq. (59) for the effective interaction energy in a many-particle electron–exciton system equals

$$\gamma=\frac{21}{32}\,\frac{e^2a^3}{\epsilon}. \tag{71}$$

6. CONCLUSION

Let us discuss possible types of experiments to detect entrainment excitons by electrons and electrons by excitons. First consider entrainment electrons by excitons. A directed flux of excitons can be obtained if we locally create excitons (for example, using a laser with cw pumping) at one of the edges of layer 1. The excitons will diffuse away from the edge of the layer in the direction opposite to the concentration gradient. During the diffusion process the excitons will partially recombine, and also interact with electrons in the second layer. The response of the electron subsystem to the exciton diffusion will either be an induced current of electrons or an induced electric field (voltage) in layer 2, which can be measured.

For the case of entrainment of excitons by electrons, the drag velocity v_{drag} can be measured using two optical probes (fibers) or apertures in opaque masks located a set distance R from one another. These probes measure the intensity of the luminescence. By measuring the time interval between maxima of the luminescence, we can determine the drag velocity of the exciton v_{drag} (this experiment is possible when a pulsed laser is used). Actually, both diffusion (connected with the concentration gradient) and drag give contributions to the measured exciton velocity. However, the role of diffusion can be isolated by measuring the velocity of exciton motion when there is no electron current dragging the excitons.

Analogous experimental setups are possible when the electrons and excitons share a common layer.

In a one-layer system it is possible for a rather shallow electron–exciton bound state to form[16] due to the polarization interaction between them. However, the predicted binding energy of this state should decrease markedly when we take into account the Pauli exchange repulsion between the free and excitonic electron and screening (the latter is especially important for exciton concentrations $n_2>1/r_1^2$, where r_1^2 is the radius of the state of a single electron bound to a single exciton). For temperatures larger than 0.2 ionization energies of this bound state, the latter gives no contribution to our kinetic equations.

The formation of electron–exciton bound states in a two-layer system turns out to be even less important because even at distances $d=100$ Å the ionization temperature of these states is less than 1 K.

These drag effects, in particular the electric field (voltage) induced by excitons in the electron layer, can be sensitive indicators of the state of the excitonic subsystem and its phase transitions. For example, the response should change markedly when a phase transition to the liquid state takes place in the two-dimensional exciton system. In particular, the appearance of moving insulating excitonic droplets can lead to electric current pulses in the electron layer.

It would be interesting to study the implications of excitonic Bose–Einstein statistics using these drag effects, in particular the appearance of an excitonic Bose–Einstein condensate. In a two-dimensional system of excitons at low temperatures, where Bose–Einstein condensation does not occur (in the thermodynamic limit) it would be interesting to investigate how the electron drag is affected by the appearance first of a local superfluid density (with uncorrelated phases) and then a global superfluid exciton density at the Kosterlitz–Thouless transition temperature. In the crossover region where the local superfluid density appears, the drag coefficients (mutual mobility and mutual diffusion) should slowly increase and be discontinuous at the Kosterlitz–Thouless transition point.

This work was supported by grants from the Russian Fund for Fundamental Research and the programs ‘‘Physics of Solid-State Nanostructures,’’ and ‘‘Fundamental Spectroscopy.’’

*)e-mail: lozovik@isan.msk.su.

---

[1] Yu. E. Lozovik and V. I. Yudson, JETP Lett. **22**, 274 (1975); Zh. Éksp. Teor. Fiz. **71**, 738 (1976) [Sov. Phys. JETP **44**, 389 (1976)].

[2] Yu. E. Lozovik, Report on the Adriatico Res. Conf. on Low-Dim. Electron Systems, ICTP (Trieste, Italy, 1976).
[3] M. B. Pogrebinskiĭ, Fiz. Tekh. Poluprovodn. **11**, 637 (1977) [Sov. Phys. Semicond. **11**, 372 (1977)].
[4] P. J. Price, Physica B **117&118**, 750 (1983).
[5] T. J. Gramila, J. P. Eisenstein, A. H. MacDonald *et al.*, Phys. Rev. Lett. **66**, 1216 (1991).
[6] U. Sivan, P. M. Solomon, and H. Shtrikman, Phys. Rev. Lett. **68**, 1196 (1992).
[7] T. J. Gramila, J. P. Eisenstein, A. H. MacDonald *et al.*, Phys. Rev. B **47**, 12957 (1993).
[8] A.-P. Jauho and H. Smith, Phys. Rev. B **47**, 4420 (1993).
[9] L. Zheng and A. H. MacDonald, Phys. Rev. B **48**, 8203 (1993).
[10] Yu. M. Sirenko and P. Vaslopoulos, Phys. Rev. B **46**, 1611 (1992).
[11] H. C. Tso, P. Vaslopoulos, and F. M. Peeters, Phys. Rev. Lett. **68**, 2516 (1992).
[12] K. Flensberg and B. Yu-K Hu, Phys. Rev. B **73**, 3572 (1994).
[13] G. Vignale and A. H. MacDonald, Phys. Rev. Lett. **76**, 2786 (1996).
[14] V. Butov, A. Zrenner, G. Abstreiter *et al.*, Phys. Rev. Lett. **73**, 304 (1994).
[15] L. D. Landau and I. M. Lifshitz, *Quantum Mechanics* (Pergamon Press, Oxford, 1977 [3rd ed.]; Nauka, Moscow, 1989 [4th ed.]).
[16] V. S. Babichenko and M. N. Kiselev, JETP Lett. **57**, 179 (1993).

Translated by Frank J. Crowne

Current–voltage characteristics of tunnel junctions between superconductors with anisotropic pairing

Yu. S. Barash and A. A. Svidzinskiĭ

P. N. Lebedev Physics Institute, Russian Academy of Sciences, 117924 Moscow, Russia

(Submitted 26 September 1996)
Zh. Éksp. Teor. Fiz. **111**, 1120–1146 (March 1997)

Charge transfer in tunnel junctions between superconductors with anisotropic singlet pairing is considered theoretically on the basis of the Eilenberger equations for the quasiclassical Green's functions. New singularities of the current–voltage characteristics, which are characteristic of the case of anisotropic pairing, are treated analytically assuming that the electrons reflect specularly from the boundaries of the tunnel barrier. All four contributions to the tunneling current are investigated. Two of them describe Josephson tunneling, and the other two contributions correspond to the quasiparticle current (the last term appears only for a variable voltage). Different dependences of the order parameter on the momentum directions in the interior of the superconductors and different orientations of the crystal axes relative to the junction plane are considered. The results of numerical calculations of the current–voltage characteristics for several particular cases are presented.
[S1063-7761(97)02903-X]

1. INTRODUCTION

Charge transfer through a tunnel junction between superconductors with isotropic pairing was considered theoretically 30 years ago. In particular, a microscopic description of the total tunneling current that flows when an external variable voltage is applied to a junction was obtained within the tunneling Hamiltonian formalism by Larkin and Ovchinnikov[1] and by Werthamer[2] (see also Ref. 3 and the references cited therein). It has recently become clear for superconductors with anisotropic pairing that measurements of the Josephson and quasiparticle currents provide important information on the structure of the superconducting order parameter on both sides of the junction, on the proximity effect, and on several other surface effects near the plane of a tunnel junction between superconductors.[4–9] Investigations in this area have attracted a great deal of attention, primarily because of the importance of determining the structure of the anisotropic superconducting order parameter for different high-temperature superconductors (see, for example, Ref. 10 and 11).

One of the characteristic features of the microscopic description of charge transfer through tunnel junctions between superconductors with anisotropic pairing is the ineffectiveness of the tunneling Hamiltonian formalism in this case.[12] This method, which can usually be employed successfully for tunnel junctions between superconductors with isotropic pairing, contains uncertainties in the case of anisotropic pairing because of the significant momentum dependence of the matrix elements describing the tunneling between the superconductors. Choosing the matrix elements independent of the momentum directions (which is permissible for isotropic s-wave pairing) leads to incorrect results for anisotropic pairing. At the same time, the momentum dependence of tunneling matrix elements cannot be obtained within this method. The use of a systematic microscopic description of the tunneling of a charge between superconductors based on a calculation of the electron propagators in the tunnel barrier is not plagued by the uncertainties inherent in the tunneling Hamiltonian formalism.

In this paper we analytically investigate the singularities on the current–voltage characteristics of tunnel junctions between superconductors with anisotropic pairing, to which an external, generally variable voltage is applied. We use a microscopic approach based on a microscopic expression for the tunneling current and on the solution of the Eilenberger equations together with the corresponding boundary conditions for quasiclassical electron propagators. In the special case of a junction between superconductors with isotropic pairing, our results coincide with those found in Ref. 1. In Ref. 1 general expressions were obtained, and the singular points of the functions $I_m(V)$ $(m=1, 2, 3, 4)$, which appear in the expression for the total tunneling current, were described [see, for example, (6) and (7) below]. The functions $I_{1,2}(V)$ play the role of amplitudes in the expression for the Josephson current, which for a constant voltage reduces to

$$I_1(V)\sin\left(\chi_1-\chi_2+2\frac{eVt}{\hbar}\right)+I_2(V)\cos\left(\chi_1-\chi_2+2\frac{eVt}{\hbar}\right).$$

The functions $I_{3,4}(V)$ describe the quasiparticle current. The function $I_4(V)$ appears only in the case of a variable voltage, while in the case of a constant voltage the quasiparticle current reduces to the single term $I_3(V)$. In particular, it was shown that at the voltage $|eV|=\Delta_1+\Delta_2$ the functions $I_1(V)$ and $I_4(V)$ diverge logarithmically, while $I_2(V)$ and $I_3(V)$ undergo jumps. Similar singularities were found for a variable voltage oscillating with a frequency $\hbar\omega=\Delta_1+\Delta_2$. The divergence of I_1 is called the Riedel peak.[13] This divergence of I_1 is associated with a corresponding singularity in the density of states of the superconductors at $\hbar\omega=\Delta$.

In the case of anisotropic pairing, in which the density of states does not exhibit divergence, the Riedel peak is, of course, smeared out. Nevertheless, as is shown below, some new characteristic singularities appear on the current–voltage characteristic under certain conditions specifically because of the anisotropy of the order parameter, some of them caused by reversal of the sign of the order parameter at the Fermi surface. The characteristic behavior of the current–voltage characteristic can depend significantly on the orientation of the crystal axes of the superconductors relative to the junction plane. Their orientation determines the spatial behavior of the order parameter near that plane. However, even in the simplest case of a uniform distribution of the superconducting order parameter on the two sides of the junction, the features of the current–voltage characteristic differ sharply from the case of isotropic pairing. In fact, for superconductors with anisotropic singlet pairing the functions $I_m(V)$ also become dependent on the momentum direction on the Fermi surface, and to find the tunneling current these functions must be integrated with some weight over the Fermi surface. For orientations of the crystal axes in which there is no suppression of the order parameter on either side of the junction, we found that nonanalytic behavior of the current–voltage characteristic is observed only for the directions of $|eV|$ which are equal to the values of the expressions $\| \Delta_2(\hat{\mathbf{p}}_2)| \pm |\Delta_1(\hat{\mathbf{p}}_1)\|$ at their extremum points. Here $\hat{\mathbf{p}}_1$ is the momentum direction of a quasiparticle impinging on the boundary of the tunnel barrier, and $\hat{\mathbf{p}}_2$ is the momentum direction of a transmitted quasiparticle, which is directly related to $\hat{\mathbf{p}}_1$ and to the form of the Fermi surfaces of the contacting metals. The expressions presented have been examined, for example, as a function of $\hat{\mathbf{p}}_1$. It was found that the singularities of the current–voltage characteristic depend strongly on the types of extrema for these expressions. This was first pointed out in Ref. 7, which treated the characteristic behavior of the current–voltage characteristic for the quasiparticle current at low temperatures in the critical orientations, in which there is no suppression of the order parameter on either side of the junction.

The situation becomes more complicated for crystal orientations in which the order parameter is significantly suppressed near the junction plane. In this case quasiparticle bound states localized near the junction plane can occur. Among them we can distinguish the bound state with zero energy, which appears under quite general conditions.[8,14–17] In the case of an opaque, specularly reflecting boundary, such a state always appears for quasiparticles whose incident and reflected momenta correspond to different signs of the order parameter in the bulk of the superconductor. Besides such universal states localized near a boundary, quasiparticle bound states with a nonzero energy, whose value is specified by the specific profile of the variation of the order parameter near the surface, can also be present.[8] Below we shall consider both types of bound states and their manifestations on the current–voltage characteristic.

2. MICROSCOPIC EXPRESSION FOR THE TUNNELING ELECTRIC CURRENT

Let us consider a tunnel junction with a transparency factor $D \ll 1$ and with a specularly reflecting plane at the potential barrier between two clean superconductors having singlet pairing. We assume that the external voltage $V(t)=\Phi_2(t)-\Phi_1(t)$ is applied to the junction. We align the x axis parallel to a normal to the junction plane, $\mathbf{n}\|\mathbf{x}$. To first order in the transparency D the microscopic expression for the tunneling current density can be written in the following form:[18]

$$j_x=-\frac{1}{8\pi^3}\int_{v_{x1}>0}\frac{d^2S_1 v_{x1}}{4\pi v_{f1}}\,\mathrm{Tr}\{\hat{\tau}_3 D(\hat{\mathbf{p}}_1)\times(\hat{g}_1^R\hat{g}_2^K+\hat{g}_1^K\hat{g}_2^A-\hat{g}_2^R\hat{g}_1^K-\hat{g}_2^K\hat{g}_1^A)_0(t,t)\}. \tag{1}$$

Here and below we set e, $\hbar$, $c=1$, where $e=-|e|$ is the charge of the electron.

In (1) the retarded, advanced, and Keldysh quasiclassical matrix propagators are taken directly on the junction boundary and should be calculated in zero order with respect to the transparency of the barrier (i.e., for a nontransparent boundary). These propagators depend on the corresponding directions $\hat{\mathbf{p}}_1$ and $\hat{\mathbf{p}}_2$ in momentum space. The subscript 1 (2) labels the left-hand (right-hand) half-space with respect to the boundary plane, v_x is the component of the Fermi velocity along the normal $\mathbf{n}$ to the junction plane. In (1) the integration is performed over the part of the Fermi surface for which $v_x>0$. The relation between the incident and transmitted Fermi momenta (i.e., between $\mathbf{p}_1$ and $\mathbf{p}_2$) is as follows. The components which are parallel to the specularly reflecting barrier plane should be equal to one another, and the magnitudes of the components along the normal to the boundary are determined from the condition that the momenta $\mathbf{p}_1$ and $\mathbf{p}_2$ lie on the corresponding Fermi surfaces. It is obvious that in the special case of superconductors with identical spherical or cylindrical Fermi surfaces (provided the cylindrical axis is parallel to the junction plane) the incident and transmitted momenta are equal, $\mathbf{p}_1=\mathbf{p}_2$.

The quantities used in (1) are found in accordance with the following example:

$$(\hat{g}_1^R\hat{g}_2^K)_0(t,t)=\int_{-\infty}^{t}dt_1\hat{g}_1^R(t,t_1)\hat{g}_2^K(t_1,t). \tag{2}$$

Here $\hat{g}_l^{R,A,K}(t,t_1)=\hat{S}_l(t)\hat{g}_l^{R,A,K}(t-t_1)\hat{S}_l^+(t_1)$ $(l=1,\,2)$, where

$$\hat{S}_l(t)=\begin{pmatrix}\exp[i\chi_l(t)/2] & 0\\ 0 & \exp[-i\chi_l(t)/2]\end{pmatrix},$$

$$\chi_l(t)=\chi_l-2\int^{t}\Phi_l(t')dt' \tag{3}$$

and χ_l is the phase of the order parameter of the lth superconductor in the junction plane when the electric potential Φ_l is zero.

Below we consider only singlet types of anisotropic pairing. The matrix propagators for them can be written in the form

$$\hat{g}=\begin{pmatrix} g & f\\ f^+ & -g\end{pmatrix}. \tag{4}$$

For the further calculations we note that the nonequilibrium effects are generally insignificant for tunnel junctions. The voltage V only shifts the Fermi levels of the electrodes relative to one another by an amount V. In addition, to find the current in first order with respect to the transparency, the Green's functions can be calculated for nontransparent half-spaces (if we neglect the tunneling of electrons through the junction). Under the conditions considered here the electron distribution functions still correspond to equilibrium, and the electric potential leads only to the appearance of the corresponding coordinate-independent terms in the phases of the superconducting order parameters and in the Green's functions on the two sides of the junction [see, for example, (3)]. The following relation then holds:

$$\hat{g}^K(\omega)=[\hat{g}^R(\omega)-\hat{g}^A(\omega)]\tanh\left(\frac{\omega}{2T}\right). \tag{5}$$

According to the general symmetry properties of propagators, $\hat{g}^A=\hat{\tau}_3(\hat{g}^R)^\dagger\hat{\tau}_3$, where $\hat{\tau}_3$ is a Pauli matrix. Let us consider the case in which the phase of the order parameter within the half-space occupied by a superconductor with an opaque boundary does not depend on the coordinates or the momentum. Then this phase corresponds to χ_l in Eq. (3), and the Green's functions should be calculated for the real order parameter, although the sign of the order parameter can depend on the momentum direction. Under these conditions we have

$$\begin{aligned}
&f^{R,A}(-\hat{\mathbf{p}},x,\omega)=-f^{+R,A}(\hat{\mathbf{p}},x,\omega),\\
&g^{R,A}(-\hat{\mathbf{p}},x,\omega)=g^{R,A}(\hat{\mathbf{p}},x,\omega),\\
&f^{R,A}(\hat{\mathbf{p}},x,-\omega)=f^{R,A*}(\hat{\mathbf{p}},x,\omega),\\
&g^{R,A}(\hat{\mathbf{p}},x,-\omega)=-g^{R,A*}(\hat{\mathbf{p}},x,\omega).
\end{aligned}$$

We also take into account in the calculations that the values of the propagator from the incident momentum $\hat{\mathbf{p}}$ and from the reflected momentum $\check{\mathbf{p}}$ coincide and $v_x(\hat{\mathbf{p}})=-v_x(\check{\mathbf{p}})$ on an opaque boundary. Therefore, when the variable voltage $V(t)=V_0+a\cos(\omega_0 t)$ is applied to the junction, we can obtain the following expression for the tunneling current from Eq. (1):

$$\begin{aligned}
j_x=\sum_{n=-\infty}^{\infty} J_n\left(\frac{a}{\omega_0}\right)\Bigg\{ & j_1(V_0+\omega_n)\\
&\times\sin\Bigg[\chi_1-\chi_2+2V_0t+\frac{a}{\omega_0}\\
&\times\sin(\omega_0 t)+\omega_n t\Bigg]+j_2(V_0+\omega_n)\\
&\times\cos\left[\chi_1-\chi_2+2V_0t+\frac{a}{\omega_0}\sin(\omega_0 t)+\omega_n t\right]\\
&+j_3(V_0+\omega_n)\cos\left[\omega_n t-\frac{a}{\omega_0}\sin(\omega_0 t)\right]\\
&+j_4(V_0+\omega_n)\sin\left[\omega_n t-\frac{a}{\omega_0}\sin(\omega_0 t)\right]\Bigg\}.
\end{aligned} \tag{6}$$

Here we have introduced the quantities:

$$\begin{aligned}
&j_m(V)=\int_{v_{x1}>0}\frac{d^2S_1}{(2\pi)^3}\,\frac{v_{x1}}{v_{f1}}\,D(\hat{\mathbf{p}}_1)I_m(V,\hat{\mathbf{p}}_1),\\
&m=1,\ 2,\ 3,\ 4,
\end{aligned} \tag{7}$$

$$\begin{aligned}
I_1(V,\hat{\mathbf{p}}_1)=\int_{-\infty}^{\infty}\frac{d\omega}{2\pi^2}\tanh\left(\frac{\omega}{2T}\right)\ \mathrm{Im}\{ & f_1^R(\omega-V)[f_2^{+R}(\omega)\\
&+f_2^{R}{}^*(\omega)]+(1\leftrightarrow 2,V\rightarrow -V)\},
\end{aligned} \tag{8}$$

$$\begin{aligned}
I_2(V,\hat{\mathbf{p}}_1)=-\int_{-\infty}^{\infty}\frac{d\omega}{2\pi^2}\tanh\left(\frac{\omega}{2T}\right)\mathrm{Re}\{ & f_1^R(\omega-V)\\
&\times[f_2^{+R}(\omega)+f_2^{R}{}^*(\omega)]-(1\leftrightarrow 2,V\rightarrow -V)\},
\end{aligned} \tag{9}$$

$$\begin{aligned}
I_3(V,\hat{\mathbf{p}}_1)=\int_{-\infty}^{\infty}\frac{d\omega}{\pi^2}&\left[\tanh\left(\frac{\omega-V}{2T}\right)-\tanh\left(\frac{\omega}{2T}\right)\right]\\
&\times\mathrm{Im}\,g_1^R(\omega)\mathrm{Im}\ g_2^R(\omega-V),
\end{aligned} \tag{10}$$

$$\begin{aligned}
I_4(V,\hat{\mathbf{p}}_1)=-\int_{-\infty}^{\infty}\frac{d\omega}{\pi^2}\tanh\left(\frac{\omega}{2T}\right)\{ & \mathrm{Re}[g_1^R(\omega-V)]\mathrm{Im}\ g_2^R\\
&\times(\omega)+\mathrm{Re}[g_2^R(\omega-V)]\mathrm{Im}\ g_1^R(\omega)\}.
\end{aligned} \tag{11}$$

The replacement $f\leftrightarrow f^+$ in formulas (8) and (9) leads to the same result for the Josephson current after integration over the momentum directions in (7).

For the special case of superconductors with isotropic pairing Eqs. (6)–(11) agree with the results of Larkin and Ovchinnikov.[1] We note that $I_1(V)$ and $I_4(V)$ are even functions of V, while I_2 and I_3 are odd. This difference enables us, in principle, to unequivocally isolate the behavior of all four contributions to the tunneling current from the results of measurements and to study it on an individual basis. Then the parity properties with respect to V just mentioned permit elimination of the replacement $V\rightarrow -V$ in the complete expressions for $I_{1,2}(V)$, which leads only to reversal of the sign in front of the corresponding term in the expression for $I_2(V)$.

Equations (8)–(11) were written out above in a form which is conserved in the more general case (compared with the case considered below) in which the phase of the complex order parameter depends on the coordinates and the momentum directions and the Green's functions satisfy the general symmetry relations

$$\begin{aligned}
&f^{+R,A}(-\hat{\mathbf{p}},x,-\omega)=-f^{R,A*}(\hat{\mathbf{p}},x,\omega),\\
&g^{R,A}(-\hat{\mathbf{p}},x,-\omega)=-g^{R,A*}(\hat{\mathbf{p}},x,\omega).
\end{aligned}$$

Here, of course, a different choice of the phase difference $\chi_1-\chi_2$ can appear in the arguments of the trigonometric functions with a corresponding change in the definition of the amplitudes $j_{1,2}$. If this phase difference is chosen dependent on the momentum direction, those trigonometric functions must appear together with $I_{1,2}(V)$ in the integral over the Fermi surface. In the general case $I_1(V)$ and $I_2(V)$ do not have the properties of evenness or oddness with respect to the direction.

For a constant voltage on the junction, setting $a=0$, from (6) we obtain

$$\begin{aligned}
j_x(V)=j_1(V)\sin(\chi_1-\chi_2+2Vt)+j_2(V)\cos(&\chi_1-\chi_2\\
&+2Vt)+j_3(V).
\end{aligned} \tag{12}$$

It follows from (12) and (6) that the characteristic singularities of the current–voltage characteristic for constant and variable voltages on the junction are given by the singular points of the functions $j_m(V)$. We use V_m to denote the singular points of these functions. Then the current–voltage characteristic for a constant voltage on the junction will have singularities only at these points. However, in the case of a variable voltage the expression for the current (6) has singularities at the voltage values $V=V_m-n\omega_0$ ($m=1, 2, 3, 4$; $n=0,\pm 1,\pm 2,\ldots$). For sufficiently large values of n ($a/\omega_0 \lesssim n$) the amplitude of the current becomes small (because of the corresponding behavior of the Bessel functions).

3. QUASICLASSICAL GREEN'S FUNCTIONS ON AN OPAQUE FLAT BOUNDARY

It follows from the preceding section that to calculate the tunneling electric current through a junction of low transparency it is sufficient to find the values of the retarded electron propagators on a flat opaque boundary of a half-space occupied by a superconductor. It is apparently impossible to find the complete analytic solution of this problem for a pairing potential that leads to an anisotropic order parameter (except for a few special orientations of the crystal axes, for which there is no suppression of the order parameter near the boundary and it does not depend on the coordinates), and a numerical calculation is required for this purpose. However, the problem is simplified significantly, if we are interested only in the singularities of the current–voltage characteristic, since finding them requires knowledge of the behavior of the Green's functions (taken on the boundary) only near the singular points. To solve this problem we use the Eilenberger equation for delayed quasiclassical propagators, which can be written for a superconductor with singlet pairing in the following manner (below we omit the subscripts on the retarded propagators):

$$(2\omega+iv_x\partial_x)f(\hat{\mathbf{p}},x,\omega)+2\Delta(\hat{\mathbf{p}},x)g(\hat{\mathbf{p}},x,\omega)=0,$$

$$(2\omega-iv_x\partial_x)f^{+}(\hat{\mathbf{p}},x,\omega)-2\Delta^{*}(\hat{\mathbf{p}},x)g(\hat{\mathbf{p}},x,\omega)=0, \qquad (13)$$

$$iv_x\partial_x g(\hat{\mathbf{p}},x,\omega)-\Delta(\hat{\mathbf{p}},x)f^{+}(\hat{\mathbf{p}},x,\omega)-\Delta^{*}(\hat{\mathbf{p}},x)f(\hat{\mathbf{p}},x,\omega)=0.$$

To be specific, we assume that the superconductor occupies the half-space $x>0$.

Apart from the self-consistency equation for $\Delta(\hat{\mathbf{p}},x)$, Eqs. (13) should be supplemented by the normalization condition

$$g^2+ff^{+}=-\pi^2 \qquad (14)$$

and the boundary conditions for the quasiclassical propagators. For an opaque, specularly reflecting surface we have

$$g(\hat{\mathbf{p}},\omega)=g(\check{\mathbf{p}},\omega)|_{x=0}, \quad f(\hat{\mathbf{p}},\omega)=f(\check{\mathbf{p}},\omega)|_{x=0},$$

$$f^{+}(\hat{\mathbf{p}},\omega)=f^{+}(\check{\mathbf{p}},\omega)|_{x=0}. \qquad (15)$$

Here $\hat{\mathbf{p}}$ is the incident momentum direction, and $\check{\mathbf{p}}$ is the reflected momentum direction.

An additional boundary (asymptotic) condition is imposed on Eqs. (13) by the behavior of the propagators in the interior of the superconductor:

$$g(\hat{\mathbf{p}},\omega)|_{x=\infty}=\frac{-\pi\omega}{\sqrt{|\Delta_\infty(\hat{\mathbf{p}})|^2-\omega^2}},$$

$$f(\hat{\mathbf{p}},\omega)|_{x=\infty}=-f^{+}(\hat{\mathbf{p}},\omega)|_{x=\infty}=\frac{\pi\Delta_\infty(\hat{\mathbf{p}})}{\sqrt{|\Delta_\infty(\hat{\mathbf{p}})|^2-\omega^2}}. \qquad (16)$$

We assume that we can choose the order parameter Δ real within a superconducting half-space with an opaque boundary (i.e., when no current passes through the boundary). We introduce the following functions:

$$f_1=\frac{1}{2}\,(f-f^{+}), \quad f_2=\frac{1}{2}\,(f+f^{+}). \qquad (17)$$

For f_1 and f_2 Eqs. (13) and (14) take on the forms

$$2\omega f_1+iv_x\partial_x f_2+2\Delta g=0,$$

$$f_2=-i\,\frac{v_x}{2\omega}\,\partial_x f_1, \qquad (18)$$

$$\partial_x g=-i\,\frac{2\Delta}{v_x}\,f_2,$$

$$g^2+f_2^2-f_1^2=-\pi^2. \qquad (19)$$

The boundary conditions at $x=0$ for $f_{1,2}$ are the same as (15), but in the bulk of the superconductor we have

$$f_1(\hat{\mathbf{p}},\omega)|_{x=\infty}=\frac{\pi\Delta_\infty(\hat{\mathbf{p}})}{\sqrt{|\Delta_\infty(\hat{\mathbf{p}})|^2-\omega^2}}, \quad f_2|_{x=\infty}=0. \qquad (20)$$

A representation of the functions g and f taken on the boundary, which is convenient for considering the singular parts of the propagators, can be obtained from these equations. To obtain this representation we introduce the following function:

$$\tilde{\Delta}(\hat{\mathbf{p}},\omega)=\frac{\int_0^\infty \Delta(\hat{\mathbf{p}},x)f_2(\hat{\mathbf{p}},x,\omega)dx}{\int_0^\infty f_2(\hat{\mathbf{p}},x,\omega)dx}. \qquad (21)$$

From the second and third equations in (18) we can easily obtain

$$g(\infty)-g(0)=-\frac{2i}{v_x}\int_0^\infty \Delta(\hat{\mathbf{p}},x)f_2(\hat{\mathbf{p}},x,\omega)dx,$$

$$f_1(\infty)-f_1(0)=\frac{2i\omega}{v_x}\int_0^\infty f_2(\hat{\mathbf{p}},x,\omega)dx. \qquad (22)$$

With consideration of this function, $\tilde{\Delta}(\hat{\mathbf{p}},\omega)$ can also be rewritten in the form

$$\tilde{\Delta}(\hat{\mathbf{p}},\omega)=-\omega\,\frac{g(\infty)-g(0)}{f_1(\infty)-f_1(0)}. \qquad (23)$$

Here we introduce the notation $g|_{x=0}=g(0)$, $g|_{x=\infty}=g(\infty)$, $f|_{x=0}=f(0)$, and $f|_{x=\infty}=f(\infty)$.

After substituting the expressions (16) and (20) for the propagators in the bulk of the superconductor into Eq. (23), we arrive at the following relation between $g(0)$, $f_1(0)$, and $\tilde{\Delta}(\hat{\mathbf{p}},\omega)$:

$$g(0)=\frac{\pi}{\omega}\,\frac{\tilde{\Delta}(\hat{\mathbf{p}},\omega)\Delta_\infty(\hat{\mathbf{p}})-\omega^2}{\sqrt{|\Delta_\infty(\hat{\mathbf{p}})|^2-\omega^2}}-\frac{\tilde{\Delta}(\hat{\mathbf{p}},\omega)}{\omega}\,f_1(0). \qquad (24)$$

A completely analogous relation can be written for the momentum direction $\check{\mathbf{p}}$. Thereafter, using the boundary conditions (15), we can write the following representation for $g(0)$ and $f_1(0)$:

$$g(0)=-\frac{\pi}{\omega}\,\frac{\tilde{\Delta}(\hat{\mathbf{p}},\omega)\tilde{\Delta}(\check{\mathbf{p}},\omega)}{\tilde{\Delta}(\check{\mathbf{p}},\omega)-\tilde{\Delta}(\hat{\mathbf{p}},\omega)}\left[\frac{\Delta_\infty(\hat{\mathbf{p}})\tilde{\Delta}(\hat{\mathbf{p}},\omega)-\omega^2}{\tilde{\Delta}(\hat{\mathbf{p}},\omega)\sqrt{|\Delta_\infty(\hat{\mathbf{p}})|^2-\omega^2}}-\frac{\Delta_\infty(\check{\mathbf{p}})\tilde{\Delta}(\check{\mathbf{p}},\omega)-\omega^2}{\tilde{\Delta}(\check{\mathbf{p}},\omega)\sqrt{|\Delta_\infty(\check{\mathbf{p}})|^2-\omega^2}}\right], \qquad (25)$$

$$f_1(0)=\frac{\pi}{\tilde{\Delta}(\hat{\mathbf{p}},\omega)-\tilde{\Delta}(\check{\mathbf{p}},\omega)}\left[\frac{\Delta_\infty(\hat{\mathbf{p}})\tilde{\Delta}(\hat{\mathbf{p}},\omega)-\omega^2}{\sqrt{|\Delta_\infty(\hat{\mathbf{p}})|^2-\omega^2}}-\frac{\Delta_\infty(\check{\mathbf{p}})\tilde{\Delta}(\check{\mathbf{p}},\omega)-\omega^2}{\sqrt{|\Delta_\infty(\check{\mathbf{p}})|^2-\omega^2}}\right]. \qquad (26)$$

It is seen from Eq. (25) that the candidates for the singular points of the propagator $g(0)$ are $\omega=0$, $\pm|\Delta_\infty(\hat{\mathbf{p}})|$, $\pm|\ \Delta_\infty(\check{\mathbf{p}})|$, and, generally speaking, the singular points of the functions $\tilde{\Delta}(\hat{\mathbf{p}},\omega)$ and $\tilde{\Delta}(\check{\mathbf{p}},\omega)$. Similarly, it is seen from Eq. (26) that the candidates for the singular points of $f_1(0)$ are $\omega=\pm|\Delta_\infty(\hat{\mathbf{p}})|$, $\pm|\Delta_\infty(\check{\mathbf{p}})|$. We note that if we are interested in the points for which $\tilde{\Delta}(\hat{\mathbf{p}},\omega)=\tilde{\Delta}(\check{\mathbf{p}},\omega)$ holds, the treatment of Eq. (24) for the momentum direction $\check{\mathbf{p}}$ yields no independent relation other than Eq. (24) for the momentum direction $\hat{\mathbf{p}}$. Thus, additional information is needed to consider this limiting case in Eqs. (25) and (26).

Let us first consider the singular parts of the propagators g_s, f_s, and f_s^+ taken on the boundary in the vicinity of $\omega=0$. It follows from (25) and (26) that $f_1(0)$ does not have a singularity at $\omega=0$, while $g(0)$ has poles at that point (provided $\Delta_\infty(\hat{\mathbf{p}})$ and $\Delta_\infty(\check{\mathbf{p}})$ have opposite signs):

$$g_s(0)=\frac{1}{\omega}\,\frac{\pi\tilde{\Delta}(\hat{\mathbf{p}},0)\tilde{\Delta}(\check{\mathbf{p}},0)}{\tilde{\Delta}(\check{\mathbf{p}},0)-\tilde{\Delta}(\hat{\mathbf{p}},0)}\,\{\mathrm{sgn}[\Delta_\infty(\hat{\mathbf{p}})]-\mathrm{sgn}[\Delta_\infty(\check{\mathbf{p}})]\}=\frac{B_g(\hat{\mathbf{p}})}{\omega}. \qquad (27)$$

It is noteworthy that besides this relation we can also find the explicit expression for $\tilde{\Delta}(\hat{\mathbf{p}},0)$ in terms of the nonuniform distribution of the order parameter. In fact, since f_1 does not have a singularity at $\omega=0$ [see (18)], the relation $g_s=\pm i f_{2,s}$ (for $\omega\to 0$) between the singular parts of the functions g and f_2 can be obtained from the normalization condition (19). Then, from the last equation of system (18) we obtain the following equation for the singular part of f_2:

$$\partial_x f_{2,s}\pm\frac{2\Delta}{v_x}\,f_{2,s}=0.$$

The solution of this equation which satisfies the boundary conditions (20) and (15) is the function

$$f_{2,s}(\hat{\mathbf{p}},x,\omega)=f_{2,s}(\hat{\mathbf{p}},0,\omega)\times\exp\left\{-\frac{2\,\mathrm{sgn}[\Delta_\infty(\hat{\mathbf{p}})]}{|v_x|}\int_0^x\Delta(\hat{\mathbf{p}},x')dx'\right\}. \qquad (28)$$

Substituting this solution into (21), we obtain

$$\tilde{\Delta}(\hat{\mathbf{p}},0)=\frac{1}{2}\,|v_x|\mathrm{sgn}[\Delta_\infty(\hat{\mathbf{p}})]\times\left[\int_0^\infty\exp\left\{-\frac{2\,\mathrm{sgn}[\Delta_\infty(\hat{\mathbf{p}})]}{|v_x|}\times\int_0^x\Delta(\hat{\mathbf{p}},x')dx'\right\}dx\right]^{-1}. \qquad (29)$$

We note that determination of the sign in (28) from the asymptotic condition in the interior of the superconductor makes it possible to fix the sign in the relation between the singular parts of the quasiclassical propagators taken near the boundary for $\omega\to 0$:

$$f_s(\hat{\mathbf{p}},x,\omega)=f_s^+(\hat{\mathbf{p}},x,\omega)=-i\ \mathrm{sgn}[v_x\Delta_\infty(\hat{\mathbf{p}})]g_s(\hat{\mathbf{p}},x,\omega). \qquad (30)$$

Thus, Eqs. (27), (29), and (30) provide a fairly general description of the singular parts of the propagators taken on the boundary in the vicinity of $\omega=0$. If we are interested only in the singularities of the current–voltage characteristic, the problem of solving the self-consistency equation for a specific pairing potential can be considered separately. The latter problem is an important part of the complete theoretical description of the current–voltage characteristic and can be solved by numerical methods.

Let us now consider the behavior of the propagators near the points $\omega=\pm|\Delta_\infty(\hat{\mathbf{p}})|$, $\pm|\Delta_\infty(\check{\mathbf{p}}|$. For this purpose, we first find the values of $\tilde{\Delta}(\hat{\mathbf{p}},\omega)$ at these frequencies. This can be done, for example, using the relation (21). In fact, Eqs. (18)–(20) give the following asymptotic behavior of $f_2(\hat{\mathbf{p}},\omega)$ as $x\to\infty$:

$$f_2(\hat{\mathbf{p}},x,\omega)\propto\exp\left(-\frac{2\sqrt{|\Delta_\infty(\hat{\mathbf{p}})|^2-\omega^2}}{|v_x|}x\right). \qquad (31)$$

Hence it is seen from (21) that in the limit $|\omega|\to|\Delta_\infty(\hat{\mathbf{p}})|$ the main contribution to the integrals in (21) is made by the region within the superconductor where the order parameter $\Delta(\hat{\mathbf{p}},x)$ is equal to its value in the interior. Taking this into account, we obtain the following relation:

$$\tilde{\Delta}(\hat{\mathbf{p}},\omega)\to\Delta_\infty(\hat{\mathbf{p}})\quad\text{for }|\omega|\to|\Delta_\infty(\hat{\mathbf{p}})|. \qquad (32)$$

Now we can examine the behavior of the functions g and f_1 in the vicinity of the points $|\omega|=|\Delta_\infty(\hat{\mathbf{p}})|$, $|\Delta_\infty(\check{\mathbf{p}})|$. Going over to the limit $|\omega|\to|\Delta_\infty(\hat{\mathbf{p}})|$ or $|\omega|\to|\Delta_\infty(\check{\mathbf{p}})|$ in Eqs. (25) and (26) and taking into account (32), we see that the divergences are reduced and that $g(0)$ and $f_1(0)$ [and, therefore, $f(0)$] have a square-root nonanalyticity of the form $\sqrt{\Delta_\infty^2(\hat{\mathbf{p}})-\omega^2}$ (or $\sqrt{\Delta_\infty^2(\check{\mathbf{p}})-\omega^2}$) at these points in first order. The orientations for which $\tilde{\Delta}(\hat{\mathbf{p}})=\tilde{\Delta}(\check{\mathbf{p}})$ holds, in which, strictly speaking, Eqs. (25) and (26) do not provide more information than (24), can be exceptions. As is generally known, if this condition holds for all the momentum directions, then there is no suppression of the order parameter, and in this special case the propagators diverge in the limit $|\omega|\rightarrow|\Delta_\infty(\hat{\mathbf{p}})|$:

$$g^R(\hat{\mathbf{p}},\omega)=-\frac{\pi\omega}{\sqrt{|\Delta(\hat{\mathbf{p}})|^2-\omega^2}},$$

$$f^R(\hat{\mathbf{p}},\omega)=\frac{\pi\Delta(\hat{\mathbf{p}})}{\sqrt{|\Delta(\hat{\mathbf{p}})|^2-\omega^2}}. \tag{33}$$

Finally, the possible existence of quasiparticle bound states with nonzero energy, localized near the boundary, should be taken into account. They can appear, for example, because of the spatial variation of the order parameter, which is suppressed near the boundary. Such a bound state can be interpreted as a bound state in the ''potential well'' formed by the order parameter.[8] Since such a bound state corresponds to a pole in the quasiclassical propagators, we can simply add the pole term to the singular parts of the propagators. If the propagators $g(0)$ and $f_1(0)$ have the same poles describing bound states with a nonzero energy, it follows from Eq. (23) and the boundary conditions (15) that $\widetilde{\Delta}(\hat{\mathbf{p}},\omega)=\widetilde{\Delta}(\check{\mathbf{p}},\omega)$ holds at the point of the pole.

With consideration of the results obtained above, the nonanalytic parts of $g(0)$ and $f(0)$ can be written in the following form:

$$\begin{aligned}g_s^R(\hat{\mathbf{p}},\omega)|_{x=0}=&\frac{B_g(\hat{\mathbf{p}})}{\omega+i\delta}+\frac{Q_g(\hat{\mathbf{p}})}{\omega-h(\hat{\mathbf{p}})\mathrm{sgn}(\omega)+i\delta}\\&+C(\hat{\mathbf{p}},\omega)\mathrm{sgn}(\omega)\sqrt{\Delta_\infty^2(\hat{\mathbf{p}})-\omega^2}\\&+C(\check{\mathbf{p}},\omega)\mathrm{sgn}(\omega)\sqrt{\Delta_\infty^2(\check{\mathbf{p}})-\omega^2}+\ \dots,\end{aligned} \tag{34}$$

$$\begin{aligned}f_s^R(\hat{\mathbf{p}},\omega)|_{x=0}=&\frac{iB_f(\hat{\mathbf{p}})}{\omega+i\delta}+\frac{iQ_f(\hat{\mathbf{p}})}{\omega-h(\hat{\mathbf{p}})\mathrm{sgn}(\omega)+i\delta}\\&+E(\hat{\mathbf{p}},\omega)\sqrt{\Delta_\infty^2(\hat{\mathbf{p}})-\omega^2}\\&+E(\check{\mathbf{p}},\omega)\sqrt{\Delta_\infty^2(\check{\mathbf{p}})-\omega^2}+\ \dots,\\ \delta\rightarrow&+0.\end{aligned} \tag{35}$$

In addition, the following relations hold:

$$B_f(\hat{\mathbf{p}})=-\mathrm{sgn}[v_x\Delta_\infty(\hat{\mathbf{p}})]B_g(\hat{\mathbf{p}})=-B_f(-\hat{\mathbf{p}}),$$

$$Q_g(\hat{\mathbf{p}})=|Q_f(\hat{\mathbf{p}})|,$$

$$Q_f(\hat{\mathbf{p}})=-Q_f^*(-\hat{\mathbf{p}}),\quad h(-\hat{\mathbf{p}})=h(\hat{\mathbf{p}}).$$

Below we shall be interested only in the values of the functions $Q_{g,f}(\hat{\mathbf{p}})$ near the poles $\omega=\pm h(\hat{\mathbf{p}})$. When the positive and negative poles are considered, it should be taken into account that $Q_g(\hat{\mathbf{p}},\omega)$ and $\mathrm{Re}\,Q_f(\hat{\mathbf{p}},\omega)$ are even functions and that $\mathrm{Im}\,Q_f(\hat{\mathbf{p}},\omega)$ is an odd function with respect to ω. Henceforth we shall not consider the dependence of $Q_{g,f}$ on ω, and we shall assume that these functions are evaluated at the positive pole.

Since there are no states with a continuous spectrum for an assigned momentum direction $\hat{\mathbf{p}}$ when $|\omega|<\min|\Delta_\infty(\check{\mathbf{p}})|,|\Delta_\infty(\hat{\mathbf{p}})|)$ holds, it is natural to impose the condition $\mathrm{Im}\,C$, $\mathrm{Im}\,Q_g=0$. In addition, for the particular orientation in which the equality $\Delta(\check{\mathbf{p}},x)=-\Delta(\hat{\mathbf{p}},x)$ holds for all x, we have $\Delta(x=0)=0$ and $f_1(0)=0$ on the superconductor boundary. Then it can be found for this orientation from the Eilenberger equations that $\mathrm{Re}\,f(0)=0$ will hold in the frequency range $|\omega|<|\Delta_\infty(\hat{\mathbf{p}})|$. Therefore, $\mathrm{Re}\,E(\hat{\mathbf{p}},\omega)+E(\check{\mathbf{p}},\omega)]=0$ holds in this frequency range.

As is shown below, the positions and types of singularities on the current–voltage characteristics of tunnel junctions, which are governed by the contribution from the quasiparticle bound states, are determined, in particular, by the extremum and nonanalytic points of the function $h(\hat{\mathbf{p}})$. The boundary conditions for propagators in the case of singlet pairing imply $h(\hat{\mathbf{p}})=h(-\check{\mathbf{p}})$. Since we also have $\hat{\mathbf{p}}=-\check{\mathbf{p}}$ for the momentum direction $\hat{\mathbf{p}}\|\mathbf{n}$, the function $h(\hat{\mathbf{p}})$ (as well as the complete propagator $g(0)$) should have an extremum in the direction in which the vector $\hat{\mathbf{p}}$ is parallel to the normal to the boundary $\mathbf{n}$ (provided there is a bound state for this value of the momentum). Similarly, we find that the function $Q_f(\hat{\mathbf{p}})$ is purely imaginary for $\hat{\mathbf{p}}\|\mathbf{n}$. In particular, it follows that for orientations of the crystal axis in which the relationship $\Delta(\check{\mathbf{p}},x)=-\Delta(\hat{\mathbf{p}},x)$ is satisfied for all x, the equality $Q_{f,g}(\hat{\mathbf{p}})=0$ will hold for $\hat{\mathbf{p}}\|\mathbf{n}$. The other characteristic points for the function $h(\hat{\mathbf{p}})$ are the momentum directions for which the quasiparticle bound states vanish near the boundary.

Figure 1 presents the results of numerical calculations of $B_g(\phi)$, $Q_g(\phi)=|Q_f(\phi)|$, $\mathrm{Re}\,Q_f(\phi)$, and $h(\phi)$. We considered a tetragonal superconductor with a cylindrical Fermi surface and an opaque, specularly reflecting flat boundary located at $x=0$ (the cylindrical z axis is parallel to the boundary plane). The pairing potential is taken in the form

$$V(\phi,\phi')=2V_d\cos(2\phi-2\phi_0)\cos(2\phi'-2\phi_0),$$

which leads to an order parameter with $d_{x_0^2-y_0^2}$ symmetry. Here ϕ is the azimuthal angle in the xy plane, measured from the direction of a normal to the boundary. The angle ϕ_0 describes the orientation of the x_0 crystal axis relative to a normal to the boundary. For the pairing potential that we selected the order parameter has the form

$$\Delta(\phi,x)=\Delta(x)\cos(2\phi-2\phi_0),$$

where $\Delta(x)$ is calculated using the self-consistency equation. The angle ϕ assigns the incident momentum direction along the quasiparticle trajectory. In the numerical calculation we took $\phi_0=\pi/9$, $T=0.45T_c$, and $\Delta_0/2T=2$, where $\Delta_0\equiv\Delta(x=\infty)$ and T_c is the critical temperature. For the orientation of the crystal axes that we selected ($\phi_0=\pi/9$) the order parameter is suppressed near the boundary to the value $\Delta(x=0)=0.28\Delta_0$. In Fig. 1 the functions $B_g(\phi)$, $Q_g(\phi)$, and $\mathrm{Re}\,Q_f(\phi)$ were normalized to $\pi\Delta_0$, while $h(\phi)$ was normalized to Δ_0. It is seen from the figure that bound states with a nonzero energy $\pm h(\phi)$ exist in the narrow range of angles $\phi\in(-0.095,0.095)$ in the vicinity of the normal to the boundary. States with zero energy exist in another, broader range of angles, in which $B_g(\phi)\neq 0$. In addition, the value of $h(\phi)$ at the maximum point is $h_m=0.7\Delta_0$, and the value of $h(\phi_{ed})$ at the cut-off points (where the bound state vanishes) is $h_{ed}=0.63\Delta_0$. We note that for the momentum

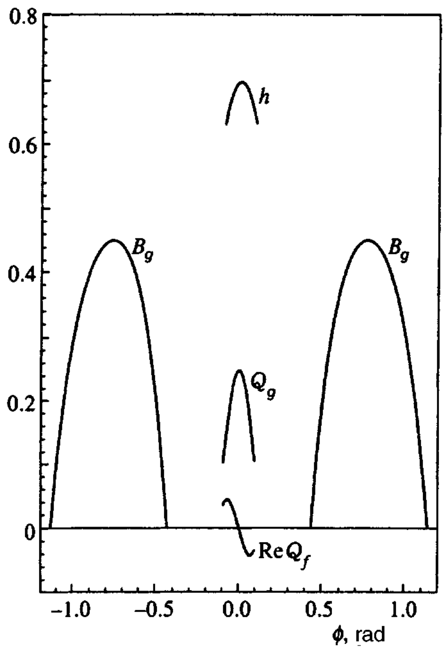


FIG. 1.

directions $\hat{\mathbf p}$ ($\phi_{ed}=\pm 0.095$), at which the bound state with a nonzero energy vanishes, $h_{ed}=\min(|\Delta_\infty(\check{\mathbf p})|,|\Delta_\infty(\hat{\mathbf p})|)$ is satisfied. This means that a discrete level splits off from the continuum spectrum at the points ϕ_{ed}. For these momentum directions the quantities $C(\hat{\mathbf p},\omega)$ and $E(\hat{\mathbf p},\omega)$ appearing in (34) and (35) (or $C(\check{\mathbf p},\omega)$ and $E(\check{\mathbf p},\omega)$, depending on which of the two quantities $\Delta_\infty(\hat{\mathbf p})|$ and $|\Delta_\infty\ (\check{\mathbf p}|$ is smaller) diverge at $|\omega|=h_{ed}$.

4. CURRENT–VOLTAGE CHARACTERISTICS FOR THE JOSEPHSON AND QUASIPARTICLE CURRENTS WHEN THE ORDER PARAMETER IS NOT SUPPRESSED ON THE BOUNDARY

The features of the current–voltage characteristic appear because the functions $I_m(V,\hat{\mathbf p}_1)$ have singular points. If these functions do not depend on the momentum directions, then, according to (6)–(11), their dependence on the voltage directly describes the behavior of the current–voltage characteristic[1,2] (see also Ref. 3 and the references cited therein). Conversely, the features of the current–voltage characteristics for superconductors with anisotropic pairing differ significantly from the singularities of the functions $I_m(V,\hat{\mathbf p}_1)$ because of the integration of these functions over the momentum directions. In this case the features of the current–voltage characteristics will correspond to the singularities of the functions $j_m(V)$. Another distinguishing feature of superconductors with anisotropic pairing, which directly influences the behavior of the Josephson and quasiparticle currents, is their sensitivity to inhomogeneities and boundaries. Suppression of the anisotropic order parameter near a boundary leads, generally speaking, not only to the appearance of the quasiparticle bound states considered in the preceding section, but also to the emergence of an order parameter with a different symmetry in the near-boundary region.[19,20]

Even in the case of the ordinary boundary conditions, in which there is no suppression of the superconducting order parameter near the junction plane, the characteristic features of the current–voltage characteristic for a junction between superconductors with anisotropic pairing differ from those in the case of a junction between superconductors with isotropic pairing. Under the condition of specular reflection from the boundary, suppression of the order parameter is not observed for orientations of the crystal axes in which the order parameter remains unchanged when the incident momentum is replaced by the reflected momentum for all momentum directions. In this section we examine the current–voltage characteristics only for such orientations. Thus, we explore the behavior of the functions $j_m(V)$ for cases in which the quasiclassical propagators have the same form (33) on an opaque boundary as in the bulk of the superconductor. Substituting these propagators into Eqs. (8)–(11), we obtain

$$I_1(V,\hat{\mathbf p}_1)=-\Delta_1(\hat{\mathbf p}_1)\Delta_2(\hat{\mathbf p}_2)\int_{-\infty}^{\infty}d\omega\,\tanh\left(\frac{|\omega|}{2T}\right)$$

$$\times\left(\frac{\theta(|\Delta_1(\hat{\mathbf p}_1)|-|\omega-V|)}{\sqrt{|\Delta_1(\hat{\mathbf p}_1)|^2-(\omega-V)^2}}\,\frac{\theta(|\omega|-|\Delta_2(\hat{\mathbf p}_2)|)}{\sqrt{\omega^2-|\Delta_2(\hat{\mathbf p}_2)|^2}}\right.$$

$$\left.+\frac{\theta(|\omega|-|\Delta_1(\hat{\mathbf p}_1)|)}{\sqrt{\omega^2-|\Delta_1(\hat{\mathbf p}_1)|^2}}\,\frac{\theta(|\Delta_2(\hat{\mathbf p}_2)|-|\omega+V|)}{\sqrt{|\Delta_2(\hat{\mathbf p}_2)|^2-(\omega+V)^2}}\right), \tag{36}$$

$$I_2(V,\hat{\mathbf p}_1)=\Delta_1(\hat{\mathbf p}_1)\Delta_2(\hat{\mathbf p}_2)$$

$$\times\int_{-\infty}^{\infty}d\omega\left[\tanh\left(\frac{\omega}{2T}\right)-\tanh\left(\frac{\omega+V}{2T}\right)\right]$$

$$\times\frac{\mathrm{sgn}(\omega)\mathrm{sgn}(\omega+V)\,\theta(|\omega|-|\Delta_1(\hat{\mathbf p}_1)|)}{\sqrt{\omega^2-|\Delta_1(\hat{\mathbf p}_1)|^2}}$$

$$\times\frac{\theta(|\omega+V|-|\Delta_2(\hat{\mathbf p}_2)|)}{\sqrt{(\omega+V)^2-|\Delta_2(\hat{\mathbf p}_2)|^2}}, \tag{37}$$

$$I_3(V,\hat{\mathbf p}_1)=\int_{-\infty}^{\infty}d\omega\left[\tanh\left(\frac{\omega-V}{2T}\right)-\tanh\left(\frac{\omega}{2T}\right)\right]$$

$$\times\frac{|\omega||\omega-V|\,\theta(|\omega|-|\Delta_1(\hat{\mathbf p}_1)|)}{\sqrt{\omega^2-|\Delta_1(\hat{\mathbf p}_1)|^2}}$$

$$\times\frac{\theta(|\omega-V|-|\Delta_2(\hat{\mathbf p}_2)|)}{\sqrt{(\omega-V)^2-|\Delta_2(\hat{\mathbf p}_2)|^2}}, \tag{38}$$

$$I_4(V,\hat{\mathbf{p}}_1)=-\int_{-\infty}^{\infty}d\omega\ \tanh\left(\frac{\omega}{2T}\right)(\omega-V)|\omega|$$

$$\times\left(\frac{\theta(|\Delta_1(\hat{\mathbf{p}}_1)|-|\omega-V|)}{\sqrt{|\Delta_1(\hat{\mathbf{p}}_1)|^2-(-V)^2}}\frac{\theta(|\omega|-|\Delta_2(\hat{\mathbf{p}}_2)|}{\sqrt{\omega^2-|\Delta_2(\hat{\mathbf{p}}_2)|^2}}\right.$$

$$\left.+\frac{\theta(|\omega|-|\Delta_1(\hat{\mathbf{p}}_1)|)}{\sqrt{\omega^2-|\Delta_1(\hat{\mathbf{p}}_1)|^2}}\ \frac{\theta(|\Delta_2(\hat{\mathbf{p}}_2)|-|\omega-V|)}{\sqrt{|\Delta_2(\hat{\mathbf{p}}_2)|^2-(\omega-V)^2}}\right). \tag{39}$$

After integration over the frequency, the singularities of the functions $I_m(V,\hat{\mathbf{p}}_1)$ appear only if the two square roots (which multiply one another in the denominators of the integrands) can vanish simultaneously. This is possible for certain frequency and voltage values. As a result, we find that the functions $I_m(V,\hat{\mathbf{p}}_1)$ can have singularities at points for which $|V|=\|\Delta_2|\pm|\Delta_1\|$. From Eqs. (36)–(39) we find the following expressions for the singular parts of the functions $I_m(V,\hat{\mathbf{p}}_1)$:

$$I_1=\frac{1}{2}\ \sqrt{|\Delta_1\Delta_2|}\mathrm{sgn}(\Delta_1\Delta_2)\left\{\left[\tanh\left(\frac{|\Delta_1|}{2T}\right)\right.\right.$$

$$\left.+\tanh\left(\frac{|\Delta_2|}{2T}\right)\right]\ln\|V|-|\Delta_1|-|\Delta_2\|-\pi|\tanh\left(\frac{|\Delta_1|}{2T}\right)$$

$$\left.-\tanh\left(\frac{|\Delta_2|}{2T}\right)|\,\theta(|V|-\|\Delta_2|-|\Delta_1\|)\right\}, \tag{40}$$

$$I_2=\frac{1}{2}\ \sqrt{|\Delta_1\Delta_2|}\mathrm{sgn}(\Delta_1\Delta_2)\mathrm{sgn}(V)$$

$$\times\left\{\mathrm{sgn}(|\Delta_1|-|\Delta_2|)\left[\tanh\left(\frac{|\Delta_1|}{2T}\right)\right.\right.$$

$$\left.-\tanh\left(\frac{|\Delta_1|+|V|\mathrm{sgn}(|\Delta_2|-|\Delta_1|)}{2T}\right)\right]\ln||V|$$

$$-||\Delta_2|-|\Delta_1|||+\pi\left[\tanh\left(\frac{|\Delta_1|}{2T}\right)+\tanh\left(\frac{|\Delta_2|}{2T}\right)\right]$$

$$\left.\times\,\theta(|V|-|\Delta_1|-|\Delta_2|)\right\}, \tag{41}$$

$$I_3=\frac{1}{2}\ \sqrt{|\Delta_1\Delta_2|}\mathrm{sgn}(V)\left\{\mathrm{sgn}(|\Delta_1|-|\Delta_2|)\left[\tanh\left(\frac{|\Delta_1|}{2T}\right)\right.\right.$$

$$\left.-\tanh\left(\frac{|\Delta_1|+|V|\mathrm{sgn}(|\Delta_2|-|\Delta_1|)}{2T}\right)\right]$$

$$\times\ln||V|-||\Delta_2|-|\Delta_1|||-\pi\left[\tanh\left(\frac{|\Delta_1|}{2T}\right)\right.$$

$$\left.\left.+\tanh\left(\frac{|\Delta_2|}{2T}\right)\right]\theta(|V|-|\Delta_1|-|\Delta_2|)\right\}, \tag{42}$$

$$I_4=-\frac{1}{2}\ \sqrt{|\Delta_1\Delta_2|}\left\{\left[\tanh\left(\frac{|\Delta_1|}{2T}\right)+\tanh\left(\frac{|\Delta_2|}{2T}\right)\right]\right.$$

$$\times\ln||V|-\left|\Delta_1|-|\Delta_2||+\pi\right|\tanh\left(\frac{|\Delta_1|}{2T}\right)$$

$$\left.-\tanh\left(\frac{|\Delta_2|}{2T}\right)\right|\theta(|V|-||\Delta_2|-|\Delta_1||)\bigg\}. \tag{43}$$

According to (6) and (7), to find the behavior of a current–voltage characteristic in the vicinity of the singularities, Eqs. (40)–(43) must be integrated over the momentum directions. After this integration, we find that the current-voltage characteristic will exhibit nonanalytic behavior at $|V|$ equal to the values of the expressions $\|\Delta_2(\hat{\mathbf{p}}_2)\pm|\Delta_1(\hat{\mathbf{p}}_1)\|$ (which can be examined, for example, as a function of $\hat{\mathbf{p}}_1$) at the extremum points. As it turns out, the characteristic behavior of the current–voltage characteristic near these extrema depends strongly on the nature of the extremum points. It is seen from a comparison of Eqs. (40)–(43) that the pairs of functions $j_1(V)$, $j_4(V)$ and $j_2(V)$, $j_3(V)$ behave identically (to within a sign) near singularities. Therefore, it is sufficient to describe only the nonanalytic points for $j_1(V)$ and $j_3(V)$.

Below it will be convenient to consider the singular points in the conductance $G=dj_x/dV$. It turns out that to find the nonanalytic behavior of $G(V)$ only the θ functions and the logarithms need be differentiated with respect to V. In the former case, a δ function appears following differentiation. Consequently, integration over the Fermi surface reduces to integration along a line on that surface. The corresponding terms $\tilde{G}_{1,3}(V)$ in the functions $G_{1,3}(V)=dj_{1,3}/dV$ can be written in the following manner:

$$\tilde{G}_1=-\mathrm{sgn}(V)\int dl\ \frac{K^-(\hat{\mathbf{p}}_1)}{|\nabla_{\hat{\mathbf{p}}_1}(|\Delta_1|-|\Delta_2|)|},$$

$$\tilde{G}_3=-\int dl\ \frac{K^+(\hat{\mathbf{p}}_1)\mathrm{sgn}(\Delta_1\Delta_2)}{|\nabla_{\hat{\mathbf{p}}_1}(|\Delta_1|+|\Delta_2|)|}. \tag{44}$$

Here l is the local coordinate along the line $|V|=\|\Delta_1(\hat{\mathbf{p}}_1)|\pm|\Delta_2(\hat{\mathbf{p}}_2)\|$ on the Fermi surface (the plus sign belongs to $\tilde{G}_3$, and the minus sign belongs to $\tilde{G}_1$). The functions $K^\pm$ are defined by

$$K^\pm(\hat{\mathbf{p}}_1)=\frac{1}{16\pi^2}\left|\tanh\left(\frac{|\Delta_1|}{2T}\right)\pm\tanh\left(\frac{|\Delta_2|}{2T}\right)\right|$$

$$\times\sqrt{|\Delta_1\Delta_2|}\mathrm{sgn}(\Delta_1\Delta_2)\ \frac{v_{x1}}{v_{f1}}\ D. \tag{45}$$

Below we shall consider different types of extrema, and we shall obtain the features of the behavior of $G_m(V)$ corresponding to them. Let the function $\|\Delta_2|-|\Delta_1\|$ have a local maximum or minimum at the point $\mathbf{p}_1=\mathbf{p}_0$ on the Fermi surface and have the following form in the vicinity of that point:

$$\|\Delta_2|-|\Delta_1\|=a\pm(b\tilde{p}_1^2+c\tilde{p}_2^2),\quad a,b,c,>0. \tag{46}$$

Here $\tilde{p}_1$ and $\tilde{p}_2$ are the local orthogonal coordinates in the vicinity of the point $\mathbf{p}_0$. Since the function $\| \Delta_2| - |\Delta_1\|$ appears in the argument of the θ function in the expression for $j_{1,4}$, as well as in the argument of the logarithm in the formula for $j_{2,3}$, we obtain two different types of singular behavior near the voltage $|V|=a$:

$$\delta G_1|_{|V|=a} = \mathrm{sgn}(\Delta_1\Delta_2)|_{\hat{\mathbf{p}}_1=\hat{\mathbf{p}}_0}\delta G_4|_{|V|=a} = \mp\frac{\pi}{\sqrt{bc}}\, K^-(\hat{\mathbf{p}}_0), \tag{47}$$

$$G_2 = \mathrm{sgn}(\Delta_1\Delta_2)|_{\hat{\mathbf{p}}_1=\hat{\mathbf{p}}_0} G_3 = \pm\frac{1}{\sqrt{bc}}\, K^-(\hat{\mathbf{p}}_0)\ln||V|-a|. \tag{48}$$

Here we have introduced the following notation for the conductance jump:

$$\delta G|_{|V|=a} = G(|V|>a) - G(|V|<a).$$

There will be similar singularities when the function $|\Delta_2|+|\Delta_1|$ has a local maximum or minimum, near which $|\Delta_2|+|\Delta_1| = a \pm (b\tilde{p}_1^2 + c\tilde{p}_2^2), a,b,c,>0$:

$$\delta G_2|_{|V|=a} = -\mathrm{sgn}(\Delta_1\Delta_2)|_{\hat{\mathbf{p}}_1=\hat{\mathbf{p}}_0}\delta G_3|_{|V|=a}$$

$$= \mp\frac{\pi}{\sqrt{bc}}\, K^+(\hat{\mathbf{p}}_0), \tag{49}$$

$$G_1 = -\mathrm{sgn}(\Delta_1\Delta_2)|_{\hat{\mathbf{p}}_1=\hat{\mathbf{p}}_0} G_4$$

$$= \pm\frac{1}{\sqrt{bc}}\, K^+(\hat{\mathbf{p}}_0)\,\mathrm{sgn}(V)\ln||V|-a|. \tag{50}$$

When the function $\|\Delta_2||\Delta_1\|$ has a saddle point, near which

$$\| \Delta_2| - |\Delta_1\| = a + b\tilde{p}_1^2 - c\tilde{p}_2^2, \quad a,b,c,>0, \tag{51}$$

the corresponding singularities in the conductance have the form

$$G_1 = \mathrm{sgn}(\Delta_1\Delta_2)|_{\hat{\mathbf{p}}_1=\hat{\mathbf{p}}_0} G_4 = \frac{1}{\sqrt{bc}}\, K^-(\hat{\mathbf{p}}_0)$$

$$\times\mathrm{sgn}(V)\ln||V|-a|, \tag{52}$$

$$G_2 = \mathrm{sgn}(\Delta_1\Delta_2)|_{\hat{\mathbf{p}}_1=\hat{\mathbf{p}}_0} G_3 = \frac{2}{\pi\sqrt{bc}}$$

$$\times\mathrm{sgn}(|V|-a)K^-(\hat{\mathbf{p}}_0)\ln^2||V|-a|. \tag{53}$$

Similarly, for the saddle point of the function

$$|\Delta_2| + |\Delta_1| = a + b\tilde{p}_1^2 - c\tilde{p}_2^2, \quad a,b,c,>0 \tag{54}$$

we obtain

$$G_2 = -\mathrm{sgn}(\Delta_1\Delta_2)|_{\hat{\mathbf{p}}_1=\hat{\mathbf{p}}_0} G_3$$

$$= -\frac{1}{\sqrt{bc}}\, K^+(\hat{\mathbf{p}}_0)\ln|a-|V||, \tag{55}$$

$$G_1 = -\mathrm{sgn}(\Delta_1\Delta_2)|_{\hat{\mathbf{p}}_1=\hat{\mathbf{p}}_0} G_4$$

$$= \frac{2\,\mathrm{sgn}(V)\mathrm{sgn}(|V|-a)}{\pi\sqrt{bc}}\, K^+(\hat{\mathbf{p}}_0)$$

$$\times\ln^2||V|-a|. \tag{56}$$

The conductance jumps described by Eqs. (47) and (49) correspond to breaks in the current–voltage characteristic. The logarithmic divergences of the conductance described by the expressions (48), (50), (52), and (55) correspond to step-type features on the current–voltage characteristic (note that G has the same sign on both sides of these singular points). Finally, the terms containing squares of logarithms, (53) and (56), describe ''beak-shaped'' features on the current–voltage characteristic. They appear in the case of saddle points of the functions $\|\Delta_2|\pm|\Delta_1\|$ after integration of the logarithmic singularities in I_m.

We now consider the case in which the expressions $\|\Delta_2|\pm|\Delta_1\|$ take extremum values on a certain line $\tilde{l}$ on the Fermi surface, rather than at isolated points. Then, for example, in the vicinity of a local maximum or minimum of the functions $\|\Delta_2|-|\Delta_1\|$ we have

$$\|\Delta_2| - |\Delta_1\| = a \pm \tilde{p}_1^2, \quad a,b,>0, \tag{57}$$

where $\tilde{p}_1$ is the local coordinate on the Fermi surface that is orthogonal to the line of extrema $\tilde{l}$.

In this case the conductance will have square-root divergences on one side of the voltage value $|V|=a$:

$$G_1 = \mathrm{sgn}(\Delta_1\Delta_2)|_{\tilde{l}}\, G_4 = -\frac{\mathrm{sgn}(V)}{\sqrt{||V|-a|}}$$

$$\times\,\theta(\pm(|V|-a))\int\limits_{\substack{||\Delta_2|-|\Delta_1||=a\\ v_{x1}>0}} d\tilde{l}\,\frac{K^-}{2\sqrt{b}}, \tag{58}$$

$$G_2 = \mathrm{sgn}(\Delta_1\Delta_2)|_{\tilde{l}}\, G_3 = \mp\frac{\theta(\mp(|V|-a))}{\sqrt{||V|-a|}}$$

$$\times\int\limits_{\substack{||\Delta_2|-|\Delta_1||=a\\ v_{x1}>0}} d\tilde{l}\,\frac{K^-}{\sqrt{b}}. \tag{59}$$

These conductance singularities corresponding to one-sided vertical tangents to the current–voltage characteristic at $|V|=a$ in each of the four terms in the expression for the total tunneling current. For example, in the case of a maximum on the $\tilde{l}$ line there are vertical tangents to the $j_1(V)$ and $j_4(V)$ curves at $|V|=a$ on the $|V|<a$ side and to the $j_2(V)$ and $j_3(V)$ curves on the $|V|>a$ side.

Similarly, if the function $|\Delta_1|+|\Delta_2|$ achieves a local maximum or minimum on the $\tilde{l}$ line

$$|\Delta_1| + |\Delta_2| = a \pm b\tilde{p}_1^2, \quad a,b>0, \tag{60}$$

the conductance has the singular behavior:

$$G_2 = -\mathrm{sgn}(\Delta_1\Delta_2)\big|_{\tilde{l}}\, G_3$$

$$= \frac{\theta(\pm(|V|-a))}{\sqrt{||V|-a|}} \int\limits_{\substack{|\Delta_1|+|\Delta_2|=a \\ v_{x1}>0}} d\tilde{l}\, \frac{K^+}{2\sqrt{b}}, \tag{61}$$

$$G_1 = -\mathrm{sgn}(\Delta_1\Delta_2)\big|_{\tilde{l}}\, G_4$$

$$= \mp \frac{\mathrm{sgn}(V)}{\sqrt{||V|-a|}}\, \theta(\mp(|V|-a)) \int\limits_{\substack{|\Delta_1|+|\Delta_2|=a \\ v_{x1}>0}} d\tilde{l}\, \frac{K^+}{\sqrt{b}}. \tag{62}$$

In the formulas written out above for the singularities of the current–voltage characteristic associated with an extremum of the difference $|\Delta_1|+|\Delta_2|$, we could not go directly to the limit $\Delta_1 = \pm\Delta_2 = \Delta$. It follows from (40)–(43) that in this particular case G_1 and G_4 do not have a singularity at $V = |\Delta_1| - |\Delta_2| = 0$, while for G_2 and G_3 near $V=0$ (or, more precisely, at $|V| \ll T$), instead of (48), (53), and (59), we obtain the expression

$$G_2 = \pm G_3 = \pm \frac{\ln|V|}{2T} \int_{v_{x1}>0} \frac{d^2S_1}{(2\pi)^3} \frac{v_{x1}}{v_{f1}} \times D(\hat{\mathbf{p}}_1) \frac{|\Delta|}{\cosh^2(|\Delta|/2T)}. \tag{63}$$

For superconductors with isotropic pairing this expression is exponentially small at low temperatures $T \ll \Delta$. Conversely, for superconductors with anisotropic pairing the expression (63) has a power-function temperature dependence at $T \ll \Delta_{\max}$. For example, if the order parameter vanishes on a certain line on the Fermi surface (when $|\Delta(\hat{\mathbf{p}})| = b|\tilde{p}_1|$ holds near that line), it follows from (63) that $G_{2,3} \propto T \ln|V|$ at low temperatures.

Figures 2 and 3 present the results of numerical calculations of the $j_m(v)(v = V/\Delta_0)$ for orientations in which there is no suppression of the order parameter on either side of the tunnel barrier. Figure 2 presents plots for a junction between a superconductor with anisotropic pairing and an isotropic superconductor: $\Delta_1 = \Delta_0 \cos(2\phi)$, $\Delta_2 = \Delta_0/2 = \mathrm{const}$. Here ϕ is the azimuthal angle in the xy plane of a tetragonal superconductor (the z axis is parallel to the junction plane). For a superconductor with d-wave pairing we assume that the Fermi surface is cylindrical. We take the transparency of the barrier in the form $D \propto \cos^2\phi$, and $\Delta_0/(2T) = 0.5$. In this case the singular points of the current–voltage characteristic will be only the maximum values of $\|\Delta_1(\phi)| \pm |\Delta_2\|$, since the minimum values of these quantities are achieved when Δ_1 equals zero. We note that in the case of $\Delta_2 = 2\Delta_0$ only the minimum of $\|\Delta_1(\phi)| - |\Delta_2\|$ would be important for this same function $\Delta_1(\phi)$. Figure 3 presents the results of numerical calculations for a junction between two identical superconductors with anisotropic pairing: $\Delta_1 = \Delta_2 = \Delta_0 \cos(2\phi)$. All the functions $j_m(v)$ were normalized to $|j_1(0)|$.

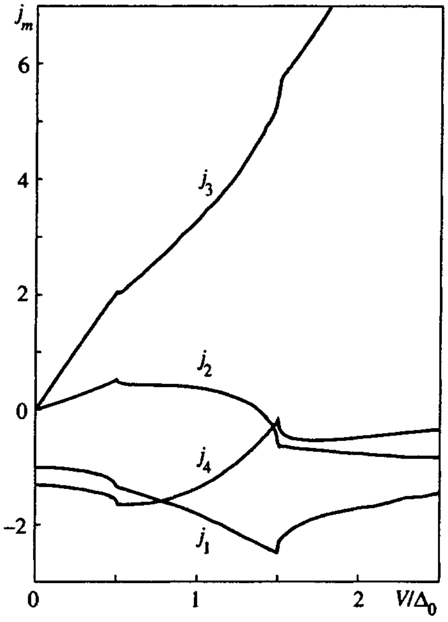


FIG. 2.

5. CURRENT–VOLTAGE CHARACTERISTIC FOR THE TUNNELING CURRENT WHEN THE ORDER PARAMETER IS SUPPRESSED ON ONE SIDE OF THE JUNCTION

Let us consider a tunnel junction between two superconductors with anisotropic pairing. Unlike the case considered in the preceding section, here we gradually vary the orientation of the crystal axis of one of the superconductors relative to the barrier plane, and we assume that the former condition $\Delta_2(\hat{\mathbf{p}}) = \Delta_2(\check{\mathbf{p}})$ holds for the other superconductor. According to (34) and (35), in an intermediate crystal orientation the

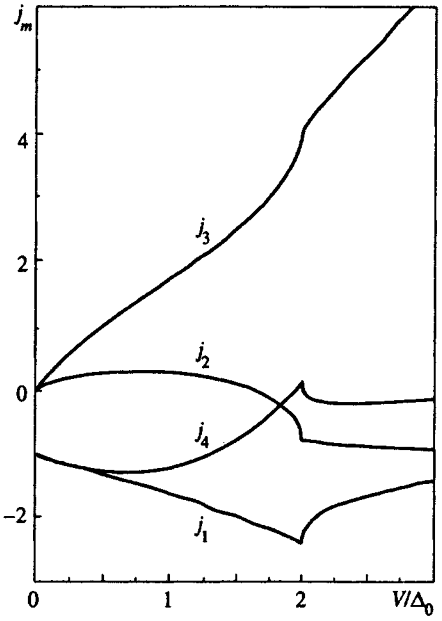


FIG. 3.

electron propagators for the first superconductor taken in the barrier plane do not have the square-root divergences characteristic of the values of these propagators in the interior of the superconductor. Therefore, the singularities which were found in the preceding section will be smoothed and will vanish at sufficiently large deviations from the initial orientation. At the same time, as we have already mentioned, some new characteristic singularities of the current–voltage characteristic appear in this case. Some of these singularities of the current–voltage characteristic are associated with the existence of regions on the Fermi surface where the order parameter $\Delta_{1,\infty}(\hat{\mathbf{p}})$ has different signs. A quasiparticle bound state with zero energy then appears near the boundary plane. Other singularities appear, if additional quasiparticle bound states with a nonzero energy occur because of the spatial variation of the order parameter.[8] As is seen from (34) and (35), in the former case the terms containing $1/\omega$ in the propagators are important. In the latter case the pole at a nonzero frequency should be considered in the propagators. Thus, we use Eqs. (34) and (35) for the singular parts of the propagators of the first superconductor and Eq. (33) for the propagators of the second superconductor.

It is significant that in the case in which quasiparticle bound states with zero energy exist on only one side of the junction (and there is no suppression of the order parameter on the other side) a singular contribution from these bound states appears only in the quasiparticle current, but not in the Josephson current. In fact, according to (30), the singular parts of the propagators associated with the pole at zero frequency satisfy the relation $f_s(\hat{\mathbf{p}})=f_s^+(\hat{\mathbf{p}})$. Conversely, the relation $f(\hat{\mathbf{p}})=-f^+(\hat{\mathbf{p}})$ holds for orientations in which there is no suppression of the order parameter. For this reason the corresponding singular parts of j_1 and j_2 vanish. However, this is not so for the bound states with a nonzero energy, since $Q_f(\hat{\mathbf{p}})$ is complex, unlike $B_f(\hat{\mathbf{p}})$, which is real ($B_f^+=B_f^*=B_f$, $Q_f^+=Q_f^*$).

Substituting (33)–(35) into (10) and (11) and then integrating over ω, we obtain the following singular contributions for I_3 and I_4:

$$\begin{aligned}
I_3(V,\hat{\mathbf{p}}_1)=&-B_{g1}(\hat{\mathbf{p}}_1)\\
&\times\tanh\left(\frac{V}{2T}\right)\frac{|V|\,\theta(|V|-|\Delta_2(\hat{\mathbf{p}}_2)|)}{\sqrt{V^2-|\Delta_2(\hat{\mathbf{p}}_2)|^2}}\\
&-\frac{Q_{g1}(\hat{\mathbf{p}}_1)\sqrt{|\Delta_2|}\operatorname{sgn}(V)}{\sqrt{2}}\\
&\times\Biggl\{\left[\tanh\left(\frac{h_1}{2T}\right)+\tanh\left(\frac{|\Delta_2|}{2T}\right)\right]\\
&\times\frac{\theta(|V|-h_1-|\Delta_2|)}{\sqrt{|V|-h_1-|\Delta_2|}}+\Biggl|\tanh\left(\frac{|\Delta_2|}{2T}\right)\\
&-\tanh\left(\frac{h_1}{2T}\right)\Biggr|\,\frac{\theta((|V|-||\Delta_2|-h_1|)\operatorname{sgn}(|\Delta_2|-h_1))}{(V|-||\Delta_2|-h_1|)\operatorname{sgn}(|\Delta_2|-h_1)}\Biggr\},
\end{aligned}\tag{64}$$

$$\begin{aligned}
I_4(V,\hat{\mathbf{p}}_1)=&B_{g1}(\hat{\mathbf{p}}_1)\\
&\times\tanh\left(\frac{|V|}{2T}\right)\frac{|V|\,\theta(|\Delta_2(\hat{\mathbf{p}}_2)|-|V|)}{\sqrt{|\Delta_2(\hat{\mathbf{p}}_2)|^2-V^2}}\\
&+\frac{Q_{g1}(\hat{\mathbf{p}}_1)\sqrt{|\Delta_2|}}{\sqrt{2}}\Biggl\{\Biggl[\tanh\left(\frac{h_1}{2T}\right)\\
&+\tanh\left(\frac{|\Delta_2|}{2T}\right)\Biggr]\frac{\theta(h_1+|\Delta_2|-|V|)}{\sqrt{h_1+|\Delta_2|-|V|}}\\
&+\left[\tanh\left(\frac{|\Delta_2|}{2T}\right)-\tanh\left(\frac{h_1}{2T}\right)\right]\\
&\times\theta\frac{((|V|-||\Delta_2|-h_1|)\operatorname{sgn}(h_1-|\Delta_2|))}{\sqrt{(V|-||\Delta_2|-h_1|)\operatorname{sgn}(h_1-|\Delta_2|)}}\Biggr\}
\end{aligned}\tag{65}$$

The terms describing the square-root analyticities in (34) and (35) are not taken into account in (64) and (65), since in the case under consideration here they lead to jumps and divergences only in derivatives of the conductance, but not in the current or the conductance itself. Because the dependences of the order parameter Δ_2 and the function h_1 of the momentum directions can be different, subsequent integration over the Fermi surface can give rise to different types of singular points on the current–voltage characteristic [the type of singularity is influenced by the behavior of the corresponding functions near the extremum (see below)]. Let us first consider the singularities of the current–voltage characteristic, which are related to the terms of the form $1/\omega$ in the expressions for the propagators of the first superconductor, and confine ourselves to two important examples. In the special case in which the second superconductor is a superconductor with isotropic pairing, the following square-root divergences will appear in the current–voltage characteristic:

$$\begin{aligned}
j_3=&-\tanh\left(\frac{V}{2T}\right)\frac{|V|\,\theta(|V|-|\Delta_2|)}{\sqrt{V^2-|\Delta_2|^2}}\\
&\times\int_{v_{x1}>0}\frac{d^2S_1}{(2\pi)^3}\,\frac{v_{x1}}{v_{f1}}\,DB_{g1}(\hat{\mathbf{p}}_1),
\end{aligned}\tag{66}$$

$$\begin{aligned}
j_4=&-\tanh\left(\frac{|V|}{2T}\right)\frac{|V|\,\theta(|\Delta_2|-|V|)}{\sqrt{|\Delta_2|^2-V^2}}\\
&\times\int_{v_{x1}>0}\frac{d^2S_1}{(2\pi)^3}\,\frac{v_{x1}}{v_{f1}}\,DB_{g1}(\hat{\mathbf{p}}_1).
\end{aligned}\tag{67}$$

The function j_3 diverges at $|V|=\Delta_2$ on the $|V|>\Delta_2$ side, and j_4 diverges on the $|V|<\Delta_2$ side. In the case $\Delta_2=0$ (an S–N junction), it follows from (66) that $G_3 \propto [T\cosh^2(V/2T)]^{-1}$. Hence at the low voltages $|V|\ll T$ we have $G_3\propto 1/T$, and we obtain an anomalous increase in the conductance at low temperatures (see also Ref. 21). Of course, such a divergence of G_3 occurs at zero temperature only for an idealized system when the factors leading to broadening of the δ peaks in the quasiparticle density of states are not taken into account.

Next, if the anisotropic order parameter Δ_2 has an extremum on a certain line on the Fermi surface, near which

$$|\Delta_2|=a\pm b\tilde{p}_1^2, \quad a,b>0, \tag{68}$$

the type of singularity on the current–voltage characteristic depends on the behavior of the function $B_{g1}(\mathbf{p}_1)$ only of this line. If $B_{g1}(\mathbf{p}_1)\neq 0$ on this line of extrema, j_3 will have a logarithmic divergence for a line of maxima, and j_4 will have such a divergence for a line of minima:

$$j_3, j_4 \propto \sqrt{\frac{a}{b}}\tanh\left(\frac{a}{2T}\right)\ln||V|-a|. \tag{69}$$

If the function $B_{g1}(\mathbf{p}_1)$ vanishes on the line of extrema and exhibits linear behavior near zero, i.e.,

$$B_{g1}(\tilde{p}_1)=\beta|\tilde{p}_1|, \tag{70}$$

then j_3 and j_4 will have one-sided vertical tangents at $|V|=a$:

$$G_3\propto\frac{\sqrt{a}}{b}\tanh\left(\frac{a}{2T}\right)\frac{\theta(|V|-a)}{\sqrt{|V|-a}},$$

$$G_4\propto\frac{\sqrt{a}}{b}\tanh\left(\frac{a}{2T}\right)\frac{\theta(a-|V|)}{\sqrt{a-|V|}}. \tag{71}$$

Now, let us move on to consider the singularities of the current–voltage characteristic appearing because of the poles in the propagators of the first superconductor at a nonzero frequency. Unlike the contribution from the bound states with zero energy, the quasiparticle bound states with a nonzero energy make a singular contribution not only to the quasiparticle current, but also to the Josephson current, even if they are present on only one side of the junction plane. This is related to the nonzero value of $\mathrm{Im}\,Q_f(\hat{\mathbf{p}})$ (while $\mathrm{Im}\,B_f(\hat{\mathbf{p}})=0$). The positions of the corresponding singular points on the current–voltage characteristic are related to the extremum values of the quantities $h_1\pm|\Delta_2|$ on the Fermi surface. For example, the presence of a line of extrema l for

$$h_1+|\Delta_2|=a\pm b\tilde{p}_1^2, \quad a,b>0 \tag{72}$$

leads to logarithmic divergences and discontinuities of the functions j_m:

$$j_1, -j_4=M^+_{f,g}\begin{cases}\ln||V|-a|,\\ \pi\theta(|V|-a),\end{cases}$$

$$-j_2, j_3=M^+_{f,g}\,\mathrm{sgn}(V)\begin{cases}\pi\theta(a-|V|),\\ \ln||V|-a|.\end{cases} \tag{73}$$

The upper (lower) rows in these formulas correspond to the upper (lower) sign in (72), and

$$M^{\pm}_g=\frac{1}{16\sqrt{2}\pi^3}\int_{v_{x1}>0}dl\,\frac{v_{x1}}{v_{f1}}\,D\sqrt{\frac{|\Delta_2|}{b}}\,Q_{g1}(\hat{\mathbf{p}}_1) \times\left|\tanh\left(\frac{h_1}{2T}\right)\pm\tanh\left(\frac{|\Delta_2|}{2T}\right)\right|, \tag{74}$$

$$M^{\pm}_f=\int_{v_{x1}>0}dl\,\frac{v_{x1}}{v_{f1}}\,D\sqrt{\frac{|\Delta_2|}{b}}\frac{\mathrm{sgn}(\Delta_2)}{16\sqrt{2}\pi^3}\,\mathrm{Im}\,Q_{f1}(\hat{\mathbf{p}}_1) \times\left|\tanh\left(\frac{h_1}{2T}\right)\pm\tanh\left(\frac{|\Delta_2|}{2T}\right)\right|. \tag{75}$$

If the function $||\Delta_2|-h_1|$ has a line of extrema

$$||\Delta_2|-h_1|=a\pm b\tilde{p}_1^2, \quad a,b>0, \tag{76}$$

we obtain the following singularities in the current:

$$j_{1,4}=M^-_{f,g}\{\mp\theta(\pm(|\Delta_2|-h_1))\ln||V|-a| +\pi\theta(\mp(|\Delta_2|-h_1))\,\theta(a-|V|)\}, \tag{77}$$

$$j_{2,3}=M^-_{f,g}\,\mathrm{sgn}(V)\{\theta(\mp(|\Delta_2|-h_1))\ln||V|-a| \pm\pi\theta(\pm(|\Delta_2|-h_1))\,\theta(a-|V|)\}. \tag{78}$$

Thus, for a positive (negative) value of $|\Delta_2|-h_1$ the functions $j_{2,3}$ ($j_{1,4}$) have a logarithmic divergence only for the line of maxima of $||\Delta_2|-h_1|$, while the functions $j_{1,4}$ ($j_{2,3}$) have such a divergence only for the line of minima of this quantity.

As was noted above, the function $h(\hat{\mathbf{p}})$ can have nonanalytic behavior, for example, in the momentum direction for which the bound state vanishes near the boundary. Taking into account this possibility, we assume that the function $h_1+|\Delta_2|$ exhibits the following nonanalytic behavior near some line on the Fermi surface:

$$h_1+|\Delta_2|=a+[b\,\theta(\tilde{p})+c\,\theta(-\tilde{p})]\tilde{p}, \quad a>0. \tag{79}$$

The special case of b(or c)$\to\infty$ corresponds to the absence of a bound state in the directions $\tilde{p}>0$ (or $\tilde{p}<0$) from this line. Then the functions j_m will have the following one-sided vertical tangents at $|V|=a$:

$$G_1, -G_4=\frac{\theta(a-|V|)}{\sqrt{a-|V|}}\,\mathrm{sgn}(V)P^+_{f,g},$$

$$G_2, -G_3=\frac{\theta(|V|-a)}{\sqrt{|V|-a}}\,P^+_{f,g}. \tag{80}$$

Here we have introduced the notation

$$P^{\pm}_g=\int_{v_{x1}>0}dl\left(\frac{1}{b}-\frac{1}{c}\right)\frac{v_{x1}}{v_{f1}}\,D\sqrt{|\Delta_2|}\,\frac{Q_{g1}(\hat{\mathbf{p}}_1)}{8\sqrt{2}\pi^3} \times\left|\tanh\left(\frac{h_1}{2T}\right)\pm\tanh\left(\frac{|\Delta_2|}{2T}\right)\right|, \tag{81}$$

$$P^{\pm}_f=\int_{v_{x1}>0}dl\left(\frac{1}{b}-\frac{1}{c}\right) \times\frac{v_{x1}}{v_{f1}}\,D\sqrt{|\Delta_2|}\,\frac{\mathrm{sgn}(\Delta_2)}{8\sqrt{2}\pi^3}\,\mathrm{Im}\,Q_{f1}(\hat{\mathbf{p}}_1) \times\left|\tanh\left(\frac{h_1}{2T}\right)\pm\tanh\left(\frac{|\Delta_2|}{2T}\right)\right|. \tag{82}$$

Similarity, if the function $|h_1-|\Delta_2||$ has the form (79) near some line, the functions j_m will have one-sided vertical

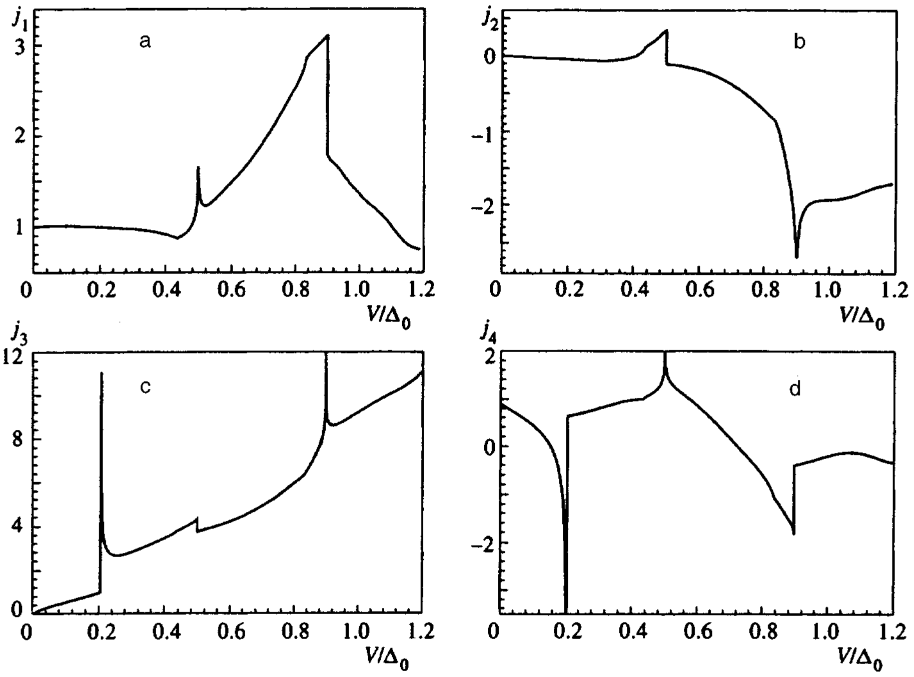


FIG. 4.

tangents at $|V|=a$:

$$G_{1,4}=\mathrm{sgn}(V)\mathrm{sgn}(|\Delta_2|-h_1)\times\frac{\theta((a-|V|)\,(|\Delta_2|-h_1))}{\sqrt{(a-|V|)\mathrm{sgn}(|\Delta_2|-h_1)}}\,P^-_{f,g}, \tag{83}$$

$$G_{2,3}=-\frac{\theta((|V|-a)\,(|\Delta_2|-h_1))}{\sqrt{(|V|-a)\mathrm{sgn}(|\Delta_2|-h_1)}}\,P^-_{f,g}. \tag{84}$$

Formulas (64) and (65) are not suitable for direct passage to the limit $\Delta_2=0$ at a finite voltage, since it was assumed during their derivation that the magnitude of the voltage is close to the corresponding singularities in these formulas. In the case of an S–N junction, it follows from (7), (10), and (11) that if h_1 has an extremum of the form $h_1=a\pm b\tilde{p}_1^2$, a, $b>0$, the conductance at $|V|=a$: will exhibit the low-temperature anomaly $G_{3,4}\propto 1/\sqrt{bT}$.

Figure 4 presents the results of numerical calculations of plots of $j_m(v)$ $(v=V/\Delta_0)$ for the case in which there is suppression of the order parameter on only one side of the tunnel barrier. A tunnel junction between a superconductor with d-wave pairing and a superconductor with isotropic s-wave pairing was considered under the following conditions:

$$\Delta_{1\infty}=\Delta_0\cos(2\phi-2\phi_0),\quad \Delta_2=0.2\Delta_0=\mathrm{const}.$$

For the superconductor with d-wave pairing we chose the same parameters as before (see Fig. 1): $\phi_0=\pi/9$, $T=0.45T_{c1}$, $\Delta_0/(2T)=2$. Here T_{c1} is the critical temperature for the superconductor with d-wave pairing. The transparency of the tunnel barrier was taken in the form $D\propto\cos^2\phi$.

The plots of j_3 and j_4 have square-root divergences at $V=\Delta_2=0.2\Delta_0$. In addition, at $V=h_m-\Delta_2=0.5\Delta_0$ the functions j_1 and j_4 have logarithmic divergences, while j_2 and j_3 have jumps. At $V=h_m+\Delta_2=0.9\Delta_0$ there are logarithmic divergences on j_2 and j_3 and jumps on j_1 and j_4. At the voltages $V=h_{\mathrm{ed}}\pm\Delta_2=0.83\Delta_0$, $0.43\Delta_0$ the current–voltage characteristics have breaks, although some of them are weakly expressed. All the functions $j_m(v)$ were normalized to $j_1(0)$.

6. CURRENT–VOLTAGE CHARACTERISTIC FOR THE TUNNELING CURRENT WHEN THE ORDER PARAMETER IS SUPPRESSED ON BOTH SIDES OF THE JUNCTION

Let us now consider the case in which the order parameter is suppressed significantly near the junction plane in both superconductors. The singular parts of the propagators on both sides of the barrier should then be described by (34) and (35). When these expressions are substituted into (8)–(11), several types of nonanalytic terms appear. They are obtained after renormalization of the pole terms to one another and from the product of the pole terms on one side and the square-root terms in (34) and (35) on the other side. After this, for the singular part of I_3, for example, we find the following expression:

$$
\begin{aligned}
I_3={}&\frac{1}{\pi}\tanh\left(\frac{V}{2T}\right)B_{g1}(\hat{\mathbf{p}}_1)\\
&\times\Bigl[\mathrm{Re}\ C_2(\hat{\mathbf{p}}_2)\sqrt{V^2-\Delta_{2\infty}^2(\hat{\mathbf{p}}_2)}\\
&\times\theta(|V|-|\Delta_{2\infty}(\hat{\mathbf{p}}_2)|)\\
&+\mathrm{Im}\ C_2(\hat{\mathbf{p}}_2)\sqrt{\Delta_{2\infty}^2(\hat{\mathbf{p}}_2)-V^2}\,\theta(|\Delta_{2\infty}(\hat{\mathbf{p}}_2)|\\
&-|V|-\frac{\pi}{2}\ Q_{g2}(\hat{\mathbf{p}}_2)\,\delta(|V|-h_2(\hat{\mathbf{p}}_2))\\
&+(\hat{\mathbf{p}}_2\rightarrow(\check{\mathbf{p}}_2)\Bigr]+\frac{1}{\pi}\ \Biggl\{\Biggl[\tanh\left(\frac{h_1(\hat{\mathbf{p}}_1)+V}{2T}\right)\\
&-\tanh\left(\frac{h_1(\hat{\mathbf{p}}_1)}{2T}\right)\Biggr]Q_{g1}(\hat{\mathbf{p}}_1)\\
&\times\Biggl[-\frac{\pi}{4}\ Q_{g2}(\hat{\mathbf{p}}_2)\,\delta(|h_1(\hat{\mathbf{p}}_1)+V|-h_2(\hat{\mathbf{p}}_2))\\
&+\mathrm{Re}\ C_2(\hat{\mathbf{p}}_2)\sqrt{(h_1(\hat{\mathbf{p}}_1)+V)^2-\Delta_{2\infty}^2(\hat{\mathbf{p}}_2)}\\
&\times\theta(|h_1(\hat{\mathbf{p}}_1)+V|-|\Delta_{2\infty}(\hat{\mathbf{p}}_2)|)+\mathrm{Im}\ C_2\\
&\times(\hat{\mathbf{p}}_2)\sqrt{\Delta_{2\infty}^2(\hat{\mathbf{p}}_2)-(h_1(\hat{\mathbf{p}}_1)+V)^2}\\
&\times\theta(|\Delta_{2\infty}(\hat{\mathbf{p}}_2)|-|h_1(\hat{\mathbf{p}}_1)+V|)-(V\rightarrow-V)\Biggr]\\
&+(\hat{\mathbf{p}}_2\rightarrow\check{\mathbf{p}}_2)\Biggr\}+(1\leftrightarrow 2).
\end{aligned}
\tag{85}
$$

Similar singularities appear in the expressions for $I_{1,2,4}$.

In the last formula we neglected the contribution which is obtained after multiplying the pole terms (where $\omega=0$) appearing in the Green's functions on both sides of the junction. When the broadening of the bound states is neglected, this contribution vanishes at a finite voltage. However, in real systems it can be important for voltages that are less than or comparable to the characteristic width of the corresponding peak in the quasiparticle density of states. For the stationary Josephson effect this contribution leads to a low-temperature anomaly in the critical current and to the possibility (at some temperature) of a phase transition from a 0 junction to a π junction.[22,23]

During the subsequent integration over the momentum directions, we assume that the order parameter $|\Delta_{2\infty}(\hat{\mathbf{p}}_2)|$ or the quantities $|h_1\pm|\Delta_{2\infty}||$, h_2, and $|h_1\pm h_2|$ have a line of extrema. Let, for example, $|\Delta_{2\infty}(\hat{\mathbf{p}}_2)|$ or $|h_1\pm|\Delta_{2\infty}||$ have a line of extrema of the form

$$
|\Delta_{2\infty}|,\ \ |h_1\pm|\Delta_{2\infty}||=a\pm b\tilde{p}^2,\quad a,b>0,
\tag{86}
$$

and let $B_1(\hat{\mathbf{p}}_1)$ or $|\Delta_{2\infty}|$ and $Q_1(\hat{\mathbf{p}}_1)$ be nonvanishing on that line. Then the conductance $G_{1,2,4}$ will have the following logarithmic singularities:

$$
\mathrm{sgn}(V)G_1,G_2,\mathrm{sgn}(V)G_4\propto\sqrt{\frac{a}{b}}\ \ln||V|-a|.
\tag{87}
$$

At the same time, the logarithmic divergence in G_3 of the form (87) appears only for a line of maxima. This is because $\mathrm{Im}\ C=0$ holds for $|\omega|<\min(|\Delta_\infty(\check{\mathbf{p}})|,|\Delta_\infty(\hat{\mathbf{p}})|)$. For the special orientation $\Delta_2(\hat{\mathbf{p}}_2,x)=-\Delta_2(\check{\mathbf{p}}_2,x)$, the logarithmic divergence in G_4 remains only for a line of minima. The square-root behavior of the propagator of the form $\sqrt{\Delta_\infty^2(\hat{\mathbf{p}})-\omega^2}$ led here to logarithmic divergence in the conductance. In other cases it can give rise to a kink in the conductance or divergence of its derivatives. In the present work we shall not consider such singularities.

If the function h_2 has a line of extrema of the form

$$
h_2=a\pm b\tilde{p}^2,\quad a,b>0,
\tag{88}
$$

the current–voltage characteristic will have the following square-root divergences:

$$
\begin{aligned}
j_1,\mp j_4={}&\frac{\pm 1}{16\pi^3}\tanh\left(\frac{|V|}{2T}\right)\ \frac{\theta(\pm(a-|V|))}{\sqrt{||V|-a|}}\\
&\times\int_{v_{x1}>0}dl\ \frac{v_{x1}}{v_{f1}}\ B_{f,g1}(\hat{\mathbf{p}}_1)\mathrm{Re}\ Q_{f,g2}(\hat{\mathbf{p}}_2)\ \frac{D}{\sqrt{b}},
\end{aligned}
\tag{89}
$$

$$
\begin{aligned}
j_2,-j_3={}&\frac{1}{8\pi^3}\tanh\left(\frac{V}{2T}\right)\ \frac{\theta(\pm(|V|-a))}{\sqrt{||V|-a|}}\\
&\times\int_{v_{x1}>0}dl\ \frac{v_{x1}}{v_{f1}}\ B_{f,g1}(\hat{\mathbf{p}}_1)\mathrm{Re}\ Q_{f,g2}(\hat{\mathbf{p}}_2)\ \frac{D}{\sqrt{b}}.
\end{aligned}
\tag{90}
$$

In this section the coefficients B_{f2} and Q_{f2}^* are taken from the corresponding expressions for f^+ of the second superconductor, while B_{f1} and Q_{f1} are taken from the expressions for f of the first superconductor.

Let the function h_2 near a certain line l have, for example, the following nonanalytic behavior:

$$
h_2=a+[b\,\theta(\tilde{p})+c\,\theta(-\tilde{p})]\tilde{p},\quad a>0.
\tag{91}
$$

This leads to logarithmic divergence of the functions $j_{1,4}$ at $|V|=a$:

$$
\begin{aligned}
j_1,j_4={}&\frac{\ln||V|-a|}{8\pi^4}\tanh\left(\frac{|V|}{2T}\right)\\
&\times\int_{v_{x1}>0}dl\ \frac{v_{x1}}{v_{f1}}\ DB_{f,g1}(\hat{\mathbf{p}}_1)\\
&\times\mathrm{Re}\ Q_{f,g2}(\hat{\mathbf{p}}_2)\left(\frac{1}{c}-\frac{1}{b}\right).
\end{aligned}
\tag{92}
$$

The functions j_2 and j_3 have jumps:

$$
\begin{aligned}
j_2,-j_3={}&\frac{1}{8\pi^3}\tanh\left(\frac{V}{2T}\right)\\
&\times\int_{v_{x1}>0}dl\ \frac{v_{x1}}{v_{f1}}\ DB_{f,g1}(\hat{\mathbf{p}}_1)\mathrm{Re}\ Q_{f,g2}(\hat{\mathbf{p}}_2)\\
&\times\left[\frac{\theta((|V|-a)b)}{|b|}+\frac{\theta((a-|V|)c)}{|c|}\right].
\end{aligned}
\tag{93}
$$

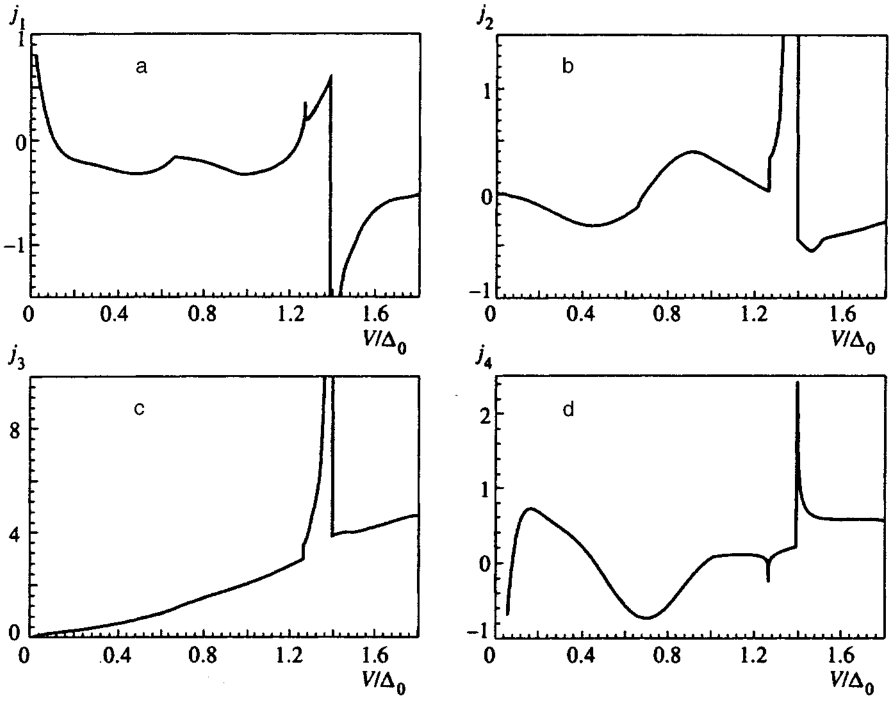


FIG. 5.

Finally, let us consider the case in which $|h_1-h_2|$ has a line of extrema

$$|h_1-h_2|=a\pm b\tilde{p}^2, \quad a,\ b>0. \tag{94}$$

Then we find that the current–voltage characteristics will have the following square-root divergences:

$$j_1,-j_4=\frac{\mp\theta(\pm(a-|V|))}{16\pi^3\sqrt{||V|-a|}}\int_{v_{x1}>0}\frac{dl}{\sqrt{b}}\,N_{f,g}, \tag{95}$$

$$j_2,-j_3=\frac{\mathrm{sgn}(V)\,\theta(\pm(|V|-a))}{8\pi^3\sqrt{||V|-a|}}\int_{v_{x1}>0}\frac{dl}{\sqrt{b}}\,N_{f,g}. \tag{96}$$

Here we have introduced the notation

$$N_{f,g}=\frac{v_{x1}}{v_{f1}}\,D\ \mathrm{Re}(Q_{f,g1}Q^*_{f,g2})\left|\tanh\left(\frac{h_1}{2T}\right)-\tanh\left(\frac{h_2}{2T}\right)\right|. \tag{97}$$

In the case in which the value of $|h_1-h_2|$ near a certain line has nonanalytic behavior of the form (91), the functions j_1 and j_4 have logarithmic divergences at $|V|=a$:

$$j_1,-j_4=\frac{\ln||V|-a|}{8\pi^4}\int_{v_{x1}>0}dlN_{f,g}\left(\frac{1}{b}-\frac{1}{c}\right), \tag{98}$$

and the functions j_2 and j_3 undergo jumps:

$$j_2,-j_3=\frac{\mathrm{sgn}(V)}{8\pi^3}\int_{v_{x1}>0}dlN_{f,g}\left[\frac{\theta((|V|-a)b)}{|b|}+\frac{\theta((a-|V|)c)}{|c|}\right]. \tag{99}$$

When h_1+h_2 has a line of extrema, all the results for the singularities are obtained from (94)–(99) after the replacements $h_2\rightarrow-h_2$, $j_1\rightarrow-j_1$, and $Q^*_{f2}\rightarrow Q_{f2}$ in these formulas. It should be added that besides the singular points described in this section there are also similar singularities, which are described by the same Eqs. (86)–(93) after the replacement $1\leftrightarrow2$.

Figure 5 presents the results of numerical calculations of the functions $j_m(v)(v=V/\Delta_0)$ when the order parameter is suppressed on both sides of the tunnel barrier. We considered a tunnel junction between two identical superconductors with d-wave pairing in the special case of a specularly reflecting junction, in which the tunnel barrier is a symmetry plane of the superconducting electrodes: $\Delta_{1\infty}(\hat{\mathbf{p}}_1)=\Delta_{2\infty}(\check{\mathbf{p}}_1)=\Delta_0\cos(2\phi-2\phi_0)$. As before (see Fig. 1), we took $\phi_0=\pi/9$, $T=0.45T_c$, $\Delta_0/(2T)=2$, $D\propto\cos^2\phi$. At the voltage $V=2h_m=1.4\Delta_0$ the plots of $j_{1,2,3,4}$ have square-root divergences. In addition, at $V=2h_{\mathrm{ed}}=1.26\Delta_0$ the plots of j_1 and j_4 have logarithmic divergences, and the plots of j_2 and j_3 have jumps. Near $V=0.635\Delta_0$ the behavior of the plots of j_1 and j_2 resembles a break. This is attributed to the contribution of the bound states with zero energy on one side of the junction and the states in the continuum spectrum of the quasiparticles on the other size of the junction. The function $\min(|\cos(2\phi-2\phi_0)|,|\cos(2\phi+2\phi_0)|)$ takes the maximum value 0.635 in the direction $\phi=\pm0.79$, in which $B_f\neq0$.

7. CONCLUSIONS

As we have shown above, a large variety of nonanalytic points can appear on the current–voltage characteristics of junctions between superconductors with anisotropic pairing when the quasiparticles are specularly reflected from the plane of the tunnel barrier. The singular behavior of the current–voltage characteristic differs significantly from the behavior typical of a junction between superconductors with isotropic s-wave pairing. Of course, in real systems all the features should be smeared out to a considered extent. In particular, the quality of the barrier plane can have an appreciable influence on the occurrence of these features. For example, instead of the divergences in the current described above, the appearance of finite peaks should be expected under real conditions. The amplitude and width of these peaks are sensitive to elastic and inelastic scattering processes, irregularities on the barrier plane, and the value of the transparency factor. Inelastic scattering processes smooth the singularities in the tunneling current for superconductors with either isotropic or anisotropic pairing. Elastic scattering processes are known to suppress superconductivity only for anisotropic pairing. All these factors, together with the quality of the plane of the tunnel barrier and the finite value of the transparency, cause broadening of the δ peaks in the quasiparticle density of states and, therefore, erode the corresponding peaks in the tunneling current. Nevertheless, the characteristic behavior of the current–voltage characteristic considered above can be observed under certain real conditions, and it can serve as a sensitive test for determining anisotropic types of pairing in superconductors, particularly for detecting reversal of the sign of the order parameter on the Fermi surface.

One of us (Yu.S.B.) sincerely thanks D. Rainer and J. A. Sauls for a fruitful discussion and for allowing us to acquaint ourselves with their work before publication, as well as H. Burkhardt for some useful remarks. This work was supported by the Russian Fund for Fundamental Research (Grant No. 96-02-16249). A.A.S. thanks the Jülich Research Center for its financial support.

[1] A. I. Larkin and Yu. N. Ovchinnikov, Zh. Éksp. Teor. Fiz. **51**, 1535 (1966) [Sov. Phys. JETP **24**, 1035 (1967)].

[2] N. R. Werthamer, Phys. Rev. **147**, 255 (1966).

[3] A. Barone and G. Paterno, *Physics and Applications of the Josephson Effect*, Wiley, New York (1982) [Russ. transl., Mir, Moscow (1984)].

[4] V. B. Geshkenbein, I. A. Larkin, and A. Barone, Phys. Rev. B **36**, 235 (1987).

[5] M. Sigrist and T. M. Rice, Soc. Jpn. **61**, 4283 (1992).

[6] A. Millis, D. Rainer, and J. A. Sauls, Phys. Rev. B **38**, 4504 (1988).

[7] Yu. S. Barash, A. V. Galaktionov, and A. D. Zaikin, Phys. Rev. B **52**, 665 (1995).

[8] L. J. Buchholtz, M. Palumbo, D. Rainer, and J. A. Sauls, J. Low Temp. Phys. **101**, 1099 (1995).

[9] M. Sigrist and T. M. Rice, Rev. Mod. Phys. **67**, 503 (1995).

[10] D. J. Van Harlingen, Rev. Mod. Phys. **67**, 515 (1995).

[11] C. Renner and Ø. Fischer, Phys. Rev. B **51**, 9208 (1995).

[12] Yu. S. Barash, A. V. Galaktionov, and A. D. Zaikin, Phys. Rev. Lett. **75**, 1676 (1995).

[13] E. Riedel, Z. Naturforsch. A **19**, 1634 (1964).

[14] C.-R. Hu, Phys. Rev. Lett. **72**, 1526 (1994).

[15] J. Yang and C.-R. Hu, Phys. Rev. B **50**, 16 766 (1994).

[16] M.M. Matsumoto and H. Shiba, J. Phys. Soc. Jpn. **64**, 4867 (1995).

[17] Y. Nagato and K. Nagai, Phys. Rev. B **51**, 16 254 (1995).

[18] A. V. Zaĭtsev, Zh. Eksp. Teor. Fiz. **86**, 1742 (1984) [Sov. Phys. JETP **59**, 1015 (1984)].

[19] L. J. Buchholtz, M. Palumbo, D. Rainer, and J. A. Sauls, J. Low Temp. Phys. **101**, 1079 (1995).

[20] M. Matsumoto and H. Shiba, J. Phys. Soc. Jpn. **64**, 3384 (1995).

[21] Y. Tanaka and S. Kashiwaya, Phys. Rev. Lett. **74**, 3451 (1995).

[22] Yu. S. Barash, H. Burkhardt, and D. Rainer, Phys. Rev. Lett. **77**, 4070 (1996).

[23] Y. Tanaka and S. Kashiwaya, Phys. Rev. B **53**, 11 957 (1996).

Translated by P. Shelnitz