Equilibrium of a high-current channel in the general theory of relativity

B. É. Meĭerovich

P. L. Kapitsa Institute of Physical Problems, Russian Academy of Sciences, 117973 Moscow, Russia
(Submitted 10 July 1996)
Zh. Éksp. Teor. Fiz. **112**, 385–399 (August 1997)

The condition for equilibrium of a high-current channel taking account of both electromagnetic and gravitational interactions of the charges with an arbitrary drift-to-light velocity ratio is derived from the equations of Einstein's general theory of relativity. The relative motion appearing between the electron and ion subsystems as a result of the current flow gives rise to an additional gravitational attraction between these subsystems. This is a relativistic effect that cannot be obtained in the Newtonian approximation. © *1997 American Institute of Physics.* [S1063-7761(97)00108-X]

1. INTRODUCTION

High intergalactic currents, which influence the structure and evolution of matter in the universe, play a fundamental role in plasma astrophysics.[1] It is of interest in this connection to extend the well-known Bennett condition for equilibrium[2] to the case when the gravitational attraction forces acting between charges are comparable to the electromagnetic forces.

The actual distortion of the metric by intergalactic currents is small, and in most cases the Newtonian approximation is valid for taking account of the gravitational forces in the equations describing the equilibrium structure of the high-current channels. The equilibrium condition for high-current channels taking account of the gravitational forces was obtained in this approximation in Ref. 3. However, the equilibrium condition derived in the Newtonian approximation is applicable only for nonrelativistic drift velocities, because in order for the Newtonian approximation itself to be applicable not only must the gravitational field be weak but in addition the motion of the matter must be nonrelativistic (Ref. 4, Secs. 87 and 99). Therefore, in order to extend the high-current equilibrium condition to relativistic drift velocities the Einstein equations beyond the limits of the Newtonian approximation must be used even in the case of a weak gravitational field. The present paper is devoted to this problem.

2. PHYSICAL FORMULATION OF THE PROBLEM

We shall treat the high-current channel as a cylindrically symmetric system consisting of two subsystems (ions and electrons) moving relative to one another with drift velocity V and not interacting with one another collisionally (but only via the fields produced by the charges themselves). This is valid if the drift velocity is much larger than the average velocities of the charges in the comoving reference frames. The efficiency of Coulomb collisions decreases rapidly with increasing relative velocity of the colliding charges. For this reason, equilibrium is established within each subsystem much more rapidly than it could be for entire plasma as a whole. The mutual friction of the subsystems can be neglected over times which are much longer than the relaxation time in the subsystems but much shorter than the current decay time due to electron–ion friction. Then it can be assumed that each system is in its own state of thermal equilibrium.

If electron–ion friction changes the drift velocity very little over the length of the current channel, the current flow can be assumed to be stationary. The drift velocity along the current channel can be maintained constant by an external longitudinal electric field. As a rule, the longitudinal electric field in high-current channels is weak compared with the magnetic field generated by the current and can be neglected when analyzing equilibrium in the system.

The condition that the velocities of the charges in the comoving reference frames of the subsystems are much lower than the drift velocity means that the relative velocities in electron–electron collisions are, as a rule, nonrelativistic. Thus, the question of the nonconservation of the number of electrons as a result of electron–positron pair production does not arise.

As a whole, our formulation of the problem of equilibrium in a high-current channel is identical to the standard formulation of the problem in the ideal magnetohydrodynamic approximation for a (generally speaking, charged) plasma, irrespective of whether or not the gravitational interaction of the charges is taken into account. Since the mean free path is much smaller than the channel radius, both the electron and ion subsystems can be treated as ideal liquids (or gases). The mutual friction between the electrons and ions is a dissipative effect that arises in the higher-order terms of an expansion in the small gradients and so we shall neglect it.

3. METRIC TENSOR

To take account of gravitational forces systematically for $V/c \sim 1$, it is necessary to solve Einstein's equations, which describe the distortion of the spatial metric g_{ik} by matter (in our case, by the charges and the electromagnetic field) and include, in part, the equations of motion of the matter itself. The spatial metric should be Galilean at distances much larger than the size of the system. In what follows, we shall employ the coordinates $x_0 = ct$, $x^1 = \ln r$, $x^2 = \varphi$, and $x^3 = z$. In these coordinates the Galilean metric has the form

$$ds^2 = (dx^0)^2 - \exp(2x^1)(dx^1)^2 - \exp(2x^1)(dx^2)^2 - (dx^3)^2,$$

 1063-7761/97/080209-08$10.00

$$g_{ik}=\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & -\exp(2x^1) & 0 & 0 \\ 0 & 0 & -\exp(2x^1) & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix}. \tag{3.1}$$

The direction of the current is a distinguished direction, and to analyze the physical nature of the equilibrium in the simplest geometry the current channel can be treated as a cylindrically symmetric, stationary, and uniform (in the direction of the current) system whose parameters all depend on only one spatial coordinate, the distance from the axis. Apart from the diagonal terms, the components $g_{0\alpha}$ of the metric tensor g_{ik} are different from zero. This is because in the presence of a constant current motion of the matter and transport of momentum by both the charges and the electromagnetic field (the Poynting vector in the general case is different from zero) occur. It will be seen from the analysis below that for $V/c\sim 1$ the contribution of $g_{0\alpha}$ to the equilibrium energy balance is of the same order as g_{00} (and g_{33}). The constant gravitational field of the high-current channel is therefore stationary, but not static (Ref. 4, Sec. 88).

We shall employ the standard notation.[4] Latin indices label four-vectors and four-tensors. Greek indices are used for three-dimensional vectors and tensors; $h=g_{00}$ is a three-dimensional scalar; $g_\alpha=-h^{-1}g_{0\alpha}$ is a three-dimensional vector; and, $\gamma_{\alpha\beta}$ is the three-dimensional metric tensor

$$\gamma_{\alpha\beta}=-g_{\alpha\beta}+g_{0\alpha}g_{0\beta}/g_{00}. \tag{3.2}$$

4. DISTRIBUTION FUNCTION

The momentum and coordinate distribution function of the charges in a state of thermal equilibrium is an integral of the motion and a relativistic invariant. For this reason, in an arbitrary reference frame the distribution function should depend only on covariant combinations of the additive integrals of motion (projections of the generalized 4-momentum P_i on the directions of the cyclic coordinates) and the 4-velocity U^i of the subsystem as whole. For our cylindrically symmetric system that is uniform in the direction of the current such a combination is P_iU^i, where

$$U^0=\frac{dx^0}{\sqrt{g_{ik}dx^idx^k}}=\left[g_{00}+2g_{0\alpha}\frac{W^\alpha}{c}+g_{\alpha\beta}\frac{W^\alpha W^\beta}{c^2}\right]^{-1/2},$$

$$U^\alpha=\frac{dx^\alpha}{\sqrt{g_{ik}dx^idx^k}}=\frac{W^\alpha}{c}\left[g_{00}+2g_{0\alpha}\frac{W^\alpha}{c}+g_{\alpha\beta}\frac{W^\alpha W^\beta}{c^2}\right]^{-1/2},$$

U^i is the 4-velocity of the subsystem relative to the laboratory coordinate system and $W^\alpha=cdx^\alpha/dx^0$. The drift velocity V equals the difference of the velocities of the subsystems: $V^\alpha=W_e^\alpha-W_i^\alpha$. The generalized 4-momentum P_i in the presence of an electromagnetic field is related to the kinematic 4-momentum p_i of a particle by $p_i=P_i-(e/c)A_i$, where A_i is the 4-potential of the electromagnetic field.

The dependence of the equilibrium distribution function on P_iU^i is determined by the particle statistics. For fermions (which electrons and protons are) this dependence has the form

$$F(X)=\{1+\exp[B+X/T]\}^{-1},\quad X=cP_iU^i, \tag{4.1}$$

where B and T are scalars. Since the distribution function is an integral of the motion, its argument $B+X/T$ does not depend on the coordinates; B is simply a constant. But the projections of the generalized 4-momentum P_i on the direction of the current and on the time axis are integrals of the motion. Therefore the ratios U^i/T, $i=0$ and 3, should not depend on the coordinates. In other words, the 4-vector $\xi^i=U^i/T$ is a Killing vector,[5] which in an arbitrary coordinate system satisfies the Killing equation

$$\xi_{i;k}+\xi_{k;i}=0. \tag{4.2}$$

The Killing equation is a consequence of the symmetry associated with the existence of cyclic coordinates in the system. It determines the coordinate dependence of the velocity, chemical potential, and temperature.

It follows from the Killing equation that the 3-velocity $W^\alpha=cdx^\alpha/dx^0$ measured in universal time x^0 does not depend on the coordinate x^1. In our case x^0 and x^3 are cyclic coordinates, and the corresponding components U^0 and U^3 are different from zero. Only two of the equations (4.2) are independent:

$$g_{00}\frac{d(U^0/T)}{dx^1}+g_{03}\frac{d(U^3/T)}{dx^1}=0,$$

$$g_{30}\frac{d(U^0/T)}{dx^1}+g_{33}\frac{d(U^3/T)}{dx^1}=0.$$

The determinant of this system is different from zero:

$$\Delta=\begin{vmatrix} g_{00} & g_{03} \\ g_{30} & g_{33} \end{vmatrix}=g_{00}g_{33}-g_{03}g_{30}=-g_{00}\gamma_{33}\neq 0.$$

Therefore we have from the Killing equations

$$\frac{d(U^0/T)}{dx^1}=0,\quad \frac{d(U^3/T)}{dx^1}=0.$$

Hence $W^3/c=U^3/U^0=\text{const}$, i.e., W^3 is indeed independent of the coordinate x^1. The coordinate dependence of the temperature is given by the formula

$$\frac{T}{U^0}=T\sqrt{g_{00}+2g_{0\alpha}\frac{W^\alpha}{c}+g_{\alpha\beta}\frac{W^\alpha W^\beta}{c^2}}=\text{const}. \tag{4.3}$$

5. ENERGY-MOMENTUM TENSOR

To calculate the energy-momentum tensor T^{ik} and other quantities with the aid of the distribution function it is necessary to know how an element of the 4-volume $d\Omega_p=dp^0dp^1dp^2dp^3$ in momentum space transforms in curvilinear coordinates. Since dp^i is a covariant vector, under a coordinate transformation its components transform as differentials of the coordinates. But in curvilinear coordinates the product $\sqrt{-g}d\Omega$, where g is the determinant of the metric tensor g_{ik} and $d\Omega=dx^0dx^1x^2dx^3$, is an invariant.

Hence it can be concluded that on integrating over the 4-volume in momentum space the product $\sqrt{-g}d\Omega_p$ remains an invariant.

Thus, we have the following invariant expression for T^{ik} as an integral of the distribution function:

$$T^{ik}=\frac{2cg}{(2\pi\hbar)^3}\int p^i p^k F(X)\,\delta(p_s p^s - m^2c^2)\sqrt{-g}dp^0dp^1dp^2dp^3. \tag{5.1}$$

Substituting the expression (4.1) into Eq. (5.1) gives the standard expression for the energy-momentum tensor

$$T^{ik}=(\mathscr{E}+\mathscr{P})U^iU^k-\mathscr{P}g^{ik}, \tag{5.2}$$

in which the energy $\mathscr{E}$ and the presssure $\mathscr{P}$ in a comoving reference frame are expressed in terms of the potentials A_i of the field:

$$\mathscr{E}=\frac{4\pi g}{(2\pi ch)^3}\int_{mc^2}^{\infty}dE\;E^2\sqrt{E^2-m^2c^4}\times\left\{1+\exp\left[\frac{E-\mu}{T}\right]\right\}^{-1}, \tag{5.3}$$

$$\mathscr{P}=\frac{4\pi g}{3}\frac{1}{(2\pi ch)^3}\int_{mc^2}^{\infty}dE(E^2-m^2c^4)^{3/2}\times\left\{1+\exp\left[\frac{E-\mu}{T}\right]\right\}^{-1}, \tag{5.4}$$

$$-\mu=BT+eA_iU^i. \tag{5.5}$$

Here g is the g-factor and μ is the chemical potential.

The equations (4.3), (5.3), (5.4), and (5.5) already incorporate the equation of state of the matter. These equations express the components of the energy–momentum tensor in terms of the potentials of the electromagnetic field and the 4-velocity of the subsystem as a whole.

The current 4-vector n^i is an integral of the distribution function

$$n^i=\frac{2}{c}\frac{g}{(2\pi h)^3}\int p^i F(X)\,\delta(p_s p^s - m^2c^2)\sqrt{-g}dp^0dp^1dp^2dp^3=nU^i. \tag{5.6}$$

Here

$$n=\frac{4\pi g}{(2\pi ch)^3}\int_{mc^2}^{\infty}dE\;E\sqrt{E^2-m^2c^4}\times\left\{1+\exp\left[\frac{E-\mu}{T}\right]\right\}^{-1}, \tag{5.7}$$

where n is the charge density in the comoving coordinate system.

The complete energy–momentum tensor of the particles is obtained from Eq. (5.2) by summing over the types of charges:

$$T^k_{P_i}=\sum_a\;[(\mathscr{E}_a+\mathscr{P}_a)U_{ai}U^k_a-\mathscr{P}_a\delta^k_i]. \tag{5.8}$$

The summation index is $a=i$, e for ions and electrons, respectively. The electric current density 4-vector equals

$$j^i=c\sum_a\;e_an^i_a=c\sum_a\;e_an_aU^i_a. \tag{5.9}$$

The total current I and charge Q per unit length of the channel are obtained from Eq. (5.9) by an invariant integration over the transverse cross section

$$\frac{I}{c}=2\pi\int_{-\infty}^{\infty}dx^1\sqrt{-g}\;\sum_a\;e_an_aU^3_a,$$

$$Q=2\pi\int_{-\infty}^{\infty}dx^1\sqrt{-g}\;\sum_a\;e_an_aU^0_a. \tag{5.10}$$

The energy–momentum tensor of the electromagnetic field equals

$$T^k_{F_i}=(-F_{il}F^{kl}+\delta^k_iF_{lm}F^{lm}/4)/4\pi.$$

In the case of a stationary system which is uniform in the direction of the current and in the azimuthal direction the components A_0 and A_3 of the field-potential 4-vector are different from zero. Correspondingly, in the electromagnetic field tensor

$$F_{ik}=A_{k;i}-A_{i;k}=\frac{\partial A_k}{\partial x^i}-\frac{\partial A_i}{\partial x^k} \tag{5.11}$$

only the components $F_{10}=-F_{01}=\partial A_0/\partial x^1$ of the electric field and the components $F_{13}=-F_{31}=\partial A_3/\partial x^1$ of the magnetic field are different from zero. It is convenient to express the energy–momentum tensor of the electromagnetic field in terms of a mixed component $F^1_0=g_{00}F^{01}+g_{03}F^{31}$ of the electric field and the covariant component F^{31} of the magnetic potential

$$T_{00}=\gamma_{11}[\,\gamma_{33}h(F^{13})^2+(F^1_0)^2]/8\pi, \tag{5.12}$$

$$T^3_0=\gamma_{11}F^1_0F^{31}/4\pi, \tag{5.13}$$

$$T^{11}=h^{-1}[\,\gamma_{33}h(F^{13})^2-(F^1_0)^2]/8\pi, \tag{5.14}$$

$$T^{22}=\gamma^{22}\gamma_{11}h^{-1}[-\gamma_{33}h(F^{13})^2+(F^1_0)^2]/8\pi, \tag{5.15}$$

$$T^{33}=\gamma^{33}\gamma_{11}h^{-1}[\,\gamma_{33}h(F^{13})^2+(F^1_0)^2]/8\pi. \tag{5.16}$$

The complete matter energy–momentum tensor appearing in Einstein's equations is the sum of the corresponding expressions for the particles and the electromagnetic field: $T^k_i=T^k_{P_i}+T^k_{F_i}$.

6. RICCI TENSOR

Despite their compact tensor form, Einstein's equations

$$R_{ik}=\frac{8\pi G}{c^4}\left(T_{ik}-\frac{1}{2}\,g_{ik}T\right), \tag{6.1}$$

are difficult to use in practice. We shall write out the complete system of equations describing stationary current flow in a cylindrically symmetric channel without assuming that the gravitational field is weak. We employ the well-known

expressions for the components R_{ik} of the Ricci tensor in the case of a constant gravitational field (Ref. 4, problem in Sec. 95):

$$R_{00}=\sqrt{h}(\sqrt{h})^{;\alpha}{}_{;\alpha}+\frac{h^2}{4}f_{\lambda\mu}f^{\lambda\mu},$$

$$R_0^{\alpha}=-\frac{1}{2}hf^{\alpha\lambda}{}_{;\lambda}-\frac{3}{4}f^{\alpha\lambda}h_{;\lambda},$$

$$R^{\alpha\beta}=P^{\alpha\beta}+\frac{h}{2}f^{\alpha\lambda}f^{\beta}{}_{\lambda}-\frac{(\sqrt{h})^{;\alpha;\beta}}{\sqrt{h}}.$$

We employ the standard notation:[4]

$$f_{\alpha\beta}=g_{\beta;\alpha}-g_{\alpha;\beta}=\frac{\partial g_{\beta}}{\partial x^{\alpha}}-\frac{\partial g_{\alpha}}{\partial x^{\beta}}$$

is an antisymmetric three-dimensional tensor and $P^{\alpha\beta}$ is a three-dimensional tensor constructed from $\gamma_{\alpha\beta}$ just as the 4-tensor R^{ik} is constructed from g_{ik}. All raising and lowering operations on the Greek indices and covariant differentiation in three-dimensional space are performed with the three-dimensional metric tensor (3.2).

In our problem the tensor $\gamma_{\alpha\beta}$ is diagonal, and existing expressions (Ref. 4, problem 2 in Sec. 92) for the tensor $P^{\alpha\beta}$ can be used. Writing $g_{00}=h=\exp(F_0)$ and representing the nonzero diagonal components of the three-dimensional metric tensor in the form

$$\gamma_{\alpha\beta}=\begin{vmatrix}\exp(2F_1)&0&0\\0&\exp(2F_2)&0\\0&0&\exp(2F_3)\end{vmatrix},$$

we obtain the following expressions for the tensor $P_{\alpha\beta}$:

$$P_{\alpha\alpha}=\sum_{\lambda\neq\alpha}\left\{F_{\alpha,\alpha}F_{\lambda,\alpha}-F_{\lambda,\alpha}^2-F_{\lambda,\alpha,\alpha}+\exp(2F_{\alpha}-2F_{\lambda})\right.$$
$$\times\left(F_{\lambda,\lambda}F_{\alpha,\lambda}-F_{\alpha,\lambda,\lambda}-F_{\alpha,\lambda}^2\right.$$
$$\left.\left.-\sum_{\mu\neq\alpha\neq\lambda}F_{\alpha,\lambda}F_{\mu,\lambda}\right)\right\},\qquad(6.2)$$

$$P_{\alpha,\beta}=\sum_{\mu\neq\alpha\neq\beta}(F_{\alpha,\beta}F_{\mu,\alpha}+F_{\beta,\alpha}F_{\mu,\beta}-F_{\mu,\beta}F_{\mu,\alpha}$$
$$-F_{\mu,\alpha,\beta}),\quad\alpha\neq\beta.\qquad(6.3)$$

In Eq. (6.2) summation is not performed over the doubly repeated index α. In the case when all quantities depend on only one coordinate the tensor $P_{\alpha\beta}$ is diagonal, since for $\alpha\neq\beta$ the differentiation in each term of the sum in Eq. (6.3) is performed with respect to two different coordinates. The index α signifies ordinary differentiation with respect to the coordinate x^{α}.

Making use of the arbitrariness in the choice of the coordinate x^1, we impose on the function F_i the additional condition[6]

$$F_1=F_0+F_2+F_3.\qquad(6.4)$$

Then the formulas greatly simplify, and we arrive at the following expressions for the nonzero components of the Ricci tensor:

$$R_{00}=\exp[2(F_0-F_1)]F_0''+(1/2)\exp(4F_0-2F_1$$
$$-2F_3)g_3'^2,\qquad(6.5)$$

$$R_0^3=-(1/2)\exp(-2F_1)[\exp\{2(F_0-F_3)\}g_3']',\qquad(6.6)$$

$$R^{11}=\exp(-4F_1)\{[2(F_2'F_3'+F_3'F_0'+F_0'F_2')-F_1'']$$
$$+(1/2)\exp(2F_0-2F_3)g_3'^2\},\qquad(6.7)$$

$$R^{22}=-\exp(-2F_2-2F_1)F_2'',\qquad(6.8)$$

$$R^{33}=\exp(-2F_3-2F_1)F_3''+(1/2)\exp(2F_0-2F_1$$
$$-4F_3)g_3'^2.\qquad(6.9)$$

A prime denotes differentiation with respect to x^1.

7. CONSERVATION LAWS

It follows from Einstein's equations that the divergence of the energy–momentum tensor

$$T_i^k=\sum_a(\mathscr{W}_aU_{ai}U_a^k-\mathscr{P}_a\delta_i^k)+(-F_{il}F^{kl}$$
$$+\delta_i^kF_{lm}F^{lm}/4)/4\pi$$

equals zero: $T_{i;k}^k=0$. We calculate $T_{i;k}^k$ directly with $\mathscr{W}_a=\mathscr{E}_a+\mathscr{P}_a$, $\mathscr{E}_a$, and $\mathscr{P}_a$ from Eqs. (5.3) and (5.4). The nonzero components U_a^k are U_a^0 and U_a^3, and the corresponding coordinates x^0 and x^3 are cyclic. Therefore $U_a^kd/dx^k=0$, $U_{a;k}^k=0$. A series of standard transformations yields

$$T_{i;k}^k=\sum_a\mathscr{W}_a(U_{ai}U_a^k)_{;k}-\mathscr{P}_{a;i}-F_{il}(F^{kl})_{;k}/4\pi.$$

Now we transform $\mathscr{P}_{a;i}=d\mathscr{P}_a/dx^i=\mathscr{P}_{a,i}$:

$$\frac{d\mathscr{P}}{dx^i}=\sum_a\frac{4\pi}{3}\frac{g_a}{(2\pi ch)^3}\int_{mc^2}^{\infty}dE(E^2-m_a^2c^4)^{3/2}$$
$$\times\frac{d}{dx^i}\left\{1+\exp\left[\frac{E-\mu_a}{T_a}\right]\right\}^{-1}.$$

The identity

$$\frac{d}{dx^i}\left\{1+\exp\left(\frac{E-\mu}{T}\right)\right\}^{-1}=\left[E\left(\frac{1}{T}\right)_{,i}-\left(\frac{\mu}{T}\right)_{,i}\right]$$
$$\times\frac{d}{dE}\left\{1+\exp\left(\frac{E-\mu}{T}\right)\right\}^{-1},$$

and integration by parts yield

$$\mathscr{P}_{a,i}=\frac{4\pi}{3}\sum_a\frac{g_aT_a}{(2\pi\hbar)^3}\int_{mc^2}^{\infty}\frac{dE}{1+\exp[(E-\mu_a)/T_a]}$$
$$\times\frac{d}{dE}\left\{(E^2-m_a^2c^4)^{3/2}\left[E\left(\frac{1}{T_a}\right)_{,i}-\left(\frac{\mu_a}{T_a}\right)_{,i}\right]\right\}.$$

Differentiating with respect to E and using the equations (5.3), (5.4), and (5.6) we obtain

$$\frac{d\mathscr{P}}{dx^i}=-\sum_a T_a\left\{\mathscr{W}_a\left(\frac{1}{T_a}\right)_{,i}-\left(\frac{\mu_a}{T_a}\right)_{,i}n_a\right\}.$$

Now, using Eqs. (5.5) and $B_{,i}=0$ we have

$$\frac{d\mathscr{P}}{dx^i}=-\sum_a T_a\left\{\mathscr{W}_a\left(\frac{1}{T_a}\right)_{,i}+e_a n_a\left[\left(\frac{U_a^k}{T_a}\right)A_{k,i}+A_k\left(\frac{U_a^k}{T_a}\right)_{,i}\right]\right\}.$$

The term $A_k(U_a^k/T_a)_{,i}$ vanishes on account of the Killing equations: $(U_a^k/T_a)_{,i}=0$. Since $U^k d/dx^k=0$, we can write $A_{k,i}U^k=(A_{k,i}-A_{i,k})U^k=F_{ik}U^k$. On the basis of Eq. (5.9), we have

$$\frac{d\mathscr{P}}{dx^i}=-\sum_a T_a\mathscr{W}_a\left(\frac{1}{T_a}\right)_{,i}-F_{ik}\frac{j^k}{c},$$

$$T^k_{i;k}=\sum_a \mathscr{W}_a T_a\left\{\frac{U_{ai,k}U_a^k}{T_a}+\left(\frac{1}{T_a}\right)_{,i}\right\}+F_{il}\left\{(F^{lk})_{;k}+\frac{4\pi j^l}{c}\right\}\frac{1}{4\pi}.$$

The relations $U_{ak;i}U_a^k=0$, $U_a^k U_{ak}=1$, $U_a^k(1/T_a)_{;k}=0$, and $U^k_{a;k}=0$ make it possible to write the sum over the types of charges in the form

$$\sum_a \mathscr{W}_a T_a U_a^k\left\{\left(\frac{U_{ai}}{T_a}\right)_{;k}+\left(\frac{U_{ak}}{T_a}\right)_{;i}\right\},$$

whence it follows by virtue of the Killing equations (4.2) that this sum vanishes. As a result we obtain

$$T^k_{i;k}=\frac{1}{4\pi}F_{im}\left\{F^{mk}{}_{;k}+\frac{4\pi j^m}{c}\right\}.$$

Therefore it follows from Einstein's equations that

$$F_{im}\{F^{mk}{}_{;k}+4\pi j^m/c\}=0. \tag{7.1}$$

It can be concluded that Maxwell's equations

$$F^{mk}{}_{;k}+4\pi j^m/c=0 \tag{7.2}$$

are wholly contained in Einstein's equations only if the determinant of the matrix of the electromagnetic field is different from zero: $\det F_{im}\neq 0$. In our case the electric and magnetic fields are perpendicular to one another, so that one of the two field invariants equals zero. This is the special case when Maxwell's equations are not a consequence of Einstein's equations.[7] Since in our case $\det F_{ik}=0$, there exists a nontrivial solution of Eq. (7.1) and hence the expressions $F^{km}{}_{;m}+4\pi j^k/c$, generally speaking, are different from zero. Of the four equations (7.1), in our case ($F_{2k}=0$, $j^1=0$, $F^{1m}{}_{;m}=0$, and the rank of the matrix F_{ik} equals 1) there actually remains only one equation

$$F_{10}(F^{0m}{}_{;m}+4\pi j^0/c)+F_{13}(F^{3m}{}_{;m}+4\pi j^3/c)=0. \tag{7.3}$$

In the presence of a relative motion of the electron and ion subsystems $F_{13}\neq 0$ holds in any reference frame. Then it follows from Eq. (7.3) that Einstein's equations hold for an arbitrary function $F^{0m}{}_{;m}+4\pi j^0/c$ if

$$F^{3m}{}_{;m}+\frac{4\pi j^3}{c}=-\left(F^{0m}{}_{;m}+\frac{4\pi j^0}{c}\right)\frac{F_{10}}{F_{13}}.$$

On the other hand, the fact that the electromagnetic field, whose potentials A_i appear explicitly in the distribution function, is produced by the charges of our system themselves is expressed mathematically by Maxwell's equations (7.2), which relate the field potentials to the charge and current density.

In summary, of the two Maxwell equations $F^{ik}{}_{;k}=-4\pi j^i/c$, $i=0,3$, one is independent and the other is a consequence of the first one and Einstein's equations.

8. COMPLETE SYSTEM OF EQUATIONS

Substituting the expressions (5.8) for the particle energy–momentum tensor, the equations (5.12)–(5.16) for the energy–momentum tensor of the electromagnetic field, and the equations (6.10)–(6.14) for the Ricci tensor into Eq. (6.1) we arrive at the following gravitational field equations:

$$\exp(2F_1-2F_0)R_{00}=F_0''+\frac{1}{2}\exp[2(F_0-F_3)]g_3'^2=\frac{8\pi G}{c^4}\times\exp(2F_1)\left\{\exp(2F_1)[\exp(2F_3)(F^{13})^2+\exp(-2F_0)(F_0^1)^2]\frac{1}{8\pi}+\sum_a\left[\frac{(1-g_3W_a/c)^2(\mathscr{E}_a+\mathscr{P}_a)}{(1-g_3W_a/c)^2-\exp(2F_3-2F_0)W_a^2/c^2}-\frac{\mathscr{E}_a-\mathscr{P}_a}{2}\right]\right\}, \tag{8.1}$$

$$\exp(2F_1)R_0^3=-\frac{1}{2}[\exp\{2(F_0-F_3)\}g_3']'=\frac{8\pi G}{c^4}\exp(2F_1)\times\left\{\sum_a\frac{(W_a/c)(1-g_3W_a/c)(\mathscr{E}_a+\mathscr{P}_a)}{(1-g_3W_a/c)^2-\exp(2F_3-2F_0)W_a^2/c^2}+\exp(2F_1)\frac{F_0^1F^{31}}{4\pi}\right\}, \tag{8.2}$$

$$\exp(4F_1)R^{11}=[2(F_2'F_3'+F_3'F_0'+F_0'F_2')-F_1'']+\frac{1}{2}\exp(2F_0-2F_3)g_3'^2=\frac{8\pi G}{c^4}\exp(2F_1)\left\{\sum_a\frac{\mathscr{E}_a-\mathscr{P}_a}{2}+\exp(2F_1)[\exp(2F_3)(F^{13})^2-\exp(-2F_0)\times(F_0^1)^2]\frac{1}{8\pi}\right\}, \tag{8.3}$$

$$\exp(2F_1+2F_2)R^{22}=-F_2''=\frac{8\pi G}{c^4}\exp(2F_1)\Bigg\{\sum_a \frac{\mathscr{E}_a-\mathscr{P}_a}{2} + \exp(2F_1)[-\exp(2F_3)(F^{13})^2 + \exp(-2F_0)(F_0^1)^2]\,\frac{1}{8\pi}\Bigg\}, \tag{8.4}$$

$$\exp(2F_1+2F_3)R^{33} = -F_3''+\frac{1}{2}\exp[2(F_0-F_3)]g_3'^2=\frac{8\pi G}{c^4} \times\exp(2F_1)\Bigg\{\exp(2F_1)[\exp(2F_3)(F^{13})^2 + \exp(-2F_0)(F_0^1)^2]\,\frac{1}{8\pi} + \sum_a\left[\frac{\exp(2F_3-2F_0)(W_a/c)^2(\mathscr{E}_a+\mathscr{P}_a)}{(1-g_3W_a/c)^2-\exp(2F_3-2F_0)W_a^2/c^2} + \frac{\mathscr{E}_a-\mathscr{P}_a}{2}\right]\Bigg\}. \tag{8.5}$$

The five equations (8.1)–(8.5) contain five functions describing the gravitational field (F_0,F_1,F_2,F_3,g_3) and two functions describing the electromagnetic field (A_0 and A_3), which enter into the equations via the expressions (5.3), (5.4), and (5.11). There are seven functions in all. The system is closed by the relation (6.4) and one Maxwell equation

$$\exp(-2F_1)[\exp(2F_1)F^{01}]' = -4\pi\sum_a e_a n_a \times\frac{\exp(-F_0)}{\sqrt{(1-g_3W_a/c)^2-\exp(2F_3-2F_0)W_a^2/c^2}}. \tag{8.6}$$

The second Maxwell equation

$$\exp(-2F_1)[\exp(2F_1)F^{31}]' = -4\pi\sum_a e_a n_a \times\frac{\exp(-F_0)W_a/c}{\sqrt{(1-g_3W_a/c)^2-\exp(2F_3-2F_0)W_a^2/c^2}} \tag{8.7}$$

is, as shown above, a consequence of this system of equations. The charge densities n_a and the temperature of the charges are expressed in terms of the same unknown functions by the equations (5.7) and (4.3).

The relations

$$F_0''+F_3''=F_1''-F_2''=\frac{16\pi G}{c^4}\exp(2F_1)\sum_a \mathscr{P}_a, \tag{8.8}$$

$$F_2'(F_3'+F_0')+F_3'F_0'+\frac{1}{4}\exp(2F_0-2F_3)g_3'^2 = \frac{8\pi G}{c^4}\exp(2F_1)\Bigg\{\sum_a \mathscr{P}_a+\exp(2F_1)[\exp(2F_3) \times(F^{13})^2-\exp(-2F_0)(F_0^1)^2]\,\frac{1}{8\pi}\Bigg\}. \tag{8.9}$$

follow from Eqs. (8.1)–(8.5). The relation (8.9) does not contain second derivatives and is essentially the first integral of the system of equations.

To relate the coordinate x^1 to the distance from the axis of the cylinder, we ''switch off'' the electromagnetic field ($F_0^1=F^{13}=0$) and ''stop'' the charges ($W_a=0$). Then

$$F_3''=F_2''\ (F_0^1=F^{13}=0,\quad W_a=0).$$

follows from Eqs. (8.4) and (8.5). If the integration constant is set equal to 1, $F_2'=F_3'+1$, then far from the matter the curvilinear cyclic coordinates x^2 and x^3 will correspond to the cylindrical coordinates φ and z, provided that the coordinate x^1 and the radius are related by $x^1=\ln r$. This corresponds to the Galilean metric (3.1). The direction $x^1\to\infty$ is the radial direction away from the current channel and $x^1=-\infty$ corresponds to the axis of the cylinder.

For an infinitely long system the metric does not become Galilean at distances much larger than the transverse dimensions. As $x^1\to\infty$ the functions F_i satisfying the equations (8.1)–(8.5) do not approach zero. The problem is that the cylindrical symmetry of a system in which there is a dependence on only one coordinate x^1 is an intermediate asymptotic property for distances from the axis which are much smaller than the length L of the current channel. At distances x^1 which are much larger than L the metric will be Galilean but at such distances the equations (8.1)–(8.5) are no longer applicable because x^3 is no longer a cyclic coordinate, and the system now depends on at least two variables (x^1 and x^3).

The finiteness of the current and charge per unit length presupposes that the integrals (5.11) converge. Then $\mathscr{E}_a$, $\mathscr{P}_a$, and n_a as functions of x^1 rapidly approach zero away from the axis:

$$\mathscr{E}_a=0,\quad \mathscr{P}_a=0,\quad n_a=0,\quad x^1\to\infty. \tag{8.10}$$

From Maxwell's equations (8.6) and (8.7) we find

$$\exp(2F_1)F^{01}=-2Q,\quad \exp(2F_1)F^{31}=-2I/c,\quad x^1\to\infty. \tag{8.11}$$

9. ENERGY BALANCE

Equilibrium is possible in the current channel if the magnetic and gravitational compression energies are balanced by the electrostatic repulsion of the space charge and the thermal expansion energy of the particles:

$$\frac{I^2}{2c^2}+\frac{G}{2}\left(\sum_a N_a m_a\right)^2=\frac{Q^2}{2}+\sum_a N_a T_a. \tag{9.1}$$

This formula was previously derived[3] under the assumption that the electrons and ions are ideal Boltzmann gases, and the

gravitational interaction of the charges is described by Newton's equations. This means that Eq. (9.1) is applicable provided that the velocity of the relative motion of the electron and ion subsystems is much smaller than the velocity of light. To extend this balance to the case $W_a \sim c$ it was necessary to go beyond the Newtonian approximation and turn to Einstein's equations.

The condition of balance in the general theory of relativity can be obtained from the first integral (8.9) by simply taking the limit $x^1 \to \infty$ in it. Since $F_0^1 = \exp(2F_0) \times (F^{01} - g_3 F^{31})$ and taking account of Eqs. (8.10) and (8.11), we find

$$\frac{1}{4}\exp(2F_0 - 2F_3) g_3'^2 + F_0' F_3' + F_2'(F_3' + F_0') = 4\frac{GI^2}{c^6}\left[\exp(2F_3) - \exp(2F_0)\left(\frac{cQ}{I} - g_3\right)^2\right]. \quad (9.2)$$

Integrating the relation (8.8), we find

$$F_0' + F_3' = \frac{8G}{c^4}\sum_a \mathscr{P}_a, \quad (9.3)$$

$$\mathscr{P}_a = 2\pi \int_{-\infty}^{\infty} dx^1 \exp(2F_1)\mathscr{P}_a, \quad (9.4)$$

$\mathscr{P}_a$ is the pressure of charges of type a, integrated over the transverse cross section of the current channel.

The first two terms on the left-hand side of Eq. (9.2) are proportional to G^2. If gravitation is completely neglected, i.e., Eq. (9.2) is reduced by a general factor proportional to G, after which G is set equal to zero and the Galilean values are substituted for F_i, then we obtain a condition for equilibrium in the current channel that is valid, as follows from our derivation, for Fermi gases of electrons and ions with an arbitrary degree of degeneracy (and not only in the approximation of ideal Boltzmann gases):

$$\sum_a \mathscr{P}_a = \frac{I^2}{2c^2} - \frac{Q^2}{2}.$$

Gravitation must be included in the general balance of the forces when the electromagnetic contribution is small and the total number of particles is enormous, so that the quantities

$$\frac{I^2}{c^2} - Q^2 \quad \text{and} \quad G\left(\sum_a N_a m_a\right)^2$$

are of the same order of magnitude. Such a situation can occur in intergalactic currents either when the current and charge are both not very large or if the compression by the magnetic field generated by the current is compensated, to a high degree of accuracy, by the electrostatic repulsion of the space charge. In the latter case the drift velocity can be relativistic, and then the Newtonian approximation is inadequate.

For simplicity, let us assume that

$$\mathscr{P}_a \ll \mathscr{E}_i. \quad (9.5)$$

Since $m_e \ll m_i$, the electrons can be relativistic and even ultrarelativistic. The gravitational field of the system is weak if $GN_a m_a/c^2 \ll 1$ and $GI^2/c^6 \ll 1$. The first parameter characterizes quantitatively the distortion of the metric by the charges and the second parameter characterizes the distortion of the metric by the electromagnetic field produced by these charges. In a weak gravitational field, for $(I/c)^2 - Q^2 \sim G(m_a N_a)^2$ the distortion of the metric by the electromagnetic field is much smaller than the distortion by the charges

$$\frac{G}{c^4}\left(\frac{I^2}{c^2} - Q^2\right) \sim \frac{(Gm_a N_a)^2}{c^4} \ll \frac{GN_a m_a}{c^2}.$$

In this case, in the expansion in terms of the parameter $1/c^2$ the distortion of the metric by the charges is an effect of order $1/c^2$ and the distortion by the electromagnetic field is of the order of $1/c^4$.

At sufficiently large distances from the axis of the current channel the functions F_0, F_3, and g_3 (but not their derivatives!) in Eq. (9.2) can be replaced by their Galilean values—zero. On account of the condition (9.5) F_2' can also be replaced by its Galilean value $F_2' = 1$. We have

$$\frac{1}{4} g_3'^2 + F_0' F_3' + \frac{8G}{c^4}\sum_a \mathscr{P}_a = 4\frac{GI^2}{c^6}\left[1 - \left(\frac{cQ}{I}\right)^2\right]. \quad (9.6)$$

In the term $F_0' F_3'$ it is sufficient to obtain both factors, just as g_3', to first order in $1/c^2$. To this accuracy, we find from Eqs. (8.1), (8.2), and (8.5)

$$F_0' = -F_3' = \frac{2G}{c^4}\sum_a \mathscr{E}_a \frac{1 + W_a^2/c^2}{1 - W_a^2/c^2}, \quad (9.7)$$

$$g_3' = \frac{8G}{c^4}\sum_a \mathscr{E}_a \frac{W_a/c}{1 - W_a^2/c^2}, \quad (9.8)$$

$$\mathscr{E}_a = 2\pi \int_{-\infty}^{\infty} dx^1 \exp(2F_1)\mathscr{E}_a,$$

where $\mathscr{E}_a$ is the energy of the subsystem of type a per unit length of the channel. Substituting the expressions (9.7) and (9.8) into Eq. (9.6) we obtain

$$\frac{G}{2}\sum_a \frac{\mathscr{E}_a}{c^2}\frac{1 + W_a/c}{1 - W_a/c}\sum_b \frac{\mathscr{E}_b}{c^2}\frac{1 - W_b/c}{1 + W_b/c} + \frac{I^2}{2c^2} = \frac{Q^2}{2} + \sum_a \mathscr{P}_a,$$

whence

$$\frac{G}{2}\left(\sum_a \frac{\mathscr{E}_a}{c^2}\right)^2 + \frac{2G\mathscr{E}_e\mathscr{E}_i}{c^6}\frac{(W_e - W_i)^2}{(1 - W_e^2/c^2)(1 - W_i^2/c^2)} + \frac{I^2}{2c^2} = \frac{Q^2}{2} + \sum_a \mathscr{P}_a. \quad (9.9)$$

The first two terms in Eq. (9.9), which are proportional to G, are the gravitational compression energy. The first term is the standard Newtonian expression for the gravitational compression energy per unit length (compare to Eq. (9.1)). The second term on the left-hand side of Eq. (9.9) describes the additional gravitational attraction arising between the subsystems as a result of their relative motion.

This effect is not contained in the Newtonian approximation. It is negligibly small in the limit $V/c \ll 1$, but it can dominate in the case of ultrarelativistic drift velocities. The equation (9.9) is valid irrespective of the degree of degeneracy of the Fermi gas of electrons and/or ions. The inequality (9.5) was used.

The gravitational attraction arising between the subsystems as a result of their relative motion in high-current channels contains an additional smallness on account of the inequality $m_e \ll m_i$. This effect should be much stronger in superfluid stars, since there the masses of the normal and superfluid components can be of the same order of magnitude, there is no mutual friction between these subsystems, and the relative velocity of the subsystems is not small.

For nonrelativistic Boltzmann gases ($\mathscr{P}_a = N_a T_a$, $\mathscr{E}_a = N_a m_a c^2$, N_a is the number of charges of type a per unit length of the current channel) the equation (9.9) gives

$$\frac{I^2}{2c^2} + \frac{G}{2}\left(\sum_a N_a m_a\right)^2 + \frac{2GN_eN_im_em_i}{c^2} \times \frac{(W_e - W_i)^2}{(1 - W_e^2/c^2)(1 - W_i^2/c^2)} = \frac{Q^2}{2} + \sum_a N_a T_a . \quad (9.10)$$

In the limit of nonrelativistic drift velocities the equation (9.10) passes into the equation (9.1).

I thank D. A. Kirzhnits for a discussion.

[1] A. I. Peratt, *Physics of the Plasma Universe*, Springer-Verlag, N.Y., 1992.
[2] W. H. Bennett, Phys. Rev. **45**, 890 (1934).
[3] B. E. Meierovich and A. L. Peratt, IEEE Trans. Plasma Sci. **20**, 891 (1992).
[4] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, Pergamon Press, N.Y. [Russian original, Nauka, Moscow, 1973].
[5] J. Ehlers, in *Relativity and Gravitation*, Gordon and Breach Science Publishers, London, 1971, p. 150.
[6] K. A. Bronnikov, J. Phys. A: Math. Gen. **12**, 201 (1979).
[7] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*, W. H. Freeman, San Francisco, CA, 1973, Sec. 20.

Translated by M. E. Alferieff

Small angle Bhabha scattering at LEP1. Wide-narrow angular acceptance

N. P. Merenkov

National Science Center Kharkov Institute of Physics and Technology, 310108 Kharkov, Ukraine
(Submitted 19 November 1996)
Zh. Éksp. Teor. Fiz. **112**, 400–428 (August 1997)

Analytical method is applied for the description of small-angle Bhabha scattering at LEP1. Inclusive event selection for asymmetrical wide-narrow circular detectors is considered. The QED correction to the Born cross section is calculated with leading and next-to-leading accuracy in second order of perturbation theory and with leading accuracy in third order. All contributions in the second order due to the photonic radiative corrections and pair production are calculated starting from essential Feynman diagrams. Third-order correction is computed by means of the electron structure function method. Second- and third-order leading corrections suitable for calorimeter event selection are investigated. Numerical results illustrate the analytical calculations. © *1997 American Institute of Physics.* [S1063-7761(97)00208-4]

1. INTRODUCTION

The small-angle Bhabha scattering (SABS) process is used to measure the luminosity of electron-positron colliders. At LEP1 an experimental accuracy on the luminosity of $\delta\sigma/\sigma<0.1\%$ has been reached.[1] However, to obtain the total accuracy, a systematic theoretical error must also be added. Accurate determination of the SABS cross section therefore directly affects some physical values measured at LEP1 experiments.[2,3] Considerable attention has therefore been recently devoted to the Bhabha process.[3–11] The accuracy that has been attained, however, is still inadequate. According to these evaluations, the theoretical estimates are still incomplete and their accuracy is far from that which is required.

The theoretical calculation of SABS cross section at LEP1 involves two slightly different problems. The first one is the description of experimental restrictions used for event selection in terms of final particles phase space. The second consists in the writing of matrix element squared with the required accuracy. There are two methods for theoretical investigation of SABS at LEP1: a method based on Monte Carlo calculation[3–5,7] and analytical method.[6,9–11]

The advantage of the Monte Carlo method is that it can model different types of detectors and event selection.[3] This method, however, cannot use the exact matrix element squared based on essential Feynman diagrams because of the infrared divergence. Therefore, some additional procedures (YFS factor exponentiation,[12] utilization of electron structure functions[13]) must be used in order to eliminate this problem and to take into account the leading contribution in the higher orders. Careful attention must be given at this point because of the possibility of double counting. In any case, the next-to-leading second-order correction remains uncertain, and this is a transparent defect of the Monte Carlo method.

The advantage of the analytical method is that it can use the exact matrix element squared. The infrared problem in the context of this approach can be solved in the usual way by taking into account the virtual, real soft, and hard photon emission and pair production in every order of perturbation theory. Any questions about double counting do not arise in analytical calculations. The defect of this method is its low mobility relative to the change in the experimental conditions for event selection. Nevertheless, the analytical calculations are very important because they allow one to check many Monte Carlo calculations for different ''ideal'' detectors.

Analytical formula for SABS cross sections at LEP1 until now were published for inclusive event selection (IES) when circular symmetrical detectors record only final electron and positron energies.[10,11] They define first- and second-order corrections to the Born cross section with leading [of the order of $(\alpha L)^n$] and next-to-leading (of the order of $\alpha^n L^{n-1}$) accuracy, as well as third-order correction with leading accuracy only. These contributions will have to be computed in order to reach the required per mille accuracy for SABS cross section at LEP1. Note that such an accuracy selects only collinear (like two-jets, final-state configuration) and semi-collinear (like three-jets configuration) kinematics.

The case of calorimeter event selection (CES) called in Ref. 3 CALO1 and CALO2 for symmetrical and wide-narrow angular acceptance, was considered by the author. The results are being prepared for publication. They include the first-order correction with leading and next-to-leading accuracy, as well as second- and third-order corrections with leading accuracy only. Thus, the CES problem of the analytical method is the calculation of next-to-leading, second-order correction.

In this paper we perform full analytical calculation for IES with the wide-narrow angular acceptance. The first- and second-order corrections are derived with next-to-leading accuracy starting from the Feynman diagrams for two-loop elastic electron-positron scattering, one-loop single-photon emission, two-photon emission and pair production. The third-order correction is obtained with leading accuracy with the help of the electron structure function method. The results for leading second- and third-order corrections in the case of CES are also given.

The contents of this paper can be outlined as follows. In Sec. 2 we introduce the ''observable'' cross section $\sigma_{\rm exp}$, with allowance for the cuts at angles and energies, and obtain the first-order correction. In Sec. 3 we investigate the

second-order corrections. They include the contributions of the processes of pair production (real and virtual) considered in Sec. 3.1 and two-photon (real and virtual) emission. In Sec. 3.2 we consider the correction due to the one-side, two-photon emission and in Sec. 3.3 we consider the correction due to opposite side, two-photon emission. The expression for second-order photonic correction is given in leading approximation only, while next-to-leading contribution to it is written in Appendix A for symmetrical and wide-narrow detectors. The latter does not contain an auxiliary infrared parameter. In Sec. 4 we derive the full, leading, third-order correction using the expansion of electron structure functions. In Sec. 5 we present the numerical results suitable for IES. In Appendix B we give some relations which were used in the analytical calculations and which are very useful for numerical calculation.

2. FIRST-ORDER CORRECTION

We introduce the dimensionless quantity

$$\Sigma=\frac{1}{4\pi\alpha^2}\,Q_1^2\sigma_{\exp}, \tag{1}$$

where $Q_1^2=\epsilon^2\theta_1^2$ (ϵ is the beam energy, and θ_1 is the minimal angle of the wide detector). The ''experimentally'' measurable cross section $\sigma_{\exp}$ is defined as follows:

$$\sigma_{\exp}=\int dx_1dx_2\Theta d^2q_1^\perp d^2q_2^\perp\Theta_1^c\Theta_2^c \times\frac{d\sigma(e^++e^-\to e^++e^-+X)}{dx_1dx_2d^2q_1^\perp d^2q_2^\perp}, \tag{2}$$

where X is undetected final particles, and $x_1(x_2)$ and $q_1^\perp(q_2^\perp)$ are the energy fraction and the transverse component of the momentum of the electron (positron) in the final state. The functions Θ_i^c take into account the angular cuts and the function Θ takes into account the cutoff on the invariant mass of the detected electron and positron:

$$\Theta_1^c=\theta(\theta_3-\theta_-)\,\theta(\theta_--\theta_1),$$
$$\Theta_2^c=\theta(\theta_4-\theta_+)\,\theta(\theta_+-\theta_2),\quad \Theta=\theta(x_1x_2-x_c),$$
$$\theta_-=\frac{|\mathbf{q}_1^\perp|}{x_1\epsilon},\quad \theta_+=\frac{|\mathbf{q}_2^\perp|}{x_2\epsilon}. \tag{3}$$

In the case of symmetrical angular acceptance

$$\theta_2=\theta_1,\quad \theta_3=\theta_4,\quad \rho=\frac{\theta_3}{\theta_1}>1,$$

but for wide-narrow acceptance

$$\theta_3>\theta_4>\theta_2>\theta_1,\quad \rho_i=\frac{\theta_i}{\theta_1}>1.$$

For numerical calculation one usually takes

$$\theta_1=0.024,\quad \theta_3=0.058,$$
$$\theta_2=0.024+\frac{0.034}{16},\quad \theta_4=0.058-\frac{0.034}{16}.$$

The first-order correction Σ_1 includes the contributions of the virtual and real soft and hard photon-emission processes

$$\Sigma_1=\Sigma_{V+S}+\Sigma^H+\Sigma_H. \tag{4}$$

The contribution due to the virtual and real soft photon (with energy less than $\Delta\epsilon$, $\Delta\ll1$) can be written as follows (in this case $x_1=x_2=1$, $\mathbf{q}_1^\perp+\mathbf{q}_2^\perp=0$):

$$\Sigma_{V+S}=2\,\frac{\alpha}{\pi}\int_{\rho_2^2}^{\rho_4^2}\frac{dz}{z^2}\left[2(L-1)\ln\Delta+\frac{3}{2}L-2\right],$$
$$L=\ln\frac{\epsilon^2\theta_1^2z}{m^2}, \tag{5}$$

where $z=\mathbf{q}_2^{\perp2}/Q_1^2$, and m is the electron mass.

The second term on the right side of Eq. (4) represents the correction due to the hard photon emission by the electron. In this case we have

$$X=\gamma(1-x_1,\mathbf{k}^\perp),\quad x_2=1,\quad \mathbf{k}^\perp+\mathbf{q}_1^\perp+\mathbf{q}_2^\perp=0,$$
$$x_c<x_1<1-\Delta. \tag{6}$$

This expression can be derived by integration of the differential cross section of single-photon emission over the region

$$\rho_2^2<z<\rho_4^2,\quad x^2<z_1=\frac{\mathbf{q}_1^{\perp2}}{Q_1^2}<x^2\rho_3^2,\quad -1<\cos\varphi<1, \tag{7}$$

where φ is the angle between the vectors $\mathbf{q}_1^\perp$ and $\mathbf{q}_2^\perp$, in the same way as it was done in Ref. 10 for the symmetrical angular acceptance. But at this point we would like to indicate the main features of the method which is used largely in the Sec. 3 and which is based on the separate calculation of the contributions due to collinear kinematics and semi-collinear kinematics.[14]

In collinear kinematics an emitted photon moves inside the cone within polar angles $\theta_\gamma<\theta_0\ll1$ centered along the electron momentum direction (initial: $\mathbf{k}\|\mathbf{p}_1$ or final: $\mathbf{k}\|\mathbf{q}_1$). In semi-collinear region a photon moves outside this cone. Because such a distinction no longer has physical meaning, the dependence on the auxiliary parameter θ_0 disappears in the total contribution. This is valid for IES and for CES.

Inside collinear kinematics it is necessary to keep the electron mass in the differential cross section

$$d\sigma=\frac{2\alpha^3s}{\pi^2q^2}\left[\frac{1+x^2}{s_1t_1}-\frac{2m^2}{q^2}\left(\frac{1}{s_1^2}+\frac{x^2}{t_1^2}\right)\right]d\Gamma,$$
$$d\Gamma=\frac{d^3q_1d^3q_2d^3k}{\epsilon_1\omega2\epsilon}\,\delta^{(4)}(p_1+p_2-k-q_1-q_2), \tag{8}$$

where $q=p_1-k-q_1$, $s_1=2(kq_1)$, $t_1=2(kp_1)$, $s=(2p_1p_2)$, and $p_1(p_2)$ is the 4-momentum of the initial electron (positron). If the photon moves inside the initial electron cone

$$s_1=x(1-x)\epsilon^2\theta_-^2,\quad t_1=-m^2(1-x)(1+\eta),$$
$$q^2=-x^2\epsilon^2\theta_-^2=-\epsilon^2\theta_+^2,$$

$$d\Gamma=\frac{m^2}{s}\,\epsilon^2\pi^2x(1-x)dx\;d\eta\;d\theta_-^2,\quad 0<\eta=\frac{\theta_\gamma^2\epsilon^2}{m^2}<\frac{\theta_0^2\epsilon^2}{m^2},\tag{9}$$

and one can derive the following expression after integration over η:

$$\sigma_{\mathbf{k}\|\mathbf{p}_1}=\frac{2\alpha^3}{Q_1^2}\int_{\rho_2^2}^{\rho_4^2}\frac{dz}{z^2}\int_{x_c}^{1-\Delta}dx\left[\frac{1+x^2}{1-x}\ln\frac{\theta_0^2\epsilon^2}{m^2}-\frac{2x}{1-x}\right]\theta(x^2\rho_3^2-z).\tag{10}$$

The right side of Eq. (10) corresponds to the contribution of the narrow strip with the width $2\sqrt{z}\lambda(l-x)$ centered around the line $z=z_1$ in (z,z_1) plane, where $\lambda=\theta_0/\theta_1$. In fact, the condition $\theta_\gamma<\theta_0$ for the initial electron cone can be formulated as follows:

$$|\sqrt{z}-\sqrt{z_1}|<\lambda(1-x),$$

$$-1<\cos\,\varphi<-1+\frac{\lambda^2(1-x)^2-(\sqrt{z_1}-\sqrt{z})^2}{2\sqrt{z_1z}}.\tag{11}$$

If photon moves inside the final electron cone

$$s_1=\frac{1-x}{x}\,m^2(1+\zeta),\quad t_1=-(1-x)\epsilon^2\theta_-^2,$$

$$q^2=-\epsilon^2\theta_-^2=-\epsilon^2\theta_+^2,$$

$$d\Gamma=\frac{m^2}{s}\,\epsilon^2\pi^2x(1-x)dxd\zeta\,\frac{d\theta_-^2}{x^2},\quad 0<\zeta<\frac{\theta_0^2\epsilon^2x^2}{m^2},\tag{12}$$

and integration over ζ leads to

$$\sigma_{\mathbf{k}\|\mathbf{q}_1}=\frac{2\alpha^3}{Q_1^2}\int_{\rho_2^2}^{\rho_4^2}\frac{dz}{z^2}\int_{x_c}^{1-\Delta}dx\left[\frac{1+x^2}{1-x}\ln\frac{\theta_0^2\epsilon^2x^2}{m^2}-\frac{2x}{1-x}\right].\tag{13}$$

The right side of Eq. (13) corresponds to the contribution of the strip with the width $2\sqrt{z}x^2(1-x)\lambda$ around the line $z_1=x^2z$ in the (z_1,z) plane. The condition $\theta_\gamma<\theta_0$ for the final electron cone can be formulated as $|\mathbf{r}|<\theta_0$, where $\mathbf{r}=\mathbf{k}/\omega-\mathbf{q}_1/\epsilon_1$, and the latter reads as

$$|\sqrt{z_1}-x\sqrt{z}|<x(1-x)\lambda,$$

$$-1<\cos\,\varphi<-1+\frac{\lambda^2x^2(1-x)^2-(\sqrt{z_1}-x\sqrt{z})^2}{2x\sqrt{zz_1}}.\tag{14}$$

Having contributions due to the collinear regions, we now must find the contribution due to the semi-collinear regions. If $m=0$ on the right side of Eq. (8), then the differential cross section suitable for this case can be written as follows:

$$d\sigma=\frac{\alpha^3d\varphi\;dz\;dz_1(1+x^2)}{\pi Q_1^2z(z_1-xz)}\left[\frac{1}{z_1+z+2\sqrt{z_1z}\,\cos\,\varphi}-\frac{x}{z_1+x^2z+2x\sqrt{z_1z}\,\cos\,\varphi}\right]dx.\tag{15}$$

When integrating the first term in the brackets on the right side of Eq. (15) one must use the restriction $\theta_\gamma>\theta_0$ or

$$|\sqrt{z_1}-\sqrt{z}|>(1-x)\lambda,\quad -1<\cos\,\varphi<1,$$

$$|\sqrt{z_1}-\sqrt{z}|<(1-x)\lambda,$$

$$1>\cos\,\varphi>-1+\frac{\lambda^2(1-x)^2-(\sqrt{z_1}-\sqrt{z})^2}{2\sqrt{zz_1}},\tag{16}$$

and for integration of the second term one must use the restriction $|\mathbf{r}|>\theta_0$ or

$$|\sqrt{z_1}-x\sqrt{z}|>x(1-x)\lambda,\quad -1<\cos\,\varphi<1,$$

$$|\sqrt{z_1}-x\sqrt{z}|<x(1-x)\lambda,$$

$$1>\cos\,\varphi>-1+\frac{\lambda^2x^2(1-x)^2-(\sqrt{z_1}-x\sqrt{z})^2}{2x\sqrt{zz_1}}.\tag{17}$$

The integration (15) over the region (16) gives

$$\sigma_a=\frac{2\alpha^3}{Q_1^2}\int_{\rho_2^2}^{\rho_4^2}\frac{dz}{z^2}\times\int_{x_c}^{1-\Delta}\frac{1+x^2}{1-x}\,dx\left[\left(\ln\frac{z}{\lambda^2}+L_2\right)\theta_3^{(x)}+L_3\overline{\theta_3^{(x)}}\right].\tag{18}$$

Analogously, the integration of the right side of Eq. (15) over the region (17) gives

$$\sigma_b=\frac{2\alpha^3}{Q_1^2}\int_{\rho_2^2}^{\rho_4^2}\frac{dz}{z^2}\int_{x_c}^{1-\Delta}\frac{1+x^2}{1-x}\,dx\left(\ln\frac{z}{x^2\lambda^2}+L_1\right).\tag{19}$$

The values L_1 which enter into Eqs. (18) and (19) are defined as follows:

$$L_1=\ln\left|\frac{x^2(z-1)(\rho_3^2-z)}{(x-z)(x\rho_3^2-z)}\right|,\quad L_2=\ln\left|\frac{(z-x^2)(x^2\rho_3^2-z)}{x^2(x-z)(x\rho_3^2-z)}\right|,$$

$$L_3=\ln\left|\frac{(z-x^2)(x\rho_3^2-z)}{(x-z)(x^2\rho_3^2-z)}\right|.$$

In addition, the following notation for the θ-functions is used:

$$\theta_3^{(x)}=\theta(x^2\rho_3^2-z),\quad \overline{\theta_3^{(x)}}=1-\theta_3^{(x)}=\theta(z-x^2\rho_3^2).$$

Thus, the Σ^H may be represented as the sum of (10), (13), (18), and (19) divided by the factor $4\pi\alpha^2/Q_1^2$ or

$$\Sigma^H=\frac{\alpha}{2\pi}\int_{\rho_2^2}^{\rho_4^2}\frac{dz}{z^2}\int_{x_c}^{1-\Delta}\frac{1+x^2}{1-x}\,[(1+\theta_3^{(x)})(L-1)+K(x,z;\rho_3,1)]dx,\tag{20}$$

$$K(x,z;\rho_3,1)=\frac{(1-x)^2}{1+x^2}\,(1+\theta_3^{(x)})+L_1+\theta_3^{(x)}L_2+\overline{\theta_3^{(x)}}L_3.$$

Here the short notation for the θ-functions is used:

$$\theta_i^{(x)}=\theta(x^2\rho_i^2-z),\quad \theta_i=\theta(\rho_i^2-z),$$

$$\overline{\theta_i^{(x)}}=1-\theta_i^{(x)},\quad \overline{\theta_i}=1-\theta_i.$$

It is easy to see that Σ^H for the wide-narrow detectors can be derived from Σ^H for symmetrical detectors (see Ref. 10) by changing the z-integrations limits

$$\int_1^{\rho^2} dz \to \int_{\rho_2^2}^{\rho_4^2} dz \tag{21}$$

and by substituting ρ_3 for ρ under integral sign.

The third term on the right side of Eq. (4) describes the photon emission by a positron. It can be derived by full analogy with Σ^H except for the restrictions on the variables z and z_1:

$$1<z<\rho_3^2, \quad x^2\rho_2^2<z_1<x^2\rho_4^2. \tag{22}$$

The contribution of collinear kinematics ($\mathbf{k}\|\mathbf{p}_2$ and $\mathbf{k}\|\mathbf{q}_2$) to the single hard photon emission cross section corresponds to the integration over the regions inside the strips with a width $2\sqrt{z}(1-x)\lambda$ and $2\sqrt{z}x^2(1-x)\lambda$, respectively. It can be written as follows:

$$\sigma_{\mathbf{k}\|\mathbf{p}_2,\mathbf{k}\|\mathbf{q}_2}=\frac{2\alpha^3}{Q_1^2}\int_1^{\rho_3^2}\frac{dz}{z^2}\int_{x_c}^{1-\Delta}\frac{1+x^2}{1-x}\,dx\left\{\left(\ln\frac{\epsilon^2\theta_0^2}{m^2}-\frac{2x}{1-x}\right)\Delta_{42}^{(x)}\left(\ln\frac{\epsilon^2\theta_0^2x^2}{m^2}-\frac{2x}{1-x}\right)\Delta_{42}\right\}, \tag{23}$$

where

$$\Delta_{42}^{(x)}=\theta_4^{(x)}-\theta_2^{(x)}, \quad \Delta_{42}=\theta_4-\theta_2. \tag{24}$$

The contribution of the semi-collinear kinematics can be derived by integration (15), taking into account the restrictions (16), (17), and (22). The latter corresponds to regions outside the narrow strips near $z_1=z$ and $z_1=x^2z$, respectively. The result is

$$\sigma_{a+b}=\frac{2\alpha^3}{Q_1^2}\int_1^{\rho_3^2}\frac{dz}{z^2}\int_{x_c}^{1-\Delta}\frac{1+x^2}{1-x}\,dx\left[\ln\frac{z}{\lambda^2}\,(\Delta_{42}+\Delta_{42}^{(x)})+\overline{L_2}\Delta_{42}^{(x)}+(\overline{L_1}-2\ln x)\Delta_{42}+\overline{L_3}(\overline{\theta_4^{(x)}}-\theta_2^{(x)})+\overline{L_4}(\overline{\theta_4}-\theta_2)\right], \tag{25}$$

where

$$\overline{L_1}=\ln\left|\frac{(z-\rho_2^2)(\rho_4^2-z)x^2}{(x\rho_4^2-z)(x\rho_2^2-z)}\right|, \quad \overline{L_2}=\ln\left|\frac{(z-x^2\rho_2^2)(x^2\rho_4^2-z)}{x^2(x\rho_4^2-z)(x\rho_2^2-z)}\right|,$$
$$\overline{L_3}=\ln\left|\frac{(z-x^2\rho_2^2)(x\rho_4^2-z)}{(x^2\rho_4^2-z)(x\rho_2^2-z)}\right|, \quad \overline{L_4}=\ln\left|\frac{(z-\rho_2^2)(x\rho_4^2-z)}{(\rho_4^2-z)(x\rho_2^2-z)}\right|. \tag{26}$$

The Σ_H is the sum of (23) and (25) divided by $4\pi\alpha^2/Q_1^2$:

$$\Sigma_H=\frac{\alpha}{2\pi}\int_1^{\rho_3^2}\frac{dz}{z^2}\int_{x_c}^{1-\Delta}\frac{1+x^2}{1-x}\,dx[(L-1)(\Delta_{42}+\Delta_{42}^{(x)})+\widetilde{K}(x,z;\rho_4,\rho_2)], \tag{27}$$

$$\widetilde{K}=\frac{(1-x)^2}{1+x^2}\,(\Delta_{42}+\Delta_{42}^{(x)})+\Delta_{42}\overline{L_1}+\Delta_{42}^{(x)}\overline{L_2}+(\overline{\theta_4^{(x)}}-\theta_2^{(x)})\overline{L_3}+(\overline{\theta_4}-\theta_2)\overline{L_4}.$$

As one can see, the auxiliary parameter θ_0 disappears in the expressions for Σ^H and Σ_H, and the large logarithm acquires the correct appearance. Thus, the separate investigation of contributions due to collinear and semi-collinear kinematics simplifies the calculations and gives a deeper understanding of the underlying physics. This approach is very important for the study of CES when it needs to describe events that belong to the electron (or positron) cluster in a different way compared with events that do not belong to it.

The different parts on the right side of Eq. (4) depend on the auxiliary infrared parameter Δ but the sum does not. It has the form

$$\Sigma_1=\frac{\alpha}{2\pi}\left\{\int_1^{\rho_3^2}\frac{dz}{z^2}\left[-\Delta_{42}+\int_{x_c}^1\left((L-1)P_1(x)(\Delta_{42}+\Delta_{42}^{(x)})+\frac{1+x^2}{1-x}\,\widetilde{K}\right)dx\right]+\int_{\rho_2^2}^{\rho_4^2}\frac{dz}{z^2}\left[-1+\int_{x_c}^1\left((L-1)P_1(x)(1+\theta_3^{(x)})+\frac{1+x^2}{1-x}\,K\right)dx\right]\right\}, \tag{28}$$

where

$$P_1(x)=\frac{1+x^2}{1-x}\,\theta(1-x-\Delta)+\left(2\ln\Delta+\frac{3}{2}\right)\delta(1-x),$$

$\Delta\to 0$.

In order to eliminate the Δ-dependence, one can use the relations

$$\int_{x_c}^1 P_1(x)dx=-\int_0^{x_c}\frac{1+x^2}{1-x}\,dx,$$
$$\int_{x_c}^1 P_1(x)\overline{\theta_3^{(x)}}dx=\overline{\theta_3^{(x_c)}}\int_{x_c}^{\sqrt{z}/\rho_3}\frac{1+x^2}{1-x}\,dx,$$
$$\int_{x_c}^1 P_1(x)\overline{\Delta_{42}^{(x)}}dx=\theta_4\overline{\theta_4^{(x_c)}}\int_{x_c}^{\sqrt{z}/\rho_4}\frac{1+x^2}{1-x}\,dx-\theta_2\overline{\theta_2^{(x_c)}}\int_{x_c}^{\sqrt{z}/\rho_2}\frac{1+x^2}{1-x}\,dx, \tag{29}$$

where $\overline{\Delta_{42}^{(x)}}=\Delta_{42}-\Delta_{42}^{(x)}$.

The right side of Eq. (28) is the full first-order QED correction to the Born SABS cross section at LEP1 for IES with switched-off vacuum polarization effect. The latter can be taken into account by inserting the quantity $[1-\Pi(-zQ_1^2)]^{-2}$ under the sign of the z-integration. (For Π see Ref. 3 and the bibliography cited there).

3. SECOND-ORDER CORRECTION

The second-order correction contains the contributions due to two-photon (real and virtual) emission and pair production. As in the symmetrical case, one needs to distinguish between the situations in which additional photons attach only one fermion line (one-side emission) and two fermion lines (opposite-sides emission) in the corresponding Feynman diagrams.

3.1. The contribution of pair production

Consider at first the contribution of the electron-positron pair production $\Sigma^{\rm pair}$ to the second-order correction:

$$\Sigma^{\rm pair}=\Sigma^{e^+e^-}+\Sigma_{e^+e^-}. \tag{30}$$

To avoid writing some formulas which have the same structure for symmetrical and wide-narrow angular acceptance we will often refer the reader to Ref. 11, in which the details of computation are given for the symmetrical case.

From Sec. 2 we can write the expression for $\Sigma^{e^+e^-}$ when the created electron-positron pair moves in the electron momentum direction, using the result of Ref. 11 for $\Sigma^{e^+e^-}$. It needs only to change the z-integration limits: $(\rho^2,1)\to(\rho_4^2,\rho_2^2)$ and substitute ρ_3 for ρ everywhere under the integral sign. We can write the result as follows:

$$\begin{aligned}\Sigma^{e^+e^-}=\frac{\alpha^2}{4\pi^2}\int_{\rho_2^2}^{\rho_4^2}\frac{dz}{z^2}\,L\Bigg\{L\Bigg(1+\frac{4}{3}\ln(1-x_c)\\ -\frac{2}{3}\int_{x_c}^1\frac{dx}{1-x}\,\overline{\theta_3^{(x)}}\Bigg)-\frac{17}{3}-\frac{8}{3}\,\zeta_2-\frac{40}{9}\ln(1\\ -x_c)+\frac{8}{3}\ln^2(1-x_c)+\int_{x_c}^1\frac{dx}{1-x}\,\overline{\theta_3^{(x)}}\Bigg(\frac{20}{9}\\ -\frac{8}{3}\ln(1-x)\Bigg)+\int_{x_c}^1[L\overline{R}(x)(1+\theta_3^{(x)})\\ +\theta_3^{(x)}C_1(x,z;\rho_3)+C_2(x)+d_2(x,z;\rho_3)]dx\Bigg\},\end{aligned} \tag{31}$$

$$\overline{R}(x)=(1+x)\Bigg(\ln x-\frac{1}{3}\Bigg)+\frac{1-x}{6x}\,(4+7x+4x^2),$$

$$\begin{aligned}C_1(x,z;\rho_3)=-\frac{113}{9}+\frac{142}{9}\,x-\frac{2}{3}\,x^2-\frac{4}{3x}-\frac{4}{3}\,(1\\ +x)\ln(1-x)+\frac{2(1+x^2)}{3(1-x)}\Bigg[2\ln\Bigg|\frac{x^2\rho_3^2-z}{x\rho_3^2-z}\Bigg|\\ -3L_{i2}(1-x)\Bigg]+\Bigg(8x^2+3x-9-\frac{8}{x}\\ -\frac{7}{1-x}\Bigg)\ln x+\frac{2(5x^2-6)}{1-x}\ln^2 x\\ +R(x)\ln\frac{(x^2\rho_3^2-z)^2}{\rho_3^4},\end{aligned}$$

$$\begin{aligned}C_2(x)=-\frac{122}{9}+\frac{133}{9}\,x+\frac{4}{3}\,x^2+\frac{2}{3x}-\frac{4}{3}\,(1+x)\ln(1\\ -x)+\frac{2(1+x^2)}{(1-x)}\,L_{i2}(1-x)+\frac{1}{3}\Bigg(-8x^2-32x\\ -20+\frac{8}{x}+\frac{13}{1-x}\Bigg)\ln x+3(1+x)\ln^2 x,\end{aligned}$$

$$R(x)=2\overline{R}(x)+\frac{2}{3}\,(1+x),$$

$$\begin{aligned}d_2(x,z;\rho_3)=\frac{2(1+x^2)}{3(1-x)}\ln\Bigg|\frac{(z-x^2)(\rho_3^2-z)(z-1)}{(z-x)^2(x^2\rho_3^2-z)}\Bigg|\\ +R(x)\ln\Bigg|\frac{(z-x^2)(\rho_3^2-z)(z-1)}{x^2\rho_3^2-z}\Bigg|.\end{aligned} \tag{32}$$

The right side of Eq. (31) does not contain infrared auxiliary parameters because it includes the contributions due to the real and virtual pair production. The contribution of the hard pair takes into account the collinear and semi-collinear kinematics.

If the created electron-positron pair is emitted in the positron momentum direction, the corresponding expression requires more modifications. The source of such modifications is the semi-collinear kinematics, as we saw in Sec. 2 for the single-photon emission.

The straightforward calculation shows that for the contribution of the semi-collinear region $\mathbf{p}_+\|\mathbf{p}_-$ [we use here notation $\mathbf{p}_\pm$ for the 3-momentum of the positron (electron)] we must include in Eq. (28) of Ref. 11 the expression

$$\begin{aligned}(\Delta_{42}+\Delta_{42}^{(x)})\ln\frac{z}{\lambda^2}+\Delta_{42}\ln\Bigg|\frac{(z-\rho_2^2)(\rho_4^2-z)}{(z-x\rho_2^2)(x\rho_4^2-z)}\Bigg|\\ +\Delta_{42}^{(x)}\ln\Bigg|\frac{(z-x^2\rho_2^2)(x^2\rho_4^2-z)}{x^2(z-x\rho_2^2)(x\rho_4^2-z)}\Bigg|+(\overline{\theta_4}\\ -\theta_2)\ln\Bigg|\frac{(z-\rho_2^2)(x\rho_4^2-z)}{(z-x\rho_2^2)(\rho_4^2-z)}\Bigg|+(\overline{\theta_4^{(x)}}\\ -\theta_2^{(x)})\ln\Bigg|\frac{(z-x^2\rho_2^2)(x\rho_4^2-z)}{(z-x\rho_2^2)(z-x^2\rho_4^2)}\Bigg|,\end{aligned} \tag{33}$$

instead of the expression in curved brackets and change the upper limit of the z-integration: $\rho\to\rho_3$.

For the contribution of the semi-collinear region $\mathbf{p}_+\|\mathbf{q}_1$ the corresponding expression is [see Eq. (33) in Ref. 11]

$$\begin{aligned}\Delta_{42}\Bigg(\ln\frac{z}{\lambda^2}+\ln\Bigg|\frac{(z-\rho_2^2)(\rho_4^2-z)}{x_2^2\rho_2^2\rho_4^2}\Bigg|\Bigg)+(\overline{\theta_4}\\ -\theta_2)\ln\Bigg|\frac{\rho_4^2(z-\rho_2^2)}{\rho_2^2(z-\rho_4^2)}\Bigg|,\end{aligned} \tag{34}$$

and for the semi-collinear region $\mathbf{p}_-\|\mathbf{p}_1$ [see Eq. (38) in Ref. 11] the corresponding expression is

$$\begin{aligned}\Delta_{42}^{(x)}\Bigg(\ln\frac{z}{\lambda^2}+\ln\Bigg|\frac{(z-x^2\rho_2^2)(x^2\rho_4^2-z)}{x_1^2x^4\rho_2^2\rho_4^2}\Bigg|\Bigg)\\ +(\overline{\theta_4^{(x)}}-\theta_2^{(x)})\ln\Bigg|\frac{\rho_4^2(z-x^2\rho_2^2)}{\rho_2^2(z-x^2\rho_4^2)}\Bigg|.\end{aligned} \tag{35}$$

In the symmetrical limit we have $\rho_3=\rho_4=\rho$, $\rho_2=1$, and

$$\begin{aligned}&\Delta_{42}\to\theta(\rho^2-z)\,\theta(z-1),\quad \Delta_{42}^{(x)}\to\theta(x^2\rho^2-z),\\ &\overline{\theta_4^{(x)}}\to\theta(z-x^2\rho^2),\quad \overline{\theta_4},\theta_2,\theta_2^{(x)}\to 0,\end{aligned} \tag{36}$$

and Eqs. (33)–(35) reduce to the corresponding expressions derived in Ref. 11.

The modification of the contributions due to virtual, real, soft and hard collinear pair production includes the change of the z-integral upper limit: $\rho\to\rho_3$ and a trivial change of the θ-functions under the integral sign: $\theta(x^2\rho^2-z)\to\Delta_{42}^{(x)}$, $1\to\Delta_{42}$. The sum of all the contributions has the form

$$\Sigma_{e^+e^-}=\frac{\alpha^2}{4\pi^2}\int_1^{\rho_3^2}\frac{dz}{z^2}\,L\Bigg\{L\Bigg[\Delta_{42}\Bigg(1+\frac{4}{3}\ln(1-x_c)\Bigg)-\frac{2}{3}\int_{x_c}^1\frac{dx}{1-x}\,\overline{\Delta_{42}^{(x)}}\Bigg]+\Delta_{42}\Bigg(-\frac{17}{3}-\frac{8}{3}\,\zeta_2-\frac{40}{9}\ln(1-x_c)+\frac{8}{3}\ln^2(1-x_c)\Bigg)+\int_{x_c}^1\frac{dx}{1-x}\,\overline{\Delta_{42}^{(x)}}\Bigg(\frac{20}{9}-\frac{8}{3}\ln(1-x)\Bigg)+\int_{x_c}^1\Bigg[L\overline{R(x)}(\Delta_{42}+\Delta_{42}^{(x)})+\Delta_{42}^{(x)}C_1(x,z;\rho_2)+\Delta_{42}(C_2(x)+\overline{d_2}(x,z;\rho_2))+(\overline{\theta_4^{(x)}}-\theta_4^{(x)})\times\Bigg(\frac{2(1+x^2)}{3(1-x)}\ln\Bigg|\frac{(x^2\rho_2^2-z)(x\rho_4^2-z)}{(x^2\rho_4^2-z)(x\rho_2^2-z)}\Bigg|+R(x)\ln\Bigg|\frac{(x^2\rho_2^2-z)\rho_4^2}{(x^2\rho_4^2-z)\rho_2^2}\Bigg|\Bigg)+(\overline{\theta_4}-\theta_4)\times\Bigg(\frac{2(1+x^2)}{3(1-x)}\ln\Bigg|\frac{(x\rho_4^2-z)(z-\rho_2^2)}{(x\rho_2^2-z)(z-\rho_4^2)}\Bigg|+R(x)\ln\Bigg|\frac{(\rho_2^2-z)\rho_4^2}{(\rho_4^2-z)\rho_2^2}\Bigg|\Bigg)\Bigg]\Bigg\},\tag{37}$$

$$\overline{d_2}(x,z;\rho_2)=\frac{2(1+x^2)}{3(1-x)}\ln\frac{(z-\rho_2^2)^2}{(z-x\rho_2^2)^2}+2R(x)\ln\frac{z-\rho_2^2}{\rho_2^2}.$$

Using Eq. (36), we can verify that the right side of Eq. (30) goes over to the corresponding expression for symmetrical angular acceptance.

3.2. The contribution of the one-side two-photon emission

In this section we give the analytical expressions for all contributions to the second-order correction which appear due to the one-side two-photon (real and virtual) emission. The master formula, which does not contain the infrared auxiliary parameter Δ, is written only for the leading approximation, and next-to-leading contribution to it is given in Appendix A.

As before, we differentiate between the radiation along the electron and positron momentum directions:

$$\Sigma_2=\Sigma^{\gamma\gamma}+\Sigma_{\gamma\gamma},\quad \Sigma^{\gamma\gamma}=\Sigma^{(S+V)^2}+\Sigma^{(S+V)H}+\Sigma^{HH},$$
$$\Sigma_{\gamma\gamma}=\Sigma_{(S+V)^2}+\Sigma_{(S+V)H}+\Sigma_{HH}.\tag{38}$$

The contribution of the virtual and real soft photons is the same for electron and positron emission:

$$\Sigma_{(S+V)^2}=\Sigma^{(S+V)^2}=\frac{\alpha^2}{\pi^2}\int_{\rho_2^2}^{\rho_4^2}\frac{dz}{z^2}\,L\Bigg[L\Bigg(2\ln^2\Delta+3\ln\Delta+\frac{9}{8}\Bigg)-4\ln^2\Delta-7\ln\Delta+3\zeta_3-\frac{3}{2}\,\zeta_2-\frac{45}{16}\Bigg].\tag{39}$$

The virtual and real soft photon correction to the single hard photon emission differs for a photon that moves along the electron momentum direction and positron direction. In the first case the corresponding contribution can be derived with the help of the result for a symmetrical detector [see Eq. (50) in Ref. 10] using the substitutions (ρ_4^2,ρ_2^2) instead of $(\rho^2,1)$ for the z-integration limits and ρ_3 instead of ρ under the integral sign. Therefore,

$$\Sigma^{(S+V)H}=\frac{\alpha^2}{2\pi^2}\int_{\rho_2^2}^{\rho_4^2}\frac{dz}{z^2}\,L\int_{x_c}^{1-\Delta}\frac{1+x^2}{1-x}\,dx\Bigg\{\Bigg(2\ln\Delta-\ln x+\frac{3}{2}\Bigg)[K(x,z;\rho_3,1)+(L-1)(1+\theta_3^{(x)})]+\frac{1}{2}\ln^2 x-\frac{(1-x)^2}{2(1+x^2)}+(1+\theta_3^{(x)})(-2+\ln x-2\ln\Delta)+\overline{\theta_3^{(x)}}\Bigg[\frac{1}{2}\,L\ln x+2\ln\Delta\ln x-\ln x\ln(1-x)-\ln^2 x-L_{i2}(1-x)-\frac{x(1-x)+4x\ln x}{2(1+x^2)}\Bigg]\Bigg\}.\tag{40}$$

In order to obtain the expression for $\Sigma_{(S+V)H}$ we must change on the right side of Eq. (39):

i) the limits of z-intergration: $(\rho_4^2,\rho_2^2)\to(\rho_3^2,1)$,

ii) $K(x,z{:}\rho_3,1)\to\widetilde{K}(x,z{:}\rho_4,\rho_2)$, $\theta_3^{(x)}\to\Delta_{42}^{(x)}$,

$$\overline{\theta_3^{(x)}}\to\overline{\Delta_{42}^{(x)}},\quad 1\to\Delta_{42}.\tag{41}$$

The contribution of two hard photons emitted in the electron momentum direction may be obtained in the same way as $\Sigma^{(S+V)H}$, using the known result for symmetrical detectors [see Eq. (54) in Ref. 10]:

$$\Sigma^{HH}=\frac{\alpha^2}{4\pi^2}\int_{\rho_2^2}^{\rho_4^2}\frac{dz}{z^2}\,L\int_{x_c}^{1-2\Delta}dx\int_{\Delta}^{1-x-\Delta}dx_1\frac{I^{HH}}{x_1(1-x-x_1)(1-x_1)^2}, \tag{42}$$

$$I^{HH}=\overline{A}\,\theta_3^{(x)}+\overline{B}+\overline{C}\,\theta_3^{(1-x_1)},$$

$$\overline{A}=\gamma\beta\left(\frac{L}{2}+\ln\frac{(x^2\rho_3^2-z)^2}{x^2(x(1-x_1)\rho_3^2-z)^2}\right)+\zeta\,\ln\frac{(1-x_1)^2(1-x-x_1)}{xx_1}+\gamma_A,$$

$$\overline{B}=\gamma\beta\left(\frac{L}{2}+\ln\left|\frac{x^2(z-1)(\rho_3^2-z)(z-x^2)(z-(1-x_1)^2)^2(\rho_3^2x(1-x_1)-z)^2}{(\rho_3^2(1-x_1)^2-z)^2(z-(1-x_1))^2(z-x(1-x_1))^2(\rho_3^2x^2-z)}\right|\right)+\zeta\,\ln\frac{(1-x_1)^2x_1}{x(1-x-x_1)}+\delta_B, \tag{43}$$

$$\overline{C}=\gamma\beta\left(L+2\ln\left|\frac{x(\rho_3^2(1-x_1)^2-z)^2}{(1-x_1)^2(\rho_3^2x(1-x_1)-z)(\rho_3^2(1-x_1)-z)}\right|\right)-2(1-x_1)\beta-2x(1-x_1)\gamma,$$

where

$$\gamma=1+(1-x_1)^2,\quad \beta=x^2+(1-x_1)^2,\quad \zeta=x^2+(1-x_1)^4,$$

$$\gamma_A=xx_1(1-x-x_1)-x_1^2(1-x-x_1)^2-2(1-x_1)\beta,$$

$$\delta_B=xx_1(1-x-x_1)-x_1^2(1-x-x_1)^2-2x(1-x_1)\gamma.$$

Unfortunately, it is impossible to give such a simple prescription as (41) in order to obtain Σ_{HH} from Eqs. (42) and (43). In the case of radiation of two hard photons along the positron momentum direction, additional detailed analysis of semi-collinear kinematics is required. All essential points of such an analysis are given in Sec. 2, and the reader can make all calculations with the help of the formulas given in Appendix B of Ref. 10. The final result is

$$\Sigma_{HH}=\frac{\alpha^2}{4\pi^2}\int_1^{\rho_3^2}\frac{dz}{z^2}\,L\int_{x_c}^{1-2\Delta}dx\int_{\Delta}^{1-x-\Delta}dx_1\times\frac{I_{HH}}{x_1(1-x-x_1)(1-x_1)^2}, \tag{44}$$

$$I_{HH}=\widetilde{A}\Delta_{42}^{(x)}+\widetilde{C}\Delta_{42}^{(1-x_1)}+\widetilde{B}\Delta_{42}+(\overline{\theta_4^{(x)}}-\theta_2^{(x)})a+(\overline{\theta_4^{(1-x_1)}}-\theta_2^{(1-x_1)})c+(\overline{\theta_4}-\theta_2)b,$$

$$a=\gamma\beta\,\ln\left|\frac{(\rho_4^2x(1-x_1)-z)(\rho_2^2x^2-z)}{(\rho_2^2x(1-x_1)-z)(\rho_4^2x^2-z)}\right|,$$

$$b=\gamma\beta\,\ln\left|\frac{(\rho_4^2(1-x_1)-z)(\rho_2^2-z)}{(\rho_2^2(1-x_1)-z)(\rho_4^2-z)}\right|,$$

$$c=\gamma\beta\,\ln\left|\frac{(\rho_4^2x(1-x_1)-z)(\rho_2^2(1-x_1)^2-z)^2(\rho_4^2(1-x_1)-z)}{(\rho_2^2x(1-x_1)-z)(\rho_4^2(1-x_1)^2-z)^2(\rho_2^2(1-x_1)-z)}\right|,$$

$$\widetilde{A}=\gamma\beta\left(\frac{L}{2}+\ln\left|\frac{(\rho_4^2x^2-z)(\rho_2^2x^2-z)}{x^2(\rho_4^2x(1-x_1)-z)(\rho_2^2x(1-x_1)-z)}\right|\right)+\zeta\,\ln\frac{(1-x_1)^2(1-x-x_1)}{xx_1}+\gamma_A,$$

$$\widetilde{B}=\gamma\beta\left(\frac{L}{2}+\ln\left|\frac{x^2(\rho_4^2-z)(\rho_2^2-z)}{(\rho_4^2(1-x_1)-z)(\rho_2^2(1-x_1)-z)}\right|\right)+\zeta\,\ln\frac{(1-x_1)^2x_1}{x(1-x-x_1)}+\delta_B,$$

$$\widetilde{C}=\gamma\beta\left(L+\ln\left|\frac{x^2(\rho_4^2(1-x_1)^2-z)^2(\rho_2^2(1-x_1)^2-z)^2}{(1-x_1)^4(\rho_4^2x(1-x_1)-z)(\rho_2^2x(1-x_1)-z)(\rho_4^2(1-x_1)-z)(\rho_2^2(1-x_1)-z)}\right|\right)-2(1-x_1)(\beta+x\gamma).$$

As one can see, the separate contributions to the right side of Eq. (38) depend on the infrared auxiliary parameter Δ but $\Sigma^{\gamma\gamma}$ and $\Sigma_{\gamma\gamma}$ do not. The elimination of Δ-dependence analytically required considerable effort. The leading terms are given below (for the next-to-leading terms see Appendix A):

$$\Sigma^{\gamma\gamma L}=\frac{\alpha^2}{4\pi^2}\int_{\rho_2^2}^{\rho_4^2}\frac{dz}{z^2}\,L^2\int_{x_c}^1dx\left[\frac{1}{2}\,(1+\theta_3^{(x)})P_2(x)+\int_x^1\frac{dt}{t}\,P_1(t)P_1\left(\frac{x}{t}\right)\theta_3^{(t)}\right], \tag{45}$$

$$\Sigma^L_{\gamma\gamma}=\frac{\alpha^2}{4\pi^2}\int_1^{\rho_3^2}\frac{dz}{z^2}L^2\int_{x_c}^1 dx\left[\frac{1}{2}(\Delta_{42}+\Delta_{42}^{(x)})P_2(x)+\int_x^1\frac{dt}{t}P_1(t)P_1\left(\frac{x}{t}\right)\Delta_{42}^{(t)}\right], \tag{46}$$

where

$$P_2(x)=P_1\otimes P_1=\int_x^1\frac{dt}{t}P_1(t)P_1\left(\frac{x}{t}\right)$$
$$=\lim_{\Delta\to 0}\left\{\left[\left(2\ln\Delta+\frac{3}{2}\right)^2-4\zeta_2\right]\delta(1-x)+2\left[\frac{1+x^2}{1-x}\left(2\ln(1-x)-\ln x+\frac{3}{2}\right)+\frac{1}{2}(1+x)\ln x-1+x\right]\theta(1-x-\Delta)\right\}, \tag{47}$$

$$\int_0^1 P_2(x)dx=0.$$

The expressions (45) and (46) are not convenient for numerical calculations. The suitable expressions can be written as follows:

$$\Sigma^{\gamma\gamma L}=\frac{\alpha^2}{4\pi^2}\left\{-2\int_{\rho_2^2}^{\rho_4^2}\frac{dz}{z^2}L^2\int_0^{x_c}P_2(x)dx-\int_{m_{23}}^{\rho_4^2}\frac{dz}{z^2}L^2\int_{x_c}^{\sqrt{z}/\rho_3}\left[P_1(x)g\left(\frac{x_c}{x}\right)+\frac{1}{2}P_2(x)\right]dx\right\}, \tag{48}$$

$$\Sigma^L_{\gamma\gamma}=\frac{\alpha^2}{4\pi^2}\left\{-2\int_{\rho_2^2}^{\rho_4^2}\frac{dz}{z^2}L^2\int_0^{x_c}P_2(x)dx-\int_{m_{14}}^{\rho_4^2}\frac{dz}{z^2}L^2\int_{x_c}^{\sqrt{z}/\rho_4}\left[P_1(x)g\left(\frac{x_c}{x}\right)+\frac{1}{2}P_2(x)\right]dx\right\}+\int_{m_{12}}^{\rho_2^2}\frac{dz}{z^2}L^2\int_{x_c}^{\sqrt{z}/\rho_2}\left[P_1(x)g\left(\frac{x_c}{x}\right)+\frac{1}{2}P_2(x)\right]dx\right\}, \tag{49}$$

where

$$g(y)=y+\frac{y^2}{2}+2\ln(1-y),\quad m_{23}=\max(\rho_2^2,x_c^2\rho_3^2),$$

$$m_{14}=\max(1,x_c^2\rho_4^2),\quad m_{12}=\max(1,x_c^2\rho_2^2).$$

The last two formulas can be derived with the help of the relations given in Appendix B. The integration over the x-variables in Eqs. (45) and (46) can be performed with the help of the formulas

$$\int^x P_2(y)dy=F_2(x),\quad \int^x P_1(y)g\left(\frac{x_c}{y}\right)dy=F_g(x),$$

$$\int^x P_1(y)dy=-g(x),\quad x<1, \tag{50}$$

$$F_2(x)=-2x-\frac{x^2}{4}+\left(x+\frac{x^2}{2}\right)\ln\frac{x^3}{(1-x)^4}+4\ln(1-x)\ln\frac{x}{1-x}+4L_{i2}(x), \tag{51}$$

$$F_g(x)=-\frac{x_c^2}{2x}+(2x+x^2)\ln x+\left(x_c+\frac{x_c^2}{2}\right)\ln\frac{x}{(1-x)^2}+\left(2x_c+\frac{x_c^2}{2}-2x-\frac{x^2}{2}\right)\ln(x-x_c)+4L_{i2}(x)+4L_{i2}\left(\frac{1-x}{1-x_c}\right),\quad x_c<x<1. \tag{52}$$

Therefore, the second-order leading contribution to the SABS cross section at LEP1 can be expressed in terms of one integral over the z-variable.

It is useful to note that for CES the leading contributions in all orders of perturbation theory take into account the emission of photons in the initial state only. Thus, the corresponding correction due to the one-side, two photon (real and virtual) emission is

$$\Sigma^{\gamma\gamma L}_{\rm CES}=-\frac{1}{8}\left(\frac{\alpha}{\pi}\right)^2\int_{\rho_2^2}^{\rho_4^2}\frac{dz}{z^2}L^2\left\{F_2(x_c)+\left[F_2\left(\frac{\sqrt{z}}{\rho_3}\right)-F_2(x_c)\right]\overline{\theta_3^{(x_c)}}\right\}, \tag{53}$$

$$\Sigma^L_{\gamma\gamma\ {\rm CES}}=-\frac{1}{8}\left(\frac{\alpha}{\pi}\right)^2\left\{\int_{\rho_2^2}^{\rho_4^2}\frac{dz}{z^2}L^2F_2(x_c)+\int_1^{\rho_4^2}\frac{dz}{z^2}L^2\left[F_2\left(\frac{\sqrt{z}}{\rho_4}\right)-F_2(x_c)\right]\overline{\theta_4^{(x_c)}}-\int_1^{\rho_2^2}\frac{dz}{z^2}L^2\left[F_2\left(\frac{\sqrt{z}}{\rho_2}\right)-F_2(x_c)\right]\overline{\theta_2^{(x_c)}}\right\}. \tag{54}$$

3.3. Second-order correction due to the opposite-side photon emission

In this section we calculate analytically the expression for

$$\Sigma^\gamma_\gamma=\Sigma^{S+V}_{S+V}+\Sigma^H_{S+V}+\Sigma^{S+V}_H+\Sigma^H_H. \tag{55}$$

The quantity Σ^γ_γ does not depend on the infrared auxiliary parameter Δ because it contains all contributions due to the virtual, real, soft and hard photon emission.

The first term on the right side of Eq. (55) takes into account only the ''opposite-side'' virtual and real soft photon corrections

$$\Sigma^{S+V}_{S+V}=\frac{\alpha^2}{\pi^2}\int_{\rho_2^2}^{\rho_4^2}\frac{dz}{z^2}L\left[L\left(4\ln^2\Delta+6\ln\Delta+\frac{9}{4}\right)-6-14\ln\Delta-8\ln^2\Delta\right]. \tag{56}$$

The contribution of one-loop virtual and real soft photon corrections to the hard single-photon emission can be written as follows:

$$\Sigma^H_{S+V}=\frac{\alpha^2}{2\pi^2}\int_{\rho_2^2}^{\rho_4^2}\frac{dz}{z^2}\Bigg[2(L-1)\ln\Delta+\frac{3}{2}L-2\Bigg]\int_{x_c}^{1-\Delta}\frac{1+x^2}{1-x}[(1+\theta_3^{(x)})(L-1)+K(x,z;\rho_3,1)], \tag{57}$$

$$\Sigma^{S+V}_H=\frac{\alpha^2}{2\pi^2}\int_{1}^{\rho_3^2}\frac{dz}{z^2}\Bigg[2(L-1)\ln\Delta+\frac{3}{2}L-2\Bigg]\int_{x_c}^{1-\Delta}\frac{1+x^2}{1-x}[(\Delta_{42}+\Delta_{42}^{(x)})(L-1)+\widetilde{K}(x,z;\rho_4,\rho_2)]dx. \tag{58}$$

In order to find the contribution of the two opposite-side hard photon emission to Σ^γ_γ, it is convenient to use the factorization theorem for the differential cross sections of two-jets processes in QED:[16]

$$\Sigma^H_H=\frac{\alpha^2}{4\pi^2}\int_0^\infty\frac{dz}{z^2}\int_{x_c}^{1-\Delta}dx_1\int_{x_c/x_1}^{1-\Delta}dx_2\,\frac{1+x_1^2}{1-x_1}\frac{1+x_2^2}{1-x_2}\times\Phi(x_1,z;\rho_3,1)\Phi(x_2,z;\rho_4,\rho_2), \tag{59}$$

$$\Phi(x,z;\rho_3,1)=(\Delta_{31}+\Delta_{31}^{(x)})(L-1)+\frac{(1-x)^2}{1+x^2}(\Delta_{31}+\Delta_{31}^{(x)})+\Delta_{31}L_1+\Delta_{31}^{(x)}L_2+(\overline{\theta_3^{(x)}}-\theta_1^{(x)})L_3+(\overline{\theta_3}-\theta_1)\ln\left|\frac{(x\rho_3^2-z)(z-1)}{(z-x)(\rho_3^2-z)}\right|, \tag{60}$$

$$\Phi(x,z;\rho_4,\rho_2)=(\Delta_{42}+\Delta_{42}^{(x)})(L-1)+\widetilde{K}(x,z;\rho_4,\rho_2),$$

$$\Delta_{31}=\theta_3-\theta_1,\quad \Delta_{31}^{(x)}=\theta_3^{(x)}-\theta_1^{(x)}, \tag{61}$$

$$\theta_1=\theta(1-z),\quad \theta_1^{(x)}=\theta(x^2-z).$$

The Δ-dependence of the separate terms on the right side of Eq. (55) can be eliminated analytically in the whole sum. The leading contribution is expressed in terms of the electron structure functions as follows:

$$\Sigma^{\gamma L}_\gamma=\frac{\alpha^2}{4\pi^2}\int_0^\infty\frac{dz}{z^2}L^2\int_{x_c}^1dx_1\int_{x_c/x_1}^1dx_2P_1(x_1)P_1(x_2)\times(\Delta_{31}+\Delta_{31}^{(x_1)})(\Delta_{42}+\Delta_{42}^{(x_2)}). \tag{62}$$

The next-to-leading contribution to Σ^γ_γ is given in Appendix A.

The form of Σ^γ_γ suitable for numerical counting can be written in terms of the functions $F_2(x)$ and $F_g(x)$ in the same manner as it was done at the end of Sec. 3.2:

$$\begin{aligned}\Sigma^{\gamma L}_\gamma=\frac{\alpha^2}{4\pi^2}\Bigg\{&-\int_{\rho_2^2}^{\rho_4^2}\frac{dz}{z^2}L^2\Bigg[4(1)F_2(x_c)+2(1)\Bigg(F_g\Bigg(\frac{\sqrt{z}}{\rho_3}\Bigg)-F_g(x_c)\Bigg)\overline{\theta_3^{(x_c)}}-\int_1^{\rho_4^2}\frac{dz}{z^2}L^2 2(1)\Bigg(F_g\Bigg(\frac{\sqrt{z}}{\rho_4}\Bigg)-F_g(x_c)\Bigg)\overline{\theta_4^{(x_c)}}+\int_1^{\rho_2^2}\frac{dz}{z^2}L^2 2(1)\Bigg(F_g\Bigg(\frac{\sqrt{z}}{\rho_2}\Bigg)-F_g(x_c)\Bigg)\overline{\theta_2^{(x_c)}}+\int_{x_c\rho_3\rho_4}^{\rho_4^2}\frac{dz}{z^2}L^2\Bigg[F_g\Bigg(\frac{\sqrt{z}}{\rho_4}\Bigg)\\&-F_g\Bigg(\frac{x_c\rho_3}{\sqrt{z}}\Bigg)+g\Bigg(\frac{\sqrt{z}}{\rho_3}\Bigg)\Bigg(g\Bigg(\frac{\sqrt{z}}{\rho_4}\Bigg)-g\Bigg(\frac{x_c\rho_3}{\sqrt{z}}\Bigg)\Bigg)\Bigg]+\int_{x_c\rho_2}^1\frac{dz}{z^2}L^2\Bigg[F_g(\sqrt{z})-F_g\Bigg(\frac{x_c\rho_2}{\sqrt{z}}\Bigg)+g\Bigg(\frac{\sqrt{z}}{\rho_2}\Bigg)\\&\times\Bigg(g(\sqrt{z})-g\Bigg(\frac{x_c\rho_2}{\sqrt{z}}\Bigg)\Bigg)\Bigg]-\int_{x_c\rho_4}^1\frac{dz}{z^2}L^2\Bigg[F_g\Bigg(\frac{\sqrt{z}}{\rho_4}\Bigg)-F_g\Bigg(\frac{x_c}{\sqrt{z}}\Bigg)+g(\sqrt{z})\Bigg(g\Bigg(\frac{\sqrt{z}}{\rho_4}\Bigg)-g\Bigg(\frac{x_c}{\sqrt{z}}\Bigg)\Bigg)\Bigg]\\&-\int_{x_c\rho_3\rho_2}^{\rho_2^2}\frac{dz}{z^2}L^2\Bigg[F_g\Bigg(\frac{\sqrt{z}}{\rho_3}\Bigg)-F_g\Bigg(\frac{x_c\rho_2}{\sqrt{z}}\Bigg)+g\Bigg(\frac{\sqrt{z}}{\rho_2}\Bigg)\times\Bigg(g\Bigg(\frac{\sqrt{z}}{\rho_3}\Bigg)-g\Bigg(\frac{x_c\rho_2}{\sqrt{z}}\Bigg)\Bigg)\Bigg]\Bigg\}.\end{aligned} \tag{63}$$

On the right side of Eq. (63) the quantities in brackets are suitable for CES, when only the initial state radiation is taken into account.

4. THIRD-ORDER CORRECTION

Within the required accuracy only the leading contribution to the third-order correction must be kept. The latter becomes more important than the next-to-leading one for LEP2 because of the increase in the energy. In order to evaluate it, one can use the iteration up to third order of the master equation for the electron structure function:[13]

$$D(x,\alpha_{\rm eff})=D^{NS}(x,\alpha_{\rm eff})+D^S(x,\alpha_{\rm eff}). \tag{64}$$

The iterative form of nonsinglet component of Eq. (64) is

$$D^{NS}(x,\alpha_{\rm eff})=\delta(1-x)+\sum_{k=1}^\infty\frac{1}{k!}\left(\frac{\alpha_{\rm eff}}{2\pi}\right)^k P_1(x)^{\otimes k},$$

$$\underbrace{P_1(x)\otimes\cdots\otimes P_1(x)}_{k}=P_1(x)^{\otimes k},$$

$$P_1(x)\otimes P_1(x)=\int_x^1P_1(t)P_1\left(\frac{x}{t}\right)\frac{dt}{t}. \tag{65}$$

Up to the third-order the singlet component of Eq. (64) is[13]

$$D^S(x,\alpha_{\rm eff})=\frac{1}{2!}\left(\frac{\alpha_{\rm eff}}{2\pi}\right)^2 R(x)$$
$$+\frac{1}{3!}\left(\frac{\alpha_{\rm eff}}{2\pi}\right)^3\left[2P_1\otimes R(x)-\frac{2}{3}R(x)\right], \quad (66)$$

where $R(x)$ is defined by Eq. (31). The effective coupling $a_{\rm eff}$ in Eqs. (64)–(66) represents the integral of the running QED constant

$$\frac{\alpha_{\rm eff}}{2\pi}=\int_0^L\frac{\alpha dt}{2\pi(1-\alpha t/3\pi)}=\frac{3}{2}\ln\left(1-\frac{\alpha L}{3\pi}\right)^{-1}. \quad (67)$$

The nonsinglet structure function describes the photon emission and pair production without allowance for the identity of final fermions, while the singlet structure function is responsible just for the identity effects.

Up to the third order the electron structure function has the form

$$D(x,L)=\delta(1-x)+\frac{\alpha L}{2\pi}P_1(x)+\frac{1}{2}\left(\frac{\alpha L}{2\pi}\right)^2\left(P_2(x)+\frac{2}{3}P_1(x)+R(x)\right)+\frac{1}{3}\left(\frac{\alpha L}{2\pi}\right)^3\left[\frac{1}{2}P_3(x)+P_2(x)+\frac{4}{9}P_1(x)+\frac{2}{3}R(x)+R^P(x)\right],$$
$$R^P(x)=P_1\otimes R(x). \quad (68)$$

For the functions $P_3(x)$ and $R^P(x)$ see Ref. 13.

The factorization form of the differential cross section[16] leads to

$$\Sigma^L=\int_0^\infty\frac{dz}{z^2}\int_{x_c}^1 dx_1\int_{x_c/x_1}^1 dx_2 C(x_1,L)C(x_2,L),$$
$$C(x_1,L)=\int_{x_1}^1\frac{dt}{t}D(t)D\left(\frac{x_1}{t}\right)\Delta_{31}^{(t)},$$
$$C(x_2,L)=\int_{x_2}^1\frac{dt}{t}D(t)D\left(\frac{x_2}{t}\right)\Delta_{42}^{(t)}. \quad (69)$$

The expansion of $C(x_1,L)$ is

$$C(x_1,L)=\delta(1-x_1)\Delta_{31}^{(x_1)}+\frac{\alpha L}{2\pi}P_1(x_1)(\Delta_{31}^{(x_1)}+\Delta_{31})+\left(\frac{\alpha L}{2\pi}\right)^2\left[C_2(x_1)(\Delta_{31}^{(x_1)}+\Delta_{31})+\int_{x_1}^1\frac{dt}{t}\Delta_{31}^{(t)}\overline{C_2}(x_1,t)\right]+\left(\frac{\alpha L}{2\pi}\right)^3\left[C_3(x_1)\times(\Delta_{31}^{(x_1)}+\Delta_{31})+\int_{x_1}^1\frac{dt}{t}\Delta_{31}^{(t)}\overline{C_3}(x_1,t)\right], \quad (70)$$

$$C_2(x)=\frac{1}{2}P_2(x)+\frac{1}{3}P_1(x)+\frac{1}{2}R(x),$$
$$\overline{C_2}(x,t)=P_1(t)P_1\left(\frac{x}{t}\right),$$
$$C_3(x)=\frac{1}{6}P_3(x)+\frac{1}{3}P_2(x)+\frac{4}{27}P_1(x)+\frac{2}{9}R(x)+\frac{1}{3}R^P(x), \quad (71)$$
$$\overline{C_3}(x,t)=P_1(t)C_2\left(\frac{x}{t}\right)+C_2(t)P_1\left(\frac{x}{t}\right),$$

and the same for $C(x_2,L)$ with the substitution x_2 instead of x_1 and $\Delta_{42}^{(x_2)}(\Delta_{42})$ instead of $\Delta_{31}^{(x_1)}(\Delta_{31})$.

Because of the θ-functions under integral sign one has to distinguish between $\int_x^1 dt t^{-1}A(t)B(x/t)\Delta_{31}^{(t)}$ and $\int_x^1 dt t^{-1}B(t)A(x/t)\Delta_{31}^{(t)}$.

In the case of CES one must take into account the initial-state radiation only. Therefore, instead of (70) we can write

$$C_{\rm CES}(x_1,L)=\Delta_{31}^{(x_1)}\left[\delta(1-x_1)+\frac{\alpha L}{2\pi}P_1(x_1)+\left(\frac{\alpha L}{2\pi}\right)^2 C_2(x_1)+\left(\frac{\alpha L}{2\pi}\right)^3 C_3(x_1)\right] \quad (72)$$

and likewise for $C(x_2,L)$.

The last step is to write the third-order contribution on the right side of Eq. (69):

$$\Sigma_3^L=\left(\frac{\alpha}{2\pi}\right)^3\int_0^\infty\frac{dz}{z^2}L^3\int_{x_c}^1 dx\left(Z_1+\int_{x_c/x}^1 dx_1 Z_2\right), \quad (73)$$
$$Z_1=(2\Delta_{42}+\Delta_{42}^{(x)}\Delta_{31}+\Delta_{31}^{(x)}\Delta_{42})C_3(x)+\int_x^1\frac{dt}{t}(\Delta_{42}^{(t)}\Delta_{31}+\Delta_{31}^{(t)}\Delta_{42})\overline{C_3}(x,t),$$
$$Z_2=[(\Delta_{31}+\Delta_{31}^{(x)})(\Delta_{42}+\Delta_{42}^{(x_1)})+(\Delta_{31}+\Delta_{31}^{(x_1)})(\Delta_{42}+\Delta_{42}^{(x)})]P_1(x)C_2(x_1)+P_1(x)\int_{x_1}^1[\Delta_{31}^{(t)}\Delta_{42}+\Delta_{42}^{(t)}\Delta_{31}+\Delta_{31}^{(x)}\Delta_{42}^{(t)}+\Delta_{42}^{(x)}\Delta_{31}^{(t)}]\frac{dt}{t}\overline{C_2}(x_1,t).$$

When writing the expressions for Z_1 and Z_2 it is assumed that $\Delta_{31}\Delta_{42}=\Delta_{42}$. In the case of CES the expressions for Z_1 and Z_2 can be written as follows:

$$Z_1=(\Delta_{42}^{(x)}\Delta_{31}+\Delta_{31}^{(x)}\Delta_{42})C_3(x),$$
$$Z_2=(\Delta_{42}^{(x)}\Delta_{31}^{(x_1)}+\Delta_{42}^{(x_1)}\Delta_{31}^{(x)})P_1(x)C_2(x_1). \quad (74)$$

Using the relations given in Appendix B we can represent the right side of Eq. (73) in the form suitable for numerical calculations as double integral over the z- and x-variables. It can be written as follows:

$$\Sigma_3^L=\Sigma_3^0+\Sigma_0^3+\Sigma_2^1+\Sigma_1^2, \quad (75)$$

where the superscript (subscript) shows the number of additional particles (real and virtual) emitted by the electron (positron). The one-side emission contributes to the right side of Eq. (75) as

$$\Sigma_3^0+\Sigma_0^3=\left(\frac{\alpha}{2\pi}\right)^3\left\{\int_{\rho_2^2}^{\rho_4^2}\frac{dz}{z^2}L^3\left[-2\int_0^{x_c}F_p(x)dx+2\int_{x_c}^1F_r(x)dx-\overline{\theta_3^{(x_c)}}\int_{x_c}^{\sqrt{z}/\rho_3}F_{pr}(x,x_c)dx\right]-\int_1^{\rho_4^2}\frac{dz}{z^2}L^3\overline{\theta_4^{(x_c)}}\int_{x_c}^{\sqrt{z}/\rho_4}F_{pr}(x,x_c)dx+\int_1^{\rho_2^2}\frac{dz}{z^2}L^3\overline{\theta_2^{(x_c)}}\int_{x_c}^{\sqrt{z}/p_2}F_{pr}(x,x_c)dx\right\},\quad(76)$$

where

$$F_p(x)=\frac{4}{3}P_3(x)+\frac{4}{3}P_2(x)+\frac{8}{27}P_1(x),$$

$$F_r(x)=\frac{4}{9}R(x)+\frac{5}{3}R^P(x),$$

$$F_{pr}(x,x_c)=\frac{1}{6}P_3(x)+\frac{1}{2}P_2(x)\left[\frac{2}{3}+g\left(\frac{x_c}{x}\right)\right]+P_1(x)\times\left[\frac{4}{27}+\frac{1}{2}f\left(\frac{x_c}{x}\right)+\frac{2}{3}g\left(\frac{x_c}{x}\right)+\frac{1}{2}r\left(\frac{x_c}{x};1\right)\right]+R(x)\left[\frac{2}{9}+\frac{1}{2}g\left(\frac{x_c}{x}\right)\right]+\frac{1}{3}R^P(x),$$

$$r(z,1)=\int_z^1R(x)dx=-\frac{22}{9}+z+z^2+\frac{4}{9}z^3-\left(\frac{4}{3}+2z+z^2\right)\ln z,$$

$$f(z)=-F_2(z).$$

In the case of CES the corresponding contribution can be derived by inserting the functions F_p^c, F_r^c, and F_{pr}^c on the right side of Eq. (76) instead of the functions F_p, F_r, and F_{pr}, respectively, where

$$F_{pr}^c(x)=C_3(x),$$

$$F_p^c(x)=\frac{1}{6}P_3(x)+\frac{1}{3}P_2(x)+\frac{4}{27}P_1(x),$$

$$F_r^c(x)=\frac{2}{9}R(x)+\frac{1}{3}R^P(x).$$

The contribution of the opposite-side emission to the right side of Eq. (75) is

$$\Sigma_2^1+\Sigma_1^2=\left(\frac{\alpha}{2\pi}\right)^3\left\{\int_{\rho_2^2}^{\rho_4^2}\frac{dz}{z^2}L^3\left[\int_0^{x_c}\left(-8P_3(x)-\frac{8}{3}P_2(x)\right)dx+4\int_{x_c}^1R^P(x)dx-\overline{\theta_3^{(x_c)}}\int_{x_c}^{\sqrt{z}/\rho_3}\left(H(x,x_c)+2g\left(\frac{x_c}{x}\right)h(x;\sqrt{z}/\rho_3)\right)dx\right]-\int_1^{\rho_4^2}\frac{dz}{z^2}L^3\overline{\theta_4^{(x_c)}}\int_{x_c}^{\sqrt{z}/\rho_4}\left(H(x,x_c)+2g\left(\frac{x_c}{x}\right)h(x;\sqrt{z}/\rho_4)\right)dx+\int_1^{\rho_2^2}\frac{dz}{z^2}L^3\overline{\theta_2^{(x_c)}}\int_{x_c}^{\sqrt{z}/\rho_2}\left(H(x,x_c)+2g\left(\frac{x_c}{x}\right)h(x;\sqrt{z}/\rho_2)\right)dx+\int_{x_c\rho_3\rho_4}^{\rho_4^2}\frac{dz}{z^2}L^3\left[\int_{x_c\rho_4/\sqrt{z}}^{\sqrt{z}/\rho_3}\left(P_1(x)G\left(\frac{x_c}{x};\frac{\sqrt{z}}{\rho_4}\right)+g\left(\frac{x_c}{x};\frac{\sqrt{z}}{\rho_4}\right)h\left(x;\frac{\sqrt{z}}{\rho_3}\right)\right)dx+(\rho_3\leftrightarrow\rho_4)\right]+\int_{x_c\rho_2}^1\frac{dz}{z^2}L^3\left[\int_{x_c\rho_2/\sqrt{z}}^{\sqrt{z}/1}\left(P_1(x)G\left(\frac{x_c}{x};\frac{\sqrt{z}}{\rho_2}\right)+g\left(\frac{x_c}{x};\frac{\sqrt{z}}{\rho_2}\right)h\left(x;\frac{\sqrt{z}}{1}\right)\right)dx+(\rho_2\leftrightarrow1)\right]-\int_{x_c\rho_3\rho_2}^{\rho_2^2}\frac{dz}{z^2}L^3\left[\int_{x_c\rho_2/\sqrt{z}}^{\sqrt{z}/\rho_3}\left(P_1(x)G\left(\frac{x_c}{x};\frac{\sqrt{z}}{\rho_2}\right)+g\left(\frac{x_c}{x};\frac{\sqrt{z}}{\rho_2}\right)h\left(x;\frac{\sqrt{z}}{\rho_3}\right)\right)dx+(\rho_3\leftrightarrow\rho_2)\right]-\int_{x_c\rho_4}^1\frac{dz}{z^2}L^3\left[\int_{x_c\rho_4/\sqrt{z}}^{\sqrt{z}/1}\left(P_1(x)G\left(\frac{x_c}{x};\frac{\sqrt{z}}{\rho_4}\right)+g\left(\frac{x_c}{x};\frac{\sqrt{z}}{\rho_4}\right)h\left(x;\frac{\sqrt{z}}{1}\right)\right)dx+(\rho_4\leftrightarrow1)\right],\quad(77)$$

where

$$g(a;b)=g(a)-g(b),\quad G(a;b)=G(a)-G(b),$$

$$G(z)=\frac{1}{2}f(z)+\frac{1}{3}g(z)+\frac{1}{2}r(z),$$

$$H(x,x_c)=P_1(x)\left[2f\left(\frac{x_c}{x}\right)+\frac{4}{3}g\left(\frac{x_c}{x}\right)+r\left(\frac{x_c}{x};1\right)\right]+g\left(\frac{x_c}{x}\right)[P_2(x)+R(x)],$$

$$h(x;\sqrt{z}/\rho)=\int_x^{\sqrt{z}/\rho}\frac{dt}{t}P_1(t)P_1\left(\frac{x}{t}\right)=\frac{1+x^2}{1-x}\left(\frac{3}{2}+2\ln\frac{(\sqrt{z}/\rho-x)(1-x)}{(1-\sqrt{z}/\rho)x}\right)-1+x-\frac{\sqrt{z}}{\rho}+\frac{x\rho}{\sqrt{z}}-(1+x)\ln\frac{\sqrt{z}}{x\rho}.$$

Note that the substitutions inside the straight brackets concern either the limits of x-integration or the expressions under the x-integral sign.

TABLE I. The SABS cross section (in nb) with first- and second-order photonic correction.

	First-order correction				Second-order correction			
x_c	BHLUMI ww	ww	nn	wn	BHLUMI ww	ww	nn	wn
0.1	166.046	166.008	130.813	134.504	166.892	166.958	131.674	134.808
0.3	164.740	164.702	129.797	133.416	165.374	165.447	130.524	133.583
0.5	162.241	162.203	128.001	131.428	162.530	162.574	128.474	131.127
0.7	155.431	155.390	122.922	125.809	155.668	155.597	123.206	125.225
0.9	134.390	134.334	106.478	107.945	137.342	137.153	108.820	109.667

In the case of CES the right side of Eq. (77) requires the following modifications: i) the coefficient at $P_3(x)$ must be reduced eight times, the coefficients at $P_2(x)$ and $R^p(x)$ must be reduced four times; ii) it must be assumed that $h=0$ and $H^c(x,x_c)$ must be inserted instead of $H(x,x_c)$, where

$$H^c(x,x_c)=P_1(x)\left[\frac{1}{2}\,f\left(\frac{x_c}{x}\right)+\frac{2}{3}\,g\left(\frac{x_c}{x}\right)+\frac{1}{2}\,r\left(\frac{x_c}{x};1\right)\right]+\frac{1}{2}\,g\left(\frac{x_c}{x}\right)[P_2(x)+R(x)].$$

5. THE NUMERICAL RESULTS

The numerical calculations were carried out for the beam energy $\epsilon=46.15$ GeV, and the limited angles of the circular detectors were taken from Eq. (3). The Born cross section

$$\Sigma_B=\frac{4\pi\alpha^2}{Q_1^2}\int_{\rho_2^2}^{\rho_4^2}\frac{dz}{z^2}\left(1-\theta_1^2\,\frac{z}{2}\right)$$

(in the symmetrical wide-wide case the limits of integration are 1 and ρ_3^2) is 175.922 nb for the *ww* angular acceptance and 139.971 nb for the *nn* and *wn* angular acceptances.

The results of our calculations of the QED correction with the switched-off vacuum polarization are shown in Tables I–III. For comparison we give also the corresponding numbers derived with the help of Monte Carlo program BHLUMI.[3]

As one can see from Table I there is an approximately constant difference on the 0.3% level between our analytical and Monte Carlo results within first-order correction. The possible reason for this effect is as follows. In the analytical calculation we systematically ignore the terms with $\theta^2 \simeq |t|/s$ as compared with unity. It is well known, however, that such terms have double logarithmic, asymptotic,[17] and parametrically equal $(\alpha|t|/\pi s)\ln^2(|t|/s)$, which is just 0.1% for the LEP1 conditions. Monte Carlo BHLUMI program, to the best of our knowledge, takes into account all first-order contributions.[18]

Table II gives the absolute values of the second-order correction to the SABS cross section, with allowance for the leading and next-to-leading contributions. The correction due to the pair production is small, in agreement with the results of the Ref. 6. The second-order photonic correction is represented as a sum of the leading contribution and next-to-leading one. As one can see, the next-to-leading part is not negligible.

Table III gives the absolute value of the leading third-order correction and the SABS cross section with all corrections obtained in this work. The third-order correction takes into account the three-photon emission and pair production which is accompanied by single-photon radiation. At large values of x_c this correction is comparable with the second-order next-to-leading correction. This effect increases under the conditions of LEP2.

6. CONCLUSIONS

In this paper we give the analytic calculation of the QED correction to SABS cross section at LEP1 for the case of inclusive event selection and wide-narrow angular acceptance. They include the leading and next-to-leading contributions in first and second orders of perturbation theory and leading contribution in third order. The leading contributions in the case of calorimeter event selection are obtained for any form of the final electron and positron cluster. The result is represented in the form of a manifold integral with definite limits, and the functions under integral sign have no physical singularities. No problem arises with infrared divergence and double counting.

The selection of essential Feynman diagrams, utilization of Sudakov's variables, which are relevant to this problem, impact factor representation of the differential cross section

TABLE II. The second-order absolute correction to the SABS cross section (in nb).

	Pair production			Two-photon emission		
x_c	ww	nn	wn	ww	nn	wn
0.1	0.007	−0.004	0.015	0.742+0.208	0.679+0.182	0.249+0.091
0.3	−0.033	−0.033	−0.020	0.546+0.199	0.556+0.171	0.069+0.098
0.5	−0.058	−0.050	−0.041	0.140+0.231	0.291+0.182	−0.314+0.134
0.7	−0.090	−0.074	−0.069	−0.027+0.234	0.117+0.187	−0.571+0.170
0.9	−0.142	−0.115	−0.115	2.961−0.142	2.458−0.116	1.822−0.090

TABLE III. Leading third-order correction and SABS cross section as obtained in this work.

	Third-order correction			SABS cross section at LEP1		
x_c	ww	nn	wn	ww	nn	wn
0.1	−0.055	−0.047	−0.006	166.910	131.623	134.817
0.3	−0.065	−0.053	−0.018	165.349	10.438	133.545
0.5	−0.036	−0.040	0.004	162.472	128.384	131.090
0.7	0.089	0.058	0.124	155.596	123.190	125.310
0.9	0.291	0.220	0.331	137.307	108.927	109.893

due to the t-channel photon exchange, and the electron structure function method and investigation of underlying kinematics were very useful in this work. We emphasize separately the simplifications connected with the impact factor representation, which allows us to represent the differential cross sections of two-jet processes in QED in factorized form. The latter allows us to use the cutoff θ-functions for the final electron and positron independently at the level of differential cross section. The calculation does not require us to go to c.m.s. of the underlying subprocess (as in Ref. 6) and avoids the corresponding complications.

At this point, we wish to comment on the analytical calculation of the leading contribution due to the photon emission and pair production carried out in Ref. 6. Authors of those articles used as the master formula for description of the QED corrections to the SABS cross section due to the initial-state radiation the representation valid for the cross sections of the Drell–Yan process,[19] electron-positron annihilation into muons (or hadrons),[20] and large-angle Bhabha scattering.[21] In this set, however, the SABS process has a very particular feature: two different scales exist only for it. The first one is the momentum transfer squared t; this scale defines the cross section. The second scale is the total c.m.s. energy squared $s=4\epsilon^2$ and $\theta^2\sim|t|/s\ll 1$ has the status of a small correction.

The t-scale physics is very simple and is defined by the peripheral interaction of the electron and positron due to the one-photon exchange, provided that the momentum transfer is strictly perpendicular: $t=-\mathbf{q}^2$. The s-scale physics is more complicated. At the Born level it is seen as a contribution of the annihilation diagram and also permits the energy and longitudinal momentum exchange for the contribution of the scattering diagram. The first-order QED correction for the s-scale cross section includes the contributions of the box diagrams, the large-angle photon emission and the up-down interference, because the eikonal representation for the scattering amplitude and the factorization form of the differential cross section break down. In the second order large-angle pair production and two-photon emission appear.

The structure function used in Ref. 6 controls the t-scale cross section only and is not related to the s-scale cross section, because physics of different scales evolves by its own laws. This is well known from the analysis of such different problems of physics as, for example, higher twist corrections in QCD[22] and turbulence phenomenon in hydrodynamics.[23]

On the other hand, only the scattered diagram contributes to the Born cross section used in Ref. 6. But every time the annihilation diagram, as compared with the scattering diagram, is neglected, one must automatically neglect θ^2, as compared with unity everywhere including the Born cross section and the experimental cuts, in order to be consistent. Taking into account these arguments, we must simplify the master formula in Ref. 6 by eliminating the terms proportional to $\xi\sim|t|/s\ll 1$ and ξ^2 in the numerator of Eq. (5) and in the cutoff restrictions. It then becames adequate to the one obtained in Ref. 10 and the one used in this work.

Numerical evaluations show a good agreement with Monte Carlo calculations within first-order correction, but an agreement for higher-order corrections will require additional efforts.

I thank E. Kuraev and L. Trentadue for fruitful discussions and critical remarks. I also thank A. Arbuzov and G. Gach for assistance with the numerical calculations. This work was supported by INTAS Grant No 93-1867.

APPENDIX A

Let us first consider the next-to-leading, second-order, Δ-independent contribution due to the one-side, two-photon emission. We first give the analytical expression for the symmetrical case, because it was not published until now. (I do not give special notation for the next-to-leading contribution to Σ, keeping in mind that only such terms are considered in this Appendix):

$$\Sigma^{\gamma\gamma}=\Sigma_{\gamma\gamma}=\frac{1}{4}\left(\frac{\alpha}{\pi}\right)^2\int_1^{\rho^2}\frac{dz}{z^2}\,LY, \tag{A1}$$

$$\begin{aligned}Y=y+\int_{x_c}^1 dx\Bigg\{A+\int_0^{1-x}dx_1\Bigg[\frac{1}{x_1}4\frac{1+x^2}{1-x}\,(\theta_\rho^{(x)}l_1+l_2)\\+\left(-1-\frac{1+x}{1-x_1}-\frac{x}{(1-x_1)^2}\right)(l_4+\theta_\rho^{(x)}l_3\\+2\,\theta_\rho^{(1-x_1)}l_5)+\frac{2(1+x)}{1-x_1}\,\theta_\rho^{(1-x_1)}\Bigg]-4\frac{1+x^2}{1-x}\,\overline{\theta_\rho^{(x)}}\\\times\Bigg[\int_{1-\sqrt{z}/\rho}^{1-x}dx_1\left(\frac{1}{x_1}\,l_5+\frac{2}{x_2}\ln\frac{x}{1-x_1}\right)\\+\int_0^{\sqrt{z}/\rho-x}\frac{dx_1}{x_1}\,l_6\Bigg]\Bigg\},\end{aligned}$$

$$y=12\zeta_3+10\zeta_2-\frac{45}{4}-16\ln^2(1-x_c)-28\ln(1-x_c),$$

$$\begin{aligned}A=(1+\theta_\rho^{(x)})\left[2(5+2x)+4(x+3)\ln(1-x)+4\frac{1+x^2}{1-x}\ln x\right]\\+2\frac{1+x^2}{1-x}\Bigg[\left(\frac{3}{2}-\ln x\right)K(x,z;\rho,1)-\frac{1}{2}\ln^2\\\times x-\frac{(1-x)^2}{2(1+x^2)}+2\ln(1-x)\Bigg(\theta_\rho^{(x)}\ln\left|\frac{x^2\rho^2-z}{x\rho^2-z}\right|\\+\ln\left|\frac{(z-1)(z-x^2)(\rho^2-z)}{(z-x)^2(x\rho^2-z)}\right|\Bigg)\Bigg]+\overline{\theta_\rho^{(x)}}\end{aligned}$$

$$\times\left[\frac{16}{1-x}\ln(1-x)+\frac{14}{1-x}-(1-x)\ln\right.$$

$$\times x+2\frac{1+x^2}{1-x}\left(-\frac{3}{2}\ln^2 x+3\ln x\,\ln(1-x)-L_{i2}(1-x)\right.$$

$$\left.-\frac{x(1-x)+4x\,\ln x}{2(1+x^2)}+\frac{(1+x)^2}{1+x^2}\ln\left|\frac{(\sqrt{z}-x\rho)}{\rho-\sqrt{z}}\right|\right.$$

$$\left.+2\ln\left|\frac{\sqrt{z}-x\rho}{\rho}\right|\ln\left|\frac{x(x\rho^2-z)}{x^2\rho^2-z}\right|\right)\Bigg],$$

$$l_1=\ln\left|\frac{(x^2\rho^2-z)(x\rho^2-z)}{(x(1-x_1)\rho^2-z)(x(x+x_1)\rho^2-z)}\right|,$$

$$l_3=\ln\left|\frac{(1-x_1)^2(1-x-x_1)(x^2\rho^2-z)^2}{x^3x_1(x(1-x_1)\rho^2-z)^2}\right|,$$

$$l_2=\ln\left|\frac{(z-x)^2(z-(1-x_1)^2)(z-(x+x_1)^2)}{(z-(1-x_1))(z-x(1-x_1))((x+x_1)-z)(x(x+x_1)-z)}\right|+\ln\left|\frac{((1-x_1)^2\rho^2-z)((x+x_1)^2\rho^2-z)(x\rho^2-z)}{((x+x_1)\rho^2-z)((1-x_1)\rho^2-z)(x^2\rho^2-z)}\right|,$$

$$l_4=\ln\left|\frac{(1-x_1)^2xx_1(z-1)(z-x^2)(z-(1-x_1)^2)^2}{x_2(z-(1-x_1))^2(z-x(1-x_1))^2}\right|$$

$$+\ln\left|\frac{(\rho^2-z)(x(1-x_1)\rho^2-z)^2}{(x^2\rho^2-z)((1-x_1)^2\rho^2-z)^2}\right|,$$

$$l_5=\ln\left|\frac{x((1-x_1)^2\rho^2-z)^2}{(1-x_1)^2(x(1-x_1)\rho^2-z)((1-x_1)\rho^2-z)^2}\right|,$$

$$l_6=\ln\left|\frac{(x\rho^2-z)((x+x_1)^2\rho^2-z)^2}{(x^2\rho^2-z)(x(x+x_1)\rho^2-z)((x+x_1)\rho^2-z)}\right|.$$

For the wide-narrow angular acceptance we need to consider only the case of the positron emission $\Sigma_{\gamma\gamma}$, because the corresponding expression for the electron emission $\Sigma^{\gamma\gamma}$ is Eq. (A1) with (ρ_4^2,ρ_2^2) as the limits of z-integration and ρ_3 instead ρ under the integral sign.

The analytical expression for $\Sigma_{\gamma\gamma}$ has the form

$$\Sigma_{\gamma\gamma}=\frac{1}{4}\left(\frac{\alpha}{\pi}\right)^2\int_1^{\rho_3^2}\frac{dz}{z^2}\,L\,A_N^W, \tag{A2}$$

$$A_N^W=y\Delta_{42}+\int_{x_c}^1 dx\left\{\Delta_{42}\left[4(4+3x)+6(x+3)\ln(1-x)\right.\right.$$

$$\left.+\left(x-1+4\frac{1+x^2}{1-x}\right)\ln x\right]+\Delta_{42}^{(x)}\left[(1-x)(3+\ln x)\right.$$

$$\left.+2(x+3)\ln(1-x)+4\frac{1+x^2}{1-x}\ln x\right]+\overline{\Delta_{42}^{(x)}}\,\frac{2}{1-x}$$

$$\times(4+(1+x)^2)\ln(1-x)$$

$$+2\frac{(1+x)^2}{1-x}\left(\theta_4\overline{\theta_4^{(x)}}\ln\left|\frac{\sqrt{z}-x\rho_4}{\rho_4-\sqrt{z}}\right|\right.$$

$$\left.-\theta_2\overline{\theta_2^{(x)}}\ln\left|\frac{\sqrt{z}-x\rho_2}{\rho_2-\sqrt{z}}\right|\right)+\frac{1+x^2}{1-x}B$$

$$+\int_0^{1-x}dx_1\left[2\frac{1+x^2}{(1-x)x_1}(\Delta_{42}^{(x)}l_{1+}+\Delta_{42}l_{2+}+(\overline{\theta_4^{(x)}}\right.$$

$$\left.-\theta_2^{(x)})l_{1-}+(\overline{\theta_4}-\theta_2)l_{2-})+\left(-1-\frac{1+x}{1-x_1}\right.\right.$$

$$\left.\left.-\frac{x}{(1-x_1)^2}\right)\left(\Delta_{42}^{(x)}\left(\ln\frac{(1-x_1)^2x_2}{x^3x_1}+l_{3+}\right)\right.\right.$$

$$+\Delta_{42}\left(\ln\frac{(1-x_1)^2xx_1}{x_2}+l_{4+}\right)+\Delta_{42}^{(1-x_1)}$$

$$\times\left(2\ln\frac{x}{(1-x_1)^2}+l_{5+}\right)+(\overline{\theta_4^{(x)}}-\theta_2^{(x)})l_{3-}+(\overline{\theta_4}$$

$$\left.-\theta_2)l_{4-}+(\overline{\theta_4^{(1-x_1)}}-\theta_2^{(1-x_1)})l_{5-}\right)$$

$$\left.+2\frac{1+x}{1-x_1}\Delta_{42}^{(1-x_1)}\right]+2\frac{1+x^2}{1-x}\theta_4\overline{\theta_4^{(x)}}$$

$$\times\left[\int_{1-\sqrt{z}/\rho_4}^{1-x}dx_1\left(\frac{1}{x_1}\widetilde{l}_6-\frac{4}{x_2}\ln\frac{x}{1-x_1}\right)\right.$$

$$\left.+\int_0^{\sqrt{z}/\rho_4-x}\frac{dx_1}{x_1}\widetilde{l}_7\right]+2\frac{1+x^2}{1-x}\theta_2\overline{\theta_2^{(x)}}$$

$$\times\left[\int_{1-\sqrt{z}/\rho_2}^{1-x}dx_1\left(\frac{1}{x_1}\widetilde{l}_6+\frac{4}{x_2}\ln\frac{x}{1-x_1}\right)\right.$$

$$\left.+\int_0^{\sqrt{z}/\rho_2-x}\frac{dx_1}{x_1}\widetilde{l}_7\right],$$

$$B=\Delta_{42}\left(-2\ln^2 x+2\ln(1\right.$$

$$\left.-x)\ln\left|\frac{x^4(z-\rho_2^2)^2(z-x^2\rho_2^2)(x^2\rho_4^2-z)(\rho_4^2-z)^2}{(z-x\rho_2^2)^3(x\rho_4^2-z)^3}\right|\right)$$

$$+\Delta_{42}^{(x)}\left(\ln^2 x+2\ln(1-x)\ln\left|\frac{(z-x^2\rho_2^2)(x^2\rho_4^2-z)}{x^4(z-x\rho_2^2)(x\rho_4^2-z)}\right|\right)$$

$$+(3-2\ln x)\widetilde{K}(x,z;\rho_4,\rho_2)+\overline{\Delta_{42}^{(x)}}\left(7-2\ln x\,\ln(1\right.$$

$$\left.-x)-2\ln^2 x-2L_{i2}(1-x)-\frac{x(1-x)+4x\,\ln x}{1+x^2}\right)$$

$$+2(\overline{\theta_4}-\theta_2)\ln(1$$

$$-x)\ln\left|\frac{(x\rho_4^2-z)^3(z-\rho_2^2)^2(z-x^2\rho_2^2)}{(\rho_4^2-z)^2(x^2\rho_4^2-z)(z-x\rho_2^2)^3}\right|+2\Bigg(\overline{\theta_4^{(x)}}$$

$$-\theta_2^{(x)}\ln(1-x)\ln\left|\frac{(z-x^2\rho_2^2)(x\rho_4^2-z)}{(x^2\rho_4^2-z)(x\rho_2^2-z)}\right|$$

$$+4\theta_4\overline{\theta_4^{(x)}}\ln\left|\frac{x\rho_4-\sqrt{z}}{\rho_4}\right|\ln\left|\frac{x(x\rho_4^2-z)}{x^2\rho_4^2-z}\right|$$

$$+4\theta_2\overline{\theta_2^{(x)}}\ln\left|\frac{\sqrt{z}-x\rho_2}{\rho_2}\right|\ln\left|\frac{z-x^2\rho_2^2}{x(z-x\rho_2^2)}\right|,$$

$$l_{1\pm}=(1\pm\hat{c})\ln\left|\frac{(z-x^2\rho_2^2)(z-x\rho_2^2)}{(z-x(1-x_1)\rho_2^2)(z-x(x+x_1)\rho_2^2)}\right|,$$

$$l_{2\pm}=(1\pm\hat{c})\left[\ln\left|\frac{(z-x\rho_2^2)^3(z-(1-x_1)^2\rho_2^2)^2(z-(x+x_1)^2\rho_2^2)^2}{(z-x^2\rho_2^2)(z-x(1-x_1)\rho_2^2)(z-x(x+x_1)\rho_2^2)(z-(1-x_1)\rho_2^2)^2(z-(x+x_1)\rho_2^2)^2}\right|\right],$$

$$l_{3\pm}=(1\pm\hat{c})\ln\left|\frac{z-x^2\rho_2^2}{z-x(1-x_1)\rho_2^2}\right|,$$

$$l_{4\pm}=(1\pm\hat{c})\ln\left|\frac{z-\rho_2^2}{z-(1-x_1)\rho_2^2}\right|,$$

$$l_{5\pm}=(1\pm\hat{c})\ln\left|\frac{(z-(1-x_1)^2\rho_2^2)^2}{(z-x(1-x_1)\rho_2^2)(z-(1-x_1)\rho_2^2)}\right|,$$

$$\tilde{l}_6=\ln\left|\frac{x^2(z-(1-x_1)^2\rho_2^2)^4}{(1-x_1)^4(z-x(1-x_1)\rho_2^2)^2(z-(1-x_1)\rho_2^2)^2}\right|,$$

$$\tilde{l}_7=\ln\left|\frac{(z-x\rho_2^2)^2(z-(x+x_1)^2\rho_2^2)^4}{(z-x^2\rho_2^2)^2(z-x(x+x_1)\rho_2^2)^2(z-(x+x_1)\rho_2^2)^2}\right|,$$

$$\tilde{l}_6=-\hat{c}\tilde{l}_6,\quad \tilde{l}_7=-\hat{c}\tilde{l}_7,$$

where $x_2=1-x-x_1$, and $\hat{c}$ is the operator of the substitution:

$$\hat{c}f(\rho_2)=f(\rho_4). \tag{A3}$$

It can be verified that in the symmetrical limit Eq. (A2) coincides with Eq. (A1).

For the opposite-side emission the next-to-leading contribution to Σ in the symmetrical case is

$$\Sigma_\gamma^\gamma=\left(\frac{\alpha}{\pi}\right)^2 L\int_0^\infty\frac{dz}{z^2}\,T, \tag{A4}$$

$$T=A\,\theta_\rho\overline{\theta_1}-\int_{x_c}^1 dx\left[\frac{1+x^2}{2(1-x)}\,N(x,z;\rho,1)+\Xi(x)\right.$$

$$\left.+\frac{\overline{\Xi}(x)}{1-x}\right]\int_{x_c/x_1}^1 dx_1\left[(1+x_1)\Xi(x_1)+\frac{2\overline{\Xi}(x_1)}{1-x_1}\right], \tag{A5}$$

where

$$A=-6-14\ln(1-x_c)-8\ln^2(1-x_c)+\int_{x_c}^1 dx\Bigg\{7(1+x)$$

$$+\frac{1+x^2}{2(1-x)}\,[3K(x,z;\rho,1)+7\overline{\theta_\rho^{(x)}}]+2\ln\frac{x-x_c}{x}\Bigg[(3$$

$$+x)(1+\theta_\rho^{(x)})+\frac{4}{1-x}\,\overline{\theta_\rho^{(x)}}+\frac{1+x^2}{1-x}\,N(x,z;\rho,1)\Bigg]$$

$$+\frac{8}{1-x}\ln\frac{x(1-x_c)}{x-x_c}. \tag{A6}$$

We introduce the following reduced notation for the θ-functions:

$$\Xi(x)=\theta_\rho\overline{\theta_1}+\theta_\rho^{(x)}\overline{\theta_1^{(x)}},\quad \Xi(x)=\theta_\rho\overline{\theta_\rho^{(x)}}-\theta_1\overline{\theta_1^{(x)}}. \tag{A7}$$

The quantity $K(x,z;\rho,1)$ in the expression for A is the K-factor for the single-photon emission, and the quantity $N(x,z;\rho,1)$ can be derived with the help of Eq. (10) in the following way:

$$N(x,z;\rho,1)=\Bigg(\tilde{K}(x,z;\rho_4,\rho_2)$$

$$\left.-\frac{(1-x)^2}{1+x^2}\,(\Delta_{42}+\Delta_{42}^{(x)})\right)\Bigg|_{\rho_4=\rho,\rho_2=1}. \tag{A8}$$

Note that $N(1,z;\rho,1)=0$.

In the wide-narrow case the corresponding formula for Σ_γ^γ may be written as follows:

$$\Sigma_\gamma^\gamma=\frac{\alpha^2}{\pi^2}\,L\int_0^\infty\frac{dz}{z^2}\,T_N^W, \tag{A9}$$

where

$$T_N^W=\tilde{A}-\frac{1}{2}\Bigg\{\int_{x_c}^1 dx\left[\frac{1+x^2}{2(1-x)}\,N(x,z;\rho_3,1)+\Xi_{31}(x)\right.$$

$$\left.+\frac{1}{1-x}\,\overline{\Delta_{31}^{(x)}}\right]\int_{x_c/x}^1 dx_1\left[(1+x_1)\Xi_{42}(x)\right.$$

$$\left.+\frac{2}{1-x_1}\,\overline{\Delta_{42}^{(x)}}\right]+\int_{x_c}^1 dx\left[\frac{1+x^2}{2(1-x)}\,N(x,z;\rho_4,\rho_2)\right.$$

$$\left.+\Xi_{42}(x)+\frac{1}{1-x}\,\overline{\Delta_{42}^{(x)}}\right]\int_{x_c/x}^1 dx_1\left[(1+x_1)\Xi_{31}(x)\right.$$

$$\left.+\frac{2}{1-x_1}\,\overline{\Delta_{31}^{(x)}}\right]\Bigg\}, \tag{A10}$$

where

$$\widetilde{A}=(-6-14\ln(1-x_c)-8\ln^2(1-x_c))\Delta_{42}$$
$$+\int_{x_c}^1 dx\Bigg\{\Delta_{42}\Bigg[7(1+x)+\frac{8}{1-x}\ln\frac{x(1-x_c)}{x-x_c}\Bigg]$$
$$+\frac{1+x^2}{2(1-x)}\Bigg[\frac{3}{2}\Delta_{42}\overline{K}(x,z;\rho_3,1)+\frac{3}{2}\Delta_{31}\widetilde{K}(x,z;\rho_4,\rho_2)$$
$$+\frac{7}{2}(\Delta_{42}\overline{\Delta}_{31}^{(x)}+\Delta_{31}\overline{\Delta}_{42}^{(x)})\Bigg]+\ln\frac{x-x_c}{x}\Bigg[(3+x)(\Delta_{31}\Xi_{42}(x)$$
$$+\Delta_{42}\Xi_{31}(x))+\frac{4}{1-x}(\overline{\Delta}_{42}^{(x)}\Delta_{31}+\overline{\Delta}_{31}^{(x)}\Delta_{42})$$
$$+\frac{1+x^2}{1-x}(\Delta_{42}N(x,z;\rho_3,1)+\Delta_{31},N(x,z;\rho_4,\rho_2))\Bigg]\Bigg\}, \tag{A11}$$

and

$$\Xi_{42}(x)=\theta_4\overline{\theta_2}+\theta_4^{(x)}\overline{\theta_2^{(x)}}=\Delta_{42}+\Delta_{42}^{(x)},$$

$$\Xi_{31}(x)=\Delta_{31}+\Delta_{31}^{(x)},\quad \overline{\Delta}_{31}^{(x)}=\Delta_{31}-\Delta_{31}^{(x)}.$$

It is obvious that in the symmetrical limit Eq. (9) coincides with (4).

APPENDIX B

Here we give some relations which are used in the analytical calculations and which may be useful for the numerical computations.

For the case of the emission along the electron momentum direction these relations are

$$\int_{\rho_2^2}^{\rho_4^2}dz\int_{x_c}^1 dx\,\overline{\theta_3^{(x)}}=\int_{\rho_2^2}^{\rho_4^2}dz\,\overline{\theta_3^{(x_c)}}\int_{x_c}^{\sqrt{z}/\rho_3}dx,$$

$$\int_{\rho_2^2}^{\rho_4^2}dz\int_{x_c}^1 dz\int_0^{1-x}dx_1\,\overline{\theta_3^{(1-x_1)}}$$
$$=\int_{\rho_2^2}^{\rho_4^2}dz\,\overline{\theta_3^{(x_c)}}\int_{x_c}^{\sqrt{z}/\rho_3}dx\int_{1-\sqrt{z}/\rho_3}^{1-x}dx_1. \tag{B1}$$

For the case of the emission along the positron direction they are

$$\int_1^{\rho_3^2}dz\int_{x_c}^1 dx[\overline{\theta_4^{(x)}}-\theta_2^{(x)}]$$
$$=\int_1^{\rho_3^2}dz\int_{x_c}^1 dx[\overline{\theta_4}-\theta_2+\theta_4\overline{\theta_4^{(x)}}+\theta_2\overline{\theta_2^{(x)}}]$$
$$=\int_1^{\rho_3^2}dz\Bigg\{(\overline{\theta_4}-\theta_2)\int_{x_c}^1 dx+\theta_4\overline{\theta_4^{(x_c)}}\int_{x_c}^{\sqrt{z}/\rho_4}dx$$
$$+\theta_2\overline{\theta_2^{(x_c)}}\int_{x_c}^{\sqrt{z}/\rho_2}dx\Bigg\},$$
$$\int_1^{\rho_3^2}dz\int_{x_c}^1 dx\int_0^{1-x}dx_1[\overline{\theta_4^{(1-x_1)}}-\theta_2^{(1-x_1)}]$$
$$=\int_1^{\rho_3^2}dz\int_{x_c}^1 dx\Bigg\{(\overline{\theta_4}-\theta_2)+\int_0^{1-x}dx_1$$
$$+\overline{\theta_4^{(x_c)}}\theta_4\int_{x_c}^{\sqrt{z}/\rho_4}dx\int_{1-\sqrt{z}/\rho_4}^{1-x}dx_1$$
$$+\overline{\theta_2^{(x_c)}}\theta_2\int_{x_c}^{\sqrt{z}/\rho_2}dx\int_{1-\sqrt{z}/\rho_2}^{1-x}dx_1\Bigg\}. \tag{B2}$$

Some additional relations arise for the case of opposite-side emission. Let us consider first the integration-limits restrictions for the product of the θ-functions in the symmetrical case

$$\theta_3\overline{\theta_3^{(x_1)}}\overline{\theta_3^{(x_2)}},\quad \theta_1\overline{\theta_3^{(x_1)}}\overline{\theta_1^{(x_2)}},\quad \theta_1\overline{\theta_1^{(x_1)}}\overline{\theta_1^{(x_2)}}. \tag{B3}$$

At first, we use Eqs. (B1) and eliminate $\overline{\theta_i^{(x_2)}}$ using the following changes: 1) $\overline{\theta_i^{(x_2)}}\to\overline{\theta_i^{(x_c/x_1)}}$, 2) the upper limit of x_2 integration in the case of $\overline{\theta_3^{(x_2)}}$ must be replaced by $\sqrt{z}/\rho_3$ and in the case of $\overline{\theta_1^{(x_2)}}$ by $\sqrt{z}$.

Thus we have three regions defined by the quantities in (z,x_1) plane:

$$\rho^2=z,\quad z=x_1^2\rho^2,\quad z=\frac{x_c^2\rho^2}{x_1^2},$$

$$1=z,\quad z=x_1^2\rho^2,\quad z=\frac{x_1}{x_c^2},$$

$$1=z,\quad z=x_1^2,\quad z=\frac{x_1^2\rho^2}{x_c^2}. \tag{B4}$$

The limits of integrations may be transformed as follows:

$$\int\ \theta_3\overline{\theta_3^{(x_1)}}\overline{\theta_3^{(x_2)}}\to\int_{x_c\rho^2}^{\rho^2}dz\int_{x_c\rho/\sqrt{z}}^{\sqrt{z}/\rho}dx_1\int_{x_c/x_1}^{\sqrt{z}/\rho}dx_2,$$
$$\int\ \theta_3\overline{\theta_1^{(x_1)}}\overline{\theta_1^{(x_2)}}\to\int_{x_c\rho}^{1}dz\int_{x_c/\sqrt{z}}^{\sqrt{z}/\rho}dx_1\int_{x_c/x_1}^{\sqrt{z}}dx_2, \tag{B5}$$

and for $\int\theta_1\overline{\theta_1^{(x_1)}}\overline{\theta_1^{(x_2)}}$ the formulas can be derived from the above formulas by setting $\rho=1$. For the wide-narrow case the prescription is similar:

$$\int\ \theta_4\overline{\theta_4^{(x_1)}}\overline{\theta_3^{(x_2)}}\to\int_{x_c\rho_3}^{\rho_4^2}dz\int_{x_c\rho_3/\sqrt{z}}^{\sqrt{z}/\rho_4}dx_1\int_{x_c/x_1}^{\sqrt{z}/\rho_3}dx_2. \tag{B6}$$

The other variants of the restrictions in the wide-narrow angular acceptance may be written as follows:

$$\int\ \theta_1\overline{\theta_2^{(x_1)}}\overline{\theta_1^{(x_2)}}\to\int_{x_c\rho_2}^{1}dz\int_{x_c\rho_3/\sqrt{z}}^{\sqrt{z}}dx_1\int_{x_c/x_1}^{\sqrt{z}/\rho_2}dx_2,$$
$$\int\ \theta_1\overline{\theta_4^{(x_1)}}\overline{\theta_1^{(x_2)}}\to\int_{x_c\rho_4}^{1}dz\int_{x_c\rho_4/\sqrt{z}}^{\sqrt{z}}dx_1\int_{x_c/x_1}^{\sqrt{z}/\rho_4}dx_2,$$
$$\int\ \theta_2\overline{\theta_2^{(x_1)}}\overline{\theta_3^{(x_2)}}\to\int_{x_c\rho_2\rho_3}^{\rho_2^2}dz\int_{x_c\rho_3/\sqrt{z}}^{\sqrt{z}/\rho_2}dx_1\int_{x_c/x_1}^{\sqrt{z}/\rho_3}dx_2. \tag{B7}$$

[1] The LEP Collaboration: ALEPH, DELP1, L3 and OPAL and the LEP Electroweak Working Group, CERN-PPE/95; B. Pietrzyk, preprint LAPP-Exp-94.18, *Invited talk at the Conf. Radiative Corrections: Status and Outlook*, Galtinburg, TN, USA, 1994; I. C. Brock *et al.*, Preprint CERN-PPE/96-89, CMU-HEP/96-04, 1996.
[2] G. Barbiellini *et al.*, *Neutrino Counting in Z. Physics at LEP*, conv. L. Trentadue, ed. by G. Altarelli, R. Kleiss, and C. Verzegnassi, CERN Report 89-08.
[3] H. Anlauf *et al.*, *Events Generator for Bhabha Scattering*, conv. S. Jadach and O. Nicrosini, Yell. Rep. CERN 96-01, 2, 229.
[4] S. Jadach, E. Richter-Was, B. F. L. Ward, and Z. Was, Comput. Phys. Commun. **70**, 305 (1992).
[5] G. Montagna *et al.*, Comput. Phys. Commun. **76**, 328 (1993); M. Cacciori, G. Montagna, and F. Piccinini, Comput. Phys. Commun. **90**, 301 (1995); CERN-TH/95-169; G. Montanga *et al.*, Nucl. Phys. B **401**, 3 (1993).
[6] S. Jadach, M. Skrzypek, and B. F. L. Ward, Phys. Rev. D **47**, 3733 (1993); S. Jadach, E. Richter-Was, B. F. L. Ward, and Z. Was, Phys. Lett. B **260**, 438 (1991).
[7] S. Jadach, E. Richter-Was, B. F. L. Ward, and Z. Was, Phys. Lett. B **353**, 349,362 (1995); S. Jadach, M. Melles, B. F. L. Ward, and S. A. Yost, Phys. Lett. B **377**, 168 (1966).
[8] W. Beenakker, F. A. Berends, and S. C. van der Marck, Nucl. Phys. B **355**, 281 (1991); W. Beenakker and B. Pietrzyk, Phys. Lett. B **304**, 366 (1993).
[9] M. Gaffo, H. Czyz, E. Remiddi, Nuovo Cimento A **105**, 271 (1992); Int. J. Mod. Phys. **4**, 591 (1993); Phys. Lett. B **327**, 369 (1994); G. Montanga, O. Nicrosini, and F. Piccinini, Preprint FNT/T-96/8.
[10] A. B. Arbuzov *et al.*, Yell. Rep. CERN 95-03, p. 369; preprint CERN-TH/95-313, UPRF-95-438; Nucl. Phys. B **485**, 457 (1997).
[11] A. B. Arbuzov, E. Kuraev, N. P. Merenkov, and L. Trentadue, Zh. Éksp. Teor. Fiz. **108**, 1164 (1995) [JETP **81**, 638 (1995)]; Preprint CERN-TH/95-241, JINR-E2-95-110.
[12] D. R. Yennie, S. C. Frautchi, and H. Suura, Ann. Phys. (N.Y.) **13**, 379 (1961).
[13] L. N. Lipatov, Yad. Fiz. **20**, 94 (1974); [Sov. J. Nucl. Phys. **20**, 48 (1974)]; G. Altarelli and G. Parisi, Nucl. Phys. B **126**, 298 (1977); M. Skrzypek, Acta Phys. Pol. B **23**, 135 (1992).
[14] N. P. Merenkov, Yad. Fiz. **48**, 1693 (1988); [Sov. J. Nucl. Phys. **48**, 1073 (1988)]; Yad. Fiz. **50**, 750 (1989) [Sov. J. Nucl. Phys. **50**, 469 (1989)].
[15] T. D. Lee and M. Nauenberg, Phys. Rev. B **133**, 1549 (1964).
[16] H. Cheng and T. T. Wu, Phys. Rev. Lett. **23**, 670 (1969); V. G. Zima and N. P. Merenkov, Yad. Fiz. **25**, 998 (1976); V. N. Baier, V. S. Fadin, V. Khoze, and E. Kuraev, Phys. Rep. **78**, 294 (1981).
[17] V. G. Gorshkov, Usp. Fiz. Nauk **110**, 45 (1973) [Sov. Phys. Usp. **16**, 322 (1973)].
[18] S. Jadach and B. W. L. Ward, Phys. Rev. D **40**, 3582 (1989).
[19] S. Drell and T. M. Yan, Phys. Rev. Lett. **25**, 316 (1970).
[20] E. A. Kuraev and V. S. Fadin, Yad. Fiz. **41**, 733 (1985) [Sov. J. Nucl. Phys. **41**, 466 (1985)].
[21] W. Beenakker, F. A. Berends, and S. C. van der Marck, Nucl. Phys. B **349**, 323 (1991).
[22] A. P. Bukhvostov, E. A. Kuraev, and L. N. Lipatov, Yad. Fiz. **38**, 439 (1983) [Sov. J. Nucl. Phys. **38**, 263 (1983)]; Yad. Fiz. **39** 194 (1984) [Sov. J. Nucl. Phys. **39**, 121 (1984)].
[23] R. K. Dodd, J. C. Eilbeck, J. D. Gibbon, and H. C. Morris, *Solitons and Nonlinear Wave Equations*, Academic Press, Inc. (Harcourt Brace Jovanovich, Publishers); S. S. Moiseyev, P. B. Rutkevich, A. V. Tur, and V. V. Yanovsky, Zh. Éksp. Teor. Fiz. **94**, 140 (1988); [Sov. Phys. JETP **67**, 78 (1988)]; A. V. Chechkin, A. V. Tur, and V. V. Yanovsky, Phys. Fluids B **4**, 3513 (1992).

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

Noise squeezing in a semiconductor laser with an inhomogeneously broadened gain line

V. V. Kozlov

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

A. S. Trifonov

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted 9 August 1996)
Zh. Éksp. Teor. Fiz. **112**, 429–440 (August 1997)

This paper is a theoretical analysis of the noise produced by a single-mode semiconductor laser. We allow for spectral hole-burning in the gain line and for the nonlinear dependence of the carrier spontaneous recombination rate on the number of carriers in the active area. We show that these processes do not inhibit the squeezing of the outgoing photon flux noise below the shot noise level at high pump currents but that the degree of squeezing decreases considerably. Finally, we establish that these processes also considerably narrow the squeezing bandwidth.
 [S1063-7761(97)00308-9]

1. INTRODUCTION

At pump currents much higher than the threshold value, the light emitted by a semiconductor laser may exhibit an essentially quantum property: it can be squeezed in the amplitude fluctuations (or noise).[1–3] The latter means that the spectrum of fluctuations of the photon flux leaving the laser cavity becomes narrower than the standard quantum limit. The spectral density of the noise appearing in the course of photocurrent detection becomes smaller than the shot noise level. This effect was first detected experimentally by Machida *et al.*[3]

Thorough measurements have shown[4] that the single-mode theory describes the photon flux noise well for pump currents ranging from the threshold value to values two to three times higher than the threshold value, i.e., to values at which amplitude squeezing can be observed. If the pump current exceeds the above value, the deviations for the values predicted by the theory become considerable. In particular, at room temperature in the free-running mode the noise of the laser radiation is considerably higher than the standard quantum limit, and usually no squeezing is observed. The usual explanation of such behavior is the presence of nonlasing sidebands.[5,6] Experiments corroborate this hypothesis: the noise was reduced below the standard quantum level by introducing selective feedback from the diffraction grating,[6–9] and increasing the Q of the central mode by injecting radiation from a high-stability external laser[10,11] or providing dispersive optical feedback.[12,13]

Inoue *et al.*[14] studied amplitude squeezing in multimode lasers both theoretically and experimentally. They found that the emerging intermode correlation leads to mutual suppression of the contributions of fluctuations of separate modes to the total intensity. In this case the noise of the multimode laser is equivalent to that of an ideal single-mode laser and light squeezing becomes possible, which is demonstrated by experiments conducted by Inoue *et al.*[14] The measurements were made at low temperatures, 20–60 K, and the homogeneity of the gain line of the laser was ensured by the high degree of alloying of the active area. The researchers called this phenomenon multimode squeezing.

It must be noted at this point that multimode squeezing differs from single-mode in many respects. The noise level registered during photodetection of squeezed multimode light may grow if not all the modes reach the photodetector or if the laser light has traveled through a medium with dispersion, e.g., an optical fiber, since squeezing is ensured by the negative correlation of fluctuations of the fundamental and side modes, while the noise of individual modes can exceed the shot noise level considerably.

Moreover, the noise can increase due to the breakdown of intermode correlations as a result of various nonlinear optical processes in the laser proper. In particular, Marin *et al.*[7] showed that the increase in the noise level registered in experiments can be explained by adding to the multimode lasing equation a term quadratic in the field strength, which models the nonlinear interaction between the modes. Satisfactory agreement with the experimental data is achieved by empirically fitting the numerical value of the nonlinear interaction coefficient. Marin *et al.*[7] also showed that the greater the side-mode suppression, the higher the degree of amplitude squeezing in the fundamental mode. At a 60-dB side-mode suppression in relation to the intensity of the fundamental mode, the nonlinear intermode interaction ceases to greatly affect the fundamental-mode noise, and such a lasing mode can be considered single-mode to high accuracy. But will a high degree of amplitude squeezing guarantee complete suppression of side modes? It is well known that nonlinear optical processes in a semiconductor laser operating at pump currents much higher than the threshold level change not only the nature of intermode interaction, but can also substantially affect the nature of single-mode lasing due to self-action effects of the fundamental mode.

The present paper studies the effect of optical inhomogeneities in a single-mode semiconductor laser on the nature of amplitude squeezing. The idea that electron–hole pairs in the active area of the laser interact simultaneously with the field of all laser modes is an approximation valid only near

the lasing threshold. For pump currents much higher than threshold, a localized drop appears in the number of occupied states in the conduction band and in the number of vacant states in the valence band, because of the finite rate of restoration of intraband equilibrium after each recombination event. By its influence on the nature of lasing, this effect has a direct analog in the theory of gas lasers, and has come to be known as spectral hole-burning. The statistical characteristics of the radiation emitted by a gas laser that allow for spectral hole-burning have been studied (see, e.g., Ref. 15). However, semiconductor lasers exhibit a number of special features, which manifest themselves, in particular, in the fact that gain depends explicitly on the number of photons in the cavity.[16]

Spectral hole-burning does not exhaust the list of nonlinear effects characteristic of semiconductor lasers. As is known, a carrier in the active area can take part in radiative recombination or relax nonradiatively. In the first case, the recombination rate is satisfactorily described by a quadratic dependence on the number of carriers. In the second, one or three carriers participate in the nonradiative process, contributing linear and cubic terms in the expression for the spontaneous decay rate (see Ref. 17).[1)] Traditionally, when studying the dynamics of a single-mode semiconductor laser, it is assumed that the spontaneous recombination rate is a linear function of the number of carriers, which is a valid approximation only in the immediate vicinity of the threshold. The applicability of the approximation is all the more justified, as in the steady-state lasing mode, the number of carriers is independent of the pump current and is equal to its threshold value. Gain saturation has a profound effect on the situation: spectral hole-burning leads to an increase in the number of carriers in the active area as the pump current grows, so that it initiates the emergence of the previously masked nonlinearity of the spontaneous decay rate. In other words, a consistent calculation of a semiconductor laser requires allowing for both types of nonlinearity and their interrelationship simultaneously.

The goal of the present research is to set up a theory of lasing of amplitude-squeezed light in a semiconductor laser when the pump current is several times the threshold value. Using a generalized model of a single-mode semiconductor laser that allows for spectral hole-burning and the nonlinear dependence of the spontaneous decay rate on the number of carriers in the active band as a starting point, we use it to describe the quantum properties of the outgoing radiation. Our main result is the conclusion that the combined contribution of both laser nonlinearities into the spectral density of the outgoing photon flux lowers the degree of squeezing at a fixed pump current exceeding the threshold value, but places no fundamental limit on the degree of squeezing as the pump current grows. This result sets the semiconductor laser apart from the gas laser (see Ref. 15), in which the excess noise resulting from inhomogeneous line broadening significantly limits the degree of squeezing and even may extinguish squeezing entirely.

2. SYSTEM OF EQUATIONS

The system of balance equations[18] describing photon–electron interactions inside the cavity has the form

$$\frac{\partial \hat{n}}{\partial t}=\hat{G}(\hat{n},\hat{N})\hat{n}-\gamma\hat{n}+\hat{R}_{\rm sp}+\hat{\Gamma}_{\gamma}(t)+\hat{\Gamma}_{g}(t), \tag{1}$$

$$\frac{\partial \hat{N}}{\partial t}=\frac{I}{q}-\hat{R}(\hat{N})-\hat{G}(\hat{n},\hat{N})\hat{n}+\hat{\Gamma}_{N}(t), \tag{2}$$

where $\hat{n}$ is the photon number operator in the cavity, $\hat{N}$ is the operator representing the number of carriers in the active area of the laser, γ is the rate of photon departure from the cavity, $\hat{R}(\hat{N})$ is the operator of spontaneous carrier relaxation, I is the injection current, q is the elementary charge, $\hat{R}_{\rm sp}$ is the operator of spontaneous emission into the laser mode, $\hat{G}(\hat{n},\hat{N})$ is the gain operator, defined as the difference between the stimulated emission and stimulated absorption rates, and $\hat{\Gamma}_{\gamma}(t)$, $\hat{\Gamma}_{g}(t)$, and $\hat{\Gamma}_{N}(t)$ are the Langevin noise sources, represented by a zero-mean Gaussian stochastic process. In the Markov approximation, the noise sources are δ-correlated (white noise):

$$\begin{aligned}&\langle\hat{\Gamma}_{i}(t)\hat{\Gamma}_{j}(t+\tau)\rangle=S_{ij}\delta(\tau), \quad i,\ j=g,\ N,\\ &\langle\hat{\Gamma}_{\gamma}(t)\hat{\Gamma}_{\gamma}(t+\tau)\rangle=S_{\gamma}\delta(\tau).\end{aligned} \tag{3}$$

Here $\hat{\Gamma}_{\gamma}(t)$ and $\hat{\Gamma}_{g}(t)$ are stochastic sources of photon noise, and $\hat{\Gamma}_{N}(t)$ is a stochastic source of carrier noise.

The noise-source spectral densities can be written as[18]

$$\begin{aligned}&S_{\gamma}=\gamma n, \quad S_{gg}=(R+A)n+R,\\ &S_{NN}=(R+A)n+\frac{N}{\tau_{e}}, \quad S_{gN}=-[(R+A)n+R],\end{aligned} \tag{4}$$

with τ_{e}^{-1} the spontaneous relaxation rate. Here and in what follows the ''hatless'' variables denote expectation values of the corresponding operators. The above equations show that the noise sources are correlated. To maintain generality, in (4) we have left the expectation value of the stimulated absorption operator $\hat{A}$, although in what follows we assume that the inversion is complete.

The fact that an electron leaves the conduction band due to recombination with a hole from the valence band leads to photon emission. This process is characterized by a radiative recombination rate $R_{\rm rad}(N)$. It is convenient to distinguish between two components in the radiative decay rate:

$$R_{\rm rad}=\beta R_{\rm rad}+(1-\beta)R_{\rm rad},$$

with the first describing radiative decay into the laser mode, and the second radiative decay into all nonlasing modes. The coefficient β can be interpreted as the efficiency of spontaneous emission into the laser mode.

On the other hand, some electrons leave the conduction band, contributing nothing to emission. The corresponding processes are characterized by a nonradiative relaxation rate $R_{\rm rel}(N)$. The total rate of spontaneous relaxation of the carriers is given by the sum of the two contributions: $R(N)=R_{\rm rad}(N)+R_{\rm rel}(N)$.

Since in a semiconductor laser two carriers, an electron and a hole, participate in radiative recombination, the radiative recombination rate $R_{\rm rad}(N)$ is proportional to the product of electron and hole concentrations, which provides a quadratic dependence of $R_{\rm rad}(N)$ on the number of carriers in the active area if the active area on the whole remains electrically neutral and the distribution of the carriers over the volume is approximately uniform. Carriers that do not participate in the emission process relax nonradiatively, providing a contribution to $R_{\rm rel}$ that is proportional to the total number of carriers. In addition, when the pump currents in semiconductor lasers are high, it is important to account for losses due to Auger relaxation, in which three carriers participate simultaneously. Summing up, we can write the final expression for $\hat{R}(\hat{N})$ as

$$\hat{R}(\hat{N})=\tau_{\rm th}^{-1}\left[a\hat{N}+\frac{b}{N_{\rm th}}\hat{N}^2+\frac{c}{N_{\rm th}^2}\hat{N}^3\right], \tag{5}$$

consistent with the results of Olshansky *et al.*[17] In the expression (5) for the spontaneous relaxation rate, the coefficient b reflects the fraction of radiative recombination processes, and the coefficients a and c the fraction of nonradiative recombination processes. The quantity $N_{\rm th}$ is defined as the number of carriers at the lasing threshold. We denote the spontaneous relaxation time at $N=N_{\rm th}$ by $\tau_{\rm th}$, and by averaging (5) we get the normalization condition for the coefficients, $a+b+c=1$.

To allow for the effect of spectral hole-burning in the gain line, we assume the following form for the dependence of the gain operator on the photon number:[16]

$$\hat{G}(\hat{n},\hat{N})=\frac{\hat{G}_0(\hat{N})}{\sqrt{1+\hat{n}/n_s}}. \tag{6}$$

The number of photons at which the gain decreases by a factor of $\sqrt{2}$ is denoted by n_s, and can be expressed in terms of intraband relaxation times:

$$n_s=\frac{\epsilon_0\hbar\overline{nn}_gV}{\mu^2\omega_0\widetilde{\Gamma}\tau_{\rm in}(\tau_c+\tau_v)},$$

where $\overline{n}$ is the mode index, n_g is the group index, V is the volume of the active area, μ is the dipole moment, ω_0 is the optical frequency, $\tau_{\rm in}$ is the polarization relaxation time, $\widetilde{\Gamma}$ is the optical clipping coefficient, and τ_c and τ_v are the interband relaxation times of the population difference between the conduction and valence bands, respectively.

According to the Einstein's relations, the rate of spontaneous emission per photon of the laser mode is equal to the rate $R_{\rm sp}$ of spontaneous emission into the laser mode. In turn, $R_{\rm sp}$ is determined by the rate $\beta R_{\rm rad}$ of radiative recombination of carriers into the laser mode. Consequently, the expression for the gain operator must be proportional to the square of the number of carriers and not, as it is often assumed, directly proportional to that number:

$$\hat{G}_0(\hat{N})=\hat{G}_N\hat{N}\,\frac{\hat{N}}{N_{\rm th}}. \tag{7}$$

We find that $\beta=G_N\tau_{\rm th}/b$, where G_N is the gain per carrier. In our discussion we ignored stimulated absorption, assuming that inversion is complete. This assumption is often made in the literature. For strong nonlinearity, when the number of carriers in the active area grows with the pump current, stimulated absorption effects weaken accordingly, with the result that the approximation is sure to be valid.

Two processes are involved in photon departure from the cavity: $\gamma=\gamma_{\rm out}+\gamma_{\rm in}$. The first is characterized by the rate of photon departure through the semitransparent mirror, $\gamma_{\rm out}$. The second incorporates internal photon losses and is characterized by the rate $\gamma_{\rm in}$.

3. STEADY-STATE SOLUTION

Let us set $\partial\hat{n}/\partial t$ and $\partial\hat{N}/\partial t$ equal to zero in Eqs. (1) and (2) and find the steady-state values for the average numbers of photons and carriers, ignoring spontaneous decay into the laser mode. The resulting system of two algebraic equations makes it possible to find the explicit dependence of the average number of carriers on the average number of photons in the cavity:

$$N=N_{\rm th}\left(1+\frac{n}{n_s}\right)^{1/4}. \tag{8}$$

Equation (8) expresses the fact that gains and losses are equal, or $G(n,N)=\gamma$, a condition that holds in a steady-state lasing mode. If we ignore the inhomogeneous nature of the broadening of a lasing line, the theory of a single-mode semiconductor laser predicts that the number of carriers is constant. For any pump current this number remains equal to its threshold value

$$N_{\rm th}=\frac{\gamma}{G_N}, \tag{9}$$

which follows from (8) when $n_s\to\infty$. Spectral hole-burning increases the number of carriers as the pump current grows. Although this dependence is extremely weak (see Eq. (8) and Fig. 1a), it determines the carrier spontaneous relaxation rate and the extent to which intralaser nonlinearities come into play; these nonlinearities may prove to be very large, due to the quadratic and cubic dependence on the number of carriers.

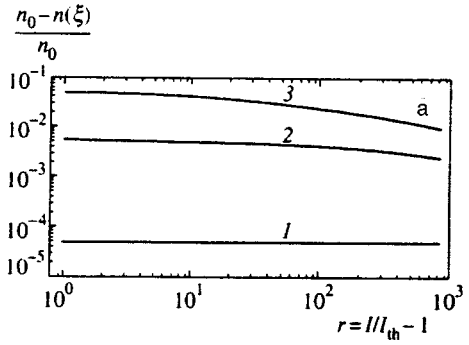
Introducing the parameter

$$r=\frac{I/q}{N_{\rm th}/\tau_{\rm th}}-1,$$

which reflects the extent to which the pump current exceeds the threshold level, we can write an equation for the average number of photons:

$$b\beta n=r+1-(1+\xi\beta n)^{1/4}[a+b(1+\xi\beta n)^{1/4}+c(1+\xi\beta n)^{1/2}]. \tag{10}$$

For purely radiative recombination we can obtain the solution of Eq. (10) explicitly:



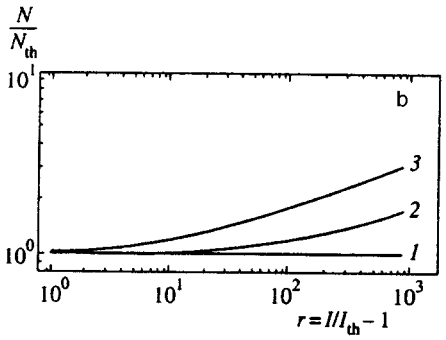


FIG. 1. (a) The relative difference in the number of photons for a laser with a weakly saturated gain line, $n(\xi=10^{-4})=n_0$, and for a laser with strong saturation $n(\xi)$. (b) Carrier number normalized to the threshold value as a function of the pump parameter: Curve *1* corresponds to $\xi=0$, curve *2* to $\xi=0.1$, and curve *3* to $\xi=0.01$.

$$n=\beta^{-1}\left[r+1+\frac{\xi}{2}-\sqrt{1+(r+1)\xi+\left(\frac{\xi}{2}\right)^2}\right].$$

In Eq. (10) we introduced a new parameter, $\xi=1/\beta n_s$, which characterizes the efficacy of spectral hole-burning. If this parameter is nonzero, the linear dependence of the number of photons in the cavity on the pump current is distorted. Note that r and β have been introduced in the same way as in the problem without spectral hole-burning.

Spectral hole-burning does not have any appreciable effect on the dependence of the phonon number on the pump current. At $\xi=0.1$, which is a typical value for semiconductor lasers, the difference from the ideal case $\xi=0$ amounts to several percentage points (see Fig. 1b).

4. PHOTON NUMBER NOISE IN CAVITY

A steady-state lasing mode is characterized by average values of the photon number n and the carrier number N, and by deviations from these values, δn and δN, which can be assumed small far from threshold ($r>0.01$). Linearizing Eqs. (1) and (2) in $\delta\hat{n}$ and $\delta\hat{N}$ and transforming to the frequency representation, we find that

$$\left(i\Omega-\frac{\partial G}{\partial n}n\right)\delta\hat{n}=\frac{\partial G}{\partial N}n\,\delta\hat{N}+\hat{\Gamma}_g(\Omega)+\hat{\Gamma}_\gamma(\Omega), \tag{11}$$

$$\left(i\Omega+\frac{\partial R}{\partial N}+\frac{\partial G}{\partial N}n\right)\delta\hat{N}=-\left(G+\frac{\partial G}{\partial n}n\right)\delta\hat{n}+\hat{\Gamma}_N(\Omega). \tag{12}$$

Knowing the explicit form of (6), (11), and (8), we can solve the above system of equations for $\delta\hat{n}$ and $\delta\hat{N}$:

$$\delta\hat{N}=\tau_e\frac{(\varepsilon-1)[\hat{\Gamma}_g+\hat{\Gamma}_\gamma]+(i\omega+\varepsilon)\hat{\Gamma}_N}{\{2[r_{\text{eff}}+\varepsilon+f\varepsilon]-\gamma\tau_e\omega^2\}+i\omega[2(r_{\text{eff}}+1+f\ )+\gamma\tau_e\varepsilon]}, \tag{13}$$

$$\delta\hat{n}=\frac{1}{\gamma}\frac{[2(r_{\text{eff}}+1+f\ )+i\gamma\tau_e\omega][\hat{\Gamma}_g+\hat{\Gamma}_\gamma]+2r_{\text{eff}}\hat{\Gamma}_N}{\{2[r_{\text{eff}}+\varepsilon+f\varepsilon]-\gamma\tau_e\omega^2\}+i\omega[2(r_{\text{eff}}+1+f\ )+\gamma\tau_e\varepsilon]}, \tag{14}$$

where $\omega=\Omega/\gamma$ is the normalized frequency, $\tau_e^{-1}=R(N)/N$ is the spontaneous relaxation rate per carrier, and

$$f=\frac{1}{2}\frac{-a(1-2\varepsilon)^{1/2}+c}{a(1-2\varepsilon)^{1/2}+b(1-2\varepsilon)^{1/4}+c}$$

is a factor determined by carrier losses in the active area; this factor vanishes in the limit of purely radiative relaxation.

To maintain generality in the problem that ignores spectral hole-burning, in (13) and (14) we introduced the parameter

$$\varepsilon=\varepsilon(n)=\frac{1}{2}\frac{\xi\beta n}{1+\xi\beta n}=\frac{1}{2}\frac{n}{n+n_s}, \tag{15}$$

which varies between zero (no saturation) and 1/2 (complete gain saturation). An important feature of this model is that the average carrier number changes only as a result of gain saturation, so that all the parameters of the problem can be expressed in terms of the average photon number and the saturation parameter. The validity of this statement is clearly demonstrated by ε:

$$\tau_e^{-1}=\tau_{\text{th}}^{-1}[a+b(1-2\varepsilon)^{-1/4}+c(1-2\varepsilon)^{-1/2}], \tag{16}$$

$$r_{\text{eff}}=\frac{(r+1)(1-2\varepsilon)^{3/4}}{a(1-2\varepsilon)^{1/2}+b(1-2\varepsilon)^{1/4}+c}-1. \tag{17}$$

The r_{eff} introduced in (17) is a generalization of the pump parameter r and reflects the extent to which the laser mode uses the pump energy. The coefficient of spontaneous emission into the laser made is modified in a similar way:

$$\beta_{\rm eff}=\frac{\beta(1-2\varepsilon)^{3/4}}{a(1-2\varepsilon)^{1/2}+b(1-2\varepsilon)^{1/4}+c}. \tag{18}$$

In terms of the quantities specified by (17) and (18), the average number of photons in a mode can be written as $n=r_{\rm eff}/b\beta_{\rm eff}$. Note that in the absence of gain saturation ($\varepsilon\to 0$), the modified parameters become constants: $\tau_e\to\tau_{\rm th}$, $r_{\rm eff}\to r$, and $\beta_{\rm eff}\to\beta$.

The spectrum of photon-number fluctuations inside the cavity is given by

$$S_{\rm in}(\Omega)=\langle \delta\hat{n}(\Omega)\,\delta\hat{n}^*(\Omega)\rangle. \tag{19}$$

Plugging in (14) yields

$$S_{\rm in}=\gamma^{-2}\left\{r_{\rm eff}^2 S_{NN}+\left[(r_{\rm eff}+f+1)^2+\left(\frac{\gamma\tau_e\omega}{2}\right)^2\right]\right.$$
$$\left.\times(S_{gg}+S_\gamma)+2r_{\rm eff}(r_{\rm eff}+f+1)S_{gN}\right\}\left\{\left[(r_{\rm eff}+\varepsilon f+\varepsilon)-\frac{\gamma\tau_e\omega^2}{2}\right]^2+\omega^2\left[(r_{\rm eff}+f+1)+\frac{\gamma\tau_e\varepsilon}{2}\right]^2\right\}^{-1}. \tag{20}$$

In the adopted notation system the noise-source spectral densities (4) have the form

$$S_\gamma=\gamma n,\quad S_{gg}=\gamma n,$$
$$S_{gN}=-\gamma n,\quad S_{NN}=\gamma n+\frac{\gamma n}{r_{\rm eff}}. \tag{21}$$

The final expression for the noise spectrum is

$$S_{\rm in}(\omega)=\frac{n}{\gamma}\,\frac{(r_{\rm eff}^2+3r_{\rm eff}+2)+2(\gamma\tau_e\omega/2)^2+2f(r_{\rm eff}+f+2)}{[(r_{\rm eff}+\varepsilon f+\varepsilon)-(\gamma\tau_e/2)\omega^2]^2+\omega^2[(r_{\rm eff}+f+1)+(\gamma\tau_e/2)\varepsilon]^2}. \tag{22}$$

5. NOISE IN THE OUTGOING PHOTON FLUX

The spectrum of fluctuations of the photon flux leaving the laser cavity through the semitransparent mirror is given by an expression similar to (19),

$$S_{\rm out}(\Omega)=\langle \delta\hat{n}_e(\Omega)\,\delta\hat{n}_e^*(\Omega)\rangle, \tag{23}$$

with the relationship between the number of photons inside the cavity and the outgoing photon flux given by[18]

$$\hat{n}_e=\gamma_{\rm out}\hat{n}-\hat{\Gamma}_{\rm out}. \tag{24}$$

Here $\hat{\Gamma}_{\rm out}$ is a fluctuation operator reflecting the fact that the photons leave the cavity at random (reflected vacuum fluctuation). It is simply related to the fluctuation operator $\hat{\Gamma}_\gamma$: $\hat{\Gamma}_\gamma=\hat{\Gamma}_{\rm in}+\hat{\Gamma}_{\rm out}$. The corresponding correlation functions are

$$\langle\hat{\Gamma}_{\rm out}(t)\hat{\Gamma}_{\rm out}(t+\tau)\rangle=\frac{\gamma_{\rm out}}{\gamma}S_\gamma=\gamma_{\rm out}n,$$
$$\langle\hat{\Gamma}_{\rm in}(t)\hat{\Gamma}_{\rm in}(t+\tau)\rangle=\frac{\gamma_{\rm in}}{\gamma}S_\gamma=\gamma_{\rm in}n. \tag{25}$$

Linearizing (24), we get

$$\hat{\delta n}_e=\gamma_{\rm out}\delta\hat{n}-\hat{\Gamma}_{\rm out}. \tag{26}$$

Plugging (26) into (23), we arrive at an expression for the photon-flux fluctuation spectrum:

$$S_{\rm out}(\Omega)=\gamma_{\rm out}^2 S_{\rm in}(\Omega)+\frac{\gamma_{\rm out}}{\gamma}S_\gamma$$
$$-2\gamma_{\rm out}\,{\rm Re}[\langle\delta\hat{n}(\Omega)\hat{\Gamma}_{\rm out}(\Omega)\rangle]. \tag{27}$$

In the limit of zero internal losses, $\gamma_{\rm out}=\gamma$, we have

$$S_{\rm out}(\omega)=\gamma n\left\{1-\left[r_{\rm eff}^2-r_{\rm eff}-2+2\varepsilon(r_{\rm eff}+1)-2f\right.\right.$$
$$\left.\times(2+f-2\varepsilon-\varepsilon f-\varepsilon r_{\rm eff})-2(1-\varepsilon)\left(\frac{\gamma\tau_e\omega}{2}\right)^2\right]$$
$$\times\left[\left[(r_{\rm eff}+\varepsilon f+\varepsilon)-\frac{\gamma\tau_e\omega^2}{2}\right]^2+\omega^2\right.$$
$$\left.\left.\times\left[(r_{\rm eff}+f+1)+\frac{\gamma\tau_e\varepsilon}{2}\right]^2\right]^{-1}\right\}. \tag{28}$$

The spectrum of low-frequency fluctuations is especially simple:

$$S_{\rm out}(0)=\gamma n\,\frac{r_{\rm eff}+2-2\varepsilon+\varepsilon^2+f[(1-\varepsilon)^2(f+4)+f-2\varepsilon^2]}{(r_{\rm eff}+\varepsilon f+\varepsilon)^2}. \tag{29}$$

Figure 2 depicts the dependence of $S_{\rm out}(0)/\gamma n$ on photon number n. The curves show that if the threshold is exceeded severalfold, the amplitude-noise squeezing decreases considerably due to gain saturation. The same tendency manifests itself if one allows for nonlinear nonradiative relaxation.

One of the most attractive features of squeezed light generated by a semiconductor laser is that the squeezing bandwidth amounts to several hundred gigahertz. A distinctive feature of this process is that spectral hole-burning in the gain line worsens the characteristics of nonclassical light, appreciably narrowing the squeezing bandwidth (see Fig. 3).

Note that although allowing for the various nonlinearities in a semiconductor laser poses additional difficulties in moving toward higher degrees of squeezing, these processes place no important limit on reduction of amplitude noise: by

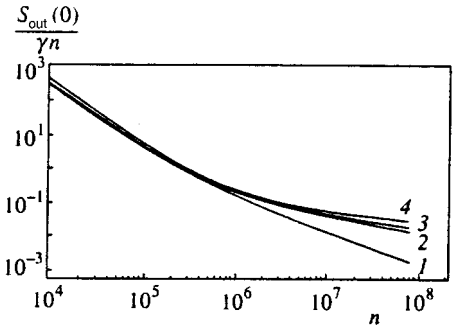


FIG. 2. The low-frequency spectrum of outgoing radiation fluctuations normalized to the shot noise level, as a function of the number of photons in the cavity in the absence of saturation, $\xi=0$ (curve *1*), and with finite saturation, $\xi=0.1$ (curves *2–4*). Curve *1* corresponds to $a=c=0$ and $b=1$, curve *3* to $a=c=0$ and $b=1$, curve *3* to $a=c=0.25$ and $b=0.5$, and curve *4* to $a=0$ and $b=c=0.5$.

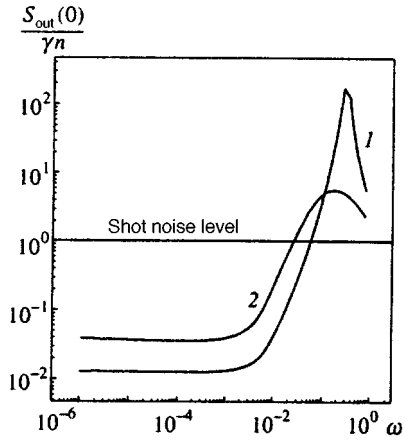


FIG. 3. The spectrum of outgoing radiation fluctuations normalized to the shot noise level, as a function of the normalized frequency ω ($\beta=5\times10^{-6}$, $\gamma\tau_e=10^3$, and $\overline{n}=10^7$): curve *1*, in the absence of saturation, $\xi=0$, $a=c=0$, and $b=1$; and curve *2*, with finite saturation, $\xi=0.1$, $a=c=0$, and $b=1$.

increasing the pump intensity, the squeezing can be made as high as desired. It must be noted, however, that higher-order nonlinearities, such as nonradiative relaxation processes involving more than five carriers, are sure to limit squeezing and do not allow the noise to fall below a certain limit.

From the practical viewpoint, the main factor that limits squeezing is still the field losses inside the active area. Let us assess the effect of these losses by writing the low-frequency fluctuation spectrum for an arbitrary γ_{out}-to-γ ratio:

$$S_{out}(0)=\gamma_{out}n\left\{1+\frac{\gamma_{out}}{\gamma}\frac{(r_{eff}+1)(2-r_{eff}-2\varepsilon)+2f(2+f-2\varepsilon-\varepsilon f-\varepsilon r_{eff})}{(r_{eff}+\varepsilon f+\varepsilon)^2}\right\}. \tag{30}$$

If γ and γ_{out} differ considerably (more than by 10–20%), the effect of spectral hole-burning on the low-frequency noise of the outgoing photon flux becomes negligible. The technology of manufacturing semiconductor lasers is developing in the direction of reducing internal field losses, with the result that the problem of how laser nonlinearities of disparate nature affect noise squeezing becomes more and more important.

6. CONCLUSION

In this paper we have proposed a modified system of Langevin–Heisenberg equations for describing the photon flux noise (fluctuations) in the single-mode lasing of a semiconductor laser when the pump current is much higher than the threshold value. We have analyzed theoretically the noise produced by a semiconductor laser with allowance for spectral hole-burning in the gain line and for a nonlinear dependence of the carrier spontaneous recombination rate on the number of carriers in the active area.

We have shown that spectral hole-burning leads to an increase in the number of carriers in the active area, which means that the dependence of carrier lifetime on the number of the carriers must be taken into account.

The basic conclusion that can be drawn from our results is that there is a tendency toward noise squeezing when the pump current is much higher than the threshold level. However, spectral hole-burning in the gain line retards this process, and squeezing occurs at higher pump-current values.

Our study corresponds to an experimental situation in which all side modes in the laser are completely suppressed. In this case, as we have demonstrated above, optical nonlinearities of the semiconductor laser are the important factors in limiting the squeezing. This limiting effect becomes especially evident when intracavity photon losses are low.

The authors would like to express their gratitude to Yu. M. Golubev and I. V. Sokolov for useful discussions and to P. A. Usachev for his constant help. The present work was partially financed through the Russian Ministry of Science by the Optics and Laser Physics Program and through a joint grant from INTAS and the Russian Fund for Fundamental Research (Grant No. 95-IN-RU-656). The work of one of the authors (V.V.K.) was made possible thanks to an INTAS stipend (Grant 93-2492-ext) and was done through a research program of the International Center for Fundamental Physics in Moscow. The same author would like to thank the Government of St. Petersburg and the Heidelberg University in Germany for financial support (a Rotary Club Stipend).

[1)]Processes involving four or more carriers have low probability and are not taken into account here.

[1] Yu. M. Golubev and I. V. Sokolov, Zh. Éksp. Teor. Fiz. **87**, 408 (1984) [Sov. Phys. JETP **60**, 234 (1984)].

[2] Y. Yamamoto and S. Machida, Phys. Rev. A **35**, 5114 (1987).
[3] S. Machida, Y. Yamamoto, and Y. Itaya, Phys. Rev. Lett. **58**, 1000 (1987).
[4] A. S. Trifonov, P. A. Usachev, M. A. Ivanov *et al.*, in *ICONO'95: Atomic and Quantum Optics: High-Precision Measurements*, S. N. Bagaev and A. S. Chirkin (eds.), Proc. SPIE (1995), Vol. 2799, p. 405.
[5] C. B. Su, J. Schlafer, and R. B. Lauer, Appl. Phys. Lett. **57**, 849 (1990).
[6] M. J. Freeman, H. Wang, D. G. Steel, R. Craig, and D. R. Scifres, Opt. Lett. **18**, 379 (1993).
[7] F. Marin, A. Bramati, E. Giacobino, T.-C. Zhang, J.-Ph. Poizat, J.-F. Roch, and P. Grangier, Phys. Rev. Lett. **75**, 4606 (1995).
[8] A. S. Trifonov, P. A. Usachev, A. V. Lutetskiy *et al.*, in *EQEC-96 Technical Digest*, Hamburg (1996), QMC6, p. 12.
[9] M. J. Freeman, H. Wang, D. G. Steel, R. Craig, and D. R. Scifres, Opt. Lett. **18**, 2141 (1993).
[10] T.-C. Zhang, J.-Ph. Poizat, P. Grelu *et al.*, Quantum Semiclassic. Opt. **7**, 601 (1995).
[11] H. Wang, M. J. Freeman, and D. G. Steel, Phys. Rev. Lett. **71**, 3951 (1993).
[12] J. Kitching, A. Yariv, and Y. Shevy, Phys. Rev. Lett. **74**, 3372 (1995).
[13] J. Kitching, A. Yariv, and Y. Shevy, in *Laser Frequency Stabilization and Noise Reduction*, Y. Shevy (ed.), Proc. SPIE (1995), Vol. 2378, p. 77.
[14] S. Inoue, H. Ohzu, S. Machida, and Y. Yamamoto, Phys. Rev. A **46**, 2757 (1992).
[15] Yu. M. Golubev and V. P. Gryaznevich, Opt. Spektrosk. **59**, 59 (1985) [Opt. Spectrosc. **59**, 89 (1985)].
[16] G. P. Agrawal, J. Appl. Phys. **63**, 1232 (1988).
[17] R. Olshansky, C. B. Su, J. Manning, and W. Powazinik, IEEE J. Quantum Electron. **QE-20**, 838 (1984).
[18] Y. Yamamoto, W. H. Richardson, and S. Machida, in *Quantum Measurements in Optics*, P. Tombesi and D. F. Walls (eds.) Plenum Press, New York (1992), Vol. 282, p. 65.

Translated by Eugene Yankovsky

Four-wave polariton scattering of light in $LiNbO_3$

G. Kh. Kitaeva, P. S. Losevskiĭ, A. A. Mikhaĭlovskiĭ, and A. N. Penin

M. V. Lomonosov Moscow State University, 119899 Moscow, Russia
(Submitted 21 October 1996)
Zh. Éksp. Teor. Fiz. **112**, 441–452 (August 1997)

Four-wave Stokes k-spectra for light scattering on polaritons in lithium niobate crystals with an Mg impurity are studied experimentally. The mechanisms for direct, cascade, coherent, and incoherent four-wave mixing of light are discussed in the course of interpreting the angular dependences of the scattered light intensity. It is shown that the dispersion of the real part of the polariton wave vector and the refractive index of the crystals at the polariton frequencies can be measured with an order of magnitude greater accuracy than by spontaneous three-wave polariton light scattering. A significant discrepancy is found between determinations of the polariton absorption coefficient from the angular spectra of three-wave scattering and four-wave scattering in terms of the model employed here.
[S1063-7761(97)00408-3]

1. INTRODUCTION

The spectroscopy of light scattered on polaritons is an extremely sensitive method for studying relatively small changes in the properties of a medium. This method can be used to study phenomena which do not show up in phonon Rayleigh scattering spectra. This occurs in those rather common cases where the frequencies of the phonons remain essentially unchanged, while only the oscillator strengths or damping constants of the phonons vary. Then the dispersion relations of the polariton states also change significantly. Effects such as small changes in the concentration of an impurity, spatial inhomogeneity, phase transitions, low-dimensionality effects, and many others, show up in the spectra of light scattered on polaritons. The dielectric characteristics of media in the neighborhood of the optical phonon frequencies can also be measured using IR-reflection spectra. When scattering effects are used, however, signal intensities can be measured in the visible, which, as a rule, is much simpler. In addition, when light-scattering spectroscopic techniques are used there are more ways of distinguishing volume and surface effects.

Three-wave spontaneous parametric light scattering on polaritons has been used successfully in studies of the properties of non-centrally symmetric crystalline structures.[1] Its applicability is limited, however, by the low scattered light intensity. This makes it especially difficult to study microscopically small objects and limits the spatial resolution of the method. Active spectroscopy based on optical four-wave scattering effects offers a way out.

The first work in this area was begun in the end of the 1960s[2] and is being carried out by various research groups to this day. A basic study of four-wave mixing processes was made in media with[3] and without[2,4–8] a center of symmetry. The case of a non-centrally-symmetric medium is the most general case for studying active spectroscopy. Here direct four-photon processes and cascade three-photon processes, associated with the nonlinear susceptibilities $\chi^{(3)}$ and $\chi^{(2)}$, respectively, contribute to the scattered light intensity. Satisfaction of the conditions of frequency and phase synchrony is important for both types of process. For the direct four-photon process they have the form

$$\omega_L-\omega_S-\omega_1+\omega_2=0, \tag{1}$$

and

$$\Delta\mathbf{k}\equiv\mathbf{k}_S-\mathbf{k}_L+\mathbf{k}_1-\mathbf{k}_2=0, \tag{2}$$

where ω are the frequencies and $\mathbf{k}$ are the wave vectors of the interacting waves in the medium, the subscripts 1, 2, and L refer to the waves incident on the medium, and the subscript S refers to the signal wave detected as it leaves the medium. Here and in the following we consider the Stokes component of the scattered light. The cascade process can be treated as a sequence of processes in which difference frequencies are generated. In the first process an idler wave is generated (associated in our case with the polariton state of the medium) owing to parametric subtraction of the frequencies of two exciting pump waves, and then the probe pump wave is scattered on this idler. The cascade process is most efficient when the synchrony conditions (1) and (2) are supplemented by the following synchrony conditions:

$$\omega_1-\omega_2=\omega_p, \tag{3}$$

and

$$\tau\equiv\mathbf{k}_1-\mathbf{k}_2-\mathbf{k}_p=0. \tag{4}$$

Here the subscripts 1 and 2 refer to the exciting radiation and p, to the polariton (idler) wave.

The interaction of these processes has been studied experimentally in GaAs, $LiNbO_3$, and $LiIO_3$ crystals, expressions for the third order susceptibility $\chi^{(3)}$ have been obtained, and its effective magnitude corresponding to both the direct and the cascade processes measured.[4–6] Interference of processes involving the different nonlinear susceptibilities shows up in an asymmetry of the scattering line profile and the appearance of additional maxima in that profile in ω- and k-space.[7] The effects have been observed experimentally during scattering of light on excitations corresponding to the upper polariton branch in the case of weak absorption at the polariton frequency.[6]

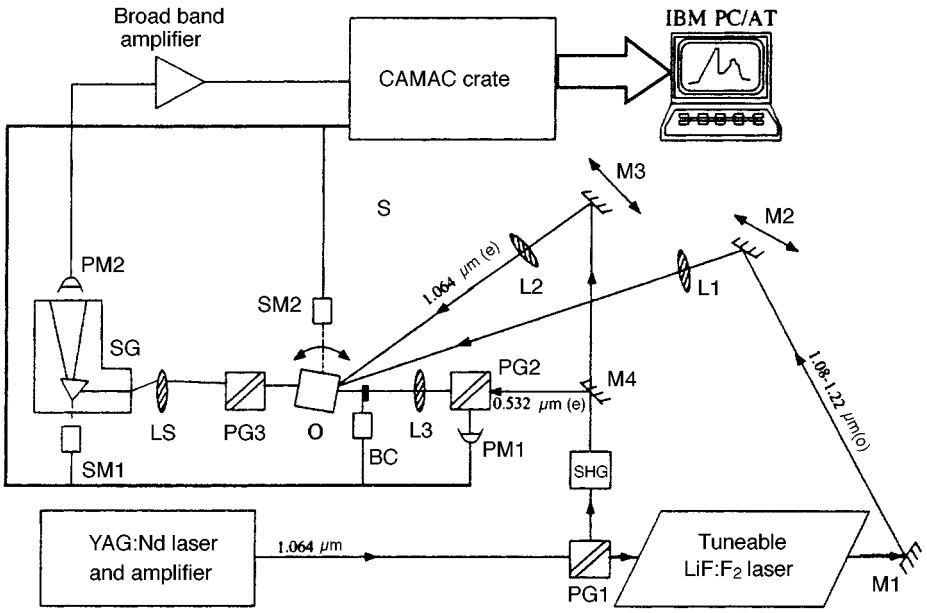


FIG. 1. A sketch of the experimental apparatus: PG1, 2, 3 Glan–Thomson polarizing prisms; M1, 2, 3 mirrors; M4 frequency selective mirror; SHG second-harmonic generating crystal; L1, 2, 3 focussing lenses; BC beam chopper; S test sample; LS three-lens system; SG spectrograph; PM1 photomultiplier detector for pump power monitor; PM2 photomultiplier detector for scattered light measurements; SM1, 2 stepping motor drives.

The purpose of this paper is to create an optimized spectroscopic system, estimate its sensitivity, and study the features of coherent Stokes four-wave scattering on polaritons when there is strong absorption at the frequency of the polariton state. This case is most typical for polariton spectroscopy and has also been little studied experimentally. We have used $LiNbO_3$ crystals with different amounts of Mg impurity as test objects.

2. EXPERIMENTAL APPARATUS AND $LiNbO_3$ SAMPLES

In most previous work, a traditional coherent antistokes scattering arrangement has been used where one of the pump waves is doubly degenerate from the standpoint of the four-wave mixing process and a signal is detected at the antistokes frequency. In our case, we have used the more general setup for a four-wave interaction in which all the waves have different frequencies and the Stokes component of the scattered light is detected. A sketch of the experimental apparatus is shown in Fig. 1. The excitation wave sources at frequencies ω_1 and ω_2 are a YAG:Nd^{3+} laser and a tuneable LiF:F_2^- laser with respective output wavelengths of $\lambda_1 = 1.064$ μm and λ_2 in the range 1.08–1.22 μm operating at a repetition rate of 1–33 Hz. The pump for the tuneable color center crystal laser is the fundamental of the YAG:Nd^{3+} laser passed through a YaG:Nd^{3+} amplifier and a Glan-Thomson polarizing prism (PG1). The probe wave is the second harmonic of the YAG:Nd^{3+} laser (wavelength $\lambda_L = 532$ nm) generated by a frequency doubler (SHG). This radiation is separated from the fundamental by a mirror with a frequency-selective reflectivity. Because near-IR sources are used to excite the polariton wave, parasitic light owing to luminescence of the test medium irradiated by the exciting radiation lies in the IR, far from the visible signal detection wavelength. The required polarization of the rays incident on the crystal is determined by the Glan–Thomson polarizing prisms PG1 and PG2. The angles of incidence of the pump waves on the sample crystal are determined by the system of mirrors M2-M4. In addition, inserting additional focussing lenses L1-L3 in the pump beams makes it possible to vary the irradiance of the pump waves in the interaction region, as well as their angular divergence. Scattered light is collected by a three-lens system LS in the plane of the inlet slit of a spectrograph SG after first passing through a Glan–Thomson polarizer PG3 which serves as an analyzer for the scattered light and cuts off the probe light that passes through the sample S.

A two-dimensional frequency–angle scattering pattern was formed at the spectrograph output. The horizontal deviation of the beam corresponded to the frequency of the scattered wave and the vertical deviation to the scattering angle in the plane of the pump wave vectors. The structure of the plate holder of the spectrograph made it possible detect the signal either photographically or electronically. In the latter case the detector is a photomultiplier PM2 operating in current (analog) mode. The signal from the tube is fed to a wide band amplifier with a controllable gain, then to a fast-gated integrating ADC in a CAMAC crate, and on to an IBM PC/AT control computer. The computer synchronizes and controls the operation of the individual system components through circuits installed in the CAMAC crate. In our apparatus, during photoelectric detection of the spectrum the photomultiplier was fixed and the slit, whose width could be varied with a micrometer screw, was positioned in front of it. The spectrum was scanned in frequency by rotating the prism in the spectrograph with a stepping motor drive SM1. Another drive, SM2, was used to rotate the crystal in the plane containing all the pump beams and this makes it possible to vary the detuning of the phase synchrony in the sample. An additional photomultiplier PM2 is used to monitor the pump power. A beam chopper BC in the path of the probe light makes it possible to subtract automatically the background owing mainly to illumination of the photodetector by radiation at the sum frequency of the two infrared lasers. The optical system is set up to detect the Stokes com-

ponent of the scattered light. This is the main difference between our spectroscopic apparatus and that used for coherent antistokes four-wave scattering.[4–6] Antistokes scattering systems have come into wider use for four-wave spectroscopy because it is possible to use frequency-degenerate pumping. (In this case we have $\omega_1=\omega_L$ and the frequency of the detected signal is $\omega_{as}=2\omega_L-\omega_2$.) This approach can be used without any restrictions for studying centrally symmetric media, when cascade scattering does not occur and all the required synchrony conditions can be met relatively easily. On the other hand, a larger number of synchrony conditions must be met in order to study polaritons by means of cascade four-wave scattering in non-centrally-symmetric crystals. To do this in many cases, degenerate pumping must be avoided and the Stokes and antistokes detection schemes must be considered on an equal basis. Then the easy compatibility of the Stokes scheme with three-wave light scattering detection becomes a major advantage. In fact, simply by turning on the exciting pumps it is possible to go easily from observing spontaneous three-photon scattering of light on polaritons to observing scattering on coherently excited states of the medium, since in both cases the scattered radiation lies within the same frequency-angle interval.

In our experiments we have used samples of $LiNbO_3$ with mass fractions of Mg impurity equal to 0.68% and 0.79%. Data on the visible and near IR refractive indices for the crystal with an impurity concentration of 0.79% were taken from Aleksandrovskiĭ *et al.*[9] and data for the crystal with a concentration of 0.68% were obtained by approximating data from the same paper and measurements of the refractive index of a congruent $LiNbO_3$ crystal with no impurity.[10] In addition, the dispersion in the refractive index of the crystals for ordinary waves in the neighborhood of the frequencies of the polariton state being studied here was determined from the spontaneous parametric light scattering spectra.

$LiNbO_3$ crystals belong to the trigonal class and have a $3m$ point symmetry. In the experiment the pumps at wavelengths $\lambda_L=0.532$ nm and $\lambda_1=1.064$ nm had extraordinary polarization and the radiation from the tuneable LiF:F_2^--laser had ordinary polarization. This determined the components of the nonlinear susceptibility tensor at which the interaction takes place: χ_{xxz} and χ_{xxy} for the quadratic nonlinear susceptibility tensor $\chi^{(2)}$, and χ_{xxzz} for the cubic nonlinear susceptibility tensor $\chi^{(3)}$. The scattering geometry is shown in Fig. 2.

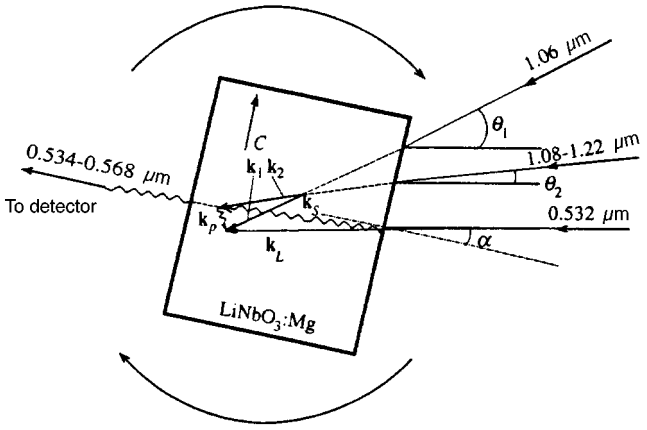


FIG. 2. Schematic illustration of the beam paths and of the mutual orientation of the wave vectors.

3. ADJUSTMENT OF THE FREQUENCY AND ANGULAR SYNCHRONY

In order to determine the dispersion relation for the polariton states $\omega_p(\mathbf{k})_p$ from four-wave scattering spectra it is necessary to ensure exact satisfaction of the frequency relations (1) and (3) and satisfaction of the equations (2) and (4) for the wave vectors of interacting waves in the medium with sufficiently small detunings $\Delta\mathbf{k}$ and τ. In practice, adjustment of the frequency and spatial (angular) synchrony involves two stages.

In the first stage, for each chosen polariton frequency value ω_p one establishes the output frequency ω_2 of the tuneable laser in accordance with Eq. (3) and the frequency ω_S for selection of the signal radiation at the output from the crystal, in accordance with Eq. (1). It should be noted that double frequency selection (at the level of the selective resonator of the tuneable laser and at the level of the spectrally selective apparatus in the detector channel) is not obligatory, but is dictated solely by the need to eliminate noisy background from the detector. In our case the measurements were done at different fixed values of the frequency difference $\omega_1-\omega_2=\omega_p$ in the range 520–578 cm^{-1}, both near the E-type phonon resonance of the medium at frequency $\omega_{TO}=580$ cm^{-1} and at some distance from it. The spectral widths of the pump lines were roughly 1 cm^{-1} for the fundamental and harmonic lines of the YAG:Nd^{3+} laser and 6 cm^{-1} for the tuneable laser. The widths of the resultant signal emission were fully consistent with the frequency structure of the pump waves.

In the second stage the spatial synchrony is adjusted. This process is evidently the most tedious phase of the measurements. For scattering in the principal plane of a uniaxial crystal there are three adjustable parameters. These are the angles determining the mutual orientation of the three pump waves and the optical axis of the crystal. Final exact adjustment must be done over any two of these three parameters. Satisfaction of condition (4) is not critical for observing direct four-photon scattering and the spatial synchronism can be adjusted through a single parameter, which makes observation of the signal much easier. However, the possibility of measuring the wave vector of the polariton is then lost.

We have measured the four-wave scattering line profiles in **k**-space. The sequence of actions was then the following: a constant frequency difference $\omega_1-\omega_2=\omega_p$ was established, the IR pump rays were directed onto the crystal at fixed angles θ_1 and θ_2 relative to the propagation direction of the probe pump, and the dependence of the signal intensity at the frequency $\omega_S=\omega_L-\omega_1+\omega_2$ was measured as a function of the rotation angle α of the crystal in the plane of the pump wave vectors. The angles α, θ_1, and θ_2 were varied about average values determined previously by solving Eqs. (1)–(4) using data on the dispersion in the refractive index of the crystals. For each fixed signal (and, therefore,

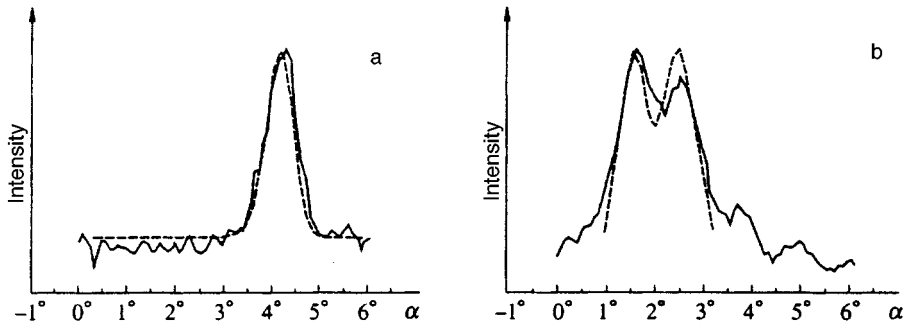


FIG. 3. Experimentally measured dependences (smooth curves) of the four-wave scattering intensity on the rotation angle of the crystal: $\theta_2=19.2°$, $\theta_1=44.0°$ (a) and $\theta_1=43.3°$ (b). The dashed curves are theoretical calculations for $\alpha_p=23\ \mathrm{cm}^{-1}$, $d=1.25$ mm, $L=1.0$ mm (a) and $L=0.8$ mm (b).

polariton) frequency the range of the solutions of the exact synchronism conditions in the angle space α, θ_1, and θ_2 forms part of a curve. Taking the possible detuning of the synchronism into account is equivalent to determining the ''spread'' in this curve. At each frequency difference $\omega_1-\omega_2=\omega_p$ a series of measurements of the line profile $I_S(\alpha)$ was made in which the mutual orientation of the probe wave and one of the IR pump waves was kept constant at the input to the crystal, while the angle of incidence of the other IR pump was varied from one measurement of the line profile to another. The four-photon scattering signal for exact adjustment of the angular synchrony was much (almost four orders of magnitude) higher than the spontaneous three-wave scattering signal. Here the spontaneous scattering signal was collected from the entire length of the sample (~ 1 cm), while the four-photon scattering signal was collected only from the region ($\sim 0.5-1$ mm long) where the pump beams overlap.

A typical individual scattering profile is shown in Fig. 3a. The scattering profile has one distinct peak with an angular width on the order of 1°. Furthermore, for certain relations between the experimentally established angles θ_1 and θ_2, supplementary peaks in the scattered light were observed with roughly the same width and comparable amplitudes. (For example, see the spectrum in Fig. 3b.)

It is more convenient to analyze the general behavior of the angular distribution of the signal intensity in terms of a series of experimentally measured line profiles. Each series of this kind of measurements of the line profile $I_S(\alpha)$ taken at a fixed angle θ_1 (or θ_2) and variable θ_2 (or, in turn, θ_1) represents a distribution $I_S(\alpha,\theta_i)$. Figure 4 shows the results of one such series of measurements where the angle of incidence θ_2 and frequency ω_2 of the tuneable laser were kept constant in the α, θ_1 plane, where α is the angle of rotation of the crystal and θ_1 is the angle of incidence of the IR wave. The points denote the positions of the peak in experimentally observed ''single humped'' and ''two humped'' $I_S(\alpha)$ curves. The range of angles α and θ_1 where we observed four-wave light scattering is shaded in the figure. When the directions of incidence of the IR wave was reoriented outside this range, the intensity of the signal fell off rapidly to a level corresponding to spontaneous three-wave parametric scattering of the probe light. The shaded region is essentially the intersection of a plane $\theta_2=\mathrm{const}$ with the region of variation of the three parameters α, θ_1, and θ_2 within which the intensity I_S is nonzero.

To answer the question of how to determine the real and imaginary parts of the polariton wave vector (and then the refractive index, absorption, and complex dielectric constant of the crystal) from the measured curves we now turn to an analysis of scattering theory.

4. ANALYSIS OF THE ANGULAR VARIATION IN THE SCATTERING INTENSITY

Among the published studies devoted to the features of four-photon coherent scattering in non-centrally-symmetric media,[4–8] the shape of the scattering profile has been analyzed in greatest detail in Ref. 7. The conditions of our experiment correspond to the case of weak pumping and strong absorption at the polariton frequency examined there.[7] The frequency–angular variation of the scattering profile should obey the following expression for the scattered light intensity I_S:[7]

$$I_S \propto \mathrm{Sinc}^2\left(\frac{\Delta kL}{2}\right)\left[\beta^2+\frac{4\tau\beta}{\alpha_p^2+4\tau^2}+\frac{1}{\alpha_p^2+4\tau^2}\right], \tag{5}$$

where τ is the wave detuning in the first stage of the cascade process, the ''heating'' of the polariton state, Δk is the wave detuning for the direct four-photon interaction process, L is the length of the region where the laser pump beams intersect in the crystal, α_p is the absorption at the polariton frequency ω_p, and $\mathrm{Sinc}(x)=\sin x/x$. The parameter β is deter-

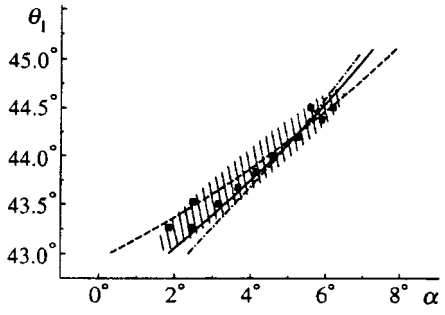


FIG. 4. Calculated zero detuning curves for the spatial synchronism in crystalline $LiNbO_3$ with an Mg concentration of 0.68% ($\omega_p=558.5\ \mathrm{cm}^{-1}$, $n_p=6.27$, $\theta_2=19.2°$). The smooth curve corresponds to synchronism of the direct process, $\Delta\mathbf{k}=0$, the dot-dashed curve to synchronism of the first stage of the cascade process, $\tau=0$ and the dashed curve to synchronism of the second state of the cascade process, $\sigma=0$. The points represent the locations of the experimentally measured scattering intensity peaks. The region in which four-wave scattering is observed is shaded.

mined by the ratio of the contributions from the direct and cascade processes to the scattered light intensity:

$$\beta = \frac{n_p c \cos \vartheta_p}{4\pi\omega_p} \frac{\chi^{(3)}}{\chi_1^{(2)}\chi_2^{(2)}}, \tag{6}$$

where ϑ_p is an angle which determines the orientation of the polariton wave vector in the crystal, n_p is the refractive index for the polariton wave in the crystal, $\chi^{(3)}$ is the effective value of the cubic susceptibility of the crystal, and $\chi_1^{(2)}$ and $\chi_2^{(2)}$ are the quadratic susceptibilities which determine the efficiencies of the first and second stages of cascade scattering.

The first term in the square brackets of Eq. (5) describes direct four-photon scattering owing to $\chi^{(3)}$. The direct scattering intensity must be greatest everywhere the equality $\Delta k = 0$ is satisfied. The smooth curve in Fig. 4 corresponds to this condition. Based on Eq. (5) we can draw the following general picture of the change in the total scattering intensity associated with tuning along the synchronism tuning curve $\Delta k = 0$. Far from the region of small detunings in the cascade process, τ, the intensity is constant and its magnitude is determined by the cubic susceptibility of the crystal at the frequencies under study here. As the region where $\tau \approx 0$ is approached, a contribution from the cascade process appears superposed on a background from the direct process.

The second term in the brackets of Eq. (5) describes the interference of the direct and cascade scattering mechanisms. Even for small $\chi^{(3)}$, when direct four-photon scattering is not observed, the interaction of these two processes may distort the scattering profile. The cascade process in itself is described by the third term in the square brackets of Eq. (5). In its pure form, for $\beta = 0$ the cascade process has a line profile given by the product of the function $\mathrm{Sinc}^2(\Delta kL/2)$ and the Lorentz factor $1/(\alpha_p^2 + 4\tau^2)$. The region where the cascade process can be observed in Fig. 4 must lie near the intersection of the curves corresponding to the condition $\Delta k = 0$ (the smooth curve) and $\tau = 0$ (dashed curve). The size of the region depends on the length L of the interaction region (i.e., on the width of the principal maximum of the function $\mathrm{Sinc}^2(\Delta kL/2)$), the absorption at the polariton frequency (i.e., the width of the Lorentz factor $1/(\alpha_p^2 + 4\tau^2)$), and the frequency–angular structure of the pump sources. In the observed scattering pattern we observed neither a direct contribution from direct four-photon scattering nor an interference structure in the scattering profiles associated with the second term in the brackets of Eq. (5). This indicates that the effective components of the cubic susceptibility are small. Our estimates yield $\beta\alpha_p < 10^{-2}$, which corresponds to $\chi^{(3)}\alpha_p / \chi_1^{(2)}\chi_2^{(2)} < 10^3\ \mathrm{cm}^{-1}$.

Incoherent cascade scattering, which is evidently also possible in our experimental arrangement, should be mentioned separately. This process reduces to spontaneous scattering of the pump probe wave on polaritons under conditions such that the population of the polariton state changes during parametric mixing of the two exciting pump waves. In this case the scattering profile must be described by the product of Lorentz factors with the same width determined only by the parameter α_p:

$$I_S \propto \frac{1}{\alpha_p^2 + 4\tau^2} \frac{1}{\alpha_p^2 + 4\sigma^2}, \tag{7}$$

where $\boldsymbol{\sigma} = \mathbf{k}_S - \mathbf{k}_L + \mathbf{k}_p$ is the detuning in the spatial synchronism for the second stage of the cascade interaction. The synchronism curve $\sigma = 0$ passes near the curves $\Delta\mathbf{k} = 0$ and $\tau = 0$ (shown in Fig. 4 as a dot-dashed curve). The fundamental qualitative difference in the line profiles for coherent and incoherent cascade scattering lies in the possibility of observing a scattering profile $I_S(\alpha)$ with two peaks. A dependence of the form (7) should not give rise to double peaks in the scattering profile. The observation of double peaks in our experiment indicates that the cascade scattering is coherent.

A comparison of the experimental data with the theoretical picture indicates that the observed scattering is purely cascade and coherent. Intense narrow single scattering peaks are observed near the intersection of the exact synchronism curve for the four-wave interaction and the synchronism curve for generating a polariton wave. With increasing distance from the intersection point the widths of the maxima increase and the intensity falls off fairly rapidly. The intensity of the four-wave scattering is substantially higher than that of the spontaneous three-wave scattering of the probe pump light only within a narrow (on the order of a few degrees) angular range of variation in θ_1. Assuming purely cascade scattering, we have calculated the interaction parameters for which each of the experimentally measured scattering line profiles is best fit by an expression of the form (5) with $\beta = 0$. The fitting parameters that were varied were the polariton absorption α_p, the length of the interaction region L, and the effective diameter of the pump beams at the focussing lenses d, since according to our optical arrangement the cross sections of the unfocussed beams of all three pump lasers were the same. Then the effective diameter of the pump beams and the focal lengths of the lenses ultimately determined the divergence of the radiation incident on the crystal. For each spectrum measured at the signal frequency, which corresponded to a polariton frequency of 558.5 cm^{-1}, we obtained essentially a single set of values for the optimal fit parameters: $\alpha_p = 23 \pm 2\ \mathrm{cm}^{-1}$, $d = 1.25$ mm, and $L = 0.8$–1.0 mm.

These values of the beam parameters d and L are in agreement with the optical layout. L could vary from spectrum to spectrum owing to incomplete overlap of the pump beam cross sections in the crystal. The rather good constancy of the fitting parameters over the entire series of measurements indicates that the model we have chosen is valid. However, the values of the polariton absorption do not agree at all with data from other methods. The results from three-photon spectroscopy differ from our value $\alpha_p = 23\ \mathrm{cm}^{-1}$ by more than an order of magnitude. A calculation of the absorption based on an oscillator model using well known data[11] on the frequencies and oscillator strengths of phonons in lithium niobate yield a value for α_p on the order of 10^3 within this frequency range. The differences between the calculated values for pure lithium niobate and the α_p measured by the three-photon scattering technique in an impure crystal are relatively small and can be explained by the presence of

the Mg impurity and the different stoichiometry of the crystals. Nevertheless, they are all considerably higher than the values we have obtained by the method described above from four-wave k-spectra. We note that previously this method for estimating the absorption has been used[12] in a study of surface polaritons in GaP. The magnitude of the imaginary part of the wave vector of the surface polariton was measured. The results were of the same order of magnitude as in our experiment. However, they were in much better agreement with values calculated on the basis of the oscillator model, although, as opposed to the case of the real part of the wave vector, they did not agree completely with the calculated results.

This disagreement casts doubt on the validity of the model used here, although it does describes all the other features of the observed picture well. Perhaps in our case the parameter α_p in Eq. (5) is not determined solely by the imaginary part of the linear susceptibility of the crystal, but depends on the power of the incident laser pump power through the nonlinear susceptibility. Another reason for the observed narrowing of the polariton scattering line may be an effect of the weak photorefractive properties of crystalline Mg:$LiNbO_3$. Regardless, the feasibility of determining the absorption coefficient of the free polariton wave from four-wave scattering spectra remains in question.

5. CONCLUSION

We have studied the processes leading to four-photon scattering of light in the case of crystalline lithium niobate, which has the following previously known parameters: refractive indices in the visible, near IR, and polariton ranges, damping coefficients in the polariton region, and the magnitudes of the quadratic polarizability. The following conclusions are important from the standpoint of future spectroscopic applications of this type of scattering:

1. The signal has been found to be much more intense (by roughly five orders of magnitude) than the signals from spontaneous three-frequency parametric light scattering in volumes on the order of 1 mm^3. This indicates that it may be possible to further reduce the test volume of the medium that is studied, and to investigate small samples and inhomogeneities in the distribution of the polariton parameters in macroscopic objects with high spatial resolution.

2. The cascade scattering mechanism has been found to predominate substantially at frequencies of $520-580\ cm^{-1}$ in lithium niobate. This is the type of four-wave interaction which is most sensitive to the parameters of the polariton state in materials. Our measurements indicate that the real part of the polariton wave vector and the refractive index of a crystal at the polariton frequency can be determined with high accuracy. In analyzing the experimental data presented in Fig. 4 we have used a value, $n_p=6.270$, for the ordinary refractive index which is more accurate than the that obtained from experiments employing three-photon scattering. A change in this value by just 0.005 leads to complete disagreement between the observed and theoretical pictures. Meanwhile, the spontaneous three-wave light scattering method in this part of the polariton spectrum is several times less accurate.

3. Determining the imaginary part of the polariton wave vector from the four-wave scattering line profile on the basis of the existing theory[4,5,7,8] leads to results that are not consistent with data from other methods for lithium niobate crystals:

The main points at which four-wave polariton spectroscopy is inferior to three-wave spontaneous polariton light scattering are:

(a) the need to satisfy a large number of synchronism conditions and, as a consequence, more complicated alignment of the spectrometer optics in order to observe the signal;

(b) the possible simultaneous occurrence of several processes and their interference which means that the initial interpretation of the spectra is not unique;

(c) the relative complexity of the experimental apparatus, including two or three laser sources, one of which is frequency-tuneable, etc.; and

(d) there is some doubt of the feasibility of measuring the imaginary parts of the polariton wave vector and the dielectric constant of the medium.

In any case, it must be recognized that four-wave spectroscopy is not currently as well developed for crystals as is the four-wave coherent spectroscopy of gases and condensed media. Three-wave spontaneous Raman scattering remains, as before, essentially the only widely used light scattering technique for studying crystals. Nevertheless, the most important advantage of four-wave spectroscopy is its much higher sensitivity. It appears that the prospects for the further development and use of this method for studying polariton states in crystals exist only in combination with the three-wave technique. Preliminary three-wave spectroscopy data on the character of the polariton dispersion greatly simplifies the procedure for adjusting the spatial synchronism in order to observe a four-wave interaction signal. In this regard using the Stokes variant of the four-wave interaction is of special interest. A Stokes scheme makes it possible to proceed in the simplest way from observing four-wave scattering to a spontaneous three-wave scattering regime. since the signal light lies in the same spectral and angular regions. Our studies show, however, that full realization of this method requires further development of scattering theory including absorption at the polariton frequency. As before, the search for new spectroscopic schemes that will ensure easier satisfaction of the spatial synchronism conditions remains a pressing problem.

We thank S. P. Kulik and M. V. Chekhova for valuable comments and help in preparing the manuscript. This work was supported by the Russian Fund for Fundamental Research (Grant No. 96-02-16336).

[1] D. N. Klyshko, *Photons and Nonlinear Optics* (Gordon and Breuh, New York, 1988).
[2] J. F. Coffinet and F. De Martini, Phys. Rev. Lett. **22**, 60 (1969).
[3] M. D. Levenson and N. Bloembergen, Phys. Rev. B **10**, 4447 (1974).
[4] E. Yablonovitch, C. Flytzanis, and N. Bloembergen, Phys. Rev. Lett. **29**, 865 (1972).
[5] J. J. Wynne, Phys. Rev. Lett. **29**, 650 (1972).
[6] Yu. N. Polivanov and A. T. Sukhodol'skiĭ, JETP Lett. **25**, 221 (1977).

[7] V. L. Strizhevskiĭ and Yu. N. Yashkir, Kvant. Elektronika **2**, 995 (1975) [Sov. J. Quantum Electron. **5**, 541 (1975)].
[8] D. N. Klyshko, Kvant. Elektronika **2**, 265 (1975) [Sov. J. Quantum Electron. **5**, 149 (1975)].
[9] A. D. Aleksandrovskiĭ, G. I. Ershova, G. Kh. Kitaeva *et al.*, Kvant. Elektronika **18**, 254 (1991) [Sov. J. Quantum Electron. **21**, 225 (1991)].
[10] D. S. Smith, H. D. Riccius and R. P. Edwin, Opt. Commun. **17**, 332 (1976).
[11] A. S. Barker Jr. and R. Loudon, Phys. Rev. A **158**, 433 (1967).
[12] F. De Martini, G. Guiliani, P. Matalony *et al.*, Phys. Rev. Lett. **37**, 440 (1976).

Translated by D. H. McNeill

Time reversal during magnetization of gas atoms by a resonant light pulse

A. I. Alekseev

Moscow State Engineering Physics Institute, 115409 Moscow, Russia
(Submitted 24 December 1996)
Zh. Éksp. Teor. Fiz. **112**, 453–469 (August 1997)

The optically induced magnetic moment of a stationary atom is calculated as a function of time and the resonance detuning $\omega-\omega_{ba}$ to within a constant factor having the dimensions of a magnetic moment based on the symmetry of an atom in the field of a resonant light pulse and symmetry with respect to time reversal including the initial conditions. The even dependence of the optically induced magnetic moment on $\omega-\omega_{ba}$ for an elliptically polarized pulse with an isotropic initial state of the atom and its odd dependence on $\omega-\omega_{ba}$ in the case of a linearly polarized pulsed with an anisotropic initial state in the atom's alignment are shown to be consequences of symmetry with respect to time reversal and the initial conditions. This behavior is retained even after passage of the light pulse, when the resulting relaxation destroys the time reversal symmetry. The optically induced magnetization of an atomic gas is found to have analogous properties. © *1997 American Institute of Physics.* [S1063-7761(97)00508-8]

1. INTRODUCTION

Optically induced nonlinear magnetization of gases and solids is widely used in basic studies of the interaction of laser light with matter (for example, gases[1–3] and solids;[4–6] see also the references given there). These studies have given rise to a number of fundamental problems. One such problem is an experimental observation[1,2] of magnetization in isotropic atomic gases by resonant linearly polarized light pulses that is not consistent with symmetry. Thus, some previously unexamined aspects of this optical phenomenon merit further study.

This article is a theoretical study of the magnetization of gas atoms by a resonant light pulse with wave vector $\mathbf{k}$ and elliptical $\mathbf{l}_{\mathbf{k}\lambda}$ ($\lambda=\pm1$) or linear $\mathbf{l}_{\mathbf{k}}$ polarization when there is symmetry with respect to time reversal, i.e., $t'\rightarrow-t'$, where t' represents the time including wave retardation. The duration τ of the light pulses is short compared to the radiative lifetime of the excited state, so that relaxation does not occur during the interaction of an atom with a light pulse in the interval $0\leqslant t'\leqslant\tau$. Here the initial state at $t'=0$ plays an important role as it determines the symmetry of the atom in the field of the light pulse, which sets a direction for the induced magnetic moment of the atom, $\boldsymbol{\mu}_{at}^{(q)}(t')$, with $q=0$ and $q=2$, respectively, for isotropic and anisotropic initial alignment states of the atom. Furthermore, a perturbative solution of the quantum mechanical equation for the density matrix including the initial conditions when symmetry with respect to time reversal is present yields an even dependence on the resonance detuning $\omega-\omega_{ba}$ for $\boldsymbol{\mu}_{at}^{(0)}(t')$ and an odd dependence for $\boldsymbol{\mu}_{at}^{(2)}(t')$. Finally, $\boldsymbol{\mu}_{at}^{(q)}(t')$ is given in the form of a product of three factors, of which the first determines the vector properties, the second has the dimensions of a magnetic moment and depends on the characteristics of the resonant transition, and the third is a universal function of t' and $\omega-\omega_{ba}$ which is independent of the atomic characteristics. Finally, the following fundamental manifestation of time reversal during magnetization of a stationary atom by a resonant light pulse is established: if the induced atomic magnetic moment is proportional to the axial vector $i[\mathbf{l}_{\mathbf{k}\lambda}\mathbf{l}_{\mathbf{k}\lambda}^{*}]$, then it is an even function of $\omega-\omega_{ba}$ with a maximum at $\omega=\omega_{ba}$. If the induced atomic magnetic moment is proportional to an axial vector which is invariant with respect to the transformation $\mathbf{k}\rightarrow-\mathbf{k}$, then it is an odd function of $\omega-\omega_{ba}$ with a maximum at $0<|\omega-\omega_{ba}|$. Here these axial vectors are determined by the symmetry of the atom in the resonant pulse field including the initial state. The behavior found here makes it possible to determine the vector properties of the optically induced magnetic moment, as well as the dependence on time and $\omega-\omega_{ba}$ without solving the equation for the density matrix in detail and without calculating the sum over the projections of the angular momenta in the formula for the optically induced magnetization. The optically induced magnetic moment is then obtained to within a common factor having the dimensions of a magnetic moment.

Within an interval $\tau<t'\leqslant\infty$ after passage of a resonant light pulse, the excited atomic state decays owing to spontaneous emission of a photon $\hbar\omega_{ba}$ and the resulting relaxation destroys the time-reversal symmetry. However, the direction of the induced magnetic moment $\boldsymbol{\mu}_{at}^{(q)}(t')$ established during the previous time interval $0\leqslant t'\leqslant\tau$ and its characteristic dependence on $\omega-\omega_{ba}$ remain the same in the following interval $\tau\leqslant\tau'\leqslant\infty$ because the function $\boldsymbol{\mu}_{at}^{(q)}(t')$ is continuous at the time $t'=\tau$. In the presence of relaxation, $\boldsymbol{\mu}_{at}^{(q)}(t')$ breaks up into the sum of two terms, each of which can be calculated to within a common factor that has the dimensions of a magnetic moment.

These results are applicable to a gas of identical atoms and the magnetic moment per unit volume is calculated as the optically induced magnetization of this gas, $\boldsymbol{\mu}_{q}(t')$, with $q=0$ and $q=2$ for isotropic and anisotropic initial states of the gas. Calculations show that the vector properties of the optically induced magnetization of the atomic gas and the characteristic dependence on $\omega-\omega_{ba}$ are the same as for an isolated atom, although the functional dependence on time

and $\omega-\omega_{ba}$ is more complicated. Thus, time reversal during magnetization of gas atoms by a resonant light pulse also shows up in the optically induced magnetization. Within the interval $0\leqslant t'\leqslant\tau$, $\boldsymbol{\mu}_q(t')$ is calculated to within a common factor having the dimensions of a magnetic moment density. Meanwhile, for $\tau\leqslant t'\leqslant\infty$, $\boldsymbol{\mu}_q(t')$ has the form of a sum of two terms with similar common factors. The uncalculated common factors in the formulas for $\boldsymbol{\mu}_q(t')$ at times $0\leqslant t'\leqslant\tau$ and $\tau\leqslant t'\leqslant\infty$ will not affect experimental studies of the vector properties of the optically induced magnetization $\boldsymbol{\mu}_q(t')$ and its dependence on time and $\omega-\omega_{ba}$ which employ measurements of the electromotive force in a detector coil in relative units.[1,2,4–6]

2. STATEMENT OF THE PROBLEM

Let us consider an atom in a center of mass coordinate system which, in the absence of a constant magnetic field, is characterized by the Hamiltonian

$$H=\sum_n \frac{1}{2m}\left[\mathbf{p}_n-\frac{e}{c}\mathbf{A}(\mathbf{r}_n,t)\right]^2+U_{\text{in}}, \tag{1}$$

where m, e, and $\mathbf{r}_n$ are the mass, charge, and radius vector of the nth electron, $\mathbf{p}_n=-i\hbar\nabla_n$ is its momentum operator, c is the speed of light in a vacuum, and U_{in} is the interaction energy of the electrons among themselves and with the nucleus which lies at the center of mass of the atom. The vector potential $\mathbf{A}(\mathbf{r}_n,t)$ is related to the electric $\mathbf{E}(\mathbf{r}_n,t)$ and magnetic $\mathbf{H}(\mathbf{r}_n,t)$ fields by the equations

$$\mathbf{E}(\mathbf{r}_n,t)=-\frac{1}{c}\frac{\partial}{\partial t}\mathbf{A}(\mathbf{r}_n,t),\quad \mathbf{H}(\mathbf{r}_n,t)=\text{curl}_n\mathbf{A}(\mathbf{r}_n,t). \tag{2}$$

The wave function Ψ of the atom satisfies the Schrödinger equation

$$i\hbar\frac{\partial\Psi}{\partial t}=H\Psi, \tag{3}$$

and the electric current density $\mathbf{j}_n$ created by the nth electron, including its spin, has the form

$$\mathbf{j}_n=\frac{e}{2m}(\Psi^*\mathbf{p}_n\Psi+\Psi\mathbf{p}_n^*\Psi^*)-\frac{e^2}{mc}\mathbf{A}(\mathbf{r}_n,t)\Psi^*\Psi+\frac{e\hbar}{m}\text{curl}_n\Psi^*\mathbf{s}_n\Psi, \tag{4}$$

where $\mathbf{s}_n$ is the spin operator of the nth electron. When the electron spin of the atom is taken into account, the wave function Ψ also depends on the spin variables. Because of the current (4), the atom has a magnetic moment

$$\boldsymbol{\mu}_{at}(t)=\frac{1}{2c}\int\sum_n[\mathbf{r}_n\mathbf{j}_n]d\mathbf{r}_n, \tag{5}$$

where, besides taking the integral over the electron coordinates, one takes the sum over the spin variables in the current $\mathbf{j}_n$.

Time-reversal symmetry exists for a given atom if the simultaneous substitutions

$$t\to-t,\quad \mathbf{A}(\mathbf{r}_n,-t)\to-\mathbf{A}(\mathbf{r}_n,t), \tag{6}$$

leave the electric field $\mathbf{E}(\mathbf{r}_n,t)$ unchanged and change the sign of the magnetic field $\mathbf{H}(\mathbf{r}_n,t)$. This symmetry shows up in the replacement of the wave function by its complex conjugate with the substitution (6); i.e.,

$$\Psi\to\Psi^*, \tag{7}$$

while the Schrödinger equation (3) is unchanged.[7] Here the electric current density (4) changes sign.[7,8] Thus, the magnetic moment (5) also changes sign with time reversal (Eqs. (6) and (7)), i.e., $\boldsymbol{\mu}_{at}(-t)=-\boldsymbol{\mu}_{at}(t)$. This last result was obtained without using the initial conditions at $t=0$, whose inclusion may lead to other features of the magnetic moment (5) associated with time reversal (Eqs. (6) and (7)).

If the atom interacts with a resonant light pulse, then its initial state at $t=0$, prior to the interaction with this pulse, will play a major role. In addition, some time after $t=0$, the relaxation owing to spontaneous emission becomes important. In order to determine the behavior of the magnetic moment (5) in this case, it is reasonable to consider the Schrödinger equation (3) in the dipole approximation, setting $\mathbf{A}(\mathbf{r}_n,t)=\mathbf{A}(0,t)$ in the Hamiltonian (1) and, therefore, in the current (4). Then, it is necessary to make a transformation

$$\Psi=\varphi\,\exp\left(\sum_n\frac{ie\mathbf{r}_n\mathbf{A}(0,t)}{\hbar c}\right), \tag{8}$$

in the Schrödinger equation (3) such that the new wave function φ including Eq. (2) satisfies another Schrödinger equation

$$i\hbar\frac{\partial\varphi}{\partial t}=(H_0-\mathbf{d}\mathbf{E}(0,t))\varphi, \tag{9}$$

containing the free atom hamiltonian

$$H_0=\sum_n\frac{\mathbf{p}_n^2}{2m}+U_{\text{in}}$$

and the interaction energy $-\mathbf{d}\cdot\mathbf{E}(0,t)$ of the dipole moment $\mathbf{d}=\Sigma_n e\mathbf{r}_n$ of the atom with the electric field $\mathbf{E}(0,t)$ at the site $\mathbf{r}_n=0$ of the nucleus (the center of mass of the atom). After the transformation (8), the formulas for the current (4) and magnetic moment (5) become much simpler:

$$\mathbf{j}_n=\frac{e}{2m}(\varphi^*\mathbf{p}_n\varphi+\varphi\mathbf{p}_n^*\varphi^*)+\frac{e\hbar}{m}\text{curl}_n\varphi^*\mathbf{s}_n\varphi,$$

$$\boldsymbol{\mu}_{at}(t)=\frac{e\hbar}{2mc}\int\sum_n\varphi^*(\mathbf{L}_n+2\mathbf{s}_n)\varphi d\mathbf{r}_n, \tag{10}$$

where $\mathbf{L}_n$ is the orbital angular momentum operator of the nth electron.

Let us assume that LS-coupling exists in an atom with zero nuclear spin with definite values of the quantum numbers J, L, and S corresponding to the operators for angular momentum $\mathbf{J}=\mathbf{L}+\mathbf{S}$, orbital angular momentum $\mathbf{L}=\Sigma_n\mathbf{L}_n$, and spin $\mathbf{S}=\Sigma_n\mathbf{s}_n$. Then the state of the atom is conveniently described by a representation in which the square of the angular momentum and its projection on the quantization axis have definite values. In this representation the magnetic moment (10) is written as[9]

$$\boldsymbol{\mu}_{at}(t) = -\mu_B \int \varphi^*(q,t) g \mathbf{J} \varphi(q,t) dq, \tag{11}$$

where

$$\mu_B = \frac{|e|\hbar}{2mc}, \quad g = 1 + \frac{J(J+1) - L(L+1) + S(S+1)}{2J(J+1)},$$

μ_B is the Bohr magneton, g is the Landé g-factor, q is the set of variables required to describe the state of the atom, and dq includes a sum over the spin variables.

The general case, where an atom moves in the laboratory coordinate system x, y, z at constant velocity $\mathbf{v}$ in an electric field $\mathbf{E}(\mathbf{r},t)$, is of practical interest. The behavior of the atom is described with the aid of the density matrix $\rho = \varphi^*(q',\mathbf{r},t)\varphi(q,\mathbf{r},t)$, where $\mathbf{r}$ is the radius vector of the center of mass of the atom and q is the set of variables for describing the atomic state in the center of mass coordinates. Given Eq. (9), the density matrix $\rho = \rho(q',q,\mathbf{r},t)$ satisfies the equation

$$\left(\frac{\partial}{\partial t} + \mathbf{v}\nabla\right)\rho = \frac{i}{\hbar}\left[(H_0' - \mathbf{d}'\mathbf{E}(\mathbf{r},t))\rho - (H_0 - \mathbf{d}\mathbf{E}(\mathbf{r},t))\rho\right], \tag{12}$$

where the hamiltonian H_0 and the dipole moment $\mathbf{d}$ are taken in the center of mass coordinate system. The magnetic moment (11) is expressed in terms of the density matrix as follows:

$$\boldsymbol{\mu}_{at}(t') = \frac{e\hbar}{2mc} \int (\mathbf{J} + \mathbf{S})\rho(q',q,\mathbf{r},t)|_{q'=q} dq. \tag{13}$$

Under time reversal and with the simultaneous substitutions $\mathbf{v} \to -\mathbf{v}$ and $\rho \to \rho^*$, Eq. (12) does not change, while the sign, but not the magnitude, of the magnetic moment (13) changes. By transforming to the matrix representation in Eq. (12) we can easily account for relaxation and the initial conditions when calculating ρ and the magnetic moment (13). This makes it possible to examine both short time intervals, within which time-reversal symmetry exists, and long time intervals, accounting for relaxation, which destroys the time-reversal condition.

3. MAGNETIZATION OF AN ATOM IN AN ISOTROPIC INITIAL STATE

Let an atom move at constant velocity $\mathbf{v}$ in a laboratory center-of-mass coordinate system x, y, z within a volume in which a resonant light pulse

$$\mathbf{E}(\mathbf{r},t) = \mathbf{l}_{\mathbf{k}\lambda} a(t') \exp(-i\omega t') + \text{c.c.}, \tag{14}$$

is propagating, where

$$t' = t - t_0 - \mathbf{k}(\mathbf{r} - \mathbf{r}_0)/\omega, \tag{15}$$

$\mathbf{l}_{\mathbf{k}\lambda}$ is the elliptical polarization vector with indices $\lambda = 1$ and $\lambda = -1$, respectively, for right- and left-hand rotation of the electric field vector $\mathbf{E} = \mathbf{E}(\mathbf{r},t)$, and $a(t')$ is the real amplitude, which is a slowly varying function compared to the exponential $\exp(-i\omega t')$. The leading edge of the light pulse (14) crosses the boundary point $\mathbf{r}_0$ of the volume at time t_0 and reaches the site $\mathbf{r}$ of the center of mass of the atom at time t. The frequency ω is close to the dipole transition frequency $\omega_{ba} = (E_b - E_a)/\hbar^{-1}$, where E_a and E_b are the energies of the ground and excited states of the free atom. The atomic state is characterized by these energies, but also by the quantum numbers J_a and J_b for the angular momentum $\mathbf{J}$ and its projections M_a and M_b on the quantization axis. Regardless of the choice of coordinate system, the vector $\mathbf{l}_{\mathbf{k}\lambda}$ has the form

$$\mathbf{l}_{\mathbf{k}1} = \mathbf{l}_{\mathbf{k}}^{(1)} \cos\psi + i\mathbf{l}_{\mathbf{k}}^{(2)} \sin\psi \tag{16}$$

for right polarization with $\lambda = 1$ and

$$\mathbf{l}_{\mathbf{k},-1} = -\mathbf{l}_{\mathbf{k}}^{(1)} \sin\psi + i\mathbf{l}_{\mathbf{k}}^{(2)} \cos\psi \tag{17}$$

for left polarization with $\lambda = -1$. Here the argument ψ takes values $0 \leqslant \psi \leqslant \pi/2$. The ratio of the axes of the polarization ellipse in Eqs. (16) and (17) is given, respectively, by $\varepsilon_1 = \sin\psi/\cos\psi$ and $\varepsilon_{-1} = \cos\psi/\sin\psi$, where $\varepsilon_1\varepsilon_{-1} = 1$. If $\psi = \pi/4$, then $\varepsilon_1 = \varepsilon_{-1} = 1$ and the polarization ellipse transforms to a circle, while Eqs. (16) and (17) describe right-circular and left-circular polarizations. If we set $\lambda = 1$ and $\psi = 0$ or $\lambda = -1$ and $\psi = \pi/2$, then Eqs. (14)–(17) describe a linearly polarized pulse. The unit vectors in Eqs. (16) and (17) satisfy the equations

$$\mathbf{k}\mathbf{l}_{\mathbf{k}}^{(1)} = \mathbf{k}\mathbf{l}_{\mathbf{k}}^{(2)} = \mathbf{l}_{\mathbf{k}}^{(1)}\mathbf{l}_{\mathbf{k}}^{(2)} = 0, \quad \mathbf{l}_{-\mathbf{k}}^{(1)} = \mathbf{l}_{\mathbf{k}}^{(1)}, \quad \mathbf{l}_{-\mathbf{k}}^{(2)} = -\mathbf{l}_{\mathbf{k}}^{(2)}, \tag{18}$$

and

$$[\mathbf{l}_{\mathbf{k}}^{(1)}\mathbf{l}_{\mathbf{k}}^{(2)}] = \mathbf{k}/k, \quad \mathbf{l}_{-\mathbf{k}\lambda} = \mathbf{l}_{\mathbf{k}\lambda}^*, \quad \mathbf{l}_{\mathbf{k}\lambda}\mathbf{l}_{\mathbf{k}\lambda'}^* = \delta_{\lambda\lambda'}. \tag{19}$$

On examining the behavior of the vectors (16) and (17), using the conditions (18) and (19) and making the substitution $\mathbf{k} \to -\mathbf{k}$ and the inversion $x \to -x$, $y \to -y$, $z \to -z$, we can show that

$$i[\mathbf{l}_{\mathbf{k}\lambda}\mathbf{l}_{\mathbf{k}\lambda}^*] = (\mathbf{k}/k)\lambda\beta \sin(2\psi), \tag{20}$$

where β is the unit pseudoscalar, which equals 1 and -1 in right and left hand coordinate systems, respectively. This last can be taken in the form $\beta = (\mathbf{l}_x \times \mathbf{l}_y)\cdot\mathbf{l}_z$, where $\mathbf{l}_x$, $\mathbf{l}_y$, and $\mathbf{l}_z$ are the unit vectors for a right hand cartesian coordinate system x, y, z. Then, along with Eq. (20), we find

$$\mathbf{l}_{\mathbf{k}\lambda}^2 = \lambda \cos(2\psi).$$

These vector and scalar products of the polarization vectors (16) and (17) are useful in a number of problems concerning reversal, as well as the substitutions $\mathbf{k} \to -\mathbf{k}$ and $\lambda \to -\lambda$, which occur in isolation or in various combinations. In these cases, besides the argument ψ we can use the ratio ε_λ of the elliptic axes of polarization with $\lambda = 1$ and $\lambda = -1$ because of the equalities

$$\sin(2\psi) = \frac{2\varepsilon_\lambda}{1+\varepsilon_\lambda^2}, \quad \cos(2\psi) = \frac{\lambda(1-\varepsilon_\lambda^2)}{1+\varepsilon_\lambda^2},$$

where on going from right (left) to left (right) polarization the quantities $\sin(2\psi)$ and $\cos(2\psi)$ do not change, since $\varepsilon_1\varepsilon_{-1} = 1$.

The pulse duration τ is short compared to the radiative lifetime γ^{-1} of the excited state of the atom, i.e.,

$$\gamma\tau \ll 1, \tag{21}$$

where

$$\gamma = 4|d_{ba}|^2 \omega_{ba}^3 / 3\hbar c^3 (2J_b + 1).$$

γ is the probability of spontaneous emission of a photon with energy $\hbar\omega_{ba}$ by the isolated atom, and d_{ba} is the reduced dipole moment.[9] The inequality (21) allows us to neglect relaxation and use Eq. (12) for studying the interaction of the atom with the pulse (14) in the time interval $0 \leqslant t' \leqslant \tau$, where, because of Eq. (15), the density matrix ρ is a function of t'. We also note that the leading edge of the pulse front (14) passes through the center of mass of the atom at time $t = t_0 + \mathbf{k}\cdot(\mathbf{r} - \mathbf{r}_0)/\omega$ or $t' = 0$, while prior to the interaction with the pulse (14), the atom is in an equilibrium isotropic state. Thus, in the JM-representation, the initial value of the density matrix $\rho = \rho(t')$ at $t' = 0$ can be written in the form

$$\rho_{M_b, M_b'}(0) = \rho_{M_b M_a}(0) = 0,$$

$$\rho_{M_a M_a'}(0) = (2J_a + 1)^{-1} \delta_{M_a M_a'}. \tag{22}$$

If the amplitude $a(t')$ is an even function, i.e., $a(-t') = a(t')$, then specifying it in the region $0 \leqslant t' \leqslant \infty$ in the form $a(t')$ for $0 \leqslant t' \leqslant \tau$ and $a(t') = 0$ for $\tau < t' \leqslant \infty$ is equivalent to specifying this amplitude on the entire time axis $-\infty \leqslant t' \leqslant \infty$. If the specified amplitude $a(t')$ depends arbitrarily on t' in the interval $0 \leqslant t' \leqslant \tau$ and $a(t') = 0$ for $\tau < t' \leqslant \infty$, then we define the amplitude $a(t')$ on the negative time half axis $-\infty \leqslant t' \leqslant 0$ formally as $a(t') = a(-t')$ for $-\tau \leqslant t' \leqslant 0$ and $a(t') = 0$ for $-\infty \leqslant t' < -\tau$. Then the field (14) and Eq. (12) do not change under the simultaneous substitutions

$$t \to -t, \quad t_0 \to -t_0, \quad \mathbf{k} \to -\mathbf{k}, \quad \mathbf{v} \to -\mathbf{v}, \quad \rho \to \rho^*, \tag{23}$$

which are accompanied by the substitution $t' \to -t'$. If the amplitude $a(t')$ is an odd function, i.e., $a(-t') = -a(t')$, then specifying it in the region $0 \leqslant t' \leqslant \infty$ is equivalent to specifying this amplitude on the entire time axis $-\infty \leqslant t' \leqslant \infty$. However, it must enter Eq. (14) in the form $a(t')\exp(i\pi/2)$ or $a(t')\exp(-i\pi/2)$ so that the substitutions (23) will not change the electric field (14). Thus, in all the following formulas, for an odd $a(t')$ it is necessary to make the substitution $a(t') \to a(t')\exp(i\pi/2)$ or $a(t') \to a(t')\exp(-i\pi/2)$.

For an atom in an electric field (14) with this amplitude $a(t')$ there is a symmetry with respect to the time-reversal (23) and, in the absence of relaxation, the magnetic moment (13) obeys the equation

$$\boldsymbol{\mu}_{at}(-t') = -\boldsymbol{\mu}_{at}(t'). \tag{24}$$

This property of the magnetic moment has been obtained without using the initial conditions (22). In this problem, however, the time $t = t_0 + \mathbf{k}\cdot(\mathbf{r} - \mathbf{r}_0)/\omega$ or $t' = 0$ is the onset time for the interaction of the atom with the field (14) in the interval $0 \leqslant t' \leqslant \tau$. Thus, here when there is symmetry with respect to time reversal (23), the origin $t' = 0$ on the time axis in the region $-\infty \leqslant t' \leqslant \infty$ is a physically distinct time. If we calculate the magnetic moment of the atom using Eqs. (12)–(22), then, besides the properties (24), we can find other characteristics of $\boldsymbol{\mu}_{at}(t')$ that result from the time reversal symmetry (23) and the initial conditions (22).

The symmetry of the atom in the electric field (14) (together with conditions (22)) implies that the optically induced magnetic moment (13) is proportional only to the axial vector (20) in this case. In order to find the other characteristics, we shall use Eq. (12) in the interval $0 \leqslant t' \leqslant \tau$ in the JM-representation:

$$\left(\frac{\partial}{\partial t} + \mathbf{v}\nabla + i\omega_{ba}\right)\rho_{M_b M_a} = \frac{i}{\hbar}\,(\mathbf{Ed}_{M_b M_a'}\rho_{M_a' M_a} - \rho_{M_b M_b'}\mathbf{Ed}_{M_b' M_a}), \tag{25}$$

$$\left(\frac{\partial}{\partial t} + \mathbf{v}\nabla\right)\rho_{M_b M_b'} = \frac{i}{\hbar}\,(\mathbf{Ed}_{M_b M_a}\rho_{M_a M_b'} - \rho_{M_b M_a}\mathbf{Ed}_{M_a M_b'}), \tag{26}$$

$$\left(\frac{\partial}{\partial t} + \mathbf{v}\nabla\right)\rho_{M_a M_a'} = \frac{i}{\hbar}\,(\mathbf{Ed}_{M_a M_b}\rho_{M_b M_a'} - \rho_{M_a M_b}\mathbf{Ed}_{M_b M_a'}), \tag{27}$$

as well as Eq. (13) in the JM-representation,

$$\boldsymbol{\mu}_{at}(t') = -\mu_B[g_a \mathbf{J}_{M_a' M_a}\rho_{M_a M_a'}(t') + g_b \mathbf{J}_{M_b' M_b}\rho_{M_b M_b'}(t')], \tag{28}$$

where g_a and g_b are the Landé g-factors for the ground and excited states of the atom. Repeated matrix subscripts denote summation.

On solving Eqs. (25)–(27) by second order perturbation theory in the field (14) and using Eq. (22), we find that the density matrices for the ground and excited levels have the same dependences on the time t' and resonance detuning $\omega - \omega_{ba} - \mathbf{k}\cdot\mathbf{v}$ when there is a Doppler shift $\mathbf{k}\cdot\mathbf{v}$ in the frequency (14) of the pulse; specifically,

$$\rho_{M_a M_a'}(t') = R_{M_a M_a'} I(t') + \text{H.c.},$$

$$\rho_{M_b M_b'}(t') = R_{M_b M_b'} I(t') + \text{H.c.}, \tag{29}$$

where

$$I(t') = \int_0^{t'} d\tau_2 \int_0^{\tau_2} d\tau_1 a^*(\tau_2) a(\tau_1) \exp[i\Delta(\tau_2 - \tau_1)], \tag{30}$$

$$\Delta = \omega - \omega_{ba} - \mathbf{kv},$$

$R_{M_a, M_a'}$ and $R_{M_b, M_b'}$ are matrices that can only be defined through detailed calculation. Substitution of the quantities (29) in Eq. (28), breaks up the magnetic moment $\boldsymbol{\mu}_{at}(t')$ into a sum of terms proportional to $I(t')$ and $I^*(t')$. These terms must contain $I(t')$ and $I^*(t')$ in a combination such that after the substitutions (23), the equality (24) will be satisfied. This implies that owing to Eq. (20) the unknown vector $\boldsymbol{\mu}_{at}(t')$ is proportional to the sum $I(t') + I^*(t')$. Thus, given Eqs. (20) and (29), the magnetic moment (28) can be written in the form

$$\boldsymbol{\mu}_{at}(t') = -(\mathbf{k}/k)\lambda\beta\,\sin(2\psi) M_0 \widetilde{X}_0(t', \Delta), \tag{31}$$

and

$$\widetilde{X}_0(t', \Delta) = (\tau a_0)^{-2}[I(t') + I^*(t')], \tag{32}$$

where a_0 is the maximum value of $|a(t')|$ in the given interval $0 \leqslant t' \leqslant \tau$, M_0 is a constant with the dimensions of a magnetic moment which depends on the characteristics of the resonant transition, and $\tilde{X}_0(t',\Delta)$ is a universal function of t' and Δ which is independent of the characteristics of the resonant transition.

The vector properties of the magnetic moment $\boldsymbol{\mu}_{at}(t')$ in Eq. (31) are caused by the symmetry of the atom in the field (14) with the initial conditions (22); this leads to an expression for the axial vector (20). Meanwhile, the even dependence of $\boldsymbol{\mu}_{at}(t')$ on the resonant detuning Δ is caused by symmetry with respect to time reversal (23) in the presence of the axial vector (20). The constant M_0 is determined by detailed calculation of the density matrices (29) in the course of solving Eqs. (25)–(27) with Eq. (22) and subsequently taking the sum in Eq. (28) over the projections of the angular momenta of the atom in the ground and excited states. This behavior for $\boldsymbol{\mu}_{at}(t')$ with a more complicated dependence on t' and Δ can also be obtained outside the framework of perturbation theory for ultrashort rectangular pulses with circular polarization, as follows from a rigorous solution[10] of the problem without relaxation.

4. MAGNETIZATION OF AN ATOM IN AN ANISOTROPIC INITIAL STATE

We now examine the magnetization of an atom by a resonant linearly polarized pulse

$$\mathbf{E}(\mathbf{r},t)=\mathbf{l}_{\mathbf{k}}a(t')\exp(-i\omega t')+\text{c.c.}, \tag{33}$$

where $\mathbf{l}_{\mathbf{k}}$ is the unit polarization vector which is independent of the transformation $\mathbf{k}\rightarrow-\mathbf{k}$, as Eqs. (16)–(19) imply. The other physical quantities in Eq. (33) are the same as in Eq. (14). Here Eqs. (25)–(27) and Eq. (28), as well as the discussion regarding the amplitude $a(t')$, are still valid for the pulse (33). According to Eq. (31), a linearly polarized pulse does not magnetize an atom in an isotropic initial state (22). Thus, we assume that prior to the interaction with the pulse (33), the atom was optically polarized by a strong resonant linearly polarized pulse with wave vector $\mathbf{k}_0$ collinear to $\mathbf{k}$ and a unit polarization vector $\mathbf{l}_0$ which is unaffected by the substitution $\mathbf{k}_0\rightarrow-\mathbf{k}_0$. Then the initial conditions for Eqs. (25)–(27) in the JM-representation take the form

$$\rho_{M_bM_a}(0)=\rho_{M_bM_b'}(0)=0, \quad \rho_{M_aM_a'}(0)=\rho^{al}_{M_aM_a'}, \tag{34}$$

where the density matrix $\rho^{al}_{M_a,M_a'}$ describes the initial optically polarized atom for $1 \leqslant J_a$, which is referred to as the atomic alignment (see Alekseev,[11] for example). Here the alignment of the atom is characterized by two orthogonal symmetry axes, of which the first is directed along $\mathbf{l}_0$ and the second is collinear with $\mathbf{k}_0$. In this case, when the atom interacts with the pulse (33), there is a single axial vector corresponding to the given symmetry which determines the direction of the magnetic moment (28) and has the form

$$[\mathbf{l}_0\mathbf{l}_{\mathbf{k}}]=(\mathbf{k}/k)\sin\varphi_{\mathbf{k}}, \tag{35}$$

where the positive direction for the angle $\varphi_{\mathbf{k}}$ is clockwise from the unit vector $\mathbf{l}_0$ to the unit vector $\mathbf{l}_{\mathbf{k}}$ on looking along $\mathbf{k}$. Thus, on making the substitution $\mathbf{k}\rightarrow-\mathbf{k}$ we have $\varphi_{-\mathbf{k}}=-\varphi_{\mathbf{k}}$, which characterizes the angle $\varphi_{\mathbf{k}}$ as a pseudoscalar. In addition, of the two orthogonal symmetry axes, only one has the specified direction $\mathbf{l}_0$, while the direction of the other axis is not specified (is equally probable). The direction of the third axis, which is orthogonal to these two, is also not specified and is equally probable. Thus, for an initially aligned atom the magnetic moment (28) should not change if one chooses $\mathbf{l}_{\mathbf{k}}$ parallel or antiparallel to the third axis. These two directions of $\mathbf{l}_{\mathbf{k}}$ correspond to angle of rotation $\varphi_{\mathbf{k}}=\pi/2$ and $\varphi_{\mathbf{k}}=-\pi/2$. Invariance of the magnetic moment (28) with respect to such rotations under these symmetry conditions can be attained after multiplying the axial vector (35) by $\cos\varphi_{\mathbf{k}}$. Thus, the unknown magnetic moment (28) in this case is proportional to the axial vector

$$(\mathbf{k}/k)\sin(2\varphi_{\mathbf{k}}). \tag{36}$$

For the initial conditions (34) the dependence of the density matrices (29) on t' and Δ is valid for other values of $R_{M_a,M_a'}$ and $R_{M_b,M_b'}$ compared to the isotropic case (22). We note also that for an initially aligned atom in the field of a linearly polarized pulse (33) there is a symmetry under time-reversal (23) which yields Eq. (24). This equality is possible only if after substituting the density matrices (29) in Eq. (28) the $I(t')$ and $I^*(t')$ enter $\boldsymbol{\mu}_{at}(t')$ in the form of the real combination $i[I^*(t')-I(t')]$. Ultimately, the magnetic moment (28) for an initially aligned atom has the form

$$\boldsymbol{\mu}_{at}(t')=-(\mathbf{k}/k)\sin(2\varphi_{\mathbf{k}})M_2\tilde{X}_2(t',\Delta), \tag{37}$$

where

$$\tilde{X}_2(t',\Delta)=i(\tau a_0)^{-2}[I^*(t')-I(t')]. \tag{38}$$

Here the constant M_2, which has the dimensions of a magnetic moment, is calculated by taking the sum in Eq. (28) over the projections of the angular momenta. The subscript 2 on M_2 and in the universal function $X_2(t',\Delta)$ is chosen so that it coincides with the rank of the polarization multipole moment characterizing the atomic alignment in Eq. (34). The vector properties of the magnetic moment in Eq. (37) are found as a consequence of the symmetry which exists during the interaction of an initially aligned atom with a linearly polarized pulse (33). In the meantime, the odd dependence on Δ in Eq. (37) is caused by the symmetry under time reversal (23) in the presence of the initial conditions (34) leading to the axial vector (36).

5. ACCOUNTING FOR RELAXATION IN THE MAGNETIZATION OF AN ATOM

For ultrashort pulses (14) and (33) which satisfy the inequality (21) there is no relaxation in the interval $0 \leqslant t' \leqslant \tau$ and the induced magnetic moments (31) and (37) are given by

$$\boldsymbol{\mu}^{(q)}_{at}(t')=-\mathbf{L}_qM_q\tilde{X}_q(t',\Delta), \quad q=0,2, \tag{39}$$

where

$$\mathbf{L}_0=(\mathbf{k}/k)\lambda\beta\sin(2\psi), \quad \mathbf{L}_2=(\mathbf{k}/k)\sin(2\varphi_{\mathbf{k}}). \tag{40}$$

On the other hand, relaxation must be taken into account for $\tau < t' \leqslant \infty$, i.e., after pulses (14) and (33) have passed.

Equation (12) is written in the JM-representation for the density matrices for the excited and ground states including relaxation in the region $\tau<t'\leqslant\infty$ as

$$\left(\frac{\partial}{\partial t}+\mathbf{v}\nabla+\gamma\right)\rho_{M_bM_b'}(t')=0, \tag{41}$$

and

$$\left(\frac{\partial}{\partial t}+\mathbf{v}\nabla\right)\rho_{M_aM_a'}(t')$$
$$=\frac{\gamma(2J_b+1)}{|d_{ba}|^2}\,\mathbf{d}_{M_aM_b}\rho_{M_bM_b'}(t')\mathbf{d}_{M_b'M_a'}, \tag{42}$$

where the density matrices are regarded as functions of the time t' including the wave retardation (15), since the initial conditions for Eqs. (41) and (42) are formulated in the t' time scale and are determined by the density matrices (29) for $t'=\tau$ in the form

$$\rho_{M_bM_b'}(\tau)=R_{M_bM_b'}I(\tau)+\text{H.c.},$$
$$\rho_{M_aM_a'}(\tau)=R_{M_aM_a'}I(\tau)+\text{H.c.} \tag{43}$$

The solution of Eqs. (41) and (42) including Eq. (43) is valid for $\tau\leqslant t'\leqslant\infty$ and has the form

$$\rho_{M_bM_b'}(t')=\rho_{M_bM_b'}(\tau)\exp[-\gamma(t'-\tau)]. \tag{44}$$

and

$$\rho_{M_aM_a'}(t')=\rho_{M_aM_a'}(\tau)+\frac{2J_b+1}{|d_{ba}|^2}\,\mathbf{d}_{M_aM_b}\rho_{M_bM_b'}(\tau)$$
$$\times\mathbf{d}_{M_bM_a'}\{1-\exp[-\gamma(t'-\tau)]\}, \tag{45}$$

where small terms of order v/c have been omitted.

After passage of the pulses (14) and (33) the excited state of the atom is quenched because of relaxation in Eq. (44). Thus, with time the magnetic moment (28) in the excited level decays to zero, while, because of Eq. (45), that at the ground state approaches a constant value. Here relaxation in $\tau<t'\leqslant\infty$ has no effect on the symmetry of the atom in the fields (14) and (33) within the earlier time interval $0\leqslant t'\leqslant\tau$ which determined the directions of the induced magnetic moments with the aid of the vectors (40). In addition, according to Eqs. (44) and (45) the dependence of the magnetic moment (28) on Δ for $\tau\leqslant t'\leqslant\infty$ is determined by the initial conditions (43) and has the same character as in the preceding interval $0\leqslant t'\leqslant\tau$. This implies that after substitution of the density matrices (44) and (45) in Eq. (28), the induced magnetic moment for $\tau\leqslant t'\leqslant\infty$ has the form

$$\boldsymbol{\mu}_{at}^{(q)}(t')=-\mathbf{L}_q\widetilde{X}_q(\tau,\Delta)\{M_q^a+M_q^{ab}+(M_q^b-M_q^{ab})$$
$$\times\exp[-\gamma(t'-\tau)]\},\quad q=0,2, \tag{46}$$

where M_q^a, M_q^{ab}, and M_q^b are constants with the dimensions of a magnetic moment which are determined only by detailed calculation. The requirement of continuity in the induced magnetic moment $\boldsymbol{\mu}_{at}^{(q)}(t')$ as a function of t' leads to the equation $M_q^a+M_q^b=M_q$.

It is clear from the continuity in the function $\boldsymbol{\mu}_{at}^{(q)}(t')$ at $t'=\tau$, which is given by Eqs. (39) and (46) in the intervals $0\leqslant t'\leqslant\tau$ and $\tau\leqslant t'\leqslant\infty$, that the vector properties of the magnetic moment (28) over the entire region $0\leqslant t'\leqslant\infty$ are determined by the symmetry during the interaction of the atom with fields (14) and (33) with the initial conditions at $t'=0$ taken into account. The characteristic dependence on Δ in Eqs. (39) and (46) is, however, determined by symmetry with respect to time reversal (23), which existed in the interval $0\leqslant t'\leqslant\tau$ without relaxation. Here for Eq. (46) in the interval $\tau<t'\leqslant\infty$ Eq. (24) loses its significance.

6. THE TIME REVERSAL EFFECT FOR OPTICALLY INDUCED MAGNETIZATION OF A GAS

We now apply Eqs. (39) and (46) to a gas of identical atoms within the volume under consideration. Then we find that the optically induced magnetization $\boldsymbol{\mu}_q(t')$ of this gas owing to elliptically polarized (14) and linearly polarized (33) pulses is given by

$$\boldsymbol{\mu}_q(t')=-L_qNM_qX_q(t',\omega-\omega_{ba}),\quad 0\leqslant t'\leqslant\tau, \tag{47}$$

and

$$\boldsymbol{\mu}_q(t')=-\mathbf{L}_qNX_q(\tau,\omega-\omega_{ba})\{M_q^a+M_q^{ab}+(M_q^b$$
$$-M_q^{ab})\exp[-\gamma(t'-\tau)]\},\quad \tau\leqslant t'\leqslant\infty, \tag{48}$$

where

$$X_q(t',\omega-\omega_{ba})=\int f(v)\widetilde{X}_q(t',\Delta)d\mathbf{v},\quad q=0,2,$$
$$f(v)=\frac{1}{(\pi^{1/2}u)^3}\exp\left(-\frac{v^2}{u^2}\right),\quad u=\left(\frac{2k_BT}{m_{at}}\right)^{1/2},$$

N is the density of atoms, $f(v)$ is the Maxwellian distribution function, u is the most probable velocity, k_B is the Boltzmann constant, T is the gas temperature, and m_{at} is the mass of an atom. The universal functions $X_q(t',\omega-\omega_{ba})$ with $q=0.2$ after integration over the velocity $\mathbf{v}$ take the form

$$X_0(t',\omega-\omega_{ba})=2(\tau a_0)^{-2}\int_0^{t'}d\tau_2\int_0^{\tau_2}d\tau_1a^*(\tau_2)a(\tau_1)$$
$$\times\exp\left[-\left(\frac{\tau_2-\tau_1}{2\tau_D}\right)^2\right]\cos[(\omega-\omega_{ba})(\tau_2$$
$$-\tau_1)], \tag{49}$$

and

$$X_2(t',\omega-\omega_{ba})=2(\tau a_0)^{-2}\int_0^{t'}d\tau_2\int_0^{\tau_2}d\tau_1a^*(\tau_2)a(\tau_1)$$
$$\times\exp\left[-\left(\frac{\tau_2-\tau_1}{2\tau_D}\right)^2\right]\sin[(\omega-\omega_{ba})(\tau_2$$
$$-\tau_1)], \tag{50}$$

where $\tau_D=1/ku$ is the Doppler relaxation time. Because of averaging over the velocity $\mathbf{v}$, the universal functions (49) and (50) have more complicated dependence on t' and $\omega-\omega_{ba}$ than Eqs. (32) and (38), but the even and odd de-

pendence on the resonance detuning remains. Furthermore, the quantities (49) and (50) also depend on the gas temperature T and atomic mass m_{at}. In the case of an ultrashort elliptically polarized pulse (14) the results for the optically induced magnetization of the gas (47)–(49) are in agreement with previous work[10] if there we set $N_a=N$, $\gamma_b=\gamma$, and $N_b-\gamma_a=0$, and use the equation

$$(\tau a_0)^2 X_0(t',\omega-\omega_{ba})=\int f(v)\left|\int_0^{t'} a(\xi)\times\exp(-i\Delta\xi)d\xi\right|^2 d\mathbf{v}.$$

In some experiments[1,2] on optically induced magnetization of a gas the resonant light pulses were passed through a cell with the test gas mounted inside a pickup coil. The variable optically induced magnetization excited an electrical current in the winding of the pickup coil and created a potential difference between its ends of

$$U(t')=-C\mathbf{n}\frac{d\boldsymbol{\mu}_q(t')}{dt'},\quad q=0,2, \tag{51}$$

where the constant C depends on the experimental circuit and the unit vector $\mathbf{n}$ is collinear with $\mathbf{k}$ and determines a fixed direction for which the magnetic induction flux is measured. Thus, $U(t')$ represents an electromotive force in the pickup coil which makes it possible to study experimentally the vector properties of the optically induced magnetization and its dependence on t', $\omega-\omega_{ba}$, γ, T, and m_{at}.

For comparison with experiment in the interval $0\leqslant t'\leqslant\tau$ it is necessary to use Eq. (47) and then the electromotive force (51) can be measured in relative units as

$$\frac{U(t')}{CNM_q}=\mathbf{nL}_q\frac{dX_q(t',\omega-\omega_{ba})}{dt'},\quad q=0,2. \tag{52}$$

Because of Eq. (48), in the other interval $\tau<t'\leqslant\infty$ we have

$$U(t')[C\tau^{-1}N(M_q^{ab}-M_q^b)]^{-1}=\mathbf{nL}_q\gamma\tau X_q(\tau,\omega-\omega_{ba})\exp[-\gamma(t'-\tau)],\quad q=0,2, \tag{53}$$

which makes possible an experimental determination of the width $\hbar\gamma$ of the excited level, as well as the dependence on τ', $\omega-\omega_{ba}$, T, and m_{at}.

The variable optically induced magnetization at $t'=0$ is zero, while at later times $\gamma^{-1}\ll t'$ it takes the asymptotic value $\boldsymbol{\mu}_q(\infty)$. Thus, after integrating both parts of Eq. (51), we find

$$\int_0^\infty U(t')dt'=-C\mathbf{n}\boldsymbol{\mu}_q(\infty),\quad q=0,2.$$

This implies that by numerically integrating the experimental plot of $U(t')$ as a function of t' for an isolated resonant light pulse (14) or (33) one can determine the universal functions $X_q(\tau,\omega-\omega_{ba})$ with $q=0,2$ as a result of the equation

$$\mathbf{nL}_q X_q(\tau,\omega-\omega_{ba})=[CN(M_q^a+M_q^{ab})]^{-1}\int_0^\infty U(t)dt. \tag{54}$$

In experiments with gases[1,2] and solids[4–6] the electromotive force (51) has been measured in relative units. Equations (52)–(54) also contain $U(t')$ in relative units. If the experimental methods of Refs. 1, 2, and 4–6 are used, then the uncomputed constants M_q, M_q^a, M_q^b, and M_q^{ab} in Eqs. (47) and (48) do not affect the experimental study of the vector properties of the optically induced magnetization $\boldsymbol{\mu}_q(t')$ with $q=0,2$ or its dependence on time and the quantities $\omega-\omega_{ba}$, γ, T, and m_{at}.

As an example, let us consider light pulses (14) and (33) with a rectangular pulse shape (amplitude)

$$a(t')=\begin{cases}a_0 & \text{for}\quad 0\leqslant t'\leqslant\tau,\\ 0 & \text{for}\quad \tau<t'\leqslant\infty,\end{cases} \tag{55}$$

as well as with a Gaussian pulse shape

$$a(t')=a_0\exp[-(t'-\tau/2)^2/(2\tau_p)^2],\quad 0\leqslant t'\leqslant\tau, \tag{56}$$

where $a(t')=0$ for $\tau<t'\leqslant\infty$ and $4\tau_p<\tau$. The characteristic time τ_p enters the spectral width $\Delta_\omega=(2\ln 2)^{1/2}/\tau_p$ of a Gaussian pulse if the inequality $4\tau_p<\tau$ is sufficiently strong (e.g., $8\tau_p\leqslant\tau$).

In the case of an inhomogeneously broadened transition $\gamma\tau_D\ll 1$ holds, and the universal function (49) at time $t'=\tau$ for a rectangular pulse (55) with $\tau=8\tau_D$ is given by

$$X_0(\tau,\omega-\omega_{ba})=(\sqrt{\pi}/4)\exp[-(\omega-\omega_{ba})^2\tau_D^2],$$

and for a Gaussian pulse (56) with $\tau=8\tau_p$, by

$$X_0(\tau,\omega-\omega_{ba})=\frac{\pi}{16}\left[1+2\left(\frac{\tau_p}{\tau_D}\right)^2\right]^{-1/2}\times\exp\left[-\frac{2\tau_p^2(\omega-\omega_{ba})^2}{1+2(\tau_p/\tau_D)^2}\right].$$

The universal function (50) is given by the formula

$$X_2(t',\omega-\omega_{ba})=2\int_0^{t'/\tau}d\xi\int_0^\xi d\eta\,\sin[\delta(\xi-\eta)]\times\exp\left\{-\varepsilon\left[\left(\xi-\frac{1}{2}\right)^2+\left(\eta-\frac{1}{2}\right)^2\right]-\left(\frac{\tau}{2\tau_D}\right)^2(\xi-\eta)^2\right\}, \tag{57}$$

where

$$\delta=(\omega-\omega_{ba})\tau,$$

$\varepsilon=0$ for the rectangular pulse (55) and $\varepsilon=(\tau/2\tau_p)^2$ for the Gaussian pulse (56). Figure 1 shows the behavior of the universal function (57) at a fixed time $t'=\tau$ for arbitrary values of $\omega-\omega_{ba}$. We consider the cases of a rectangular pulse (55) with $\varepsilon=0$ and a Gaussian pulse (56) with $\tau=8\tau_p$ and $\varepsilon=16$. The strong inequality $\gamma\tau_D\ll 1$ for an inhomogeneously broadened transition made it possible to study the two limiting cases of a substantial Doppler effect with

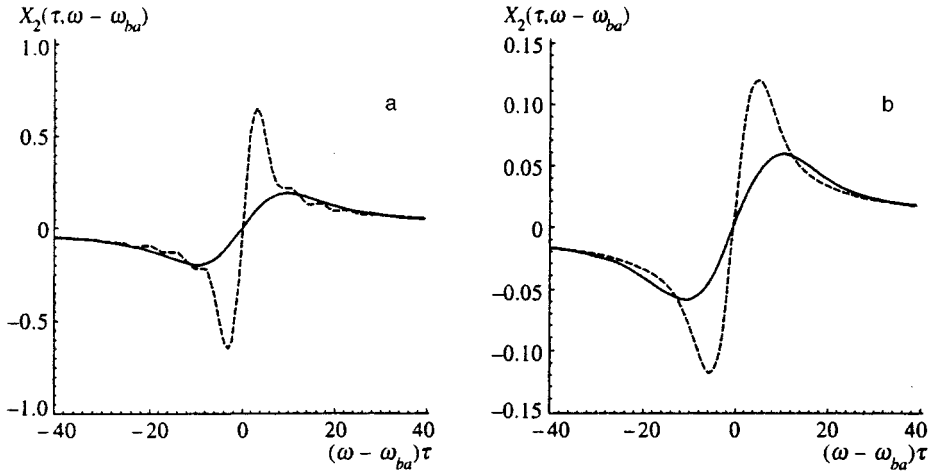


FIG. 1. The universal function (57) at a fixed time $t'=\tau$ and arbitrary resonant detunings $\omega-\omega_{ba}$. The smooth curve corresponds to including the Doppler effect with $\tau=10\tau_D$. The dashed curve corresponds to the case $10\tau=\tau_D$, when the Doppler effect can be neglected. (a) $\varepsilon=0$ for a rectangular pulse (55); (b) $\varepsilon=16$ for a Gaussian pulse with $\tau=8\tau_p$ (56).

$\tau=10\tau_D$ and a negligible Doppler effect with $10\tau=\tau_D$. The maximum of the odd function $X_2(\tau,\omega-\omega_{ba})$ for the rectangular pulse (55) with $\varepsilon=0$ in the absence of the Doppler effect ($10\tau=\tau_D$) occurs at $\omega-\omega_{ba}=3.2/\tau$, while with a strong Doppler effect ($\tau=10\tau_D$), we have $\omega-\omega_{ba}=9.9/\tau$. A similar analysis for a Gaussian pulse with $\tau=8\tau_p$ and $\varepsilon=16$ leads to a resonant detuning of $\omega-\omega_{ba}=5.3/\tau$ and $\omega-\omega_{ba}=10.7/\tau$, respectively. Evidently, the Doppler effect reduces and broadens the main dome-shaped maximum in these curves for both pulse shapes (55) and (56). Furthermore, the behavior of these curves shows how $X_2(\tau,\omega-\omega_{ba})$ changes with the shape of the pulse.

Figure 2 shows the dependence of the universal function (57) on time t' for a fixed resonant detuning $\omega-\omega_{ba}$ corresponding to the maximum of $X_2(\tau,\omega-\omega_{ba})$ as an odd function of $\omega-\omega_{ba}$ for a rectangular pulse (55) with $\varepsilon=0$ and a Gaussian pulse (56) with $\tau=8\tau_p$ and $\varepsilon=16$. These fixed values of $\omega-\omega_{ba}$ have been taken from Fig. 1. For both pulses (55) and (56) the Doppler effect significantly retards the rise in $X_2(t',\omega-\omega_{ba})$ with the passage of time. In addition, the shape of the curves depends strongly on the pulse shape.

For experimental studies of the dependence of the optically induced magnetization $\boldsymbol{\mu}_q(t')$ with $q=0,2$ on the resonance detuning $\omega-\omega_{ba}$ with the same initial conditions (22) or (34), it is necessary to pass a series of light pulses (14) or (33) with fixed polarization and amplitude, but different frequencies ω in the neighborhood of the transition frequency ω_{ba}, sequentially through the test gas. Here the light pulses must satisfy the requirement of perturbation theory that the preceding pulse not change the initial conditions (22) or (34) for the next pulse in the chosen approximation. Then $\boldsymbol{\mu}_q(t')$ with $q=0,2$ for each subsequently passing light pulse will be given by the same formula (47) for $0\leqslant t'\leqslant\tau$ and (48) for $\tau\leqslant t'\leqslant\infty$, where the initial time $t'=0$ is associated with the arrival of this pulse at a point $\mathbf{r}$ inside the gas volume. In addition, the light pulses must be separated by a time longer than γ^{-1} so that the preceding pulse does not distort $\boldsymbol{\mu}_q(t')$ with $q=0,2$ for the next pulse. In this arrangement for scanning the frequency ω, the preceding pulse creates a residual constant optically induced magnetization, as implied by Eq. (48). However, the latter does not contribute to the electromotive force (51) as a function of t' for the next pulse, so it does not show up in experiments on the optically induced magnetization owing to an individual sequentially passing pulse (see Eqs. (51)–(54)).

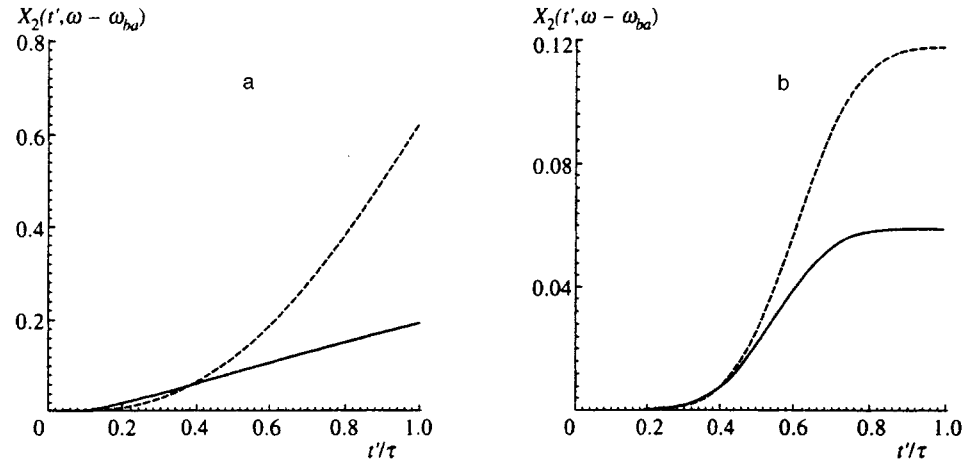


FIG. 2. The universal function $X_2(t',\omega-\omega_{ba})$ (57) as a function of the time $t'=\tau$ for fixed resonant detunings $\omega-\omega_{ba}$ corresponding to the maximum of $X_2(\tau,\omega-\omega_{ba})$ as the odd function of $\omega-\omega_{ba}$ shown in Fig. 1. The physical significance of the smooth and dashed curves is as in Fig. 1. (a) $\varepsilon=0$ for a rectangular pulse (55); (b) $\varepsilon=16$ for a gaussian pulse with $\tau=8\tau_p$ (56).

7. DISCUSSION

The above analysis yields the following general behavior. If the magnetic moment $\boldsymbol{\mu}_{at}(t')$ induced in a stationary atom by a resonant light pulse is proportional to the axial vector (20), then it is an even function of the resonant detuning $\omega-\omega_{ba}$ with a maximum at $\omega=\omega_{ba}$. If the magnetic moment $\boldsymbol{\mu}_{at}(t')$ is proportional to an axial vector which is invariant under the transformation $\mathbf{k}\rightarrow-\mathbf{k}$ of the wave vector, then $\boldsymbol{\mu}_{at}(t')$ is an odd function of $\omega-\omega_{ba}$ with a maximum at $0<|\omega-\omega_{ba}|$. This behavior is of a fundamental character, since it is a consequence of the symmetry of an atom in the field of a resonant light pulse (14) or (33), as well as symmetry with respect to time reversal (23) including initial conditions of the form (22) or (34) at $t'=0$. This behavior represents an effect of time reversal during the magnetization of an atom by a resonant light pulse.[12] In a number of cases this effect can be used to determine the vector properties of the induced magnetic moment $\boldsymbol{\mu}_{at}(t')$ and its dependence on t' and $\omega-\omega_{ba}$ to within a common factor that has the dimensions of a magnetic moment, without detailed perturbation theory calculations in the absence of relaxation at times $0\leqslant t'\leqslant\tau$, where the light pulse duration τ obeys Eq. (21). In the following interval $\tau<t'\leqslant\infty$, where relaxation takes place by spontaneous decay of the excited state of the atom after passage of the light pulse, the vector properties of the induced magnetic moment $\boldsymbol{\mu}_{at}(t')$ and its characteristic dependence on $\omega-\omega_{ba}$ are preserved and are a consequence of its behavior in the previous interval $0\leqslant t'\leqslant\tau$ where relaxation did not occur. In the interval $\tau\leqslant t'\leqslant\infty$, $\boldsymbol{\mu}_{at}(t')$ contains two uncomputed common factors with the dimensions of a magnetic moment.

If these results are applied to a gas of identical atoms within a fixed volume, we find that with time reversal (23) the optically induced magnetization of this gas by elliptically (14) or linearly (33) polarized pulses has the same properties as the magnetic moments of individual atoms, but with a more complicated functional dependence on $\omega-\omega_{ba}$ for both even and odd functions. This means that the effect of time reversal on the magnetization of atoms by a resonant light pulse also shows up in the optically induced magnetization of a gas. It is noteworthy that the uncomputed common factors with the dimensions of a magnetic moment density have no effect on experimental studies of the vector properties of the optically induced magnetization or its dependence on $\omega-\omega_{ba}$, τ, γ, T, and m_{at} if the electromotive force on the pickup coil (51) is measured in relative units.

In order to emphasize the fundamental nature of this effect of time reversal during the magnetization of atoms, let us consider an atomic gas in the field of an ultrashort resonant light pulse (14) propagating parallel to a constant magnetic field $\mathbf{H}$. An atom in the fields (14) and $\mathbf{H}$ without relaxation is symmetric reversal with respect to time (23) with the simultaneous transformation $\mathbf{H}\rightarrow-\mathbf{H}$.[7] In this case,[13] the optically induced magnetization $\boldsymbol{\mu}(t',\mathbf{H})$ of the gas for an isotropic initial state (22) breaks up into two terms:

$$\boldsymbol{\mu}(t',\mathbf{H})=\boldsymbol{\mu}_1(t',\mathbf{H})+\boldsymbol{\mu}_2(t',\mathbf{H}).$$

Here the first term $\boldsymbol{\mu}_1(t',\mathbf{H})$ is proportional to the axial vector (20), while the second term $\boldsymbol{\mu}_2(t',\mathbf{H})$ is collinear with $\mathbf{H}$. Under time reversal (23) with the simultaneous transformation $\mathbf{H}\rightarrow-\mathbf{H}$ in the absence of relaxation we obtain[13] $\boldsymbol{\mu}(-t',-\mathbf{H})=-\boldsymbol{\mu}(t',\mathbf{H})$, where the first term $\boldsymbol{\mu}_1(t',\mathbf{H})$ with the axial vector (20) is an even function of the resonance detuning $\omega-\omega_{ba}$. As opposed to this, the second term $\boldsymbol{\mu}_2(t',\mathbf{H})$ is invariant with respect to the wave vector substitution $\mathbf{k}\rightarrow-\mathbf{k}$ and is an odd function of $\omega-\omega_{ba}$. Thus, the effect of time reversal during magnetization of gas atoms in a resonant light pulse found here also shows up in a constant magnetic field.

[1] A. M. Badalyan, A. A. Dabagyan, M. E. Movsesyan *et al.*, Izv. Akad. Nauk SSSR, Ser. Fiz. **43**, 304 (1979).
[2] A. A. Dabagyan, M. E. Movsesyan, and R. E. Movsesyan, JETP Lett. **29**, 534 (1979).
[3] A. I. Alekseev and N. A. Korotkova, Zh. Éksp. Teor. Fiz. **108**, 118 (1995) [JETP **81**, 63 (1995)].
[4] J. P. van der Ziel and N. Bloembergen, Phys. Rev. A **138**, 1287 (1965).
[5] V. A. Zuikov, I. S. Bibkov, S. B. Karamyshev, and V. V. Samartsev, Laser Phys. **5**, 147 (1995).
[6] R. G. Usmanov and E. P. Khaĭmovich, Opt. i Spektr. **79** 378 (1995) [Opt. Spectrosc. **79**, 348 (1995)].
[7] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, 3rd ed. (Pergamon, Oxford, 1977).
[8] V. B. Berestetskii, E. M. Lifshitz, and L. P. Pitaevskii, *Relativistic Quantum Theory*, Pergmon Press, Oxford (1971).
[9] I. I. Sobel'man, *Introduction to the Theory of Atomic Spectra* (Pergamon, Oxford, 1972).
[10] A. I. Alekseev, Opt. i Spektr. **75**, 842 (1993) [Opt. Spectrosc. **75**, 499 (1993)].
[11] A. I. Alekseev, Zh. Éksp. Teor. Fiz. **111**, 63 (1997) [JETP **84**, 35 (1997)].
[12] A. I. Alekseev, JETP Lett. **65**, 242 (1997).
[13] A. I. Alekseev, Zh. Éksp. Teor. Fiz. **104**, 2954 (1993) [JETP **77**, 371 (1993)].

Translated by D. H. McNeill

Ionization of a two-electron atom in a strong electromagnetic field

O. V. Ovodova, A. M. Popov, and O. V. Tikhonova

D. V. Skobeltsyn Research Institute for Nuclear Physics, M. V. Lomonosov Moscow State University, 119899 Moscow, Russia

(Submitted 7 February 1997)
Zh. Éksp. Teor. Fiz. **112**, 470–482 (August 1997)

A one-dimensional model of a helium atom in an intense field of a femtosecond electromagnetic pulse has been constructed using the Hartree technique. ''Exact'' calculations have been compared to the approximations of ''frozen'' and ''passive'' electrons. A nonmonotonic dependence of the single-electron ionization probability on the radiation intensity has been detected. Minima in the ionization probability are due to multiphoton resonances between different atomic states due to the dynamic Stark effect. We suggest that the ionization suppression is due to the interference stabilization in this case. © *1997 American Institute of Physics.* [S1063-7761(97)00608-2]

1. INTRODUCTION

Dynamics of ionization of atomic and molecular systems in intense electromagnetic fields of a femtosecond duration has been discussed in numerous publications.[1–4] By the present time, processes of over-threshold ionization, mechanisms of multiphoton, tunneling, and over-barrier ionization, generation of multiple-charge ions in laser field, Coulomb explosion of molecules and other effects have been studied. In theoretical studies of interaction between an atomic system and electromagnetic field, numerical integration techniques have become more important in recent years in the context of the problem of a quantum system in a field of an electromagnetic wave.[4] Significant technical difficulties, however, compel the researchers to limit their studies to simpler one-dimensional models. Although a one-dimensional approximation does not allow investigations of some physical phenomena, such as angular distribution of electrons due to photoionization, effects related to polarization of electromagnetic field, etc., it helps, nonetheless, in understanding the dynamics of a real atomic system under an intense electromagnetic field. Qualitative agreement between calculations of ionization of one- and three-dimensional hydrogen atoms was demonstrated by Su *et al.*[5] and Kulander *et al.*[6] Similar data for a negative hydrogen ion H^- were obtained by Popov *et al.*[7] A three-dimensional ionization model of a multielectron atom (xenon) in an approximation in which all electrons but one are frozen was studied by Kulander.[8] Note that calculations on the modern Cray-T3D supercomputer yielded exact solutions to the five-dimensional Schrödinger equation describing a helium atom in an electromagnetic field,[9] but in this case processing and interpretation of calculations present an another enormous problem.

Therefore, numerical models using various approximations should be applied to interaction between electromagnetic radiation and real atomic systems in the limit of intense field in order to take into account all electrons in an atom, their interaction with each other and with external electromagnetic field during a laser pulse. Considerable interest has been aroused recently by calculations of the dynamics of two-electron, one-dimensional atomic systems obtained in various approximations.[10–12] The most consistent approach to the analysis of such systems is numerical integration of the two-particle Schrödinger equation on a two-dimensional grid.[10,11]

An accurate solution of such a problem would also require a lot of CPU time, therefore analysis of a two-electron system based on time-dependent Hartree–Fock equations[12] is practicable. As is known, in the Hartree (Hartree–Fock) technique the full wave function $\Psi(x_1,x_2,t)$ of the system is expressed in the form of a product of two one-particle wave functions $\psi_1(x_1,t)$ and $\psi_2(x_2,t)$ of both electrons (in the Hartree–Fock technique the full wave function should be also made symmetric or antisymmetric with respect to electron exchange). The motion of each electron is determined by the electrostatic potential of the nucleus and an averaged potential generated by the other electron. It is noteworthy that the quantitative difference between Hartree and Hartree–Fock calculations is probably negligible, since the values of exchange integrals due to the antisymmetrization of the electron wave function with respect to exchange of electrons is small in comparison to the energy of the direct electrostatic interaction.

This paper reports on a study of the dynamics of a one-dimensional helium atom based on the system of time-dependent Hartree equations in the range of radiation intensity $P=10^{-4}-0.3$ (hereinafter we use the atomic system of units), which corresponds to conditions of one-electron ionization. Our data are compared to calculations in approximations of ''passive'' and ''frozen'' electrons. We have obtained a nonmonotonic dependence of the one-electron ionization probability, and causes of this nonmonotonic behavior are discussed in the paper.

2. ONE-DIMENSIONAL MODEL OF HELIUM ATOM

We have selected model potentials of interaction between the electrons and the nucleus and between electrons in the form similar to that suggested by Grobe and Eberly:[10]

$$V(x_i)=-\frac{Z}{\sqrt{\alpha^2+x_i^2}}, \quad i=1,2, \tag{1}$$

 1063-7761/97/080257-07$10.00

TABLE I. Energy levels of one-dimensional helium atom.

State	Hartree calculations (this work) ε_1	ε_2	$\varepsilon=\varepsilon_1+\varepsilon_2$	Exact calculations[11] ε	Hyperspherical function approximation[11] ε
(1,1)	−1.1035	−1.1035	−2.207	−2.238	−2.185
(1,2)	−1.217	−0.500	−1.717	−1.704	−1.725
(1,3)	−1.319	−0.284	−1.603	−1.626	−1.629
(1,4)	−1.369	−0.184	−1.553	−1.567	−1.570
(1,5)	−1.402	−0.125	−1.527	−1.545	−1.542
(1,15)	−1.466	−0.017	−1.483	-	-
(2,2)	−0.512	−0.512	−1.024	−1.045	-
(2,3)	−0.599	−0.267	−0.866	-	-
(2,4)	−0.663	−0.179	−0.842	-	-

$$V_{12}=\frac{1}{\sqrt{\alpha^2+(x_1-x_2)^2}}, \tag{2}$$

where $Z=2$ is the nucleus charge, x_1 and x_2 are the coordinates of the first and second electrons, α is the cut-off parameter (in what follows, we assume that $\alpha=a_0=1$, where a_0 is the Bohr radius).

Thus, stationary atomic states can be determined by solving the eigenvalue problem

$$H\Phi(x_1,x_2)=\varepsilon\Phi(x_1,x_2), \tag{3}$$

where

$$H=H_1+H_2+V_{12}, \tag{4}$$

$$H_i=-\frac{1}{2}\frac{\partial^2}{\partial x_i^2}+V(x_i), \quad i=1,2, \tag{5}$$

and ε is the energy of the stationary state.

Good approximations to an exact solution of Eq. (3) are solutions based on the Hartree or Hartree–Fock methods. In the Hartree method, the solution of Eq. (3) is expressed in the form of a product of one-particle wave functions:

$$\Phi(x_1,x_2)=\varphi_1(x_1)\varphi_2(x_2), \tag{6}$$

where functions φ_1 and φ_2 satisfy the one-particle Schrödinger equation with a potential formed by the nucleus and delocalized second electron:

$$(H_i+\delta V_i)\varphi_i(x_i)=\varepsilon_i\varphi_i(x_i), \quad i=1,2, \tag{7}$$

where

$$\delta V_{12}=\int \frac{|\varphi_{2,1}(x)|^2dx}{\sqrt{\alpha^2+(x_{2,1}-x)^2}}, \tag{8}$$

ε_i is the ith electron energy, and the total energy is $\varepsilon=\varepsilon_1+\varepsilon_2$.

Table I lists electron energies of several stationary states $(n,m)\equiv(\varphi_n^{(m)}(x_1),\ \varphi_m^{(n)}(x_2))$ obtained by solving equation system (7). Here $\varphi_n^{(m)}(x)$ is the wave function of the nth stationary state of one of the electrons, given that the second electron is on the level m. It is important to note that functions $\varphi_n^{(m)}(x)$ and $\varphi_n^{(k)}(x)$ $(k\neq m)$ are different, although they are very similar for all large m and k larger than unity. The difference between these two functions is due to the differences in configurations of electrostatic potentials formed by an electron is different states. In our calculations, it proved important only in the state (1,1).

Note that the spectrum of one-electron excited states $(1,n)$, where $n\geqslant 2$, is similar to the spectrum of such states in a real He atom; moreover, all states with both electrons in excited states are auto-ionization states. Table I also lists energies which correspond to exact solutions of Eq. (3) and calculations based on the hyperspherical approximation.[11] It is clear that our results are in good agreement with exact calculations. As an example, Fig. 1 shows wave functions of the ground (1,1) and excited (1,3) states.

3. HELIUM ATOM IN THE FIELD OF AN ELECTROMAGNETIC WAVE

In the dipole approximation, the dynamics of a two-electron system in the field of an electromagnetic wave is described by the equation

$$i\frac{\partial}{\partial t}\Psi(x_1,x_2,t)=(H_1+H_2+V_{12})\Psi(x_1,x_2,t) -(x_1+x_2)E(t)\Psi(x_1,x_2,t), \tag{9}$$

where $E(t)$ is the wave electric field, and operators H_1, H_2, and V_{12} are given by Eqs. (4) and (2).

In the Hartree approximation, the wave function $\Psi(x_1,x_2,t)$ determined by time-dependent equation (9) can be expressed in the form of a product of time-dependent one-electron wave functions:

$$\Psi(x_1,x_2,t)=u(x_1,t)v(x_2,t). \tag{10}$$

By using the variational technique for the time-dependent Schrödinger equation,

$$\delta J=\left\langle \delta\Psi\left| i\frac{\partial}{\partial t}-H_1-H_2-V_{12}+(x_1+x_2)E\right|\Psi\right\rangle + \left\langle \Psi\left| i\frac{\partial}{\partial t}-H_1-H_2-V_{12}+(x_1+x_2)E\right|\delta\Psi\right\rangle=0,$$

we obtain an equation system for one-electron wave functions u and v:

$$i\frac{\partial u}{\partial t}=(H_1-x_1E)u+\langle v|V_{12}|v\rangle u+u\beta_1(t),$$

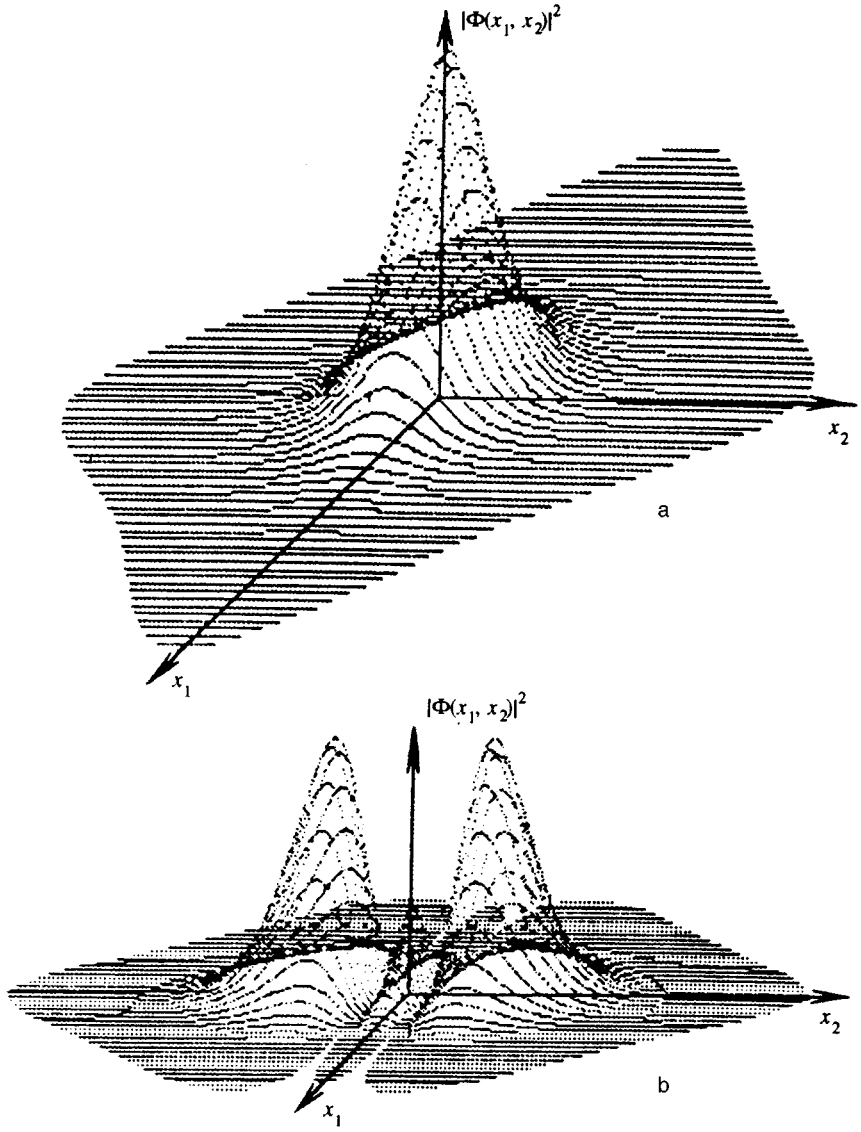


FIG. 1. Probability density distribution $|\Phi(x_1,x_2)|^2$ (a) in the ground and (b) excited (1,3) states of He atom calculated by the Hartree technique.

$$i\frac{\partial v}{\partial t}=(H_2-x_2E)v+\langle u|V_{12}|u\rangle v+v\beta_2(t), \tag{11}$$

where β_1 and β_2 are functions of only time and are independent of coordinates, therefore they modify only phases of the one-electron wave functions. Using the substitution

$$u(x_1,t)=\psi_1(x_1,t)\exp\left[-i\int\beta_1(t)dt\right],$$

$$v(x_2,t)=\psi_2(x_2,t)\exp\left[-i\int\beta_2(t)dt\right],$$

we obtain an equation system for new one-electron functions ψ_1 and ψ_2, which describe motion of each electron in a self-consistent field generated by the nucleus and the other electron. With due account of the electromagnetic wave field, these equations have the form

$$i\frac{\partial}{\partial t}\psi_1(x_1,t)=[H_1+\delta V_1(\psi_2,x_1,t)]\psi_1(x_1,t)$$

$$-x_1E(t)\psi_1(x_1,t),$$

$$i\frac{\partial}{\partial t}\psi_2(x_1,t)=[H_2+\delta V_2(\psi_1,x_2,t)]\psi_2(x_2,t)$$

$$-x_2E(t)\psi_2(x_2,t), \tag{12}$$

where $H_{1,2}$ and $\delta V_{1,2}$ are given by Eqs. (5) and (8), respectively, and operator $-x_iE(t)$ describes interaction between the ith electron and field of the electromagnetic wave.

It is noteworthy that the difference between the original (u,v) and new (ψ_1,ψ_2) one-particle wave functions is only in the phase factor, whose presence or absence does not affect calculations of populations of different states and the ionization probability, which is the ultimate aim of our study.

We assumed that at the initial moment the system was in one of its stationary states, which can be also calculated using the Hartree technique.

Calculations were performed for the radiation frequency $\omega=0.18$, which corresponds to radiation of the KrF excimer laser, in the intensity range $P=E^2=10^{-4}-0.3$, i.e., $3\times10^{12}-10^{16}$ W/cm^2. The electric field was described by the following functions of time:

$$E(t)=\begin{cases} E_0(t/\tau)\cos(\omega t), & t\leqslant\tau=2T, \\ E_0\cos(\omega t), & 2T\leqslant t\leqslant 12T, \\ E_0(1-t/\tau)\cos(\omega t), & 12T\leqslant t\leqslant 14T, \end{cases}$$

where $T=2\pi/\omega$ is the optical cycle duration.

In some cases, in addition to trapezoid pulses, the effect of Gaussian pulses was investigated:

$$E(t)=E_0\cos(\omega t)\exp\left[-\frac{1}{2}\left(\frac{t-t_0}{\tau}\right)^2\right],$$

where t_0 is the moment of the intensity peak.

Calculations were performed on a uniform spatial grid of dimension $150a_0$, with 6000 nodes for each electron, by the technique described in Koonin's monograph.[13] The step of integration over time was $0.01T$. In calculating potentials driving each electron, the charge distribution due to the other electron was taken from the previous time step. In the ''frozen'' electron approximation, we assumed that the wave function of one of the electrons was identical to the wave function of the initial state except for a phase factor, so that the electrostatic potential generated by this electron was constant with time.

In interpreting the results, the helium atom wave function

$$\Psi(x_1,x_2,t)=\psi_1(x_1,t)\psi_2(x_2,t)$$

was expanded in terms of stationary states of the initial atomic Hamiltonian ($\varphi_n^{(m)}(x_1)\varphi_m^{(n)}(x_2)$):

$$\Psi(x_1,x_2,t)=\sum_{n,m} C_{nm}(t)\varphi_n^{(m)}(x_1)\varphi_m^{(n)}(x_2) \times\exp[-i(\varepsilon_n+\varepsilon_m)t]. \quad (13)$$

From the formal viewpoint, expansion (13) is valid at an arbitrary moment of time t. But when the laser pulse is on, coefficients C_{nm} have the sense of populations of atomic states as long as wave functions of stationary states are not distorted by the electromagnetic wave field. In calculating coefficients C_{nm}, we found that in the intensity range of 10^{-4} to 0.3 only states with one excited electron are populated, and the corresponding excitation probabilities are determined by the expression

$$W_{1n}=|C_{1n}|^2 =\left|\int \Psi^*(x_1,x_2,t)\varphi_1^{(n)}(x_1)\varphi_n^{(1)}(x_2)dx_1dx_2\right|^2. \quad (14)$$

The probability of detecting the system in states with both electrons excited proved to be negligible. This property is due to the high excitation threshold of such states, which is considerably higher than the threshold of the one-electron continuum.

In our calculations, we took into account $N=15$ different states (1,1)–(1,15), since in the discussed range of intensities the probabilities of the higher states population are negligible. The total ionization probability was calculated by the formula

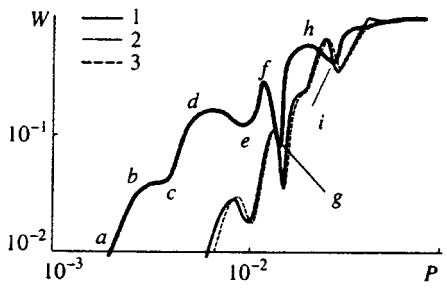


FIG. 2. Ionization probability of a He atom initially in the ground state as a function of the radiation intensity: (*1*) ''exact'' calculation, and (*2*) ''frozen'' and (*3*) passive electron approximations.

$$W_I=1-\sum_{n=1}^{N}|C_{1n}|^2. \quad (15)$$

4. RESULTS OF COMPUTER SIMULATIONS

4.1. Applicability of the ''frozen'' electron approximation

Figure 2 shows calculations of helium atom ionization as a function of radiation intensity $P=E^2$ for the case when atom was in the ground state (1,1) before the laser pulse. Curve *1* corresponds to a solution of full equation system (12) and curve *2* to the ''frozen'' electron approximation, in which the wave function of one electron corresponds to the initial stationary state. In this case, the dynamics of the second (active) electron is determined by the potential due to the nucleus, static potential formed by the ''frozen'' electron, and the field of the electromagnetic wave. The graph shows that the ''frozen'' electron approximation underestimates the ionization probability because in the ''frozen'' electron model the action of the electromagnetic wave on one electron is ignored, as a result, the average separation between the electrons is larger and the Coulomb energy smaller than in the full model. This assertion is confirmed by the calculation in which the electric field was ''switched off'' in one of Eq. (12) (model of a passive electron). In this case, the wave function of the passive electron varied with the electric potential formed by the active electron. Corresponding calculations are shown by curve *3* in Fig. 2. One can see that the models of passive and ''frozen'' electron yield almost identical curves of $W_I(P)$.

To conclude this section, note that state (1,1) is the only one in which the ''frozen'' electron approximation yields unsatisfactory results. It turned out that in the cases of states (1,2) and (1,3) the error in calculations of ionization rates in the ''frozen'' electron model is within $\sim 1\%$ throughout the entire studied range of radiation intensity.

4.2. Nonmonotonic intensity dependence of ionization probability

An important feature of the resulting $W_I(P)$ function is its nonmonotonic behavior in the intensity range far from saturation. Similar maxima and minima on the $W_I(P)$ curve also occur in the case of other initial atom states (Fig. 3). Note that such nonmonotonic curves were also obtained in experiments.[14–18] These features were attributed[14–17] to mul-

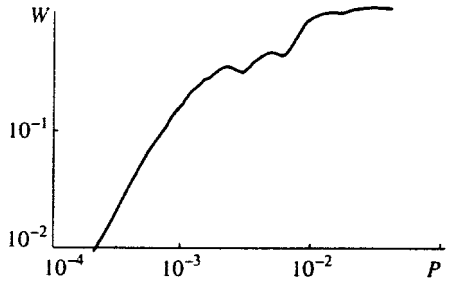


FIG. 3. Ionization probability of a helium atom initially in state (1,2) versus the radiation intensity.

tiphoton resonance in intermediate states due to dynamic Stark effect caused by the laser field. In sufficiently long (nanosecond) laser pulses with relatively low radiation intensities, $P=10^9-10^{11}$ W/cm^2, these resonances result in an abrupt increase in the rate of population of excited states. which considerably accelerates the ionization process owing to a lower photon multiplicity of transitions between the discrete spectrum and continuum.

In studying ionization of atoms of noble gases (Ar, Kr, and Xe) by femtosecond pulses generated by a titanium–sapphire laser at a radiation intensity $P=10^{13}-10^{15}$ W/cm^2, Chin *et al.*[18] detected a decrease in the ion yield in certain ranges of the radiation intensity in comparison with the nonresonant function $W_I \propto P^N$, where $N=[I/\omega]+1$ is the multiplicity of the multiphoton absorption and I is the ionization potential. This decrease in the ionization rate was attributed[18] to dynamic resonances in the process of laser action and suppression of ionization of excited atoms. For example, in multiphoton ionization of Xe atoms, the dynamic Stark effect can cause resonant population of the $6f$ state, which includes seven degenerate levels stable against the ionization through the continuum. Given this mechanism of the ionization suppression, one can expect a larger population of the $6f$ state. Unfortunately, the population of excited states as a function of radiation intensity was not studied in experiments.[18]

We assume that the nonmonotonic curves of the ionization probability as a function of the laser field intensity obtained in our calculation and in Ref. 18 can be interpreted in terms of one mechanism based on the interference stabilization, which was discussed in detail by Dubrovskiĭ *et al.*[19]

The results of Ref. 19 can be generalized to the case of multiphoton ionization of an atom in an intense field if there is a resonance in one of the Rydberg states. The ionization probability as a function of frequency is characterized in this case by sharp minima at frequencies close to the resonant transition between the ground state and one of the Rydberg states modified by an intense laser field. The action by the intense field on higher states induces transitions of the λ-type and leads to interference stabilization of Rydberg states.[20]

The presence of stable Rydberg levels modified by a strong field is the cause of an abrupt drop in the ionization probability in the case of a resonant transition to one of such states. A similar situation probably obtains in our computer simulations. For example, on sections ab, cd, and ef of the curves in Fig. 2, the ionization probability depends on the intensity as $W \propto P^5$, which corresponds to the multipliity of

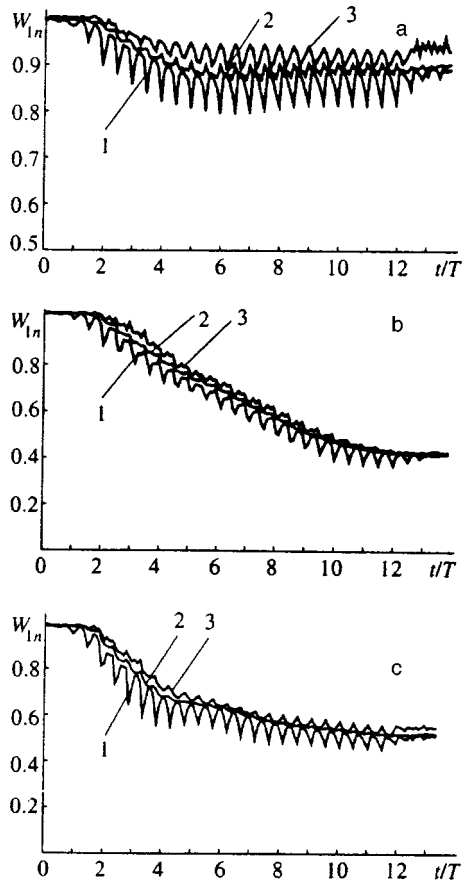


FIG. 4. Population dynamics of various states of the discrete spectrum under laser radiation at intensities of (a) 0.0142, (b) 0.0196, and (c) 0.0264: (*1*) population of state (1,1); (*2*) sum of populations of states (1,1) and (1,2); (*3*) ionization probability.

the multiphoton absorption. Deviations from this function on section bc, de, and fg means that the ionization rate is lower at some values of the field intensity. As will be shown below, each minimum in the ionization probability is determined by the condition of resonance with a high-energy one-electron state, with due account of the dynamic Stark shift. Note that the mechanism proposed by Dubrovskiĭ *et al.*[19] is feasible in sufficiently strong fields whose amplitude satisfies the condition

$$E_0/\omega^{5/3} \sim 1. \tag{16}$$

In our case, $\omega=0.18$, and fields with $E_0 \geqslant 0.01$ satisfy criterion (16) of a strong electromagnetic field. Another argument in favor of the interference stabilization[19] can be the two-exponential character of the population of atomic states as a function of time at laser field intensities corresponding to minima on the ionization curve. The dynamics of population in various states when the laser pulse is on for different intensities is shown in Fig. 4. The curves in Figs. 4a and 4c correspond to minima labeled by e and g in Fig. 2, and Fig. 4b to maximum f in Fig. 2.

The data plotted in Fig. 4 for intensities corresponding to minima in Fig. 2 indicate that ionization takes place only during the first four cycles of the optical field, and subsequently it almost stops, whereas at the radiation intensity

TABLE II. Matrix elements of the dipole moment operator in one-dimensional helium atom.

n	$m=1$	$m=2$	$m=3$
2	0.866	-	-
3	0	−1.962	-
4	0.131	0	3.508
5	0	−0.434	0
6	0.0669	0	0.539
7	0	−0.220	0
8	0.0412	0	0.268
9	0	−0.139	0
10	0.0265	0	0.147

corresponding to point f in Fig. 2, ionization continues at an approximately constant rate throughout the laser pulse. In our opinion, this feature of the ionization dynamics is an indication in favor of the interference mechanism of the ionization suppression at field intensities corresponding to points c, e, g, and i. Moreover, at the intensities corresponding to these points in Fig. 2, states (1,3)–(1,5) have the highest population among the high-energy states. In all cases, the population of these states reaches 2–4%, and is several times the population of other excited states $(1,n)$ $(n\geqslant 6)$. These population numbers are achieved in about four to five optical cycles after the laser pulse onset. This allows us to suppose that it is the population of these states that leads to the atom stabilization at intensities corresponding to points c, e, g, and i in Fig. 2.

A similar situation takes place for other initial states, in particular (1,2). This state is ionized by two-photon absorption, which corresponds to the initial part of the curve of $W_I(P)$ in Fig. 3. At field intensities $\geqslant 2\times 10^{-3}$ the ionization is also slower, and nonmonotonic sections occur on the curve of $W_I(P)$.

4.3. Perturbation theory estimates of minima positions on the curve

In order to verify the hypothesis about the stabilizing effect of nonmonotonic sections on the curves of $W_I(P)$, we checked estimates of Stark shifts of different states. It is known that the shift of the nth stationary state due to high-frequency electric field in the lowest order of perturbation theory is given by the formula[21]

$$\Delta\varepsilon_n=\frac{1}{4}\sum_{m\neq n}|d_{nm}E|^2\left(\frac{1}{\omega_{nm}-\omega}+\frac{1}{\omega_{nm}+\omega}\right)$$
$$=-\frac{1}{4}\alpha_n E^2, \tag{17}$$

where d_{nm} is the matrix element of the dipole moment, ω_{nm} is the frequency of the transition between states with indices n and m, and α_n is the atom polarizability.

Taking into account in Eq. (17) only excited states $\varphi_n^{(1)}(x_1)\varphi_1^{(n)}(x_2)$ and the similarity between $\varphi_k^{(n)}(x_1)$ and $\varphi_m^{(n)}(x_2)$ $(k,m>1)$, we obtain

TABLE III. Polarizability of stationary states of one-dimensional helium atom for $\omega=0.18$.

n	1	2	3	4	5
α	3.616	−27.43	0.1432	−7.768	−31.42

$$d_{nm}=\int \varphi_n^{(1)}(x_1)\varphi_1^{(n)}(x_2)(x_1+x_2)$$
$$\times\varphi_m^{(1)}(x_1)\varphi_1^{(m)}(x_2)dx_1dx_2$$
$$\simeq\int \varphi_n^{(1)}(x_1)x_1\varphi_m^{(1)}(x_1)dx_1. \tag{18}$$

Values of some matrix elements are listed in Table II, and polarizabilities α_n of several low-lying states in Table III. Calculations of polarizabilities can be more accurate if continuum states are included in Eq. (17). Using the data of Table III, we have calculated radiation intensities at which multiphoton resonances between states (1,1) or (1,2) and higher states occur. The inclusion of the lowest excited state (1,2) is due to its high population in intense laser field (Fig. 4).

Table IV lists positions of ionization probability minima on curves shown in Fig. 2, as well as pairs of resonant states which cause ionization rate minima and the corresponding resonance multiplicities. The table also lists detunings from the resonances calculated by the perturbation theory taking into account the Stark effect. For resonances corresponding to minima c, e, and g, the detuning is within 10% of the photon energy, and for resonance i the calculation error is larger and amounts to $\simeq 20\%$. Note that Table IV lists only dipole-allowed transitions. One can clearly see a good correlation between different data, i.e., minima on the curve of $W_I(P)$ can be put into correspondence with multiphoton resonances due to the dynamic Stark effect.

The data presented in this paper indicate that the physical cause of the ionization deceleration in this case is interference stabilization of coherently populated excited atomic states when one of them is in a multiphoton resonance with the ground (or lowest excited) state, in accordance with the mechanism suggested by Dubrovskiĭ *et al.*[19] The most interesting of them is resonance i $(P=0.0256)$, which is probably related to the overlap between levels (1,2) and (1,4). In this case, the ionization suppression mechanism suggested by Fedorov *et al.*[20] as an explanation of stabilization of Rydberg atoms is fully consistent with our description of the phenomenon, but, unlike Fedorov *et al.*, we have connected the states (1,2) and (1,4) with continuum via two-photon transitions.

TABLE IV. Radiation intensities at which multiphoton resonances occur.

Number	Intensity	Transition	Resonance multiplicity	Detuning
c	3.6(−3)	(1,1)→(1,5)	4	8.1(−4)
e	9.0(−3)	(1,1)→(1,2)	3	1.9(−2)
g	1.44(−2)	(1,2)→(1,5)	1	4.2(−3)
i	2.56(−2)	(1,2)→(1,4)	0	3.8(−2)

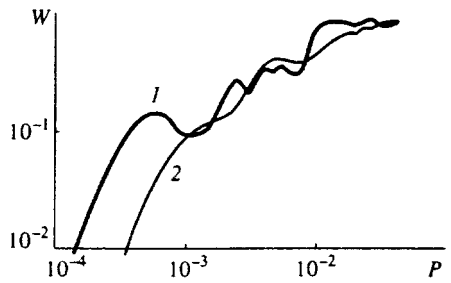


FIG. 5. Ionization probability of state (1,3) of a He atom under a laser pulse of (*1*) trapezoid and (*2*) Gaussian shapes vs laser field intensity.

Note that the resonance position in the intensity range $P \geqslant 0.1$ cannot be calculated using the perturbation theory by Eq. (17) because in this case the difference between shifts of atomic levels is larger than the initial separation between them. This circumstance, in particular, leads to a large detuning for resonance i.

Resonances also manifest themselves in the form of nonmonotonic sections on the curve of $W_I(P)$ in the case of a Gaussian pulse (Fig. 5). In this case, nonmonotonic sections are smoother in comparison to trapezoid pulses, and the curve has a wavy shape, like the experimental curve in Ref. 18. Note than in experiments[18] the curve is also smoothed owing to integration of the signal over the intensity profile in the focal region, which is also Gaussian. Therefore, we assume that the real shape of the curve of the ionization probability as a function of intensity in experiments[18] is nonmonotonic, and nonmonotonic features can be interpreted in terms of the approach proposed in Ref. 19.

5. CONCLUSIONS

On the basis of time-dependent Hartree equations, we have constructed the one-dimensional model of a two-electron atom in an intense electromagnetic field. In all states, except the ground state, the effect of the optical field on the inner electron can be neglected. For the ground state, the ''frozen'' electron model underestimates the ionization rate, especially at low intensities. The curve of the ionization probability as a function of power density is nonmonotonic because of resonances between different states of the discrete spectrum due to the dynamic Stark effect. We assume that minima on the curves of $W_I(P)$ can be interpreted in terms of the interference mechanism of atom stabilization.

This authors gratefully acknowledge discussions of the results of this work with M. V. Fedorov and his valuable remarks.

This work was supported by the Russian Fund for Fundamental Research (Project 96-02-19286).

[1] N. B. Delone and M. V. Fedorov, Usp. Fiz. Nauk **158**, 215 (1989) [Sov. Phys. Usp. **32**, 500 (1989)].
[2] M. V. Fedorov, *Electron in a Strong Optical Field* [in Russian], Nauka, Moscow (1991).
[3] N. B. Delone and V. P. Krainov, *Multiphoton Processes in Atoms*, Springer-Verlag, New York–Berlin (1993).
[4] E. A. Volkova and A. M. Popov, Laser Phys. **5**, 1065 (1995).
[5] Q. Su, J. H. Eberly, and J. Javanainen, Phys. Rev. Lett. **64**, 862 (1990).
[6] K. C. Kulander, K. J. Schafer, and J. L. Krause, Phys. Rev. Lett. **66**, 2601 (1991).
[7] A. M. Popov, O. V. Tikhonova, and E. A. Volkova, Laser Phys. **5**, 1029 (1995); **5**, 1184 (1995).
[8] K. C. Kulander, Phys. Rev. A **38**, 788 (1988).
[9] K. T. Taylor, J. S. Parker, D. Dundas, and E. Smith, in *Proc. VII Int. Conf. On Multiphoton Processes (ICOMP-VII), Abstracts*, Garmisch-Partenkirchen, Germany (1996); J. Parker, K. T. Taylor, C. W. Clark, and S. Blodgett-Ford, J. Phys. B **29**, L33 (1996).
[10] R. Grobe and J. H. Eberly, Phys. Rev. A **48**, 4664 (1993).
[11] A. I. Artemiev, R. Grobe, and J. H. Eberly, in *Proc. SILAP-IV*, H. G. Müller and M. V. Fedorov (eds.), Kluwer, Dordrecht (1996), p. 285.
[12] M. S. Pindzola, D. C. Griffin, and C. Bottcher, Phys. Rev. Lett. **66**, 2305 (1991); M. S. Pindzola, P. Gavras, and T. W. Gorczyca, Phys. Rev. A **51**, 3999 (1995).
[13] S. Koonin, *Computational Physics*, Addison-Wesley, Reading, MA (USA) (1986).
[14] D. T. Alimov and N. B. Delone, Zh. Éksp. Teor. Fiz. **70**, 29 (1976) [Sov. Phys. JETP **43**, 15 (1976)].
[15] J. Morellec, D. Normand, and G. Petite, Phys. Rev. A **14**, 300 (1976).
[16] G. Mainfray and C. Manus, Appl. Opt. **19**, 3934 (1980).
[17] N. B. Delone and M. V. Fedorov, Trudy FIAN **115**, 42 (1980).
[18] S. L. Chin, A. Talebpour, Y. Liang, S. Augst, and C. Y. Chien, in *Proc. SILAP-IV*, H. G. Müller and M. V. Fedorov (eds.), Kluwer, Dordrecht (1996); A. Talebpour, C. Y. Chien, and S. L. Chin, submitted to J. Phys. B (1997).
[19] Yu. V. Dubrovskiĭ, M. Yu. Ivanov, and M. V. Fedorov, Zh. Éksp. Teor. Fiz. **99**, 411 (1991) [Sov. Phys. JETP **72**, 228 (1991)].
[20] M. V. Fedorov and A. M. Movsesian, J. Phys. B **21**, L155 (1988); M. V. Fedorov, Laser Phys. **3**, 219 (1993).
[21] A. S. Davydov, *Quantum Mechanics* [in Russian], Nauka, Moscow (1973).

Translation was provided by the Russian Editorial office.

Theory of single-mode lasing in coherent quantum cascade lasers

V. F. Elesin

Moscow State Engineering Physics Institute, 115409 Moscow, Russia
(Submitted 21 February 1997)
Zh. Éksp. Teor. Fiz. **112**, 483–498 (August 1997)

A theory is developed for steady-state single-mode lasing in coherent quantum-well cascade lasers. This laser model is an example of a strictly quantum mechanical problem in which approximate kinetic approaches are not used to account for dissipative scattering processes. Exact wave functions are found for the system in weak and strong electromagnetic fields, so that the output power and frequency can be determined as functions of the coherent pump current and system parameters. It is shown that for pumping by monoenergetic electrons the power has a nonlinear (root) dependence and tends to saturate in strong fields. It is predicted that the coherent pumping efficiency may be increased by adjusting the energy of the pump electrons, which will lead to a linear power dependence, a high efficiency, and low threshold currents. A population inversion is found not be a necessary condition for lasing in the coherent laser. In particular, in the high field regime the population of the lower level exceeds that of the upper, while in the optimally adjusted regime they are the same. © *1997 American Institute of Physics.* [S1063-7761(97)00708-7]

1. INTRODUCTION

In 1971 Kazarinov and Suris[1] proposed a new type of semiconductor laser in which the radiative transitions take place between size-quantized levels (subbands). After almost a quarter century this proposal was realized in nanostructures where the principal elements are two quantum wells with working levels (subbands) in each[2] (''oblique transitions'') or a single quantum well with two working levels (''vertical transitions'').[3] Pumping to the upper working level takes place by resonant tunnelling.

These lasers, which are known as quantum cascade lasers, have some important advantages: the possibility of tuning the output from the infrared to the submillimeter range, a weak temperature dependence for the threshold current, etc.

Quantum cascade lasers have a number of important features, such as one type of charge (unipolarity), the same signs on the subband masses, elimination of the prohibition on inter-subband nonradiative transitions, etc. Another fundamental feature of these lasers is that, in general, the resonant tunnelling which provides the pumping is coherent.

Resonant tunnelling can be coherent or incoherent, depending on the structure parameters and the temperature. If the times $\tau_{\rm ph}$ for dissipative relaxation of the electrons which destroy the coherence are less than the reciprocal widths Γ_j^{-1} of the size-quantization levels, then incoherent tunnelling takes place.[4] In this case, the tunnelling is modelled by an external source which supplies electrons with a fixed power.

This kind of approach has been applied to quantum cascade lasers elsewhere.[1,5–7] It can be used to describe several of the features noted above. Thus, for example, a kinetic theory for the quantum cascade laser has been developed[6] that describes interactions of the electrons with optical phonons which lead, in particular, to a high threshold current. It has been shown that it is possible to choose a local energy overpopulation regime in which the threshold current can be made 1–2 orders of magnitude lower by using the reabsorption effect for the optical phonons.

At the same time, when the opposite inequality $\tau_{\rm ph}>\Gamma_j^{-1}$ holds, as may occur in quantum wells and wires and, especially, in quantum dots, it is necessary to take the coherence of resonant tunnelling into account. As a first step it is natural to consider the theory of quantum cascade lasers in general without considering scattering processes, i.e., the so-called coherent approximation. A laser of this type (for concreteness, a coherent laser) is of considerable interest. In fact, at first coherent resonant tunnelling can ensure more efficient pumping, since electrons accumulate in the well owing to interference effects. Second, a coherent laser is an interesting physical object in which, as we shall show below, lasing can take place without the involvement of dissipative processes. Processes of the latter sort are necessary to realize an emission event in laser theory for large-sized objects.[8,9] Thus, we might expect new effects and behavior to show up during operation of a coherent laser. It should also be noted that this type of laser is an open system which depends strongly on the boundary conditions and is in a current state.

The purpose of this paper is to develop a theory for steady-state lasing of a coherent laser and to calculate the power and frequency of the electromagnetic field functions of the coherent pump current and system parameters. We have examined a simple model which allows an analytic solution and can be used to present the results in an analytic form over a wide range of fields, from weak to strong. Lasing takes place in a single quantum well (dot) (see Fig. 1) with two working levels (subbands) having energies $\mathscr{E}_{2R}$ and $\mathscr{E}_{1R}$, the difference between which determines the frequency $\hbar\omega$ of the electromagnetic field. (For brevity, in the following we shall speak of a well and levels.) Electrons with energies $\mathscr{E}\simeq\mathscr{E}_{2R}$ enter the well from the left at a constant velocity, undergo a radiative transition to the level with energy $\mathscr{E}_{1R}$, and leave the well by tunnelling (or again undergo a transition to level 2).

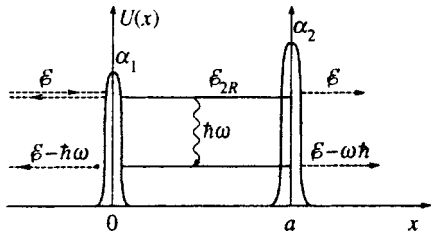


FIG. 1.

It is impossible to find the wave functions of an open system in an electromagnetic field using the expansion in eigenfunctions of the field-free Hamiltonian that is customary in laser theory,[9] because they are nonorthogonal. We have employed another approach which involves seeking stationary solutions of the Schrödinger equation which satisfy the appropriate boundary conditions. This method has been used previously for problems involving resonant tunnelling in alternating fields and for resonant-tunnelling diodes,[10–12] but, as a rule, in first-order perturbation theory with respect to the field. When applied to lasers it can be used only to find the gain coefficient at the lasing threshold. In this paper we find the wave functions without using perturbation theory and construct a theory of coherent lasers for weak and strong fields.

Our model differs substantially from the model used in a recent paper,[13] which appears to be the first devoted to the theory of the coherent laser. It dealt with a specific nonstationary model (referred to as a microlaser model by the author) in which an electron wave packet falls into a quantum well and emits a field quantum. The resulting numerical solutions make it possible to follow the emission process and are of some interest. They, however, cannot (as the author acknowledges) be used to find the power, frequency, and threshold current, i.e., all of the quantities which are needed to describe a real experiment.

This article is constructed as follows: in Sec. 2 the model is described and the basic equations are formulated. The wave functions and polarization currents are found in a general form in Sec. 3 and for the important limiting case of high barriers in Sec. 4. Section 5 is devoted to an analysis of lasing in a coherent laser with monoenergetic pumping and Sec. 6, to lasing in a coherent laser with a Fermi pump distribution.

In the following we shall examine the one-dimensional case: in the absence of scattering processes it is easy to generalize to include lateral motion. The shifts in the energy levels owing to buildup of charge in the well are assumed small compared to Γ_j and the temperature is assumed to be zero.

2. DESCRIPTION OF THE MODEL: BASIC EQUATIONS

In order to determine the basic behavior, we study the following model of a coherent quantum cascade laser. Figure 1 shows a one-dimensional quantum well with barriers at the points $x=0$ and $x=a$. The parameters of the well are chosen so that the two lower levels with energies $\mathscr{E}_{1R}$ and $\mathscr{E}_{2R}$ differ by an amount roughly equal to $\hbar$ times the frequency of the electromagnetic field; i.e., $\hbar\omega \simeq \mathscr{E}_{2R}-\mathscr{E}_{1R}=\hbar\omega_{21}$. A steady-state flux of electrons with a density proportional to q^2 and energy $\mathscr{E}$ roughly equal to $\mathscr{E}_{2R}$ is incident on the system from the left.

An electromagnetic field, which can be regarded with good accuracy as classical,

$$E_x(z,t)=E(t)\sin(kz)\cos(\omega t+\varphi(t)), \tag{1}$$

is emitted as the electrons undergo a transition from the upper level 2 to the lower level 1.

In accordance with experiment,[2,3] here we assume that the field is polarized perpendicular to the plane of the well (i.e., along the x axis), while the wave vector is directed in the plane (along the z axis). An optical cavity of length L isolates these modes. We restrict ourselves to considering single mode operation. The equations for the slowly varying field amplitude $E(t)$ and the phase $\varphi(t)$ have the form[8,9]

$$\frac{\partial E}{\partial t}=-\frac{E}{2\tau_0}-\frac{2\pi}{\kappa}J_c(k), \tag{2}$$

$$\frac{\partial\varphi}{\partial t}E+(\omega-\Omega)E=-\frac{2\pi}{\kappa}J_s(k), \tag{3}$$

and

$$J_{c,s}(k)=\int_0^a dx\ e^{ikx}J_{c,s}(x), \tag{4}$$

where $J_c(k)$ and $J_s(k)$ are the Fourier components of the polarization currents coincident in phase with the field (J_c) and shifted by $\pi/2$ relative to the field (J_s) which describe the interlevel transitions, τ_0 is the lifetime of a photon in the cavity, $\Omega=kc$ are the eigenfrequencies of the cavity, κ is the dielectric constant, and c is the speed of light. The currents $J_c(x)$ and $J_s(x)$ are expressed in terms of the wave function $\Psi(x,t)$ of the system which obeys the following Schrödinger equation:

$$i\frac{2m^*}{\hbar}\frac{\partial\Psi}{\partial t}=-\frac{\partial^2\Psi}{\partial x^2}+U(x)\Psi+\hat{V}(x,t)\Psi, \tag{5}$$

where m^* is the effective mass of the electron and

$$U(x)=\alpha_1\delta(x)+\alpha_2\delta(x-a),\quad \alpha_i=\frac{2m^*\widetilde{\alpha}_i}{\hbar^2},$$

is the potential energy of the barriers. The last term in Eq. (5),

$$\hat{V}(x,t)\Psi=\frac{2ei}{\hbar}A_x(t)\frac{\partial}{\partial x}\Psi$$

describes the interaction of the electrons with the electromagnetic field and $A_x(t)$ is the vector potential in the Coulomb gauge, which is nonzero in the well. This form of interaction is preferable for nonlocalized wave functions to Ex.[13] Expressing $A_x(t)$ in terms of the field amplitude $E(t)$, we rewrite the last term in the form

$$\hat{V}\Psi=V(e^{i\omega t}-e^{-i\omega t})\frac{\partial}{\partial x}\Psi,\quad V=-\frac{eE}{\omega}. \tag{6}$$

The system of Eqs. (2), (3), and (5) should be supplemented by the general expression for the current,

$$J(x,t)=-\frac{ie}{2m^*}\left[\Psi^*\frac{\partial\Psi(x,t)}{\partial x}-\text{c.c.}\right]$$

and boundary conditions for $\Psi(x,t)$:

$$\Psi(-0,t)=\Psi(+0,t),\quad \Psi(-a,t)=\Psi(+a,t),$$
$$\Psi'(+0,t)-\Psi'(-0,t)=\alpha_1\Psi(0,t),\quad \Psi'=\partial\Psi/\partial x,$$
$$\Psi'(+a,t)-\Psi'(-a,t)=\alpha_2\Psi(a,t). \tag{7}$$

The fluxes of electrons incident from the left and of reflected electrons will be taken into account below (Eq. (12)) after a specific form of $\Psi(x,t)$ has been chosen.

It should be noted that in Eq. (5) we have left out terms that are quadratic in $A(t)$. This approximation, which is usually used in laser theory,[9] is also valid here in terms of the parameter $V/p=eE/\omega p\ll 1$, where p is the electron momentum (see Eq. (21)).

3. WAVE FUNCTIONS AND POLARIZATION CURRENTS IN A COHERENT LASER

We seek a steady-state solution of Eq. (5) in the form of the series

$$\Psi(x,t)=\sum_n \exp\left[-it\left(\frac{\mathscr{E}}{\hbar}+n\omega\right)\right]\psi_n(x),\quad n=0,\pm 1,..., \tag{8}$$

where the functions $\psi_n(x)$ satisfy the system of equations

$$(\mathscr{E}+n\omega)\psi_n(x)+\psi_n''(x)=V[\psi_{n-1}'(x)-\psi_{n+1}'(x)], \tag{9}$$

where $\psi_n(x)$ describes the state with energy $\mathscr{E}+n\omega$. (Here $\mathscr{E}$ and ω are the energy and frequency multiplied by $2m^*$ and in the following we set $\hbar=c=1$.)

It is well known that during the lasing process the main resonant contribution is from two levels.[14] In our case these are the upper level with energies $\mathscr{E}_{2R}$ and the lower level with $\mathscr{E}_{1R}$, which correspond to the wave functions $\psi_0(x)$ and $\psi_{-1}(x)$. Thus, the wave function (8) reduces to two terms

$$\Psi(x,t)\simeq\psi_0(x)e^{-i\mathscr{E}t}+\psi_{-1}(x)e^{-it(\mathscr{E}-\omega)}, \tag{10}$$

where $\psi_0(x)$ and $\psi_{-1}(x)$ satisfy the system of equations

$$\mathscr{E}\psi_0(x)+\psi_0''(x)=V\psi_{-1}',$$
$$(\mathscr{E}-\omega)\psi_{-1}(x)+\psi_{-1}''(x)=-V\psi_0' \tag{11}$$

with the boundary conditions

$$\psi_n(0)\left(1-\frac{\alpha_1}{ip_n}\right)+\frac{\psi_n'(0)}{ip_n}=2q\,\delta_{n0},$$
$$\psi_n(a)\left(1-\frac{\alpha_2}{ip_n}\right)-\frac{\psi_n'(a)}{ip_n}=0,\quad p_n=\sqrt{\mathscr{E}+n\omega}. \tag{12}$$

These last are obtained from Eq. (7) on substituting Eq. (10) and taking into account the constant flux of electrons, incident from the left with energy $\mathscr{E}$, which is proportional to q, and the loss of electrons from the well with energies $\mathscr{E}$ and $\mathscr{E}-\omega$. The steady-state solution (10) is valid when the field reaches its steady-state value in a system with a constant electron flux.

With the aid of Eq. (10) the expressions for the currents $J_c(x)$ and $J_s(x)$ can be written in terms of the functions $\psi_0(x)$ and $\psi_{-1}(x)$:

$$J_c(x)=-\frac{ie}{2m^*}[(\psi_0^*\psi_{-1}'+\psi_{-1}^*\psi_0')-\text{c.c.}],$$
$$J_s(x)=\frac{e}{2m^*}[(\psi_{-1}^*\psi_0'-\psi_0^*\psi_{-1}')+\text{c.c.}]. \tag{13}$$

We note that the Fourier components $J_c(x)$ and $J_s(k)$ can be replaced simply by integrals, since $ka\ll 1$, i.e.,

$$J_{c,s}(k)\simeq\frac{1}{a}\int_0^a dx\; J_{c,s}(x). \tag{14}$$

Usually (see Refs. 10 and 11 and the references given there) one seeks a perturbation theory solution in V of the form (10) by substituting it in the initial equation (5) and restricting oneself to the first approximation. However, because of the system of Eqs. (11) derived here and the chosen form of the interaction (7), we can find an exact solution to Eqs. (11) without turning to perturbation theory.

A solution of Eqs. (11) can be sought in the form

$$\psi_n(x)=A_ne^{\gamma x},\quad n=0,-1, \tag{15}$$

where the eigenvalues γ satisfy the equation

$$\gamma^4+2\gamma^2\left(\mathscr{E}+\frac{V^2-\omega}{2}\right)+\mathscr{E}^2-\mathscr{E}_\omega=0 \tag{16}$$

and are equal to

$$\gamma_{1,2}=\pm\gamma_-,\quad \gamma_{3,4}=\pm\gamma_0,$$
$$\gamma_-=i\sqrt{\mathscr{E}+\frac{V^2-\omega}{2}-\sqrt{\left(\frac{V^2-\omega}{2}\right)^2+\mathscr{E}V^2}},$$
$$\gamma_0=i\sqrt{\mathscr{E}+\frac{V^2-\omega}{2}+\sqrt{\left(\frac{V^2-\omega}{2}\right)^2+\mathscr{E}V^2}}. \tag{17}$$

We write the general solution to Eq. (11) in the form

$$\psi_n(x)=\sum_{j=1}^4 A_{nj}\exp(\gamma_jx). \tag{18}$$

The coefficients A_{0j} and A_{-1j} are related by an equation which follows from Eq. (12):

$$A_{-1j}=\varepsilon_jA_{0j},\quad \varepsilon_j=\frac{\gamma_j\overline{V}}{\mathscr{E}-\omega+\gamma_j^2},\quad \overline{V}=\frac{eE}{\omega}. \tag{19}$$

Substituting Eqs. (18) and (19) into the boundary conditions (12), we arrive at a system of algebraic equations for the A_{nj}:

$$\sum_{j=1}^4 A_{0j}(1-\beta_j)=2q,\quad \sum_j A_{0j}\exp(\gamma_ja)(1-\widetilde{\beta}_j)=0,$$

$$\sum_j A_{0j}(1-\beta_{-j})\varepsilon_j=0, \quad \sum_j A_{0j}\exp(\gamma_j a)(1-\tilde{\beta}_{-j})\varepsilon_j=0, \tag{20}$$

where

$$\beta_j=\frac{\alpha_1-\gamma_j}{ip}, \quad \beta_{-j}=\frac{\alpha_1-\gamma_j}{ip_-}, \quad \tilde{\beta}_j=\frac{\alpha_2+\gamma_j}{ip},$$

$$\tilde{\beta}_{-j}=\frac{\alpha_2+\gamma_j}{ip_-}, \quad p=\sqrt{\mathscr{E}}, \quad p_-=\sqrt{\mathscr{E}-\omega}. \tag{21}$$

The solution of the system of Eqs. (20) can be written in the form

$$A_{0j}=\frac{2q(-1)^{j+1}}{\Delta}\sum_{l\neq j}\sum_{\substack{m\neq j,l\\ s\neq j,l}}\exp(\gamma_l a)$$
$$\times(1-\tilde{\beta}_l)\Lambda_{lj}\Delta_{ms}(p_-)\varepsilon_m\varepsilon_s, \tag{22}$$

where

$$\Lambda_{lj}=\begin{cases}-1 & \text{for} \quad l=3, \quad j=1,2,\\ -1 & \text{for} \quad l=2, \quad j=3,4,\\ +1 & \text{otherwise.}\end{cases}$$

Here Δ is the determinant of the system (20):

$$\Delta=\varepsilon_1\varepsilon_2\Delta_{12}(p_-)\Delta_{34}(p)+\varepsilon_1\varepsilon_4\Delta_{14}(p_-)\Delta_{23}(p)$$
$$+\varepsilon_2\varepsilon_3\Delta_{23}(p_-)\Delta_{14}(p)-\varepsilon_1\varepsilon_3\Delta_{13}(p_-)\Delta_{24}(p)$$
$$-\varepsilon_2\varepsilon_4\Delta_{24}(p_-)\Delta_{13}(p)+\varepsilon_3\varepsilon_4\Delta_{34}(p_-)\Delta_{12}(p), \tag{23}$$

where

$$\Delta_{ml}(p)=(1-\beta_m)(1-\tilde{\beta}_l)\exp(\gamma_l a)$$
$$-(1-\tilde{\beta}_m)(1-\beta_l)\exp(\gamma_m a),$$
$$\Delta_{ml}(p_-)=(1-\beta_{-m})(1-\tilde{\beta}_{-l})\exp(\gamma_l a)$$
$$-(1-\tilde{\beta}_{-m})(1-\beta_{-l})\exp(\gamma_m a). \tag{24}$$

Thus, Eqs. (17), (18), (22), and (23) give a general solution of the system (11) and (12) which is valid all the way to high fields. By substituting Eq. (18) in Eqs. (13) and (14), we can easily find the currents $J_s(k)$ and $J_c(k)$ and obtain, along with Eqs. (2) and (3), closed equations for determining the field and phase (lasing frequency ω). These equations make it possible to analyze the dependence of the threshold current, power, and lasing frequency on the structure parameters. It is easy to generalize these results to other types of structure.

4. WAVE FUNCTIONS AND POLARIZATION CURRENTS IN THE LIMIT OF HIGH BARRIERS

Resonant tunnelling shows up most clearly if the barrier penetration is low, so that the level widths Γ_1 and Γ_2 are small compared to the energies $\mathscr{E}_{1R}$ and $\mathscr{E}_{2R}$. This happens if the inequalities

$$\alpha_1/p\gg 1, \quad \alpha_2/p\gg 1, \tag{25}$$

are satisfied. In the following we shall assume that they are.

In addition, in the general expressions obtained above it is necessary to take into account the limitations imposed on the field amplitude: $V/p\ll 1$. (See Sec. 2.) Given the smallness of the parameter V/p, for γ_0, γ_-, and ε_j we have

$$\gamma_0\simeq ip\sqrt{1+\frac{V^2}{\omega}}, \quad \gamma_-\simeq ip_-\sqrt{1-\frac{V^2}{\omega}}, \tag{26}$$

and

$$\varepsilon_{1,2}\simeq\pm\frac{i\omega}{\overline{Vp_-}}, \quad \varepsilon_{3,4}\simeq\mp\frac{ip\overline{V}}{\omega},$$

$$\varepsilon_1\varepsilon_2\simeq\frac{2}{\widetilde{V}^2}, \quad \varepsilon_1\varepsilon_3\simeq\frac{p}{p_-}, \tag{27}$$

where

$$\widetilde{V}^2=\frac{2V^2p_-^2}{\omega^2}\ll 1. \tag{28}$$

The combination $\varepsilon_m\varepsilon_l$ appears in the determinant (23) in way such that a factor $\sim 1/\widetilde{V}^2$ is in front of the first term, ~ 1 is in front of the next four, and $\sim\widetilde{V}^2$ is in front of the last. The last term can be omitted to the accuracy which we accept. The first term in the determinant is the product of partial determinants, $\Delta_{12}(p_-)\Delta_{34}(p)$, which have a resonant character. For $\widetilde{V}=0$, $\alpha_j/p\gg 1$, and $\alpha_j/p_-\gg 1$, we can write the expression for $\Delta_{34}(p)$ near the resonance in the form

$$\Delta_{34}(p_R+\delta_p)=\left(2-\frac{\alpha}{ip}\right)\left(2-\frac{\alpha\delta}{ip}\right)\exp(-ipa)$$
$$+\frac{\alpha^2\delta}{p^2}\exp(ipa)\simeq\exp(-ip_0a)$$
$$\times\left[-\frac{2(1+\delta^2)}{\delta}+\frac{2i\alpha^2\delta a}{p_0^2}\,\delta p\right], \tag{29}$$

where

$$\delta p=p-p_R, \quad p_R=\frac{\alpha p_0}{\alpha+(1+\delta)/a\delta}, \quad p_0=\frac{\pi n}{a},$$

$$n=1,2,..., \quad \alpha_1=\alpha, \quad \alpha_2=\alpha\delta. \tag{30}$$

A similar expression for $\Delta_{12}(p_-)$ is obtained on making the substitution $p\to p_-$.

For the following discussion it is convenient to write $\Delta_{34}(p)$ and $\Delta_{12}(p_-)$ in the following form:

$$\Delta_{34}(p)\simeq\frac{ia\alpha^2\delta}{p^3}\exp(-ip_0a)\tilde{\Delta}_{34}, \quad \tilde{\Delta}_{34}=\mathscr{E}-\mathscr{E}_{2R}+i\Gamma_2,$$

$$\Delta_{12}(p_-)\simeq\frac{ia\alpha^2\delta}{p_-^3}\exp(-ip_-a)\tilde{\Delta}_{12},$$

$$\tilde{\Delta}_{12}=\mathscr{E}-\omega-\mathscr{E}_{1R}+i\Gamma_1, \tag{31}$$

where

$$\Gamma_j=\frac{2p_j^3(1+\delta^2)}{a\alpha^2\delta^2},\quad p_2\equiv p,\quad p_1\equiv p_-,\quad \mathscr{E}_{jR}=p_{jR}^2, \tag{32}$$

where Γ_j is the width of the electronic levels. Evidently, we must set $\widetilde{V}=0$ in the determinants Δ_{ml} with $\widetilde{V}^2$. In the principal approximation with respect to α/p they are equal to

$$\Delta_{ml}(p_{1R})\simeq-\frac{2\alpha^2\delta}{p_-^2},\quad \Delta_{ml}(p_{2R})\simeq-\frac{2\alpha^2\delta}{p^2} \tag{33}$$

and are α^2/p^2 times greater than the resonant values of Δ_{12} and Δ_{34}. Thus, corrections $\propto\widetilde{V}^2$ for $\Delta_{12}\Delta_{34}$ make a small contribution with respect to p^2/α^2 to the term proportional to $\widetilde{V}^2$. As a result, we arrive at the following expression for the determinant:

$$\Delta=-\frac{2a^2\alpha^4\delta^2}{\widetilde{V}^2p_1^2p_2^3}\,\widetilde{\Delta}(\lambda),\quad \lambda^2=\frac{16p_1p_2\widetilde{V}^2}{a^2},$$

$$\widetilde{\Delta}(\lambda)=\{[(\mathscr{E}-\omega-\mathscr{E}_{1R})+i\Gamma_1][\mathscr{E}_{2R}-\mathscr{E}-i\Gamma_2]+\lambda^2\}. \tag{34}$$

It is easy to see from Eq. (34) that the electromagnetic field can have a significant effect on resonant tunnelling if

$$\lambda^2>\Gamma_1\Gamma_2,\quad \widetilde{V}>\frac{p_1p_2(1+\delta^2)}{2\alpha^2\delta^2}. \tag{35}$$

Since $\alpha/p\gg1$, the condition (35) is satisfied simultaneously with the inequality $\widetilde{V}\ll1$. (See Eq. (28).) We shall refer to a field which satisfies the condition (35) as a strong field (cf. Ref. 9).

We now find the coefficients A_{0j} and $A_{-1,j}$. First of all, it should be noted that, as can be seen from Eqs. (22) and (27), the wave function $\psi_0(x)$ (as well as $\psi_{-1}(x)$) is a superposition of eigenfunctions with momenta $p(\exp(\pm ipx))$ and functions $\exp(\pm ip_-x)$ induced as a result of interlevel transitions. This is why the approach involving an expansion in the eigenfunctions of the unperturbed problem usually employed in laser theory does not work.

As Eq. (27) implies, the principal parameters A_{0j} in terms of the parameter $\widetilde{V}^2$ are A_{03} and A_{04}, which are related to lowest order in α/p by

$$A_{03}\simeq-A_{04},\quad A_{03}=-\frac{2qp_2^2\widetilde{\Delta}_{12}(p_1)\exp(-ipa)}{\alpha p_1a\widetilde{\Delta}(\lambda)}, \tag{36}$$

so that

$$\psi_0(x)\simeq2iA_{03}\sin(p_2x). \tag{37}$$

If we include the correction to next order in α/p in the expression for A_{03}, then instead of Eq. (37) we have

$$\psi_0(x)\simeq\frac{2q\widetilde{\Delta}_{12}(p_1)p_2^2}{\alpha p_1a\widetilde{\Delta}(\lambda)}\left\{-\sin[p_2(x-a)]+\frac{p_2}{2\alpha\delta}\exp[ip_2(x-a)]\right\}. \tag{38}$$

The second term in Eq. (38) describes the intra-subband current. It is important for calculating the transmission and reflection coefficients, but can be omitted from the currents J_c and J_s because it is small ($\sim p/\alpha$) compared to the first terms.

The main contribution to ψ_{-1} is from the coefficients A_{-11} and A_{-12}, which are given in the principal approximation with respect to α/p by

$$A_{-11}\simeq-A_{-12}, \tag{39}$$

and

$$A_{-11}\simeq-\frac{8\sqrt{2}q\widetilde{V}p_2^3}{\alpha a^2p_1\widetilde{\Delta}(\lambda)}, \tag{40}$$

so that

$$\psi_{-1}(x)=2iA_{-11}\sin(p_1x)=-\frac{16\sqrt{2}iq\widetilde{V}p_2^3}{\alpha a^2p_1\widetilde{\Delta}(\lambda)}\sin(p_1x). \tag{41}$$

These expressions are valid for fields $\widetilde{V}\ll p/\alpha$. If $p/\alpha<\widetilde{V}<1$, then it is easy to show, using Eq. (22), that a term proportional to $\widetilde{V}^2\alpha^2/p^2$ is added to the numerator of the expression for $\psi_0(x)$.

It should be noted that similar expressions for ψ_n and Δ in the resonance case $\delta p=0$ have been obtained recently[12] by another method. Perturbation theory with respect to the field was applied to solving Eq. (5) and then a series expansion was made with the principal terms in α/p retained.[12] The convergence condition for the series places a limit on the field amplitude. Since a Hamiltonian of the form Ex for the interaction with the field was used,[12] a direct comparison is difficult. The expressions given there[12] are approximately valid for fields satisfying the condition $\lambda^2<\Gamma_1\Gamma_2$, i.e., weak fields in the sense of Eq. (35). The wave functions found in Ref. 12 were used to calculate the transmission and reflection coefficients.

Substituting Eqs. (37) and (41) in Eq. (13), for the currents $J_c(k)$ and $J_s(k)$ we find

$$J_{c,s}(k)=-\frac{2ieM_{12}}{m^*}[A_{03}^*A_{-11}\mp A_{03}A_{-11}^*]i^{(l\mp1)/2}, \tag{42}$$

which, with Eqs. (36) and (40), take the form

$$J_c(k)=\widetilde{V}Q\eta\frac{\Gamma_1\Gamma_2}{|\widetilde{\Delta}(\lambda)|^2}, \tag{43}$$

$$J_s(k)=\widetilde{V}Q\eta\frac{\Gamma_2(\mathscr{E}-\omega-\mathscr{E}_{1R})}{|\widetilde{\Delta}(\lambda)|^2}, \tag{44}$$

where

$$M_{12}=\frac{1}{a}\int_0^a dx[p_1\sin(p_2x)\cos(p_1x)-p_2\sin(p_1x)\cos(p_2x)], \tag{45}$$

and

$$Q=\frac{q^2p}{m^*},\quad \eta=\frac{64M_{12}e^2p_1^2\delta^2}{\omega^2a^2(1+\delta^2)}, \tag{46}$$

and Q is the coherent pump current.

5. SINGLE-MODE LASING IN A QUANTUM CASCADE LASER WITH MONOENERGETIC ELECTRON PUMPING

We shall assume that the electrons incident on the upper working level have the same energy $\mathscr{E} \simeq \mathscr{E}_{2R}$. This situation is realized if the Fermi energy $\mathscr{E}_F$ of the electrons in the emitter is small compared to the level with Γ_j. After substitution of Eqs. (43) and (44) in Eqs. (2) and (3) in the stationary case, the latter take the form

$$1 = \widetilde{Q}\,\frac{\Gamma_1\Gamma_2}{|\widetilde{\Delta}(\lambda)|^2}, \tag{47}$$

and

$$\omega - \Omega = \widetilde{Q}\,\frac{\Gamma_2(\mathscr{E} - \omega - \mathscr{E}_{1R})}{\tau_0|\widetilde{\Delta}(\lambda)|^2}, \tag{48}$$

where

$$\widetilde{Q} = 4\pi\tau_0 Q\,\eta/\kappa, \tag{49}$$

and $\widetilde{\Delta}(\lambda)$ is given by one of Eqs. (34).

We find the threshold pump currents $\widetilde{Q}_{th}$ and lasing frequency ω by setting $\lambda = 0$ in Eqs. (47) and (48). Taking the electron energy to be $\mathscr{E}_{2R}$, i.e., $\mathscr{E} = \mathscr{E}_{2R}$, we find an expression for the threshold current

$$\widetilde{Q}_{th} = \Gamma_2[\Gamma_1^2 + (\omega_{21} - \omega)^2]/\Gamma_1 \tag{50}$$

and for the frequency

$$\omega = \frac{S\omega_{21} + \Omega}{S+1}, \quad S = \frac{1}{\Gamma_1\tau_0}. \tag{51}$$

The ''stabilization coefficient'' S, given by the ratio of the width of the cavity mode $(1/\tau_0)$ to the width of the emission line,[8] can vary over wide limits, depending on the parameters of the structure and the cavity length. If, for example, we take[2,3] $\Gamma_1 \approx 10^{-12}\ \mathrm{s}^{-1}$ and $\tau_0 \approx 10^{-11}\ \mathrm{s}^{-1}$, then $S \approx 10$ and the frequency ω is mainly determined by the frequency ω_{21}. Since the separation between the cavity modes is small and the inequality $\Gamma_1 \geqslant (\omega_{21} - \Omega)/(1+S)$ may be satisfied, the threshold current is

$$\widetilde{Q}_{\mathrm{th}} = \Gamma_1\Gamma_2, \quad Q_{th} = \Gamma_1\Gamma_2\,\frac{\omega^2 a^2(1+\delta^2)\kappa}{4\pi\tau_0 \cdot 64 M_{12} e^2 p_1^2} \tag{52}$$

and is determined by the product of the level widths.

Therefore, in order to reduce the threshold current it is necessary to choose parameters which lead to a reduction in Γ_j. Equation (52) can be used to analyze the dependence of Q_{th} on other parameters as well, such as the frequency, well dimensions, etc.

It is easy to show that above the lasing threshold, Eq. (51) for the frequency remains valid, i.e., the frequency does not depend on the pump current, while the reduced laser power $P_\omega = \lambda^2$ is given by

$$P_\omega = \widetilde{\Gamma}(\sqrt{\widetilde{Q}} - \widetilde{\Gamma}), \quad \widetilde{\Gamma} = \sqrt{\Gamma_1\Gamma_2}. \tag{53}$$

We note, first of all, that the power P_ω depends nonlinearly on $\widetilde{Q}$, so that the rate of rise of the power falls off with rising $\widetilde{Q}$ and $P_\omega(\widetilde{Q})$ tends to saturate.

In addition, it is clear that for high pump levels the power is proportional to the product $\sqrt{\Gamma_1\Gamma_2}$. Thus, the power depends in a nontrivial way on the width $\widetilde{\Gamma}$: small values of $\widetilde{\Gamma}$ correspond to small threshold currents, and to lower powers, and vice versa. Therefore, there is an optimum (with respect to efficiency) which is attained for $\widetilde{Q} = 4\widetilde{\Gamma}^2$, at which the power is given by

$$P_\omega = \widetilde{\Gamma}^2, \quad \lambda = \widetilde{\Gamma}. \tag{54}$$

The reason for the reduced efficiency and rate of rise of the power is related to the fact that for fields $\lambda \geqslant \widetilde{\Gamma}$ the conditions for coherent tunnelling are violated (see below) and the flux of electrons incident from the source experiences strong reflection. This is confirmed by a calculation of the reflection coefficient.

Let us assume that the electron energy $\mathscr{E}$ differs from $\mathscr{E}_{2R}$ by an amount ξ, i.e.,

$$\mathscr{E} = \mathscr{E}_{2R} + \xi. \tag{55}$$

We can see that adjusting ξ changes the lasing regime radically. To avoid cumbersome formulas, we begin the analysis with the case $\Gamma_1 = \Gamma_2 = \Gamma$, assuming that $\omega_{21} - \omega \ll \Gamma$.

We find an expression for the lasing power from Eq. (47):

$$P_\omega \equiv \lambda^2 = \xi^2 + \Gamma(\sqrt{\widetilde{Q} - 4\xi^2} - \Gamma). \tag{56}$$

For $\xi = 0$, Eq. (56) transforms to Eq. (53). For $\xi \neq 0$, the threshold current increases with rising ξ:

$$\widetilde{Q}_{th} = (\Gamma^2 + \xi^2)^2/\Gamma^2. \tag{57}$$

At the same time, above the threshold a detuning leads to an increase in the field λ. Thus, there is an optimum value of ξ_0 for a given pump current which can be determined from the equation

$$\frac{\partial\lambda^2}{\partial\xi_0} = 2\xi_0\left(1 - \frac{2\Gamma}{\sqrt{\widetilde{Q} - 4\xi_0^2}}\right) = 0. \tag{58}$$

Equation (58) has the following solutions:

$$\xi_0 = 0, \quad \xi_0^2 = \widetilde{Q}/4 - \Gamma^2. \tag{59}$$

The first solution corresponds to a peak power within the pump range $1 < \widetilde{Q}/\Gamma^2 < 4$, where the second solution does not exist $(\xi_0^2 < 0)$. When the condition

$$\widetilde{Q} > 4\Gamma^2 \tag{60}$$

is satisfied, the second solution appears. It corresponds to a peak P_ω (and for the zero solution the maximum is replaced by a minimum). For $\xi = \xi_0$ the power λ^2 as a function (56) of $\widetilde{Q}$ takes the form

$$\lambda^2 = \widetilde{Q}/4. \tag{61}$$

It can be seen from Eq. (61) that the power rises linearly with $\widetilde{Q}$ and is independent of Γ. Note that the linear regime sets in (noting Eq. (60)) when the field amplitude λ exceeds the damping Γ, i.e.,

$$\lambda > \Gamma. \tag{62}$$

Thus, if we increase the detuning in accordance with Eq. (59) as $\widetilde{Q}$ increases, then the lasing power is maximal. This effect can be attributed to splitting of the levels in a strong field and the appearance of a gap equal to λ.[9,14] Thus, if the energy of the incident electrons is fixed (e.g., $\xi=0$), then the energy of the pump electrons will deviate from resonance and they will be reflected. If, on the other hand, $\xi=\xi_0$ is increased in proportion to λ (see Eq. (59)), then the detuning is compensated. This interpretation is confirmed by a relationship between ξ_0 and λ obtained from Eq. (59) together with Eq. (61):

$$\xi_0^2=\lambda^2-\Gamma^2. \tag{63}$$

The existence of a regime with $\xi_0 \neq 0$ suggests a way of increasing the efficiency of a coherent laser in which the pump source has a simultaneously varying current and voltage related by Eq. (59). Here it should be kept in mind that the laser enters a regime where the power is independent of the damping Γ for $\widetilde{Q}>4\Gamma^2$. Thus, the optimum solution is to choose the minimum possible damping Γ. It is also clear that the drift from resonance can be compensated partially by a nonmonoenergetic electron distribution. This case will be examined in the next section.

It is easy to generalize these results to the case $\Gamma_1 \neq \Gamma_2$. For the reduced power, instead of Eqs. (56) and (57) we have

$$\lambda^2=\xi^2+\widetilde{\Gamma}\left(\sqrt{\widetilde{Q}-\frac{\xi^2(\Gamma_1+\Gamma_2)^2}{\widetilde{\Gamma}^2}}-\widetilde{\Gamma}\right) \tag{64}$$

and the threshold pump level is

$$\widetilde{Q}_{th}=\frac{\xi^4+\xi^2(\Gamma_1+\Gamma_2)^2+\widetilde{\Gamma}^4}{\widetilde{\Gamma}^2}. \tag{65}$$

The optimum tuning

$$\xi_0^2=\frac{\widetilde{Q}\,\widetilde{\Gamma}^2}{(\Gamma_1+\Gamma_2)^2}-\frac{(\Gamma_1+\Gamma_2)^2}{4} \tag{66}$$

ensures a linear dependence of λ^2 on the pump $\widetilde{Q}$,

$$\lambda^2=\frac{\widetilde{Q}\,\widetilde{\Gamma}^2}{(\Gamma_1+\Gamma_2)^2}-\frac{(\Gamma_1-\Gamma_2)^2}{4}, \tag{67}$$

if the pump and the field satisfy the inequalities

$$\widetilde{Q}>\frac{(\Gamma_1+\Gamma_2)^4}{4\widetilde{\Gamma}^2}, \quad \lambda>\widetilde{\Gamma}. \tag{68}$$

The coherent laser has yet another feature which distinguishes it from ordinary lasers: a population inversion is not required for it to operate. In fact, we can find the ratio of the populations in the upper and lower levels using Eqs. (38) and (41):

$$I=\frac{\int_0^a|\psi_0|^2dx}{\int_0^a|\psi_{-1}|^2dx}=\frac{\widetilde{\Gamma}^2}{\lambda^2}\,s, \quad s\simeq 1. \tag{69}$$

This implies that for a low pump current and weak field, i.e.,

$$\lambda<\widetilde{\Gamma}, \tag{70}$$

there is a population inversion ($I>1$). However, if the pump and field become strong, the population in the lower level exceeds that in the upper level, i.e., there is no population inversion. It can be shown that in the optimum tuning regime, $\xi=\xi_0$, we have $I\approx 1$ according to Eq. (59), i.e., the populations are equal. The reason is as follows: in a coherent laser the electrons entering the well interact coherently with the field over their lifetime Γ^{-1}. In an incoherent laser scattering processes shift the phase and destroy the coherence of the interaction with the field, so the transition probability depends on the difference in populations. It can also be shown that the population in the upper level does not change as the pumping level is varied (at least for fields $\widetilde{V}<\alpha/p$), while that in the lower level increases in proportion to $\lambda/\widetilde{\Gamma}$.

Note that Eq. (69) has been obtained with a model in which the field is localized within the well. Since the field is delocalized, and exists outside the well, it is necessary to take absorption and emission of the field outside the well into account. However, these nonresonant transitions are small in terms of the parameter p^2/α^2.

6. NONMONOENERGETIC PUMPING

If the Fermi energy $\mathscr{E}_F$ in the emitter region is sufficiently high compared to $\widetilde{\Gamma}$, then it becomes necessary to average the currents (43) and (44) over the energy. We use the customary formulas[10,11] for the one-dimensional case:

$$J=\rho\int_{-\mathscr{E}_F/2}^{\mathscr{E}_F/2}J(\mathscr{E})d\mathscr{E}, \tag{71}$$

where ρ is the density of the state. Including the motion in the lateral plane leads to an additional factor

$$J\sim\int_{-\mathscr{E}_F/2}^{\mathscr{E}_F/2}J(\mathscr{E})\left(\frac{\mathscr{E}_F}{2}-\mathscr{E}\right)d\mathscr{E}. \tag{72}$$

In general, the applicability of Eqs. (71) and (72) to the coherent approximation is not evident. However, to illustrate the basic behavior we limit ourselves to Eq. (71). We shall again assume that $\omega\sim\omega_{21}\ll\Gamma$. Then, on substituting Eqs. (43) and (44) in Eq. (71), we obtain

$$J_c(k)=\widetilde{V}Q\,\eta\widetilde{\Gamma}^2\rho\int_{-\mathscr{E}_F/2}^{\mathscr{E}_F/2}\frac{d\mathscr{E}}{|\lambda^2-(\mathscr{E}+i\Gamma_1)(\mathscr{E}+i\Gamma_2)|^2}. \tag{73}$$

Since we studied monoenergetic pumping in Sec. 5, we shall now find J_c for the opposite limiting case, assuming that $\mathscr{E}_F\gg\Gamma_j,\lambda$. Taking the limits of integration as infinite, from Eq. (73) we obtain

$$J_c=\frac{\widetilde{V}Q\,\eta\widetilde{\Gamma}^2\pi\rho}{(\Gamma_1+\Gamma_2)(\lambda^2+\widetilde{\Gamma}^2)} \tag{74}$$

and, therefore, for the laser power we have

$$P_\omega=\widetilde{\Gamma}^2\left[\frac{\pi\widetilde{Q}\rho}{\Gamma_1+\Gamma_2}-1\right]. \tag{75}$$

It is important to note that the power depends linearly on the pump current even without ''self-tuning.'' This happens because, since the Fermi energy $\mathscr{E}_F$ is large, electrons at the

resonance energy always exist. It is easy to show (by taking the exact integral in Eq. (73)) that for $\lambda > \mathscr{E}_F$ a root dependence again occurs in Eq. (53). It should also be pointed out that, as opposed to Eqs. (61) and (67), in Eq. (75) the power depends significantly on Γ_j and ρ at high pump levels.

We can also find the lasing frequency for nonmonoenergetic pump electrons from the equation

$$\omega - \Omega = \frac{\tilde{Q}\Gamma_2 \rho}{\tau_0} \times \int_{-\infty}^{\infty} \frac{d\mathscr{E}(\mathscr{E} + \omega_{21} - \omega)}{|[(\mathscr{E} + (\omega_{21} - \omega) + i\Gamma_1](-\mathscr{E} - i\Gamma_2) + \lambda^2|^2}. \tag{76}$$

If we set $\Gamma_1 = \Gamma_2$, the integral of the first term is equal to zero because of the even denominator (as can be seen easily by making the variable substitution $\mathscr{E} = \mathscr{E}' - (\omega_{21} - \omega)/2$). Thus, the expression for the frequency is again given by Eq. (51) with $S = 1/2\tau_0\Gamma$.

In the general case of $\Gamma_1 \neq \Gamma_2$, the current J_s after integration over energy is given (for $\omega_{21} - \omega \ll \Gamma_1$) by

$$J_s = \frac{\pi \tilde{V} Q \eta 2\tilde{\Gamma}^2(\omega_{21} - \omega)\rho}{(\Gamma_1 + \Gamma_2)^2(\lambda^2 + \tilde{\Gamma}^2)}. \tag{77}$$

Substituting Eq. (77) in Eq. (3) and using Eq. (75), we find that ω is again given by Eq. (51) with $S = 1/\tau_0(\Gamma_1 + \Gamma_2)$.

7. DISCUSSION OF RESULTS. CONCLUSION

As expected (see the Introduction), a coherent laser with coherent pumping has a number of important features. First of all, there is no close link to a population inversion. Depending on the pumping current and power, the relationships between the upper and lower levels can be arbitrary. In the optimum operating regime for a coherent laser with tuning ξ_0, electrons enter the upper level, undergo one radiative transition (for arbitrary field λ), and leave the quantum well. This sort of behavior can be compared to self-transparency,[9] with the important difference that in our case the electron stays in the lower level.

If, on the other hand, there is no tuning, monoenergetic pumping leads to a nonlinear (root) dependence for the laser power on the pump current. A sharp deviation from a linear dependence is observed when the field λ exceeds the level width $\tilde{\Gamma}$. This prediction of the model also (as in the tuning regime) can be verified experimentally as, incidentally, can the linear dependence of the power in the case of nonmonochromatic pumping. Naturally, the most efficient coherent pump is monoenergetic with energy tuning in accordance with Eq. (59).

We would also like to emphasize that the electron damping required for generating an electromagnetic field is introduced in this model by the boundary conditions. Thus, there is no need to invoke approximate kinetic approaches to account for dissipative scattering processes. In this way, we have an example of a model for steady-state lasing as a strictly quantum mechanical problem which, at the same time, can be used to find the power and frequency.

The model studied here has been simplified as much as possible in order to reveal the basic features of the coherent laser. We assume, however, that with a certain choice of parameters that will ensure satisfaction of the condition $\tau_{ph} > \Gamma_j^{-1}$ (see the Introduction), the model is qualitatively applicable to quantum cascade lasers based on quantum wells, wires, and, especially, dots. A quantitative theory must include scattering processes, shifts in the levels owing to charge accumulation in the well, and the real barrier structure.

The author thanks Yu. V. Kopaev for numerous discussions and valuable comments.

This work has been supported by the Russian Fund for Fundamental Research (Grant No. 96-02-17363-a), the Russian MNTP Scientific Council on the Physics of Solid State Nanostructures (Project 1-092/4), and INTAS (Project No. 93-1703-ext).

[1] A. Kazarinov and R. Suris, Fiz. Tekhn. Poluprovodn. **5**, 242 (1971) [Sov. Phys. Semicond. **5**, 207 (1971)].
[2] J. Faist, F. Capasso, D. Sivco *et al.*, Science **264**, 553 (1994).
[3] J. Faist, F. Capasso, D. Sivco *et al.*, Appl. Phys. Lett. **66**, 538 (1994).
[4] S. Luryi, Appl. Phys. Lett. **47**, 490 (1985).
[5] A. F. Kazarinov and R. A. Suris, Fiz. Tekhn. Poluprovodn. **6**, 135 (1972) [Sov. Phys. Semicond. **6**, 109 (1972)]; **7**, 488 (1973) [Sov. Phys. Semicond. **7**, 347 (1973)]; J. Faist, F. Capasso *et al.*, Electron. Lett. **29**, 2230 (1993); F. Capasso *et al.*, Quantum. Electron. **22**, 1853 (1986); A. Kastulsky *et al.*, Appl. Phys. Lett. **59**, 2636 (1991); M. Helm *et al.*, Phys. Rev. Lett. **63**, 74 (1989).
[6] V. F. Elesin and Yu. Kopaev, Solid State Commun. **96**, 987 (1995); V. F. Elesin and Yu. V. Kopaev, Zh. Éksp. Teor. Fiz. **108**, 2186 (1995) [JETP **81**, 1192 (1995)].
[7] V. Gorpinkel, S. Luryi, and B. Gelmont, in *Proc. IEEE, Japan* (1996), p. 94.
[8] W. E. Lamb, Phys. Rev. **134**, 1429 (1964).
[9] V. M. Galitskiĭ and V. F. Elesin, *Resonant Interaction of Electromagnetic Fields with Semiconductors* [in Russian], Énergoatomizdat, Moscow (1986).
[10] R. K. Mains and G. I. Haddad, J. Appl. Phys. **64**, 3564 (1988); **64**, 5041 (1988).
[11] H. C. Lju, Phys. Rev. B **43**, 12538 (1991); A. B. Pashkovskiĭ, Zh. Éksp. Teor. Fiz. **109**, 1779 (1996) [JETP **82**, 959 (1996)].
[12] E. I. Golant and A. B. Pashkovskiĭ, JETP Lett. **63**, 590 (1996).
[13] M. Kira, Phys. Rev. B **53**, 15789 (1995).
[14] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, 3rd ed. (Pergamon, Oxford, 1977).

Translated by D. H. McNeill

Anomalously high kinetic energy of charged macroparticles in a plasma

A. P. Nefedov, A. G. Khrapak, S. A. Khrapak, O. F. Petrov, and A. A. Samaryan

Scientific Research Center for the Thermal Physics of Impulsive Interactions, Russian Academy of Sciences, 127412 Moscow, Russia

(Submitted 23 January 1997)
Zh. Éksp. Teor. Fiz. **112**, 499–506 (August 1997)

A number of recent experimental investigations of nonideal plasmas containing macroparticles have revealed an anomalous increase in that part of the kinetic energy of these macroparticles that corresponds to their random motion. In this paper a model is proposed for the dynamic motion of charged macroparticles that explains this phenomenon. Calculations based on this model are compared with experimental results. © *1997 American Institute of Physics.* [S1063-7761(97)00808-1]

1. INTRODUCTION

Systems consisting of a plasma plus solid charged particles with sizes ranging from below a micron to several microns are commonly encountered in space (planetary rings, interstellar clouds, comet tails, the ionosphere of the Earth, etc.). Macroparticles are also observed in the low-temperature laboratory plasmas used in plasma sputtering and etching, in the manufacture of microelectronic components, etc. The spontaneous generation of ordered structures made up of macroparticles observed in recent experiments on various types of laboratory plasmas[1–7] implies that such systems can be used as macroscopic models of real microscopic crystalline structures. The obvious advantage of using a plasma with microparticles (compared, e.g., with colloidal solutions) to model microsystems is the relative ease of obtaining and studying such systems, as well as the smallness of the times they require to relax to an equilibrium state and respond to changes in external parameters. These facts explain the recent surge of interest in studying the properties of plasmas with macroparticles.

One question that has been intensely studied recently concerns the dynamic behavior of the macroparticles. In a low-temperature weakly ionized laboratory plasma, macroparticles having a large surface area efficiently exchange energy with atoms of the neutral component. Therefore it is customary to assume that the macroparticles are characterized by the temperature of the neutral gas, and that their motion is subject to the laws of Brownian motion. A convenient way to clarify the nature of the behavior of these particles qualitatively is to introduce the autocorrelation function of the particle velocity $\Phi(\tau)=\langle v(0)v(\tau)\rangle$, where the angle brackets denote averaging over an ensemble. The simplest theory of Brownian motion that does not include hydrodynamic interactions (which are important only in extremely viscous media) predicts exponential decay of $\Phi(\tau)$ with a characteristic time τ_s,[8] defined as the time for slowing down of a macroparticle due to the viscosity of the medium: $\tau_s=M/\beta$ (here M is the mass of the macroparticle and β is its coefficient of friction in the medium). Thus,

$$\Phi(\tau)=\langle v_T^2\rangle\exp(-\tau/\tau_s). \tag{1}$$

Here $\langle v_T^2\rangle=3T_n/M$ is the mean square thermal velocity of the particle (where T_n is the temperature of the neutral gas in energy units). The mean square displacement of the particle within a time τ, in accordance with Ref. 9, is given by the expression

$$\langle \Delta r^2(\tau)\rangle = 2\int_0^\tau (\tau-t)\Phi(t)\,dt$$

$$= 2\langle v_T^2\rangle\tau_s\left\{\tau+\tau_s\left[\exp\left(-\frac{\tau}{\tau_s}\right)-1\right]\right\}. \tag{2}$$

From Eq. (2) it is easy to determine the value of the mean square displacement of a macroparticle in two limiting cases. For $\tau\gg\tau_s$ we have

$$\langle \Delta r^2(\tau)\rangle = 2\langle v_T^2\rangle\tau_s\tau = 6D_0\tau, \tag{3}$$

where $D_0=(1/3)\langle v_T^2\rangle\tau_s=T_n/\beta$ is the diffusion coefficient for a Brownian particle. In the opposite case $\tau\ll\tau_s$

$$\langle \Delta r^2(\tau)\rangle = \langle v_T^2\rangle\tau^2. \tag{4}$$

Equations (3) and (4) reflect the diffusive character of the macroparticle motion at large times and its essentially ballistic character for small times.

The most recent investigations of an ideal plasma with macroparticles show that under certain conditions the temperature can exceed the temperature of a neutral gas by factors of a hundred and even a thousand (here and in what follows we understand the macroparticle temperature to mean the temperature corresponding to the kinetic energy of their random motion).[7,10–12] This phenomenon is observed at low pressures and has not been explained. The temperature itself is determined in the following way: the motion of the macroparticles in a horizontal plane is recorded by a video camera, and then the appropriate processing of the video image is carried out. Based the projection of the mean square motion of the particle in a prespecified direction within the time τ between frames, the mean-square velocity is calculated to be $\langle v_x^2\rangle=(\Delta x^2)/\tau^2$. The temperature is introduced through the relation $2(T_{mes}/2)=(M/2)(\langle v_x^2\rangle+\langle v_y^2\rangle)$. Of course it goes without saying that this procedure itself needs an appropriate justification. Thus, even for ordinary noninteracting Brownian particles the interpretation of the experi-

mental results depends on the ratio of the braking time of a particle in the medium τ_s to the time interval between frames τ (see Eqs. (3) and (4)).

In this paper we present a model of the dynamic behavior of charged macroparticles in a plasma that can explain the considerable departure of the temperature of dust particles from that of the neutral component as the pressure in the system falls. We will compare the results obtained in the framework of this model with the experimental results available at the present time. The conditions for its applicability are also discussed.

2. DYNAMIC BEHAVIOR OF CHARGED MACROPARTICLES IN A PLASMA

Consider a system consisting of negatively charged macroparticles with a charge Z (in units of the electron charge), electrons, singly charged positive ions, and neutral gas. Let n_p, n_e, n_i, n_n be the corresponding concentrations of macroparticles, electrons, ions, and neutral gas atoms. The system is assumed to be quasineutral, so that

$$Zn_p + n_e = n_i. \tag{5}$$

Moreover, we will assume that the plasma is weakly ionized, i.e., $n_n \gg n_e, n_i$. We further assume that each of the subsystems of electrons, ions, and neutral gas are in a state of thermodynamic equilibrium with temperatures T_e, T_i, and T_n respectively. Since the mobility of electrons and ions considerably exceeds that of the macroparticles, the distribution of electrons and ions in a quasi-uniform plasma will adiabatically follow the distribution of macroparticles. When a small perturbation in the macroparticle density $n_p(r)$ arises, an electric field $\mathbf{E} = -\nabla\varphi$ appears that attempts to return the system to a uniform distribution. Let us write Boltzmann distributions for the electrons and ions:

$$n_e(r) = n_e^0 \exp\left[\frac{e\varphi(r)}{T_e}\right], \tag{6}$$

$$n_i(r) = n_i^0 \exp\left[-\frac{e\varphi(r)}{T_i}\right]. \tag{7}$$

Combining Eqs. (6) and (7) with the condition of quasineutrality (Eq. 5), we obtain

$$E = \frac{Z}{e}\left(\frac{n_i^0}{T_i} + \frac{n_e^0}{T_e}\right)^{-1} \nabla n_p. \tag{8}$$

The macroparticle flux is written in the form

$$I_p = -\frac{ZeE}{3} n_p - D_0 \nabla n_p. \tag{9}$$

Substituting (8) into (9) we obtain

$$I_p = -D_0\left[1 + \frac{T_i}{T_n}\frac{Z^2 n_p^0}{(1 + T_i/T_e)n_e^0 + Zn_p^0}\right]\nabla n_p. \tag{10}$$

After taking the divergence of both sides of Eq. (10) and using the equation of continuity for the macroparticle density, we finally obtain a diffusion equation

$$\frac{\partial n_p}{\partial t} + D_p \nabla^2 n_p = 0, \tag{11}$$

where the diffusion coefficient for the charged macroparticles has the form

$$D_p = D_0\left[1 + \frac{T_i}{T_n}\frac{Z^2 n_p^0}{(1 + T_i/T_e)n_e^0 + Zn_p^0}\right]. \tag{12}$$

Obviously, the mechanism we are discussing is analogous to ambipolar diffusion in an electron–ion plasma. Note that, whereas in an ordinary low-temperature plasma the coefficient of ambipolar diffusion can be significantly larger than the diffusion coefficient of ions only in the presence of hot electrons, in the present case we may expect a significant increase in D_p compared to D_0 under certain conditions even for an isothermal plasma ($T_e = T_i = T_n$) due to the large value of the charge Z. It is interesting to consider various limiting expressions for the diffusion coefficient of charged macroparticles in a plasma. Setting $Z = 0$, we have $D_p = D_0$, as we should expect for uncharged macroparticles. Reducing n_p to zero, we also obtain $D_p = D_0$; this shows that the increase in the macroparticle diffusion coefficient is essentially a collective effect. It is also easy to obtain an expression for D_p in an isothermal system consisting of positively charged macroparticles, electrons emitted by them, and neutral gas. This expression was discussed in Refs. 13 and 14 (in this case we simply need to replace Z by $-Z$). In complete agreement with the results of Refs. 13 and 14 we obtain $D_p = D_0(1 + Z)$.

It is well known that D_p can exceed D_0 in dilute suspensions of interacting particles,[9] which in most ways are analogous to systems of plasmas with macroparticles. The dynamic interactions of macroparticles can be represented approximately in the following way: let us write the velocity of each macroparticle in the form of a sum $v_p = v_T + v_d$, where v_T is the velocity connected with imbalance in collisions with atoms of the surrounding medium (the ''Brownian'' force), while v_d is the velocity caused by the interaction between the macroparticles (the drift velocity in the field that arises as the macroparticle density is perturbed). Let us also introduce the the drift velocity correlation time τ_I. Under the assumption that v_d and v_T are not correlated, we can write by analogy with Eq. (1)[9]

$$\Phi(\tau) = \langle v_T^2\rangle \exp(-\tau/\tau_s) + \langle v_d^2\rangle \exp(-\tau/\tau_I). \tag{13}$$

For $\tau \gg \tau_I$, Eq. (2) implies the following expression for the mean square displacement of the macroparticles:

$$\langle \Delta r^2\rangle = 6(D_0 + D_c)\tau = 6D_p\tau, \tag{14}$$

where

$$D_c = \frac{1}{3}\langle v_d^2\rangle \tau_I \tag{15}$$

is the so-called coefficient of collective diffusion, whose appearance is explained by the interaction between the macroparticles. For another limiting case $\tau \ll \tau_s$ we have

$$\langle \Delta r^2\rangle = (\langle v_T^2\rangle + \langle v_d^2\rangle)\tau^2. \tag{16}$$

It is obvious that correctly defining the time τ_I presents some problems. Here we define it as the time required to travel the average distance a between macroparticles in the plane of observation while moving at the drift velocity:

$$\tau_I^2(\langle v_{dx}^2\rangle+\langle v_{dy}^2\rangle)=a^2. \tag{17}$$

In suspensions of interacting particles the difference between τ_s and τ_I is very large, four to five orders of magnitude for particles of micron size.[9] Therefore, although D_c can greatly exceed D_0, the following condition usually holds:

$$\langle v_d^2\rangle \ll \langle v_T^2\rangle. \tag{18}$$

In this case the average kinetic energy of the macroparticles

$$K=\frac{3}{2}T_n+\frac{M\langle v_d^2\rangle}{2} \tag{19}$$

is determined exclusively by the temperature of the surrounding medium. In a plasma with macroparticles no such difference between τ_s and τ_I can arise, so that Eq. (18) may not be satisfied and the subsystem of macroparticles will not be in thermodynamic equilibrium with the neutral component. If the temperature of the subsystem of macroparticles is defined through the relation $K=3T_p/2$, after rewriting Eq. (12) in the form $D_p=D_0+D_c$, using Eq. (15), and the definition of τ_I in Eq. (17), we obtain

$$T_p=T_n+\frac{2MD_0^2}{a^2}\left(\frac{D_c}{D_0}\right)^2, \tag{20}$$

where in agreement with Eq. (12) the ratio D_c/D_0 is defined by the expression

$$\frac{D_c}{D_0}=\frac{T_i}{T_n}\frac{Z^2n_p^0}{(1+T_i/T_e)n_e^0+Zn_p^0}. \tag{21}$$

Under certain conditions the second term on the right side of Eq. (20) can exceed T_n. This also will correspond to a difference between the temperatures of the macroparticles and the neutral component. This difference arises from the internal fields that appear in a low-density system due to the spatial separation of charges. As the system pressure decreases, D_0 increases, which qualitatively explains the phenomenon of increased macroparticle kinetic energy observed in Refs. 7 and 10–12.

3. COMPARISON OF THE MODEL WITH EXPERIMENT

For a more detailed comparison of this model with experimental results we will use the paper by Melzer *et al.*,[10] which contains the most complete information on the parameters of our system. In this paper the authors observed the formation of crystal-like macroparticle structures in an rf He plasma at low pressures. The crystal structure forms near the boundary of the cathode sheath, where the force of gravity is balanced by an electric field. The macroparticles used in these experiments were spherical monodispersed particles of radius $R=4.7\ \mu$m and mass $M=6.7\times10^{-10}$ gm. Melzer *et al.* observed a phase transition from a crystalline state to a gaslike state as the pressure of the gas dropped from 120 to 40 Pa at constant discharge power. They explained the phase transition by invoking an increase in the effective macroparticle temperature, which in turn decreases the interaction parameter Γ, defined as the ratio of the electrostatic interaction energy between neighboring particles to their kinetic energy: $\Gamma=Z^2e^2/aT_p$. For values of the interaction parameter larger than a critical value Γ_c the system should be in the crystalline state, while as Γ decreases the system of macroparticles enters a liquid state. Debye screening in the plasma causes the value of Γ_c to depend on the ratio of the averaged interparticle distance to the Debye radius: $k=a/\lambda_D$. The case $k=0$ corresponds to the simplest model of a single-component plasma. In this model $\Gamma_c\simeq170$.

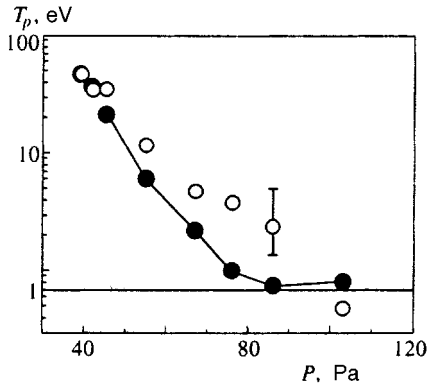


FIG. 1. Macroparticle temperature as a function of system pressure: ●—values obtained in experiment; ○—calculated using the model given here

Our scheme for measuring the particle temperature was described in the Introduction. In our model, this approach is correct only in the limit $\tau\ll\tau_s$, in which case the real kinetic energy of the macroparticles is measured according to Eqs. (16) and (19). The time between frames in these measurements was $\tau=20$ ms. Figure 1 shows the temperature of the macroparticles measured in this way for various values of the pressure. The horizontal straight line in Fig. 1, which corresponds to a temperature of $\sim$ 0.7 eV, is the smallest value of macroparticle temperature that can be determined by this video apparatus. This instrumental limit is connected with the discrete structure of the video image.[10]

Let us now discuss our choice of system parameters to be substituted into Eq. (20). We take $T_e\gg T_n=T_i=300$ K. The average distance between macroparticles is practically independent of the pressure and equals $a=450\ \mu$m,[10] which corresponds to a macroparticle density $n_p=a^{-3}\approx10^4$ cm^{-3}. In the range of pressures under study, the mean free path of He atoms greatly exceeds the radius of the macroparticles. Correspondingly, the coefficient of resistance of the medium is determined (under the assumption of complete accommodation) by the expression[15]

$$\beta=\frac{4\pi}{3}\left(4+\frac{\pi}{2}\right)R^2P\sqrt{\frac{m_n}{2\pi T_n}}, \tag{22}$$

where P is the neutral gas pressure and m_n is the mass of the He atom.

Now, by substituting into Eqs. (20) and (21) the macroparticle charge and ion concentration obtained in experiment at various pressures, we obtain the dependence of the mac-

roparticle temperature on pressure (within the framework of our model) that is applicable to the conditions of the experiment in Ref. 10. This dependence is also shown in Fig. 1. The quantity $\tau_I \approx a^2/2D_c$ obviously is a minimum at the lowest pressure in the system, and comes to $\sim$ 0.2 s at 40 Pa, while τ_s in this range of pressures varies from 20 to 60 ms. Thus, the condition $\tau_s > \tau$ is fulfilled; consequently, the procedure applied to measure the kinetic energy of the macroparticles is correct.

The results of calculations using our model are found to be in satisfactory agreement with the experimental results. It is unreasonable to expect more precise quantitative agreement for a large number of reasons. First of all, we note that the discussion given in this paper of the dynamic behavior of macroparticles is correct when the parameter Γ is not too large (i.e., for gaslike and liquidlike states of the macroparticle system), which corresponds to the region of low pressures in the experiments under study. In this sense, the use of Eq. (20) for all ranges of pressure may be considered to be an extrapolation of the model outside the limits of its region of applicability. The good qualitative agreement of the calculated macroparticle temperature with the experimental results at low pressures and its merely qualitative agreement at higher pressures (see Fig. 1) is also a reflection of this fact. Some of the discrepancy between the model and experiment can also be explained by experimental error in determining the macroparticle charge and ion concentration. For example, uncertainties in the charge measurements due to the method used in these experiments can be up to 50 leads to still larger uncertainties in the macroparticle temperature calculated using Eq. (20). The error corresponding to this is shown in Fig. 1 for one of the calculated points. Finally, we cannot avoid pointing out that there is a certain arbitrariness in the definition of τ_I.

Besides the considerations mentioned above, the following circumstance suggests that certain refinements are needed. Our observations during the particle motion are made near the boundary of the cathode sheath (a layer of positive bulk charge and electric field). In this region it is possible that the quasineutrality condition Eq. (5) is not strictly satisfied. Moreover, in this region T_i can exceed T_n somewhat due to ion drift towards the cathode in the electric field.

Figure 2 shows the dependence of the interaction parameter Γ on pressure in a system with temperature T_p calculated according to Eq. (20). This function clearly demonstrates that, despite some increase in the interaction between dust particles as the neutral gas pressure decreases (their charge increases while the screening decreases), the interaction parameter rapidly decreases due to the increase in macroparticle kinetic energy. This behavior of the parameter Γ is in complete agreement with the melting of the quasicrystalline structure observed in the experiments as the pressure falls in the discharge from 120 to 40 Pa.

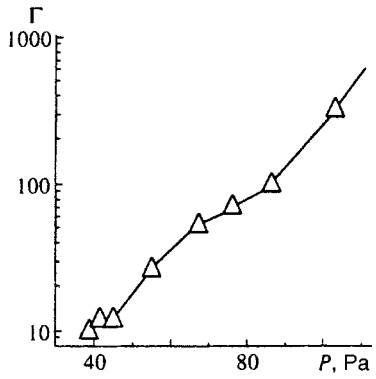


FIG. 2. Dependence of the interaction parameter Γ between macroparticles on system pressure

4. CONCLUSION

In this paper we have described a model for the dynamic behavior of macroparticles in a plasma that explains the anomalous increase in macroparticle kinetic energy observed in experiments at low pressures of the neutral component. The reason for this increase in temperature is the drift motion of dust particles in the internal fluctuating fields. The model described here is in satisfactory agreement with the experimental results. Possible reasons for certain discrepancies between calculated and experimental points have been discussed.

This paper was partially supported by the Russian Fund for Fundamental Research (project No. 97-02-17565).

[1] J. Chu and I. Lin, Phys. Rev. Lett. **72**, 4009 (1994).
[2] H. Thomas, G. E. Morfill, V. Demmel *et al.*, Phys. Rev. Lett. **73**, 652 (1994).
[3] A. Melzer, T. Trottenberg, and A. Piel, Phys. Lett. A **191**, 301 (1994).
[4] Y. Hayashi and K. Tachibana, Jpn. J. Appl. Phys., Part 1 **33**, 804 (1994).
[5] V. Fortov, A. Nefedov, O. Petrov *et al.*, Phys. Lett. A **219**, 89 (1996).
[6] V. E. Fortov, A. P. Nefedov, O. F. Petrov *et al.*, JETP Lett. **63**, 187 (1996).
[7] V. E. Fortov, A. P. Nefedov, V. M. Torchinskiĭ *et al.*, JETP Lett. **64**, 92 (1996).
[8] P. A. Weitz, D. J. Pine, P. N. Pusey *et al.*, Phys. Rev. Lett. **63**, 1747 (1989).
[9] P. N. Pusey and R. J. A. Tough, in *Dynamic Light Scattering: Applications of Photon Correlation Spectroscopy*, R. Pecora ed. (Plenum Press, New York, 1985), p. 85.
[10] A. Melzer, A. Homann, and A. Piel, Phys. Rev. E **53**, 2757 (1996).
[11] H. M. Thomas and G. Morfill, Nature (London) **379**, 806 (1996).
[12] J. B. Pieper and J. Goree, Phys. Rev. Lett. **77**, 3137 (1996).
[13] D. I. Zhukovitskiĭ and I. T. Yakubov, Teplofiz. Vys. Temp. **23**, 842 (1985) [in Russian].
[14] I. T. Iakubov and A. G. Khrapak, Sov. Technol. Rev. B: Therm. Phys. **2**, 269 (1989); D. I. Zhukovitskiĭ, A. G. Khrapak, and I. T. Yakubov, in the book: *Plasma Chemistry*, Vol. II, p. 130 [in Russian], B. M. Smirnov ed. (Energoizdat, Moscow, 1984).
[15] E. M. Lifshits and L. P. Pitaevskiĭ, *Physical Kinetics*, p. 86 (Pergamon Prtess, Oxford, 1981) [Russ. original Nauka, Moscow, 1979].
[16] Th. Trottenberg, A. Melzer, and A. Piel, Plasma Sources Sci. Technol. **4**, 450 (1995).

Translated by Frank J. Crowne

Radiative forces acting on point sources moving in a stratified fluid

V. A. Gorodtsov and S. N. Reznik

Institute of Problems in Mechanics, Russian Academy of Sciences, 117526 Moscow, Russia

Yu. A. Stepanyants

Institute of Applied Physics, Russian Academy of Sciences, 603600 Nizhniĭ Novgorod, Russia
(Submitted 13 January 1997)
Zh. Éksp. Teor. Fiz. **112**, 507–523 (August 1997)

Both components of the total radiative force acting on uniformly moving mass sources in a stratified incompressible fluid are calculated. These forces are related to the emission of internal gravity waves. One component of the radiative force is directed opposite to the direction of motion. This is the well-known wave drag force. Different useful representations of this component are derived for several typical point sources (monopole, dipole, and others). The other component, directed perpendicular to the direction of the motion, is calculated for the first time. The results obtained are extended to the case of rotating media, where internal inertial–gravity waves exist. © *1997 American Institute of Physics.* [S1063-7761(97)00908-6]

1. INTRODUCTION

It is well known that a uniform motion of force or mass sources in a stratified liquid is accompanied by the emission of internal waves. The wave fields and radiative forces which are generated by and act on such moving sources have been calculated in many works; see, for example, Refs. 1–12 and the references cited there (it is impossible to list here all works on this subject). The emission of internal waves is analogous to the Cherenkov emission of electromagnetic waves by charges moving uniformly with superluminal velocity in dielectrics. On account of radiative energy losses a retarding force due to wave drag acts on a moving source. In anisotropic media, however, apart from this force the total radiative force can have another component directed perpendicular to the velocity of the source.[13]

In most works having to do with the force characteristics arising with wave emission some modification of the problem of the horizontal or vertical motion of sources in a stratified medium was studied. In this case there is a definite symmetry with respect to the direction of gravity, so that in these problems the radiative force due to the emission of internal waves reduced only to wave drag and was calculated on the basis of energy considerations: The power $\mathbf{F}_{wd}\cdot\mathbf{V}_s$ dissipated per unit volume equals the energy flux W carried away from a given volume by the internal waves. However, if the source moves at an angle to the horizontal direction, then the spectrum and intensity of the radiation to the left and right of the source are, generally speaking, different as a result of the anisotropic nature of the internal waves. As an illustration, Fig. 1 shows a shadowgraph (kindly provided by E. Ya. Sysoeva and Yu. D. Chashechkin) that depicts the field of internal waves which are emitted by an obliquely moving sphere in a stratified fluid. It is clearly seen in this photograph that the characteristic wavelengths of the waves on the two sides of the trajectory of the sphere are appreciably different. On account of the fact that the radiation emitted from the source and the momentum carried away by the waves are not the same on different sides of the trajectory, apart from a retarding force a lateral recoil force should also act on the source. As a result, the total radiative force should contain two components: a wave drag force and a lateral force, which, evidently, cannot be calculated from energy considerations because it performs no work. It is obvious that, just like the wave drag force, it can be found directly by calculating the total momentum flux carried away by the emitted internal waves. In addition, if the wave drag force has already been found (for example, by means of energy considerations), the lateral force can then be easily calculated by a different method as well — using the formulas which will be presented and justified below. This method has already been successfully employed to calculate the lateral forces acting on uniformly moving sources emitting Rossby waves in the β-plane and drift waves in plasma.[13]

In the present paper the total radiative force acting on uniformly moving monopole and dipole mass sources in a stratified medium are derived. It is shown that the different computational methods give identical results.

2. BASIC EQUATIONS; GENERAL RELATIONS

For simplicity, we shall consider a mass source moving uniformly in an incompressible stratified fluid (as is well known,[12] the calculations for a force source reduce to calculations for an effective mass source). We shall assume that the fluid is unbounded and exponentially density-stratified $\rho_0(z)=e^{-N^2z/g}$, so that the Brunt–Väisälä frequency satisfies $N=\text{const}$. The basic equations of hydrodynamics in the Boussinesq approximation for small disturbances reduce to the system

$$\hat{L}\frac{\partial \mathbf{v}_\perp}{\partial t}=\left(\frac{\partial^2}{\partial t^2}+N^2\right)\frac{\partial}{\partial t}\nabla_\perp m, \quad \hat{L}w=\frac{\partial^3 m}{\partial z\,\partial t^2},$$

$$\hat{L}\left(\frac{\rho}{\rho_0}\right)=-\frac{N^2}{g}\frac{\partial^2 m}{\partial t\,\partial z}, \quad \hat{L}\left(\frac{p}{\rho_0}\right)=-\left(\frac{\partial^2}{\partial t^2}+N^2\right)\frac{\partial m}{\partial t}, \quad (1)$$

 1063-7761/97/080276-09$10.00

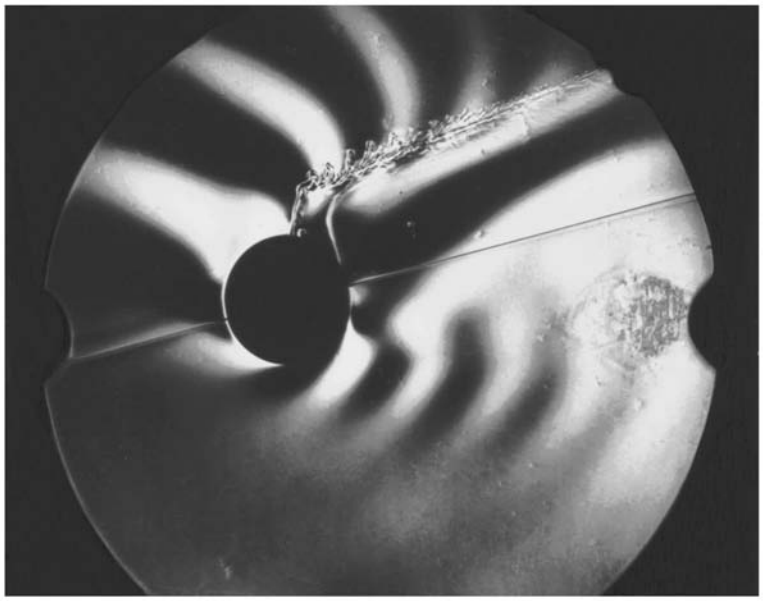


FIG. 1. Shadowgraph of the field of internal waves emitted by an obliquely moving sphere in a stratified fluid.

where $m(t,\mathbf{r})=m(\mathbf{r}-\mathbf{V}_s t)$ is the mass source determining the flow rate of the fluid, the source velocity $\mathbf{V}_s$ is assumed to be constant, $\hat{L}\equiv \partial^2/\partial t^2\,\Delta+N^2\Delta_\perp$ is a linear operator containing the three-dimensional Laplacian Δ and its two-dimensional part $\Delta_\perp$ corresponding to the x, y horizontal plane, and $\mathbf{v}=\{\mathbf{v}_\perp, w\}$ and $\mathbf{v}_\perp=\{u,v\}$.

From the system of equations (1) for the pressure field, a formal solution expressed in terms of the Green's function $G(t,\mathbf{r})$ is easily found for the stationary problem with the aid of the Fourier transform

$$P(\omega,\mathbf{k})=i\omega(N^2-\omega^2)G(\omega,\mathbf{k})m(\omega,\mathbf{k}),$$

$$p(t,\mathbf{r})=-\left(\frac{\partial^2}{\partial t^2}+N^2\right)\frac{\partial}{\partial t}\times\int\int m(t',\mathbf{r}')G(t-t',\mathbf{r}-\mathbf{r}')dt'd\mathbf{r}'. \qquad (2)$$

Knowing the pressure field it is easy to calculate the energy losses per unit time (dissipated power) which are due to the emission of internal waves:

$$W=\int p(t,\mathbf{r})m(t,\mathbf{r})d\mathbf{r}=\frac{1}{(2\pi)^{s+2}}\int\int\int P(\omega,\mathbf{k})m(\omega',-\mathbf{k})\times e^{-i(\omega+\omega')t}d\omega' d\omega d\mathbf{k},$$

where $s=2$ or 3 is the dimension of the space.

In the case of uniform motion the Fourier transform of the source is proportional to a delta function: $m_0(\mathbf{k})\delta[\omega-(\mathbf{k}\cdot\mathbf{V}_s)]$. As a result we obtain

$$W=-\frac{1}{(2\pi)^s}\times\int\int\omega(N^2-\omega^2)\mathrm{Im}\{G(\omega,\mathbf{k})\}\times|m_0(\mathbf{k})|^2\delta[\omega-(\mathbf{k}\cdot\mathbf{V}_s)]d\mathbf{k}d\omega. \qquad (3)$$

As one can see, the magnitude of these energy losses is determined only by the imaginary part of the Fourier transform of the Green's function (here and below we employ the retarded Green's function, which satisfies the causality condition $G(t-t',\mathbf{r}-\mathbf{r}')|_{t<t'}=0$). In the Boussinesq approximation for a uniformly stratified fluid we have

$$\mathrm{Im}\ G(\omega,\mathbf{k})=-\pi\,\mathrm{sign}(\omega)\,\delta(\omega^2k^2-N^2k_\perp^2). \qquad (4)$$

Now the equation (3) can be rewritten in the form

$$W=\frac{\pi}{(2\pi)^s}\int\int|\omega|(N^2-\omega^2)|m_0(\mathbf{k})|^2\delta[\omega-(\mathbf{k}\cdot\mathbf{V}_s)]\delta(\omega^2k^2-N^2k_\perp^2)d\mathbf{k}d\omega. \qquad (5)$$

This representation reflects the wave character of the losses, since the integrand is concentrated (because of the second delta function) on the dispersion surface of the internal waves

$$\omega^2k^2=N^2k_\perp^2 \quad \text{or} \quad \omega=\pm Nk_\perp/k.$$

Furthermore, the radiation wave field is stably related to the uniformly moving source, so that the radiation frequency is proportional to the longitudinal component of the wave vector (the first delta function in Eq. (5))

$$\omega=(\mathbf{k}\cdot\mathbf{V}_s) \quad \text{or} \quad \mathbf{c}\equiv\frac{\omega}{k^2}\mathbf{k}=\mathbf{V}_s.$$

For this reason, the possible wave vectors of the emitted waves are determined by the simple system of equations

$$\begin{cases}\omega=(\mathbf{k}\cdot\mathbf{V}_s),\\ \omega^2k^2=N^2k_\perp^2.\end{cases} \qquad (6)$$

It follows from the second equation of this system that the range of frequencies of the emitted waves is limited: $\omega^2 \leqslant N^2$.

3. PLANAR PROBLEM FOR INTERNAL WAVES

The system of relations (6) makes it possible to express in the two-dimensional case the two possible wave vectors $\mathbf{k}_+$ and $\mathbf{k}_-$ of the radiation as a function of the frequency ω (frequency parameterization):

$$\mathbf{k}_\pm = \left(\frac{\omega}{V_\pm}, \pm \frac{\sqrt{N^2-\omega^2}}{V_\pm} \right), \quad V_\pm \equiv V_{sx} \pm V_{sz} \frac{\sqrt{N^2-\omega^2}}{\omega}. \tag{7}$$

An alternative angular parametrization of these wave vectors $\mathbf{k}_\pm$ is found as follows. Let us switch to polar coordinates $k_x = k \cos\theta$ and $k_z = k \sin\theta$ and express the components of the velocity of the source in terms of the magnitude of the velocity and the slope angle with respect to the horizontal axis x: $V_{sx} = V_s \cos\varphi$ and $V_{sz} = V_s \sin\varphi$. Then, using the relation $\omega = (\mathbf{k}\cdot\mathbf{V}_s)$, we write

$$\mathbf{k} = \frac{(\omega \cos\theta, \omega \sin\theta)}{V_{sx}\cos\theta + V_{sz}\sin\theta} = \frac{\omega}{V_s} \frac{(\cos\theta, \sin\theta)}{\cos(\theta-\varphi)}. \tag{8}$$

Using the dispersion relation $\omega = \pm N \cos\theta$ we obtain the desired angular parametrization

$$\mathbf{k}_\pm = \pm \frac{N \cos\theta}{V_s} \frac{(\cos\theta, \sin\theta)}{\cos(\theta-\varphi)}. \tag{9}$$

Thus, we obtain an expression for the wave number and the limit on the angles and frequencies of the emitted internal waves:

$$k = \frac{N}{V_s} \left| \frac{\cos\theta}{\cos(\theta-\varphi)} \right|, \quad \frac{\omega}{\cos(\theta-\varphi)} > 0. \tag{10}$$

The inequality follows from the equation $\omega = kV_s \cos(\theta-\varphi)$ and the positivity of the wave number.

To determine the contributions of these waves more accurately, it is necessary to transform the products of the two delta functions appearing in the integrand for the dissipated power W.

In the case of frequency parameterization we have

$$\delta[\omega - (\mathbf{k}\cdot\mathbf{V}_s)]\,\delta(\omega^2 k^2 - N^2 k_x^2) = \frac{H(N^2-\omega^2)}{2\omega^2\sqrt{N^2-\omega^2}} [\delta(\mathbf{k}-\mathbf{k}_+) + \delta(\mathbf{k}-\mathbf{k}_-)], \tag{11}$$

where $H(\xi)$ is the Heaviside function.

Now the equation for the dissipated power acquires the form of the following expansion in terms of the frequency (frequency parametrization of the wave vectors $\mathbf{k}_\pm$ is assumed)

$$W = \frac{1}{8\pi} \int_{-N}^{N} \frac{\sqrt{N^2-\omega^2}}{|\omega|} (|m_0(\mathbf{k}_+)|^2 + |m_0(\mathbf{k}_-)|^2) d\omega. \tag{12}$$

In the case of a point source ($|m_0(\mathbf{k})|^2 = 1$) the integral in Eq. (12) diverges logarithmically at low frequencies. To eliminate the divergence it is sufficient to introduce a form factor of the form $|m_0(\mathbf{k})|^2 = |f(\omega/V_s)|^2 \to 0$ as $\omega \to 0$ (we recall that $\mathbf{k}_\pm \cdot \mathbf{V}_s = \omega$. For this form factor the dissipated power

$$W = \frac{1}{4\pi} \int_{-N}^{N} |f(\omega/V_s)|^2 \sqrt{\frac{N^2}{\omega^2} - 1}\; d\omega \tag{13}$$

becomes finite and does not depend on the direction of motion of the source. The wave drag force $F_{wd} = W/V_s$ in this case also does not depend on the direction of motion of the source. The mass source under consideration is nonlocal in the direction of its motion:

$$m(t,\mathbf{r}) = \frac{\delta(r_\perp)}{2\pi} \int f(k_\parallel) e^{ik_\parallel (r_\parallel - V_s t)} dk_\parallel = \delta(r_\perp) \times F(r_\parallel - V_s t).$$

A point dipole oriented in the direction of motion is also a mass source of this type

$$m(t,\mathbf{r}) = -(\mathbf{l}\cdot\nabla)\,\delta(\mathbf{r}-\mathbf{V}_s t) = \frac{l}{V_s} \frac{\partial}{\partial t} \delta(\mathbf{r}-\mathbf{V}_s t), \tag{14}$$

where $\mathbf{l} = l\mathbf{V}_s/V_s$ is the dipole moment and $m(\omega,\mathbf{k}) = -2\pi i(\omega l/V_s)\,\delta[\omega - (\mathbf{k}\cdot\mathbf{V}_s)]$. The form factor for this force is $f(\omega/V_s) = -il\omega/V_s$.

For a more general form of nonlocalization the dissipated power depends on the direction of motion.

We shall now transform the product of delta functions in order to obtain an angular parametrization

$$\delta[\omega - (\mathbf{k}\cdot\mathbf{V}_s)]\,\delta(\omega^2 k^2 - N^2 k_x^2) = \frac{\delta(\omega)\,\delta(k^2)}{N^2\cos^2\theta} + \frac{V_s}{2N^3} \frac{|\cos(\theta-\varphi)|}{|\cos\theta|^3} \delta\left[k - \frac{N}{V_s} \left| \frac{\cos\theta}{\cos(\theta-\varphi)} \right| \right] \times \left\{ H\left[\frac{\cos\theta}{\cos(\theta-\varphi)} \right] \delta(\omega - N\cos\theta) + H\left[-\frac{\cos\theta}{\cos(\theta-\varphi)} \right] \delta(\omega + N\cos\theta) \right\}. \tag{15}$$

Except for the term with $\delta(\omega)\,\delta(k^2)$, which does not contribute to the losses, the final expression gives the desired angular parametrization of the frequency and wave number. Substituting this expression into the general equation for the losses (5) we obtain

$$W = \frac{N}{8\pi} \int_0^{2\pi} \frac{\sin^2\theta}{|\cos\theta|} \times m_0\left(\frac{N\cos^2\theta}{V_s\cos(\theta-\varphi)}, \frac{N\cos\theta\sin\theta}{V_s\cos(\theta-\varphi)} \right) \Bigg|^2 d\theta. \tag{16}$$

In this form, in accordance with what has been said above, the dependence on the direction of motion (on the angle φ) disappears for a nonlocalization of longitudinal type

$$|m_0(\mathbf{k})|^2 = |f[(\mathbf{k}\cdot\mathbf{V}_s)/V_s]|^2 = \left| f\left(\frac{N}{V_s}\cos\theta \right) \right|^2,$$

and the dissipated power and the wave drag force remain finite in the limit $f\to 0$ with $\cos\theta\to 0$.

In the case of a mass dipole consisting of nonlocal source and sink separated by a distance $2|\mathbf{a}|$ and moving in an arbitrary direction with a nonzero angle of attack (i.e. $\mathbf{a}\times\mathbf{V}_s\neq 0$), we have

$$m(t,\mathbf{r})=m_0(\mathbf{r}-\mathbf{a}-\mathbf{V}_s t)-m_0(\mathbf{r}+\mathbf{a}-\mathbf{V}_s t),$$

$$m(\omega,\mathbf{k})=-4\pi i m_0(\mathbf{k})\sin(\mathbf{k}\cdot\mathbf{a})\delta[\omega-(\mathbf{k}\cdot\mathbf{V}_s)] \tag{17}$$

and one need only make the substitution

$$|m_0(\mathbf{k})|^2\to 4\sin^2(\mathbf{k}\cdot\mathbf{a})|m_0(\mathbf{k})|^2 \tag{18}$$

in the equations presented above for the energy losses. Then the angular expansion for the dissipated power acquires the form

$$W=\frac{N}{2\pi}\int_0^{2\pi}\frac{\sin^2\theta}{|\cos\theta|}\sin^2\left[\frac{Na}{V_s}\cos\theta\frac{\cos(\theta-\alpha)}{\cos(\theta-\varphi)}\right]|m_0(\mathbf{k})|^2 d\theta, \tag{19}$$

$$\mathbf{k}=\frac{N}{V_s}\left|\frac{\cos\theta}{\cos(\theta-\varphi)}\right|(\cos\theta,\sin\theta),$$

where α is the inclination angle of the dipole with respect to the x axis. Here the limit $|m_0(\mathbf{k})|^2\to m_0^2=$ const corresponds to a dipole with spatially separated point source and sink. Hence an equation can also be obtained, in particular, for the angular expansion of the dissipated power for a point dipole ($a\to 0$, $m_0\to\infty$, $l=2m_0a=$ const $<\infty$) moving at some angle of attack ($\alpha=\varphi\neq 0$):

$$W=\frac{N^3l^2}{8\pi V_s^2}\int_0^{2\pi}\sin^2\theta|\cos\theta|\frac{\cos^2(\theta-\alpha)}{\cos^2(\theta-\varphi)}d\theta. \tag{20}$$

Similarly, we have for the frequency representation of the dissipated power

$$W=\frac{1}{2\pi}\int_{-N}^{N}\sqrt{\frac{N^2}{\omega^2}-1}\sum_{\pm}\sin^2(\mathbf{k}_{\pm}\cdot\mathbf{a})|m_0(\mathbf{k}_{\pm})|^2 d\omega$$

$$=\frac{1}{2\pi}\int_{-N}^{N}\sqrt{\frac{N^2}{\omega^2}-1}\sum_{\pm}\sin^2\left(\omega\frac{a_x\omega\pm a_z\sqrt{N^2-\omega^2}}{V_{sx}\omega\pm V_{sz}\sqrt{N^2-\omega^2}}\right)$$

$$\times|m_0(\mathbf{k}_{\pm})|^2 d\omega. \tag{21}$$

The interference factor $\sin^2(\ldots)$ appearing here ordinarily guarantees convergence of the integral at low frequencies, for example, for a dipole oriented in the same direction as its velocity ($\mathbf{a}\|\mathbf{V}_s$), in which case it reduces to $\sin^2(\omega a/V_s)$. Then the result for W and F_{wd} with $|m_0(\mathbf{k})|^2=1$ is independent of the direction of motion. On the other hand, in the particular case of a dipole in horizontal motion with the angle of attack ($V_{sz}=0$, $a_z\neq 0$) the interference factor, which assumes the simplified form

$$\sin^2\left(\frac{a_x}{V_{sx}}\omega\pm\frac{a_z}{V_{sx}}\sqrt{N^2-\omega^2}\right),$$

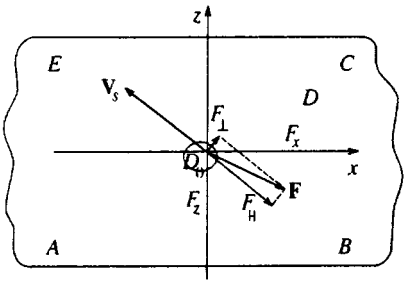


FIG. 2. Geometry of fluid flow around a body in a system of coordinates in which the source is stationary.

no longer eliminates the nonintegrable singularity in Eq. (21) at low frequencies. Here regularization likewise requires that the source be nonlocal, and once again nonlocalization of a longitudinal type is sufficient:

$$|m_0(\mathbf{k})|^2=|f(\omega/V_s)|^2\to 0 \text{ as } \omega\to 0.$$

The expression (21) can be written somewhat more simply with the substitution $\omega\to-\omega$,

$$W=\frac{1}{\pi}\int_{-N}^{N}\sqrt{\frac{N^2}{\omega^2}-1}\sin^2\left(\omega\frac{a_x\omega+a_z\sqrt{N^2-\omega^2}}{V_{sx}\omega+V_{sz}\sqrt{N^2-\omega^2}}\right)$$

$$\times|m_0(\mathbf{k}_+)|^2 d\omega. \tag{22}$$

Then we have in the limiting case of a point dipole

$$W=\frac{1}{4\pi}\int_{-N}^{N}|\omega|\sqrt{N^2-\omega^2}\left(\frac{l_x\omega+l_z\sqrt{N^2-\omega^2}}{V_{sx}\omega+V_{sz}\sqrt{N^2-\omega^2}}\right)^2 d\omega.$$

Trigonometric substitutions reduce this equation to the form (20).

Knowing the dissipated power, we can easily find in each specific case the wave drag force, which is directed opposite to the velocity vector of the source. Specifically, for a point dipole moving without an angle of attack the wave drag force is $\mathbf{F}_{wd}=-\rho_0 l^2N^3\mathbf{V}_s/6\pi V_s^4$. However, this force is only a part of the total radiative force $\mathbf{F}_r=\mathbf{F}_\parallel+\mathbf{F}_\perp$, where $\mathbf{F}_\parallel\equiv\mathbf{F}_{wd}$.

Let us now consider the question of the calculation of the radiative force $\mathbf{F}_\perp$ perpendicular to the direction of motion. Since this force performs no work, it cannot be calculated on the basis of the law of conservation of energy, as done above for the wave drag force. A different approach is required here. One possible approach is developed in Ref. 14, where an expression is found for the horizontal force acting on a body in a stratified shear flow. In a frame of reference moving together with the source, the results of Ref. 14 transfer *in toto* to our case with a substantial simplification, since the fluid flow does not contain a velocity shear at infinity. The approach developed consists essentially of the following.

As is well known, the hydrodynamic force acting on a unit length of a cylindrical body in a fluid is determined as the negative of the sum of the momentum flux carried away by the internal waves from a given volume (see Fig. 2) to

infinity and the change in the momentum per unit time inside this volume. The x-component of the force can be expressed as[15]

$$F_x = -\frac{\partial}{\partial t}\int\int_D \rho_0 u\,ds + \int\int_{ABCE} \Pi_{xk} n_k\,ds, \tag{23}$$

where Π is the momentum flux tensor, $\mathbf{n}$ is a unit vector normal to the surface of the cylinder, and the regions D and $ABCE$ are shown in Fig. 2.

If the body in the flow is replaced by a mass source $m(x,z)$, then from Euler's equation and the continuity equation follows an equation for the x-component of the velocity

$$\frac{\partial u}{\partial t} + \frac{\partial u^2}{\partial x} + \frac{\partial uw}{\partial z} + \frac{1}{\rho_0}\frac{\partial p}{\partial x} = um. \tag{24}$$

Integrating the equation (24) over the contour $ABCE$ (see Fig. 2) we obtain

$$\frac{\partial}{\partial t}\int\int_{D+D_0} \rho_0 u\,ds - \int\int_{ABCE} \Pi_{xk} n_k\,ds = \int\int_{D+D_0} \rho_0 um\,dx\,dz. \tag{25}$$

Letting the outer boundary of the fluid volume D approach ∞, we obtain in the limit an expression for the horizontal component of the radiative force (see Ref. 14)

$$F_x = -\int_{-\infty}^{+\infty}\int \rho_0 um\,dx\,dz. \tag{26}$$

The Fourier transform of the vertical component of the velocity of the disturbance is easily found from the second equation of the system (1), where the operator $\partial/\partial t$ must be replaced by $U_s\partial/\partial x + W_s\partial/\partial z$, as a result of which we find

$$\tilde{w} = \frac{-ik_z\tilde{m}}{k^2 - (N/(U_s + k_z W_s/k_x) + i\varepsilon)^2},$$

where U_s and W_s are the horizontal and vertical components of the flow velocity $\mathbf{V}_s$ and the formal symbol ε is, in accordance with Lighthill's rule,[16] an infinitesimal quantity reflecting the causal character of the response of the medium.

The Fourier transform of the horizontal component of the velocity of the disturbance can be found from the continuity equation

$$\tilde{u} = \frac{\tilde{m}}{ik} - \frac{k_z}{k_x}\tilde{w}.$$

Then we obtain for the horizontal component of the force in the case of a point dipole oriented parallel to the direction of motion

$$F_x = \frac{\rho_0}{2\pi^2}\mathrm{Re}\left(\int_0^\infty dk_x \int_{-\infty}^{\infty} \frac{k_z}{k_x}\tilde{w}\tilde{m}\,dk_z\right)$$
$$= \frac{\rho_0 l^2 N^3}{4\pi V_s^3}\int_{-\pi/2}^{\pi/2} \frac{(\sin\theta\cos\theta)^2}{\cos(\theta-\varphi)}\,d\theta.$$

Here we employed the property of the Fourier transform of a real quantity $\tilde{f}(-k) = \tilde{f}^*(k)$.

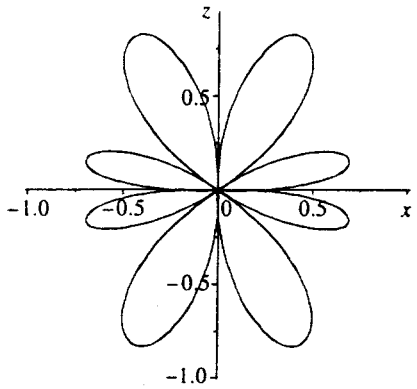


FIG. 3. Normalized transverse force $F_\perp/|F_\parallel|$ versus the inclination angle φ of the trajectory with respect to the horizontal axis x for a dipole source in polar coordinates. $F_\parallel = -\rho_0 l^2 N^3/6\pi V_s^3$.

We find the transverse radiative force (perpendicular to the direction of motion) from the relation (see Fig. 2)

$$F_\perp = \frac{F_x - F_\parallel\cos\varphi}{\sin\varphi} = -\frac{\rho_0 l^2 N^3}{4\pi V_s^3}$$
$$\times \int_{-\pi/2}^{\pi/2} \sin^2\theta\cos\theta\tan(\theta-\varphi)\,d\theta. \tag{27}$$

Here the expression (20) with $\alpha = \varphi$ was used for $F_\parallel$.

We note that the expression (27) for $F_\perp$ is identical to the expression calculated by the method of Ref. 13 using the concept of wave momentum. By analogy to quantum mechanics the wave momentum density can be written in the form $\mathbf{p} = \overline{N}\mathbf{k}$, where $\overline{N} = \overline{S}/|\mathbf{V}_g|$ is the specific number of quasiparticles and $\overline{S}$ is the wave action. Then the force due to the wave momentum flux passing through the surface of a cylinder of unit height is determined in polar coordinates by the integral

$$\mathbf{F} = \int_S \mathbf{p}|\mathbf{V}_g|\,ds = \int_{-\pi/2}^{\pi/2} \mathbf{k}\overline{S}(\theta) r\,d\theta.$$

The product of the transverse component of the wave vector and the wave action

$$\overline{S} = \frac{\rho_0 l^2 N^3 \sin^2\theta\cos\theta}{8\pi V_s^3 k_\parallel r}$$

appears in the integrand in the equation (27).

The integral in the equation (29) cannot be expressed in terms of elementary functions, but it can be easily calculated numerically. The dependence of $F_\perp$ on φ has the form shown in Fig. 3. It is obvious that for a source moving along the x and z axes the lateral force vanishes on account of symmetry while the wave drag force is independent of angle and remains finite (if the angle of attack for the dipole equals zero). It is interesting, however, that in each range of angles $0 + n\pi/2 < \varphi < (n+1)\pi/2$, where $n = 0$, 1, 2, and 3, there exists another direction $\varphi_0 + n\pi/2$ in which $F_\perp$ vanishes. Numerical calculations using the equation (27) give the angle $\varphi_0 \approx 0.655 \approx 37.5°$. Analysis of this formula shows that each such direction is stable with respect to small variations of the

angle, while the directions $\theta=0+n\pi/2$ are unstable. In other words, if the source tilts to the right or left away from the direction characterized by the angle φ_0, then the lateral force strives to restore the source to its initial orientation. However, if the source tilts away from the horizontal or vertical directions, then the lateral force pushes it farther away from these directions.

4. INERTIAL–GRAVITY WAVES

Taking account of the rotation of the medium together with stratification for inertial–gravity (or gravity–gyroscopic) waves automatically eliminates the ''infrared'' divergence of the integrals, even in the case of point sources, because zero frequencies are cut off (these frequencies fall outside the admissable frequency range). We shall write the basic equations of hydrodynamics in the Boussinesq approximation for small disturbances due to the motion of a point source in an incompressible stratified fluid rotating around the vertical axis:

$$\frac{\partial \mathbf{V}}{\partial t}+2\Omega\nabla z\times\mathbf{V}=-\nabla p-\rho g\nabla z,$$

$$\frac{\partial \rho}{\partial t}+\frac{N^2}{g}\mathbf{V}\cdot\nabla z=0,\quad \nabla\cdot\mathbf{V}=m. \tag{28}$$

Since the Coriolis force performs no work, the energetics of the wave motion remains the same as in a nonrotating fluid. One need only substitute into the expression for the dissipated power $W=\int pm\, d\mathbf{r}$ the new dependence of the pressure on the mass source

$$\frac{\partial}{\partial t}\hat{L}p=-\left(\frac{\partial^2}{\partial t^2}+N^2\right)\left(\frac{\partial^2}{\partial t^2}+4\Omega^2\right)m,$$

$$-i\omega p(\omega,\mathbf{k})=-(N^2-\omega^2)(4\Omega^2-\omega^2)G(\omega,\mathbf{k})m(\omega,\mathbf{k}), \tag{29}$$

$$\hat{L}\equiv\frac{\partial^2}{\partial t^2}\Delta+N^2\Delta_\perp+4\Omega^2\frac{\partial^2}{\partial z^2}$$

and use the retarded Green's function of the modified operator $\hat{L}$

$$G(\omega,\mathbf{k})=[(\omega+i\varepsilon)^2k^2-N^2k_\perp^2-4\Omega^2k_z^2]^{-1},$$

$$\text{Im } G(\omega,\mathbf{k})=-\pi\,\text{sign}(\omega)\,\delta(\omega^2k^2-N^2k_\perp^2-4\Omega^2k_z^2). \tag{30}$$

Finally we obtain

$$W=\frac{\pi}{(2\pi)^s}\int\int|m_0(\mathbf{k})|^2\frac{(N^2-\omega^2)(\omega^2-4\Omega^2)}{|\omega|}\delta[\omega$$

$$-(\mathbf{k}\cdot\mathbf{V}_s)]\delta(\omega^2k^2-N^2k_\perp^2-4\Omega^2k_z^2)d\mathbf{k}\,d\omega. \tag{31}$$

Here the integrand is concentrated at the surface given by the argument of two delta functions, i.e., the surface described by the system of equations

$$\omega=(\mathbf{k}\cdot\mathbf{V}_s),\quad \omega^2k^2=N^2k_\perp^2+4\Omega^2k_z^2. \tag{32}$$

The dispersion equation now limits the possible frequencies to the interval between 2Ω and N (for $N>2\Omega$ this interval will be $N>\omega>2\Omega$ and for $N<2\Omega$, the opposite, $N<\omega<2\Omega$). Therefore both high and low frequencies are cut off when rotation is taken into account.

In the two-dimensional problem we have

$$\omega=k_xV_{sx}+k_zV_{sz},\quad k_x\sqrt{N^2-\omega^2}=\pm k_z\sqrt{\omega^2-4\Omega^2}.$$

Now it is easy to find the frequency parametrization of the possible radiation wave vectors

$$k_x^\pm=\frac{\omega}{V_\pm},\quad k_z^\pm=\pm\frac{\omega}{V_\pm}\sqrt{\frac{N^2-\omega^2}{\omega^2-4\Omega^2}},$$

$$V_\pm\equiv V_{sx}\pm V_{sz}\sqrt{\frac{N^2-\omega^2}{\omega^2-4\Omega^2}}. \tag{33}$$

We note that the wave vectors $\mathbf{k}^\pm$ introduced here are somewhat different from the wave vectors $\mathbf{k}_\pm$ used previously in the limit $\Omega=0$.

The angular parametrization is obtained similarly

$$\omega=\pm\sqrt{N^2\cos^2\theta+4\Omega^2\sin^2\theta},$$

$$k=\pm\frac{\sqrt{N^2\cos^2\theta+4\Omega^2\sin^2\theta}}{V_s\cos(\theta-\varphi)}. \tag{34}$$

Here we call attention to the value $k=0$. Since the dispersion equation

$$k^2(\omega^2-N^2\cos^2\theta-4\Omega^2\sin^2\theta)=0$$

has a separable form, it can also admit a nonwave solution with $k^2=0$.

The product of the two delta functions in the integrand for the dissipated power can now be transformed as follows:

$$\delta[\omega-(\mathbf{k}\cdot\mathbf{V}_s)]\delta(\omega^2k^2-N^2k_x^2-4\Omega^2k_z^2)$$

$$=\frac{1}{2|\omega|\sqrt{(N^2-\omega^2)(\omega^2-4\Omega^2)}}H\left(\frac{N^2-\omega^2}{\omega^2-4\Omega^2}\right)$$

$$\times\left[\delta\left(k_x-\frac{\omega}{V_+}\right)\delta\left(k_z-\frac{\omega}{V_+}\sqrt{\frac{N^2-\omega^2}{\omega^2-4\Omega^2}}\right)\right.$$

$$\left.+\delta\left(k_x-\frac{\omega}{V_-}\right)\delta\left(k_z+\frac{\omega}{V_-}\sqrt{\frac{N^2-\omega^2}{\omega^2-4\Omega^2}}\right)\right]$$

$$=\frac{H[(N^2-\omega^2)(\omega^2-4\Omega^2)]}{2|\omega|\sqrt{(N^2-\omega^2)(\omega^2-4\Omega^2)}}$$

$$\times[\delta(\mathbf{k}-\mathbf{k}^+)+\delta(\mathbf{k}-\mathbf{k}^-)]. \tag{35}$$

Making use of this, we obtain from Eq. (31)

$$W=\frac{1}{8\pi}\int H\left(\frac{N^2-\omega^2}{\omega^2-4\Omega^2}\right)\sqrt{\left(\frac{N^2}{\omega^2}-1\right)\left(1-\frac{4\Omega^2}{\omega^2}\right)}$$

$$\times\sum_\pm|m_0(\mathbf{k}^\pm)|^2d\omega. \tag{36}$$

Since low frequencies have been cut off, here the energy losses are finite even for a point source ($|m_0(\mathbf{k})|^2=1$):

$$W=\frac{1}{4\pi}\int \sqrt{\left(\frac{N^2}{\omega^2}-1\right)\left(1-\frac{4\Omega^2}{\omega^2}\right)}H\left(\frac{N^2-\omega^2}{\omega^2-4\Omega^2}\right)d\omega. \tag{37}$$

This integral can be expressed in terms of complete elliptic integrals,[8] making it easy to find the wave drag force as well.

In another special case of an extended dipole consisting of point source and sink which is moving in an arbitrary direction at the angle of attack we have

$$m_0(\mathbf{k})=-2im_0\ \sin(\mathbf{k}\cdot\mathbf{a}),$$

$$W=\frac{m_0^2}{2\pi}\int \sqrt{\left(\frac{N^2}{\omega^2}-1\right)\left(1-\frac{4\Omega^2}{\omega^2}\right)}H\left(\frac{N^2-\omega^2}{\omega^2-4\Omega^2}\right)$$
$$\times\sum_{\pm}\ \sin^2\left(\omega\frac{A_{\pm}}{V_{\pm}}\right)d\omega, \tag{38}$$

$$A_{\pm}\equiv a_x\pm a_z\sqrt{\frac{N^2-\omega^2}{\omega^2-4\Omega^2}}.$$

For a zero angle of attack ($\mathbf{a}\parallel\mathbf{V}_s$) the last factor in the integrand assumes the simplified form $2\sin^2(\omega a/V_s)$, so that the results for the dissipated power and the wave drag force are once again independent of the direction of motion.

We now transform the product of delta functions, making use of the angular parametrization,

$$\delta[\omega-(\mathbf{k}\cdot\mathbf{V}_s)]\delta(\omega^2k^2-N^2k_x^2-4\Omega^2k_z^2)=\delta[\omega$$
$$-kV_s\ \cos(\theta-\varphi)]\delta[k^2(\omega^2-N^2\cos^2\theta$$
$$-4\Omega^2\sin^2\ \theta)]=\frac{\delta(\omega)\delta(k^2)}{N^2\cos^2\theta+4\Omega^2\ \sin^2\ \theta}$$
$$+\frac{\delta\left[k-\dfrac{\omega}{V_s\cos(\theta-\varphi)}\right]}{2k^2V_s|\cos(\theta-\varphi)|\sqrt{N^2\cos^2\theta+4\Omega^2\ \sin^2\ \theta}}$$
$$\times[\delta(\omega-\sqrt{N^2\cos^2\theta+4\Omega^2\ \sin^2\ \theta})$$
$$+\delta(\omega+\sqrt{N^2\cos^2\theta+4\Omega^2\ \sin^2\ \theta})]. \tag{39}$$

At this stage we drop the term containing $\delta(\omega)\delta(k^2)$ (it will be discussed below). Then we have for the dissipated power

$$W=\frac{(N^2-4\Omega^2)^2}{8\pi}\int_0^{2\pi}\frac{\sin^2\ \theta\ \cos^2\theta}{(N^2\cos^2\theta+4\Omega^2\sin^2\theta)^{3/2}}$$
$$\times\left|m_0\left[\frac{\cos\ \theta\sqrt{N^2\cos^2\theta+4\Omega^2\ \sin^2\theta}}{V_s\cos(\theta-\varphi)},\right.\right.$$
$$\left.\left.\frac{\sin\ \theta\sqrt{N^2\cos^2\theta+4\Omega^2\sin^2\theta}}{V_s\cos(\theta-\varphi)}\right]\right|^2 d\theta. \tag{40}$$

We must set $|m_0(\ldots)|^2=1$ for a point source and

$$|m_0(\ldots)|^2=4m_0^2$$
$$\times\sin^2\left[\frac{\cos(\theta-\alpha)}{\cos(\theta-\varphi)}\ \frac{a}{V_s}\sqrt{N^2\cos^2\theta+4\Omega^2\sin^2\theta}\right]$$

for a dipole.

When the angle of attack of the dipole becomes zero, the function $|m_0(\ldots)|$ and the dissipated power no longer depend on the direction of motion. Then we obtain in the limit of a point dipole

$$W=\frac{(N^2-4\Omega^2)^2l^2}{8\pi V_s^2}\int_0^{2\pi}\frac{\sin^2\theta\ \cos^2\theta}{\sqrt{N^2\ \cos^2\theta+4\Omega^2\sin^2\theta}}d\theta. \tag{41}$$

The discarded term with $\delta(\omega)\delta(k^2)$ gives for the expression for the dissipated power a total correction of the form

$$\frac{1}{4\pi}\int_0^{2\pi}d\theta\int_0^{\infty}k\ dk\int\ |m_0(\mathbf{k})|^2$$
$$\times\frac{(N^2-\omega^2)(\omega^2-4\Omega^2)\delta(\omega)\delta(k^2)}{|\omega|(N^2\cos^2\theta+4\Omega^2\sin^2\theta)}d\omega,$$

which for a point source reduces to a product of a logarithmically diverging integral over frequency regularized by nonlocalization and an indeterminate integral over wave number $\int_0^{\infty}k\delta(k^2)dk$. If $\delta(k^2)$ is replaced by $\delta(k^2-\varepsilon)$, the last integral becomes 1/2 and if $\delta(k^2)$ is replaced by $\delta(k^2+\varepsilon)$, the integral vanishes. There arises the question of how to calculate this integral correctly.

The answer is that rejecting the Boussinesq approximation leads to a regularization that is close to the latter regularization. This is seen from the following expression:

$$\delta[k^2\omega^2-k^2(N^2\ \cos^2\theta+4\Omega^2\sin^2\theta)]$$
$$\to\delta[(k^2+k_0^2)\omega^2-k^2(N^2\cos^2\theta+4\Omega^2\sin^2\theta)],$$

where $k_0^2=N^2/2g$. The regularization $\delta(k^2)\to\delta(k^2+\varepsilon)$ is also supported by the fact that for a different transformation of the product of the generalized functions in terms of (35) to (37) without such an indeterminacy, a result differing from the expression (40) only by a trigonometric substitution for the integration variable is obtained.

The lateral component of the radiative force can once again be calculated by the methods presented above. An expression for the lateral component of the radiative force acting on a dipole can be obtained in a manner completely similar to the case of purely gravity waves:

$$F_{\perp}=-\frac{\rho_0l^2N^3}{4\pi V_s^3}\left(1-4\frac{\Omega^2}{N^2}\right)^2$$
$$\times\int_{-\pi/2}^{\pi/2}\frac{\sin^2\theta\ \cos^2\theta\ \tan(\theta-\varphi)}{\sqrt{\cos^2\theta+4\Omega^2\sin^2\theta/N^2}}d\theta. \tag{42}$$

Apart from a dimensional factor in front of the integral, the expression for $F_{\perp}$ for a monopolar source is distinguished by the fact that the denominator in the integrand contains, instead of a square root, the expression $(\cos^2\theta+4\Omega^2\sin^2\theta/N^2)^{3/2}$.

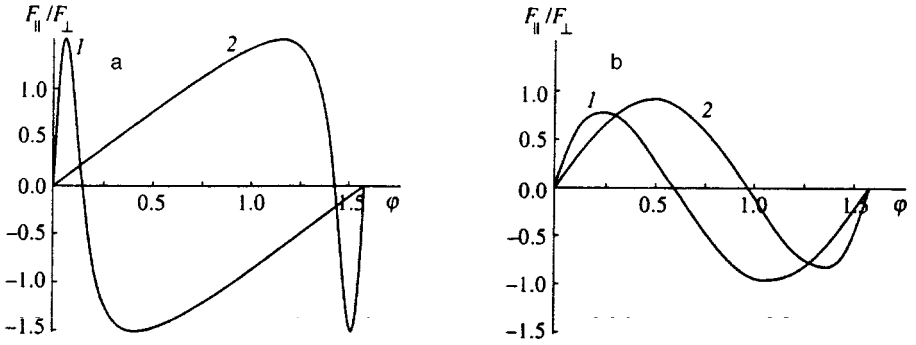


FIG. 4. Normalized transverse force $F_\perp/|F_\parallel|$ versus the inclination angle φ of the trajectory with respect to the horizontal axis x for monopole (a) and dipole (b) sources in a Cartesian coordinate system for different values of the parameter $2\Omega/N$: 1 — $2\Omega/N=0.1$, 2 — $2\Omega/N=10$.

Plots of the lateral force $F_\perp$ normalized to the wave drag force $F_\parallel$ are displayed in Fig. 4 in Cartesian coordinates ($0\leqslant\varphi\leqslant\pi/2$) for a monopole (a) and dipole (b) with different values of the parameter Ω/N. The form of the plots is qualitatively the same as for gravity waves, except that the value of φ_0 for which $F_\perp$ vanishes increases with the parameter Ω/N, approaching some constant value $\varphi_l>\pi/4$.

In the limiting case when the Coriolis parameter (the doubled rotational frequency of the fluid) equals the Brunt–Väisälä frequency ($2\Omega=N$), the total radiative force vanishes on account of the absence of wave motions in the volume of the fluid. This follows formally from the equations obtained for both components of the radiative force (see, for example, Eq. (42)).

5. THREE-DIMENSIONAL PROBLEM

The starting general equations (5) for internal waves and (31) for inertial–gravity waves are suitable for arbitrary spatial dimension s. Setting $s=3$, we confine our attention to waves of the first type, since for $s=3$ taking account of rotation does not give any advantage with respect to the convergence of the integrals and it complicates the exposition.

In spherical coordinates (it should be noted that in contrast to the two-dimensional case the angle θ is now measured not from the horizontal direction but rather from the vertical direction)

$$k_x=k\sin\varphi\sin\theta,\quad k_y=k\cos\varphi\sin\theta,$$

$$k_z=k\cos\theta,\quad k_\perp=k\sin\theta,$$

$$V_{sx}=V_s\sin\gamma,\quad V_{sy}=0,\quad V_{sz}=V_s\cos\gamma,$$

$$\mathbf{k}\cdot\mathbf{V}_s=kV_s\cos(\mathbf{k}\cdot\mathbf{V}_s)=kV_s[\cos\theta\cos\gamma+\sin\theta\sin\gamma\sin\varphi]$$

we have the angular parametrization

$$\delta[\omega-(\mathbf{k}\cdot\mathbf{V}_s)]\,\delta(\omega^2k^2-N^2k_\perp^2)=\frac{\delta\{k-N|\sin\theta|/[V_s|\cos(\mathbf{k}\cdot\mathbf{V}_s)|]\}}{2Nk^2V_s|\sin\theta\cos(\mathbf{k}\cdot\mathbf{V}_s)|}\times\{H[\cos(\mathbf{k}\cdot\mathbf{V}_s)]\,\delta(\omega-N\sin\theta)+H[-\cos(\mathbf{k}\cdot\mathbf{V}_s)]\,\delta(\omega+N\sin\theta)\}.$$

We note that in the three-dimensional case the presence of the compensating factor k^2 in the integration volume element eliminates any question concerning the legitimacy of the substitution $\delta(k^2\ldots)=(1/k^2)\,\delta(\ldots)$.

Finally, we obtain for the dissipated power

$$W=\frac{N^2}{16\pi^2V_s}\int_0^{2\pi}d\varphi\times\int_0^{\pi}\frac{\sin\theta\cos^2\theta}{|\cos\theta\cos\gamma+\sin\theta\sin\gamma\sin\varphi|}\times|m_0(\mathbf{k})|^2|_{k=\kappa}\,d\theta,\tag{43}$$

where

$$\kappa=\frac{N|\sin\theta|}{V_s|\cos\theta\cos\gamma+\sin\theta\sin\gamma\sin\varphi|}.$$

Here, in contrast to the two-dimensional problem, the contribution from large wave numbers is logarithmically large. The integrals can now be regularized by using a form factor of the form $|m_0(\mathbf{k})|^2=|F(k)|^2\to0$ as $k\to\infty$ or cutting off the integral over wave numbers at large k, which in the angular representation is equivalent to the constraint $|\cos\theta\cos\gamma+\sin\theta\sin\gamma\sin\varphi|>\varepsilon|\sin\varphi|$. The dipolar character of the source does not save us from divergence, and the spatial extent of the source (nonlocality) must be taken into account.

Further, for a point source with a regularizing form factor we easily find an expression for the lateral radiative force

$$F_\perp=\frac{N^2}{16\pi^2V_s^2}\int_0^{2\pi}d\varphi\times\int_0^{\pi}\frac{\sin\theta\cos^2\theta}{|\cos\theta\cos\gamma+\sin\theta\sin\gamma\sin\varphi|}\times\frac{k_\perp}{k_\parallel}|m_0(\mathbf{k})|^2|_{k=\kappa}\,d\theta,\tag{44}$$

$$k_\parallel=k(\cos\theta\cos\gamma+\sin\theta\sin\gamma\sin\varphi),$$

$$k_\perp^2=k^2-k_\parallel^2.$$

The last two integrals likewise cannot be expressed in terms of elementary functions, but they can be calculated numerically.

6. CONCLUSIONS

The results obtained above show that the motion of sources in stratified media is accompanied by the emission of internal waves, as a result of which radiative forces arise.

Since internal waves possess anisotropic properties in the vertical plane, generally speaking, such waves are generated by sources with different efficiency to the left and right of the trajectory of the source. Therefore the radiative force which arises contains in the general case two components: a longitudinal component which is directed opposite to the velocity of the source — this is the retarding wave drag force — and a transverse force. The calculation of the transverse component of the radiative force for internal waves in the present work is the first such calculation (a similar force was previously calculated for Rossby waves in the β-plane and for drift waves in plasma[13]).

The calculations showed that in the simplest cases of monopolar and dipolar sources (moving without an angle of attack in the latter case) the wave drag force does not depend on the direction of motion, while the lateral force depends strongly on the angle between the velocity vector and the horizontal axis. This component of the source vanishes for motion in the horizontal plane, in the vertical direction, and at some fixed angle with respect to the horizontal plane. Curiously, the motion of a source in this direction is stable with respect to small variations of the direction of the source velocity: the lateral force which appears is of a restoring character. In the case of motion in the horizontal or vertical directions, small variations of the direction of the velocity result in the appearance of a deflecting lateral force, so that such motion is unstable.

The results obtained were extended to three-dimensional motion as well as to the case of a rotating stratified media, where the motion of a source gives rise to the emission of inertial–gravity waves.

Knowing the radiative force makes it possible to estimate the back effect of the emitted waves on the moving sources. This makes it possible to study self-consistent problems and to calculate the trajectories of sources.

This work was supported by the Russian Fund for Fundamental Research (Projects Nos. 96-05-64476 and 96-01-00585), INTAS (Grant No. 94-4057), MTs FPIN (Grant No. 97-2-05), and MOPO (Grant No. 95-0-13.0-69 and others).

We thank Yu. I. Troitskaya for helpful suggestions and discussions.

[1] S. F. Dotsenko, Mor. Gidrofiz. Issled., Sevastopol', No. 1, 22 (1973).
[2] V. P. Goncharov, V. A. Krasil'nikov, and V. I. Pavlov, Izv. Akad. Nauk SSSR, Fiz. Atmos. Okeana **12**, 1310 (1976).
[3] V. P. Dokuchaev and I. S. Dolina, Izv. Akad. Nauk SSSR **13**, 655 (1977).
[4] V. D. Lipovskiĭ, Izv. Vyssh. Uchebn. Zaved., Radiofiz., part 1, **23**, 159 (1980); Izv. Vyssh. Uchebn. Zaved., Radiofiz., part 2, **23**, 655 (1980).
[5] V. D. Lipovskiĭ, Izv. Akad. Nauk SSSR **17**, 1134 (1981).
[6] V. A. Gorodtsov and É. V. Teodorovich, Izv. Akad. Nauk SSSR, Mekh. Zhid. Gaza, No. 2, 77 (1981).
[7] G. I. Grigor'ev, *Emission of Internal Gravity Waves by Mass Sources Moving with Variable Acceleration*, Preprint No. 164, NIRFI, Gor'kiĭ (1983).
[8] V. A. Gorodtsov and É. V. Teodorovich in *N. E. Kochin and the Development of Mechanics* [in Russian], Nauka, Moscow (1984), p. 131.
[9] V. A. Borovikov, Yu. V. Vladimirov, and M. Ya. Kel'bert, Izv. Akad. Nauk SSSR, Fiz. Atmos. Okeana **20**, 526 (1984).
[10] I. V. Sturova in *Nonlinear Problems of the Theory of Surface and Internal Waves* [in Russian], Novosibirsk, Nauka (1985), p. 200.
[11] K. A. Bezhanov, A. T. Onufriev, and A. M. Ter-Krikorov, Izv. Akad. Nauk SSSR, Mekh. Zhid. Gaza, No. 5, 86 (1987).
[12] Yu. A. Stepanyants, I. V. Sturova, and É. V. Teodorovich, Itogi Nauki Tekhniki, Mekh. Zhid. Gaza, VINITI **21**, 93 (1987).
[13] Yu. A. Stepanyants and A. L. Fabrikant, Zh. Éksp. Teor. Fiz. **102**, 1517 (1992) [Sov. Phys. JETP **75**, 818 (1992)].
[14] S. N. Reznik and Yu. I. Troitskaya, Izv. Ross. Akad. Nauk, Fiz. Atmos. Okeana 1, 133 (1996).
[15] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics*, Pergamon Press, N. Y. [Russian original, Nauka, Moscow, 1984].
[16] M. J. Lighthill, *Waves in Fluids*, Cambridge University Press, N. Y., 1978 [Russian trans., Mir, Moscow, 1981].

Translated by M. E. Alferieff

Temperature hysteresis of the change in the cholesteric pitch and surface anchoring in thin planar layers

H. Zink

Laboratorium voor Akoestick on Thermische Fysica, Katholiska Universiteit Leuwen, B-301 Leuwen, Belgium

V. A. Belyakov

L. D. Landau Institute of Theoretical Physics, Russian Academy of Sciences, 117334 Moscow, Russia
(Submitted 30 January 1997)
Zh. Éksp. Teor. Fiz. **112**, 524–536 (August 1997)

Hysteresis of the temperature of jumps in the cholesteric pitch in planar layers of cholesteric liquid crystals (CLCs) has been observed in the temperature dependence of the optical transmission spectra measured in a 4.8 μm thick specimen of a 60% chiral racemic mixture of CE6. The temperature difference for the pitch jumps during heating and cooling was equal to about 0.1 °C near 40.3 °C. No difference in the temperature of the pitch jumps during heating and cooling was observed for an 18 μm thick specimen. A theoretical description of this hysteresis is presented on the basis of a continuum theory of elasticity of CLCs that also takes account of the surface anchoring of the CLC. Different possible mechanisms leading to a change in the director configuration in the layer during a pitch jump and their correspondence to the jump mechanism and hysteresis which occur in the experimental specimens are discussed.
 [S1063-7761(97)01008-1]

1. INTRODUCTION

In previous publications[1–3] we reported optical transmission spectroscopy measurements of the pitch of cholesteric liquid crystals (CLCs) and pitch jumps in thin planar cholesteric layers as a function of temperature and we reported measurements of the deviation of the director orientation at the specimen surface from the alignment direction at the jump temperature. A continuum theory of elasticity of CLCs taking account of surface anchoring gave a quite good quantitative description of the measurements. Subsequent measurements showed temperature hysteresis in the process of the restructuring of the director configuration in a layer in the course of a pitch jump. The hysteresis was observed to depend strongly on the specimen thickness in the measurements. It is also clear from general considerations that the hysteresis phenomena are a direct consequence of the strong anchoring interaction at the surface of the layer.

In the present paper we report on detailed optical transmission spectroscopy measurements of the cholesteric pitch in planar CLC layers of different thickness. A theoretical discussion is presented of the corresponding hysteresis phenomena and the correspondence of different possible mechanisms leading to a rearrangement of the configuration of the director field in the layer in the course of a jump in the pitch of the helix to processes which occur in the experimental specimens. For the mechanism whereby the director at the surface of the layer slips through a potential barrier, the parameters of the temperature hysteresis of the cholesteric pitch are calculated on the basis of a continuum theory of elasticity of CLCs, taking account of the surface anchoring forces for a Rapini model anchoring potential.

2. MEASUREMENT RESULTS

To determine the temperature dependence of the cholesteric pitch we measured with a temperature resolution of 0.01 °C the transmission spectra of thick (18 μm) and thin (4.8 μm) planar specimens consisting of a 60% chiral racemic mixture of CE6 (see Ref. 1, where a more detailed description of the specimens is given). The spectra were measured in the entire temperature interval of existence of the cholesteric phase and the temperature dependence of the parameters of the cholesteric in a layer, including also the temperature variations of the pitch, was determined by comparing the measured spectra with calculations performed on the basis of the approach described in Refs. 4 and 5 .

Thick specimen ($d=18$ μm)

For this specimen the number N of half-turns of the cholesteric helix over the specimen thickness d varies between 67 and 134 in the entire temperature interval of existence of the cholesteric. Figure 1a displays the measured temperature dependence of the Bragg wavelength Λ_B and the pitch p. The values of Λ_B and p were determined from the measured transmission spectra on the basis of a description presented in Refs. 4 and 5 of the optics of cholesteric layers. The filled symbols refer to measurements with increasing temperature and the open symbols refer to measurements with decreasing temperature. Assuming that the temperature dependence of the pitch is the same as that of the Bragg wavelength, the measured values of Λ_B and p, with the exception of points close to the region of the isotropic phase, can be described by the relations[6,7]

$$\Lambda_B=\lambda_0+b(T-T_{sc})^{\nu}, \quad \lambda_0=\lambda_{00}[1+a(T-T_{sc})], \qquad (1)$$

where T_{sc} is the temperature of the smectic–cholesteric

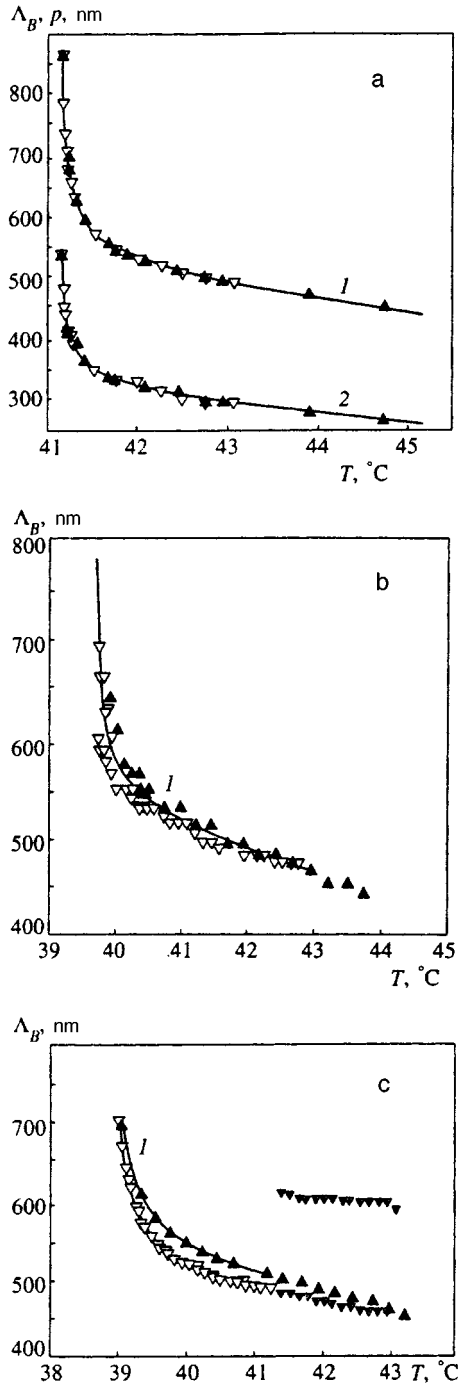


FIG. 1. Temperature dependence of the Bragg wavelength (*1*) and pitch (*2*) for 18 μm (a) and 4.8 μm (b — T_{sc}=39.65 °C, c — T_{sc}=38.98 °C) thick specimens. Filled symbols — for increasing temperature, open symbols — for decreasing temperature. The measurement results for the blue phase are also presented in Fig. 1c (small triangles).

phase transition, ν is the critical exponent, and λ_0 is the Bragg wavelength in the absence of a smectic transition. In the experiment, no difference was observed in the values of the quantities measured with increasing and decreasing temperature. To within the experimental accuracy,[7] T_{sc} has the same value in different series of measurements. The effect of surface anchoring on the value of the pitch and the Bragg wavelength is negligibly small on account of the large number N of half-turns in the layer. For this reason, we take the temperature dependence of the pitch and Bragg wavelength to be the same as in a bulk cholesteric. The temperature dependence of the pitch, displayed in Fig. 1, is described by Eq. (1) with the values of the adjustable parameters presented in Table I.

TABLE I.

Parameters for the specimen with $d=18$ μm	
T_{sc}, °C	41.13
ν	−0.6
λ_{00}, nm	302
a, °C^{-1}	0.042
b	30.05

Thin specimen ($d=4.8$ μm)

For this specimen the number N of half-turns of the cholesteric helix over the thickness d of the specimen ranges between 22 and 40 in the entire temperature range in which the cholesteric exists. Surface effects become noticeable in this case.

Figure 1b displays the temperature dependence of the Bragg wavelength with increasing and decreasing temperature. The main difference from the thick specimen is in the value of the transition temperature, T_{sc}=39.65 °C. The value of T_{sc} is found to vary from one series of measurements on a specimen to another, but it is always lower than in the thick specimen. Taking T_{sc}=39.65 °C and the same values of the remaining parameters as in the thick specimen, we obtained the dependence represented by the solid curve displayed in Fig. 1b. It is found that the deviations of the experimental points from the computed curve are larger than in the case of the thick specimen. The points obtained with increasing temperature lie slightly higher than the points corresponding to decreasing temperature. However, this is not always so: As an example, results in which a large difference is observed in measurements performed with increasing and decreasing temperature and T_{sc}=38.98 °C are displayed in Fig. 1c.

As shown earlier,[2,3] in the general case the measured transmission spectra are described by a theory with a nonintegral number of half-turns of the cholesteric helix over the thickness of the specimen, the difference in the number of half-loops from integers corresponding to the deviation of the director from the alignment direction at the surface by some angle Φ. In the general case, the measured spectra are described by a weighted superposition of two theoretical spectra for $N+\Delta N$ and $N+1-\Delta N$ half-turns, where $\Delta N=2\Phi/\pi$. For example, the data on the temperature evolution of Φ, N, the weighting factor w, and the Bragg wavelengths Λ_B for four temperatures in a series of measurements with increasing temperature show that a pitch jump from $N=28$ to $N=29$ occurs between the temperatures 40.37 °C and 40.39 °C.

A jump from $N=30$ to $N=29$ with decreasing temperature was observed at 40.32 °C and a jump from $N=31$ to $N=30$ was observed at 40.37 °C.[2,3] These results demonstrate the presence of temperature hysteresis in the experimental specimen.

3. TEMPERATURE HYSTERESIS OF THE PITCH JUMP IN THE CONTINUUM THEORY OF ELASTICITY

In the present section we present an analysis of the hysteresis phenomena for temperature variations of the cholesteric pitch in planar CLC layers based on the continuum theory of elasticity taking account of molecular anchoring forces at the surfaces of the layer. The condition found in Ref. 2 for a jump in the cholesteric pitch, resulting in a restructuring of the director configuration from a cholesteric helix with N half-turns to a helix with $N+1$ half-turns over the thickness of the layer, corresponds to the temperature at which the free energies F for these configurations are equal to one another. However, in order for the transition between director configurations to occur, in the course of the structural transformation the system must overcome an energy barrier separating the initial and final states. This is the reason why it is possible to observe hysteresis phenomena in structural transitions with a change in the number of helix half-turns over the thickness of the cholesteric layer. Specifically, hysteresis is manifested as a difference in the temperatures of the pitch jump during heating and cooling of the cholesteric layer.

It is obvious that the corresponding hysteresis phenomena depend primarily on the surface anchoring interaction at the boundaries of the layers. In the absence of surface anchoring there is no temperature hysteresis in pitch changes for a planar layer, since the pitch simply equals its equilibrium value for the bulk cholesteric.

For hysteresis, the mechanism whereby the director configuration in the layer is actually restructured is very important. For example, a transition from a helix with N half-turns to a helix with $N+1$ half-turns over the thickness of the layer can occur continuously by rotations of the director at the surfaces of the layer, resulting in deflections of the director away from the alignment direction at the surfaces. Another possibility is due to the sharp fluctuation-induced variations in the director structure in the volume of the cholesteric layer or some kind of instabilities of the cholesteric helix followed by relaxation of the director structure that arises in the layer to an equilibrium configuration.

In the first possibility the system must overcome a potential barrier separating configurations with N and $N+1$ half-turns of the director in the layer. Such a transition occurs without strong local disturbances of the director configuration at any point over the thickness of the layer when the temperature variations of the parameters of the superheated (supercooled) specimen reduce the potential barrier between the configurations to zero (of course, the transition can be facilitated by temperature fluctuations of the director orientation at the surface of the specimen).

The second possibility presupposes strong local distortions of the cholesteric helix. For this reason, the initial configuration cannot transform by means of smooth and weak distortions into the final configuration. This is why the transformation process in this case cannot be described on the basis of the standard continuum theory of elasticity of cholesterics, where the spatial derivatives of the director field of order no higher than second appear and only in a squared form. The most likely situation is that transformations of this kind are due to the existence of singularities (defects) for some time in the director field of the layer. However, processes of this kind fall outside the scope of the present work. For this reason, in what follows we shall limit the quantitative analysis of the problem to the first case only, i.e., the potential barrier between configurations with N and $N+1$ half-turns of the director in the layer is overcome continuously in the course of the transformation.

We start by finding the temperature behavior of Φ, the deflection angle of the director orientation at the surface of the cholesteric layer from the alignment direction. Following Ref. 2, we write the free energy of the layer in the form

$$F(\Phi)=2W_s(\Phi)+\frac{K_{22}}{2}\left[\frac{2\pi}{p_{N+\Delta N}}-\frac{2\pi}{p(T)}\right]^2 d, \tag{2}$$

where K_{22} is the elastic torsional modulus, $W_s(\Phi)$ is the surface anchoring potential, $p(T)$ is the equilibrium value of the pitch in a bulk CLC, $p_{N+\Delta N}$ is the value of the pitch in the cholesteric layer, $N+\Delta N$ is the number of half-turns of the cholesteric helix over the thickness of the layer (N is assumed to be an integer), and d is the thickness of the layer. As a simplification, we assumed that the alignment direction and the surface anchoring force are identical for both surfaces of the layer and therefore $\Delta N=2\Phi/\pi$.

It follows from Eq. (2) that the angle Φ as a function of temperature, i.e., $\Phi(p(T))$, is determined by the equation

$$\frac{\partial W_s(\Phi)}{\partial\Phi}+K_{22}\left[\frac{2\pi}{p_{N+\Delta N}}-\frac{2\pi}{p(T)}\right]=0. \tag{3}$$

Here it should be emphasized that for a fixed temperature T or, equivalently, for a fixed pitch $p(T)$ there exist several solutions of Eq. (3) which correspond to different numbers N of half-turns in the layer. For infinitely strong surface anchoring the number of the solutions is infinite. For finite anchoring forces the number of possible solutions is finite and limited by the condition $\Phi<\Phi_c$, where the critical angle Φ_c is determined by the form of the anchoring potential $W_s(\Phi)$. The physical meaning of the critical angle Φ_c is very simple — this is the deflection angle of the director from the alignment direction at the surface for which the potential barrier between the configurations with N and $N+1$ half-turns vanishes. Naturally, only one of the solutions with different N corresponds to the minimum value of the free energy (2) and is stable (we denote the corresponding value of N by N_0). All other solutions are metastable, since the corresponding free energies are greater than the minimum value corresponding to N_0. The temperature of the transition between configurations differing by one half-turn in the layer is determined by the temperature (the pitch $p(T)$) for which $\Phi=\Phi_c$. Therefore a study of hysteresis in the present approach reduces to determining the temperatures T for different values of N (or, equivalently, the values of the pitches $p(T)$) that correspond to Φ_c for a pitch jump with N changing by ±1.

The computed temperature dependences of the deflection angle Φ for a Rapini model surface anchoring potential (see, for example, Ref. 8) are displayed in Fig. 2.

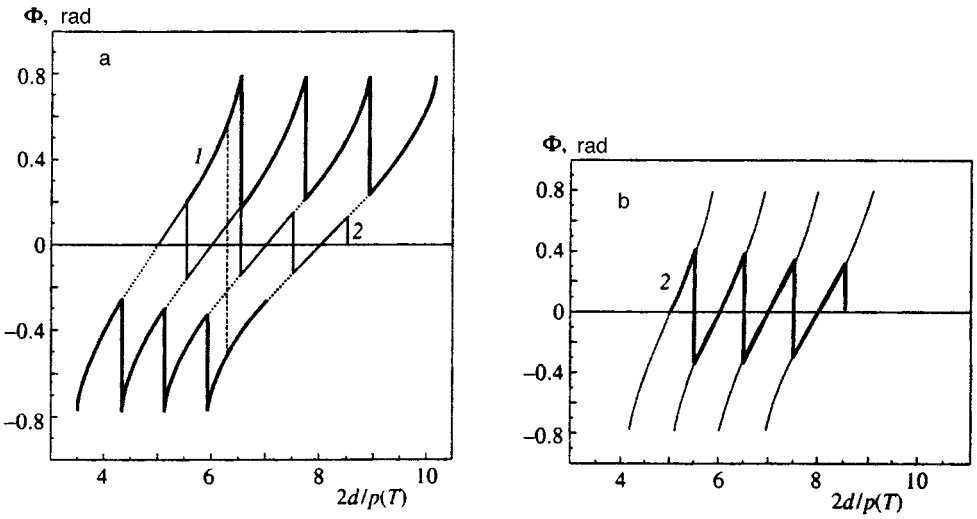


FIG. 2. Computed temperature dependences of the deflection angle of the director on the surface of the layer for strong (a) and weak (b) surface anchorings ($S=5$ and $S=15$, respectively) in the region of equilibrium numbers of half-turns of the helix over the thickness of the layer $N=5$, 6, 7, and 8 for the Rapini surface anchoring potential. The curve above the abscissa refers to the case of increasing temperature and the curve below the abscissa refers to the case of decreasing temperature; *1* — hysteresis present, *2* — hysteresis absent.

The potential barrier between the director configurations in layers with N and $N+1$ half-turns is determined by the expression

$$B(N_0,T)=F(N_0,\Phi_c)-F(N_0,\Phi(T),T), \tag{4}$$

where the equilibrium value of the deflection angle $\Phi(T)$ for temperature T is determined by the solution of Eq. (3) and the free energies $F(\Phi_c,T)$ and $F(\Phi(T),T)$ in Eq. (4) are determined by the expression (2). The temperature dependences of the height of the barrier potential which were calculated for the Rapini anchoring potential are displayed in Fig. 3.

The equations and computational results presented above make it possible to determine the parameters of the jump in the cholesteric pitch during heating and cooling, i.e., to describe quantitatively the temperature hysteresis in the variations of the cholesteric pitch in the layer. The transition temperature for a jump from N to $N+1$ half-turns can be obtained from Eq. (3) by making the substitution $\Phi=\Phi_c$

$$\frac{1}{p(T)_{N\to N+1}}=\frac{1}{2\pi K_{22}}\frac{\partial W_s(\Phi)}{\partial\Phi}\bigg|_{\Phi=\Phi_c}+\frac{1}{2d}\left(N+\frac{2\Phi_c}{\pi}\right). \tag{5}$$

The temperature of the jump from N to $N-1$ half-turns, which is found in a similar manner, equals

$$\frac{1}{p(T)_{N\to N-1}}=\frac{1}{2\pi K_{22}}\frac{\partial W_s(\Phi)}{\partial\Phi}\bigg|_{\Phi=\Phi_c}+\frac{1}{2d}\left(N-\frac{2\Phi_c}{\pi}\right). \tag{6}$$

Thus, the temperature hysteresis for a jump occurring during cooling (from $N+1$ to N half-turns) relative to a jump occurring during heating (from N to $N+1$ half-turns) is determined by the relation

$$\frac{1}{p(T)_{N\to N+1}}-\frac{1}{p(T)_{N+1\to N}}=\frac{1}{\pi K_{22}}\frac{\partial W_s(\Phi)}{\partial\Phi}\bigg|_{\Phi=\Phi_c}+\frac{1}{2d}\left(\frac{4\Phi_c}{\pi}-1\right). \tag{7}$$

In obtaining the relation (7), we made the natural assumption that the anchoring potential W_s is an even function of the angle Φ. This assumption leads to the following equation for the critical angle Φ_c:

$$W_s(\Phi_c)=\frac{1}{2}W_s\left(\frac{\pi}{2}\right)=\frac{1}{2}W_{\max}, \tag{8}$$

where $W_{\max}$ means the maximum value of the anchoring potential.

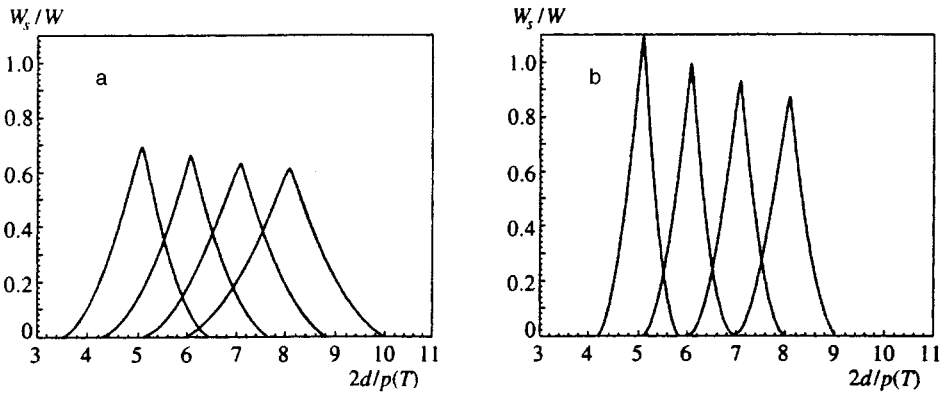


FIG. 3. Computed temperature dependence of the height of the potential barrier separating the director configurations in a layer for a Rapini surface anchoring potential. The director configurations differ by one half-turn in the region of the equilibrium numbers of half-turns of the helix over the thickness of the layer $N=5$, 6, 7, 8 for strong (a) and weak (b) surface anchorings ($S=5$ and $S=15$, respectively).

The jump partially removes the stress in the cholesteric helix before the jump, and for this reason after the jump the deflection angle Φ decreases and satisfies the next condition at the jump point: $|\Phi_{N\pm1}|<|\Phi_c|$. If there is no hysteresis, then $\Phi_{N\pm1}=-\Phi_j$, where Φ_j is the deflection angle at the jump point.[6] It is helpful to note that the expressions presented above do not involve any assumptions about the form of the anchoring potential (except that it is an even function of the deflection angle). Therefore, under conditions when the approach developed above is applicable for describing the experimental results on hysteresis phenomena in pitch jumps in a planar layer, the equations derived make it possible to obtain direct model-independent information about the surface anchoring potential $W_s(\Phi)$.

4. HYSTERESIS IN THE RAPINI MODEL POTENTIAL

To make quantitative theoretical predictions it is necessary to have an explicit form of the anchoring potential. In what follows, the general relations obtained in the preceding section are employed to describe hysteresis in the Rapini model for the anchoring potential, i.e., in the above relations the anchoring potential was given by the relation $W_s=-W\cos^2(\Phi/2)$.

The results of the corresponding calculations of the temperature dependence of Φ — the deflection angle of the director from the alignment direction at the surface of the layer — are presented in Fig. 2 for several values of the number N of half-turns of the director across the layer. Figure 3 displays the computed temperature dependences of the potential barrier between the director configurations in the layer which differ by one half-turn of the helix in the layer. The computed curves make it possible to determine the hysteresis parameters as well as the temperature behavior of the cholesteric pitch in the layer in the absence of hysteresis. To make the computational results more general and applicable to different specific experimental situations, the deflection angle Φ is presented in the form of a function of $2d/p(T)$, where $p(T)$ is the equilibrium value of the cholesteric pitch at temperature T for a bulk cholesteric. If the temperature dependence $p(T)$ is known (see, for example, Fig. 1), then the curves can be easily represented in the form of the dependence $\Phi(T)$. The admissable values of Φ in the figures are bounded by the condition

$$-\frac{\pi}{4}<\Phi<\frac{\pi}{4}, \tag{9}$$

since it follows from the formula (8) for the Rapini potential that $\Phi_c=\pi/4$.

As one can see from Fig. 2, the temperature behavior of Φ and hysteresis depend strongly on the value of the parameter $S=4\pi K_{22}/Wp$ and the thickness of the specimen. Specifically, for $\Phi=\Phi_c$, i.e., at the jump point, the difference between the number $2d/p(T)$ of half-turns across the layer for the equilibrium value of the cholesteric pitch and N half-turns in an unstressed director configuration (for $\Phi=0$), i.e., $\Delta N_j=N-2d/p(T)$, found from Eq. (3) is described by the expression

$$|\Delta N_j|=\frac{Wd}{2\pi K_{22}}+\frac{1}{2}=\frac{N}{S}+\frac{1}{2}. \tag{10}$$

We note that if the temperature dependence of the pitch $p(T)$ is known from experiment (see, for example, Fig. 1), then the equation (10) can be used to determine the temperature of the pitch jump in the slip mechanism for the director to overcome the surface anchoring potential barrier.

Substituting into Eq. (7) the Rapini anchoring potential we obtain for the jump hysteresis according to the number of half-turns

$$\Delta N_h=2d\left[\frac{1}{p(T)_{N\to N+1}}-\frac{1}{p(T)_{N+1\to N}}\right]=\frac{Wd}{\pi K_{22}}=2\frac{N}{S}. \tag{11}$$

For the Rapini potential the expression (4) for the height of the potential barrier with positive values of the deflection angle of the director, determined from the relation (3), assumes the form

$$\frac{B(T)}{W}=\frac{1}{2}\left\{\cos(2\Phi(T))+\frac{\pi S}{2N}\left[1-\frac{4\Phi(T)}{\pi}\right]\times\left[N+\frac{1}{4}+\frac{\Phi(T)}{\pi}-\frac{2d}{p(T)}\right]\right\}. \tag{12}$$

For strong anchoring (in the calculations $S=5$), the curves of the temperature dependence of the deflection angle Φ show that if the pitch jump occurs by the mechanism whereby the director on the surface slips through a potential barrier (when Φ reaches the value $\pi/4$), then after the jump the director configuration in the layer does not correspond to the configuration with the minimum free energy for the jump temperature. Therefore this configuration is strongly distorted, since $|\Phi_j^a|>|\Phi_j|$ holds, where Φ_j is the deflection angle at the jump point in the absence of hysteresis and Φ_j^a is the deflection angle of the director actually realized after the jump.

For an intermediate anchoring force (in the calculations $S=10,15$), as a result of the director slipping through the barrier the director configuration can be transferred into a state with a minimum free energy ($|\Phi_j^a|<|\Phi_j|$) in the case when the number of half-turns of the cholesteric helix over the thickness of the layer is sufficiently small. However, for large N the director configuration in the layer once again becomes metastable ($|\Phi_j^a|>|\Phi_j|$).

Analysis of the hysteresis phenomena on the basis of the present mechanism whereby on the surface the director slips through a barrier for the parameters of the problem which correspond to the experiment performed ($S\approx20$, $N\approx30$) leads to the following results. The difference between the number of half-turns across the layer with $\Phi=\Phi_c$ for the equilibrium value of the pitch and N, i.e., ΔN_j, should be of order 3, and after the jump, when $\Phi=\Phi_j^a$, it should equal approximately 2. This means that in the present jump mechanism, in which the director at the surface slips through a barrier, the director configuration in the layer is strongly distorted and metastable not only immediately before but also after the pitch jump.

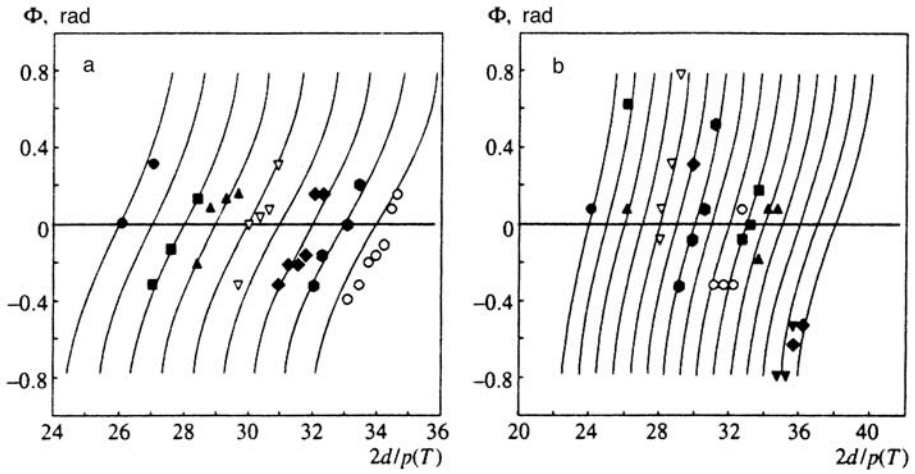


FIG. 4. Temperature dependence of the deflection of the director on the surface of the layer from the alignment direction (different values correspond to different series of measurements) measured with cooling (a) and heating (b) of 4.8 μm thick specimen and calculated for $S=25$.

5. DISCUSSION

Definite conclusions concerning the mechanism leading to restructuring of the director configuration in the layer in the course of the experimentally observed pitch jumps can be drawn by comparing the experimentally observed hysteresis with the results of the theoretical analysis presented above. As the calculations for large N or even for small N with strong surface anchoring show, the continuum theory predicts that in addition to a stable configuration of the director several metastable configurations differing by the number N of half-turns of the cholesteric helix can also exist in the layer (see, for example, in Fig. 2a the intersection of the calculated curves by a straight line parallel to the y axis near $2d/p(T)\approx 6$).

If such metastable configurations were present in the layer for a sufficiently long time, the experimental spectra would consist of a superposition of more than two spectra with different values of N. Moreover, structural changes should be observed in the director configuration in a layer with N changing by more than 1. However, the experimental measurements performed do not show that these possibilities are realized.

Additional conclusions about the mechanism of the observed jumps can be drawn by comparing the calculated curves with the prejump temperature dependences of the deflection angle Φ of the director at the specimen surface from the alignment direction. The fit of the theoretical curves $\Phi(T)$ with respect to the parameter S for temperatures preceding the pitch jump to the measured values of this quantity (Fig. 4) demonstrates a reasonable agreement between theory and experiment for $S=25$ during both cooling and heating of the specimen. If the value of S found in this manner from experiment is used to calculate hysteresis, for example, for $N=30$, then we obtain from Eq. (11) $2d/p(T_j^h)-2d/p(T_j^a)=2.4$, where $p(T_j^h)$ and $p(T_j^a)$ are the equilibrium values of the pitch of a bulk cholesteric for the temperatures of a jump during heating and cooling, respectively. An estimate of the same quantity from the measured values of the temperatures of the pitch jumps and the temperature changes of the equilibrium pitch in a bulk cholesteric (see Fig. 1b) gives for a 4.8 μm thick specimen a much smaller value, equal to only 0.5. This means that the observed hysteresis is much smaller than predicted on the basis of the continuum theory for the mechanism whereby on the surface the director slips through a potential barrier.

Thus, it must be concluded that in the specimen studied the restructuring of the configuration of the director in a layer at the temperature of a pitch jump does not occur by means of continuous changes in the orientation of the director at the surface in the absence of strong disturbances of the cholesteric helix in the volume of the layer. This means that some kind of fluctuations of the director in the layer or instability of the director configuration prevent the deflection angle Φ of the director from reaching its critical value Φ_c ($\pi/4$ for the Rapini potential). It should be expected that these factors influence the appearance of hysteresis in the pitch jump especially effectively if several metastable configurations of the director with $|\Phi_j^a|<|\Phi_c|$ exist in the layer.

We shall now briefly discuss some possible physical mechanisms leading to temperature hysteresis in the jump of a cholesteric pitch for the hysteresis parameters observed in the present work. The conclusion drawn above agrees with general physical ideas about the mechanism of structural transformations. From general considerations it follows that if the energy of the system undergoing a structural transition is insufficient for overcoming the potential barrier separating the states drawn into the transition, then the transition mechanism may either be fluctuations or some type of instability.

Let us begin with a discussion of the effect of fluctuations on the pitch jump mechanism examined above, i.e. with a clarification of how fluctuations of the orientation of the director at the surface of the layer can alter the theoretical results presented above. It is obvious that director fluctuations can make it easier for the director to overcome the potential barrier but they do not change the foregoing picture of the restructuring of the director configuration in the layer (excluding, possibly, the case of very weak surface anchoring). Thus, in most cases the fluctuations of the director orientation at the surface will lead only to a change in the jump temperature (in the direction of decreasing hysteresis). Such fluctuations of the director can be taken into account quantitatively by a method similar to that used in Refs. 8 and 9 to take account of fluctuations of the director in nematic layers.

Fluctuations of the director in the volume of the layer can lead to results which are qualitatively different from those examined above. Such fluctuations can initiate the development of an instability of the cholesteric helix in a layer with very strong local disturbances in the director field (most likely with the formation of defects in the director field). As such instabilities grow, they can assist the director configuration in the layer to reach its equilibrium form.

To get an idea of the type of director fluctuations in a layer that can most effectively restructure the director configuration in the layer in the course of a pitch jump, we shall consider several types of infinitesimal disturbances of the configuration of the director field in the layer. Since the director configuration in the final stable state reached as a result of a pitch jump is known, it is entirely natural to consider director fluctuations that can most effectively lead to the known configuration of the stable final state. We know from the analysis performed above that the stable final director configuration reached as a result of a pitch jump is modulated with a period different from both the equilibrium pitch $p(T)$ of the bulk cholesteric and the period determined by the number of half-turns in the initial state, i.e. $2d/N$. For this reason, let us see how the free energy of the metastable initial configuration of the director in the layer changes when an infinitesimal disturbance of the form $\Delta\cos(qz-\Phi_0)$, where Δ is a small quantity and q is the wave vector of the modulation, is imposed on it. Substituting this disturbance into the expression for the free energy (2) shows that the changes produced in the free energy by such a disturbance in the approximation linear in Δ can be both positive and negative, depending on the ratios of q and Φ_0 and the corresponding parameters describing the disturbed metastable state. It is obvious that a negative change of the free energy does not mean that the metastable configuration will be destroyed as a result of the corresponding disturbance, since this question cannot be resolved in the linear approximation. Nonetheless, a negative change in the free energy can be regarded as an indication of the fact that fluctuations of the corresponding type can be substantially intensified.

The foregoing analysis shows that if some scalar quantity characterizing the cholesteric, for example, the elastic modulus K_{22}, is modulated, then the free energy of the disturbance does not depend critically on the value of the modulation wave vector q. However, if the local orientations of the director are modulated, for example, by spatial modulation of the pitch, then a critical value of the modulation vector $q=4d/N$ (which corresponds to a modulation period equal to approximately half the cholesteric pitch in the layer) does appear. The changes in the free energy for this value of the modulation period grow more rapidly than for other periods.

Thus, it can be conjectured that fluctuations of this type in the volume promote a transition of the metastable configuration of the director in the layer into a stable state. In this case, the extra (or missing) number of director turns over the thickness of the layer can be eliminated (accumulated) in the volume of the layer as a result of strong changes in the configuration of the director field on a spatial scale of the order of the cholesteric pitch in the layer. If this pitch jump mechanism is indeed realized in the experimental specimens, then scattering of light by fluctuations with the indicated wave vector should be intensified in them at a temperature immediately preceding the jump in the cholesteric pitch in the layer.

6. CONCLUSIONS

The hysteresis observed in the present work in the temperature of the cholesteric pitch jumps and the theoretical analysis of the problem performed in this connection do not answer all questions concerning the temperature hysteresis in pitch changes for thin planar cholesteric layers. However, for the experimental specimens studied they permit eliminating from possible mechanisms of the pitch jump the one whereby the director on the surface of the layer slips through the potential barrier of the anchoring forces. The analysis performed above showed that the pitch jump mechanism due to slipping of the director through a potential barrier on the surface is most realistic for thin specimens and weak surface anchoring forces. We emphasize that a study of the temperature dependence of the pitch and its hysteresis under conditions when slipping of the director through a surface force barrier occurs can be used for direct experimental investigation of the potential of the surface anchoring forces without invoking any ansatz about the form of this potential.

The mechanism responsible for the pitch jump in the experimental specimens requires further investigation. Specifically, it is necessary to study the pretransition fluctuations in the specimen and to determine the nature of the superposition (observed in the transmission spectra) of two director configurations differing by one half-turn of the cholesteric helix over the thickness of the layer. For example, there arises the question of whether or not the observed superposition is a consequence of temporal fluctuations or the spatial nonuniformity of the director configuration over the surface of the layer. On the whole, the results presented above demonstrate the advantages of the new optical spectroscopic approach used in the present work for studying surface anchoring in cholesterics.

This work is supported by the Russian Fund for Fundamental Research (Project No. 97-02-16505).

[1] H. Zink and V. A. Belyakov, Mol. Cryst. Liq. Cryst. **265**, 445 (1995).
[2] H. Zink and V. A. Belyakov, JETP Lett. **63**, 43 (1996).
[3] H. Zink and V. A. Belyakov, Mol. Cryst. Liq. Cryst. **282**, 17 (1996).
[4] V. A. Belyakov and V. E. Dmitrienko, Sov. Sci. Rev. A Phys., edited by I. M. Khalatnikov, **13**, 1 (1989).
[5] V. A. Belyakov, *Diffraction Optics of Complex-Structured Periodic Media*, Springer-Verlag, N. Y., 1992, Chap. 4 [Russian original, Nauka, Moscow, 1988].
[6] R. S. Pindak, C. H. Huang, and J. T. Ho, Phys. Rev. Lett. **32**, 43 (1974).
[7] H. Zink and W. Van Dael, Liq. Cryst. **14**, 603 (1993).
[8] L. M. Blinov and V. G. Chigrinov, *Electrooptic Effects in Liquid Crystal Materials*, Springer-Verlag, N. Y. (1994), Chap. 3.
[9] B. Ya. Zel'dovich and N. Tabiryan, Zh. Eksp. Teor. Fiz. **81**, 1738 (1981) [Sov. Phys. JETP **54**, 922 (1981)].
[10] T. Ya. Marusiĭ, Yu. A. Reznikov, V. Yu. Reshetnyak *et al.*, Zh. Éksp. Teor. Fiz. **91**, 851 (1986) [Sov. Phys. JETP **64**, 502 (1986)].

Translated by M. E. Alferieff

Tunneling conductivity oscillations in a magnetic field in metal–insulator–narrow-gap-HgCdTe structures: the energy spectrum and spin–orbit splitting of $2D$ states

G. M. Min'kov, O. É. Rut, and A. V. Germanenko

Institute of Physics and Applied Mathematics, Ural University, 620083 Ekaterinburg, Russia
(Submitted 12 July 1996; resubmitted 31 December 1996)
Zh. Éksp. Teor. Fiz. **112**, 537–550 (August 1997)

We study tunneling conductivity oscillations in a magnetic field in narrow-gap p-HgCdTe–oxide–metal (Yb, Al) structures. In tunnel structures with Yb we detect two types of tunneling conductivity oscillations. The first is related to the crossing of the Landau levels of two-dimensional ($2D$) states localized in the surface quantum well of the semiconductor, and has an energy E_F+eV, where E_F is the Fermi energy of the semiconductor and V is the bias voltage; the second has an energy E_F. We find that in such structures with an asymmetric quantum well there is strong spin–orbit splitting in the spectrum of the $2D$ states. In p-HgCdTe–oxide–Al tunnel structures the surface potential is much weaker and only oscillations of the first type are observed. We find that in such structures there is only one spin state of the $2D$ carriers, while the second is pushed into the continuous spectrum because of strong spin–orbit coupling. To analyze the experimental results we calculate the spectrum of $2D$ states localized in the surface quantum well in a semiconductor with a Kane dispersion law. We find that all the experimental results are in good agreement with the results of calculations. Finally, we discuss the features of ''kinematically coupled'' states in an asymmetric quantum well.
 [S1063-7761(97)01108-6]

1. INTRODUCTION

The tunneling between the electron systems of different dimensionality, $2D-3D$,[1–3] $2D-2D$,[4,5] etc. has lately attracted much attention. On the one hand, the interest is due to the prospects of developing a new generation of semiconductor nanostructures and devices based on these structures that employ such phenomena as resonant tunneling and Coulomb ''blockade.'' On the other hand, research into tunneling conductivity and its dependence on voltage, temperature, magnetic field strength, and other external agents provides a unique possibility for studying the details of the energy spectrum of the current carriers of reduced dimensionality,[1,2] the role of electron–electron interaction,[6,7] and the special features of tunneling processes.

This paper presents the results of investigations into the energy spectrum of $2D$ states, either occupied by $2D$ current carriers or vacant, localized in a surface quantum well at the narrow-gap p-HgCdTe–oxide junction. A specific feature of $2D$ states of electrons in narrow-gap materials is that even at low concentrations of $2D$ electrons their average energy is at least of the order of the energy gap, i.e., the interaction of the nearest bands is strong (this means that the electron wave function is multicomponent), and in calculating the spectrum of $2D$ states we must be sure to take this into account.[8]

Another feature of the these structures is that the surface quantum well is asymmetric. This lifts the spin degeneracy of the $2D$ states for finite longitudinal quasimomenta even in the absence of an external magnetic field.[9] The splitting is the result of spin–orbit coupling with the smooth part of the potential and with the semiconductor–insulator interface, and only by using the multiband model of the energy spectrum can we analyze both contributions. Generally speaking, the crystal lattice of A^3B^5 and A^2B^6 semiconductors has no inversion center, a feature that can also lift spin degeneracy,[10,11] but in narrow-gap materials the splitting of the spectrum of light particles (electrons or light holes) that results from this asymmetry is small compared to energies accessible in experiments.

The main experimental data on spin–orbit splitting in the spectrum of $2D$ states have been obtained primarily by studying the Shubnikov–de Haas oscillations,[12,13] from the de Haas–Van Alphen effect,[14] and by employing voltage–capacitance spectroscopy methods.[15,16] These approaches make it possible to do measurements over a broad range of $2D$ carrier concentrations but provide information only about states at the Fermi level.

The method of tunneling spectroscopy in a quantizing magnetic field, which we used, makes it possible to study the spectrum of $2D$ states not only at the Fermi energy E_F of the semiconductor but within a broad range extending above and below the Fermi energy. The possibilities of tunneling spectroscopy are extended when a modified method proposed in Ref. 17 is used.

The first to investigate $2D$ states by the method of tunneling spectroscopy in a quantizing magnetic field was Tsui,[18,19] who studied, within a broad range of bias voltages, the tunneling conductivity oscillations in n-InAs in a magnetic field, i.e., oscillations related to the tunneling to the ground and excited subbands of $2D$ states and determined by the corresponding effective masses of the carriers and their energy dependence. However, in InAs the spin-to-orbital splitting ratio for the electron states is moderate ($\hbar\omega_c/g\mu_B B\approx 0.18$), and no effects associated with spin were found in these experiments. Band-to-band tunneling in metal–insulator–semiconductor (InSb, InAs) structures with an inversion layer at the surface was studied by Muller and

Kunze,[20,21] who observed a section of negative differential resistivity in these structures and tunneling conductivity oscillations in a magnetic field. But all this research revealed no tunneling conductivity effects associated with spin, despite the fact that in InSb, the spin-to-orbital splitting ratio for the electron states is close to 0.5 ($\hbar\omega_c/g\mu_B B \approx 0.34$).

In this paper we present the results of research into tunneling conductivity oscillations in tunnel structures manufactured from a narrow-gap semiconductor p-HgCdTe. We have found that in tunnel structures with a strong surface potential that attracts electrons, the tunneling conductivity oscillations in a magnetic field perpendicular to the surface are related to tunneling to size-quantized states localized in the surface quantum well. We have also found that the asymmetry of the quantum well leads to strong spin–orbit coupling of the spectrum of $2D$ states, so that over a certain small range of the potential, only one of the branches of $2D$ states can be split by spin–orbit coupling.

2. EXPERIMENTAL DETAILS

We studied the differential conductivity and its voltage derivative as functions of the magnetic field strength and voltage in p-HgCdTe–oxide–metal (Yb, Al) tunnel structures. The tunnel junctions were prepared from single-crystal specimens of p-$Hg_{1-x}Cd_xTe$ ($0.17<x<0.2$) with a concentration of uncompensated acceptors $N_A-N_D=(0.5-2)\times10^{18}$ cm^{-3}. The tunnel-transparent oxide was formed by irradiating the specimen, pre-etched in a 10% solution of bromine in butanol, with UV radiation for 10 to 15 minutes. After that Yb or Al was spray-deposited on the specimen through a mask, followed by Pb. The use of ytterbium, a metal with a small work function, made it possible to manufacture tunnel junctions with a deep surface quantum well at the p-HgCdTe–oxide interface. In tunnel structures with Al the surface potential well is much shallower.

For each tunnel structure that was studied, we determined the band gap (E_g) and the dispersion law for the bulk states of the semiconductor from the tunneling conductivity oscillations in a magnetic field $\mathbf{B}\perp\mathbf{n}$, where $\mathbf{n}$ is a unit vector normal to the tunnel junction. In this orientation the magnetic field does not quantize the spectrum of $2D$ states, and the tunneling conductivity oscillations are related to the tunneling to the Landau levels of the bulk states. The method is described in detail in Refs. 22 and 23.

In this paper we list the results of experiments involving three tunnel structures: two Yb–oxide–HgCdTe structures with $E_g=50$ meV (structures 1 and 2), and one structure Al–oxide–HgCdTe with $E_g=45$ meV (structure 3). The concentration of uncompensated acceptors in these structures amounted to 8×10^{17} cm^{-3}. At such a concentration, which is close to the Mott transition, the Fermi level is at the top of the valence band,[24] so that in a comparison with the results of calculations we will assume that $E_F=0$ (here and in what follows the energy is measured from the top of the valence band in the bulk of the semiconductor). The results obtained for other tunnel structures were found to be similar.

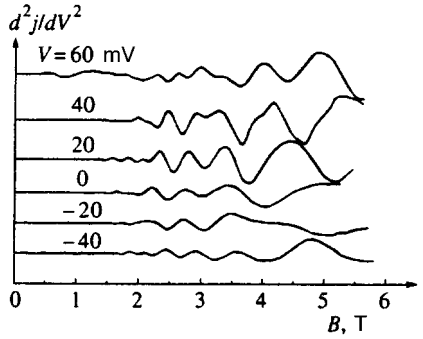


FIG. 1. Oscillations of d^2j/dV^2 in a magnetic field $\mathbf{B}\|\mathbf{n}$ under different bias voltages for structure 1 ($T=4.2$ K).

3. TUNNELING CONDUCTIVITY OSCILLATIONS IN *p*-HgCdTe–OXIDE–Yb STRUCTURES

The typical curves representing the dependence of d^2j/dV^2 on the magnetic field strength for structure 1 in the orientation $\mathbf{B}\|\mathbf{n}$ are depicted in Fig. 1 for several values of the bias voltage. We see that oscillations exist both at positive bias voltages, which corresponds to an upward shift in the Fermi quasilevel of the metal by eV with respect to the Fermi quasilevel of the semiconductor (Fig. 2a), and at negative bias voltages. The observed oscillations are linear combinations of several types of oscillation, which can be separated by performing a Fourier analysis of d^2j/dV^2 as a function of $1/B$ (Fig. 3). We see that at large voltage biases the Fourier spectra exhibit four resolvable maxima. As the bias voltage decreases, these maxima move close to each other in such a way that at $V=0$ they merge into two maxima.

The dependence of the position of the fundamental magnetic fields $B_f=(\Delta(1/B))^{-1}$, where $\Delta(1/B)$ is the period of oscillations in the reciprocal magnetic field, on the bias voltage for structure 1 over the entire bias-voltage range is depicted in Fig. 4. Clearly, there are two different types of oscillation. First, in type I the fundamental magnetic fields increase with increasing bias voltage, while in type II they

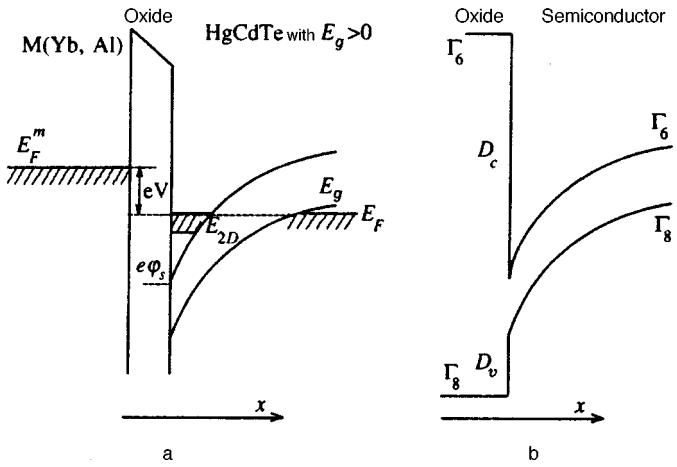


FIG. 2. (a) Energy diagram of a tunnel junction with a bias voltage V applied. (b) Energy diagram of the insulator–semiconductor structure used in the calculations.

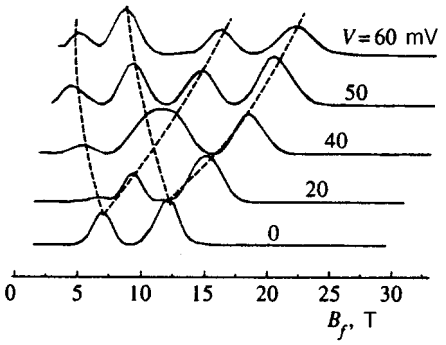


FIG. 3. Fourier spectra of tunneling conductivity oscillations in structure 1 at different bias voltages in a magnetic field **B**‖**n**. The dashed curves indicate the shifts of the maxima.

decrease. Second, in each structure studied, we observed two branches of each oscillation type, which we denote by k_{I}^{+} and by k_{I}^{-} for type I and k_{II}^{+} and k_{II}^{-} for type II. The characteristic curve representing the dependence of the fundamental magnetic fields on the angle θ between **B** and **n** are close to $1/\cos\theta$, and indicate that all observed oscillations for **B**‖**n** are related to the quantization of the spectrum of $2D$ states, i.e., to tunneling from the metal to the Landau levels of the $2D$ states when $V>0$, and from the Landau levels of the $2D$ states to the metal when $V<0$.

Similar oscillations were observed in our studies[25] of tunneling junctions manufactured from gapless p-HgCdTe. In that paper we discussed in detail the possible tunneling conductivity oscillations in a magnetic field in MIS structures with $2D$ electrons localized in a surface quantum well. We found that when the depth of the potential well (and hence the energy of the $2D$ states) depends on the applied voltage, two types of oscillation emerge, one when the Landau levels of the $2D$ states cross the Fermi level of the metal (type I), and the other when the Landau levels of the $2D$ states cross the Fermi level of the semiconductor (type II).

We observed these types in our structures (Fig. 4). The oscillations of both types are periodic in the reciprocal magnetic field, and it is precisely for this reason that they separate so well under Fourier transformations of d^2j/dV^2 vs. $1/B$ dependences; in a semiconductor with an isotropic dispersion law, their period is determined by the value of the longitudinal (i.e., parallel to the junction plane) quasimomentum k at an energy $E=E_F+eV$ for type-I oscillations and at an energy $E=E_F$ for type-II oscillations:

$$B_f^{-1}(V)=\frac{2e}{c\hbar k^2(E)}. \tag{1}$$

Knowing the values of the fundamental fields of type-II oscillations, we can find the concentration of $2D$ electrons:

$$n_{2D}=n^{+}+n^{-}=\frac{(k_{\mathrm{II}}^{2})^2+(k_{\mathrm{II}}^{-})^2}{4\pi}.$$

Thus, the decrease in B_f with increasing bias voltage (Fig. 4) corresponds to a decrease in the concentration of the $2D$ electrons as the potential well becomes shallower. This is why the dependence of the period of type-I oscillations on the bias voltage (Fig. 4) does not directly yield the dispersion law for $2D$ states, since for each bias voltage, k corresponds to a potential well with a specific depth.

Note that type-I tunneling conductivity oscillations are observed even at negative bias voltages, down to $V=-40$ mV. In this bias-voltage range the oscillations are related to tunneling from $2D$ states against the valence-band background (Fig. 2a), and their presence is proof that the ''resonant'' smearing is not strong enough for the oscillation pattern to become entirely blurred. The calculations of ''resonant'' smearing done by Sobkowicz[8] show that in this situation the ''smearing'' is indeed weak.

More information about the spectrum of $2D$ states can be obtained by using a modification of the tunneling spectroscopy method proposed in Ref. 17. The modification allows, by sending short pulses through the system, to change the charge of the centers localized at the barrier or at the semiconductor–insulator interface, which makes it possible to change the surface potential φ_s and determine k_i^2 for fixed energies E_F+eV and E_F, but for different values of φ_s. The dependence of $k_{\mathrm{I}}^{+,-}$ and $k_{\mathrm{II}}^{+,-}$ on the pulse amplitudes at a 45-mV bias voltage is depicted in Fig. 5. We see that positive pulses decrease the values of k_i, while negative pulses increase k_i. The reason is that by reducing the depth of the potential well, positive pulses raise the bottom of the $2D$ subband and hence reduce the concentration of $2D$ electrons, which determines the periods of type-II oscillations, and also reduce the energy E_F+eV by k_i^2, which means that the periods of type-I oscillations are reduced as well. Figure 6 depicts the dependence of k_i^2 at $V=45$ mV on the concentration of $2D$ carriers, which is changed by inputting pulses for two tunnel structures that differ in the initial values of the surface potential. We see that in structure 2, at a minimum concentration of the $2D$ carriers (a condition achieved at the maximum amplitude of positive pulses), there is only one frequency of type-II oscillations. As we will shortly see, at such a surface potential and at an energy equal to E_F, this corresponds to the cases in which there is only one spin state of the $2D$ carriers.

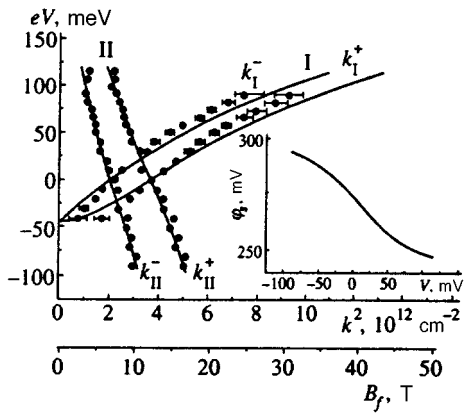


FIG. 4. Fundamental fields of tunneling conductivity oscillations for structure 1 as functions of the bias voltage: subscript I corresponds to type-I oscillations and subscript II, to type-II oscillations; k^{+} and k^{-} correspond to oscillations related to tunneling to different spin states. The solid curves represent the results of calculations described in the main text. Inset shows the dependence of the surface potential on the bias voltage.

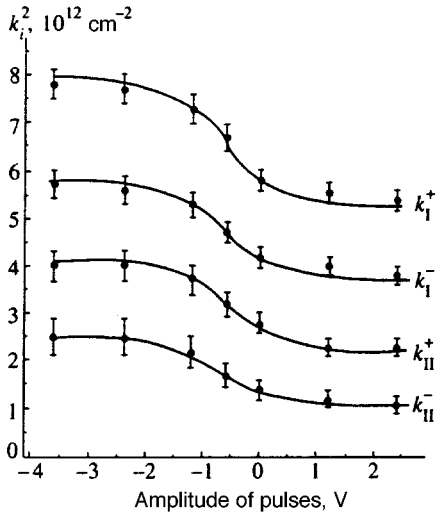


FIG. 5. Quasimomenta at a bias voltage of 45 mV as functions of the amplitude of pulses 10-μs long with a repetition rate of 20 Hz (structure 1).

4. ANALYSIS OF EXPERIMENTAL RESULTS

To analyze the experimental results we adopt the two-band Kane model, which provides a good description of the spectrum of bulk states of HgCdTe (the distance from band Γ_8 to band Γ_7 in HgCdTe is 1 eV, which is much larger than the band gap and the energy of the $2D$ states, so that we ignore the contribution of the interaction with the Γ_7 band). To calculate the spectrum of $2D$ states in this model we must select an insulator model. We believe the best model is the one proposed by Sobkowicz,[8] who assumed that the insulator has the same band structure as the semiconductor (i.e., a valence band Γ_8 and a conduction band Γ_6) but a larger band gap (Fig. 2b). The model contains two parameters characterizing the semiconductor–insulator interface, D_c and D_v, which are the discontinuities in the Γ_6 and Γ_8 bands, respectively. A detailed description of the spectrum of $2D$ states can be found in Refs. 25 and 26. Research into gapless HgCdTe has shown that the most suitable values of the parameters D_c and D_v for the HgCdTe–oxide interface are 2 and 1 eV, respectively.[25] We use the same values when calculating the spectrum of $2D$ states in HgCdTe with $E_g>0$.

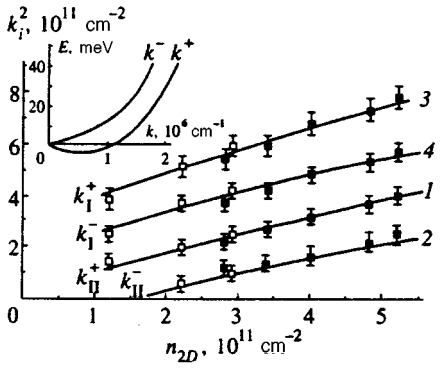


FIG. 6. Quasimomenta (at a bias voltage of 45 mV) determined from the periods of oscillations of type I and II, as functions of the concentration of $2D$ electrons for structure 1 (■) and for structure 2 (□). The solid curves represent the results of calculating the quasimomentum of $2D$ states at $E=E_F$ (curves *1* and *2*) and $E=E_F+eV$ (curves *3* and *4*) (see the description in the main text). Inset shows the dispersion law for the $2D$ states at a surface potential corresponding to a concentration $n_{2D}=1.2\times10^{11}$ cm^{-2} of the $2D$ carriers.

To calculate the spectrum of $2D$ states we must also know the surface potential $\varphi(x)$, which can be found by solving the Poisson equation,

$$\frac{d^2\varphi}{dx^2}=-\frac{e}{\kappa\kappa_0}\left((N_A-N_D)\,\vartheta(L-x)+\sum_{2D}|\Psi(x)|^2\right), \quad (2)$$

where κ is the dielectric constant, L is the width of the region occupied by the space charge, $\vartheta(L-x)$ is the Heaviside step function, $\Psi(x)$ is the $2D$ electron wave function, and summation is over all occupied $2D$ states.

Figure 6 shows that, depending on the bias voltage, the $2D$ electron concentration in the structures studied varies within the range $n_{2D}=(1-5)\times10^{11}$ cm^{-2}, and the charged acceptor concentration in the region with the space charge, within the range $(N_A-N_D)L\approx(0.8-1.5)\times10^{12}$ cm^{-2}, so that to a first approximation we can ignore the first term on the right-hand side of Eq. (2). In this case $\varphi(x)=\varphi_s(1-x/L)^2$, where φ_s is the value of the potential at the boundary, and

$$L=\left(\frac{2\kappa\kappa_0\varphi_s}{e(N_A-N_D)}\right)^{1/2}.$$

The results of numerical calculations in which $\varphi_s(V)$ is the only adjustable parameter are depicted in Fig. 4 by solid curves, where the inset shows the φ_s vs. V dependence used in the calculations. We see that for all bias voltages, the calculated values of the longitudinal quasimomenta at the energies E_F+eV and E_F are in good agreement with values found from the periods of type-I and type-II oscillations, respectively. Here the two branches on each oscillation type correspond to the spin-split ground $2D$ subband. One would assume that the two branches correspond to the ground and excited $2D$ subbands. However, our calculations have revealed that if we assume this to be true, the ratio of the quasimomenta of the two branches at fixed bias voltage is considerably higher than the experimental value of this ratio. For instance, at $eV=0$ the calculated ratio is eight and not two, which is the value yielded by the experiment (Fig. 4).

The results obtained by the modified tunneling spectroscopy method (Fig. 6) allow for a comparison with the calculated results without using φ_s as an adjustable parameter. Indeed, we calculate k^+ and k^- at energies E_F and E_F+eV as functions of φ_s. Then, having eliminated φ_s, we set up k^+ and k^- as functions of the $2D$ electron concentration $n_{2D}=((k_{II}^+)^2+(k_{II}^-)^2)/4\pi$.

The experimental results are plotted in Fig. 6 precisely to these scales. We see that they fit the theoretical results (solid curves) well over the entire range of $2D$ electron concentrations from 1×10^{11} to 5.5×10^{11} cm^{-2}. Note that at low $2D$ electron concentrations ($n_{2D}<1.5\times10^{11}$ cm^{-2}), when only one frequency of type-II oscillations is observed, calculations predict the existence of only one branch of $2D$ states at $E=E_F$.

The inset in Fig. 6 shows the calculated dispersion law for $2D$ states at a $2D$ electron concentration of

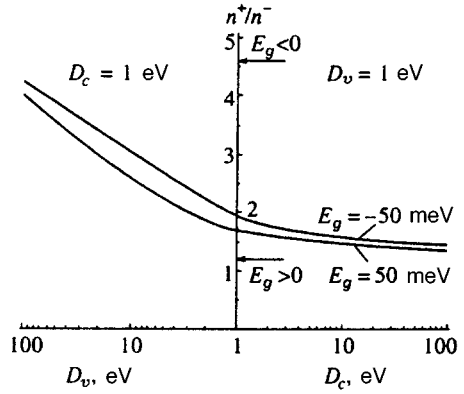


FIG. 7. Calculated value of n^+/n^- as a function of the band discontinuities at the semiconductor–insulator interface at $n_{2D}=1\times10^{12}$ cm^{-2} for semiconductors with E_g either positive or negative. The arrows indicate the values of n^+/n^- calculated with zero boundary conditions.

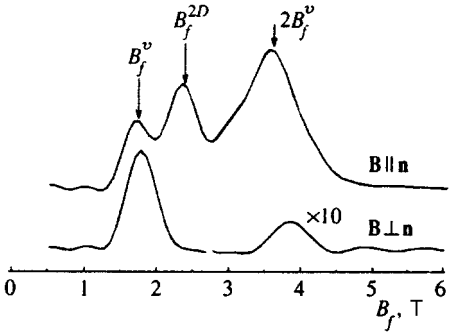


FIG. 8. Fourier spectra of tunneling conductivity oscillations in structure 3 at a bias voltage $V=75$ mV.

1.2×10^{11} cm^{-2}. Clearly, the k^+ branch has a minimum at $k\neq0$, so that at $E=E_F$ there can only be states belonging to this branch.

What is the effect of the parameters D_c and D_v, which are used in the adopted theoretical model and characterize the semiconductor–insulator interface? To answer this question we calculated the ratio n^+/n^-, determined by the magnitude of spin–orbit splitting, as a function of D_c and D_v for a fixed total $2D$ carrier concentration $n_{2D}=1\times10^{12}$ cm^{-2} for semiconductors with $E_g=+50$ meV and $E_g=-50$ meV (Fig. 7). We see that the splitting of the $2D$ states in these semiconductors is much more sensitive to D_v than to D_c. Here for $D_c>D_v$, in the semiconductor with $E_v>0$ the value of n^+/n^- is close to that calculated in the often-used model of zero boundary conditions,[16] while in the semiconductor with $E_g<0$ the two values differ considerably. The reason is that zero boundary conditions correspond in gapless semiconductors to an infinite discontinuity in the Γ_8 band ($D_v=\infty$; see Ref. 25), while zero boundary conditions in semiconductors with $E_g>0$ correspond to an infinite discontinuity in the Γ_6 band ($D_c=\infty$).

Within this model it is easy to allow for the screening of the surface potential by $2D$ electrons. Simultaneous solution of the Schrödinger and Poisson equations by successive approximations reveals that for the structures studied here, the corrections to k^+ and k^- do not exceed the experimental error.

Measurements of the tunneling conductivity oscillations in a magnetic field in both structures with Yb at the maximum positive and negative pulses, which (as noted earlier) change the surface potential, were conducted over the entire range of bias voltages. The results proved to be in good agreement with those of calculations.

The spin–orbit splitting of the spectrum of $2D$ electrons in a narrow-gap semiconductors HgCdTe was measured by Radantsev[15] and Wollrab *et al.*[27] at various $2D$ carrier concentrations, but at fixed energy equal to the Fermi energy of the semiconductor. The experimental results listed in these papers fit our model well.

These data show (Fig. 4) that in narrow-gap semiconductors, the spin–orbit splitting of the spectrum of $2D$ states in an asymmetric surface quantum well is large—of order 20–40 meV for $k>10^6$ cm^{-1}. This should lead to a situation at small values of φ_s in which, when the bottom of the band of $2D$ states is near the bottom of the conduction band of the bulk semiconductor, there are localized $2D$ states with only one spin, while states with the other spin are pushed into the continuous spectrum. Such a situation was observed for $2D$ states over a broad energy range in experiments involving *p*-HgCdTe–oxide–Al tunnel structures, in which φ_s is much smaller than in structures with Yb.

5. TUNNELING CONDUCTIVITY OSCILLATIONS IN *p*-HgCdTe–OXIDE–Al STRUCTURES

The Fourier spectra of tunneling conductivity oscillations in structure 3 for two orientations of the magnetic field, $\mathbf{B}\perp\mathbf{n}$ and $\mathbf{B}\|\mathbf{n}$, are depicted in Fig. 8. Clearly, for $\mathbf{B}\perp\mathbf{n}$, when the spectrum of $2D$ states is not quantized and the tunneling conductivity oscillations are related to tunneling only to bulk Landau levels, the Fourier spectrum exhibits only two harmonics, B_f^v and $2B_f^v$. The harmonic $2B_f^v$ is related to the spin splitting of Landau levels. The fact that this harmonic has a small amplitude for $\mathbf{B}\perp\mathbf{n}$ in a narrow-gap semiconductor in which spin splitting is nearly half the cyclotron splitting means that with this orientation of the magnetic field, the principal contribution to the oscillations is provided by tunneling to states with one spin. The reason for such a difference in the probabilities of tunneling into different spin states of the Landau levels in a narrow-gap semiconductor will be discussed in a separate paper.

When $\mathbf{B}\|\mathbf{n}$, in addition to oscillations with the frequencies of the previously discussed case, there is a harmonic with a fundamental field B_f^{2D} in the range $B_f^v<B_f^{2D}<2B_f^v$ (the large amplitude of oscillations with the frequency $2B_f^v$ corresponds to the case of close probabilities of tunneling to the two spin sublevels of the bulk Landau levels for this orientation of magnetic field). The angular dependence of the fundamental frequencies (Fig. 9) suggests that the fundamental fields B_f^v and $2B_f^v$ are independent of the angle, with only the amplitude of the $2B_f^v$ harmonic rapidly decreasing as θ grows from 0° to 90°. The B_f^{2D} harmonic closely resembles $1/\cos\theta$, which point to its relation to tunneling to $2D$ states.[1)]

The dependence of k^2 on V for structure 3 obtained via a Fourier analysis of the tunneling conductivity oscillations

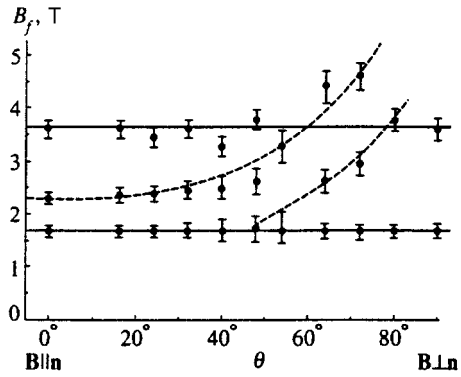


FIG. 9. Angular dependence of the fundamental fields of tunneling conductivity oscillations of structure 3 at a bias voltage $V=75$ mV. The solid lines correspond to oscillations caused by tunneling to bulk states, and the dashed curves to $2D$ states.

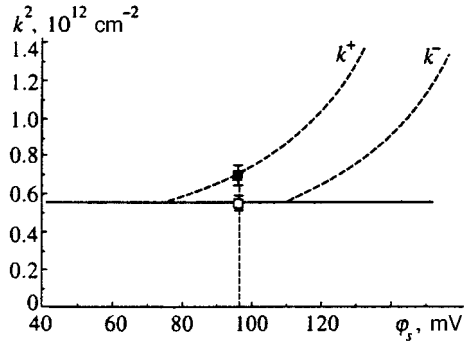


FIG. 11. k^2 as a function of the surface potential calculated at an energy of 75 meV: the solid line represents bulk states and the dashed curves the $2D$ states. ■ and □ stand for the corresponding experimental values of k^2 at a bias voltage $V=75$ mV.

(Fig. 10) shows that the bottom of the band of $2D$ states is near the bottom of the conduction band, i.e., above the Fermi energy, so that this structure has no $2D$ carriers, and oscillations at frequency B_f^{2D} correspond to tunneling to vacant $2D$ states. Note that in this structure, for **B**$\|$**n** there is only one branch of $2D$ states over the entire energy range (Fig. 10).

Using the adopted theoretical model, we can estimate the range of surface potential over which there is only one spin state of the $2D$ carriers. Figure 11 depicts the calculated k^2 vs. φ_s dependence for $2D$ states at an energy of 75 meV in a tunnel structure with parameters E_g and N_A-N_D corresponding to structure 3. We see that at this energy, the first spin state of the $2D$ carriers separates at $\varphi_s\approx 75$ mV, and the second at $\varphi_s\approx 110$ mV. Thus, in the 75–110 mV range of surface potentials, there can be only one spin state of $2D$ carriers. We assume that this is the state observed in structure 3 at **B**$\|$**n**. Figure 11 shows that at $V=75$ mV this structure has $\varphi_s=95$ mV.

The dispersion law for the electron $2D$ states at $\varphi_s=95$ mV calculated with the adopted model is shown in Fig. 10. We see that in this tunnel structure, in contrast to the structure with Yb, the experimentally determined values of k^2 in the 55–120 mV range of bias voltages are in good agreement with the dispersion law calculated with a V-independent value of the surface potential. The reason is that the surface state density in structure 3 is probably higher than in structures 1 and 2, so that the variation of the surface potential resulting from variations in the bias voltage over the moderately broad range of values of V in which we observed tunneling conductivity oscillations is small. Figure 10 also shows that in this structure the second spin state of $2D$ electrons only makes itself felt at $k^2>1.5\times 10^{12}$ cm^{-2} (the state also exists near $k=0$, but depicting it on this scale is difficult).

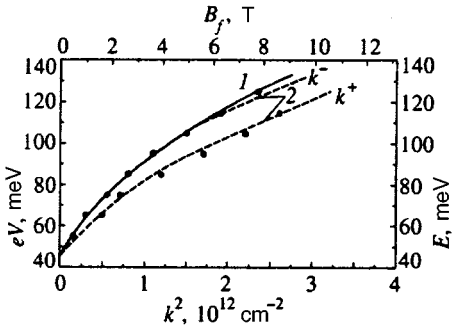


FIG. 10. B_f and k^2 as functions of the bias voltage. The dots represent the experimental values of B_f obtained with the orientation **B**$\|$**n** for structure 3. Curve *1* represents the dispersion law for bulk states calculated in the Kane model with parameters $E_g=45$ meV and $P=8\times 10^{-8}$ eV cm; curves *2* represent the dispersion laws for the $2D$ states split by the spin–orbit coupling (the curves were calculated at $\varphi_s=95$ meV).

Within the energy range studied here, this state must have values of k that are close to those of bulk states (Fig. 10), with the result that it is impossible to distinguish the oscillations with the respective frequencies in experiments. As noted earlier, such a state could appear in oblique magnetic fields for $\theta>50°$ (Fig. 9). It would be useful to be able to vary the surface potential in structure 3, but employing the modified method with additional pulses has led to no noticeable variation of φ_s in this structure.

The possibility of $2D$ states appearing only at high quasimomenta in semiconductors with a highly nonparabolic spectrum has been discussed in Refs. 28 and 29. Such states became known as kinematically coupled. Both papers examine the spinless problem. In real semiconductors with a highly nonparabolic spectrum (InSb and HgCdTe), allowance for spin leads to additional features in kinematically coupled states. This can clearly be seen in Fig. 12, which shows the dispersion laws for the bulk and $2D$ states calculated in the Kane model for two values of the surface potential. When φ_s is low, there can be $2D$ states with only one spin, and they appear at $k\neq 0$, i.e., they are kinematically coupled (Fig. 12a). When φ_s is high, these states exist at all values of k, and the states with the second spin are kinematically coupled (Fig. 12b). As Fig. 10 shows, this is the situation in structure 3.

It may seem inconsistent that all the experimental results concerning the spectrum of $2D$ states were obtained by analyzing oscillations in a magnetic field, while the theoretical calculations were performed for zero magnetic field. Here, to compare the results of calculations with experimental results, we used Eq. (1), i.e., we tacitly assumed that the rules of

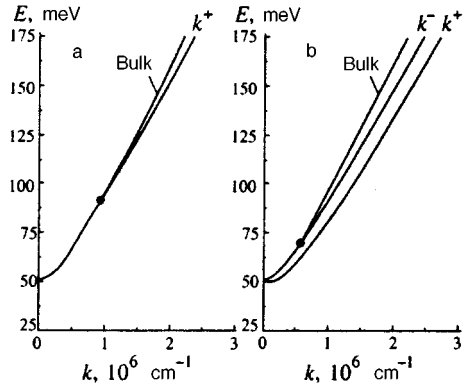


FIG. 12. Dispersion laws for bulk and $2D$ states calculated at $N_A=10^{18}$ cm^{-3}, $E_g=50$ meV, $D_c=2$ eV, and $D_v=1$ eV and at two values of the surface potential: (a) $\varphi_s=80$ mV, and (b) $\varphi_s=120$ mV.

semiclassical quantization of the spectrum of $2D$ states in a magnetic field are valid for such narrow-gap semiconductors. To confirm the validity of such an assumption, we calculated the Landau levels of $2D$ states in the Kane model in the same way as we did in Ref. 30. The results of calculations at a fixed surface potential $\varphi_s=300$ mV are depicted in Fig. 13a. The solid and dashed curves indicate the position of the two series of Landau levels corresponding to the different spin states of the $2D$ carriers. At fixed energy, the levels in each series are not strictly periodic in the reciprocal magnetic field, but the values of k determined from the average period of each series via Eq. (1) yield, to high accuracy, the dispersion law for the corresponding branch of the $2D$ states calculated in zero magnetic field (Fig. 13b). Thus, within an important range of energies and magnetic fields, the rule of semiclassical quantization in a magnetic field is satisfied to high accuracy.

6. CONCLUSION

We have used tunneling spectroscopy in a magnetic field to study the energy spectrum of $2D$ states localized in a surface quantum well of the narrow-gap semiconductor HgCdTe. The high surface potential in p-HgCdTe–oxide–Yb tunnel structures creates a subband of $2D$ electron states whose bottom is below the Fermi energy.

Two types of tunneling conductivity oscillations in a magnetic field are observed in such structures. The periods of these oscillations are determined, at energies E_F and E_F+eV, by the component of the quasimomentum of the $2D$ states parallel to the surface. This makes it possible to study the energy spectrum of $2D$ states over a broad energy range. We have found that in such structures with an asymmetric potential well there is strong spin–orbit splitting of the spectrum of $2D$ states.

In p-HgCdTe–oxide–Al tunnel structures the surface potential is much weaker, the bottom of the subband of $2D$ states is above the Fermi energy of the semiconductor, and the tunneling conductivity oscillations in a magnetic field are related to tunneling to vacant $2D$ states. The presence of only one branch in the spectrum of $2D$ states in these structures results from the fact that due to the strong spin–orbit coupling the second spin state is pushed into the continuous spectrum.

To analyze the experimental results we have calculated the energy spectrum of $2D$ states localized in the surface quantum well in a semiconductor with a Kane dispersion law. We have found that all the experimental data are in good agreement with the theoretical results.

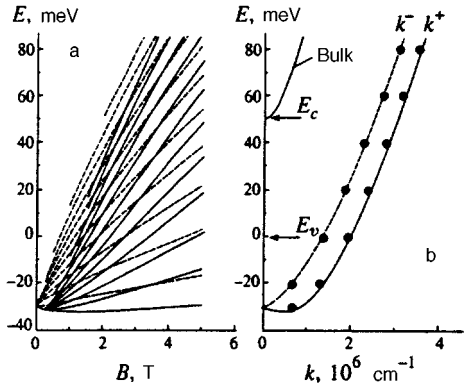


FIG. 13. (a) Landau levels of $2D$ states calculated in the Kane model (details in the main text). (b) Curves representing the dispersion laws for bulk and $2D$ states split by spin–orbit coupling (calculations done with zero magnetic field). The dots correspond to the values of k determined from the positions of the Landau levels in Fig. 13a. The following parameters were used in the calculations: $N_A=10^{18}$ cm^{-3}, $E_g=50$ meV, $D_c=2$ eV, $D_v=1$ eV, and $\varphi_s=300$ mV.

[1)] When $\theta>50°$, another angle-dependent component can be detected; possibly, this component is related to the appearance in inclined fields of a second spin state of $2D$ carriers. We know of no theoretical calculations of the spectrum of $2D$ carriers in a semiconductor with a complicated spectrum in an arbitrary inclined field, so that our experimental results could serve as a stimulus to such calculations.

[1] D. C. Tsui, Phys. Rev. B **8**, 2657 (1973).
[2] G. M. Min'kov, O. É. Rut, V. A. Larionova, and A. V. Germanenko, Zh. Éksp. Teor. Fiz. **105**, 719 (1994) [JETP **78**, 384 (1994)].
[3] R. C. Ashoori, J. A. Lebens, N. P. Bigelow, and R. H. Silsbee, Phys. Rev. B **48**, 4616 (1993).
[4] K. M. Brown, N. Turner, J. T. Nicholls, E. H. Linfield, M. Pepper, D. A. Ritchie, and G. A. C. Jones, Phys. Rev. B **50**, 15 465 (1994).
[5] J. P. Eisenstein, L. N. Preiffer, and K. W. West, Surf. Sci. **305**, 393 (1994).
[6] J. P. Eisenstein, L. N. Preiffer, and K. W. West, Phys. Rev. Lett. **69**, 3804 (1992).
[7] B. L. Altshuler and A. G. Aronov, in *Electron–Electron Interactions in Disordered Systems*, A. L. Éfros and M. Pollak (eds), North-Holland, Amsterdam (1985), p. 1.
[8] P. Sobkowicz, Semicond. Sci. Technol. **5**, 183 (1990).
[9] Yu. L. Bychkov and É. I. Rashba, JETP Lett. **39**, 78 (1984).
[10] G. Dresselhaus, Phys. Rev. **100**, 580 (1955).
[11] R. Eppenga and M. F. H. Schuurmans, Phys. Rev. B **37**, 10 923 (1988).
[12] J. Luo, H. Munekata, F. F. Fang, and P. J. Stiles, Phys. Rev. B **38**, 10 142 (1988).
[13] B. Das, D. C. Miller, S. Datta, R. Reifenberger, W. P. Hong, P. K. Bhattacharya, J. Singh, and M. Jaffe, Phys. Rev. B **39**, 1411 (1989).
[14] H. L. Stoermer, T. Haavasoja, V. Narayanamurti *et al.*, J. Vac. Sci. Technol. B **1**, 423 (1984).
[15] V. F. Radantsev, Zh. Éksp. Teor. Fiz. **96**, 1793 (1989) [Sov. Phys. JETP **69**, 1012 (1989)].
[16] V. F. Radantsev, T. I. Deryabina, G. I. Kulaev, and E. L. Rumyantsev, Phys. Rev. B **53**, 15 756 (1996).
[17] G. M. Min'kov, A. V. Germanenko, V. V. Kruzhaev, O. É. Rut, and V. A. Larionova, JETP Lett. **62**, 330 (1995).
[18] D. C. Tsui, Phys. Rev. B **4**, 4438 (1971).

[19] D. C. Tsui, Phys. Rev. B **8**, 2657 (1973).
[20] J. Muller and U. Kunze, Semicond. Sci. Technol. **8**, 705 (1993).
[21] U. Kunze, Z. Phys. B **80**, 47 (1990).
[22] L. P. Zverev, V. V. Kruzhaev, and G. M. Min'kov, Zh. Éksp. Teor. Fiz. **80**, 1163 (1981) [Sov. Phys. JETP **53**, 595 (1981)].
[23] L. P. Zverev, V. V. Kruzhaev, and G. M. Min'kov, Fiz. Tverd. Tela (Leningrad) **26**, 2943 (1984) [Sov. Phys. Solid State **26**, 1778 (1984)].
[24] I. M. Tsidil'kovskiĭ, G. I. Harus, and N. S. Shelushinina, Adv. Phys. **34**, 43 (1985).
[25] G. M. Minkov, A. V. Germanenko, V. A. Larionova, and O. E. Rut, Phys. Rev. B **54**, 1841 (1996).
[26] A. V. Germanenko, G. M. Minkov, V. A. Larionova, O. E. Rut, C. R. Becker, and G. Landwehr, Phys. Rev. B **52**, 17 254 (1995).
[27] R. Wollrab, R. Sizman, F. Koch *et al.*, Semicond. Sci. Technol. **4**, 491 (1989).
[28] R. E. Doezema and H. D. Drew, Phys. Rev. Lett. **57**, 762 (1986).
[29] An-zhen Zhang, J. Slinkman, and R. E. Doezema, Phys. Rev. B **244**, 10 752 (1991).
[30] G. M. Minkov, A. V. Germanenko, V. A. Larionova, and O. E. Rut, Semicond. Sci. Technol. **10**, 1578 (1995).

Translated by Eugene Yankovsky

Superradiance by a system of highly polarized exchange-coupled nonequivalent spins

D. A. Kostarov, N. P. Fokina, and K. O. Khutsishvili

I. Dzhavakhishvili Tbilisi State University, 380028 Tbilisi, Republic of Georgia
(Submitted 4 August 1996; resubmitted 13 January 1997)
Zh. Éksp. Teor. Fiz. **112**, 551–563 (August 1997)

We propose a model of radiofrequency (rf) superradiance by a system of interacting nonequivalent spins in a point specimen. In contrast to the rf superradiance observed and described earlier, here spin–spin coupling acts as the interaction with the cavity. To be definite, we examine the spins of two isotopes of a metal that are coupled by the Ruderman–Kittel interaction. The analysis of such a system when the magnetization of one spin species is inverted shows that the system can have one resonance frequency and two different decay times, instead of two resonance frequencies and one decay time in the usual situation. When such ''repulsion'' of decay times occurs and the absolute values of the spin polarizations are large, transverse magnetization increases and exhibits features characteristic of superradiance. Finally, we calculate the parameters of this superradiance: the voltage across the terminals of an rf pickup coil, the pulse length, the delay time, and the superradiant intensity. © *1997 American Institute of Physics.* [S1063-7761(97)01208-0]

1. INTRODUCTION

In recent years there has been an upsurge of interest in the magnetic resonance spectra of two species of exchange-coupled highly polarized nonequivalent spins.[1–5] Such spins can be found in two metal isotopes coupled by the Ruderman–Kittel interaction[1–4] or in magnetic ions occupying two magnetically nonequivalent positions in the lattice and coupled by the exchange interaction.[5] Systems of this kind exhibit merging–splitting and suppression–enhancement effects in the magnetic resonance lines, which means that the position and intensity of the lines do not correspond to the Larmor frequencies and relative isotope concentrations. The reason in both cases is the presence of an interaction of the exchange type $(H_{\rm ex})_{ab}$ between the two spin species, a and b, an interaction that in the case of highly polarized spins leads to a complicated dependence of the frequencies and line intensities on the spin polarization. By studying these spectra it is possible to establish, to high accuracy, the values of exchange constants and even determine, among other things, their sign.

It was also found that the nonlinearity of the equations of motion of the magnetizations, due to the interaction $(H_{\rm ex})_{ab}$, can affect the dynamics of the highly polarized spin system. Specifically, numerical solution of the equations of motion of two isotopes, ^{63}Cu and ^{65}Cu, with allowance for $(H_{\rm ex})_{ab}$ and with one magnetization inverted, leads to an avalanche-type buildup of the transverse magnetization pulse.[4] Similar transverse magnetization pulses were observed by Bosiger *et al.*,[6] Kiselev *et al.*,[7] and Bazhanov *et al.*[8] in a system of nuclear spins of one species with preestablished high negative polarization (these pulses were detected by measuring the voltage across the terminals of the coil containing the specimen). In Refs. 8–10 it was shown that these were superradiant pulses from the inverted nuclear spins interacting with the alternating field in the cavity (in the radio-frequency (rf) front end). The phenomenon mentioned in Ref. 4 differs from the effect reported in Refs. 6–8 in that there is no coupling between the spins and an external cavity.

We believe that in the case in which the isotope magnetizations can move, the interaction of one spin species with the other mimics the interaction between spins and the field in a cavity. The importance of studying such superradiance stems from the fact that in this way, coherent pulses can be generated in a frequency range where no real cavity will operate (e.g., in the x-ray range), not only with spins, but any two-level particles acting as emitters. Furthermore, superradiance in a cavityless two-component system of two level particles can serve as a source of information about coupling constants for the interactions between different particle species and about the characteristics of the two-level particles.

At this point it should be noted that in 1954, Dicke[11] predicted the possibility of superradiance by a system of mutually correlated (through the radiation field) inverted two-level particles in a point sample ($V \ll \lambda^3$, where V is the specimen's volume, and λ is the wavelength of the radiation). Until now, however, this phenomenon has been observed either in a finite specimen ($V \gg \lambda^3$; see Ref. 12) or in a specimen placed in an external cavity; in both cases correlation is induced by the radiation field, which obeys Maxwell's equations. Note that in interpreting the superradiance data, Maxwell's equations were adopted for the field in the cavity.[13–15,9,10] The reason is that in a finite specimen, mode selection[15] is achieved by ''selecting'' the proper shape of the specimen, which is equivalent to the presence of a cavity, and mode selection in the rf range is done by using an ordinary rf front end. Naturally, the theoretical description of superradiance in both cases is roughly the same.

In the present paper we propose a model of superradiance by interacting nonequivalent spins (two-level particles) that initially have high polarizations of both signs and are not coupled to a cavity. The mechanism of inducing correlations differs from the one mentioned earlier; more precisely, it originates in the interaction of one particle species with the internal (molecular) field generated by the other species. Ex-

 1063-7761/97/080300-07$10.00

amples of such particles are the nuclear spins of metal isotopes (^{107}Ag and ^{109}Ag in silver, ^{203}Tl and ^{205}Tl in thallium, and ^{63}Cu and ^{65}Cu in copper)[1–4,16,17] coupled by the Ruderman–Kittel interaction, characterized by close gyromagnetic ratio values and equal (to order of magnitude) concentrations, and kept at extremely low spin temperatures T_s ($\hbar\omega_{\rm res} \geqslant k_B T_s$), which are, however, higher than the nuclear magnetic order temperature. Since below we discuss the case of nuclear spins of ^{63}Cu and ^{65}Cu, here are some relevant data: the gyromagnetic ratios are $\gamma_{63}=11.3$ MHz T^{-1} and $\gamma_{65}=12.1$ MHz T^{-1}, the relative concentrations (natural abundances) are $x_{65}=0.31$ and $x_{63}=0.69$, and the nuclear magnetic ordering temperature is about 60 nK. The characteristic experimental conditions used by the Helsinki[1,2,17] and Bayreuth[3,4] groups in their studies of these metals in the paramagnetic state were as follows: a conduction electron temperature T_e of the order of hundreds of microkelvins, a spin temperature of the order of several microkelvins, and constant magnetic fields of roughly 10 mT. At such lattice temperatures, the nuclear spin–lattice relaxation time T_1 is much greater than the reciprocal linewidth T_2 of stationary NMR; in particular, at $T_e=400$ μK and $T_2 \sim 10^{-5}$ s we have $T_1 \sim 1$ hr (Ref. 17), i.e., $T_1 \gg T_2$, so that spin–lattice relaxation can be ignored. We describe the relaxation of the transverse components of the isotope magnetizations in the Bloch approximation, since we know of no data that contradicts this approximation for these metals (see Refs. 1, 2 and 17).

We also note that these processes are transient, i.e., they take place in time intervals much shorter than the time of cross relaxation between the isotopes (we are interested in times of the order of fractions of a microsecond, while the cross–relaxation time for the temperatures involved amounts to tens of milliseconds). The problem of selecting the values of the constant magnetic field for the expected effect to occur is discussed below.

2. DERIVATION OF THE BASIC EQUATIONS

To be specific, we take a spin system consisting of two species of nuclear spins (a and b) of two metal isotopes with gyromagnetic ratios $\gamma_a > \gamma_b$ and place it in a constant magnetic field with induction $\mathbf{B}_0 \| \mathbf{z}$, with the spins coupled by the Ruderman–Kittel interaction. The Hamiltonian of such a system has the form

$$H = -\hbar\omega_a I_a^z - \hbar\omega_b I_b^z - \sum_{i,j} J_{ij} I_{ai} I_{bj}, \tag{1}$$

where $\omega_{a,b}=\gamma_{a,b}B_0$, $I_{a,b}$ are the isotope spin operators, and J_{ij} is the Ruderman–Kittel coupling constant. We write the equations of evolution of the magnetization components obtained in the semiclassical approximation via the Hamiltonian (1) as follows:

$$\begin{aligned}
\dot m_b^x &= \omega_b^0 m_b^y - \gamma_b J_0 m_b^z m_a^y - \frac{m_b^x}{T_2},\\
\dot m_b^y &= -\omega_b^0 m_b^x + \gamma_b J_0 m_b^z m_a^x - \frac{m_b^y}{T_2},\\
\dot m_b^z &= \gamma_b J_0 (m_b^x m_a^y - m_b^y m_a^x),
\end{aligned} \tag{2}$$

where

$$\omega_b^0 = \omega_b + \gamma_b J_0 m_a^z, \quad \omega_a^0 = \omega_a + \gamma_a J_0 m_b^z; \tag{2'}$$

$m_{a,b}^\alpha = \gamma_{a,b}\hbar n \langle I_{a,b}^\alpha \rangle$, $\alpha = x,y,z$; n is the spin number density; $J_0 = z_n/\hbar^2\gamma_a\gamma_b n$ (it is assumed that only the z_n nearest neighbors of the ith spin contribute to the sum $\Sigma_j J_{ij} = z_n J$); and T_2 is the spin–spin relaxation time. The experiments of Oja *et al.*,[1] Hakonen *et al.*,[2] and Estrom *et al.*[17] show that the two isotopes have close values of T_2. As is known,[16] for the isotopes of the above metals $z=12$; here, for instance, each of the 12 neighbors of a selected spin may be ^{65}Cu with probability x_{65} or ^{63}Cu with probability x_{63}.

The equations for m_a^α can be obtained from Eqs. (2) by interchanging the indices: $a \leftrightarrow b$. Thus, the problem is described by a system of six first-order differential equations.

It is well known that the generation process requires a source of energy in the system. We assume that this requirement is fulfilled by population inversion of one spin species. In this case, obviously, the stored energy can affect the interaction between the two spin species and hence the nature of spin precession. In particular, the spectra of frequencies[18] and decay times may differ from those in the ordinary situation. To elucidate these features we find the spectrum of the spin system under consideration by plugging $m_{a,b}^+ = \tilde m_{ab}^+ \exp(-i\Omega t)$ into the equations

$$\begin{aligned}
\dot m_a^+ &= -i\omega_a^0 m_a^+ + i\gamma_a J_0 m_{a0}^z m_b^+ - \frac{m_a^+}{T_2},\\
\dot m_b^+ &= -i\omega_n^0 m_b^+ + i\gamma_b J_0 m_{b0}^z m_a^+ - \frac{m_b^+}{T_2},
\end{aligned} \tag{3}$$

which follow from Eqs. (2) if we replace m_a^z and m_b^z by their initial values m_{a0}^z and m_{b0}^z (by using the operators $m_{a,b}^+ = m_{a,b}^x + i m_{a,b}^y$ we are able to reduce the system of equation to two first-order differential equations). This yields the following equations for the precession amplitudes:

$$\begin{aligned}
&\left(\Omega - \omega_a^0 + \frac{i}{T_2}\right)\tilde m_a^+ + \gamma_a J_0 m_{a0}^z \tilde m_b^+ = 0,\\
&\gamma_b J_0 m_{b0}^z \tilde m_a^+ + \left(\Omega - \omega_b^0 + \frac{i}{T_2}\right)\tilde m_b^+ = 0.
\end{aligned} \tag{4}$$

The condition that the system (4) have a nontrivial solution leads to the following equation for the spectrum:

$$\begin{aligned}
&\Omega^2 - \Omega\left(\omega_a^0 + \omega_b^0 - \frac{2i}{T_2}\right) - \frac{1}{T_2^2} - \gamma_a\gamma_b J_0 m_{a0}^z m_{b0}^z\\
&\quad - \frac{i}{T_2}(\omega_a^0 + \omega_b^0) + \omega_a^0\omega_b^0 = 0,
\end{aligned} \tag{5}$$

whose solution has the form

$$\Omega_{a,b} = \frac{\omega_a^0 + \omega_b^0}{2} - \frac{i}{T_2} \pm \frac{1}{2}\sqrt{(\omega_{ab}^0)^2 + 4\gamma_a\gamma_b J_0^2 m_{a0}^z n_{b0}^z}. \tag{6}$$

Since at $J_0=0$ and $\omega_{ab}^0 = \omega_a^0 - \omega_b^0 > 0$ (from now on we assume this inequality to be valid) the solution (6) yields $\Omega_- = \omega_b^0$ and $\Omega_+ = \omega_a^0$, we conclude that the ''plus'' corresponds to an a-like precession mode and the ''minus'' to a b-like precession mode.

We now allow for the fact that one magnetization is inverted, as a result of which $4\gamma_a\gamma_b J_0^2 m_{a0}^z m_{b0}^z$ becomes negative, and for sufficiently high degrees of spin polarization exceeds $(\omega_{ab}^0)^2$ in absolute value, i.e.,

$$-4\gamma_a\gamma_b J_0^2 m_{a0}^z m_{b0}^z > (\omega_{ab}^0)^2. \tag{7}$$

Then for $\Omega_{a,b}$ we have

$$\Omega_a = \Omega_G - i\Gamma_-, \quad \Omega_b = \Omega_G - i\Gamma_+, \tag{8}$$

where

$$\Omega_G = \frac{\omega_a^0 + \omega_b^0}{2}, \quad \Gamma_\pm = \frac{1}{T_2} \pm \frac{1}{2}\sqrt{4\omega_T^2 - (\omega_{ab}^0)^2},$$
$$\omega_T^2 = -\gamma_a\gamma_b J_0^2 m_{a0}^z m_{b0}^z. \tag{9}$$

Thus, in conditions where inequality (7) is valid, the system of coupled magnetizations is characterized by a single frequency Ω_G at which the system can emit radiation, as we will show shortly. This result agrees with that for the frequency of oscillations, obtained in Ref. 18, of a coupled system consisting of dipoles and a cavity with resonance frequencies ω_n and ω_c, respectively, and decay times T_2 and τ_c, respectively:

$$\Omega_G = \left(\frac{\omega_n}{2\tau_c} + \frac{\omega_c}{T_2}\right)\left(\frac{1}{2\tau_c} + \frac{1}{T_2}\right)^{-1},$$

with $\Omega_G = (\omega_n + \omega_c)/2$ at $1/2\tau_c = 1/T_c$ (in our system the noninteracting isotopes have equal decay times).

Equations (8) and (9) show that condition (7) leads to a situation in which instead of two coupled-oscillator modes with below-threshold frequencies

$$(\omega_{a,b})_{\text{b.th}} = \frac{\omega_a^0 + \omega_b^0}{2} \pm \frac{1}{2}\sqrt{(\omega_{ab}^0)^2 + 4\gamma_a\gamma_b J_0^2 m_{a0}^z m_{b0}^z} \tag{9$'$}$$

and equal decay times T_2^{-1} (which is the case with noninverted magnetization), we must introduce two magnetization vectors belonging to spins a and b, precessing with equal frequencies Ω_G, and characterized by different decay rates $\Gamma_\pm$. Thus, if the precession of the magnetization vectors $\mathbf{m}_a$ and $\mathbf{m}_b$ of a spin system close to equilibrium is a superposition of two precessions with normal below-threshold frequencies $(\omega_a)_{\text{b.th}}$ and $(\omega_b)_{\text{b.th}}$ and equal decay rates T_2^{-1}, after a transition to a state with a high value of the energy stored by the inverted spin species (condition (7)), the spin system is described by two magnetization vectors, $\mathbf{m}_a$ and $\mathbf{m}_b$, with equal frequencies but different decay rates. We see, therefore, that when ω_{ab}^0 is positive and condition (7) is met, the evolution of the spins with the lower partial frequency (b-spins) proceeds with the decay rate Γ_+ and that of the spins with the higher partial frequency (a-spins) with the decay rate Γ_-, i.e., a-spins decay more slowly than b-spins.[1)]

To study the transient process described by the nonlinear differential equations (2), we seek their solution by the method of slowly varying amplitudes[21] (here we employ the fact that the time constant of the amplitude variation is much larger than the precession period). As in the description of this method in Ref. 21, we begin with two second-order equations for m_a^x and m_b^x without decay, which has already been taken into account by the spectrum (6), and two first-order equations for m_a^z and m_b^z. We seek the solution of these equations in the form

$$m_a^x = \frac{\tilde{m}_a^x(t)}{2}\exp[-i(\Omega_G - i\Gamma_-)t] + \text{c.c.},$$
$$m_b^x = \frac{\tilde{m}_b^x(t)}{2}\exp[-i(\Omega_G - i\Gamma_+)t] + \text{c.c.}, \tag{10}$$

where the $\tilde{m}_{a,b}^x(t)$ are slowly varying complex-valued amplitudes. Allowing for the fact that the inequalities $|\dot{\tilde{m}}_{a,b}^x| \ll \Omega_G |\tilde{m}_{a,b}^x|$ represent the conditions for slow variation of the amplitudes, we arrive at the following equations for the slowly varying amplitudes:

$$\dot{\tilde{m}}_a^x = i\frac{\omega_{ab}^0}{2}\tilde{m}_a^x - \Gamma_-\tilde{m}_a^x - i\gamma_a J_0 m_a^z \tilde{m}_b^x \exp[(\Gamma_- - \Gamma_+)t],$$
$$\dot{\tilde{m}}_b^x = -i\frac{\omega_{ab}^0}{2}\tilde{m}_b^x - \Gamma_+\tilde{m}_b^x - i\gamma_b J_0 m_b^z \tilde{m}_a^x \times \exp[(\Gamma_+ - \Gamma_-)t]. \tag{11}$$

We also write the equations for $\dot{m}_a^z$ and $\dot{m}_b^z$ in the form

$$\dot{m}_a^z = \frac{\gamma_a J_0 i}{2}[(\tilde{m}_a^x)^* \tilde{m}_b^x - \tilde{m}_a^x (\tilde{m}_b^x)^*],$$
$$\dot{m}_b^z = \frac{\gamma_b J_0 i}{2}[(\tilde{m}_a^x)^* \tilde{m}_b^x - \tilde{m}_a^x (\tilde{m}_b^x)^*]. \tag{12}$$

If $\Gamma_+ \gg \Gamma_-$, we can say that the rapidly decaying b-spins adjust to the evolution of the slowly decaying a-spins, so that we can adiabatically exclude the evolution of the b-spins, to which end we must nullify the derivative of the x-component of the magnetization of b-spins:

$$\dot{\tilde{m}}_b^x = 0 \tag{13}$$

(we have used what Haken called the subordination principle[22]). The above equality leads to the condition $\dot{m}_b^z = 0$, i.e., $m_b^z = m_{b0}^z$, and combining (11) with (13) yields

$$\tilde{m}_b^x = -\frac{\gamma_b J_0 m_b^z}{\omega_{ab}^0/2 + i\Gamma_+}\tilde{m}_a^x \exp[(\Gamma_+ - \Gamma_-)t]. \tag{14}$$

This equality determines the enhancement of the complex-valued precession amplitude of the b-spins with respect to the precession amplitude of the a-spins. We assume that $m_{b0}^z > 0$ and $m_{a0}^z = -|m_{a0}^z| < 0$. Indeed, if the b-spins move so rapidly that the a-spins are unable to follow, the b-spins do not evolve independently. Then only the a-spins can release the stored energy in the course of the rotation of the magnetization vector. This means that the a-spins must be the inverted spins. Plugging the frequencies (2$'$) into (7), we can easily see that the inequality (7) is true for $m_{a0}^z < 0$ only for negative J_0 (note that in most metals $J_0 < 0$). We now plug (14) into the first equation in (11) and write the complex-valued amplitudes $\tilde{m}_{a,b}^x$ in the form of the product of a real-valued amplitude and a phase factor:

$$\tilde{m}_a^x = a(t)\exp[i\varphi(t)], \quad \tilde{m}_b^x = ib(t)\exp[i\vartheta(t)].$$

Then

$$\dot{a}+i\dot{\varphi}a=i\left(\frac{\omega_{ab}^0}{2}+i\Gamma_-\right)a+\frac{i\gamma_a\gamma_bJ_0^2m_{b0}^z}{(\omega_{ab}^0/2)^2+\Gamma_+^2}\times\left(\frac{\omega_{ab}^0}{2}+i\Gamma_+\right)m_a^z a. \tag{15}$$

Separating the real and imaginary parts in Eq. (15), we obtain

$$\dot{a}=-\Gamma_- a-\frac{1}{T_R}\frac{m_a^z a}{|m_{a0}^z|}, \tag{16}$$

$$\dot{m}_a^z=\frac{1}{T_R}\frac{a^2}{|m_{a0}^z|}, \tag{17}$$

$$\dot{\varphi}=\frac{\omega_{ab}^0}{2}\left(1+\frac{1}{T_R\Gamma_+}\frac{m_a^z}{|m_{a0}^z|}\right), \tag{18}$$

where

$$\frac{1}{T_R}=\frac{\gamma_a\gamma_bJ_0^2|a_{a0}^z|m_{b0}^z}{(\omega_{ab}^0/2)^2+\Gamma_+^2}\Gamma_+ \tag{19}$$

is similar to the reciprocal radiative decay time in the problem of spin–cavity interaction.[23] Note that the phase $\varphi(t)$ changes little in the course of a time interval that is the reciprocal of Γ_+, i.e., $\dot{\varphi}\Gamma_+^{-1}\ll 1$.

Equation (18) implies that this is true if $\omega_{ab}^0/2\Gamma_+\ll 1$. Thus, Eqs. (16)–(18) are true if the system of particles of the two species has two small parameters,

$$\frac{|\Gamma_-|}{\Gamma_+}\ll 1 \quad \text{and} \quad \frac{\omega_{ab}^0}{2\Gamma_+}\ll 1. \tag{20}$$

Interestingly, the separation of the general precession of the particles at the mean frequency leads to a situation in which the rate of variation of the precession phase is proportional to the difference of the partial Zeeman frequencies.

Let us estimate the shift in the isotope precession phases. Comparing the ratio

$$\frac{\tilde{m}_a^x}{\tilde{m}_b^x}=\frac{a}{b}e^{i\delta}, \quad \delta=\varphi-\vartheta-\frac{\pi}{2},$$

with (14), we find the quantities needed for subsequent calculations:

$$\cos^2\delta=\frac{4\Gamma_+^2}{(\omega_{ab}^0)^2+4\Gamma_+^2}, \tag{21}$$

$$\tan\delta=\frac{\omega_{ab}^0}{2\Gamma_+}. \tag{22}$$

Summarizing, we can say that in this section we proceeded, under conditions (20), from the description of the motion of two isotope magnetizations by six first-order differential equations to a description by two first-order differential equations and the expression (18) for the lacking slow variable φ.

3. CALCULATING GENERATION PARAMETERS FOR AN EXCHANGE-COUPLED SYSTEM OF NONEQUIVALENT SPINS

Solving the system of equations (16) and (17) is easy because it is similar to solving the system of equations for spins coupled with the cavity under the condition that $\tau_c\ll T_c$, where τ_c is the ringing time of the rf front end.[9] To obtain the solution we first examine the initial stage in the temporal development of the magnetization $\mathbf{m}_a$, when $m_a^z\approx-|m_{a0}^z|$. Then

$$\dot{a}\approx\left(\frac{1}{T_R}-\Gamma_-\right)a,$$

which leads to the following condition of buildup of a:

$$\frac{1}{T_R}-\Gamma_->0. \tag{23}$$

Note that Γ_- can be either positive or negative. When Γ_- is negative, the inequality (23) holds automatically. For this reason we consider the cases $\Gamma_->0$ and $\Gamma_-<0$ separately.

First we assume that Γ_- is positive. This condition is ensured by magnetization values such that

$$\omega_T^2\leqslant\left(\frac{\omega_{ab}^0}{2}\right)^2+\left(\frac{1}{T_2}\right)^2.$$

Generation occurs when

$$\frac{1}{T_R}\geqslant\Gamma_-, \tag{24}$$

where the equality corresponds to the threshold. Writing (24) explicitly and allowing for the fact that the phase varies slowly, $(\omega_{ab}^0/2)^2\ll\Gamma_+^2$, we obtain the generation threshold in the form

$$\omega_T^2=\frac{1}{2}\left[\left(\frac{\omega_{ab}^0}{2}\right)^2+\left(\frac{1}{T_2}\right)^2\right]. \tag{25}$$

Note that (25) automatically leads to the following inequality:

$$\frac{1}{T_2}<2\sqrt{\left(\frac{\omega_{ab}^0}{2}\right)^2+\gamma_a\gamma_bJ_0^2m_{a/0}^zm_{b0}^z},$$

which guarantees that the equilibrium lines (9′) belonging to the different isotopes are resolved (this condition is needed so that one line can be selectively inverted).

But if

$$\omega_T^2>\left(\frac{\omega_{ab}^0}{2}\right)^2+\left(\frac{1}{T_2}\right)^2,$$

then $\Gamma_-<0$, and as noted earlier, the condition for buildup of amplitude a is met automatically. For large negative Γ_-, however, the decay-rate repulsion condition ($\Gamma_+\gg|\Gamma_-|$) ceases to be met. The final conclusion is that generation in the transient regime can occur at such initial magnetization values that ω_T^2 lands in the interval

$$\frac{1}{2}\left[\left(\frac{\omega_{ab}^0}{2}\right)^2+\left(\frac{1}{T_2}\right)^2\right]\leqslant\omega_T^2\leqslant\left(\frac{\omega_{ab}^0}{2}\right)^2+\left(\frac{1}{T_2}\right)^2,$$

i.e., the most favorable condition for generation is when the excess above threshold is a little more than twofold.

Comparing this result with the superradiance pattern of a system of spins coupled with the cavity, we can say that now one spin species acts as a cavity in relation to the other species. Equations (16) and (17) imply that here the internal field generated by the b-spins, $|\tilde{h}_{\rm mol}|=|2J_0\tilde{m}_b^x|$, acts as the cavity field[9] in relation to the a-spins. Just as in Ref. 9, this field, according to (14), is proportional to $\tilde{m}_a^x$ and leads to a rotation of the magnetization vector of the a-spins, and in the process there is an avalanche-type buildup of $|\tilde{m}_a^x|$.

But there are also considerable differences here: the field does not obey Maxwell's equations in a cavity, but is due entirely to spin interaction; also, while in the case of spins coupled to a cavity the separation of the rapidly decaying oscillator occurs on the basis of an inequality, $\tau_c \ll T_2$ or $\tau_c \gg T_2$, which contains the decay times of uncoupled oscillators,[9,18,24] when there are two spin species the corresponding decay rates in the uncoupled state are equal, but due to the strong coupling one rate increases while the other decreases.

Another difference is that while for a system consisting of spins and a cavity two types of generation are possible, (a) for $\tau_c \ll T_2$ similar to generation by an ammonia maser,[9,18] and (b) for $\tau_c \gg T_2$ similar to generation by an optical laser,[18,24] for a system of two spin species only maser-type generation is possible. This is to be expected since the reason for the effect being studied is the spin–spin coupling, which causes coherence to emerge. In the case of laser-type generation, however, the slow variables develop over time intervals much longer than T_2, when the spin–spin interaction (both secular and nonsecular) has long since ceased.

When the inequalities (23) hold, the solution of the system of equations (16) and (17) has the form

$$a=\pm|m_{a0}^z|(1-T_R\Gamma_-)\operatorname{sech}\left[(1-T_{Da})\left(\frac{1}{T_R}-\Gamma_-\right)\right], \tag{26}$$

$$m_a^z=-|m_{a0}^z|T_R\Gamma_-+|m_{a0}^z|(1-T_R\Gamma_-) \times\tanh\left[(t-T_{Da})\left(\frac{1}{T_R}-\Gamma_-\right)\right], \tag{27}$$

$$\dot{\varphi}=\frac{\omega_{ab}^0}{4T_R\Gamma_+}(1-T_R\Gamma_-)\tanh\left[(t-T_{Da})\left(\frac{1}{T_R}-\Gamma_-\right)\right]. \tag{28}$$

Thus, we have established that after the evolution of the rapidly decaying spins has been adiabatically eliminated (with $|\Gamma_-|\ll\Gamma_+$ and $\omega_{ab}^0\ll 2\Gamma_+$), the equations for the slowly varying amplitudes of the slowly decaying spins describe a motion similar to the fall of an ''overdamped'' pendulum from the upper unstable state of equilibrium; only in our case the magnetization of one spin species ''falls'' from the inverted state toward the z axis. When this slowly varying amplitude is plugged into (10), we see that this ''fall'' is superimposed on a precession with frequency $\Omega_G+\dot{\varphi}$, the frequency of oscillation, where the shift $\dot{\varphi}$ is determined by the difference of partial frequencies, ω_{ab}^0.

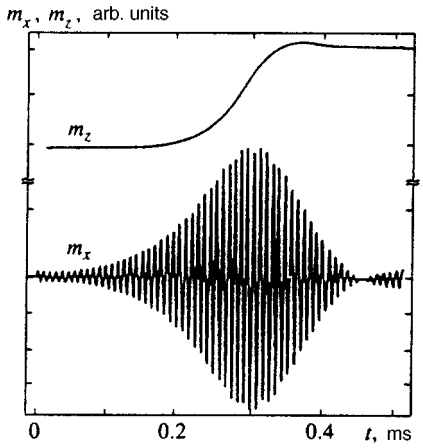


FIG. 1.

The sign of a in (26) is determined by the sign of the initial value, $a(0)$. The initial values of the transverse magnetizations are determined by the conditions for inverted magnetization (pumping). When the spins interact with a cavity, the initial negative value of the longitudinal magnetization of the spins is created by the dynamic cooling method.[6–8] Only Nyquist noise in the rf front end can serve as a source of finite initial transverse magnetization in this case (as shown by Sleater *et al.*,[25] such noise is much more effective than radio-frequency spontaneous emission).

In our problem of superradiance generation by two species of exchange-coupled spins, approximately antiparallel magnetizations can be achieved by, say, applying coherent pulses. Here the angle of rotation of the magnetization vector of the a-spins must be close to π (but must not be exactly equal to π), while that of the second magnetization must be close to zero. Such angle values guarantee the existence of finite initial values of the transverse magnetizations.

Equation (26) describes an $|a|$-buildup pulse with a sharp peak in time shaped in a form of a hyperbolic secant, which is characteristic of superradiance. The maximum value of $|a|$ is attained at $t=T_{D\alpha}$, and the time $t=0$ corresponds to the moment when the generation threshold is reached; $T_{D\alpha}$ is called the delay time, and its value can be estimated from the initial condition

$$|a(0)|=|m_{a0}^z|(1-T_R\Gamma_-)\operatorname{sech}\left[T_{Da}\left(\frac{1}{T_R}-\Gamma_-\right)\right]. \tag{29}$$

By analogy with ordinary superradiance,[8–10] we assume that $T_{D\alpha}\gg T_R$. The hyperbolic cosine can then be replaced by an exponential, so that we obtain at the following estimate:

$$T_{Da}\approx\frac{T_R}{1-T_R\Gamma_-}\ln\left[\left|\frac{m_{a0}^z}{a(0)}\right|(1-T_R\Gamma_-)\right]. \tag{30}$$

Note that in Ref. 4 the pulse of transverse magnetization m_x shaped in the form of a hyperbolic secant (Fig. 1) was obtained by computer simulation of the development of the magnetization of two copper isotopes, after the ^{65}Cu isotope, having the larger gyromagnetic ratio, was deflected by 179°

and the ^{63}Cu isotope by 1° with respect to the z axis. In this setting our model leads to the same conclusion.

The superradiance described above can be observed by measuring the voltage across the terminals of an rf pickup coil. When the rf front end is tuned precisely to the oscillating frequency, this voltage is[9]

$$U=-\eta_0\mu_0 SNQ(\dot{m}_a^x+\dot{m}_b^x), \tag{31}$$

where η_0 is the degree to which the specimen fills the coil, N is the number of turns in the coil, S is the coil's cross-sectional area, Q is the Q-factor of the rf front end, and $\mu_0=4\pi\times10^{-7}$ H m^{-1}. Plugging in $\dot{m}_b^x=0$ and $m_a^x=a\cos[\Omega_G t+\varphi(t)]$, we obtain the following expression for the envelope of the superradiance signal:

$$|U_{SR}^{\rm env}|=\eta_0\mu_0 SNQ(\Omega_G+\dot{\varphi})|m_{a0}^z|(1-T_R\Gamma_-)\times\operatorname{sech}\left[(t-T_R\Gamma_-)\left(\frac{1}{T_R}-\Gamma_-\right)\right]. \tag{32}$$

We see that in reality the system must oscillate at the frequency $\Omega_G+\dot{\varphi}$, with the frequency shift given by (28). We can now use (32) to calculate the FWHM of the buildup pulse of the superradiance signal:

$$\Delta t_{SR}\approx\frac{T_R\ln(2+\sqrt{3})}{1-T_R\Gamma_-}. \tag{33}$$

The superradiant intensity can be obtained by negating the rate of variation of the Zeeman energy of the a-spins precessing at frequency Ω_G and evolving according to Eq. (17):

$$I_{SR}=\frac{\Omega_G}{\gamma_a}\dot{m}_a^z V, \tag{34}$$

where V is the specimen's volume. Plugging the right-hand side of Eq. (17) with the known value (26) of the real-valued amplitude a into (34), we find that

$$I_{SR}=\frac{\Omega_G V|m_{a0}^z|}{\gamma_a T_R}(1-T_R\Gamma_-)^2\times\operatorname{sech}^2\left[(t-T_{Db})\left(\frac{1}{T_R}-\Gamma_-\right)\right]. \tag{35}$$

Now let us establish how the generation parameters depend on the number of emitters. To this end we first plug J_0^2, $\cos^2\delta$, $m_{a0}^z=\gamma_a\hbar n x_a p_{a0} I$, and $m_{b0}^z=\gamma_b\hbar n x_b p_{b0} I$ into (19), where x_a and x_b are the relative isotope concentrations, and I is the spin value. Combining the result with (20), we get

$$\frac{1}{T_R}=z_n^2\left(\frac{J}{\hbar}\right)^2\alpha\Gamma_+^{-1},\quad \alpha=x_a x_b p_{b0}|p_{a0}|I^2.$$

For sufficiently small values of $T_R\Gamma_-$ (the criterion for smallness is given below) and at $t=T_{D\alpha}$, the superradiant intensity becomes

$$I_{SR}=\frac{\hbar N_{\rm sp}x_a|p_{a0}|I\Omega_G}{T_R}=\hbar N_{\rm sp}\Omega_G x_a|p_{a0}|I\frac{\alpha z_n^2}{\Gamma_+}\left(\frac{J}{\hbar}\right)^2,$$

where $N_{\rm sp}$ is the total number of spins. The length of the superradiance pulse is

$$\Delta t\approx\frac{\Gamma_+}{\alpha z_n^2(J/\hbar)^2}.$$

If we draw an analogy with a superradiance process in which mode selection is done in some manner, we should have $\Delta t_{SR}\propto n^{-1}$ and $I_{SR}\propto N_{\rm sp}n$ (n is the active particle concentration), which is characteristic of superradiance. But since in our case feedback needed for generation emerges because of short-range two-particle interaction, the pulse intensity is proportional to $N_{\rm sp}z_n^2$ rather than $N_{\rm sp}n$ and the pulse width is accordingly inversely proportional to z_n^2 rather than n.

As noted in Refs. 8 and 9, $T_R\Gamma_-$ can be considered small if the excess above the threshold is twofold, i.e., $T_R\Gamma_-=1/2$. When pumping (the attained value of the product of magnetizations) is such that this condition is met, we have $\Gamma_-=0$, and at higher pumping levels $\Gamma_-<0$. It is at small negative values of Γ_- that generation becomes especially prominent.

What we have said above refers to the case where ω_{ab}^0 is positive. But if $\omega_{ab}^0<0$, we must only perform the substitution $\Gamma_+\leftrightarrow\Gamma_-$, i.e., the a-spins decay more rapidly than the b-spins. Generation occurs if the b-spins were initially inverted and the degrees of polarization of the two spin-species are high (naturally, condition (7) still holds, but now it is true for positive J_0). Similarly, we can decouple the equations for the magnetizations, solve the system of differential equations for b and m_b^z, and write the corresponding values for $\dot{\vartheta}$ and a. All the superradiance parameters (30)–(35) remain valid if we interchange the indices, $a\leftrightarrow b$.

4. CONCLUSION

We have shown that if in a magnetic field B_0 there are two metal isotopes with gyromagnetic ratios γ_a and γ_b, relative concentrations x_a and x_b, and spins I, with each spin coupled to the nearest neighbors of the other species by the Ruderman–Kittel interaction with a coupling constant J, one isotope inverted, and the polarizations p_{a0} and p_{b0} satisfying the inequality

$$4I^2 x_a x_b|p_{a0}p_{b0}|z_n^2\left(\frac{J}{\hbar}\right)^2>(\omega_{ab}^0)^2,$$

then

(1) their spectrum, which in equilibrium consists of two frequencies (9′), reduces to a single frequency Ω_G given by (9); and

(2) their decay rates, which in equilibrium equal T_2^{-1}, become different, $\Gamma_\pm$ (see Eqs. (9)).

If in these conditions the system is pumped above the generation threshold, which means that

$$\frac{1}{2}[\{\omega_{ab}^0\}^2+4T_2^{-2}]\leqslant 4I^2 x_a x_b|p_{a0}p_{b0}|z_n^2\left(\frac{J}{\hbar}\right)^2\leqslant(\omega_{ab}^0)^2+4T_2^{-2},$$

and with J negative the inverted spins are those with the higher partial frequency and

$$(\omega_{ab}^0)^2<4T_2^{-2},$$

then

(3) this spin system generates a pulse at Ω_G, which has almost all the features of superradiance (a delay time, an envelope shaped like a hyperbolic secant, and a narrow signal, $\Delta t \ll T_D$).

There is a significant difference between our case and ordinary superradiance, i.e., the dependence of radiant intensity on the number of active particles is stronger by a factor z_n^2 in comparison to the stationary case, while in ordinary superradiance this increase is N_{sp}-fold. But still, on the basis of our studies, we use the term superradiance for this transient generation process.

We have also calculated the parameters of this superradiance process (the voltage across the terminals of an rf pickup coil, the pulse width, and the superradiant intensity).

A similar approach can be adopted to obtain superradiance in any interaction-coupled systems consisting of two particle species without using external cavities or mode-selection schemes.

The authors would like to express their gratitude to V. A. Atsarkin for discussing the results and for useful remarks. They would also like to mention the invaluable comments on the manuscript done by the late L. L. Buishvili.

[1)]A similar shift in the relaxation times can be expected when the spins interact with a cavity. Indeed, a shift in the spin–spin relaxation time in the superradiance process at large nuclear polarizations was observed by Bazhanov *et al.*,[8] Flepp,[19] and Badii *et al.*[20]

[1] A. S. Oja, A. J. Annila, and Y. Takano, Phys. Rev. Lett. **65**, 1921 (1990).
[2] P. J. Hakonen, K. K. Nummila, and R. T. Vuorinen, Phys. Rev. B **45**, 2196 (1992).
[3] G. Eska and E. Schuberth, Jpn. J. Appl. Phys. **26**, Suppl. No. 3, 435 (1987).
[4] G. Eska, in *Proc. Conf. on Quantum Fluids and Solids, Gainesville, 1989*, G. G. Ihas, and Y. Takano (eds.), AIP Conf. Proc. **194.**
[5] Shin-ichi Kuroda, M. Motokawa, and M. Date, J. Phys. Soc. Jpn. **61**, 1036 (1992).
[6] P. Bosiger, E. Brun, and D. Meier, Phys. Rev. B **18**, 671 (1978).
[7] Yu. F. Kiselev, A. F. Prudkoglyad, A. S. Shumovskiĭ, and V. I. Yukalov, Zh. Éksp. Teor. Fiz. **94**, No. 2, 344 (1988) [Sov. Phys. JETP **67**, 413 (1988)].
[8] N. A. Bazhanov, D. S. Bulyanitsa, A. I. Zaĭtsev, A. I. Kovalev, V. A. Malyshev, and E. D. Trifonov, Zh. Éksp. Teor. Fiz. **97**, 1995 (1990) [Sov. Phys. JETP **70**, 1128 (1990)].
[9] N. P. Fokina, K. O. Khutsishvili, and S. G. Chkhaidze, Zh. Éksp. Teor. Fiz. **102**, 1013 (1992) [Sov. Phys. JETP **75**, 552 (1992)].
[10] A. V. Andreev, Usp. Fiz. Nauk **160**, No. 12, 1 (1990) [Sov. Phys. Usp. **33**, 997 (1990)].
[11] R. H. Dicke, Phys. Rev. **93**, 99 (1954).
[12] N. Skribanovitz, I. P. Hermann, J. C. MacGillivray, and M. S. Feld, Phys. Rev. Lett. **30**, 309 (1973).
[13] J. C. MacGillivray and M. S. Feld, Phys. Rev. A **14**, 1169 (1976).
[14] R. Bonifacio and L. A. Lugiato, Phys. Rev. A **11**, 1507 (1975).
[15] Yu. V. Naboĭkin, V. V. Samartsev, P. V. Zinov'ev, and N. B. Silaeva, *Coherence Spectroscopy of Molecular Crystals* [in Russian], Nauka Dumka, Kiev (1986).
[16] Yu. S. Karimov and I. F. Shchegolev, Zh. Éksp. Teor. Fiz. **41**, 1082 (1961) [Sov. Phys. JETP **14**, 772 (1962)].
[17] J. P. Estrom, J. F. Jacquinot, M. T. Loponen, J. K. Soini, and P. Kumar, Physica B & C **98**, 45 (1979).
[18] R. H. Pantell and H. E. Puthoff, *Fundamentals of Quantum Electronics*, Wiley, New York (1969).
[19] L. Flepp, PhD Thesis, Zürich University (1991).
[20] R. Badii, E. Brun, M. Finardi, L. Flepp, R. Holzner, J. Parisi, C. Reyl, and J. Simonet, Rev. Mod. Phys. **66**, 1389 (1994).
[21] N. N. Bogolyubov and Yu. A. Mitropol'skiĭ, *Asymptotic Methods in the Theory of Nonlinear Oscillations* [in Russian], 4th ed., Nauka, Moscow (1974).
[22] H. Haken, *Advanced Synergetics. Instability Hierarchies of Self-Organizing Systems and Devices*, Springer, Berlin (1983).
[23] R. M. Pomerantsev, V. M. Ryzhkov, and G. V. Skrotskiĭ, *Physical Bases of Quantum Magnetometry* [in Russian], Nauka, Moscow (1972).
[24] N. P. Fokina, K. O. Khutsishvili, S. G. Chkhaidze, A. M. Lomidze, Fiz. Tverd. Tela (St. Petersburg) **37**, 1910 (1995) [Phys. Solid State **37**, 1040 (1995)].
[25] T. Sleater, E. L. Hahn, C. Hilbert, and J. Clarke, Phys. Rev. Lett. **55**, 1742 (1985).

Translated by Eugene Yankovsky

Spin-wave resonances in nonuniformly strained films of $FeBO_3$

L. E. Svistov

Institute of Crystallography, Russian Academy of Sciences, 117333 Moscow, Russia

V. L. Safonov

Russian ''Kurchatov Institute'' Center for Science, 123182 Moscow, Russia

K. R. Khachevatskaya

University of Buenos Aires, Argentina
(Submitted 11 October 1997)
Zh. Éksp. Teor. Fiz. **112**, 564–573 (August 1997)

Spin-wave resonances are investigated in thin films of the antiferromagnet $FeBO_3$ with magnetic anisotropy of the easy-plane type. It is observed that high-order resonances are observed only when nonuniform stresses are created in the sample. In nonuniformly stressed samples the antiferromagnetic resonance lines are broadened, and against the background of this broadening spin-wave resonances are visible whose positions are well described within the free-particle approximation. Using this method it is possible to resolve resonances with wave numbers $\simeq 1.5\times 10^5$ cm^{-1}. The presence of strong uniaxial elastic stresses is established experimentally in nonuniformly strained films. A contact mechanism for exciting spin-wave resonances is discussed. © *1997 American Institute of Physics.* [S1063-7761(97)01308-5]

1. INTRODUCTION

In an ideal infinite crystal, linear excitation of spin waves is possible only when the pump field frequency ω_p coincides with the frequency of some spin wave ω_k (the law of conservation of energy) and the pump field wave vector k coincides with the wave vector of the same spin wave (the law of conservation of quasimomentum). The values of the wave vector of an electromagnetic microwave field are not large ($\leqslant 10^2$ cm^{-1}); therefore, linear excitation of spin waves with large wave numbers is impossible.

The presence of boundaries in a real sample causes the spectrum of spin waves to become discrete. Consequently, a nonzero coupling appears between a uniform microwave field and the large-wave-vector spin-wave modes of the sample. The efficiency of this resonant excitation should decrease with increasing values of the wave vector, and the microwave field should couple only to modes in which the magnetic moment averaged over the sample oscillates. For a sample in the form of a film this condition implies that only modes with odd numbers of half-waves couple to the magnetic microwave field. The efficiency of this coupling falls off in inverse proportion to the wave number k_z. For this reason, it is not usually considered possible to observe spin-wave resonances with large wave numbers $\sim 10^5$ cm^{-1}. The study of spin waves with these large wave vectors is made possible by creating a static magnetic field that is highly nonuniform along the sample, by creating artificial nonuniformities in the magnetic properties with respect to sample thickness (see, e.g., Refs. 1 and 2), or by nonlinear methods.[3]

In this paper we investigate the phenomenon of linear excitation of large-wave-number spin waves by a microwave magnetic field. A similar linear excitation of spin-wave modes was observed previously in the antiferromagnet $MnCO_3$ with easy-plane magnetic anisotropy.[4] However, in that paper the authors used samples with irregular shapes, and thus were unsuccessful in deciphering the observed resonances or analyzing the efficiency of their excitation.

We have investigated spin-wave resonances in thin films of the antiferromagnet $FeBO_3$ with the easy-plane type of magnetic anisotropy ($T_N = 348$ K). The grown faces of the films coincided with the plane of easy magnetization and were optically smooth. In order to observe spin-wave resonances the mean-free path of magnons λ must exceed the film thickness d. We studied films with thicknesses $\simeq 10-40$ μm. According to the data of Ref. 5, whose authors determined the magnon lifetime from the threshold for paramagnetic excitation of spin waves with wave vectors $\simeq 10^5$ cm^{-1}, the condition $\lambda > d$ should be satisfied at temperatures below 100–150 K.

The primary goal of this work is to study the excitation mechanism for spin-wave resonances and to determine whether it is possible to study their relaxation properties in terms of the linewidth.

The spectrum of the low-frequency magnon branch ω_k in $FeBO_3$ has the form

$$\omega_k = \gamma[H(H+H_D)+H_\Delta^2+(\alpha k)^2]^{1/2}, \tag{1}$$

where $\gamma = g\mu_B = 17.8\times 10^9$ s$^{-1}\cdot$kOe^{-1} is the gyromagnetic ratio, H is the static magnetic field, $H_D \simeq 100$ kOe is the Dzyaloshinskiĭ field, $H_\Delta^2 = 2H_E H_{me}$, $H_\Delta \simeq 1.9$ kOe, $H_E \simeq 2.6\times 10^6$ Oe is the exchange field, H_{me} is the magnetoelastic anisotropy field, and $\alpha \simeq 0.8\times 10^{-3}$ Oe$\cdot$cm is the nonuniform exchange constant. The values of these constants correspond to a range of temperatures far from the Néel temperature, i.e., $T \ll T_N$.

Our studies were made using standard EPR spectrometers from the Bruker Company in the microwave range 8 mm to 3 cm. The angular dependence was measured using a industrial-type EPR spectrometer.

 1063-7761/97/080307-06$10.00

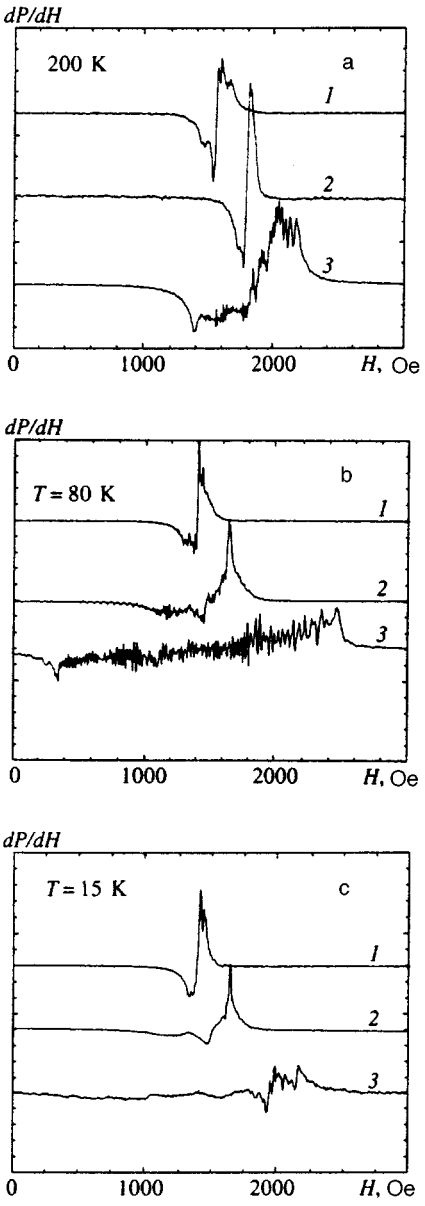


FIG. 1. Traces of the derivative of the absorption with respect to field (dP/dH) as a function of H for values of the temperature 200 K (a), 80 K (b), and 15 K (c). The traces shown in the figures and denoted by the number 1 correspond to an essentially unstrained sample glued to a holder by its ends. The number 2 denotes antiferromagnetic resonance traces taken for the same sample whose surface is coated by a thin layer of dilute glue. The antiferromagnetic resonance traces denoted by the number *3* were obtained for the same sample glued along a plane to a plastic holder.

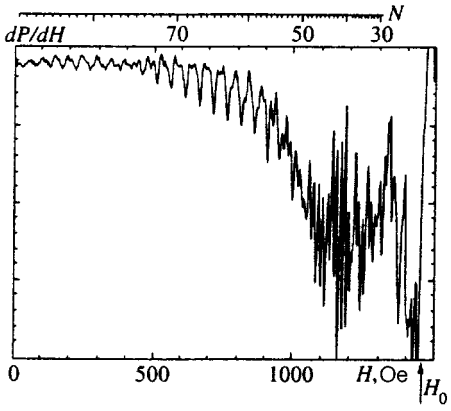


FIG. 2. A fragment of the trace of an absorption curve from a sample coated with a thin layer of glue at a temperature of 80 K (Fig. 1b, curve *2*). The arrows show the positions of spin- wave resonances obtained by calculating the values of the wave vector $k_z=(\pi/d)n$ using Eq. (1), where $d=16$ μm is the film thickness.

2. EXPERIMENTAL RESULTS AND DISCUSSION

2.1. Analysis of experimental data and formulation of a model

Figure 1 shows a trace of the derivative of the absorption with respect to field, dP/dH, versus H for various values of the temperature and various methods of attaching the sample. In these figures the traces denoted by the number *1* correspond to a practically unstressed sample glued to a holder at its ends. For this type of gluing we observed a narrow antiferromagnetic resonance line, and no spin-wave resonances were observed that corresponded to large wave numbers (i.e., in the range of small magnetic fields) over the entire temperature range 4.2–300 K.

The number *2* indicates antiferromagnetic resonance traces taken on the same sample whose surface was coated by a thin layer of dilute glue. The traces denoted by the number *3* were taken on the same sample glued over the plane to a holder made of plastic.

Note that at room temperature (300 K) the position and shape of the antiferromagnetic resonance line depends only slightly on the method of fastening the sample. As the temperature decreases, because the coefficients of thermal expansion of the sample, the glue, and the plastic substrate are all different the surface of the sample undergoes strain, which significantly broadens the antiferromagnetic resonance line. Superposed on this broadened antiferromagnetic resonance line we observed narrow resonance lines whose density and intensity increase near the resonance field H_0. The fine structure we observe has a simple explanation (given below) only for the case of the weak strain caused by the thin layer of glue. The strain caused by gluing the sample to the plastic substrate probably gives rise to bending strain, which is nonuniform throughout the sample thickness. This makes an explanation of the fine structure difficult, and hence it will not be discussed here.

In Fig. 2 we show a fragment of an antiferromagnetic resonance line trace for a sample coated with a thin film of glue at a temperature of 80 K (see Fig. 1b, curve *2*). On the same figure the arrows indicate the results of calculating the positions of spin-wave resonances using Eq. (1) with values of the wave number $k_z=\pi n/d$, where $d=16$ μm is the film thickness. Good agreement between the position of the lines and the calculated positions is obvious, up to $n=80$. This resonance corresponds to a wave number $k_z\simeq 1.5\times 10^5$ cm^{-1}. Spin-wave resonances with $n\leqslant 20$ near H_0 cannot be resolved, because the distances between them become comparable to the width of an individual line. The excitation efficiency of modes with even and odd numbers of half-waves n are roughly the same in the vicinity of the

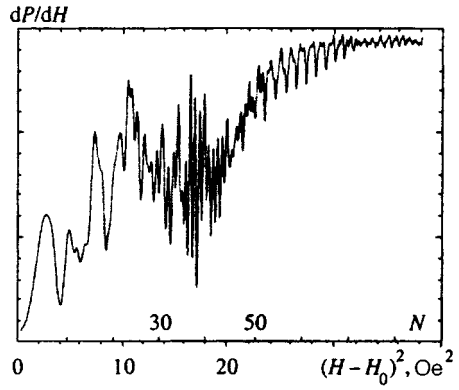


FIG. 3. The same trace of the absorption curve as in Fig. 2, but now $(H_0-H)^2$ is plotted along the abscissa. $H_0=1450$ Oe is the antiferromagnetic resonance field for an unstrained sample. In these coordinates, microwave resonances based on Eq. (1) should be equidistant.

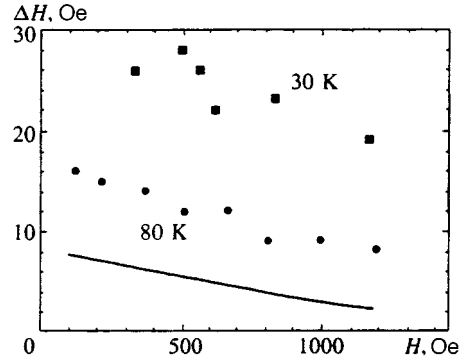


FIG. 4. Dependence of the spin-wave linewidth on the corresponding values of the resonance field at two temperatures $T=30$ K and $T=80$ K. The solid curve shows the results of calculations.

antiferromagnetic resonance field. However, far from H_0, each second resonance is considerably weaker than its neighbors.

Figure 3 shows the same fragment of antiferromagnetic resonance line trace, but the quantity $(H_0-H)^2$ is plotted along the abscissa. Here $H_0\simeq 1450$ Oe is the antiferromagnetic resonance field for an unstrained sample. In these coordinates, spin-wave resonances should be equidistant, which is confirmed by the figure. Note that there is a certain arbitrariness in choosing the value of H_0 within the width of the antiferromagnetic resonance line; however, for a free sample this has only a small effect on the distance between neighboring spin-wave resonances with large wave numbers.

Figure 4 shows the dependence of the spin-wave resonance linewidths on values of the corresponding resonance fields at two temperatures, $T=30$ K and $T=80$ K. It is clear that the linewidths decrease as the resonance field H_0 is approached. We assume that the spin-wave resonance linewidths are determined by relaxation processes rather than inhomogeneous broadening. As an upper estimate we find the inverse lifetime of a standing spin-wave τ_m^{-1}, which comes to 85×10^6 s^{-1} for modes corresponding to a magnetic field $H=120$ Oe. The inverse lifetime of spin waves with a frequency of 18 GHz obtained by measuring the threshold field for their parametric excitation[5] in bulk samples comes to 35×10^6 s^{-1} for the same values of T and H. In previous studies (see, e.g., Refs. 6 and 7) it was shown that in $FeBO_3$ it is the three-particle interactions of magnons with phonons that are most efficient at low temperatures. Using the frequency and field dependence of τ_m^{-1} given in these papers, and the value obtained in experiments on parametric excitation of magnons, we were able to calculate the expected dependence of the spin- wave resonance linewidth on magnetic field. The results of these calculations are shown in Fig. 4 by the solid curve. It is clear that magnon-phonon processes are satisfactorily described by the dependence of τ_m^{-1} on magnetic field. The portion of the spin-wave resonance linewidth that does not depend on magnetic field can be related to inhomogeneous broadening or to other relaxation mechanisms.

Note that the increase in the spin-wave resonance linewidth with decreasing temperature is also in agreement with the results of Ref. 5, in which a minimum in τ_m^{-1} was observed at a temperature around 60 K. It is in this range of temperatures (60–80 K) that the fine structure is most clearly evident in the form of well resolved narrow lines.

In Fig. 5 we show traces of the signal proportional to the power passing through the resonator as a function of the value of the static field. These traces were made on an unstrained sample and on a sample coated with a thin layer of glue. It is clear that in the second case the antiferromagnetic resonance lines are broadened, at least for three lines of different intensities. On the same figure we show the angular dependence of the position of the antiferromagnetic resonance line for an unstrained sample and the most intense line of a sample covered with glue. The position of the antiferromagnetic resonance line of an unstrained sample exhibits a 60° anisotropy (as in Ref. 5). In the case of an unstrained sample, each of the three well resolved lines exhibits a 180° anisotropy. The latter fact probably indicates a division of the sample into regions (in each of which a uniaxial compression is present in the easy plane of the crystal.

The effect of uniaxial compression on the spin-wave spectrum in antiferromagnets with the easy-plane type of magnetic anisotropy was studied experimentally and theoretically in Refs. 8 and 9. It was shown that the effect of uniaxial compression in the magnetization plane is equivalent to an effective anisotropy field H_p in the easy plane of the crystal directed perpendicular to the applied stress. The value of H_p is proportional to the value of the uniaxial compression p and depends on the elastic and magnetoelastic constants. (A calculation of the spectrum of spin waves in a crystal with uniaxial compression is given in the Appendix.)

Figure 6 shows calculated field dependence of the antiferromagnetic resonance frequency for various mutual orientations of the field H and the axis of elastic stress at two values of the stress. On the same figure, the solid curves show the field dependence of the antiferromagnetic resonance frequency of an unstrained crystal. A nonuniform strain along the sample surface causes the gap in the spin-wave spectrum to vary at various positions in the sample. Hence we observe regions in which the conditions for antiferromagnetic resonance are satisfied over a wide range of

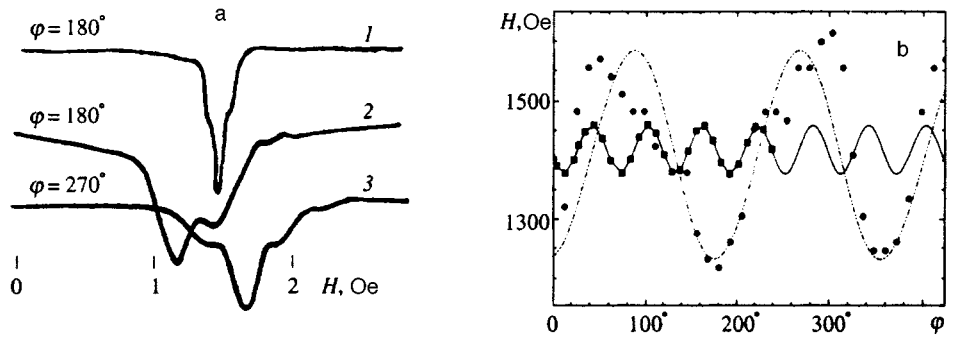


FIG. 5. a) Traces of the signal proportional to the microwave power passing through the resonator as a function of the value of the static field, taken for an unstrained sample (*1*) and for a sample coated with a thin layer of glue (*2, 3*) at a temperature of 4.2 K; b) angular dependence of the position of the antiferromagnetic resonance line: □—unstrained sample, ○—the most intense line from a sample coated with glue. The solid and dashed curves are the results of fitting to functions of the forms $H_0 = A + B\cos(3\varphi + C)$ and $H_0 = A + B\cos(2\varphi + C)$ respectively.

fields H. This simple model explains the observed broadening of the antiferromagnetic resonance line caused by gluing the sample.

The increase in efficiency of excitation of spin-wave resonances under conditions of nonuniform strain of the sample can also be explained in terms of this model. Oscillations corresponding to large wave numbers are excited by a contact mechanism in regions adjacent to those in which the antiferromagnetic resonance conditions are satisfied.

This mechanism causes all of the neighboring regions to be excited. Moreover, each of the regions will exhibit spin-wave resonance at values of the field corresponding to the magnetoelastic gap in that region. Since the uniaxial strains caused by gluing the film have a random character, it is natural to expect a random superposition of spin-wave resonances from these regions. However, regions in which hydrostatic compression takes place (or where there is no strain at all) will exhibit spin-wave resonances at the same values of the field H determined by Eq. (1). We are probably observing resonances excited in these regions in our experiments in the form of fine structure in the absorption line.

An argument in favor of the contact mechanism for exciting spin-wave resonance is the fact that resonances are observed both for even and odd mode numbers.

2.2. Testing the model

A key step in building the model described above is dividing the sample into regions with differing uniaxial stresses. In order to test this hypothesis we investigated the antiferromagnetic resonance spectrum with the intent of observing additional absorption lines from regions subjected to uniaxial stress.

A characteristic feature of the spin wave spectrum in an antiferromagnet subjected to uniaxial compression **p** is a decreasing dependence of the antiferromagnetic resonance frequency on the magnitude of the static field H for the parallel orientation $\mathbf{p} \| \mathbf{H}$ up to a critical value of the field H_c determined by the value of the uniaxial stress (see Fig. 6 and the Appendix). The lowest antiferromagnetic frequency at the field H_c will become equal to the resonant frequency of an unstrssed crystal in zero field. For $FeBO_3$ this equals 4.5 GHz; therefore, we searched for antiferromagnetic resonance signals from those regions of the sample in which uniaxial compression is parallel to the static field at a frequency of 9.3 GHz.

Since the decreasing branch of the antiferromagnetic resonance in iron borate had not been experimentally studied previously, at the beginning we investigated samples with artificially created uniaxial compression, which was created in a thin layer of $FeBO_3$ ($d = 16$ μm) in the following way. First the sample was glued to a thin cover glass (with thickness 300 μm) using epoxy resin. Then the structure was glued to a carrier that provided bending of the substrate, as illustrated in the inset to Fig. 7. This bending of the substrate corresponds to uniform uniaxial compression along the thickness of a thin sample. The entire structure was placed in the microwave resonator of a spectrometer in the 3-centimeter range.

Figure 7 shows traces of the absorption spectrum (dP/dH) for various values of the angle between the direc-

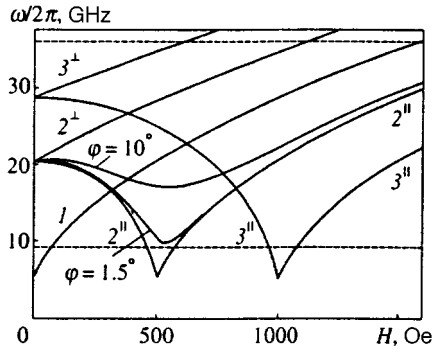


FIG. 6. Computed dependence of the antiferromagnetic resonance frequency on the value of the static field H for a unstressed sample (curve *1*) and for a sample subjected to uniaxial compression (curves 2 and 3). The values of uniaxial stress correspond to $H_c = 500$ Oe and 1000 Oe respectively for curves *2* and *3*. The labels ∥ and ⊥ correspond to parallel and mutually perpendicular orientations of the external field **H** and the compression axis **p**. The two remaining curves correspond to dependence of the antiferromagnetic resonance frequency on the value of the static field for the small angles 1.5° and 10° between **H** and **p**, respectively; $H_c = 500$ Oe. The dashed curves show the working frequencies for the spectrometers used in this paper.

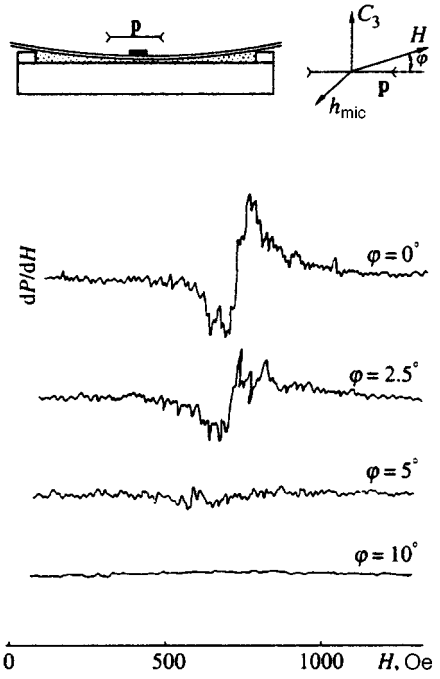


FIG. 7. Traces of absorption spectra dP/dH from a field H for various values of the angle between the directions of the static field and the compression axis: $\varphi = 0°$, 2.5°, 5°, and 10°. The microwave oscillator frequency was 9.3 GHz, $T = 80$ K. The inset shows a sketch of the substrate that creates uniaxial compression in the thin-film sample.

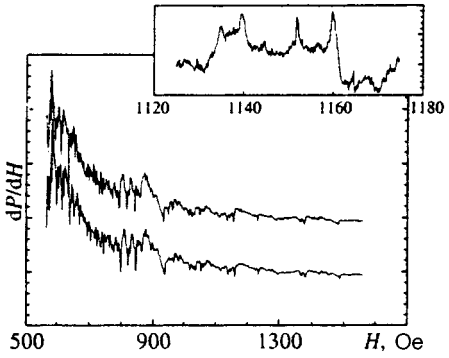


FIG. 8. Traces of absorption spectra dP/dH versus field H taken in a single experiment for a sample coated with a thin layer of glue. The inset shows a fragment of the trace shown on a larger scale. The microwave oscillator frequency was 9.3 GHz, $T = 80$ K.

tions of the static field and the compression axis. Noteworthy is the fact that the absorption lines, as we expected (see Fig. 6) lie within a narrow interval of angles. In this case the resonant field should be close to the field H_c associated with uniaxial compression. An estimate of the value of the uniaxial compression created in our experiments gives $p \sim 8 \times 10^8$. Moreover, near the fundamental resonance many narrow lines appear whose density and intensity falls off as we deviate from the field H_c, and also as the angle between the compression axis and the field **H** increases.

Figure 8 shows the trace of an absorption line for the same sample whose surface is coated with a thin film of glue, but without uniaxial compression. In this case we also observe many narrow absorption lines whose density and intensity decreases as the magnetic field increases. These lines can be observed up to a field $\simeq 2$ kOe. These lines are quite reproducible during a single experiment, but after heating and subsequent cooling the pattern of resonances is completely changed. The inset shows a fragment of the trace in an expanded scale.

Note that the region of magnetostatic and surface modes of an unstrained sample is at least bounded by the field $H_0 \pm 4\pi M_0 H_D / H_E$. The antiferromagnetic resonant field H_0 for an unstrained sample equals 60 Oe according to Eq. (1), and $4\pi M_0 H_D / H_E = 240$ Oe. Thus, in an ideal crystal in the range of fields $H > 300$ Oe we should not be generating magnetic excitations with a frequency of 9.3 GHz.

The set of narrow lines observed in this experiment (see Figs. 7 and 8) is naturally explained by resonances from various regions of the crystal with uniaxial stresses created by the glue that differ in magnitude. This agrees with the model described above.

It is interesting to note that in samples with a clean surface we also observe narrow resonances, only their density was considerably smaller and they were located in regions of considerably smaller fields (200–300 Oe). This probably indicates that in natural crystals strong nonuniform elastic stresses exist that possibly arise during the growth process.

3. CONCLUSION

The analysis given in this paper allows us to formulate our main results:

1. In a nonuniformly strained film of $FeBO_3$ spin-wave resonances were observed with record high values of wave numbers corresponding to $k_z \simeq 1.5 \times 10^5$ cm^{-1}.

2. The width of the spin-wave resonance lines is in good agreement with the inverse lifetime of the spin waves obtained from measuring the threshold for parametric excitation of magnons in bulk samples. This attests to the negligibly small losses that occur when spin waves are reflected from the crystal boundaries.

3. A contact mechanism has been proposed for the excitation of spin-wave resonances in nonuniformly stressed films. Using this model, we can qualitatively explain the high efficiency of excitation of spin-wave resonances with large wave numbers.

4. We have shown experimentally the presence of strong uniaxial stresses in a nonuniformly strained film.

5. We have experimentally observed the decreasing branch of the antiferromagnetic resonance spectrum near the spin-orientation phase-transition field in a uniaxially strained crystal of $FeBO_3$.

The authors are grateful for useful discussions with N. M. Kreĭnes, L. A. Prozorova, and A. I. Smirnov. They thank V. N. Selezneva for providing the high-quality $FeBO_3$ single crystals. This work was carried out with the support of the Russian Fund for Fundamental Research (Project No. 96-02-16575).

APPENDIX

Let us write the free-energy density of an antiferromagnet, including magnetic, elastic, and magnetoelastic subsystems, in the form:[9]

$$\mathscr{F}=\mathscr{F}_m+\mathscr{F}_e+\mathscr{F}_{me}.$$

where

$$\mathscr{F}_m=2M_0\left[H_E m^2+\frac{H_A}{2}\,l_z^2-H_D(m_x l_y-m_y l_x)-m_x H\right],$$

$$\mathscr{F}_e=\frac{1}{2}\,\mathscr{C}_{11}(u_{xx}^2+u_{yy}^2)+\frac{1}{2}\,\mathscr{C}_{33}u_{zz}^2+\mathscr{C}_{12}u_{xx}u_{yy}$$
$$+\mathscr{C}_{13}(u_{xx}+u_{yy})u_{zz}+2\mathscr{C}_{44}(u_{yz}^2+u_{xz}^2)+2\mathscr{C}_{66}u_{xy}^2$$
$$-\sigma_{xx}u_{xx}-\sigma_{yy}u_{yy}-2\sigma_{xy}u_{xy},$$

$$\mathscr{F}_{me}=\mathscr{B}_{11}(u_{xx}l_x^2+u_{yy}l_y^2)+\mathscr{B}_{12}(u_{yy}l_x^2+u_{xx}l_y^2)$$
$$+\mathscr{B}_{33}u_{zz}l_z^2+2.\mathscr{B}_{44}(u_{yz}l_y+u_{xz}l_x)l_z$$
$$+2.\mathscr{B}_{66}u_{xy}l_x l_y.$$

Here M_0 is the saturation magnetization of one sublattice, H_A is the anisotropy field, and **l**, **m** are the unit vectors for antiferromagnetism and magnetization ($m^2+l^2=1$, $\mathbf{m}\cdot\mathbf{l}=0$), u_{ij} is the strain tensor, and $\mathscr{C}_{ij}$, $\mathscr{B}_{ij}$ are respectively the elastic and magnetoelastic constants. The mechanical stress tensor σ_{ij} is determined by the uniaxial compression **p**, whose axis deviates from the x axis by an angle φ. Since in this model we assume isotropy of all properties in the basal plane, the following relations hold between the constants: $2\mathscr{C}_{66}=\mathscr{C}_{11}-\mathscr{C}_{12}\simeq 3\times 10^{12}$ erg/cm^3, $\mathscr{B}_{66}=\mathscr{B}_{11}-\mathscr{B}_{12}\simeq 1.7\times 10^{7}$ erg/cm^3.

The ground state of the magnetic subsystem is determined by the angle φ and the equilibrium angle ψ by which the vector **m** in the basal plane deviates from the field **H** directed along the x axis. The part of the free energy that depends on angle has the form

$$\Delta\mathscr{F}^{(0)}=-\frac{M_0}{2H_E}\{[H_D+H\cos\psi]^2-H_E H_p\cos[2(\psi+\varphi)]\},$$

where $H_p=(\mathscr{B}_{11}-\mathscr{B}_{12})p/M_0(\mathscr{C}_{11}-\mathscr{C}_{12})$ is the effective field created by the one-sided pressure. Minimizing $\Delta\mathscr{F}_0$ we find ψ. Thus, for $\varphi=0$ it is known[9] that $\cos\psi=HH_D/(2H_E H_p-H^2)$ for $H<H_c$ and that $\psi=0$ in the opposite case. The critical field H_c is determined from the relation $H_c(H_c+H_D)=2H_E H_p$. For $\varphi=\pi/2$ we obtain $\psi=0$. We were unable to find ψ analytically for $0<\varphi<\pi/2$. In this case the minimum of the free energy can be determined by a numerical method.

The following expression gives the antiferromagnetic resonance frequency of the low-frequency branch of the spectrum:

$$\omega_0=\gamma\{H[H_D+H\cos\psi]\cos\psi+2H_E$$
$$\times[H_{me}-H_p\cos 2(\psi+\varphi)]\}^{1/2},$$

where $H_{me}=(\mathscr{B}_{11}-\mathscr{B}_{12})^2/M_0(\mathscr{C}_{11}-\mathscr{C}_{12})$. The dependence of the antiferromagnetic resonance frequency on magnetic field for several values of φ is shown in Fig. 6.

[1] A. G. Gurevich and G. A. Melkov, *Magnetic Oscillations and Waves* (CRC Press, Boca Raton, FL, 1996).
[2] N. E. Zilberman, A. G. Temiryazen, and M. P. Tikhomirova, Zh. Éksp. Teor. Fiz. **108**, 281 (1995) [JETP **81**, 151 (1995)].
[3] B. Y. Kotyuzhanskii and L. A. Prozorova, Space Sci. Rev. **A13**, 1–131 (1990).
[4] B. Y. Kotyuzhanskii, L. E. Svistov, and H. Benner, J. Phys. C **3**, 9253 (1991).
[5] B. Y. Kotyuzhanskiĭ and L. A. Prozorova, Zh. Éksp. Teor. Fiz. **81**, 1913 (1981) [Sov. Phys. JETP **54**, 1013 (1981)].
[6] B. Y. Kotyuzhanskiĭ, L. A. Prozorova, and L. E. Svistov, Zh. Éksp. Teor. Fiz. **92**, 238 (1987) [Sov. Phys. JETP **65**, 134 (1987)].
[7] A. S. Mikhaĭlov and A. V. Chubukov, Zh. Éksp. Teor. Fiz. **86**, 1401 (1984) [Sov. Phys. JETP **59**, 819 (1984)].
[8] A. S. Borovik-Romanov and E. G. Rudashevskiĭ, Zh. Eksp. Teor. Fiz. **47**, 2095 (1964) [Sov. Phys. JETP **20**, 1407 (1964)].
[9] V. G. Bar'yakhtar and E. A. Turov, in *Spin Waves and Magnetic Excitations*, A. S. Borovik-Romanov and S. K. Sinha, eds. (Elsevier Sci. Publ., Amsterdam 333, 1988).

Translated by Frank J. Crowne

Quantitative conversion spectroscopy of the ultrasoft isomeric transition of uranium-235 and the electronic structure of uranium oxides

A. D. Panov

Russian Scientific Center ''Kurchatov Institute,'' 123182 Moscow, Russia
(Submitted 17 October 1996)
Zh. Éksp. Teor. Fiz. **112**, 574–595 (August 1997)

Combined measurements of conversion electron spectra and the decay constant (76.5 eV, $(1/2)^+ \rightarrow (7/2)^-$) of the $E3$-isomeric transition of the uranium-235 nucleus have been performed with collection of the isomer atoms on an indium surface. The conversion spectra are interpreted as corresponding to a mixture of two different oxides A and B of uranium, one of which (A) is similar to UO_2, and the other (B) consists of a uranium–oxygen cluster based on the linear uranyl group O–U–O. From a set of mixed experimental spectra conversion spectra have been found corresponding to the chemical states A and B of the isomer atoms, and the variation of the absolute intensities of the conversion lines has been quantitatively investigated for them by varying the chemical composition of the isomer atoms and the ratio between the intensities of various conversion lines of the B spectrum. Experimental ratios between the intensities of the conversion lines are compared with the expected ratios in accordance with the distribution of the $6p$ electron density in the uranyl group. It is concluded that the experimental data agree with the calculation and that abrupt violations of proportionality of the partial probabilities of conversion of the electron density near the nucleus are absent. In accordance with the hypothesis of proportionality of the partial probabilities of conversion, an experimental estimate is given of the degree of localization of the deep-lying uranium $6p_{1/2}$ shell during formation of the chemical bond in the uranyl group: around 70% of the $6p_{1/2}$ electron density remains in the quasi-atomic uranium shell and around 30% is transferred to hybrid molecular orbitals.
[S1063-7761(97)01408-X]

1. INTRODUCTION

As theoretical and experimental studies have shown, the internal conversion electron spectra of nuclear reactions can be extraordinarily informative about the electronic structure of matter if an energy resolution on the order of 1 eV is achieved in such spectra. A study of variations of the shape of the conversion spectrum associated with a change in the chemical environment of the converting atom would be of special interest. High-resolution conversion spectroscopy, used to investigate the electronic structure of matter, has acquired the name conversion electron spectroscopy.[1] However, as a consequence of a number of experimental difficulties it appears that so far it has been possible to extend the technique of conversion electron spectroscopy only to the soft conversion transitions of five nuclei: ^{235}U, ^{99}Tc [see references in Ref. 1 (review)], ^{119}Sn, ^{73}Ge (Ref. 2), and ^{201}Hg (Ref. 3). The ultrasoft isomeric transition of the ^{235}U nucleus occupies a special place among these objects.

The ^{235}U nucleus has the excited state $(1/2)^+$ (Refs. 4 and 5), whose energy according to the most recent data is only 76.5 ± 0.4 eV (Ref. 6). This is a long-lived isomeric state (^{235m}U), and decay of the nucleus to the ground state $(7/2)^-$ is realized via the almost completely converted $E3$ transition (conversion coefficient $\sim 10^{21}$, Ref. 7) with half-life around 26 min.[4,5] The conversion process is energetically resolved for the filled subshells $(6s_{1/2})^2$, $(6p_{1/2})^2$, $(6p_{3/2})^4$, and for the valence electrons of the uranium atom.

The isomeric transition of ^{235m}U is the softest of the nuclear transitions that can be reliably observed at present; therefore, the conversion process of this transition possesses a number of unique properties (e.g., the possibility of manifesting strong interference effects, see below). Only one case is known in which the nuclear excitation energy is less than for ^{235m}U, and that is ^{229m}Th with an excitation energy of 3.5 ± 1.0 eV (Ref. 8); however, there are still no direct experimental observations of the decay of this isomer, and its decay mechanism is unknown. All other known nuclear transitions, which, in principle, can be observed by conversion spectroscopy, have energies not less than 1 keV (Ref. 1). Let us touch on the nature of the low-lying $(1/2)^+$ state of the ^{235}U nucleus.

The ^{235}U nucleus is a strongly deformed ($\delta \approx 0.25$, Ref. 9, p. 125) odd nucleus. The lower energy diagram of this nucleus is qualitatively well interpreted as a set of rotation bands based on single-particle states of the unpaired nucleon in the self-consistent field of the deformed even core (Ref. 9, Ch. 5, Sec. II). In a spherically symmetric potential the characteristic separations between the single-particle levels are on the order of 1 MeV with degeneracy in the magnetic quantum number m_j. For violation of spherical symmetry the degeneracy in m_j is removed, and each state splits into several sublevels, with the magnitude of the splitting rapidly growing with growth of the deformation of the potential. Starting with deformations $\delta \sim 0.1$, the sublevels belonging to various initial (nlj)-subshells of the spherically symmetric potential begin to intersect; hence the possibility arises of forming very closely situated single-particle states.

The $(7/2)^-$ ground state of the ^{235}U nucleus and the isomeric state $(1/2)^+$ according to the experimental data of Ref. 10 are interpreted as the initial states of rotational bands constructed, respectively, on the $[743]^{7-}_{2}$ and $[631]^{1+}_{2}$ orbitals of the neutron in the field of the even core of ^{234}U. Therefore, it may be surmised that the proximity of the $(7/2)^-$ and $(1/2)^+$ states of uranium-235 is explained by an overlap of the magnetic sublevels of the various initial (nlj)-shells of the spherically symmetric potential. This assumption is confirmed by calculations of single-particle states of the neutron in the Nilson[11] and Saxon–Woods[12] potentials, which are qualitatively similar. According to these calculations, the $[743]^{7-}_{2}$ and $[631]^{1+}_{2}$ states are associated, respectively, with the $1j_{15/2}$ and $3d_{5/2}$ states of the initial spherically symmetric potential. In the spherically symmetric potential the $1j_{15/2}$ and $3d_{5/2}$ states are separated in energy by more than 1 MeV. However, for deformations $\delta=0.2-0.4$ the sublevels $[743]^{7-}_{2}$ and $[631]^{1+}_{2}$ approach each other and almost overlap. Herein lies the qualitative explanation of the proximity of the $(7/2)^-$ and $(1/2)^+$ states. Indeed, the picture is somewhat more complicated, since these states are purely single-particle states—they may contain a noticeable admixture of collective excitations.[12,13]

The principles of the experimental technique of conversion electron spectroscopy applied to ^{235m}U conversion were laid down in Ref. 14, and in Ref. 15 results were obtained by means of a new technique which are very important for the theory of the chemical bond of heavy elements (formation of inner valence molecular orbitals—molecular orbitals with the participation of deep, completely filled atomic shells) but which are of a qualitative character. It would be of extreme interest to make the transition to quantitative studies of the electronic structure of uranium compounds using conversion electron spectroscopy, and this is the main aim of the present paper.

To refine the statement of the problem, it is necessary to say a few words about the correspondence between current theory and experiment in the conversion spectroscopy of the uranium-235 isomer. Ground-breaking works in the theory of conversion of the uranium isomer are those of Grechukhin and Soldatov.[13,16] First of all, these papers established the abrupt difference in probability of conversion of the $6p_{1/2}$ and $6p_{3/2}$ electrons of uranium (binding energy, respectively, around 30 eV and 20 eV, Ref. 17). Second, they investigated the not entirely trivial question of spatial localization of the conversion process of ^{235m}U. The degree of localization of the transition is determined by the rate of convergence of integrals of the form $\langle f|r^{-4}|i\rangle$ (Refs. 13 and 16) as a function of the radius R of the region of integration. Here $|i\rangle$ is the initial state of the electron in an atomic orbit and $|f\rangle$ is its final state in the continuum. In Ref. 16 it was shown in numerical calculations that for the $6p$ and $6d$ electrons of uranium, integrals of this type converge to within 1% of their total value within a region of radius $0.1a_0$, where a_0 is the Bohr radius. A sphere of radius $0.1a_0$ may be taken as the region of localization of the conversion process of ^{235m}U. And finally, for the case of conversion in a single uranium atom it was shown that with very high accuracy (better than 1%) the partial probabilities of conversion on the $6p_{1/2}$ and $6p_{3/2}$ electrons of uranium are proportional to the square of the corresponding wave functions at the nucleus and depend on neither the binding energy[13] nor the configuration of valence shells of the atom.[16] It is specifically this property of the conversion spectrum of the uranium-235 isomer that is important for investigating the electronic structure of matter. Estimates have been given using the approach developed in Refs. 13, 16, and 18 of possible chemical variations of the decay constant of the uranium isomer. Variations of the decay constant of the uranium isomer were studied experimentally in Refs. 19–25. The scale of chemical variations of the decay constant observed in these studies (3–5%) is somewhat smaller than the maximum possible variations of the decay constant theoretically predicted.[13,16,18] This might have been expected since quite exotic configurations of the electronic shell of the uranium atom were considered, which apparently are not realized in experiment. In other respects it may be stated that at this stage experiment is in reasonable agreement with theory.

References 13 and 16 did not take into account variations in the amplitudes of the continuum states of the conversion electron in the conversion zone due to elastic scattering of the electron by atoms of its environment and subsequent interference. This effect was considered in a paper by Dobretsov.[26] Instead of the small corrections to the conversion probability which might have been expected, in a number of cases the interference effect turns out to be quite large. Thus, according to Ref. 26, in the conversion of the uranium isomer in the atomic cluster UAg_{12} corrections to the partial probabilities of conversion on individual atomic orbitals reach 89% of their unperturbed values, and the value and sign of the correction oscillate with variation of the energy of the orbital and with variation of the distance to the atoms of the environment. In this case the direct connection between the line intensities of the conversion spectrum and the electron densities at the nucleus established earlier in Refs. 13 and 16 is completely lost.

However, the results of Ref. 26 apparently do not completely agree with the experimental facts. Indeed, although Ref. 26 deals only with decay of the uranium isomer in an environment of silver atoms, the predicted interference effect should undoubtedly also manifest itself in many other cases. It should lead to characteristic variations of the decay constant of the uranium isomer on a scale of 20% or even greater (up to 47%, Ref. 26), which significantly exceeds the values observed in experiment. The experimentally observed variations of the decay constant most likely have a purely chemical magnitude, and interference effects apparently do not manifest themselves in any way.

It should be noted, however, that variations of the decay constant in the given case are not a very good test of the theory. The possibility cannot be excluded that interference gives large contributions to the partial probabilities of conversion belonging to individual atomic orbitals where these contributions have opposite sign and mutually cancel so that the decay constant varies only slightly. Observations of the chemical variations of the intensities of separate lines of the conversion spectrum would give much more detailed data.

These considerations lead to the following refinement of the formulation of the problem of the present work. First of all, it is necessary to develop a technique that would allow one to experimentally measure the absolute values of the partial probabilities of conversion belonging to individual atomic or molecular orbitals or ratios of these probabilities. The quantities obtained with the help of the new experimental technique could then be compared with the expected electron density distribution on the uranium nucleus or with chemical variations of the electron density in order to elucidate to what extent the experimental data support proportionality between the partial probabilities of conversion and the electron densities.[13,16] If indications upholding such a proportionality are obtained, then we must explain why the interference effects predicted in Ref. 26 not show up in experiment. Finally, if an understanding is reached here, then conversion spectroscopy of the uranium isomer may be used in practice to investigate the electron density distribution in uranium-containing molecules and clusters.

2. EXPERIMENTAL TECHNIQUE AND SAMPLES

To prepare samples containing the ^{235m}U isomer, the technique developed in Refs. 14 and 15 is used in the present work with minor modifications and additions. The ^{235m}U isomer was obtained as the product of α decay of ^{239}Pu in $^{239}PuO_2$ deposited as a thin layer (~ 20 μm/cm^2) on the inner surface of a platinum hemisphere of radius 40 mm (Ref. 27). The ^{235m}U recoil atoms were collected by an electric field onto metal substrates placed near the center of the hemisphere. All experimental results presented in this paper were obtained by collecting the isomer in an air atmosphere at a pressure of 70 GPa. It was found that the composition of the gaseous medium in the accumulation chamber only weakly affects the chemical state of the isomer atoms in the samples, as may be assessed from the form of the measured conversion spectra. The working region of the sample was bounded by a teflon mask with a rectangular opening with dimensions 0.5×5 mm^2. The typical collection time of an isomer atom was 40 min, and the discharge current could be regulated within the limits 1–5 nA. The potential difference between the sample and the plutonium hemisphere was around 1 kV (minus on the sample). The composition of the surfaces of the samples after isomer collection was monitored by x-ray electron spectroscopy and by α-activation of ^{239}Pu. Contamination of the surface by plutonium or teflon decomposition products was not observed. The total activity of the sample immediately after termination of isomer collection was usually on the order of 10^5 uranium isomer conversion decays per second. The samples were transferred in air to the electron spectrometer after unsealing of the accumulation chamber.

In the collection of uranium isomer onto the surface of a metallic sample the isomer atoms do not at first reach the metal surface, but first reach the film of hydrocarbon contaminants sorbed in the metal surface. The characteristics of the hydrocarbon film for copper substrates which have been taken through a typical cycle of uranium isomer collection were examined in Ref. 28 by x-ray electron spectroscopy. The film thickness amounts to around 15 Å, the total atomic density of oxygen and carbon is roughly equal to the atomic density of copper, i.e., 8.5×10^{22} cm^{-3}, and the carbon and oxygen atoms are present in different concentrations (within the limits of statistical error). The uranium isomer atoms succeed in braking completely in the hydrocarbon film, where with a probability of nearly 100% they bind with oxygen.[29]

To measure the conversion spectra of ^{235m}U we used a production-model NR5950A electron spectrometer[30] with a pre-acceleration system built by the authors of Ref. 14. The nominal working energy range of the spectrometer is $300-1500$ eV. Analysis of the electron energies is properly performed with the help of a spherical electrostatic analyzer (180°) having a constant potential difference between the electrodes. The measurements are performed simultaneously over an energy interval approximately 10 eV in width located near 115 eV. Before entering the spherical analyzer, the electrons are focused and braked to the necessary energy with the help of a preparatory electron-optical system consisting of four electron lenses. If it is required to measure a spectral interval wider than 10 eV, then the spectrum is scanned by varying the braking potential. In order to fit the uranium isomer conversion spectrum (which extends from 0 to roughly 75 eV) into the working range of the spectrometer, an accelerating voltage of -500 eV relative to the entrance to the first (grounded) electron lens of the preparatory electron-optical system is applied to the sample. The samples are placed in the spectrometer chamber, which has a working vacuum on the order of 10^{-9} Torr, through a transfer drawer with differential evacuation. The procedure of placing a sample in the spectrometer takes 1–2 min. The conversion spectra were measured in energy intervals of 20 and 50 eV and written to the 256-channel analyzer of the spectrometer. The energy resolution of the spectrometer is better than 1 eV (Ref. 14) and does not depend on the electron energy. The procedure of energy calibration of the conversion spectra is described in Sec. 4.

3. QUANTITATIVE ANALYSIS OF CONVERSION SPECTRA OF URANIUM-235

Figure 1 displays two experimental uranium-isomer conversion spectra measured with the help of the NR5950A spectrometer that are typical for uranium-isomer collection onto metal substrate surfaces. An obstacle to correct analysis of the spectra is presented by two types of distortions of the shape of the initial conversion spectrum. The first type of distortion is associated with inelastic scattering of the conversion electrons in the sample. Each conversion peak generates a low-energy tail consisting of inelastically scattered and secondary electrons. By adding together, these tails create an intense background of complex shape on which the conversion peaks are imposed (Fig. 1). Another distorting factor consists of instrument effects, the main one of which is the dependence of the effective geometrical factor of the spectrometer on the electron energy (variability of the spectrometer transmission function). Thus, mathematical processing of the spectra should consist first of correction of the experimental spectra for inconstancy of the spectrometer transmission function, then of subtracting out the inelastic

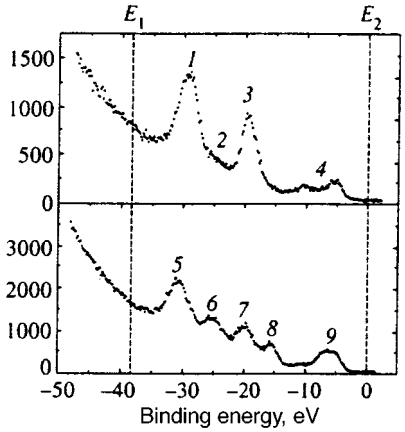


FIG. 1. Experimental conversion spectra of the uranium isomer for isomer collection on an indium surface. The spectra differ in the chemical state of the isomer (see Sec. 4) and correspond approximately to the extreme possibilities realized in experiment. The binding energy of the electrons corresponding to the conversion peaks is plotted along the abscissa; the kinetic energy increases from left to right. (E_1,E_2) is the informative spectral interval. The labels on the lines in the figure are as follows: *1*—''$6p_{1/2}$,'' *2*—''O2*s*,'' *3*—''$6p_{3/2}$,'' *4*—*VB*, *5*—''$6p_{1/2}$,'' *6*—''O2*s*,'' *7*—''$6p_{3/2}(l)$,'' *8*—''$6p_{3/2}(r)$,'' *9*—*VB*.

background, and finally in determining the areas (which we also call intensities) of the widths[1)] and positions of the lines. The latter problem is solved here by decomposing the spectra into lines of Gaussian shape by using the maximum likelihood method. The adequacy of the decomposition was monitored by the quantity χ^2/N (N is the number of degrees of freedom, χ^2/N should be close to 1, which in fact was always the case); the covariance matrix of the fitting parameters was calculated by the Monte Carlo method. Let us comment on the subtraction of the inelastic background and correction for instrument distortions.

We introduce the following terminology and notation. We call the spectral interval (E_1,E_2) occupied by true conversion lines the informative interval (Fig. 1). We denote by $I(E)$ the true conversion spectrum, by $Y(E)$ the sample spectrum (i.e., the spectrum $I(E)$ distorted by inelastic processes), and by $Q(E)$ the observed spectrum (i.e., the spectrum $Y(E)$ distorted by instrument factors). Here E denotes the kinetic energy of the electrons. By spectra we understand the corresponding probability distributions normalized in some convenient way. Accordingly, all equalities derived below must be understood as accurate to within arbitrary constant factors. We denote by $a(\xi)$ the instrument function of the spectrometer, responsible for the smearing of the spectrum by the instrument resolution. This function is assumed to be normalized to unity and not depend explicitly on the electron energy at entrance to the spectrometer. The tilde above the spectrum symbol denotes the corresponding spectrum smeared by the instrument resolution, i.e., $\tilde{Y}(E)=\int a(E-E')Y(E')dE'$, etc. Thus, assuming that $T(E)$ varies slowly in comparison with the width of the instrument function, it can be shown that

$$Q(E)=T(E)\tilde{Y}(E), \tag{1}$$

where $T(E)$ is the spectrometer transmission function, and

$$Y(E)=k(E)I(E)+\int_E^{E_2} I(E')R(E',E-E')dE'. \tag{2}$$

In Eq. (2) $k(E)$ is the probability of escape of a conversion electron with initial energy E from the sample without inelastic scattering, and $R(E',E-E')$ describes the shape of the inelastic tail for the conversion electrons with initial energy E'. We introduce the new functions

$$F(E)=k(E)I(E), \quad P(E',E-E')=[k(E')]^{-1} \times R(E',E-E'). \tag{3}$$

Assuming that the dependence of the function $P(E',E-E')$ on its first argument is weak in comparison with the width of the instrument function, it is not hard to show that

$$\tilde{Y}(E)=\tilde{F}(E)+\int_E^{E_2}\tilde{F}(E')P(E',E-E')dE'. \tag{4}$$

The function $\tilde{F}(E)$ has the following physical meaning: this is the spectrum of the electrons that have escaped from the sample without inelastic scattering (i.e., it is the elastic signal), smeared by the instrument resolution. Correspondingly, Eq. (4) also has a simple physical meaning, namely that the total signal leaving the sample ($\tilde{Y}$) is the elastic signal ($\tilde{F}$) plus the inelastic background (the integral), and the procedure for solving Eq. (4) for the spectrum $\tilde{F}$ if the spectrum $\tilde{Y}$ is known is to subtract out the inelastic background from the conversion spectrum.

Obviously, the spectrum $\tilde{F}(E)$, generally speaking, does not coincide with the true conversion spectrum $I(E)$. However, analyzing to first order we may limit the discussion to the spectrum $\tilde{F}(E)$ in place of $I(E)$ for the following reason. As Monte Carlo studies of the process of inelastic scattering of conversion electrons in a hydrocarbon film have shown[31] and as follows from experimental dependences of the electron mean free path on energy,[32] the function $k(E)$ is proportional with good accuracy to $I(E)$. The variability of $k(E)$ (on the order of 10% of the mean value over the informative spectral interval) is rooted in the systematic error of determining the line intensities.

Equations (1) and (4) make it possible from the observed experimental spectrum $Q(E)$ first to find the spectrum $\tilde{Y}(E)$ and then from the spectrum $\tilde{Y}(E)$ to find the elastic spectrum $F(E)$ if only the spectrometer transmission function $T(E)$ and the scattering function $P(E',E-E')$ are known. The procedure for determining the transmission function $T(E)$ is described in detail in Ref. 33. According to Ref. 33, the transmission function can be approximated by an exponential

$$T(E)\simeq\exp(\alpha E), \quad \alpha=-(1.69\pm0.05\pm0.17)\cdot10^{-2}\ \text{eV}^{-1}. \tag{5}$$

To determine the scattering function it is necessary to study in detail the process of electron scattering in matter in terms of some microscopic model. However, reliable quantitative calculations of this kind are very complicated and probably cannot be carried out at the present time. Instead, it was shown in Ref. 34 that it is possible to introduce an effective phenomenological procedure for determining the scattering function. The following nontrivial circumstance

suggests a suitable scattering function. It turns out that the low-energy side of the informative spectral interval (E_1, E_2) (Fig. 1) always adjoins a segment of inelastic background which grows nearly exponentially in the direction of lower electron kinetic energies. This allows us to choose the scattering function to be of the form

$$P(E-E') = \alpha_0 \exp[-\mu(E-E')]. \quad (6)$$

The constants a_0 and μ are easily determined from the amplitude of the spectrum at the point E_1 and the slope of the inelastic tail. Therefore, the problem of solving Eq. (4) becomes completely determined (the equation is easily solved by the method of simple iterations).

The question of systematic errors which can arise in the spectral line intensities in connection with the use of an approximate shape of the scattering function (6) was examined in two different ways in Refs. 34 and 31. In Ref. 34 an upper-bound estimate of the systematic errors was performed. The parameters of the scattering function were varied to the maximum extent possible, and it was observed how the conversion line intensities vary as a result. It was found that the mean variations of the intensities over all the lines used in the study and over all types of investigated extreme variations of the shape of the scattering function amounted to 0.7% of the total area of the conversion spectrum, and the maximum variations, to 2%. In Ref. 31 the procedure for subtracting out the background, based on the solution of Eq. (4) with scattering function (6), was applied to model spectra obtained from the initial spectra by modeling the passage of electrons through a hydrocarbon film using the Monte Carlo method. Since the exact form of the model elastic signal $F(E)$ in this case is exactly known, by comparing the result of background subtraction with the model elastic spectrum, one can estimate the systematic errors introduced into the line intensities by the implemented background subtraction algorithm. In this study the mean perturbation of the line intensities amounted to 0.3% of the area of the entire spectrum, and the maximum perturbation, to 0.7%. Generalizing the results of both methods, we may conclude that the algorithm of subtracting out the inelastic background can introduce an error in the spectral line intensities on the order of 1% of the total area of the spectrum, which amounts to roughly 5% of the intensities of the most important lines of the conversion spectrum of the uranium-235 isomer.

In Ref. 33 an analysis was performed of the total systematic error in the partial conversion probabilities obtained in the above-described procedure of mathematical spectral processing. Systematic errors associated with errors in the determination of the spectrometer transmission function (5), the approximate nature of the scattering function (6), and the approximation $k(E) = \text{const}$ were taken into account. It was found that the maximum relative error of the absolute values of the partial probabilities of conversion amounts to 12% for the most important spectral lines[2)] (''$6p_{1/2}$,'' ''O2s,'' and ''$6p_{3/2}$,'' and the valence band VB). The error in the ratios of the partial probabilities of conversion can reach 25%, but this is relative to the case where we are comparing line intensities at opposite ends of the informative spectral interval, the distance between which amounts to roughly 25 eV. As the distance between the lines decreases, the error decreases proportionally. And finally, the error in the ratios of the intensities of lines belonging to different conversion spectra (i.e., corresponding to different chemical states of the uranium isomer) but corresponding to the same kinetic energy is less than 13%.

4. EXPERIMENT

There are a number of factors that make the preparation of samples with the uranium-235 isomer in a prescribed and controlled chemical state and in quantities sufficient for measurement of high-resolution conversion spectra a very difficult task. The main difficulties are connected with the very short mean free path of the soft conversion electrons, which hinders the use of samples prepared by chemical means. Another substantial complication is the fact that it is impossible to monitor the chemical state of the uranium isomer in the thin surface layer in any way independent of conversion spectroscopy, since we are talking here about fractions of a picogram of uranium in such a sample. In this connection, the main idea of the approach used below is to use those chemical states of the uranium isomer in which it naturally appears in the preparation of samples under the simplest and most reproducible conditions. It is only necessary to find those experimental conditions in which it is possible to obtain different conversion spectra and then in as detailed a way as possible to interpret the corresponding chemical states.

We have found that such states can be realized by using indium substrates to collect the ^{235m}U isomer. When collecting uranium isomer on an indium substrate by varying the discharge current in the accumulation chamber within the limits of 1 to 4 nA it is possible to alter the chemical state of uranium isomer within wide limits, as is indicated by the substantial change in the shape of the conversion spectrum. Some typical uranium isomer spectra, obtained in the collection of isomer on an indium surface for various discharge currents in the sample preparation chamber, are shown in Fig. 2 (here and in the remainder of this work the spectra are shown after correcting for instrument distortions and subtracting out the inelastic background). For spectrum a in Fig. 2 the discharge current was the smallest (1 nA), for spectrum d it was the largest (4 nA), and for the other samples it was intermediate. The form of the spectra in Fig. 2 suggests that the intermediate spectra (b and c) can be obtained simply as weighted sums of only two spectra $f_A(E)$ and $f_B(E)$, whose forms are similar, respectively, to spectra a and d in Fig. 2. We will refer to this conjecture as the hypothesis of two-component mixing. Our problem is to carry out a quantitative check of this hypothesis, to separate out the pure spectra f_A and f_B, and finally, to associate them with some chemical states A and B of the atoms of the uranium isomer. It is convenient to start with the last question.

Figure 3 shows two conversion spectra of uranium isomer near the ultimate possibilities realized in experiments with isomer collection onto an indium surface. It follows from the discussion of the structure of the samples in Sec. 2 that the chemical states A and B are uranium oxides. Indeed,

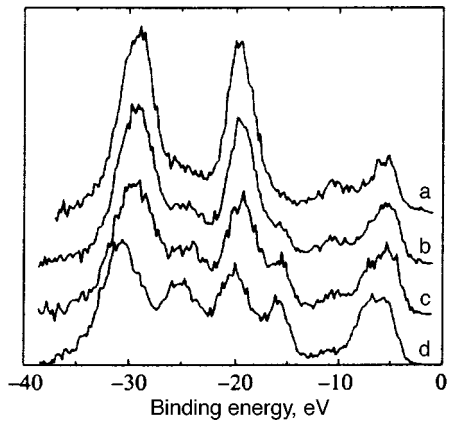


FIG. 2. Uranium-isomer conversion spectra recorded for collecting the isomer atoms on an indium surface at different discharge currents in an isomer collection chamber. Spectra a, b, c, and d correspond to currents of approximately 1, 2.5, 3.5, and 4 nA.

this is confirmed by the fact that the ''O2*s*'' line, which can be interpreted only as corresponding to the molecular orbital formed by hybridization of the $6p$ electrons of uranium and the $2s$ electrons of oxygen, is present in both spectra in Fig. 3 (Ref. 15).

In regard to the oxide corresponding to the spectrum f_B (similar to spectrum b in Fig. 3), we may draw some quite confident conclusions. The doublet structure in the region of the ''$6p_{3/2}$'' peak unambiguously indicates that we are dealing here with an atomic cluster based on the linear uranyl group O–U–O.[17,35] Moreover, from the separation between the components of the doublet (around 4.3 eV) it is possible to determine the distance between the oxygen and uranium atoms in uranyl to be 1.73 ± 0.04 Å (Ref. 36). It is well known that in the formation of complexes in which uranyl participates, the uranyl group preserves a relative independence, and in the equatorial plane it may contain 4, 5, or 6 additional oxygen or halogen atoms.[37] These are all reasons for assuming that the f_B spectrum corresponds to the uranyl group coordinated by 4–6 oxygen atoms in the equatorial plane. Such an environment is similar to the environment of uranium in γ-UO_3, where the uranyl group is coordinated with four oxygen atoms. From the distance between the components of the ''$6p_{3/2}$'' doublet it is also possible to estimate the distance from the uranium atom to the equatorial oxygen atoms as 2.40 ± 0.04 Å (Ref. 36).

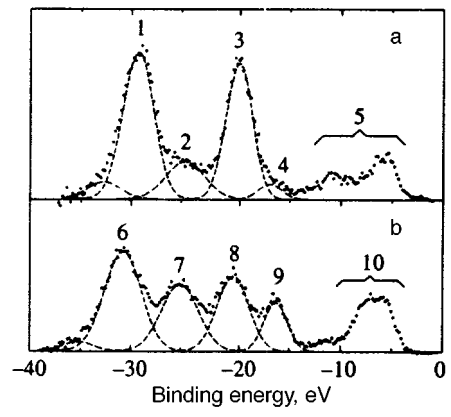


FIG. 3. Decomposition of uranium-isomer conversion spectra obtained for isomer atom collection on an indium surface. Annotation of the lines in the figure is as follows: *1, 6*—''$6p_{1/2}$,'' *2, 7*—''O2*s*,'' *3, 8*—''$6p_{3/2}(l)$,'' *4, 9*—''$6p_{3/2}(r)$,'' *5, 10*—*VB*. Spectrum *a* corresponds to the lowest experimentally observed degree of oxidation of the uranium isomer, and spectrum *b*, to the highest.

In regard to the oxide corresponding to the f_A spectrum (similar to spectrum a in Fig. 3), it is difficult to draw such definite conclusions. In spectrum a in Fig. 3 a weak trace of the right-hand component of the ''$6p_{3/2}$'' uranium line is visible, ''$6p_{3/2}(r)$.'' However, the intermediate situation, when the right-hand component of the ''$6p_{3/2}$'' peak is present but has low intensity, does not correspond to any known simple oxide of uranium. A small admixture of the ''$6p_{3/2}(r)$'' component in experimental spectrum a in Fig. 3 is more simply explained by an admixture of the chemical state B to state A. Thus, in the chemical state A splitting of the ''$6p_{3/2}$'' uranium line is entirely absent; consequently, the uranium–oxygen cluster corresponding to state A has higher symmetry than the cluster corresponding to state B. Oxygen environments of uranium are known which can lead to such a result. This happens, for example, for the simple crystalline oxides UO_2 and UO with symmetry O_h (Ref. 38). Besides the higher symmetry of cluster A in comparison with cluster B, we may note a considerably weaker bond between uranium and oxygen in cluster A since the line of the conversion spectrum corresponding to the hybrid orbital U$6p$–O$2s$ is significantly less pronounced. Just such behavior is to be expected for UO_2 and UO. All this indicates some kind of analogy between the chemical state A and uranium dioxide or monoxide.

By the relative intensity of the conversion line we mean the ratio of the area under the line to the area of the entire spectrum. Since the ''$6p_{3/2}(r)$'' line (Fig. 3) is present in the f_B spectrum and absent in the f_A spectrum (according to the proposed interpretation of the f_A spectrum), it is clear that the relative intensity of the ''$6p_{3/2}(r)$'' line in the mixed experimental spectrum characterizes the corresponding mixing coefficient of the A and B states. Moreover, if the relative intensity of the ''$6p_{3/2}(r)$'' line in the pure f_B spectrum is known (we denote it by P_B^0), then by an appropriate choice of the constructed procedure for extrapolating from the data set of experimental mixed spectra it should be possible to separate out the pure f_A and f_B spectra.

The magnitude of P_B^0 can be found from the uranium isomer conversion spectra measured during isomer collection onto a copper surface. Such spectra were examined in Ref. 33. The following two facts were established. First, during collection of isomer onto a copper surface the spectra always contain five characteristic peaks similar to the situation in Fig. 3b, i.e., similar to the spectrum f_B, where by varying the conditions of isomer collection it is not possible to change the shape of the spectra. Second, a specially constructed statistical analysis of the shape of the spectra showed that no other variations of the shape of the spectra from one measurement to the next are observed besides those associated with the Poisson electron counting statistics. These results may be interpreted as follows. Copper surfaces

offer especially favorable conditions for formation of uranyl clusters; therefore the conversion spectrum observed for collection onto a copper surface is the spectrum of the pure B state. We will call this spectrum $f_{B(\mathrm{Cu})}$. We identify the relative intensity of the ''$6p_{3/2}(r)$'' line in the f_B spectrum with the corresponding quantity in the experimental $f_{B(\mathrm{Cu})}$ spectra. In Ref. 33 the result $P_B^0=(10.3\pm0.3)\%$ was obtained.

We next determine the relative intensity P of the ''$6p_{3/2}(r)$'' line for different mixed spectra from the indium surface and introduce the quantity

$$k=P/P_B^0. \tag{7}$$

Thus, in the examination of mixed spectra we proceed from the assumption that the value $k=0$ corresponds to the pure f_A spectrum and the value $k=1$ corresponds to the pure f_B spectrum.

We normalize the experimental conversion spectra to the corresponding total probabilities of conversion per unit time, which in the case of the uranium-235 isomer coincide with the decay constant of the isomer level. Let (E_i,E_f) be an arbitrary segment of the conversion spectrum. We consider a conversion spectrum f_k with the parameter k: $0\leqslant k\leqslant 1$. We denote by P_k, P_A, and P_B the integrals over the interval (E_i,E_f) respectively for the spectra f_k, f_A, and f_B. It is not hard to show that

$$P_k\simeq(1-k)P_A+kP_B, \tag{8}$$

where the deviation from exact equality in formula (8) may be neglected thanks to the small difference in decay constants of the uranium isomer in states A and B. We call the dependence of the area of a fixed segment of the spectrum on the parameter k the kinetic curve. The prediction of linearity of the kinetic curves (8), following from the hypothesis of two-component mixing, can be checked experimentally.

To use the formalism introduced above, the conversion spectra of the uranium isomer should be measured simultaneously with the decay constants. The technique of simultaneous measurement of the conversion spectrum and the decay constant was introduced by myself in Ref. 39. The essence of this technique is to measure the conversion spectrum in two stages, in two successive time intervals while the decay constant is determined by comparing the areas of the corresponding spectra calculated from the informative spectral interval (E_1,E_2) (Fig. 1). This guarantees that the decay constant and the conversion spectrum correspond exactly to the same chemical state of the uranium isomer.

Two series of measurements with collection of uranium isomer onto an indium surface constitute the experimental data base of the present work. In the first series, which we will call series T, thirty measurements of conversion spectra with simultaneous measurement of the decay constants were carried out. For technical reasons in this series of measurements there were time losses on the order of one half-life of the isomer; therefore, the series-T spectra are not very well statistically robust. In order to compensate for this deficiency of series T, a second series of measurements was carried out,

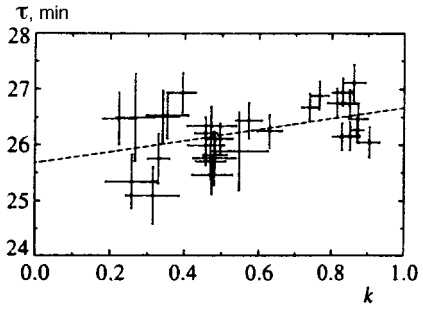


FIG. 4. Experimental points $\tau(k)$ for experimental series T and linear approximation by the maximum-likelihood method.

which we will call series R, in which 29 spectra without simultaneous measurement of the decay constants with concomitant time losses were measured.

On the basis of the series-T experimental data we constructed an empirical dependence $\lambda(k)$, where λ is the decay constant and k is the parameter defined by formula (7). An empirical dependence was found in the form of a linear function from the experimental points $\lambda_{\mathrm{exp}}(k_{\mathrm{exp}})$ with the help of the maximum likelihood method. Figure 4 plots the corresponding experimental points and a straight-line fit (for convenience the corresponding half-lives τ rather than the decay constants λ are plotted along the ordinate). The dependence so obtained was then used to normalize the spectra of series R as well as series T.

We extended the obtained empirical dependence $\lambda(k)$ to $k=0$ and $k=1$, which gave the half-lives of the uranium isomer corresponding to the pure states A and B:

$$\tau_A=25.67\pm0.21\ \mathrm{min},\quad \tau_B=26.68\pm0.12\ \mathrm{min},$$

$$(\tau_B-\tau_A)/\tau_A=(3.9\pm1.2)\%.$$

It makes sense to compare these values with the half-lives of the uranium isomer measured by de Mevergnies with massive samples of the chemically prepared compounds $^{235m}UO_2$ and γ-$^{235m}UO_3$ (Ref. 23):

$$\tau(UO_2)=24.68\pm0.26\ \mathrm{min},\quad \tau(UO_3)=26.06\pm0.08\ \mathrm{min},$$

$$[\tau(UO_3)-\tau(UO_2)]/\tau(UO_2)=(5.6\pm1.1)\%.$$

It is evident that the same trend of the half-life obtains both between UO_2 and γ-UO_3 and between the states A and B, and that the relative values of the half-lives in fact do not differ within the limits of experimental error. This confirms the reasonableness of interpreting state A as an analog of UO_2 and state B as an analog of γ-UO_3.

To check the hypothesis of two-component mixing, experimental kinetic curves were constructed for four segments of the conversion spectrum corresponding to the main conversion lines: ''$6p_{1,2}$,'' ''O2s,'' ''$6p_{3/2}(l)$,'' and VB. Figure 5 plots the kinetic curves for the series-R measurements (for series T the curves have a similar form, but with larger statistical errors). Using the maximum likelihood method, we fit the kinetic data with straight lines. In the course of the fit we calculated the quantity

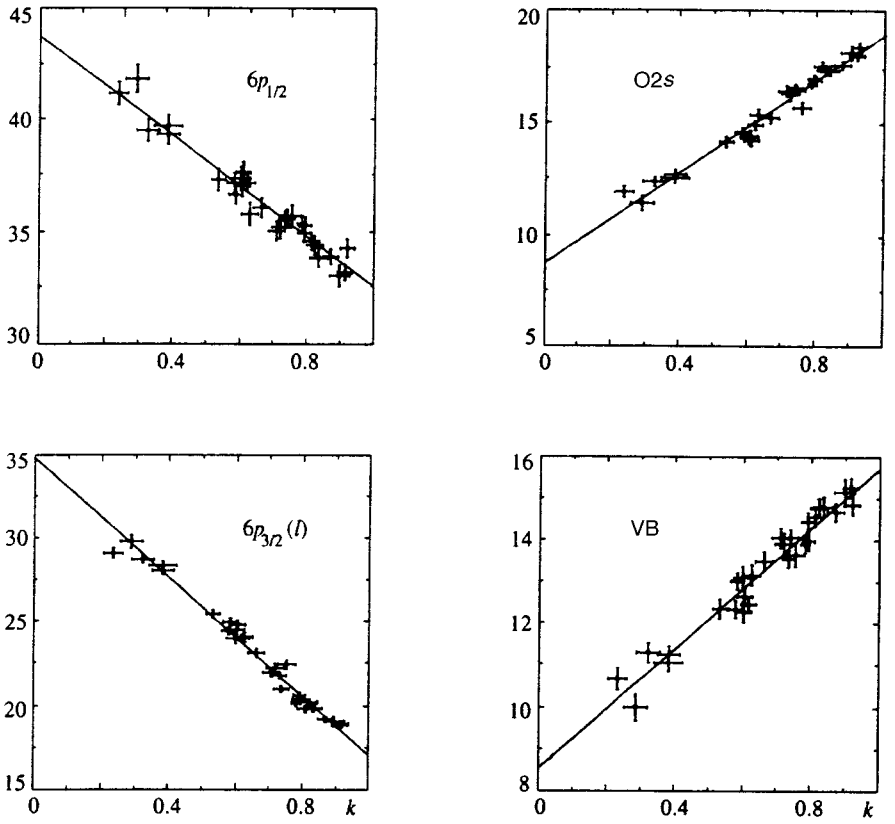


FIG. 5. Kinetic curves for experimental series *R*. The quantities plotted along the ordinates are the partial probabilities of conversion belonging to the indicated spectral regions (%).

$$\chi_c^2=\sum_i \frac{(a+bk_i-\lambda_i)^2}{\sigma\lambda_i^2+b^2\sigma k_i^2}, \tag{9}$$

which is an analog of the χ^2 parameter of the more familiar linear regression analysis. In formula (9) a and b are the fitting parameters of the linear dependence. Values of χ_c^2/N, where N is the number of degrees of freedom, are given in Table I for both series of measurements—T and R. The standard deviation for all values is equal to 0.26. From Table I it can be seen that there is not one significant deviation of χ_c^2/N from unity, i.e., there are no indications of deviations of the kinetic curves from a linear dependence. Thus, the hypothesis of two-component mixing is found to be well confirmed.

Formula (8) makes it possible to find the spectra f_A and f_B corresponding to the set of experimental mixed spectra in the maximum-likelihood sense. Toward this end, it is sufficient to construct the kinetic curves for each channel of the spectrum, to fit them with straight lines using the maximum likelihood method (8), and to extend each such straight line to $k=0$ and $k=1$. In order to track the reproducibility of the results, we processed the series-R and series-T spectra in this way separately. Figure 6 shows the pure spectra f_A and f_B, found from the mixed spectra of series R. The spectra found from series T have essentially the same form, but somewhat higher statistical spreading of the amplitudes. Fragments of the decompositions of the spectra into Gaussian lines corresponding to the region of inner valence molecular orbitals are shown in Fig. 6 by dashed lines.

Table II presents data on the structure of these spectra. The standard deviations of the line parameters, calculated by the Monte Carlo method, are indicated; systematic errors are not shown, but can be determined as in Sec. 3. For convenience, the partial probabilities of conversion are given in units in which $\lambda_B\equiv 100$. To convert to absolute probability values, it is necessary to multiply the corresponding data in

TABLE I. Check of the hypothesis of two-component mixing of the spectra.

	Spectral region			
Series	U6$p_{1/2}$	O2s	U6$p_{3/2}(l)$	*VB*
T	0.89	0.78	1.23	1.23
R	0.87	1.05	1.24	1.19

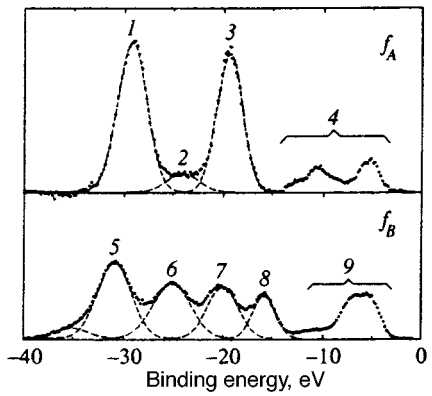


FIG. 6. Pure f_A and f_B spectra found from the series of experimental mixed spectra *R*. Annotation of the lines: *1*—"6$p_{1/2}$," *2*—"O2s," *3*—"6$p_{3/2}$," *4*—*VB*, *5*—"6$p_{1/2}$" (1*S*1*u*), *6*—"O2s" (2*S*1*u*), *7*—"6$p_{3/2}(l)$" (1*S*3*u*), *8*—"6$p_{3/2}(r)$" (3*S*1*u*), *9*—*VB*.

TABLE II. Structure of pure f_A and f_B spectra. Each cell of the table contains (from top down): the partial probability of conversion in relative units, the line position (binding energy) in eV, and the line width in eV. The statistical errors are indicated in parentheses. Where the error is not indicated explicitly, it is approximately one significant digit. Rows $\langle A\rangle$ and $\langle B\rangle$ correspond to data averaged over series T and R.

	Line				
Spectrum	$6p_{1/2}$	O2s	$6p_{3/2}(l)$	$6p_{3/2}(r)$	VB
	44.6(0.4)	7.1(0.4)	37.7(0.4)	0	15.1(0.2)
$A(T)$	28.5	23.7	18.6	-	-
	3.3	4.4(0.4)	3.1	-	-
	46.4(0.4)	6.2(0.5)	37.2(0.4)	0	15.2(0.2)
$A(R)$	29.3	24.0	19.4	-	-
	3.5	3.8(0.4)	3.0	-	-
	31.5(0.2)	22.1(0.3)	16.4(0.3)	10.5(0.1)	19.2(0.2)
$B(T)$	30.6	24.8	19.9	15.3	-
	3.6	4.3	3.5	2.7	-
	30.7(0.3)	22.4(0.4)	16.9(0.3)	10.4(0.1)	19.0(0.1)
$B(R)$	30.8	25.0	19.9	15.8	-
	4.0	4.5	3.6	2.7	-
	45.5(0.3)	6.7(0.4)	37.5(0.3)	0	15.2(0.1)
$\langle A\rangle$	28.9(0.4)	23.9(0.2)	19.0(0.4)	-	-
	3.4(0.1)	4.1(0.3)	3.1(0.1)	-	-
	31.3(0.2)	22.2(0.2)	16.7(0.2)	10.5(0.1)	19.0(0.1)
$\langle B\rangle$	30.7(0.1)	24.9(0.1)	19.9(0.1)	15.6(0.3)	-
	3.8(0.2)	4.4(0.1)	3.5(0.1)	2.7(0.1)	-
	31.4(0.5)	23.6(0.9)	16.0(0.6)	10.3(0.3)	18.8(0.2)
Cu(B)	31.0	25.1	19.9	15.6	-
	4.0	5.0(0.4)	3.4	3.0	-

Table II by 2.598×10^{-4} min. As a reference point for determining the binding energies, we used the position of the ''$6p_{3/2}(l)$'' line in the f_B spectrum, for which we adopted the value 19.89 eV. This energy is the averaged binding energy of the ''U$6p_{3/2}(l)$'' line found from x-ray electron spectroscopy data, determined from a set of 15 different uranyl compounds.[40] For convenience, the last line of Table II gives the breakdown of the $f_{B(\mathrm{Cu})}$ spectrum based on the data of Ref. 33. Notice that there is beautiful agreement between the parameters of the $f_{B(\mathrm{Cu})}$ and f_B spectra. This confirms that the B state on the indium surface and the state in which the uranium isomer always appears for collection onto a copper surface are indeed one and the same chemical state of the isomer as was assumed in the treatment of the series-T and R mixed spectra.

5. DISCUSSION: COMPARISON OF EXPERIMENT WITH CALCULATIONS OF THE ELECTRONIC STRUCTURE OF URANYL

It is well known that because the uranyl group possesses relative independence (stability) in the formation of the chemical bond, the main features of the x-ray electron spectra of uranyl compounds in the region of electron binding energies $\lesssim$40 eV are determined specifically by the presence of this group.[17,35] This should hold to an even larger extent in the case of the conversion spectrum of uranium isomer. Therefore, to identify the lines of the f_B spectrum and to compare the results of experiment with the first-order theory we can apply the results of a calculation of the electronic structure of the isolated uranyl group. Here it is necessary to use relativistic calculations, since in the nonrelativistic calculations the $6p_{1/2}$ and $6p_{3/2}$ electrons do not differ while in the uranium atom they are separated in binding energy by the spin–orbit interaction by 10 eV and, what is most important, their single-electron conversion probabilities differ by more than a factor of two.[13]

A number of relativistic calculations of the electronic structure of uranyl are known.[41–45] All of them qualitatively reproduce the structure of the conversion spectrum f_B reasonably well in the region of the inner valence molecular orbitals (the lines ''$6p_{1/2}$'' through ''$6p_{3/2}(r)$,'' Fig. 6) and agree with each other both in the order of sequence of the lines and in the makeup of the molecular orbitals in the sense of the method of linear combination of atomic orbitals (LCAO). Figure 6 identifies the lines of the f_B spectrum with the molecular orbitals of uranyl. The notation used corresponds to that in Refs. 44 and 45. Table III gives the structure of the inner valence molecular orbitals of uranyl in the terminology of the LCAO method according to Refs. 44 and 45.

It is clear directly from the shape of the f_A spectrum (Fig. 6) that the U6p electrons are much more weakly delocalized in the chemical state A of the uranium isomer than in state B. The intensity of the ''O2s'' line is so weak that it may be assumed that the ''$6p_{1/2}$'' orbital of compound A has essentially an atomic character and is therefore occupied by two U$6p_{1/2}$ electrons. Assuming a quasi-atomic nature of the ''$6p_{1/2}$'' orbital of compound A and using the results of the calculation of the electronic structure of uranyl given in Refs. 44 and 45, we can estimate how much the intensity of the ''$6p_{1/2}$'' conversion line in the f_B spectrum should vary in comparison with the f_A spectrum.

Reference 16 proposed an interpretation of the conversion spectra of soft nuclear transitions in terms of a ''filling of the standard states in the conversion zone.'' In order to utilize the data calculated by the LCAO method[44,45] for a

TABLE III. Electronic structure of uranyl according to Refs. 44 and 45 and estimates of net occupancies of the molecular orbitals by the atomic electrons of uranium. The eigenenergy is normalized to the energy value -19.9 eV for the $1S3u$ orbital.

Spectral line f_B	Orbital	Energy, eV	Makeup of orbital (full Mulliken occupancies)	Overlap occupancy	Net occupancy
''$6p_{1/2}$''	$1S1u$	-33.3	$1.58(6p_{1/2}) + 0.06(6p_{3/2}) + 0.36(\mathrm{O}2s)$	0.292	$1.44(6p_{1/2}) + 0.05(6p_{3/2})$
''O2s''	$2S1u$	-26.0	$0.33(6p_{1/2}) + 1.09(6p_{3/2}) + 0.48(\mathrm{O}2s)$	0.240	$0.30(6p_{1/2}) + 1.00(6p_{3/2})$
none	$2S1g$	-23.7	$0.40(6d) + 1.58(\mathrm{O}2s)$	0.460	$0.17(6d)$
''$6p_{3/2}(l)$''	$1S3u$	-19.9	$1.91(6p_{3/2}) + 0.09(\mathrm{O}2p)$	0.096	$1.86(6p_{3/2})$
''$6p_{3/2}(r)$''	$3S1u$	-13.5	$0.06(6p_{1/2}) + 0.34(6p_{3/2}) + 0.72(\mathrm{O}2p) + 0.83(\mathrm{O}2s)$	0.008	$0.06(6p_{1/2}) + 0.34(6p_{3/2})$

comparison with the line intensities of the conversion spectrum, it is necessary to relate the formalism of Ref. 16 with concepts characteristic of the LCAO method.

Let some molecular orbital ϕ of uranyl in the LCAO representation have the form

$$\phi=\sum_l c_l^{\mathrm{U}}\chi_l^{\mathrm{U}}+\sum_m c_m^{\mathrm{O}}\chi_m^{\mathrm{O}}, \tag{10}$$

where χ_l^{U} are the atomic wave functions of uranium, χ_m^{O} are the atomic wave functions of oxygen, c_l^{U} and c_m^{O} are the coefficients of the linear combination. Since the probability of conversion on a neighboring atom is negligibly small,[16] to describe conversion it is sufficient to retain the terms in expression (10) belonging to the atomic orbitals of uranium. We thus obtain an expansion of exactly the same form as was investigated in Ref. 16. For this case it has been shown that the single-electron probability of conversion from the molecular orbital ϕ is given by

$$P_\phi^1=\sum_l |c_l^{\mathrm{U}}|^2 P_l^{\mathrm{U}}, \tag{11}$$

where P_l^{U} are the single-electron probabilities of conversion for the atomic orbitals of uranium. Using formula (11), we can write the total probability of conversion from the molecular orbital ϕ as

$$P_\phi=\sum_l N_l P_l^{\mathrm{U}}, \tag{12}$$

where the quantities N_l are the net occupancies of the molecular orbital ϕ by the χ_l^{U} electrons in the classification of occupancies given by Mulliken.[46]

Applying formula (12) and the assumption of the quasiatomic nature of the ''$6p_{1/2}$'' line in the f_A spectrum, we find for the ratio of intensities of the ''$6p_{1/2}$'' lines in the f_B and f_A spectra

$$\frac{P_B(\text{''}6p_{1/2}\text{''})}{P_A(\text{''}6p_{1/2}\text{''})}=\frac{N_B(6p_{1/2})P(6p_{1/2})+N_B(6p_{3/2})P(6p_{3/2})}{2P(6p_{1/2})}, \tag{13}$$

where $P(6p_{1/2})$ and $P(6p_{3/2})$ are the single-electron probabilities of conversion on the $6p_{1/2}$ and $6p_{3/2}$ electrons in the isolated uranium atom, and $N_B(6p_{1/2})$ and $N_B(6p_{3/2})$ are the atomic net occupancies of the $1S1u$ molecular orbitals of uranyl.

We estimated the net occupancies from the data of Refs. 44 and 45 in terms of the total occupancies and overlap occupancies (Table III). Instead of the single-electron probabilities of conversion we used the single-electron conversion factors $w_e(6p_{1/2})$ and $w_e(6p_{3/2})$—dimensionless quantities proportional to the single-electron probabilities of conversion, which were calculated for the conversion transition of the uranium isomer in Ref. 13:

$$w_e(6p_{1/2})=4.81\cdot 10^5, \quad w_e(6p_{3/2})=2.16\cdot 10^5. \tag{14}$$

Calculations were performed in Ref. 13 with atomic wave functions in the Thomas–Fermi–Dirac and Hartree–Fock–Slater models. The values in Eqs. (14) are averages over these two models.

Using the data in Table III for the conversion factors (14) and the net occupancies used in formula (13), we find

$$\left.\frac{P_B(\text{''}6p_{1/2}\text{''})}{P_A(\text{''}6p_{1/2}\text{''})}\right|_{\mathrm{theor}}=0.73, \tag{15}$$

whereas from the experimental data listed in Table II we obtain

$$\left.\frac{P_B(\text{''}6p_{1/2}\text{''})}{P_A(\text{''}6p_{1/2}\text{''})}\right|_{\mathrm{exper}}=0.69\pm 0.01\pm 0.09. \tag{16}$$

The experimental (16) and theoretical (15) intensity ratios are found to be in beautiful agreement.

TABLE IV. Expected (''theoretical'') and experimental relative intensities of the conversion lines belonging to the inner valence molecular orbitals. The total errors (including systematic) are indicated in parentheses.

	''$6p_{1/2}$'' ($1S1u$)	''O2s'' ($2S1u$)	''$6p_{3/2}(l)$'' ($1S3u$)	''$6p_{3/2}(r)$'' ($3S1u$)
Calculation	44.9	23.0	25.7	6.5
Experiment	38.8(4.7)	27.5(3.3)	20.7(2.5)	13.0(1.6)

Analogously, it is possible to find the expected ratios of the line intensities of the f_B spectrum and compare them with the corresponding experimental values. The result of such a comparison is given in Table IV, where the sum of the line intensities belonging to the inner valence molecular orbitals is normalized to 100, both for the experimental and the theoretical values. In all cases, except for the ''$6p_{3/2}(r)$'' line, there is reasonable agreement between the experimental and calculated intensities. For the ''$6p_{3/2}(r)$'' line the calculated intensity is half the experimental value. However, the authors of Ref. 44 noted that taking the crystal field of the equatorial ligand-atoms of uranyl into account may increase the total occupancy of the $3S1u$ orbital by the $6p$ electrons from 0.4 (see Table III) to 0.8. Obviously, this would roughly double the estimate of the intensity of the ''$6p_{3/2}(r)$'' line in the f_B spectrum, as a result of which the agreement would improve. I will not attempt to compare theory and experiment in the valence-band region (VB, Fig. 6) since the accuracy of present-day calculations of the electronic structure in this region is not high.

Thus, reasonable agreement is observed between the expected and experimental intensities of the conversion lines in the conversion spectrum of uranium isomer in the region where calculations of the electronic structure appear to possess satisfactory accuracy. By virtue of the semi-quantitative character of the calculations, it is still difficult to speak of real agreement of theory with experiment; rather it may be stated that abrupt anomalies in the intensities of the conversion lines which might be linked with interference phenomena accompanying elastic scattering of conversion electrons are not revealed. We may conclude that the results obtained here do not contradict the assumption that the partial probabilities of conversion of uranium isomer remain proportional to the electron density at the nucleus as follows from the works of Grechukhin and Soldatov[13,16] and as is indirectly indicated by data on variations of the uranium isomer decay constant.[19–25] In the LCAO approximation this means that the partial probabilities of conversion are proportional to the atomic net occupancies of the molecular orbitals by the electrons that take part in conversion [formula (12)].

If we accept this conclusion, then the question arises why the strong interference effects in the conversion probabilities predicted in Ref. 26 are not manifested in experiment. It is obviously necessary to search for mechanisms that can lead to the suppression of interference of conversion electrons. I considered this question in Ref. 29. In this work, I showed that it is possible to point to at least two mechanisms which were not taken into account in Ref. 26 and which can in general lead to suppression of interference effects in conversion spectra: thermal vibrations of atoms and inelastic scattering of conversion electrons. Qualitative estimates in Ref. 29 showed that one can expect a roughly fivefold suppression of the contribution of interference effects to the partial probabilities of conversion of uranium-235 isomer, which implies a radical suppression of the effects predicted in Ref. 26. Consequently, the probabilities of conversion can remain proportional to the electron density at the nucleus with an accuracy of 5–10%, and the original version of the theory of Grechukhin and Soldatov holds with the same accuracy.[13,16]

In this light, a decrease in the intensity of the ''$6p_{1/2}$'' line in the f_B spectrum in comparison with the f_A spectrum by roughly 30% may be considered as an indication of migration of U$6p_{1/2}$ electrons with binding energy around 30 eV to the hybrid molecular orbitals. With the help of formula (12) we can obtain an estimate of the number of electrons that have migrated over to the hybrid molecular orbitals (in the terminology of the LCAO method). We obtain 0.6 electron for the electrons that have migrated over and 1.4 electrons for those remaining in the initial ''quasi-atomic'' orbital.

Note that except for conversion spectroscopy of uranium-235 isomer, at present there exists no experimental method that allows one to obtain such detailed information about the partial occupancies of the inner molecular orbitals of heavy-element compounds. Potential competitors of conversion electron spectroscopy in this region are x-ray or ultraviolet electron spectroscopy and x-ray emission spectroscopy. However, photoelectron spectroscopy does not possess selectivity to the contributions of different types of electrons to the molecular orbitals; therefore, generally speaking, the line intensities of the photoelectron spectrum do not give direct information about the structure of the corresponding molecular orbital. X-ray emission spectra, in principle, contain the same information, but in the case of heavy elements, in particular uranium, they have such a complicated structure[47] that it is still difficult to pose the question of quantitative analysis of such spectra.

It should be noted that conversion spectroscopy of uranium isomer still faces a number of complicated unresolved problems. The most urgent among them are the construction of a quantitative theory of conversion which would permit an accurate estimate of the contribution to conversion both of interference effects and effects suppressing interference, and on the experimental plane, the development of a technique for preparing samples with uranium isomer in a prescribed atomic environment. On the other hand, it has been possible to solve a number of fundamental questions in the way of transforming conversion spectroscopy of uranium isomer into a fully capable, quantitative method for investigating the electronic structure of matter, and on this path the first substantial results have been obtained.

ACKNOWLEDGMENTS

The author expresses his gratitude to V. I. Zhudov and B. V. Odinov for making it possible to use to experimental setup which they built (the apparatus for collecting uranium-235 isomer atoms and the pre-acceleration system for the

NR5950A spectrometer). The author also expresses his gratitude to V. N. Gerasimov, V. Yu. Dobretsov, and V. I. Zhudov for helpful discussions, and D. V. Grebennikov for help with this work. The author is grateful to Dmitri Petrovich Grechukhin and to the adviser of this work, Vladimir Mikhaĭlovich Kulakov, for their constant support and for advice and deeply regrets that they were not able to see the fruition of this work.

This work was carried out with the partial financial support of the International Science Foundation (Grant No. N9E000), the International Science Foundation and the Russian Government (Grant No. N9E300), and the Russian Fund for Fundamental Research (Project No. 93-03-18537). The author expresses his gratitude to these organizations.

[1)]By line widths we everywhere understand here the full-widths at half maximum.

[2)]We put spectral line symbols of the type ''$6p_{1/2}$'' in quotation marks because the spectral lines correspond not to pure atomic electrons, but to hybrid molecular orbitals.

[1] V. N. Gerasimov and V. M. Kulakov, Izv. Akad. Nauk SSSR, Ser. Fiz. **54**, 1705 (1990).
[2] J. W. Petersen, G. Weyer, E. Holzschuh, and W. Kundig, ISOLDE Collaboration, Phys. Lett. A **146**, 226 (1990).
[3] V. N. Gerasimov, D. V. Grebennikov, V. M. Kulakov, S. K. Lisin, and V. V. Kharitonov, Preprint IAÉ-5815/2 (1994).
[4] F. Azaro and I. Perlman, Phys. Rev. **107**, 318 (1957).
[5] J. R. Huizenga, C. L. Rao, and P. W. Engelkemeir, Phys. Rev. **107**, 319 (1957).
[6] A. D. Panov, in *International Conference on Nuclear Physics (XLVIth Conference on Nuclear Spectroscopy and the Structure of the Atomic Nucleus)*, St. Petersburg, Russia (1996), p. 340.
[7] M. C. Michel, F. Azaro, and I. Perlman, Bull. Amer. Phys. Soc. Ser. II **2**, 394 (1957).
[8] R. G. Helmer and C. W. Reich, Physica C **49**, 1845 (1994).
[9] A. Bohr and B. R. Mottelson, *Nuclear Structure*, Vol. 2 (Benjamin, New York, 1975).
[10] J. Almeida, T. von Egidy, P. H. M. van Assche, H. G. Börner, W. E. Davidson, K. Schreckenbach, and A. I. Nameson, Nucl. Phys. A **315**, 71 (1979).
[11] I.-L. Lamm, Nucl. Phys. A **125**, 504 (1969).
[12] F. A. Gareev, S. P. Ivanova, L. A. Malov, and V. G. Soloviev, Nucl. Phys. A **171**, 134 (1971).
[13] D. P. Grechukhin and A. A. Soldatov, Yad. Fiz. **23**, 273 (1976) [Sov. J. Nucl. Phys. **23**, 143 (1976)].
[14] V. I. Zhudov, A. G. Zelenkov, V. M. Kulakov, V. I. Mostovoĭ, and B. V. Odinov, JETP Lett. **30**, 516 (1979).
[15] D. P. Grechukhin, V. I. Zhudov, A. G. Zelenkov, V. M. Kulakov, B. V. Odinov, A. A. Soldatov, and Yu. A. Teterin, JETP Lett. **31**, 592 (1980).
[16] D. P. Grechukhin and A. A. Soldatov, VANT, Ser. Jad. Konst., No. 2(56), 36 (1984).
[17] B. W. Veal, D. J. Lam, W. T. Carnall, and H. R. Hoekstra, Phys. Rev. B **12**, 5651 (1975).
[18] D. P. Grechukhin and A. A. Soldatov, Yad. Fiz. **38**, 1397 (1983) [Sov. J. Nucl. Phys. **38**, 849 (1983)].
[19] H. Mazaki and S. Shimizy, Phys. Rev. **148**, 1161 (1966).
[20] M. Neve de Mevergnies, Phys. Lett. B **26**, 615 (1968).
[21] M. Neve de Mevergnies, Phys. Rev. Lett. **23**, 422 (1969).
[22] M. Neve de Mevergnies, Phys. Rev. Lett. **29**, 1188 (1972).
[23] M. Neve de Mevergnies and P. Del Marmol, Phys. Lett. B **49**, 428 (1974).
[24] V. I. Zhudov, V. K. Kulakov, and B. V. Odinov, in *Abstracts of the XXIXth Conference on Nuclear Spectroscopy and the Structure of the Atomic Nucleus* (Nauka, Leningrad, 1979), p. 563.
[25] V. I. Zhudov, V. M. Kulakov, B. V. Odinov, and A. D. Panov, VANT, Ser. Vad. Konst., No. 4(58), 3 (1984).
[26] V. Yu. Dobretsov, Zh. Éksp. Teor. Fiz. **102**, 774 (1992) [Sov. Phys. JETP **75**, 424 (1992)].
[27] V. I. Zhudov and D. V. Grebennikov, VANT, Ser. Yad.-Fiz. Issled., No. 11(19), 100 (1990).
[28] A. D. Panov, Poverkhnost' No. 10, 21 (1995).
[29] A. D. Panov, Preprint IAÉ-5901/2 (1995).
[30] M. A. Kelly and C. E. Tyler, Hewlett-Packard J. **24**, 2 (1973).
[31] A. D. Panov, Preprint IAÉ-5954/2 (1995).
[32] M. P. Seach and W. A. Dench, Surf. Interface Anal. **1**, 2 (1979).
[33] A. D. Panov, Preprint IAÉ-5960/2 (1996).
[34] A. D. Panov, Preprint IAÉ-5926/2 (1996).
[35] A. L. Gubskiĭ, A. P. Kovtun, Yu. A. Teterin, and A. S. Baev, Preprint IAÉ-4053/12 (1984).
[36] V. I. Nefedov, Yu. A. Teterin, T. Raĭkh, and Kh. Niche, Dokl. Akad. Nauk **348**, 634 (1996).
[37] L. V. Volod'ko, A. I. Komyak, and D. S. Umreĭko, *Uranyl Compounds. Spectra and Structure*, Vol. 1 [in Russian] (Belorussian State Univ. Press, Minsk, 1981).
[38] V. A. Gubanov, A. Rosen, and D. E. Ellis, J. Inorg. Nucl. Chem. **41**, 975 (1979).
[39] A. D. Panov, VANT, Ser. Obshch. Yad. Fiz., No. 4(40), 79 (1987).
[40] Yu. A. Teterin and A. S. Baev, *X-Ray Photo-Electron Spectroscopy of Compounds of the Light Actinides: Review* [in Russian] (State Committe on the Use of Atomic Energy of the USSR, Moscow, 1986).
[41] C. Y. Yang, K. H. Jonson, and J. A. Horsley, J. Chem. Phys. **68**, 1001 (1978).
[42] J. H. Wood, M. Boring, and S. B. Woodruff, J. Chem. Phys. **74**, 5225 (1981).
[43] P. F. Walch and D. E. Ellis, J. Chem. Phys. **65**, 2387 (1976).
[44] V. A. Glebov and V. S. Nefedov, Koord. Khim. **7**, 1664 (1981).
[45] V. A. Glebov and V. S. Nefedov, Koord. Khim. **7**, 1673 (1981).
[46] R. S. Mulliken, J. Chem. Phys. **23**, 1833 (1955).
[47] Yu. A. Teterin, V. A. Terekhov, and K. E. Ivanov, Dokl. Akad. Nauk **345**, 356 (1995) [Phys. Dokl. **40**, 195 (1995)].

Translated by Paul F. Schippnick

Spin dynamics in solid dilute paramagnets

F. S. Dzheparov, I. V. Kaganov, and E. K. Khenner

Institute of Theoretical and Experimental Physics, 117259 Moscow, Russia
(Submitted 25 October 1996)
Zh. Éksp. Teor. Fiz. **112**, 596–617 (August 1997)

The free-induction decay signals for the two-spin clusters, the ''matrix,'' and the entire spin system as a whole are calculated for a real dipole–dipole interaction on the basis of a method for the cluster expansion of correlation functions previously developed by the authors. The intracluster interactions, which generate a discrete spectrum, are taken into account exactly, and the interactions outside the clusters are taken into account on the basis of the Anderson-Weiss-Kubo (AWK) theory of phase relaxation. The corresponding shape functions of the resonance line are calculated. Before the clusters are separated, an AWK analysis in the limit of slow fluctuations of the local fields is in satisfactory agreement with the exact results obtained in the Anderson model of dipolar interactions, but a similar analysis for realistic fluctuation rates and $Dt>1$ leads to a significant qualitative difference: $\exp(-Dt)$ transforms into $\exp(-\sqrt{D_1 t})$. After the clusters are separated, this difference is obliterated. The free-induction decay is almost exponential up to $Dt\sim 10$, and nonexponential behavior of the long-time asymptote appears sooner. This finding is not altered significantly when the analysis is complicated by raising the maximum rank of the clusters to 3 or more. © *1997 American Institute of Physics.* [S1063-7761(97)01508-4]

1. INTRODUCTION

The theory of the magnetic spectroscopy of multispin systems utilizes a relatively small number of fundamental correlation functions and their spectral densities, which completely describe the dynamics of such systems. In the early stages of the formulation of the theory of magnetic resonance, the entire discussion was basically confined to studying the line shape and spin-lattice relaxation, which are exhaustively described by two or three such functions.

The situation has changed drastically.[1–4] In particular, several properties of multispin correlation functions, which specify, for example, multispin coherencies[2] and the rates of flip-flop transitions[4] in nuclear spin systems, are accessible to direct measurement. On the other hand, experimental[5,6] and theoretical[7,8] investigations of spectral and spatial transfer in magnetically dilute ESR systems stimulated the development of a novel cluster expansion method,[3] which utilizes multispin correlation functions of physically separated groups of spins, or clusters, and, in principle, permits the formulation of an adequate theory for such processes.

An exact calculation of cluster correlation functions for a real interaction is still impossible. Therefore, bearing in mind the problems of the spin dynamics of magnetically dilute systems, in the present work we turned our attention first to the exact solution of the Anderson model,[9] which presently serves as the main source of reliable information for magnetically dilute systems (see, for example, Refs. 10 and 11 and the references cited therein). The Anderson model does not take into account the influence of temporal fluctuations of the local fields on the processes in the system. Therefore, we next introduced and examined a very simple model, which takes this phenomenon into consideration in the spirit of the Anderson-Weiss-Kubo (AWK) theory[12,13] and is based on a cluster analysis.[3] In this way we obtained explicit integral expressions for the free-induction decay function and discovered a disparity between the predictions of the Anderson model and the new, more realistic model. We note, however, that the results of the two approaches are consistent in the limit of slow fluctuations and for the total free-induction decay signal.

Our analysis is based on the following arguments. Cluster expansion can be described in the following manner (a very brief description is given here; for further details see Ref. 3). A cluster of rank l is a group of l spins which, for purely geometric reasons, interact more strongly with one another than with any spin outside of the group. To make the division unequivocal, orthogonal clusters, in which each spin is found in no more than one such cluster, are introduced. Expansion in orthogonal clusters is accomplished in the following manner: k, i.e., the highest rank of the clusters considered in the problem, is fixed on the basis of physical arguments, and all the clusters of that rank and then of rank $k-1$, etc., are separated. The part of the system remaining after separation of all the clusters is called the ''matrix.'' Quasihomogeneity of the matrix, which is understood to be the possibility of describing its thermodynamics in the continuous-medium approximation, was proposed in Ref. 3 as the physical criterion for selecting k. Before the clusters are separated, the dipolar specific heat is infinite in this approximation, but after cluster separation, the dipolar specific heat of the matrix becomes finite and is determined by interactions at moderate distances. The second moment of the absorption line exhibits completely analogous behavior. Therefore, it can be presumed that the time for establishing equilibrium in the matrix is of the order of the characteristic time of the flip-flop transitions and that it corresponds to the phase relaxation time approximately as in regular systems. The fluctuating local fields induced by the environment on the spins of the matrix and on the clusters are considerably

closer to a normal random process than are the fields on the spins before separation of the clusters. The intracluster interactions can be taken into account exactly.

We next focus a considerable amount of attention on the transverse correlation function $G(t)$, i.e., the free-induction decay function, which is classical in the theory of magnetic resonance. Its spectral density gives the absorption line shape function, and the cluster expansion of this density expresses the absorption line shape function relatively independently of the evolving parts of the system.

We restrict ourselves to the high-temperature approximation, since the specific details of unordered systems are manifested quite fully already in this limit. A series of calculations is performed in the Anderson model for an arbitrary spin. This is of interest for possible comparison with the results of a numerical simulation, which is generally performed for classical spins because of the restricted scope of present-day computing possibilities.[14,15]

2. CONFIGURATIONAL AVERAGING OF CORRELATION FUNCTIONS

We assume everywhere below that the spin system studied consists of impurity spins that are randomly distributed throughout the lattice.

In the occupation number representation[16] and in the high-temperature approximation the transverse correlation function has the form[1]

$$G(t)=\langle S^+(t)\rangle/\langle S^+(0)\rangle=\langle\langle S^+(t)S^-\rangle_0\rangle_c/\langle\langle S^+S^-\rangle_0\rangle_c. \tag{1}$$

Here $\langle\ldots\rangle_0=\mathrm{Tr}(\ldots)/\mathrm{Tr}1$, $S^{\pm}\equiv S^x\pm iS^y=\Sigma_j n_j S_j^{\pm}$, n_j is the occupation number of site j ($n_j=0$ or 1), $\langle\ldots\rangle_c$ denotes configurational averaging, and

$$S^+(t)=\exp\{iH_dt\}S^+\ \exp\{-iH_dt\},$$

where H_d is the Hamiltonian of the spin-spin interaction. In the Anderson model

$$H_d=\frac{3}{2}\sum_{i<j}A_{ij}n_in_jS_i^zS_j^z,$$

where $A_{ij}\equiv\hbar\gamma^2(1-3\cos^2\vartheta_{ij})/r_{ij}^3$ is the dipolar interaction. A more general interaction, which is realized under typical conditions, is described by the secular part of the dipolar Hamiltonian and has the form

$$H_d=\frac{1}{2}\sum_{i<j}A_{ij}n_in_j(3S_i^zS_j^z-\mathbf{S}_i\mathbf{S}_j).$$

According to its meaning, $G(t)$ is an additive single-particle function and is therefore expanded into a sum of contributions from clusters of rank l:

$$G(t)=\sum_{l=1}^{k}G_l(t),$$

where $G_l(t)\equiv\langle\langle S_l^+(t)\rangle_0\rangle_c/\langle\langle S^+(0)\rangle_0\rangle_c$.

Here and below k is the highest rank chosen for the orthogonal clusters,[3] and the indices i and j label the lattice sites.

The cluster expansion of an additive single-particle operator A has the form

$$A\equiv\sum_i n_i\alpha_i=\sum_{l=1}^{k}\sum_{i_1<\ldots<i_l}Z_l(\mathbf{r}_{i_1},\ldots,\mathbf{r}_{i_l})\sum_{p=1}^{l}\alpha(\mathbf{r}_{i_p}), \tag{2}$$

where Z_l is the occupation number of an orthogonal cluster located at the sites $\mathbf{r}_{i_1},\ldots,\mathbf{r}_{i_l}$ (Ref. 3), and $l=1$ and $Z_1(\mathbf{r})$ refer to a spin from the matrix and its occupation number. In the Anderson model

$$n_0S_0^+(t)=n_0\ \exp(iH_dt)S_0^+\exp(-iH_dt)$$
$$=n_0S_0^+\prod_{j\neq0}\exp\left(\frac{3}{2}iA_{0j}S_j^zn_jt\right);$$

therefore,

$$n_0\langle S_0^+(t)S^-\rangle_0=\frac{2}{3}n_0s(s+1)\prod_{j\neq0}\{1+n_j[B_j(t,s)-1]\}, \tag{3}$$

where s is the spin of the particles in the system,

$$B_j(t,s)\equiv\frac{1}{2s+1}\sum_{\sigma=-s}^{s}\exp\left(\frac{3}{2}itA_{0j}\sigma\right)$$
$$=\frac{\sin[(3/2)tA_{0j}(s+1/2)]}{(2s+1)\sin((3/4)tA_{0j})}.$$

As usual, during the derivation of (3) that the identity $\varphi(n_x)\equiv\varphi(0)+n_x[\varphi(1)-\varphi(0)]$ for the arbitrary function $\varphi(n_x)$ was taken into account. The continuous-medium approximation corresponds to the limit where the relative concentration of spins satisfies $f\equiv\langle n_j\rangle_c\to0$, but the density $C=fn$ (n is the number density of the lattice sites) remains finite, and the sums over the lattice are replaced by integrals over all space. Thus,

$$G_l(t)=\left\langle\frac{n^{l-1}N}{(l-1)!}\int d\mathbf{r}_2\ldots d\mathbf{r}_lZ_l(0,\mathbf{r}_2,\ldots,\mathbf{r}_l)\right.$$
$$\left.\times\langle S_0^+(t)S_0^-\rangle_0\right\rangle_c\frac{1}{\langle\langle S^+S^-\rangle_0\rangle_c}. \tag{4}$$

Here N is the number of sites in the crystal.

With consideration of (3) it is seen that the configurational average reduces to the expression

$$\left\langle Z_l(0,\mathbf{r}_2,\ldots,\mathbf{r}_l)\prod_{j\neq0}\{1+n_j[B_j(t,s)-1]\}\right\rangle_c.$$

We first calculate

$$\left\langle\widetilde{Z}_l(0,\mathbf{r}_2,\ldots,\mathbf{r}_l)\prod_{j\neq0}\{1+n_j[B_j(t,s)-1]\}\right\rangle_c,$$

where $\widetilde{Z}_l$ is the occupation number of a nonorthogonal cluster of rank l. Taking into account the uncorrelated character of the distribution of the spins among the sites, in the continuous-medium approximation we obtain

$$\left\langle \tilde{Z}_l(0,\mathbf{r}_2,\ldots,\mathbf{r}_l)\prod_{j\neq 0}\{1+n_j[B_j(t,s)-1]\}\right\rangle_c$$

$$=\langle \tilde{Z}_l(0,\mathbf{r}_2,\ldots,\mathbf{r}_l)\rangle_c \prod_{m=2}^{l} B_{\mathbf{r}_m}(t,s)$$

$$\times\exp\left\{-C\int_{\mathbf{q}\notin V_{0\ldots\mathbf{r}_l}} d\mathbf{q}[1-B_{\mathbf{q}}(t,s)]\right\}, \tag{5}$$

where $V_{0\ldots\mathbf{r}_l}$ is the excluded volume of the cluster $\tilde{Z}_l$, the region within which there must be no other spins besides the spins comprising the cluster, so that the condition for its existence would not be violated.

Taking into account that $Z_l=0$ and $\tilde{Z}_l\neq 0$ holds only if $\tilde{Z}_l$ appears in the composition of a cluster of higher rank, we have the recurrence relation

$$Z_l(\mathbf{r}_1,\ldots,\mathbf{r}_l)=\tilde{Z}_l(\mathbf{r}_1,\ldots,\mathbf{r}_l) - \sum_{m=1}^{k-l}\sum_{\substack{k_1<\ldots<k_m\\ k_j\notin V_{\mathbf{r}_1\ldots\mathbf{r}_l}}} Z_{l+m}(\mathbf{r}_1,\ldots,\mathbf{r}_l,\mathbf{r}_{k_1}\ldots\mathbf{r}_{k_m}). \tag{6}$$

Hence follow the corresponding formulas for $\langle\ldots\rangle_c$ and the formal transition to the continuous-medium limit. If two arguments in the occupation number of a cluster coincide, it is equal to zero.

Equation (6) permits the calculation of $\langle Z_l\rangle_c$ successively from the maximum value $l=k$ to $l=1$, since

$$\langle \tilde{Z}_l(\mathbf{r}_1,\ldots,\mathbf{r}_l)\rangle_c = f^l\exp\{-fnV_r\}$$

in the limit $f\ll 1$.[3]

Using Eqs. (4)–(6), we can calculate the contribution $G_l(t)$ of a cluster of any rank to the correlation function.

3. CLUSTER EXPANSION OF SHAPE FUNCTIONS AND FREE-INDUCTION DECAY IN THE ANDERSON MODEL

We stipulate that the superscript on $G(t)$ and its Fourier transform $g(\Delta)$ denotes the maximum rank k of the clusters separated. As usual, the Anderson linewidth

$$D=\frac{2\pi^2}{3\sqrt{3}}\hbar C\gamma^2$$

determines the characteristic time scale of the problem. The normalization condition

$$\sum_{l=1}^{k} G_l^k(0)=1$$

is presumed. Then $G_l^k(0)$ has the meaning of the probability that a spin appears in an orthogonal cluster of rank l, if the maximum rank of the clusters separated equals k.

If we take $k=1$, i.e., if the system is considered as a whole without separating clusters, $Z_1(\mathbf{r})=n_r$ holds and

$$G_1^1(t)=\exp(-F_sDt), \tag{7}$$

where

$$F_s\equiv\frac{\int d\mathbf{r}[1-B_r(t,s)]}{\int d\mathbf{r}[1-B_r(t,1/2)]}=\frac{2}{2s+1}\sum_{\sigma=-s}^{s}|\sigma|.$$

We recall that all the correlation functions considered are invariant under time reversal. Therefore, it is assumed here and below that $t\geqslant 0$. Calculating F_s, for a half-integer spin and an integer spin we obtain, respectively,

$$F_s=s+\frac{1}{2},\quad F_s=\frac{2s(s+1)}{2s+1}=\left(s+\frac{1}{2}\right)\left[1-\frac{1}{(2s+1)^2}\right]. \tag{8}$$

Thus, the behavior of the free-induction decay is qualitatively identical for any value of the spin. For $s=1/2$, Eq. (7) coincides with the familiar expression,[9] while in the limit $s\to\infty$ we have $G_1^{1\mathrm{cl}}(t)=\exp(-DLt)$, where $L\hbar=\hbar\sqrt{s(s+1)}$ is the angular momentum of the particle. The corresponding relation (7) for the absorption line shape function describes a Lorentzian line with a half-width DF_s.

We note that the classical limit (8) can also be obtained by direct integration followed by averaging of the equation of motion of a classical spin in the Anderson model

$$\dot{\mu}_m^+=\frac{3}{2}{\sum_j}' iA_{mj}L_j^z\mu_m^+n_j$$

for the component $\mu_m^+\equiv\mu_m^x+i\mu_m^y$ of the magnetic moment in the mth lattice site with consideration of the constancy of $\mu_j^z=\gamma L_j^z$. Here and below the prime on the summation sign (differentiation) signifies the absence of a diagonal term with A_{mm}. The averaging over the configurations is performed just as in the quantum case, and the thermodynamic averaging reduces to the averaging of $\mathbf{L}_j$ over the directions.

Let us now consider the situation in which the maximum cluster rank is $k=2$ (we separate pairs and thus take into account the inhomogeneity existing in the system). Using (4)–(6), we obtain

$$G_2^2(t)=C\int d\mathbf{r}B_r(t,s)\exp\left\{-CV_{0r}-C\times\int_{q\notin V_{0r}}d\mathbf{q}[1-B_q(t,s)]\right\}$$

$$=\exp(-F_sDt)C\int d\mathbf{r}B_r(t,s)\times\exp\left\{-C\int_{q\in V_{0r}}d\mathbf{q}B_q(t,s)\right\}. \tag{9}$$

It was shown in Ref. 3 that the excluded volume of a two-spin cluster can be represented in the form

$$V_{0r}=\frac{4\pi}{3\kappa}\frac{\overline{\chi}}{\chi_{0r}}r^3,$$

where $\chi_{0r}\equiv 1-3\cos^2\vartheta_{0r}$,

$$\overline{\chi}\equiv\int_0^1 d\xi|1-3\xi^2|=\frac{4}{3\sqrt{3}},$$

and $\kappa\equiv G_2^2(0)=0.58\pm 0.01$ is the fraction of spins in pairs.

The exact calculation of the asymptote $G_2^2(t)$ is hampered by the complicated configuration of the volume V_{0r}; however, to obtain a semiquantitative result we make the replacements $r^3 \to r^3|\chi_{0r}|$ and $q^3 \to q^3|\chi_{0r}|$ in the second integral in Eq. (9) and then assume that the volume obtained from the excluded volume after these transformations has the shape of a sphere. We note that the comparatively short distance between the spins of the pair forming the excluded volume relative to the characteristic dimension of that volume renders trivial the question of the position of its center, which we place at zero. After elementary integration over the angular variables and replacement of the radial variables, which simplify the cosine arguments, for $s=1/2$ we obtain

$$G_2^2(t)=\exp(-Dt)\frac{2Dt}{\pi}\int_0^\infty \frac{dx}{x^2}\cos x \times\exp\left\{-\frac{2Dt}{\pi}\int_{\kappa x}^\infty \frac{dy}{y^2}\cos y\right\}. \tag{10}$$

We stress that in (10) $\cos y$ and $\cos x$ originate directly from $B_r(t,1/2)$ and $B_q(t,1/2)$. In the limit $Dt \gg 1$, the asymptotic expression (10) is specified by the saddle point $\kappa x_0=\pi/2$ and has the form

$$G_2^2(t)\sim \frac{4\kappa}{\pi}\sqrt{Dt}\,\exp[-Dt(1-\delta)]cos\frac{\pi}{2\kappa},$$

where $\delta \equiv 2/(5\pi) \ll 1$. Another calculation method, which takes into account the angular dependences in (9) and takes the excluded volume V_{0r} itself as a sphere, gives an even smaller value of δ. Since $G(t)=G_2^2(t)+G_1^2(t)=\exp(-Dt)$, at large times the functions $G_2^2(t)$ and $G_1^2(t)$ in the Anderson model decrease with identical rates and are opposite in sign.

We find the Fourier transform of $G_2^2(t)$, i.e., the shape functions of the pairs:

$$g_2^2(\Delta)\equiv\frac{1}{\pi}\,\mathrm{Re}\int_0^\infty G_2^2(t)e^{i\Delta t}dt.$$

Substituting (9) therein and calculating the integral over t with consideration of the fact that $B_r(t,s)=b(r/t^{1/3},s)$, we obtain

$$g_2^2(\Delta)=\frac{1}{\pi}C\int d\mathbf{r}b(r,s)\left\{C^2\left\{V_{0r}+\int_{q\notin V_{0r}}d\mathbf{q}[1-b(q,s)]\right\}^2-\Delta^2\right\}\times\left\{C^2\left\{V_{0r}+\int_{q\notin V_{0r}}d\mathbf{q}[1-b(q,s)]\right\}^2+\Delta^2\right\}^{-2}. \tag{11}$$

Making the replacement

$$V_{0r}+\int_{q\notin V_{0r}}d\mathbf{q}[1-b(q,s)] \to \sqrt{V_{0r}^2+\left\{\int d\mathbf{q}[1-b(q,s)]\right\}^2},$$

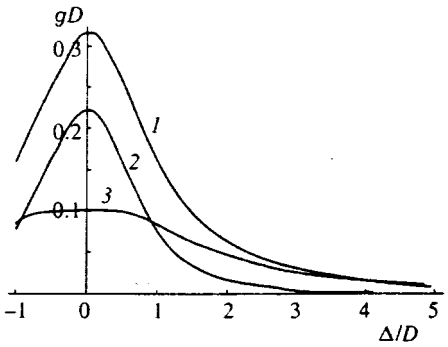


FIG. 1. Cluster expansion of the absorption line shape function for $s=1/2$: *1* — g_1^1D, *2* — g_1^2D, *3* — g_2^2D.

which is exact for small and large values of r (and introduces, as will be shown below, a small error when $\Delta/D \leqslant 1$), for $x=1/2$ we obtain

$$g_2^2(\Delta)=\frac{1}{\pi D}\exp\left(-\frac{\pi_1}{\sqrt{1+\Delta^2/D^2}}\right)\left(1+\frac{\Delta^2}{D^2}\right)^{-2}\times\left(\frac{\Delta^2}{D^2}+\frac{1}{\pi_1}\sqrt{1+\frac{\Delta^2}{D^2}}\right), \tag{12}$$

where $\pi_1 \equiv 2/(\pi\kappa)=1.1$. The same expression is obtained through a similar approximation directly from (10). In the classical limit, instead of (12) we have

$$g_2^{2\mathrm{cl}}(\Delta)=\frac{1}{2\pi\pi_1^2 DL}\left\{1-\exp\left[-\frac{2\pi_1}{\sqrt{1+\Delta^2/(DL)^2}}\right]-\frac{2\pi_1\Delta^2}{(DL)^2}\times\exp\left[-\frac{2\pi_1}{\sqrt{1+\Delta^2/(DL)^2}}\right]\left[1+\frac{\Delta^2}{(DL)^2}\right]^{-3/2}\right\}. \tag{13}$$

We performed a direct numerical analysis of the representation (11) for an isotropic interaction (when A_{0r} does not depend on the angle ϑ_{0r}) and $s=1/2$. The result was qualitatively consistent with (12), and the quantitative disparity was less than 10% and was concentrated in the region $\Delta \leqslant D$. The analogous quantities $G_1^2(t)$ and $g_1^2(\Delta)$ for the matrix are clearly obtained by subtracting (12) and (13) from the Fourier transform (7).

The behavior of the shape functions of the pairs, the matrix, and the system as a whole for the cases of $s=1/2$ and $s=\infty$ are presented in Figs. 1 and 2.

The absorption line (12) is nonzero at its center and has a broad plateau (with a weakly expressed minimum) at $\Delta=0$. In the classical limit the line shape function (13) has only one maximum at its center. From the physical standpoint this difference is determined to a considerable degree by the shape of the envelope of the lines corresponding to a particular state of the nearby spin: for $s \geqslant 1$ the absorption line has only one maximum at its center. It follows from (12) and (13) that the pairs deplete the wing of the line (compare Ref. 3).

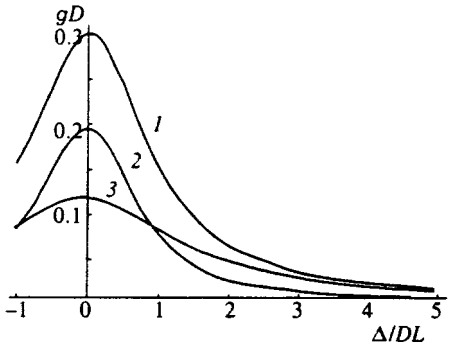


FIG. 2. Cluster expansion of the absorption line shape function. Classical limit: *1* — $g_1^{1\mathrm{cl}}DL$, *2* — $g_1^{2\mathrm{cl}}DL$, *3* — $g_2^{2\mathrm{cl}}DL$.

4. CONSIDERATION OF TEMPORAL FLUCTUATIONS IN THE APPROXIMATION OF A NORMAL RANDOM PROCESS

If the free-induction decay signal (3) in the Anderson model is approximated by a Gaussian function with an exact second moment under the assumption that

$$\langle S_j^+(t)S_j^-\rangle_0=\frac{2}{3}s(s+1)\exp\left(-\frac{1}{2}M_{2j}t^2\right), \tag{14}$$

$$M_{2j}\equiv\frac{3}{4}s(s+1)\sum_m{}' A_{mj}^2 n_m,$$

then after performing configurational averaging we find that

$$G(t)=\prod_j{}'\left\{1+f\left[\exp\left(-\frac{3}{8}s(s+1)A_{0j}^2t^2\right)-1\right]\right\}$$
$$=\exp\left(-\sqrt{\frac{2}{\pi}}\,DtF_s\right),\quad f\ll 1. \tag{15}$$

These formulas were previously obtained for $s=1/2$ in Refs. 7 and 8. The argument of the exponential function in them for free-induction decay is very close to the correct one [see Eq. (7)], being smaller than it by a factor of 1.25. This accuracy is surprisingly high for such a rough approximation.

As we know, Gaussian temporal decay corresponds to a normal distribution of the static local fields. If the temporal fluctuations of the local fields are approximated by a normal random process in the spirit of the AWK approximation, it is found that

$$\langle S_0^+(t)S_0^-\rangle_0=\frac{2}{3}s(s+1)\exp[-\Phi_0(t)], \tag{16}$$

where

$$\Phi_0(t)=\int_0^t d\tau(t-\tau)\langle\omega_{l0}(\tau)\omega_{l0}\rangle_0,$$

$\omega_{l0}=(3/2)\Sigma_m'A_{m0}n_mS_m^z$ is the local field at spin 0. We define the correlation function $K(t)$ of the local field $\omega_{l0}(t)$ by the relation

$$\langle\omega_{l0}(t)\omega_{l0}\rangle_0=M_{20}K(t),\quad K(0)=1. \tag{17}$$

Equations (16) and (17) differ from the standard version of the AWK theory in that M_{20} and $K(t)$ depend on the realization of the random distribution of the impurities in the lattice.

The local fields fluctuate because of flip-flop transitions, and, thus, $K(t)$ characterizes their statistical properties. Since the fluctuation rates of the spins depend on the features of their interaction, the function $K(t)$ is different for different configurations of the spins in the lattice. However, this dependence on the configuration is weaker than the dependence of M_{20} on the configuration. If the weaker dependence is neglected, the function $K(t)$ is assumed to be independent of the configuration, and we define

$$I(t)=\int_0^t d\tau(t-\tau)K(\tau),$$

relation (16) can be averaged over the configurations, and we find that[3]

$$G(t)=\exp[-2DF_s\sqrt{I(t)/\pi}]. \tag{18}$$

For

$$t\gg\tau_{c_0}\equiv\int_0^\infty d\tau K(\tau)$$

it therefore follows that

$$\ln G(t\gg\tau_{c_0})=-2DF_s\sqrt{t\tau_{c_0}/\pi},$$

in agreement with the conclusions in Ref. 17, which were drawn on the basis of other arguments.

The AWK approximation is valid if a local field is the sum of a large number of approximately identically distributed terms. The cluster separation method leads to the inclusion of the most singular part of the interactions in the intracluster interactions, and it is thus taken into account exactly. Therefore, after separation of the clusters, the local fields induced at the spins of the matrix and in the clusters by the surrounding spins become far closer to a normal process, and the results obtained by combining the AWK theory and cluster expansion should be far more exact than (15). Below we shall formulate these results and test them by comparing them in the limit $\tau_cD\gg 1$ with the exactly solved Anderson model.

We separate the pairs from the system and consider a 2-cluster composed of the spins S_0 and S_1. The Hamiltonian of this pair is

$$H_p=A_{01}\left(S_0^zS_1^z-\frac{1}{4}S_0^+S_1^--\frac{1}{4}S_0^-S_1^+\right)+\omega_{l0}(t)(S_0^z+S_1^z).$$

Here, as before, we neglect the dependence of the field induced by the spins outside the pair on the coordinate of each spin within the cluster.

For $s=1/2$ the equations of motion for spins 0 and 1 of the pair are easily integrated. As a result, we obtain

$$\langle S_0^+(t)S^-\rangle_0=\frac{1}{2}\cos\left(\frac{3}{4}A_{01}t\right)\exp[-M_{20}I(t)], \tag{19}$$

and Eq. (16) remains valid for a matrix spin.

In this section we neglect the fluctuations of the correlator $K(t)$. The more general case is considered in the next section.

We substitute (19) into (4) and perform configurational averaging just as in (5) with consideration of the fact that the occupation number of site j by a spin from a 2-cluster is

$$n_j^c \equiv \sum_q Z_2(r_j, q),$$

and the occupation number of the same site by a matrix spin is $n_j^m \equiv n_j - n_j^c$. As a result, for the functions describing free-induction decay from the pairs and from the matrix we obtain the expressions:

$$G_2^2(t) = C\int d\mathbf{r}\,\cos\frac{3A_{0r}t}{4}\exp(-CV_{0r}) \times\exp\left\{-C\int_{q\notin V_{0r}} d\mathbf{q}\left(1-\exp\left[-\frac{9}{16}A_q^2 I(t)\right]\right)\right\},$$

$$G_1^2(t) = \exp\left[-\frac{2}{\sqrt{\pi}}D\sqrt{I(t)}\right] - C\int d\mathbf{r} \times\exp\left[-\frac{9}{16}A_{0r}^2 I(t) - CV_{0r}\right] \times\exp\left\{-C\int_{q\notin V_{0r}} d\mathbf{q}\left(1-\exp\left[-\frac{9}{16}A_q^2 I(t)\right]\right)\right\}. \tag{20}$$

Let us discuss the properties of the correlation function $K(t)$ defined by Eq. (17). It is composed of the longitudinal components of the spin operators. The correlator

$$q(t) \equiv \langle\langle S_i^z(t) S_i^z\rangle_0\rangle_c / \langle\langle (S_i^z)^2\rangle_0\rangle_c$$

behaves like $\exp(-Dt/3)$ at small Dt, as follows from an analysis of the first term in the concentration expansion. The function $K(t)$ decreases with the characteristic time τ_c, where τ_c^{-1} is the fluctuation rate. In ordered systems it satisfies $\tau_c D \sim 4$ (see Ref. 18 and the references cited therein), so that we should expect the same relation in the central part of the line for the flip-flop transitions in the matrix and in the pairs.

The sum $\Sigma_{i=1}^{N_s} S_i^z n_i$ (summation over the spins) is an integral of motion; therefore, it is natural to expect that the long-time asymptote will be diffusive, i.e., $q(t \gg \tau_c) \sim (Dt)^{-3/2}$. At intermediate values of t (in the range $1/2 \geqslant q(t) \geqslant 0.1$) the behavior of $q(t)$ is apparently satisfactorily expressed by the Förster exponential function $\exp(-U\sqrt{Dt})$, where the constant satisfies $U \sim 1$.[16] The behavior of $K(t)$ is qualitatively similar to the behavior of $q(t)$. Therefore, the influence of the diffusive tail on $I(t)$ is small, at least in three-dimensional systems. The influence of the initial, purely exponential segment is also insignificant. Therefore, we shall henceforth generally assume in the numerical evaluations that $K(t) = \exp(-\sqrt{t/\tau_c})$.

As a result, nonzero values of τ_c^{-1} are displayed only for $tD \sim 10$ and thus do not strongly influence the line shape.

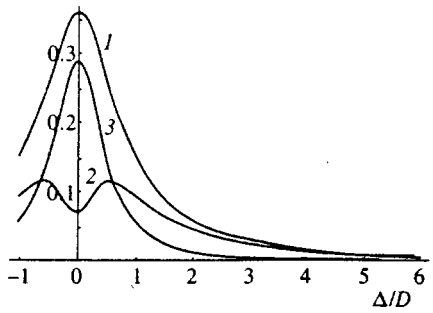


FIG. 3. Cluster expansion of the absorption line shape function for $\tau_c = 4/D$ and $s = 1/2$: *1* — gD, *2* — $g_2^2 D$, *3* — $g_1^2 D$.

For $\tau_c^{-1} = 0$ we have $I(t) = t^2/2$. Taking the Fourier transform (2), integrating over t, and leaving the principal terms with respect to r in the numerator and the denominator of the resulting integrands, just as in the derivation of (12), after some relatively simple transformations we obtain

$$g_2^2(\Delta) = \frac{D}{\pi D_\delta^2}\exp\left(-\frac{\pi_1 D/D_\delta}{\sqrt{1+\Delta^2/D_\delta^2}}\right)\left(1+\frac{\Delta^2}{D_\delta^2}\right)^{-2} \times\left(\frac{\Delta^2}{D_\delta^2}+\frac{1}{\pi_1}\sqrt{\frac{2}{\pi}}\sqrt{1+\frac{\Delta^2}{D_\delta^2}}\right), \tag{21}$$

$$g_1^2(\Delta) = \frac{1}{\pi}\frac{D_\delta}{D_\delta^2+\Delta^2} - \frac{1}{\pi^2\pi_1}\int_0^\infty \frac{dy}{\sqrt{y}} \times\exp\left(-\frac{\pi_1^2}{2}y\right)\frac{D^2+y(D_\delta^2-\Delta^2)}{[D^2+y(D_\delta^2+\Delta^2)]^2}.$$

Here $D_\delta \equiv \sqrt{2/\pi}D$ is the decay rate of $G_1^1(t) = \exp(-D_\delta t)$ in the model under consideration without cluster separation.[3] In the limit $\Delta \gg D$ it is not difficult to see that $g_1^2(\Delta) \sim 1/\Delta^3$ and $g_2^2(\Delta) = D/\pi\Delta^2$, i.e., the pairs deplete the wing of the line, and the asymptote of $g(\Delta)$ coincides with that in the Anderson model. The function (21) also practically coincides with the function obtained in the Anderson model in the central part of the line. To illustrate the influence of $\tau_c^{-1} \neq 0$ on the line shape function, we calculate $g(\Delta) = g_2^2(\Delta) + g_1^2(\Delta)$, i.e., the absorption line shape function of the system, in the approximation of a normal random process for $\tau_c^{-1} = D/4$ and $s = 1/2$. The corresponding plots are presented in Fig. 3. It is seen that the presence of fluctuations has some influence on the center of the line. As expected, the slower decay of $G(t)$ leads to an increase in $g(0)$ in comparison with the case of no fluctuations. The behavior of $g_2^2(\Delta)$ near 0 represents an intermediate case between $\tau_c^{-1} = 0$ and $\tau_c^{-1} = \infty$. In the latter case $g_2^2(0) = 0$.[3]

Transforming the integration variables and the excluded volume in (20) as we did to obtain Eq. (10), we obtain explicit integral representations for the free-induction decay functions:

$$G_2^2(t) = \frac{2}{\pi}\kappa D\sqrt{I(t)}\,\exp\left[-\frac{2}{\sqrt{\pi}}D\sqrt{I(t)}\right]$$

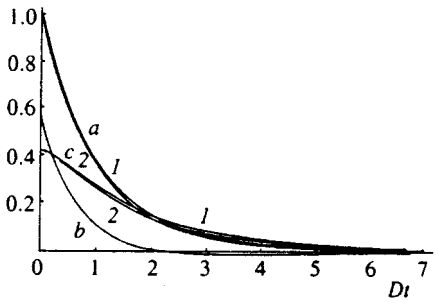


FIG. 4. Cluster expansion of the free-induction decay: *1* — $\tau_c=4/D$, *2* — Anderson model; a — total signal, b — signal from pairs, c — signal from the matrix.

$$\mathrm{Re}\int_0^\infty \frac{dx}{x^2}\exp\left\{\frac{t}{\kappa\sqrt{I(t)}}\times\left[ix-\frac{2\kappa}{\pi}\frac{DI(t)}{t}\int_x^\infty\frac{dy}{y^2}\exp(-y^2)\right]\right\}, \tag{22}$$

$$G_1^2(t)=\exp\left[-\frac{2}{\sqrt{\pi}}D\sqrt{I(t)}\right]\left\{1-\frac{2\kappa}{\pi}D\sqrt{I(t)}\int_0^\infty\frac{dx}{x^2}\times\exp\left[-\frac{x^2}{\kappa^2}-\frac{2}{\pi}D\sqrt{I(t)}\int_x^\infty\frac{dy}{y^2}\exp(-y^2)\right]\right\}.$$

Plots of the functions (22) and their sum $G(t)$ for the Förster form of $K(t)$ and $\tau_c D=4$ are presented in Fig. 4. For comparison, the figure also presents plots of these functions in the Anderson model.

Let us investigate their long-time asymptotes. The asymptote $G_1^2(t)$ can be found using the Laplace method by representing the integral appearing in it in the form

$$\int_0^\infty d\psi[\psi'(x)]^{1/\kappa^2-1}\exp\left[-\frac{2}{\pi}D\sqrt{I(t)}\psi\right],$$

where

$$\psi(x)\equiv\int_0^x dy\,\exp\left(-\frac{1}{y^2}\right),$$

and it has the form

$$G_1^2(t\gg\tau_c)\sim\exp\left[-\frac{2}{\sqrt{\pi}}D\sqrt{I(t)}\right]\left\{1-\kappa\Gamma\left(\frac{1}{\kappa^2}\right)\times\left[\frac{D\sqrt{I(t)}}{\pi}\right]^{1-1/\kappa^2}[\ln(D\sqrt{I(t)})]^{3(1/\kappa^2-1)/2}\right\}. \tag{23}$$

We recall that $I(t\gg\tau_c)\sim 2t\tau_c$ for finite τ_c and $I(t\to\infty)\sim t^2/2$ for $\tau_c^{-1}=0$.

If $\tau_c^{-1}=0$ holds, the asymptote of $G_2^2(t)$ can be found by representing the integral appearing in this function in the form

$$\mathrm{Re}\int_0^\infty d\psi\frac{d\zeta}{d\psi}\exp\left(-\frac{\sqrt{2}}{\pi}Dt\psi\right),$$

where $\psi(x)$ is defined precisely as in the derivation of (23) and

$$\zeta(x)\equiv\int_0^x dx\,\exp\left\{\frac{i\sqrt{2}}{\kappa x}\right\}.$$

As a result

$$G_2^2(Dt\gg1,\tau_c^{-1}=0)\sim-\frac{\kappa^2 Dt}{\pi}\exp\left(-\sqrt{\frac{2}{\pi}}Dt\right)\times\sin\left(\frac{\sqrt{2\ln Dt}}{\kappa}\right)[\ln(Dt)]^{-1}. \tag{24}$$

Comparing (23) and (24) with (7) and the asymptotic expression (10), we can see good agreement between the Anderson model and the approximation of a normal random process in the limit of small fluctuations. In the range $0\leqslant Dt\leqslant 10$ the difference between the results of these two models is so small that it is not visible on the plots; however, the relative difference does not decrease as t increases and thus leads to an appreciable difference between the corresponding shape functions near $\Delta=0$.

The main difference between the approximation of a normal random process for $\tau_c=\infty$ and the Anderson model, which is exact in this case, is confined to the difference between the decay rates of the free-induction decay signal of the pairs, which amounts to

$$\left(\left(1-\frac{2}{5\pi}\right)-\sqrt{\frac{2}{\pi}}\right)\left(\frac{2}{\pi}\right)^{-1/2}=9.4\%.$$

The functions $G_2^2(t)$ coincide to within a uniform accuracy of 5% in the range $0\leqslant Dt\leqslant 10$. The Fourier transforms $g_2^2(\Delta)$ exhibit the greatest difference at $\Delta=0$ (the relative difference is 7%). It decreases with increasing Δ and becomes less than 1.5% when $\Delta\geqslant D$. The total free-induction decay coincides even more closely with the Anderson model. Thus, the separation of 2-clusters led to a radical change in the accuracy of the result $G(t)=\exp(-\sqrt{2/\pi}Dt)$, which follows from (18) when $\tau_c^{-1}=0$.

To find the asymptotic form of $G_2^2(t)$ at finite values of τ_c, we represent the integral appearing in it (after the replacements of variables $x\to 1/x$ and $y\to 1/y$) in the form

$$\mathrm{Re}\int_0^\infty dx\,\exp\left[\sqrt{\frac{t}{2\kappa^2\tau_c}}f(x)\right],$$

where

$$f(x)\equiv\frac{i}{x}-\frac{4\kappa D\tau_c}{\pi}\int_0^x dy\,\exp\left(-\frac{1}{y^2}\right).$$

Here

$$G_2^2(t\to\infty)\sim\frac{4\cdot 2^{1/4}}{\sqrt{\pi}}\kappa^{3/2}D\tau_c^{3/4}t^{1/4}\times\exp\left(-\frac{2\sqrt{2}}{\sqrt{\pi}}D\sqrt{t\tau_c}\right)$$

$$\times \mathrm{Re}\left\{ \exp\left(f(x_0) \sqrt{\frac{t}{2\kappa^2 \tau_c}} \right) \sqrt{\frac{-1}{f''(x_0)}} \right\}. \quad (25)$$

Here x_0, i.e., the saddle point of the function $f(x)$, is the stationary point farthest from 0 which satisfies the equation

$$\frac{i}{x^2} + \frac{4\kappa D\tau_c}{\pi} \exp\left(-\frac{1}{x^2} \right) = 0.$$

The asymptotic expression (25) can easily be found in the limit $D\tau_c \to 0$. Then

$$x_0 = \frac{1}{2} \sqrt{\frac{\pi}{\kappa D\tau_c}} \frac{1-i}{\sqrt{2}},$$

and

$$G_2^2(Dt \gg 1) \sim (2\pi\kappa^3 Dt)^{1/4} \cos\left(\sqrt{\frac{4Dt}{\pi\kappa}} - \frac{\pi}{8} \right) \times \exp\left(-\sqrt{\frac{4Dt}{\pi\kappa}} - 2\sqrt{\frac{2}{\pi}} D\sqrt{t\tau_c} \right). \quad (26)$$

Thus, damping of the free-induction decay function of the pairs occurs even in the limit of infinitely large fluctuations, i.e., due to the configurational averaging of the motions in the pairs. As τ_c increases, the point x_0 monotonically approaches 0, and for $D\tau_c = 4$ it equals $x_0 = 0.845 - 0.331i$. In this case

$$G_2^2(t \to \infty) \sim 1.42 (Dt)^{1/4} \cos(0.779\sqrt{Dt}) \times \exp(-3.43\sqrt{Dt}). \quad (27)$$

When the fluctuation rate decreases further, x_0 moves very slowly toward the real axis:

$$x_0 \sim \exp\left\{ -\frac{i\pi/4}{\ln(D\tau_c)} \right\} / \sqrt{\ln(D\tau_c)}.$$

In this case

$$f(x_0) \sim i\sqrt{\ln(D\tau_c)} - \frac{\pi/4}{\sqrt{\ln(D\tau_c)}}.$$

A comparison of (23) and (27) gives the asymptotic expression

$$G(t) \sim \exp\left(-2\sqrt{\frac{2}{\pi}} D\sqrt{t\tau_c} \right)$$

for $t \gg \tau_c$ (compare Ref. 3). The function approaches this asymptote at $t/\tau_c \sim 6$, before which its decay is approximately exponential.

Thus, the two characteristic times observed in the system specify, respectively, the short- and long-time asymptotes of the free-induction decay functions.

5. LOCAL-FIELD FLUCTUATION RATE SPECTRUM

In calculating the longitudinal correlator $K(t)$ of the local fields at a particular spin we have hitherto neglected the dependence of $K(t)$ on the local environment of that spin. Let us analyze this dependence to evaluate our approximation and refine the form of the function.

The spin flips belonging to the wing of the resonance line (mainly of the spins in pairs) are caused primarily by two processes: a two-spin process in pairs found in the same segment of the resonance line with a width of order D and a four-spin process, which couples spins on the wing of the line with spins at its center. The latter process is of a higher order, but it is resonant relative to the dipolar Hamiltonian of an isolated pair and makes it possible to transfer energy over large distances through the spectrum.[3,5] The characteristic time $\tau_c(\omega)$ of the flip-flop transitions producing fluctuations of the local field from the spins on the wing of the line is greater than $\tau_c(0)$ in the center of the line and increases as the resonant frequency ω of the spins increases relative to the center of the line. When $\omega \gg D$ holds, ω is, in fact, the eigenfrequency of the pair to which the spin under consideration belongs.

The matrix element of a two-spin process is, in order of magnitude, A_{0x}, where 0 and x are the spatial positions of two mutually resonant pairs, and the matrix element of a four-spin process is A_{0q}^2/ω, where q is the coordinate of a spin in the matrix or a pair relative to the center of the line.[3] We consider $\omega \gg D$, since $\tau_c(D)$ is of the order of τ_c, i.e., the characteristic time of the flip-flop transitions in the central portion of the line. The characteristic interaction of pairs with the matrix is greater in magnitude than the interpair interaction,[3] and thus the four-spin interpair process can be neglected, since it does not differ fundamentally from the pair-matrix process.

We use W_m to denote the rate of the pair-matrix process and W_p to denote the rate of the pair-pair process. Then the spin-flip rate of a spin in a pair is

$$W = W_m + W_p = U_m \sum_{q \neq r}{}' \frac{A_{0q}^2 A_{0r}^2}{D\omega^2} Z_1(q) Z_1(r) + U_p \sum_x{}' \frac{A_{0x}^2}{D} Z_{2\omega}(x). \quad (28)$$

Here U_m and U_p are coefficients of order unity, and $Z_{2\omega}(x)$ is the occupation number of site x by a pair with the frequency ω. We neglect the dimensions of the pair in comparison to the distance from it to other objects, and we assume that the characteristic phase relaxation rate, which determines W_m and W_p, coincides with D and is identical throughout the entire system [thus, Eq. (28) is the next approximation after the one considered in the preceding section from the standpoint of the spread in the rates of processes occurring in the system]. The characteristic time of the flip-flop transitions is

$$\tau_c(\omega) = \left\langle \int_0^\infty dt \, \exp(-Wt) \right\rangle_c,$$

and we write the correlator of the local fields from spin 0 as

$$K_0(t) \equiv \exp(-Wt) = \exp(-W_p t) \exp(-W_m t) \equiv K_{0p}(t) K_{0m}(t).$$

We use $\langle e^{-Wt}\rangle_c \approx \langle K_{0p}\rangle_c \langle K_{0m}\rangle_c$ as an estimate. Here we take the configurational average at a fixed frequency. Without taking into account the correlations in the relative positions of the pairs, for the first exponential function it is not difficult to obtain

$$\langle K_{0p}(t)\rangle_c = \langle \exp(-W_p t)\rangle_c \approx \exp\left(-U_1 \frac{D^2}{\omega^2}\sqrt{Dt}\right), \quad (29)$$

where $U_1 > 0$ is a constant.

To calculate the asymptote of $\langle \exp(-W_m t)\rangle_c$ we first note that because the characteristic pair–matrix interaction is greater in magnitude than the pair–pair interaction[3] and the interaction of the spins on the wing is relatively weak, the replacement $Z_1(q) \to n_q$ in (28) introduces only a small error. The condition $q \neq r$ is a short-range correlation and thus can be neglected in finding the long-time asymptote; however this correlation is strong (is a restraint); therefore, we shall conduct a more systematic analysis.

Let us consider the quantity

$$\tau_m(\alpha) = \lim_{\varepsilon \to 0} \left\langle \int_0^\infty dt\, \exp(i\varepsilon t + t^\alpha - W_m t) \right\rangle_c, \quad (30)$$

where Re $\varepsilon > 0$, Im $\varepsilon > 0$, and $0 \leqslant \alpha < 1$. The integral under the limit sign converges along any contour in the upper right-hand quadrant of the complex plane. We represent

$$W_m = U_m \sum_{q \neq r}{}' \frac{A_{0q}^2 A_{0r}^2}{D\omega^2} n_q n_r$$

as

$$W_m = \frac{U_m}{D\omega^2}\left(\sum_q{}' A_{0q}^2 n_q\right)^2 - \frac{U_m}{D\omega^2}\sum_q{}' A_{0q}^4 n_q,$$

write

$$\exp\left[-\frac{U_m}{D\omega^2}\left(\sum_q{}' A_{0q}^2 n_q\right)^2\right]$$

as a Gaussian integral to obtain an exponent that is linear with respect to the occupation numbers, turn the integration contour in (30) by $+\pi/2$, carry out the configurational averaging, and perform the replacement of the integration variables. As a result, in the limit $\varepsilon \to 0$ we obtain

$$\tau_m(\alpha) = \sqrt{\frac{i}{4\pi}}\int_0^\infty dt\, \exp[(it)^\alpha] t^{1/6} \int_{-\infty}^\infty dy\, \exp\left\{ t^{1/3}\left[\frac{i}{4}y^2 - \frac{C U_m^{1/4}}{D^{1/4}\sqrt{\omega}}\int d\mathbf{q}\left(1 - \exp\left(\frac{iA_q^4}{t^{1/3}} - iA_q^2 y\right)\right)\right]\right\}.$$

The asymptote of the inner integral is $\exp(-U_1 t^{1/3})$ and Re $U_1 > 0$. Thus, the limiting value for the existence of $\tau_m(\alpha)$ equals $\alpha = 1/3$. Therefore, the long-time asymptote of the correlation function of the local fields due to the four-spin interaction is

$$\ln\langle K_{0m}(t)\rangle_c = \ln\langle \exp(-W_m t)\rangle_c = -U_2\left(\frac{D^3 t}{\omega^2}\right)^{1/3}, \quad (31)$$

where U_2 is a constant and Re $U_2 > 0$. Comparing it with (29) we see that the characteristic flip-flop transition time $\tau_c(\omega) \sim \omega^2$ on the wing of the resonance line.

The correlator of the local fields is

$$\langle \omega_l(t)\omega_l\rangle = \sum_{jq} n_j n_q \langle \omega_l^j(t)\omega_l^q\rangle,$$

where ω_l^j is the field from the spin at site j. We assume that the fields from different spins do not correlate. Then

$$\langle \omega_l(t)\omega_l\rangle = \sum_j n_j \langle \omega_l^j(t)\omega_l^j\rangle = \sum_j{}' \frac{9}{16} A_{0_j}^2 n_j K_j(t), \quad K_j(0) = 1. \quad (32)$$

The distribution function of the static local fields in the system under consideration $\varphi(\omega) = \langle \delta(\omega^2 - M_2)\rangle_c$ has the form

$$\varphi(\omega) = \frac{D}{\pi\omega^3}\exp\left(-\frac{D^2}{\pi\omega^2}\right), \quad \int_0^\infty d\omega^2 \varphi(\omega) = 1. \quad (33)$$

Writing (32) in the form

$$\langle \omega_l(t)\omega_l\rangle = \sum_{j\omega}{}' \frac{9}{16} A_{0j}^2 n_{j\omega} K_j(t),$$

where $n_{j\omega}$ is the occupation number of site j by a spin with a static local field ω on it, neglecting the short-range spatial correlations of the $n_{j\omega}$ with one another, and bringing the integration variables into dimensionless forms, instead of (22) we obtain

$$G_2^2(t) = \exp\left(-\frac{2}{\sqrt{\pi}} D\langle\langle\sqrt{I_j(t)}\rangle_c\rangle_\omega\right)\frac{2\kappa}{\pi}\sqrt{Dt} \times \mathrm{Re}\int_0^\infty \frac{dx}{x^2}\exp\left\{\sqrt{Dt}\left(i\kappa^{-1}x - \frac{2}{\pi}\int_x^\infty \frac{dy}{y^2}\left\langle\left\langle \exp\left(-\frac{DI_j(t)}{t}y^2\right)\right\rangle_c\right\rangle_\omega\right)\right\}, \quad (34)$$

$$G_1^2(t) = \left(-\frac{2}{\sqrt{\pi}} D\langle\langle\sqrt{I_j(t)}\rangle_c\rangle_\omega\right)\left(1 - \frac{2}{\pi}\int_0^\infty \frac{dx}{x^2} \times \langle \exp[-D^2 x^2 I(x,t)]\rangle_c \times \exp\left\{-\frac{2}{\pi}\int_{\kappa x}^\infty \frac{dy}{y^2}\langle\langle \exp[-y^2 D^2 I_j(t)]\rangle_c\rangle_\omega\right\}\right).$$

Here

$$I_j(t) \equiv \int_0^t d\tau (t-\tau) K_j(\tau),$$

$I(x,t)$ is defined as $I_j(t)$ provided j is a spin with a local field Dx on it [x is the dimensionless frequency $x = 3A_{0r}/4D$ in (20)], $\langle \ldots \rangle_c$ indicates that the configurational averaging is carried out with a fixed value of the static local field ω, and

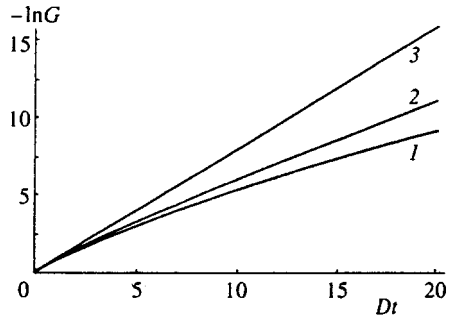


FIG. 5. Characteristic form of $-\ln G(t)$: *1* — case of a constant flip-flop transition time, *2* — slowing on the wing of the resonance line, *3* — $\tau_c=\infty$.

$$\langle X(\omega)\rangle_\omega \equiv \int_0^\infty d\omega^2 \varphi(\omega) X(\omega).$$

We are interested mainly in the long-time asymptotes (34), since the purpose of this section was to take into account the presence of slowly fluctuating objects, which begin to influence the behavior of the free-induction decay function specifically at large times.

In the expression for $G_2^2(t)$ in (34) the main contribution to the quantity

$$\left\langle \left\langle \exp\left\{ -\frac{DI_j(t)}{t} y^2 \right\} \right\rangle_c \right\rangle_\omega$$

is made by the configurations and frequencies for which $I_j(t)$ is smallest, i.e., when the spin generating the function $I_j(t)$ is located at the center of the line. The specific weight of these configurations is of order unity, so that the asymptote of the integral in (34) in the expression for $G_2^2(t)$ is the same as in (22). The integral in the parentheses in expression (34) for $G_1^2(t)$ decreases as t increases, as is seen already from the fact that consideration of slow objects leads to an effective increase in $I(t)$, and this accelerates the decay of the integral under consideration [see Eq. (23)].

To estimate $\langle\langle\sqrt{I_j(t)}\rangle_c\rangle_\omega$ we note that by introducing configurational averaging under the radical we increase this function only slightly (the asymptote with respect to t remains the same, as follows, for example, from the constrained character of the relative dispersion of $\tau_c(\omega)$, which is uniform with respect to ω). Thus,

$$\frac{2}{\sqrt{\pi}} D\langle\langle\sqrt{I_j(t)}\rangle_c\rangle_\omega \sim J(t) \equiv \frac{2}{\sqrt{\pi}} D \int_0^\infty d\omega^2 \varphi(\omega) \times \sqrt{\int_0^t d\tau (t-\tau)\langle e^{-W\tau}\rangle_c}.$$

Using (29) and (31), we find that the main contribution to the asymptote with respect to t is made by the region of large ω, i.e., the term in $\langle e^{-W\tau}\rangle_c$ that is associated with the four-spin process. As a result, when $Dt\to\infty$,

$$\langle\langle\sqrt{I_j(t)}\rangle_c\rangle_\omega \sim K\sqrt{Dt}\,\ln(Dt). \tag{35}$$

Here K is a constant, $K>0$, and the function approaches this asymptote at $t/\tau_c > 10^1$ due to the slow decay of the function $\langle K_{0m}(t)\rangle_c$. Before that, $\langle\langle\sqrt{I_j(t)}\rangle_c\rangle_\omega \sim Dt$. Figure 5 shows the time dependence of $J(t)$, which qualitatively describes the behavior of $-\ln G(t)$. As characteristic cases we choose

$$\langle e^{-W\tau}\rangle_c = \exp(\langle -\sqrt{\tau/\tau_{c0}}\rangle),$$

$$\langle e^{-W\tau}\rangle_c = \exp\{-[t/\tau_c(\omega)]^{1/3}\},$$

i.e., the cases of a constant characteristic time for the flip-flop transitions and characteristic slowing of the latter on the wing of the resonance line. Here $\tau_c(\omega)=T_{c0}(1+\omega^2/D^2)$, and $\tau_{c0}D=4$, as previously.

It is clear from the foregoing that at very long times the behavior of $G_1^2(t)$ and $G_2^2(t)$ is close to (23) and (27), and at moderate times the total exponential factor in these functions decreases approximately as $\exp(-UDt)$, $U\sim 1$, i.e., more rapidly than in the model with a constant fluctuation rate. Hence it follows that at short and moderate times the behavior of $G(t)$ is described well by the Anderson model and that it becomes an exponential function of a square root only in the limit of very long times. We also note here that, as is seen from (34), the presence of a variable component with a fluctuation rate $1/\tau_c \sim D/4$ in the local field is sufficient for the presence of oscillations in the signal of the pairs, and just this component determines the character of the oscillations. This is because they appear as a consequence of the configurational averaging of the motions in the spatially constrained pairs, and the medium outside each pair provides for monotonic damping of the free-induction decay signal and suppression of the contribution of the intrapair motions to it to a greater extent, the smaller is the fraction of the fluctuating component in the field induced by it. The more rapidly a spin fluctuates, the smaller is its contribution to the free-induction decay signal. Each spin in a restricted volume is spatially ''smeared,'' and its contribution to the free-induction decay signal is determined by the regions which make a minimal contribution; the contribution of the remaining regions decays rapidly. This is a specific effect of an irregular system. The farther the spins are from the point where the field from them is observed, the greater is the probability of the appearance of a spin among them that produces a slowly fluctuating field from a large part of the region where it is smeared, which determines the damping rate of $G(t)$. This also leads to a difference in the influence of the spread of the fluctuation rates in the system on the damping associated with the presence of infinite distances and on the oscillations associated with finite distances.

6. EVALUATION OF THE CONTRIBUTION OF CLUSTERS OF HIGHER RANKS TO THE CORRELATION FUNCTIONS

The performance of calculations in a higher order of cluster expansion (the separation of triads) does not alter the results, at least when there are fluctuations of the local fields in the system. In fact, comparing (5) and (20), we see that the expression for the free-induction decay function of triads has the form

$$G_3^3(t)\sim\exp\left[-\frac{2}{\sqrt{\pi}}D\sqrt{I(t)}\right]C^2\int d\mathbf{r}_1 d\mathbf{r}_2\,\cos\left(\frac{3}{4}A_{\mathbf{r}_1}t\right)$$
$$\times\cos\left(\frac{3}{4}A_{\mathbf{r}_2}t\right)\exp\left\{-C\int_{q\in V_{0\mathbf{r}_1\mathbf{r}_2}}d\mathbf{q}\right.$$
$$\left.\times\exp\left[-\frac{9}{16}A_q^2 I(t)\right]\right\}, \tag{36}$$

where $V_{0\mathbf{r}_1\mathbf{r}_2}$ is the excluded volume of a spin triad. Taking into account that

$$2C\int_{q\in V_{0\mathbf{r}_1\mathbf{r}_2}}d\mathbf{q}\,\exp\left[-\frac{9}{16}A_q^2 I(t)\right]>C\int_{q\in V_{0\mathbf{r}_1}}d\mathbf{q}$$
$$\times\exp\left[-\frac{9}{16}A_q^2 I(t)\right]+C\int_{q\in V_{0\mathbf{r}_2}}d\mathbf{q}$$
$$\times\exp\left[-\frac{9}{16}A_q^2 I(t)\right],$$

after a replacement of variables we obtain

$$G_3^3(t)<\exp\left(-\frac{2}{\sqrt{\pi}}D\sqrt{I(t)}\right)\left[C\int d\mathbf{r}\,\cos\left(\frac{3}{4}A_{\mathbf{r}_2}\frac{t}{2}\right)\right.$$
$$\left.\times\exp\left\{-C\int_{q\in V_{0\mathbf{r}}}d\mathbf{q}\,\exp\left(-\frac{9}{16}A_q^2\frac{I(t)}{4}\right)\right\}\right]^2.$$

The expression in square brackets equals the integral in the formula for G_2^2 in (22) to within a numerical coefficient with the replacements $t\to t/2$ and $\tau_c\to\tau_c/2$ at large t. However, it is seen from (25)–(27) and the ensuing equations at the end of Sec. 4 that the asymptote of this integral has the form $\exp(-\beta\sqrt{Dt})$, where $\beta>0$ depends weakly on τ_c. Thus,

$$G_3^3(t)\sim G_2^2(t)\exp[-\beta(\sqrt{2}-1)\sqrt{Dt}],$$

and in the long-time limit the contribution of the triads can be neglected. The use of the expressions from this section for $G_l^k(t)$ does not alter this conclusion. This reasoning is incorrect only in the case of the complete absence of fluctuations in the system. Then there is no damping due to averaging of the intracluster motions over the configurations and, as is seen from a comparison of (20) and (24), the order k of the cluster expansion introduces the power function $(Dt)^k$ in front of the main exponential function $\exp(-\sqrt{2/\pi}Dt)$, and all the orders as a whole can lead to renormalization of the argument of this exponential function, which is natural, if we recall that the Anderson model with $G(t)=\exp(-Dt)$ is exact in this case.

On the other hand, at small and moderate t the contribution of the triads is suppressed by their small concentration,[3] so that they can be taken into account in the shape function of the resonance line in the form of a numerical correction to the coefficient in the asymptotes of the pairs on the distant wing, but there only the depletion of the wing by clusters, which is achieved already when $k=2$, is generally significant.

7. CONCLUSIONS

In the present work we have quantitatively analyzed the main problem in the theory of the shape function and free-induction decay in a magnetically dilute spin system both on the basis of the exactly solvable Anderson model and in the approximation of a normal random process for local fields. In the limit of slow fluctuations the two approaches practically coincide. The temporal fluctuations of the spins on the wing of the absorption line are created mainly by a four-spin process, which appears in second-order perturbation theory, the corresponding correlation function of the local field decaying considerably more slowly than the functions usually employed to describe such processes in spatially regular media.[18] In the system there are very slow fluctuations, which do not make a significant contribution to the long-time asymptotes due to their small statistical weight, but lead to drastic slowing of the approach to the asymptote. As a result, the total free-induction decay function of the system at considerable times can be described by the Anderson model.

The present result is very important and nontrivial. The fact is that a Lorentzian line usually appears in the spin kinetics of spatially regular systems in the limit of short correlation times, i.e., fast fluctuations of the local fields. This assertion is valid for almost any type of interaction and any dimension of the space. In magnetically dilute systems and in the Anderson model the mechanism for the appearance of a Lorentzian line is totally different. Here, in contrast, an infinitely long correlation time and a Lorentzian line appear, only if the interaction weakens with the distance as $1/r^d$, where d is the dimension of the space. It is known that a Lorentzian line does not undergo motional narrowing, if the motion has the character of uncorrelated collisions. In reality, the temporal fluctuations of local fields are smooth, and, therefore, the faithfulness of the Anderson model for long times is remarkable.

A model which is simpler than the one used in the present work is based on the assumption that the field at each spin can be approximated by a normal random process and that only the second moment depends on the configuration.[3] The expression then obtained for the free-induction decay is

$$G(t)=\exp\left[-\sqrt{\frac{2}{\pi}}D\sqrt{2I(t)}\right]. \tag{37}$$

In this model the Lorentzian line changes shape and continues to narrow strongly as the fluctuation rate $1/\tau_c$ of the local fields increases. In the limit of slow fluctuations Eq. (37) agrees satisfactorily with the Anderson model; however, in the case of realistic values of τ_c the result of this approach at moderate times diverges from the conclusions of the cluster analysis, which are far closer to the Anderson result.

The total free-induction decay function is exponential at short and moderate times and becomes an exponential function of a square root at large times, the time when the asymptote is approached being significantly dependent on the form of the correlation function of the local fields from the slowest objects in the system. Strong slowing of the approach of the free-induction decay to the asymptote actually occurs.

The development of cluster expansion methods has made it possible to show that the regular procedure for isolating strongly interacting groups of spins from a system leads to a faithful description of many phenomena and that the second-order expansion is already sufficient, since the third-order expansion produces only a slight correction.

The existing theories of free-induction decay do not claim to describe it at moderate and long times. The treatment that we have performed predicts the character of free-induction decay at practically all times, making it possible to experimentally test the correspondence of the model constructed in the present work to real processes in dilute spin systems.

[1] A. Abragam and M. Goldman, *Nuclear Magnetism: Order and Disorder*, Clarendon Press, Oxford (1982) [Russ. transl., Mir, Moscow (1984)].
[2] R. R. Ernst, G. Bodenhausen, and A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford (1987) [Russ. transl., Mir, Moscow (1990)].
[3] F. S. Dzheparov and E. K. Khenner, Zh. Éksp. Teor. Fiz. **104**, 3667 (1993) [JETP **77**, 753 (1993)].
[4] Yu. G. Abov, A. D. Gul'ko, F. S. Dzheparov *et al.*, Fiz. Elem. Chastits At. Yadra **26**, 1654 (1995) [Phys. Part. Nuclei **26**, 692 (1995)].
[5] V. A. Atsarkin, G. A. Vasneva, and V. V. Demidov, Zh. Éksp. Teor. Fiz. **91**, 1523 (1986) [Sov. Phys. JETP **64**, 898 (1986)].
[6] É. V. Avagyan, V. A. Atsarkin, G. A. Vasneva, Zh. Éksp. Teor. Fiz. **85**, 1790 (1983) [Sov. Phys. JETP **58**, 1041 (1983)].
[7] D. A. Drabold and P. A. Fedders, Phys. Rev. B **37**, 3440 (1988).
[8] F. S. Dzheparov and É. B. Fel'dman, Izv. Akad. Nauk SSSR, Ser. Fiz. **52**, 455 (1988).
[9] P. W. Anderson, Phys. Rev. **82**, 342 (1951).
[10] F. S. Dzheparov and E. K. Khenner, Zh. Éksp. Teor. Fiz. **96**, 1844 (1989) [Sov. Phys. JETP **69**, 1042 (1989)].
[11] E. B. Feldman and S. Lacelle, J. Chem. Phys. **104**, 2000 (1996).
[12] P. W. Anderson and P. R. Weiss, Rev. Mod. Phys. **25**, 269 (1953).
[13] R. Kubo, J. Phys. Soc. Jpn. **17**, 1100 (1962).
[14] M. Tanaka, Y. Aoyama, and Y. Yamamoto, J. Phys. Soc. Jpn. **59**, 32 (1990).
[15] A. G. Demenov and E. K. Henner, in *Extended Abstracts of 27th Congress Ampere*, Kazan (1994), p. 196.
[16] F. S. Dzheparov, Zh. Éksp. Teor. Fiz. **99**, 982 (1991) [Sov. Phys. JETP **72**, 546 (1991)].
[17] F. S. Dzheparov, A. A. Lundin, and T. N. Khazanovich, Zh. Éksp. Teor. Fiz. **92**, 554 (1987) [Sov. Phys. JETP **65**, 314 (1987)].
[18] F. S. Dzheparov and S. V. Stepanov, in *Muons and Pions in Matter, Proceedings of the 3rd International Symposium on Muon and Pion Interactions with Matter*, Dubna (1995), p. 247.

Translated by P. Shelnitz

Galvanomagnetic phenomena in layered conductors

V. G. Peschanskiĭ

B. I. Verkin Low-Temperature Physicotechnical Institute, Ukrainian National Academy of Sciences, 310164 Kharkov, Ukraine

(Submitted 11 November 1996)
Zh. Éksp. Teor. Fiz. **112**, 618–627 (August 1997)

The dependence of the resistance and the Hall field in a layered conductor with a quasi-two-dimensional electron energy spectrum of arbitrary shape on the magnitude and orientation of the magnetic field in relation to the layers is analyzed. It is found that when current flows perpendicular to the layers, the resistance of the specimen strongly depends on the angle ϑ between the normal and the vector of a strong magnetic field. The Kapitza law is shown to hold within a fairly broad range of magnetic fields in the plane of the layers, i.e., the resistance increases linearly with the magnetic field strength. The Hall field proves to be insensitive to the emergence of open sections of the Fermi surface, and the Hall constant in strong magnetic fields is the same for any orientation of the magnetic field and the current. © *1997 American Institute of Physics.* [S1063-7761(97)01608-9]

The search for new superconducting materials is to a great extent responsible for the interest in low-dimensional conductors. Before the discovery of metal-oxide superconductors attention was focused on superconductors of organic origin characterized by a layered or filament-like structure with a significant anisotropy in electrical conductivity in the normal (nonsuperconducting) state. Undoubtedly, the behavior of such superconductors in a magnetic field is of interest. In contrast to metals, here a magnetic field may have no effect on low-dimensional conductors and galvanomagnetic effects inherent in common metals may be enhanced.

In 1928, Kapitza[1] discovered a remarkable phenomenon, the linear increase with magnetic field strength of the resistance of metals at liquid-air and liquid-carbon-monoxide temperatures. To detect this effect Kapitza had to build magnets in which the magnetic field was as high as 30–50 T. At the same time, the University of Leiden had the facilities to study the behavior of substances at even lower temperatures, which made it possible to enhance the effectiveness of weaker fields due to the increase in the mean free path of the charge carriers. Shubnikov and de Haas discovered, however, that at liquid hydrogen temperatures the resistance of bismuth single crystals (the most perfect substance at that time) exhibited a much more complicated dependence (compared to the linear increase with magnetic field) of the resistance on the magnetic field strength,[2] while at liquid helium temperatures the resistance revealed a strong oscillatory dependence on the reciprocal magnetic field strength (the Shubnikov–de Haas effect).[3] The discovered low-temperature oscillatory effect was found to be common to all metals, but most vividly the Shubnikov–de Haas effect manifested itself in layered conductors of the type of tetrathiafulvalene salts and tetraselentetracene halogens.[4–13] It would be interesting to see to what extent these layered conductors are more convenient for studying the Kapitza effect.

In the 1930s the linear increase in the resistance of metals with the magnetic field strength discovered by Kapitza did not agree with the main principles of the electron theory of metals since, according to the Onsager principle of the symmetry of the kinetic coefficients,[14] the resistance of conductors must be an even function of the magnetic field strength. The first attempt to explain the results of Kapitza's experiments was made only in 1958 (Ref. 15). It was found that the fact that the Onsager principle does not contradict the linear dependence of the resistance of metals on $\mathbf{H}$ is due to the complicated dependence of the energy ε on the quasimomentum $\mathbf{p}$ of the charge carriers. The main characteristic of the electron energy spectrum, the Fermi surface $\varepsilon(\mathbf{p})=\varepsilon_F$, is open for almost all metals, so that the orbits of the conduction electrons with the Fermi energy ε_F in a magnetic field for which $\mathbf{p}\cdot\mathbf{H}=\text{const}$ pass through many unit cells of the momentum space. The period of revolution of the conduction electrons along such highly elongated orbits, $T=2\pi/\Omega$, can be much shorter than the mean-free-path time τ, no matter how strong the magnetic field. As a result, averaging the resistance of a polycrystalline wire over the different orientations of the crystals and hence over the various electron orbits leads to a linear dependence of the wire's resistance on the magnitude of a strong magnetic field in which $\Omega_0\tau\gg 1$, where Ω_0 is the maximum frequency of revolution of a Fermi electron in the magnetic field.[15,16] If the thickness of a polycrystalline specimen of a metal with an open Fermi surface is much larger than the crystallite sizes, the specimen's resistance in a strong magnetic field is proportional to $H^{4/3}$ (Refs. 17 and 18), i.e., the ρ vs. H dependence is close to linear. For a single-crystal specimen, the spectrum of all possible frequencies of revolution of the Fermi conduction electrons also extends from zero to Ω_0, provided that the Fermi surface has saddle points. In this case there may be self-intersecting orbits, traveling along which an electron cannot complete a revolution. However, the fraction of the electron orbits close to the self-intersecting one, for which the period T is longer than the mean-free-path time, is proportional to $\exp(-\Omega_0\tau)$, since the period, regarded as function of $p_H=\mathbf{p}\cdot\mathbf{H}/H$, becomes logarithmically divergent as a self-intersecting orbit is approached. As a result, in a very narrow range of magnetic field strengths,[19] the complicated dependence of ρ on H is

rapidly replaced, as the field gets stronger, by a quadratic dependence or even by saturation of the resistance.

In quasi-two-dimensional conductors, the period of revolution of the charge carriers in a magnetic field is weakly dependent on the momentum projection p_H, which is the reason for the large increase in the magnetoresistance oscillation amplitude. In contrast to metals, where only a small fraction of the charge carriers with a cross-sectional area of the Fermi surface close to the extremal value participate in the formation of the Shubnikov–de Haas effect,[20–22] in quasi-two-dimensional conductors almost all the charge carriers with the Fermi energy contribute to the quantum oscillatory effects. There is every reason to believe that in such conductors the number of conduction electrons near a self-intersecting orbit, for which $T>\tau$, is much higher than in ordinary metals. These electrons probably provide the leading contribution to conductivity in a much broader range of strong magnetic fields ($\Omega_0\tau\gg 1$), and averaging over the various frequencies of revolution of the charge carriers leads to a result that differs strongly from the one obtained for metals.

Most organic superconductors, which are layered structures, appear to have a quasi-two-dimensional electron energy spectrum, with the electrical conductivity along layers being much higher than the electrical conductivity along the normal **n** to the layers. Many layered conductors, including the tetraselentetracene halogens and tetrathiafulvalene salts mentioned earlier, exhibit metallic conductivity even across the layers. This justifies, in describing the electron process in such conductors, the use of the concept of quasiparticles that carry the electron charge e and are similar to conduction electrons in metals. The energy of the charge carriers in such conductors,

$$\varepsilon(\mathbf{p})=\sum_{n=0}^{\infty}\varepsilon_n(p_y,p_z)\cos\frac{anp_x}{\hbar}, \tag{1}$$

is weakly dependent on the momentum projection $p_x=\mathbf{p}\cdot\mathbf{n}$, and the Fermi surface is a mildly corrugated cylinder with, possibly, small closed cavities belonging to anomalously small groups of charge carriers. Here a is the layer separation, $\hbar$ is Planck's constant, the maximum value of the function $\varepsilon_1(p_y,p_z)$ on the Fermi surface is $\eta\varepsilon_F\ll\varepsilon_F$, and the maximum values of $\varepsilon_n(p_y,p_z)$ with $n\geqslant 2$ are even smaller. This property is characteristic of the dispersion law of the charge carriers in the tight-binding approximation, when the overlap of the wave functions of the electrons belonging to different layers is negligible.

We will now analyze the dependence of the magnetoresistance and the Hall field on the magnitude and orientation of the magnetic field with very general assumptions concerning the shape of the quasi-two-dimensional electron energy spectrum of the layered conductor.

We find the relationship between the current density and the electric field,

$$j_i=\sigma_{ik}E_k \tag{2}$$

by solving the kinetic equation for the charge-carrier distribution function

$$f(\mathbf{p})=f_0(\varepsilon)-eE_i\psi_i(\mathbf{p})\frac{\partial f_0(\varepsilon)}{\partial\varepsilon}.$$

In an approximation linear in the weak electric field E the equation is

$$\frac{e}{c}\mathbf{v}\times\mathbf{H}\,\frac{\partial\psi_i}{\partial p}-\hat{W}_{\mathrm{col}}\{\psi_i\}=v_i, \tag{3}$$

where $\mathbf{v}=\partial\varepsilon/\partial\mathbf{p}$ and $f_0(\varepsilon)$ are the velocity and equilibrium Fermi distribution function of the conduction electrons, and c is the speed of light in vacuum. In the τ-approximation for the collision integral $\hat{W}_{\mathrm{col}}$, in which $\hat{W}_{\mathrm{col}}\{\psi_i\}=-\psi_i/\tau$, the solution of Eq. (3) has the form

$$\psi_i(t,p_H,\varepsilon)=\int_{-\infty}^{t}dt'\,v_i(t',p_H,\varepsilon)\,\exp\frac{t'-t}{\tau}. \tag{4}$$

For the variables in the momentum space we have taken the constants of motion ε and p_H and t, the time of motion of a charge in the magnetic field $\mathbf{H}=(H\sin\vartheta,0,H\cos\vartheta)$, according to the equations

$$\begin{aligned}\frac{\partial p_x}{\partial t}&=\frac{ev_y}{c}H\cos\vartheta,\\ \frac{\partial p_y}{\partial t}&=(v_z\sin\vartheta-v_x\cos\vartheta)\frac{eH}{c},\\ \frac{\partial p_z}{\partial t}&=-\frac{ev_y}{c}H\sin\vartheta.\end{aligned} \tag{5}$$

By using the solution (4) of the kinetic equation we can easily find the components of the electrical conductivity tensor:

$$\sigma_{ik}=-\frac{2e^3H}{c(2\pi\hbar)^3}\int d\varepsilon\,\delta(\varepsilon-\varepsilon_F)\times\int dp_H\int_0^T dt\,v_i(t)\int_{-\infty}^{t}dt'v_k(t')\exp\frac{t'-t}{\tau}=\langle v_i\psi_k\rangle. \tag{6}$$

Next we assume that the Fermi surface of the layered conductor with a quasi-two-dimensional electron energy spectrum is a mildly corrugated cylinder that is ''open'' along the p_x axis. For ϑ finite, all sections of this cylinder by the plane $\mathbf{p}\cdot\mathbf{H}=\mathrm{const}$ are closed trajectories and are almost indistinguishable when $\eta\ll 1$. Only at $\vartheta=0$ do open trajectories appear in the momentum space, and along these trajectories the charge carriers move with a period

$$T=\frac{2\pi m^*c}{eH}=\frac{c}{eH}\int_0^{2\pi\hbar/a}\frac{dp_x}{v_y}. \tag{7}$$

The average value of the velocity v_y over a period is finite:

$$\overline{v_y}=\frac{1}{T}\int_0^T dt\,v_y(t)=\frac{2\pi\hbar ec}{aHT}=\frac{\hbar}{am^*}, \tag{8}$$

with the various directions of drift of the charge carriers fanning out over the entire yz plane. Here the components σ_{yy} and σ_{zz} of the electrical conductivity tensor coincide in

order of magnitude with the electrical conductivity along the layers in the absence of a magnetic field. However, the components σ_{ij} in which at least one index coincides with x are negligible when $\eta \ll 1$, while in a strong magnetic field they decrease as H grows, since $\overline{v_x}=0$.

On the boundary section of the Fermi surface, $p_H=p_s$ (here p_s is the momentum projection along the magnetic field at the saddle point), which separates the open electron trajectories and a small region of closed orbits, the period becomes infinite and the charge carriers moving along orbits close to this section contribute considerably to the components σ_{xj} and σ_{ix}, which we calculate below.

The period of electron motion along orbits with p_H close to p_s is extremely long, since most of the time the electrons are near the saddle point $\mathbf{p}_s=(0,0,p_s)$ of the constant-energy surface, where $v_x=v_y=0$. In the immediate vicinity of the section $p_H=p_s$ the projection v_y of the electron velocity is a complicated function of t, but far from this section we can ignore the small t-dependent corrections in parameter η in the expression for $v_y(t)$. This yields the following expression for the period of the charge-carrier motion:

$$T(p_H)=\frac{2\pi\hbar c}{aeHv_y(0)}=\frac{2\pi v_F}{\Omega_0 v_y(0)}, \tag{9}$$

which for $v_y \ll v_F=\varepsilon_F a/\hbar$ may become comparable to the mean-free-path time. As a result, the contribution of the charge carriers with a small velocity projection v_y to σ_{xx} is the leading one. If we ignore the small t-dependent corrections in the expression for v_y, then p_x becomes a linear function of the time of the motion of the charge in the magnetic field, and the charge's velocity

$$v_x=-\sum_{n=1}^{\infty}\frac{an}{\hbar}\varepsilon_n(p_y,p_z)\sin(n\Omega t) \tag{10}$$

is, in the leading approximation in parameter η, a harmonic function of t, i.e., is determined primarily by the first term in (10). We can now easily calculate the component σ_{xx} of the conductivity tensor in this approximation:

$$\sigma_{xx}=\frac{2e^2\tau}{(2\pi\hbar)^3}\int 2\pi m^* dp_z \sum_{n=1}^{\infty}\frac{\{\varepsilon_n(p_y,p_z)an/\hbar\}^2}{1+(n\Omega\tau)^2}. \tag{11}$$

In the immediate vicinity of the section $p_H=p_s$ we must refine the numerators in (11) by replacing them with $|v_x^n|^2$. Since p_x for these charge carriers is a complicated function of t, the Fourier transforms of their velocity v_x^n do not necessarily decrease as the number n grows, and the contribution to the asymptotic behavior of σ_{xx} of the electron from a small neighborhood of the saddle point of the Fermi surface is not limited only to the first harmonics in the expansion of $v_x(t)$ in a Fourier series. The maximum value of the velocity of electron motion along the y axis along the boundary section $p_z=p_s$ is equal in order of magnitude to $\eta^{1/2}v_F$, and the t-dependence can be ignored only when $v_y(0) \gg \eta^{1/2}v_F$.

We write σ_{xx} in the form $\sigma_{xx}=\sigma_{xx}^{(1)}+\sigma_{xx}^{(2)}+\sigma_{xx}^{(3)}$, where $\sigma_{xx}^{(1)}$ allows for the contribution of charge carriers for which velocity v_x has the form (10), the second term allows for the contribution to σ_{xx} of the conduction electrons with open orbits, with $v_y \leqslant \eta^{1/2}v_F$, and the last term

$$\sigma_{xx}^{(3)}=\frac{4e^2\tau}{(2\pi\hbar)^3}\int_{p_s}^{p_0}dp_z 2\pi m^*(p_z)\sum_{n=1}^{\infty}\frac{|v_x^n|^2}{1+(n\Omega\tau)^2} \tag{12}$$

allows for the contribution to σ_{xx} of charge carriers with closed orbits. Here p_0 is the maximum value of p_z at the reference point $\mathbf{p}_0=(\pi\hbar/a,0,p_0)$ of the Fermi surface.

For $\gamma_0=1/\Omega_0\tau > \eta^{1/2}$, integration with respect to p_z over the small interval where $\eta^{1/2}v_F < v_y \ll v_F$ yields

$$\sigma_{xx}^{(1)}=\sigma_0\eta^2\gamma_0, \tag{13}$$

where σ_0 coincides in order of magnitude with the electrical conductivity along the layers in the absence of a magnetic field.

Near the reference point of the Fermi surface the cyclotron effective mass of the charge carriers, m^*, is proportional to $\eta^{-1/2}$, and increases as we get closer to the section $p_H=p_s$, on which it becomes infinite. When $\eta^{1/2} \ll \gamma_0$, the conduction electrons belonging to the closed sections of the Fermi surface do not have enough time to complete a full revolution along their orbits, and to high accuracy their contribution to σ_{xx} can be written as follows:

$$\sigma_{xx}^{(3)}=\frac{4e^2\tau}{(2\pi\hbar)^3}\int 2\pi m^*(p_z)dp_z\overline{v_x^2}=\sigma_0\eta^{5/2}. \tag{14}$$

Here and in what follows we ignore unimportant numerical factors of order unity in the formulas for σ_{xx}.

Thus, for $\eta^{1/2} \ll \gamma_0 \ll 1$, the small fraction of conduction electrons with open orbits, which slowly move along the y axis, provide the leading contribution to σ_{xx}, a contribution much larger than that of all the other charge carriers on the Fermi surface.

As the magnetic field gets stronger, the number of conduction electrons with $T>\tau$ decreases, and the contribution to σ_{xx} of charge carriers with orbits extremely close to the section $p_H=p_s$ becomes significant. We can easily find the dependence on p_H of the period of motion of charge carriers near a self-intersecting orbit by expanding the energy in powers of p_y. Keeping only the first two terms in Eq. (1), we obtain

$$\varepsilon(\mathbf{p})=\varepsilon_0(0,p_z)+\frac{p_y^2}{2m_1}+\varepsilon_1(0,p_z)\cos\frac{ap_x}{\hbar}. \tag{15}$$

If we now use Eq. (7), we arrive at an expression for $T(p_H)$:

$$T(p_z)=\frac{\hbar c}{aeH}\left\{\frac{m_1}{\varepsilon_1(0,p_z)}\right\}^{1/2}\int_0^{\pi}d\alpha(\xi^2+\sin^2\alpha)^{-1/2}, \tag{16}$$

where $\xi^2=\{\varepsilon_0(0,p_s)-\varepsilon_0(0,p_z)+\varepsilon_1(0,p_s)-\varepsilon_1(0,p_z)\}\times[2\varepsilon_1(0,p_z)]^{-1}$.

For $\xi \ll 1$, the period of the charge-carrier motion,

$$T(p_z)\simeq\Omega_0^{-1}\eta^{-1/2}\ln\frac{1}{\xi}, \tag{17}$$

diverges logarithmically, and the contribution to σ_{xx} of the electron from a small neighborhood near the boundary section of order $\Delta p_z = p_s - p_z \simeq p_s \eta$ has the form

$$\sigma_{xx}^{(2)} = \sigma_0 \eta^{5/2} \gamma_0^2 \int_1^\infty du \frac{u^3 \exp\{-u\}}{u^2 \gamma_0^2 + \eta} \simeq \sigma_0 \frac{\gamma_0^2 \eta^{5/2}}{\gamma_0^2 + \eta}. \quad (18)$$

While for $\eta^{1/2} \ll \gamma_0$ the contribution to σ_{xx} of the charge carriers belonging to a small region of order $p_0 \eta$ in the vicinity of a self-intersecting section of the Fermi surface is negligible and $\sigma_{xx} \simeq \sigma_{xx}^{(1)}$, in the opposite limiting case the contribution of these charge carriers is of the same order of magnitude as that of all the other conduction electrons. It can easily be demonstrated that $\sigma_{xx}^{(1)}$ for $\eta > \gamma_0^2$ also decreases with increasing magnetic field strength in proportion to γ_0^2. As a result, when the magnetic field is strong, i.e., $\gamma_0 \leqslant \eta^{1/2}$, we have

$$\sigma_{xx} \simeq \sigma_0 \eta^{3/2} \gamma_0^2. \quad (19)$$

The behavior of the Hall field in layered conductors also differs from that in metals. To make the results of the present investigation more graphic, we examine the galvanomagnetic effects in a conductor with a simple model dispersion law for the charge carriers:

$$\varepsilon(\mathbf{p}) = \frac{p_y^2 + p_z^2}{2m} - \eta \frac{v_F \hbar}{a} \cos \frac{a p_x}{\hbar}, \quad (20)$$

which uses the approximation of loosely bound (almost free) charge carriers in the planes of the layers. Since the leading contribution to the electrical conductivity across the conductor layers is provided by the charge carriers with small v_y and in calculating the galvanomagnetic characteristics the dependence of $\varepsilon_n(p_y, p_z)$ on p_z at $\vartheta = 0$ can be ignored, the analysis of galvanomagnetic phenomena that follows is of a fairly general nature.

Using the equations of motion of a charge in a magnetic field (Eqs. (5)) at $\vartheta = 0$ and the dispersion law (20), we obtain

$$\psi_x = \gamma(\psi_y - v_y \tau), \quad (21)$$

where $\gamma = mc/eH\tau$. This leads to

$$\sigma_{xx} = \gamma \sigma_{xy}, \quad \sigma_{yx} = \gamma(\sigma_{yy} - \sigma_0), \quad \sigma_{xy} = -\sigma_{yx}, \quad (22)$$

and the matrix form of the tensor σ_{ij} becomes

$$\sigma_{ij} = \begin{pmatrix} \sigma_{xx} & \gamma^{-1}\sigma_{xx} & 0 \\ -\gamma^{-1}\sigma_{xx} & \sigma_0 - \gamma^{-2}\sigma_{xx} & 0 \\ 0 & 0 & \sigma_0 \end{pmatrix}, \quad (23)$$

while for the resistance tensor (which is the inverse of (23)) we have

$$\rho_{ij} = \begin{pmatrix} \sigma_{xx}^{-1} - \sigma_0^{-1}\gamma^{-2} & -(\gamma\sigma_0)^{-1} & 0 \\ -\gamma^{-1}\sigma_0^{-1} & \sigma_0^{-1} & 0 \\ 0 & 0 & \sigma_0^{-1} \end{pmatrix}. \quad (24)$$

Clearly, to a high degree of accuracy we can assume that the resistivity of the conductor perpendicular to the layers, ρ_{xx}, is approximately $1/\sigma_{xx}$ and increases linearly with the magnetic field strength when $\eta^{1/2} \ll \gamma \ll 1$. The Hall field

$$\mathbf{E}_{\text{Hall}} = R\mathbf{j} \times \mathbf{H} \quad (25)$$

is also proportional to H, while the Hall constant R is inversely proportional to the entire volume inside the Fermi surface. In metals the Hall constant behaves similarly only for open sections of the Fermi surface.

The absence of magnetoresistance $\Delta\rho = \rho(H) - \rho(0)$ when the current flows along layers is due to the quadratic dispersion of the charge carriers in the yz plane. When the dependence of the carrier energy on p_y and p_z is more complicated, the resistance increases with magnetic field strength and reaches a saturation plateau, just as it does in ordinary metals. However, in contrast to metals, $\Delta\rho$ in quasi-two-dimensional conductors is extremely low and vanishes at $\eta = 0$. The reason is that only the projection of the magnetic field on the normal to the layers has a strong effect on the dynamics of the charge carriers, and this projection is nil at $\vartheta = 0$.

For ϑ finite but smaller than η, the electrical conductivity along the layers in a magnetic field differs little from the conductivity in the absence of a magnetic field, and as long as $\gamma_0 \geqslant \eta$ the formulas (13) and (19) for σ_{xx} remain valid. However, for $\vartheta \gg \eta$ there are no self-intersecting orbits, and the resistance of the specimen perpendicular to the layers becomes saturated in realistic magnetic fields.

If we use the expression (20) for the dispersion law of the charge carriers, the components of the resistance tensor are

$$\rho_{ij} = \begin{pmatrix} \frac{1}{\sigma_{xx}} - \frac{\cos^2\vartheta}{\sigma_0(\gamma_0^2 + \sin^2\vartheta)} & \frac{H\cos\vartheta}{Nec} & 0 \\ -\frac{H\cos\vartheta}{Nec} & \frac{1}{\sigma_0} & \frac{H\sin\vartheta}{Nec} \\ 0 & -\frac{H\sin\vartheta}{Nec} & \frac{1}{\sigma_0} \end{pmatrix}, \quad (26)$$

where N is the charge carrier number density. The above matrix ρ_{ij} is valid for all values of $\gamma_0 = mc/eH\tau$ and all finite values of ϑ. Here the Hall constant is equal to $1/Nec$ for all orientations of the magnetic field and current density with respect to the conductor layers. This unique feature of the Hall effect in quasi-two-dimensional conductors with an arbitrary dispersion law for the charge carrier holds only in strong magnetic fields, where the frequency of revolution of the electrons along closed orbits is much lower than their collision rates.

The conductor's resistance to a current flowing in the direction perpendicular to the layers is determined primarily by the inverse of the component σ_{xx} of the conductivity tensor. The weak dependence of p_H on the energy and velocity of the charge carriers leads to a situation in which σ_{xx} becomes highly sensitive to the orientation of the strong magnetic field. For $\vartheta \gg \gamma_0$ the expansion of σ_{xx} in a power series in η begins at least with quadratic terms. However, for some values of angle ϑ the asymptotic behavior for small η of the electrical conductivity across the layers,

$$\sigma_{xx}(\eta,H)$$
$$= \frac{2e^2H}{c(2\pi\hbar)^3}\int_0^{2\pi\hbar\sin\vartheta/a} dp_H\left(1-\exp\left\{-\frac{T}{\tau}\right\}^{-1}\right)$$
$$\times\int_0^T dt\int_{t-T}^{t} dt'\sum_{n,m}\varepsilon_n(t,p_H)\varepsilon_m(t',p_H)$$
$$\times\sin\left\{\frac{an}{\hbar}\left(\frac{p_H}{\sin\vartheta}-p_z(t,p_H)\cot\vartheta\right)\right\}$$
$$\times\sin\left\{\frac{am}{\hbar}\left(\frac{p_H}{\sin\vartheta}-p_z(t',p_H)\cot\vartheta\right)\right\}$$
$$\times\exp\frac{t'-t}{\tau}, \tag{27}$$

can vary considerably.

For $\vartheta\gg\eta$ the projections $p_i(t,p_H)=p_i(t)+\Delta p_i(t,p_H)$ of the electron momentum are weakly dependent on p_H, and in calculating the asymptotic expression for σ_{xx} we can drop the Δp_i in (27). As a result, for small η/ϑ and $\gamma_0\ll\vartheta$ we have

$$\sigma_{xx}=\frac{e^2\tau am^*(\vartheta)\sin\vartheta}{8\pi^3\hbar^4}\sum_{n=1}^{\infty} n^2|I_n(\vartheta)|^2$$
$$+\sigma_0\eta^2\left\{\eta^2 f_1(\vartheta)+\left(\frac{\gamma_0}{\sin\vartheta}\right)^2 f_2(\vartheta)\right\}, \tag{28}$$

where

$$I_n(\vartheta)=\frac{1}{T}\int_0^T dt\ \varepsilon_n(t)\exp\left\{\frac{ian}{\hbar}p_z(t)\cot\vartheta\right\}, \tag{29}$$

and $f_1(\vartheta)$ and $f_2(\vartheta)$ are functions of order unity that depend on the specific form of the dispersion law for the charge carriers. These two functions become important only at values $\vartheta=\vartheta_c$ at which $I_1(\vartheta)$, the leading term in the sum over n, vanishes.

Clearly, $I_n(\vartheta)$ has a large number of zeros. When $\cot\vartheta\gg1$, the leading contribution to $I_n(\vartheta)$ is provided by the integral in (29) over the small section of the electron orbit where v_y is small. Using the stationary-phase method, we get

$$I_n(\vartheta)=\frac{\varepsilon_n(t_1)}{T}\left|\frac{2\pi\hbar c}{anv_y'(t_1)\cot\vartheta}\right|^{1/2}\cos\left\{\frac{anD_p\cot\vartheta}{2\hbar}-\frac{\pi}{4}\right\}. \tag{30}$$

Here D_p is the diameter of the Fermi surface along the p_z axis, and the prime denotes differentiation with respect to t at the stationary phase point, where $v_y(t_1)=0$.

If the terms in the sum over n in (28) decrease rapidly enough with increasing n that the $I_n(\vartheta)$ with $n\geqslant2$ are smaller than $v_F\eta\gamma_0/\sin\vartheta$, then at $\vartheta=\vartheta_c$ and $\eta<\gamma_0/\sin\vartheta\ll1$ the resistance perpendicular to the layers increases quadratically with the magnetic field strength, and reaches the saturation plateau at a value of order $\sigma_0^{-1}\eta^{-4}$ only when the magnetic fields are stronger, whereupon $\gamma_0\ll\eta\sin\vartheta$.

The use of the τ-approximation in the collision integral proves to be quite sufficient for studying the galvanomagnetic characteristics of layered conductors of organic origin that are tetrathiafulvalene (TTF) salts, in which the range of strong magnetic fields ($\gamma_0\ll1$) is attainable.

The fact that the specific angular dependence of the magnetoresistance of such conductors was observed in experiments[4,5,10] convincingly supports the existence of the discussed orientational effect: significant variation in the asymptotic behavior of the magnetoresistance for certain orientations of the strong magnetic field in relation to the layers.

In layered high-T_c oxide-cuprate superconductors, the mean free paths of the charge carriers in the normal (nonsuperconducting) state are moderate and achieving the conditions for a strong magnetic field is difficult. In a weak magnetic field the magnetoresistance strongly depends on the mechanisms of relaxation of the charge carriers. To interpret the anomalous behavior of resistance observed in some high-T_c bismuth superconductors (the nonmonotonic dependence of the resistance on the temperature and the negative magnetoresistance across the layers in a longitudinal magnetic field), in the collision we must consistently take into account the relaxation processes that occur in the electron system. By allowing for the fluctuation mechanism of electrical conductivity Dorin *et al.*[23] and Balestrino *et al.*[24] gave a qualitative explanation of these anomalies.

Nevertheless, in purer high-T_c superconductors, one expects that at low temperature and for $\eta^{1/2}\lesssim\gamma_0\ll1$ the magnetoresistance perpendicular to the layers increases linearly with the transverse magnetic field, irrespective of the dependence of the collision integral on the temperature and magnetic field strength, since the asymptotic expression for the magnetoresistance across the layers in this range of magnetic fields is independent (just as the Hall field is) of the mean-free-path time of the charge carriers.

[1] P. Kapitza, Proc. R. Soc. London, Ser. A **129**, 358 (1928).
[2] L. V. Shubnikov and W. J. de Haas, Leiden Commun. **19**, 207f (1930).
[3] W. J. de Haas, J. W. Blom, and L. V. Shubnikov, Physica (Utrecht) **2**, 907 (1930).
[4] M. V. Kartsovnik, V. N. Laukhin, V. I. Nizhankovskiĭ, and A. A. Ignat'ev, JETP Lett. **47**, 363 (1988).
[5] M. V. Kartsovnik, P. A. Kononovich, V. N. Laukhin, and I. F. Shchegolev, JETP Lett. **48**, 541 (1988).
[6] I. D. Parker, D. D. Pigram, R. H. Friend, M. Kurmo, and P. Day, Synth. Met. **27**, A387 (1988).
[7] K. Oshima, T. Mori, H. Inokuchi, H. Urayama, H. Yamochi, and G. Sato, Phys. Rev. B **38**, 938 (1988).
[8] N. Toyota, T. Sasaki, K. Murata, Y. Honda, M. Tokumoto, H. Bando, N. Kinoshita, H. Anzai, T. Ishiguro, and Y. Muto, J. Phys. Soc. Jpn. **57**, 2616 (1988).
[9] W. Kang, G. Montambaux, J. R. Cooper, D. Jerome, P. Batail, and C. Lenoir, Phys. Rev. Lett. **62**, 2559 (1989).
[10] M. V. Kartsovnik, P. A. Kononovich, V. N. Laukhin, S. I. Pesotskiĭ, and I. F. Shchegolev, Zh. Éksp. Teor. Fiz. **97**, 1305 (1990) [Sov. Phys. JETP **70**, 735 (1990)].
[11] I. F. Shchegolev, P. A. Kononovich, V. M. Kartsovnic, V. N. Laukhin, S. S. Pesotskii, B. Hilti, and C. W. Mayer, Synth. Met. **39**, 357 (1990).
[12] M. Tokumoto, A. G. Swanson, J. S. Brooks, C. C. Agosta, S. T. Hannahs, N. Kinoshita, H. Anzai, and J. R. Anderson, J. Phys. Soc. Jpn. **59**, 2324 (1990).
[13] R. Yagi, Y. Iye, T. Osada, and S. Kagoshima, J. Phys. Soc. Jpn. **59**, 3069 (1990).
[14] L. Onsager, Phys. Rev. **37**, 405 (1931).

[15] I. M. Lifshits and V. G. Peschanskiĭ, Zh. Éksp. Teor. Fiz. **33**, 1251 (1957) [Sov. Phys. JETP **6**, 962 (1958)].
[16] J. Ziman, Philos. Mag. **3**, 1117 (1959).
[17] H. Stachowiak, Acta Phys. Pol. **26**, 217 (1964).
[18] Yu. A. Dreĭzin and A. M. Dykhne, JETP Lett. **14**, 66 (1971).
[19] V. G. Peschansky, J. A. Roldan Lopez, and Toji Gnado Jao, J. Phys. I **1**, 1469 (1991).
[20] I. M. Lifshits and A. M. Kosevich, Zh. Éksp. Teor. Fiz. **29**, 730 (1955) [Sov. Phys. JETP **2**, 636 (1956)].
[21] M. Ya. Azbel', Zh. Éksp. Teor. Fiz. **39**, 878 (1960) [Sov. Phys. JETP **12**, 608 (1961)].
[22] L. Onsager, Philos. Mag. **43**, 1006 (1952).
[23] V. V. Dorin, R. A. Klemm, A. A. Varlamov, A. I. Buzdin, and D. V. Livanov, Phys. Rev. B **48**, 12 951 (1993).
[24] G. Balestrino, E. Meloni, and A. A. Varlamov, JETP Lett. **61**, 833 (1995).

Translated by Eugene Yankovsky

Frequency mixing in a bistable system in the presence of noise

A. N. Grigorenko, P. I. Nikitin, and G. V. Roshchepkin

Institute of General Physics, Russian Academy of Sciences, 117942 Moscow, Russia
(Submitted 15 November 1996)
Zh. Éksp. Teor. Fiz. **112**, 628–642 (August 1997)

The dynamics of a bistable system resulting from the action of several periodic perturbations and noise is studied. The response of the system at combination harmonics and the changes in the response caused by high-frequency modulation and a constant field that removes the degeneracy in the bistable system are calculated. The mixing of harmonics is investigated experimentally in an iron garnet thin film. It is shown that the higher harmonics exhibit behavior characteristic of stochastic resonance and that the form and symmetry of the dependence of harmonics on the constant field removing the degeneracy can be used to determine the characteristics of a bistable system with respect to its response to external periodic modulation. © *1997 American Institute of Physics.* [S1063-7761(97)01708-3]

Stochastic resonance, a noise-induced increase in the response of a bistable system to periodic modulation, has attracted a great deal of attention on the part of researchers.[1–10] Their interest is attributed to the fact that under stochastic resonance noise increases not only the response of the system, but also the signal-to-noise ratio.

Let us recall the main features of stochastic resonance. A system in a bistable potential with two degenerate minima is usually considered. In the presence of a small harmonic force the response of the system (the total spectral density of the generalized coordinate of the system in a certain frequency range measured at the frequency of the external force) is determined by the local ''stiffness'' of the potential minimum in which the system is found. When there is noise, i.e., a random force that causes the system to undergo transitions between the minima, the response of the system is determined by a certain global ''effective'' stiffness of the potential, which can be much smaller than the local stiffness of the minimum. Thus, the addition of noise to a bistable system causes an increase in the response of the system to harmonic modulation and can lead to an increase in the signal-to-noise ratio.

Such behavior is not described by the extensively studied linear-response theory. In fact, the total force acting on a system under the conditions of stochastic resonance consists of a small periodic force and a fairly large random force and is, therefore, not small. At the same time, after averaging over the noise component, a linear function of the modulation force can be isolated in the response, and the nonlinear system can be replaced by a linear system with parameters that depend on the noise level. If the noise is created by thermal fluctuations and the system is in thermodynamic equilibrium, fluctuation-dissipation theory and linear-response theory can be used.[9]

Let us consider a system whose response to a static force f is given by the formula $X=X(f)$ as the simplest example of such a replacement. We assume that the force acting on the system consists of a small periodic force $f_{\sim}$ and a fairly large random force $f_{\rm ran}$ and that the frequencies are such that the total response is given, as before, by the formula $X=X(f_{\sim}+f_{\rm ran})$. Then

$$X=\sum_{n,m}\frac{X^{(m)}}{m!}C_m^n f_{\sim}^n f_{\rm ran}^{m-n}.$$

Averaging and isolating the part that is linear with respect to the modulation force, we obtain

$$X_{\sim}=\sum_m \frac{X^{(2m+1)}}{(2m)!}\langle f_{\rm ran}^{2m}\rangle f_{\sim}=k f_{\sim},$$

where the constant k depends on the noise parameters. When the noise signal is weak, the response is determined by the differential $X_{\sim}\approx X' f_{\sim}$ and is local. When the noise force is large, the response is determined by the global characteristic functions $X(f)$ and the features of the potential. The main difference between stochastic resonance and the elementary case lies in the need to average the nonlinear equations of motion.[3,5,7,8] Nevertheless, stochastic resonance is one of the few nonlinear problems in which it is possible to obtain fairly accurate analytical results. We cite Kapitsa's pendulum[11] as a fundamental example of such problems.

Stochastic resonance has been investigated experimentally in various bistable systems: lasers,[4] electrical circuits,[3,6] a uniaxial ferromagnet,[12] etc. A resonance theory,[5,7–9] which includes such concepts as the resonance phase,[13] the Q factor,[14] and the time of the first transition,[15] has been developed. However, the attention of those investigators was focused on harmonic modulation, and questions concerning the mixing of frequencies in a bistable system in the presence of noise were addressed only comparatively recently.[16–19] Continuing the investigation begun in Ref. 14, in this paper we shall thoroughly examine the mixing of harmonics in a bistable potential. The interest in this question is due to the possibility of obtaining a stochastic resonance curve with a high Q factor (i.e., a stochastic resonance peak of high amplitude). Actually, in the case of stochastic resonance the minimum of the signal-to-noise ratio is determined by the response associated with the dynamics of the system within one potential well. At the same time, there is no mixing of harmonics in a parabolic potential; therefore, the smaller it is, the lower is the noise level in the system. Special attention will be focused on the asymmetric case, in which the potential minima are different. The layout of the paper is as fol-

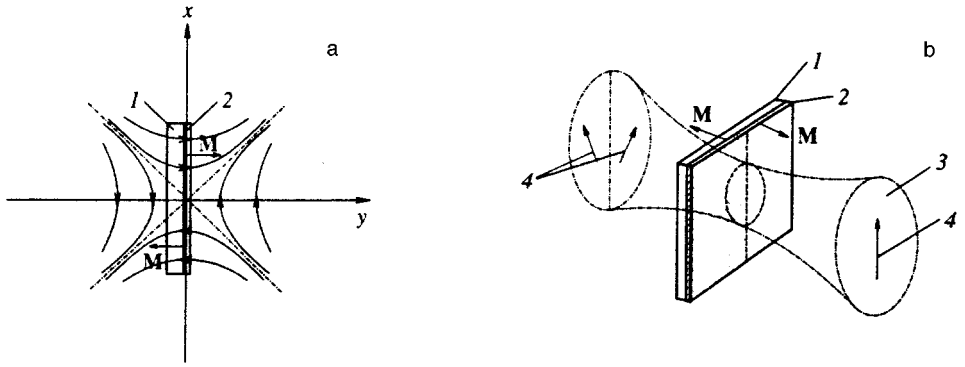


FIG. 1. Diagrams of the position of the sample in a quadrupole magnetic field (a) and the focusing of laser radiation on the sample (b): *1* — substrate, *2* — iron garnet film, *3* — polarized laser radiation beam, *4* — light polarization directions.

lows. In Sec. 2 we describe the experimental setup that was used to study the mixing of harmonics in a ferromagnetic system. Section 3 presents the principal experimental results. The theory of the response of a bistable system subjected to the action of two harmonic forces in the presence of noise is presented in Sec. 4. Then Sec. 5 compares the theory with experiment, and the main results and conclusions are given in Conclusions.

2. DESCRIPTION OF THE OBJECT OF INVESTIGATION AND THE SETUP

The mixing of harmonics under the conditions of stochastic resonance was studied experimentally in a bistable magnetic system in the form of a local portion of a domain boundary in a thin easy-axis iron garnet film. It is known[20] that the motion of domain boundaries in iron garnet films consists of jumps of local sections. These jumps appear because of the interaction of the boundary with pinning sites. Thus, a local portion of a domain boundary fixed by an external field between two nearby pinning sites is a very simple bistable system. The characteristics of the motion of domain boundaries, the parameters of the pinning sites, and an analysis of an ensemble of magnetic defects in iron garnet films are presented in greater detail in Refs. 21 and 22.

Films of the iron garnets $(LuBi)_3(FeGa)_5O_{12}$ and $(GdTlBi)_3(FeGa)_5O_{12}$ with thicknesses from 7 to 30 μm, a saturation magnetization from 80 to 500 G, and an anisotropy field from 1 to 1.8 kOe were investigated. The results presented in this paper were obtained on a $(LuBi)_3(FeGa)_5O_{12}$ film with the following characteristics: a saturation magnetization $4\pi M_s = 80$ G, an anisotropy field $H_a = 1800$ Oe, a thickness $h = 30$ μm, and a stripe period $P = 25$ μm.

A local portion of an isolated domain boundary with a typical dimension ~ 10 μm, which was immobilized between two nearby pinning sites separated by a distance of ~ 0.4 μm, served as a bistable system. The position of the local portion in the film served as the generalized coordinate, and a bistable potential appeared because of the interaction of the domain boundary with the pinning sites.

To form an isolated boundary the iron garnet film was placed in a magnetic field gradient equal to 2000 Oe/cm, which was created by a special system of magnets.[23] Figure 1a shows the magnetizing of a film in a quadrupole field with the formation of a straight isolated boundary in a saddle line of the field. Such a geometry makes it possible to avoid the interactions between the numerous domain boundaries usually present in uniaxial films. In addition, by displacing the film relative to the saddle line of the gradient field, we could obtain a straight isolated domain boundary at any desired place in the film.

It should be noted that because of the presence of a large set of pinning sites, the local portion of the domain boundary moved in an effective bistable potential only when magnetic fields weaker than 0.5 Oe were applied. When the fields were stronger, motion took place in a multistable potential.

The displacement of the local portion was recorded by a standard magneto-optical method with transillumination. A block diagram of the setup is presented in Fig. 2. A beam of the linearly polarized output of He–Ne laser *1* with a wavelength $\lambda = 0.63$ μm and a polarization plane (the *yz* plane) perpendicular to the plane of the figure was focused by the objective lens *2* with a focal length $F = 1.5$ cm onto the surface of iron garnet film *3*, which was deposited epitaxially on a transparent gallium-gadolinium garnet substrate *4*.

The iron garnet film was immersed in a quadrupole mag-

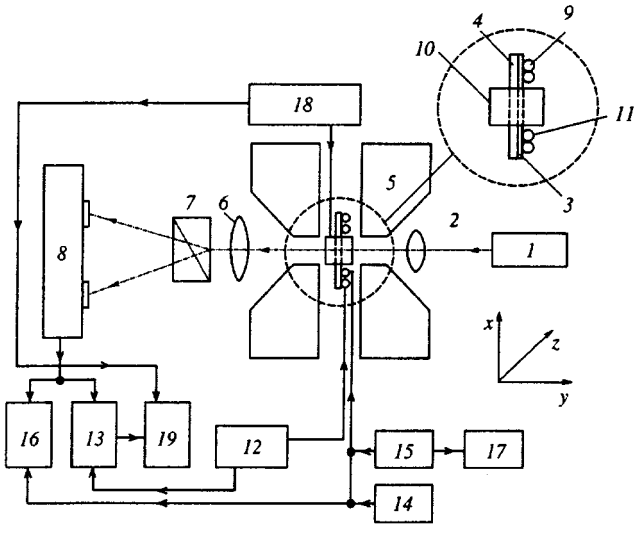


FIG. 2. Block diagram of the setup: *1* — laser, *2* — objective lens, *3* — iron garnet film, *4* — substrate, *5* — system of magnets, *6* — objective lens, *7* — Wollaston prism, *8* — photodetection device, *9, 10, 11* — coils for creating magnetic fields, *12* — noise generator, *13* — selective nanovoltmeter, *14, 15* — generators of periodic signals, *16, 17* — oscillographs used to monitor the main signals, *18* – source of a constant magnetic field, *19* — recorder.

netic field formed by system of magnets *5*. When the beam of polarized radiation passed through the sample, the polarization vector turned in accordance with the direction of $\mathbf{M}_s$ (see Fig. 1b). As a result, the original light beam split into two with oppositely oriented polarization vectors. Objective lens *6* with a focal length $F=16$ cm projected the image of the illuminated portion of the film onto the plane of the photodetection device *8*. After passage through the Wollaston prism *7*, the polarized beams were separated in space according to their directions of propagation and entered different photodiodes. Apart from the two photodiodes, the photodetection device included a noise-subtraction circuit and an amplifier, whose output was fed into the model 233 selective nanovoltmeter *13* and oscillograph *16*. Thus, the measured signal describes the variation of the optical flux appearing upon displacement of the domain boundary, and in our case it was directly proportional to the displacement.

Coil *9* served as a source of a magnetic noise field. The pair of coils *10* and *11* served as a source of harmonic magnetic fields. All the coils were located in the immediate vicinity of the film, as is shown in Fig. 2. Coil *10*, which created a field that varies slowly with a constant rate, was positioned above the film, and the turns of coil *9* and of coil *11*, which was a generator of a harmonic field, ringed the portion of the film below the focal spot of the light beam. When the coils were oriented this way, the perturbing fields and the noise field were directed along the *y* axis perpendicularly to the surface of the film, and the motion of the domain boundary was parallel to the *x* axis. For the purpose of reducing the interference caused by vibrations when small changes in the magneto-optical signals were measured, the entire structure was placed on an optical bench.

During the experiments a magnetic noise field was excited by an electrical signal formed by noise generator *12*. For monitoring and recording purposes, the same signal was fed into selective nanovoltmeter *13*. Model G4-153 generators *14* and *15* served as sources of harmonic signals. For monitoring and adjustment purposes, these signals were simultaneously fed into oscillographs *16* and *17*. Coil *10* was powered by dc generator *18*, which formed a slowly varying signal of both polarities, so that the field of coil *10* increased continuously from -55 mOe to $+55$ mOe. The optical system was aligned and the instruments were adjusted before the experiments were begun. The adjustments served to achieve the following principal goals.

1) The focal spot must be small enough that averaging of the response signal over many local portions of the domain boundary would not occur. This was characterized by the shape of the signal: when the driving field varies smoothly, the signal should be stepped.[20]

2) The domain boundary must divide the illuminated portion of the film under the focal spot into two symmetric parts (sectors), as is shown in Fig. 1b. Then, all other conditions being equal, the motion of the domain boundary would induce the strongest response signal.

3) The location of the domain boundary in the gradient magnetic field must be confined between two potential wells of the same depth.

Fulfillment of the last condition was determined in the following manner. The noise field was switched on in the presence of a harmonic field, causing an abrupt increase in the response signal, which was recorded by the selective nanovoltmeter *13*. Then a weak constant current was supplied to coil *10* using generator *18*. The resultant magnetic field displaced the domain boundary along the *x* axis, and a decrease in the response signal was consequently observed. Observation of a decrease in the signal to the same level when the direction of the current is reversed and, accordingly, the domain boundary is displaced in the opposite direction, would attest to fulfillment of the condition.

4) The Wollaston prism *7* must be oriented so that the signal from the photodetection device in the form of a noise track would have a minimal amplitude in the absence of a driving field.

Adjustment of the instruments involved establishing the parameters of the noise and perturbing signals that are required to carry out a specific experiment.

3. EXPERIMENT

In the first part of the experiments the response of a local portion of a domain boundary to an applied harmonic field in the presence of a random magnetic field was measured. The response was the spectral density of a magneto-optical signal, which was measured using a selective nanovoltmeter with a selectivity of 54 dB. The frequency dependence of the response contained a narrow peak, which was associated with the external harmonic field, and a smooth curve, which was associated with the noise fields. The contribution to the response caused by the external harmonic force at the frequency of that force was regarded as the measured signal in our case, while the contribution to the response caused by the random force was regarded as noise. Under the conditions investigated the noise characteristics of the response (the noise) did not vary in the absence of a modulating field. Thus, the noise could be measured as the response of the system in the absence of modulation. We note once again that the magneto-optical signal was directly proportional to the displacement of the local portion of the domain boundary, i.e., the generalized coordinate of the bistable system investigated.

The spectral density N of the magnetic noise field served as the noise characteristic. It was determined from the spectral current density in the noise coil at the frequency of the external harmonic field and was expressed in $\mu\text{Oe/Hz}^{1/2}$. The noise generator permitted variation of the noise parameter N in the range from 15 $\mu\text{Oe/Hz}^{1/2}$ to 110 $\mu\text{Oe/Hz}^{1/2}$. The amplitude of the harmonic field was assigned in the range from 5 mOe to 50 mOe, and its frequency was set in the range from 200 Hz to 7 kHz.

Figure 3 shows plots of the signal and the noise, which were parametrized by the resistance assigning the noise parameter in the noise generator and measured with a short integration time. Figure 4 presents the dependence of the signal-to-noise ratio on N. We note that, as N varies, the signal-to-noise ratio at first decreases, then begins to increase, and reaches a local maximum at a certain value of N, after which it decreases. The presence of such a maximum also characterizes the presence of magnetostochastic reso-

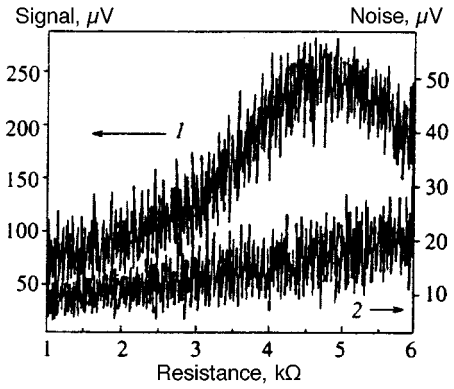


FIG. 3. Dependence of the signal (*1*) and the noise (*2*) on the resistance specifying the noise parameter. The amplitude of the variable field was 9 mOe, and its frequency was 260 Hz.

nance. Thus, a domain boundary fixed between two nearby pinning sites does, in fact, exist under stochastic resonance conditions.

We next studied the dependence of the response of the system on the symmetry of the potential. A constant magnetic field perpendicular to the plane of the film was applied to remove the degeneracy of the minima of the bistable potential. At a sufficiently strong constant field the system became monostable, and the local portion of the domain boundary was immobilized on a microdefect. This led to destruction of the stochastic resonance conditions and, as a consequence, to a decrease in the response to the applied variable field.

Figure 5 presents the response curves measured at a modulating field frequency equal to 4 kHz as a function of the amplitude of the constant magnetic field that removes the degeneracy. The parameter used here was N, which was varied from 15 μOe/Hz$^{1/2}$ to 63 μOe/Hz$^{1/2}$. As is seen from Fig. 5, the curves descend monotonically for both field directions. As the noise force increases, the response decay curve broadens. This broadening can be characterized by ΔH, i.e., the width of the range of fields in which the response is equal to or greater than 0.5 of the height of the maximum of the signal. Figure 6 shows the dependence of ΔH and of the maximum amplitude S of the signal $[S(H=0)]$ on the square of the noise parameter N^2, which is proportional to the noise level in the system. As we see, the $S(N^2)$ curve has

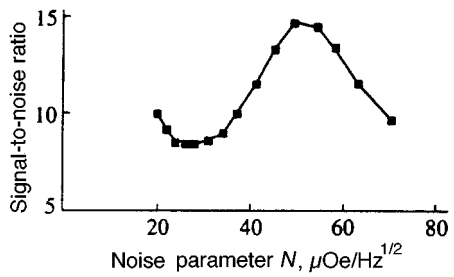


FIG. 4. Stochastic resonance curve for the fundamental harmonic. The amplitude of the variable magnetic field was 9 mOe, and its frequency was 260 Hz.

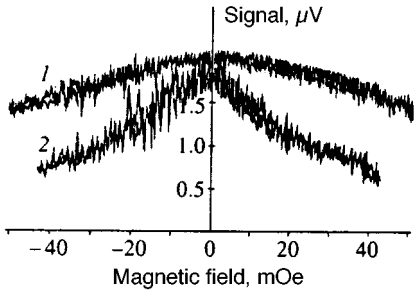


FIG. 5. Dependence of the signal on the constant magnetic field that makes the minima of the bistable potential unequal for $N=63$ μOe/Hz$^{1/2}$ (*1*) and $N=27$ μOe/Hz$^{1/2}$ (*2*). The amplitude of the harmonic field was 18 mOe, and its frequency was 4 kHz.

a maximum at $N^2=1400$ μOe/Hz$^{1/2}$, and the plot of $\Delta H(N^2)$ rises linearly with increasing N^2 over the entire range of variation.

The bulk of the experiments were devoted to the mixing of different harmonics, i.e., motion of a domain boundary under the action of the two harmonic fields H_1 and H_2 with frequencies f_1 and f_2. The mixing of harmonics was investigated both in a bistable potential with equal minima and with application of a constant field that removes the degeneracy. The dependence of the response on the noise force had the typical form for stochastic resonance. Figure 7 presents the stochastic resonance curve for an odd harmonic appearing under the action of two harmonic fields with the frequencies $f_1=3$ kHz and $f_2=500$ Hz and the amplitudes $H_1=20$ mOe and $H_2=25$ mOe, which was measured at the frequency $f_1+2f_2=4$ kHz. This dependence on the noise parameter was measured in the absence of a constant field. Even harmonics of the response were not observed under these conditions. When a constant field was applied, even harmonics appeared. Figure 8 presents the stochastic resonance curve for the even harmonic f_1+f_2, which was measured with the following parameters: $f_1=3$ kHz, $H_1=30$ mOe, $F_2=500$ Hz, $H_2=30$ mOe, and the constant field $H=10$ mOe. There was special interest in the dependence of the signals of higher harmonics on the constant field that makes the potential minima unequal. For this purpose, the noise level was fixed, and the constant field was varied in the range from -55 mOe to 55 mOe. Figure 9 shows the depen-

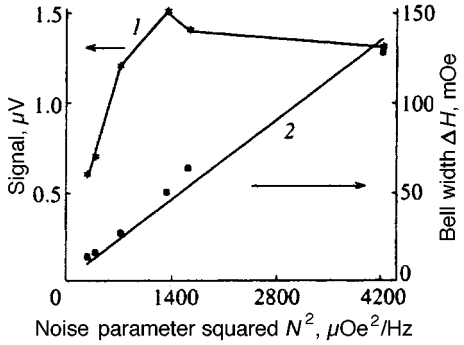


FIG. 6. Dependence of the width ΔH of the bell-shaped curve (*1*) and the maximum of the signal S (*2*) on the square of the noise parameter. The amplitude of the harmonic field was 18 mOe, and its frequency was 4 kHz.

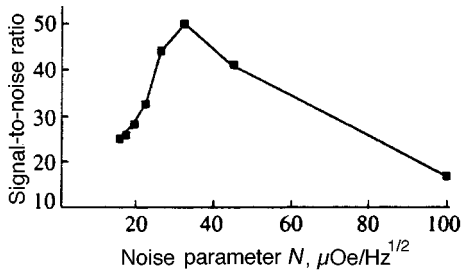


FIG. 7. Stochastic resonance curve for an odd harmonic. The amplitudes of the harmonic fields were $H_1=H_2=20$ mOe, and the frequencies were $f_1=3$ kHz and $f_2=500$ Hz. The signal detection frequency was $f=f_1+2f_2=4$ kHz.

dence of the signal of an even harmonic obtained for the noise parameter $N=27$ μOe/Hz$^{1/2}$ and for the following parameters of the harmonic fields: $f_1=4$ kHz, $H_1=7.7$ mOe, $f_2=1$ kHz, and $H_2=11.2$ mOe. We note the minimum of the response at zero field and two symmetric maxima. When the noise parameter was increased, the distance between the maxima of this curve increased. The response at the frequency of the second harmonic had a similar form. The odd harmonics were characterized by a different form of the dependence of the response on the constant field that destroys the symmetry of the system. This dependence, which is shown in Fig. 10, had a maximum at zero intensity of the constant field. The dependence was plotted for $f=4$ kHz and $H=16$ mOe, the recording frequency was equal to $3f=12$ kHz, and the noise parameter was $N=27$ μOe/Hz$^{1/2}$. The width of the curve shown in Fig. 10 also increased as the noise parameter increased.

4. THEORY

The one-dimensional motion of a system in a bistable potential $U_0(x)$ in the presence of a modulation force $F(t)=f_1\cos(\omega_1 t)+f_2\cos(\omega_2 t)$, a constant force h, which removes the degeneracy of the potential, and a random noise force $\xi(t)$ is described by the Langevin equation

$$\frac{dx}{dt}=-\frac{dU_0(x)}{dx}+F(t)+h+\xi(t). \tag{1}$$

Our problem is to find the response of the system to external modulation, i.e., the spectral density of the coordinate of the

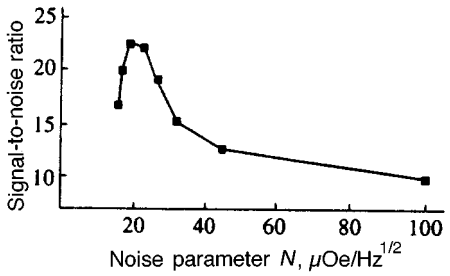


FIG. 8. Stochastic resonance curve for an even harmonic. The amplitudes of the harmonic fields were $H_1=30$ mOe and $H_2=21$ mOe, and the frequencies were $f_1=2.2$ kHz and $f_2=500$ Hz. The detection frequency was $f=f_1+f_2=2.7$ kHz, and the constant magnetic field that removes the degeneracy was equal to 10 mOe.

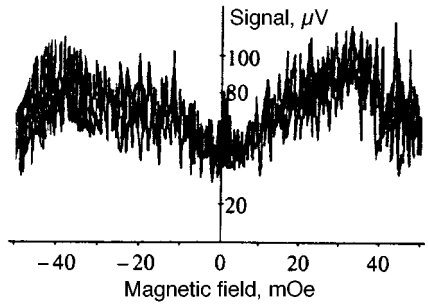


FIG. 9. Dependence of the signal of an even harmonic on the constant magnetic field that makes the minima of the bistable potential unequal. The amplitudes of the harmonic signals were $H_1=7.7$ mOe and $H_2=11.2$ mOe, the frequencies were $f_1=4$ kHz and $f_2=1$ kHz, the detection frequency was $f=f_1+f_2=5$ kHz, and the noise parameter was $N=27$ μOe/Hz$^{1/2}$.

system in a certain frequency range measured at the frequency of the external field or a combination harmonic. To find the response sought, we utilize an approach in which the continuous dynamics of the system is replaced by discrete jumps between the potential minima.[5]

Such a replacement is possible when the noise level is not excessively high, so that the system spends most of its time near the potential minima. We define $n_+(t)$ as the probability of finding the system near the right-hand minimum, while $n_-(t)$ is the probability of finding the system near the left-hand minimum:

$$n_+(t)=1-n_-(t)=\int_c^\infty p(x,t)dx, \tag{2}$$

where c is the position of the potential maximum and $p(x,t)$ is the probability density for the system studied. Then the equation describing the dynamics of the populations n_+ and n_- can be written in the form

$$\frac{dn_+}{dt}=-\frac{dn_-}{dt}=W_-(t)n_--W_+(t)n_+, \tag{3}$$

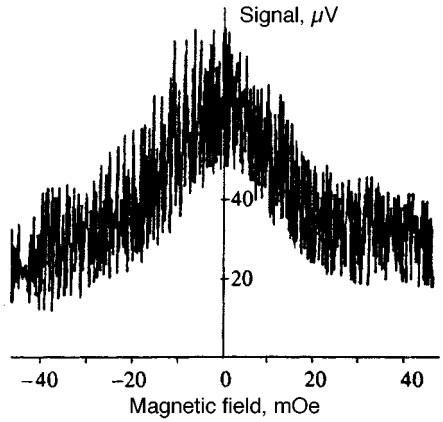


FIG. 10. Dependence of the signal of the third harmonic on the constant magnetic field that makes the minima of the bistable potential unequal. The amplitudes of the harmonic signal was $H_1=16$ mOe, the frequency of the field was $f_1=4$ kHz, the detection frequency was $f=3f_1=12$ kHz, and the noise parameter was $N=27$ μOe/Hz$^{1/2}$.

where $W_+(t)$ and $W_-(t)$ are the probabilities of transitions from the right-hand and left-hand minima per unit time, respectively. The distribution density $P(x,t)$ of the discrete system is determined by the relation

$$P(x,t)=n_+(t)\,\delta(x-x_+)+n_-(t)\,\delta(x-x_-), \tag{4}$$

where x_+ and x_- are selected so that the mean values of x and x^2 calculated using the functions $p(x,t)$ and $P(x,t)$ would coincide.

In the absence of modulation $[F(t)=0]$ the stationary distribution is given by the equation

$$p(x,t)=\mathcal{N}\exp[-U_0(x)/D], \tag{5}$$

where $\mathcal{N}$ is a normalization constant and D is the noise level. We assume that $+a$ and $-a$ are the minima of the symmetric bistable potential and that the expansion of the potential near the minima has the form

$$U_0(x)\approx U_0(\pm a)+k(x\pm a)^2/2.$$

Then, if

$$D\ll ka^2, \tag{6}$$

the system does, in fact, spend most of its time near the minima, and the stationary value of the population is found as

$$n_{+0}=N\int_c^\infty \exp\left[-\frac{U_0(x)-hx}{D}\right]dx. \tag{7}$$

This gives

$$n_{\pm 0}=\frac{\exp(\pm ha/D)}{2\cosh(ha/D)}. \tag{8}$$

In this case the main contribution to the mean values of the moments

$$\langle x^n\rangle=N\int_{-\infty}^\infty x^n\exp\left[-\frac{U_0(x)-hx}{D}\right]dx \tag{9}$$

is also made by the potential minima. Performing the calculations, we obtain

$$\langle x\rangle\approx a\tanh(ha/D), \tag{10}$$

$$\langle x^2\rangle-\langle x\rangle^2\approx a^2[\cosh^2(ha/D)]^{-1},$$

and, accordingly $x_\pm=\langle x\rangle\pm 2n_{\mp 0}a$, so that $x_+-x_-=2a$ even in the presence of the constant force h.

Within the same approximation of a low noise level [Eq. (6)], we can write a formula for the transition frequency:[24]

$$W_\pm=\nu_0\exp\left[-\frac{U(c)-U(a_\pm)}{D}\right], \tag{11}$$

where $2\pi\nu_0=\sqrt{|U''(a_\pm)U''(c)|}$, and $a_\pm$ and c are the positions of the potential minima and maximum, respectively.

Since we assume that the noise is not excessively great [see Eq. (6)], the transition frequency $W_\pm$ is much smaller than the characteristic frequency of motion of the system at a minimum $\omega_0=k$, which corresponds to the establishment of local equilibrium. Therefore, at frequencies of the external fields which satisfy the condition

$$\omega_{1,2}\ll\omega_0 \tag{12}$$

we can use the adiabatic approximation, under which the function $W_\pm(t)$ is found by substituting the instantaneous values of the positions of the potential maximum into Eq. (11). This leads to the following formula for $W_\pm$:

$$W_\pm=\nu_K\exp\left(\frac{\mp(h+F)a-q(h+F)^2/k}{D}\right), \tag{13}$$

where

$$\nu_K=\nu_0\exp\left[-\frac{U(0)-U(a)}{D}\right]$$

is the Kramers frequency, and q is a numerical coefficient of order unity, which depends on the form of the potential. The problem consisting of (3) and (13) cannot be solved exactly. Let us examine two limiting cases: 1) both forces are small and 2) one force is small, while the other is arbitrary, but has a high frequency.

4.1. Weak modulation

In this case we can neglect the quadratic term in the denominator of the exponential function (13) and expand all the quantities in series in the small parameter $\varepsilon=Fa/D$: $n_+=n_{+0}+n_{+1}+n_{+2}+\ldots$, $W_\pm=W_{\pm 0}+W_{\pm 1}+W_{\pm 2}+\ldots$. This gives a hierarchy of equations, the first of which specifies the stationary density, and the others,

$$\frac{dn_{+1}}{dt}=\frac{W_{-1}W_{+0}-W_{+1}W_{-0}}{W_{+0}+W_{-0}}-(W_{+0}+W_{-0})n_{+1}, \tag{14.1}$$

$$\frac{dn_{+2}}{dt}=-(W_{+1}+W_{-1})n_{+1}-(W_{+0}+W_{-0})n_{+2}, \tag{14.2}$$

describe the resonance and the mixing of harmonics in an asymmetric potential. The signal is found either from the correlation function $\langle x(t+\tau)x(t)\rangle$ or from formal calculations of x_ω and $|x_\omega|^2$.

The two approaches give the following formulas for the spectral density and the signal-to-noise ratio in the case of $h\neq 0$, $f_1\neq 0$, and $f_2=0$:

$$S(\omega)=\left(4n_{+0}n_{-0}-\frac{2A^2}{B^2+\omega_1^2}\right)\frac{4a^2B}{B^2+\omega^2}+\frac{4\pi a^2A^2}{B^2+\omega_1^2}\times\delta(\omega_1-\omega), \tag{15}$$

where

$$A=\frac{2W_{+0}W_{-0}}{W_{+0}+W_{-0}}\frac{fa}{D},\quad B=W_{+0}+W_{-0}=2\nu_K\cosh\left(\frac{ha}{D}\right),$$

and

$$R=\frac{\pi f^2a^2}{D^2}\frac{\nu_0\exp(-\Delta U/D)}{2\cosh(ha/D)\Delta f}=\frac{R_0}{\cosh(ha/D)}, \tag{16}$$

where Δf is the width of the band detected and R_0 is the signal-to-noise ratio in the absence of the constant force h. At a frequency $\omega_1 \ll B$ the dependence of the signal can be found as

$$S = S_0[\cosh^4(ha/D)]^{-1}, \tag{17}$$

where S_0 is the signal when $h=0$. The mixing of harmonics is found by solving Eqs. (14). In the case of ω_1, $\omega_2 \ll W_{+0}+W_{-0}$ the signal-to-noise ratio for higher harmonics is determined by the following equations:

$$R_2(\omega=\omega_1+\omega_2) = \frac{\pi f_1^2 f_2^2 a^6}{D^4 \Delta f} \frac{\nu_0 \exp(-\Delta U/D)}{\cosh(ha/D)} \tanh^2\left(\frac{ha}{D}\right), \tag{18.1}$$

$$R_3(\omega=\omega_1+2\omega_2) = \frac{2\pi f_1^2 f_2^4 a^8}{D^6 \Delta f} \frac{\nu_0 \exp(-\Delta U/D)}{\cosh(ha/D)} \times \tanh^2\left[1+3\tanh^2\frac{ha}{D}\right]. \tag{18.2}$$

(We note that an error was made in Ref. 14 in calculating the signal of mixed harmonics.) Thus, $R_{n+1} \sim (fa/D)^{2n} R_0$, where R_0 is the signal-to-noise ratio for the fundamental harmonic, and, therefore, the higher harmonics also exhibit stochastic resonance. In addition, all the harmonics depend on the constant field h, which makes the minima unequal, with the characteristic scale D/a.

4.2. Fast modulation

Let us assume that the force f_2 is not small and that the frequency ω_2 is large: $\omega_2 \gg \nu_K, \omega_1$. (Under these conditions ω_2 is still smaller than the fundamental oscillation frequency of the system ω_0, so that the adiabatic approximation is valid.) The fast and slow components can be isolated in the response, and averaging can be performed over the fast component. This leads to an equation for the slow component $n_{1+}(t)$:

$$\frac{dn_{1+}}{dt} = \langle W_-(t)\rangle - [\langle W_+(t)+W_-(t)\rangle] n_{1+}, \tag{19}$$

where $\langle W\rangle$ denotes averaging of Eq. (13) over the fast component. Performing the calculations, we obtain

$$\frac{1}{G}\frac{dn_{1+}}{dt} = W_-(t) - [W_+(t)+W_-(t)] n_{1+}, \tag{20}$$

where $W_+(t)$ and $W_-(t)$ correspond to the transition probabilities for the case in which only the force f_1 acts, and

$$G(\alpha) = \frac{1}{2\sqrt{\pi\beta}} \int_{-\infty}^{\infty} J_0(\alpha') \exp\left[-\frac{(\alpha'-\alpha)^2}{4\beta}\right] d\alpha', \tag{21}$$

where $\alpha = f_2 a/D$, $\beta = q f_2^2/kD$, and J_0 is a Bessel function of order zero.

Equation (21) describes the stochastic resonance for the harmonic ω_1. The presence of a strong variable high-frequency field of fairly large amplitude leads to renormalization of the fundamental frequency $\nu_0 \to G\nu_0$. This gives a signal-to-noise ratio similar to the usual expression for stochastic resonance:

$$R(\omega_1) \cong \frac{\pi f_1^2 a^2}{D^2 \Delta f} \frac{G\nu_0 \exp(-\Delta U/D)}{2\cosh(ha/D)}. \tag{22}$$

However, as follows from Eq. (21), G depends on the noise level D. It is easy to obtain the following asymptotes for G: $G \to 0$ as $D \to \infty$ and $G \approx \exp(f_2 a/D)$ as $D \to 0$. Thus, when the noise level is low, the high-frequency force f_2 leads to a decrease in the effective barrier height to $\Delta U \cong \Delta U - f_2 a$, confirming the observation made in Ref. 16.

5. COMPARISON OF EXPERIMENT AND THEORY

The motion of a domain boundary considered in Sec. 3 is described by the equation

$$\frac{dx}{dt} = \mu\left[-\frac{dW_0}{dx} + H(t) + H + H_{\text{ran}}\right], \tag{23}$$

where μ is the mobility of the domain boundaries, $W_0(x)$ is the bistable potential of the magnetic microdefects, $H(t)$ is the modulation magnetic field, H is the constant field, and H_{ran} is the noise field. A comparison of (1) and (23) allows us to conclude that the noise level D is specified by the expression

$$D = (\mu N)^2, \tag{24}$$

where N is the noise parameter used in Sec. 3, and the constant field h is found as $h = \mu H$. The height of the potential barrier U_0 can be estimated as $U_0 = \mu H_c a \approx 4\times 10^{-3}$ cm^2/s, where $\mu = 10^3$ cm$(\text{s}\cdot\text{Oe})^{-1}$ is the mobility of the domain boundaries in the film under investigation, $H_c = 0.2$ Oe is the field at which the domain boundaries break away from the pinning sites, which was close to the coercive field of the film, and $2a \approx 0.4$ μm is the distance between the potential minima. The characteristic frequency of motion of the boundary is found as $\nu_0 \approx \mu H_c/a = 10^7$ Hz, and the Kramers frequency ν_K varied in our experiments from 0 to 6×10^6 Hz.

The experimental data are in good agreement with the theory. In fact, the position of the stochastic resonance maximum for the fundamental harmonic corresponds to the condition $2D_{\max} = U_0$, whence $N_{\max} = \sqrt{H_c a/2\mu} = 45$ μOe/Hz$^{1/2}$, which coincides with the measured value $N_{\max} = 49$ μOe/Hz$^{1/2}$ (see Fig. 4). In Figs. 7 and 8 there are also peaks for the signal-to-noise ratio at combination frequencies. This means that there is an optimal noise level in the system, at which the signal-to-noise ratio at the higher harmonics reaches a maximum. The experimental range of variation of the noise did not allow us to detect an increase in the signal-to-noise ratio at low noise levels, but a comparison of Figs. 7 and 8 with Fig. 4 suggests that the Q factor of the resonance is higher for the mixed harmonics (see Ref. 14).

The dependence of the harmonics on the constant field h, which makes the minima unequal, is most interesting from the standpoint of practical applications. The field h leads to the loss of one of the most important conditions for stochastic resonance, viz., bistability, and diminishes the response at

the fundamental and higher harmonics, as is demonstrated by Figs. 5, 9, and 10, as well as Eqs. (15), (16), and (18). This permits the use of stochastic resonance to detect not only periodic, but also slowing varying fields. Under these circumstances the field h appears in the basic relations only in the combination ha/D, and the dependence of the harmonics on h consequently has a single identical scale D/a. Thus, if we know only one parameter a (which characterizes the distance between the degenerate minima) and measure the width of the decay curve of odd harmonics or the distance between the maxima of even harmonics, we can determine the value of the noise level in the system from the dependence of the response to external modulation.

The dependence of the width ΔH of the bell-shaped curve of the signal shown in Fig. 5 on the square of the noise parameter N^2, which is proportional to the noise level D [see Eq. (24)], is presented in Fig. 6. As is seen from Fig. 6, the width ΔH is, in fact, directly proportional to D. Knowing the noise level in the system and ΔH, we can estimate the distance between the minima of the bistable system. This gives $2a \cong \mu N^2/\Delta H \sim 0.5$ μm, which is consistent with the value $2a = 0.4$ μm obtained from static measurements.

In addition, it is important to note the symmetry of the even and odd harmonics as a function of the constant field h. The odd harmonics have a maximum at $h=0$, and the even ones have a minimum (see Figs. 9 and 10). The distance between the maxima of the even harmonics, as the measurements show, also increases with the noise and can be approximated by the formula $\Delta H \cong \mu N^2/a$. Thus, indirect measurements of the response of the system to periodic modulation enable us to draw conclusions regarding the tunneling characteristics of the system.

6. CONCLUSIONS

We have shown theoretically and experimentally that the combination harmonics appearing when two periodic forces act on a bistable system in the presence of noise exhibit behavior that is characteristic of stochastic resonance. The dependence of the signals at the fundamental and mixed harmonics on the constant field that makes the minima unequal has been investigated, and it has been shown that the dependence sought has a variation scale proportional to the noise level. The laws derived have been tested on a very simple bistable magnetic system, viz., a local portion of a domain boundary immobilized between two nearby microdefects in a thin iron garnet film. Stochastic resonance has been detected at mixed harmonics (of second and third order), and the dependence of the harmonics on the constant field that makes the minima of the bistable potential unequal has been measured. The symmetry of the harmonics and the parameters of their variation are in qualitative agreement with the theory developed.

The most important consequence of this research is the possibility of determining the fluctuation parameters of a bistable system (the noise level) by measuring the response of the system to periodic modulation (the modification of the response when a constant field that makes the minima unequal is applied).

This work was supported by the Russian Fund for Fundamental Research (Grants 96-02-18956 and 96-02-19608), as well as by the international association INTAS (Grant 94-1720).

[1] R. Benzi, A. Sutera, and A. Vulpiani, J. Phys. A: Math. Gen. **14**, L453 (1981).
[2] R. Benzi, G. Parisi, A. Sutera, and A. Vulpiani, Tellus **34**, 10 (1982).
[3] S. Fauve and F. Heslot, Phys. Lett. A **97**, 5 (1983).
[4] B. McNamara, K. Wiesenfeld, and R. Roy, Phys. Rev. Lett. **60**, 2626 (1988).
[5] B. McNamara and K. Wiesenfeld, Phys. Rev. A **39**, 4854 (1989).
[6] L. Gammaitoni, F. Marchesoni, E. Menichella-Saetta, and S. Santucci, Phys. Rev. Lett. **62**, 349 (1989).
[7] R. Fox, Phys. Rev. A **39**, 4148 (1989).
[8] P. Yung and P. Hanggi, Phys. Rev. A **41**, 2977 (1990).
[9] M. I. Dykman, P. V. E. McClintock, R. Mannella, and N. G. Stocks, JETP Lett. **52**, 141 (1990).
[10] A. N. Grigorenko, V. I. Konov, and P. I. Nikitin, JETP Lett. **52**, 593 (1990).
[11] L. D. Landau and E. M. Lifshitz, *Mechanics*, 2nd ed., Pergamon Press, Oxford–New York (1969).
[12] A. N. Grigorenko, P. I. Nikitin, A. N. Slavin, and P. Zhou, J. Appl. Phys. **76**, 6335 (1994).
[13] M. I. Dykman, R. Mannella, P. V. E. McClintock, and N. G. Stocks, Phys. Rev. Lett. **68**, 2985 (1992).
[14] A. N. Grigorenko, P. I. Nikitin, and G. V. Roschepkin, J. Appl. Phys. **79**, 6113 (1996).
[15] L. Gammaitoni, F. Marchesoni, and S. Santucci, Phys. Lett. **195**, 116 (1994).
[16] S. Vohra and L. Fabiny, Phys. Rev. E **50**, R2391 (1994).
[17] R. Bartussek, P. Jung, and P. Hanggi, Phys. Rev. E **49**, 3930 (1994).
[18] V. Shneidman, P. Jung, and P. Hanggi, Phys. Rev. Lett. **72**, 2682 (1994).
[19] P. Jung and P. Talkner, Phys. Rev. E **51**, 2640 (1995).
[20] A. N. Grigorenko, S. A. Mishin, A. M. Prokhorov, and E. G. Rudashevskiĭ, Fiz. Tverd. Tela (Leningrad) **30**, 2948 (1988) [Sov. Phys. Solid State **30**, 1699 (1988)].
[21] A. N. Grigorenko, S. A. Mishin, and E. G. Rudashevskiĭ, Fiz. Tverd. Tela (Leningrad) **31** (8), 107 (1989) [Sov. Phys. Solid State **31**, 1337 (1989)].
[22] A. N. Grigorenko, S. A. Mishin, and E. G. Rudashevskiĭ, Pis'ma Zh. Tekh. Fiz. **14**, 1722 (1988) [Sov. Tech. Phys. Lett. **14**, 749 (1988)].
[23] A. N. Grigorenko, O. A. Larin, S. A. Mishin, and E. G. Rudashevskiĭ, Prib. Tekh. Eksp. **(1)**, 163 (1988).
[24] C. W. Gardiner, *Springer Series in Synergetics, Vol. 13: Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*, 2nd. ed., Springer-Verlag, Berlin (1983) [Russ. transl., Mir, Moscow (1986)].

Translated by P. Shelnitz

Breakdown of the linear current regime in periodic structures

A. M. Satanin*) and S. V. Khor'kov

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

V. V. Skuzovatkin

Institute of Microstructure Physics, Russian Academy of Sciences, 603600 Nizhniĭ Novgorod, Russia
(Submitted 31 December 1996)
Zh. Éksp. Teor. Fiz. **112**, 643–660 (August 1997)

We study the effect of a weak nonlinearity in media on the linear regime of current flow in two-dimensional periodic structures with two equal component concentrations. We find that the asymptotic behavior of the electric field and current as functions of the distance between the angles in heterogeneous media is determined by the parameter $h=\sigma_2/\sigma_1$ (here σ_1 and σ_2 are the linear conductivities of the cells) and the external magnetic field B. This dependence leads to divergence of the higher-order moments of field and current at certain critical values h_c and B_c and to divergence of the response functions related to the higher-order moments. For square cells the effective nonlinear conductivity diverges at $h \leqslant h_c$, with $h_c=(\sqrt{2}-1)^2$. For structures of general shape we find the dependence of h_c on the angles and the external magnetic field. We show that for a given structure the linear regime of current flow in the system can be reversibly transformed into a nonlinear one by varying the magnetic field strength. The critical field B_c is approximately determined from the condition $\omega_c\tau \sim 1$, where ω_c and τ^{-1} are, respectively, the cyclotron frequency and the collision rate. Finally, we discuss the feasibility of detecting these effects experimentally. © *1997 American Institute of Physics.* [S1063-7761(97)01808-8]

1. INTRODUCTION

Much effort, both theoretical and experimental, has gone into the study of the electrical conductivity of random media and periodic structures.[1] Two-dimensional systems, in view of the importance of dual symmetry,[2] have received special treatment. The deep analogy between two-dimensional random systems and periodic lattices has repeatedly been noted,[2–4] so that they can be combined into a single class of systems with universal behavior in the critical region, i.e., near the metal–insulator transition. Two-dimensional systems are also of interest from the experimenter's viewpoint because nanotechnology make it possible to manufacture films of a given composition and films with a given texture.

The effective electrical conductivity σ^{eff} of a two-component film with conductivities σ_1 and σ_2 and concentrations p and $1-p$, respectively, can be written as $\sigma^{\text{eff}}=\sigma_1 f(p,h)$, where $h=\sigma_2/\sigma_1$, and f is a function with universal behavior in the critical region.[2,5] If $h=0$, there exists a critical concentration p_c at which the metal–insulator transition occurs and the conductivity σ^{eff} vanishes. For a two-dimensional random medium $p_c=0.5$. At this concentration the correlation radius, which characterizes the connected region of the conducting component, becomes infinite.

Periodic structures with two equal component concentrations behave like disordered films. The simplest example is provided by periodic networks whose cells form a checkerboard pattern; the conductivities of the cells are σ_1 and σ_2. Since the component concentrations are equal, the system is in the critical region, i.e., at the threshold of the metal–insulator transition. At $p=p_c$ the effective conductivity of a disordered film and a texture depends on the parameter h; for the case at hand this dependence has the form $\sigma^{\text{eff}}=\sigma_1\sqrt{h}$ (see Ref. 2). This makes it possible to assign these systems to a single universal class.

Another important property of such systems is related to the distribution of electric fields and currents in them. Dykhne[2] found that the one-point correlators of the field and current averaged over the surface area of the disordered film, $\langle \mathbf{e}^2 \rangle$ and $\langle \mathbf{j}^2 \rangle$, diverge at small values of h. Periodic structures also exhibit a similar property.[4]

We believe that it is important to generalize these results to nonlinear random media and periodic structures. That nonlinear effects are large in the critical region has been demonstrated in the experiments of Gefen *et al.*,[6] who detected anomalous behavior of the critical current and the critical nonlinearity field in disordered gold films near the metal–insulator transition as functions of the parameter p. Nonlinear effects as functions of p were studied in Refs. 7–11. It was found that higher-order correlators of the electric field and current provide ample information about the microstructure of heterogeneous media.[7,8] The effective nonlinear conductivity, the $1/f$ noise coefficient, and the third-harmonic amplitude were expressed in terms of the correlators $\langle \mathbf{e}^4 \rangle$ and $\langle \mathbf{j}^4 \rangle$ (see Refs. 8–10). Anomalous behavior in the critical region of the nonlinear conductivity in disordered films with two equal component concentrations was discussed in Refs. 12 and 13, where it was found that the nonlinear conductivity and the related correlators $\langle \mathbf{e}^4 \rangle$ and $\langle \mathbf{j}^4 \rangle$ diverge at small h. The critical properties of textures have yet to be studied.

In the present paper we show that the current flow regime in checkerboard textures differs drastically from that in

 1063-7761/97/080351-09$10.00

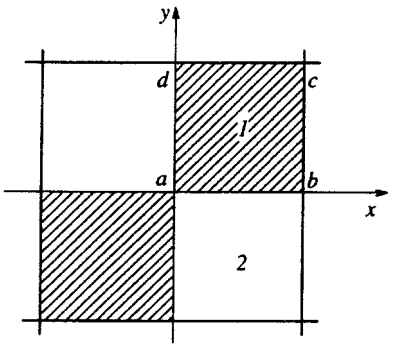


FIG. 1. A fragment of a two-dimensional periodic structure with cell conductivities σ_1 and σ_2.

a random medium. It was found that in view of the anomalous behavior of current and field near corners, the linear current flow regime breaks down at a finite value of the parameter h. This critical value depends on the geometry of the texture and the external magnetic field. We study the dependence of the critical nonlinearity field and the critical current on the parameters of the system. Some results of the present research have been presented in a brief communication.[14]

The plan is as follows. Section 2 is devoted to the basic equations and the generalization of the Aharony–Stroud–Hui relationships for nonlinear conductivity to locally anisotropic media. In Sec. 3 we discuss an exact solution of the linear-problem equations for the electric field for a checkerboard structure, and analyze the nature of the field singularities and the possible consequences for the nonlinear problem. In Sec. 4 we calculate the effective nonlinear conductivity and show that the nonlinear conductivity diverges at a finite ratio of the linear conductivities of the components. This result is generalized in Sec. 5 to oblique lattices, while in Sec. 6 we study the behavior of the system in an external magnetic field. We devote Sec. 7 to a discussion of the role that singularities in textures play in the formation of the response functions, the fact that the critical behavior of the higher-order moments of field and current are of a nonuniversal nature, and the feasibility of observing the predicted instability. In Appendix A we discuss some details of calculating the nonlinear conductivity. Finally, in Appendix B we elaborate on the problem of calculating the electric field near a corner of an oblique lattice.

2. THE BASIC EQUATIONS

The geometry of the structure we will examine is depicted in Fig. 1. The hatched squares are associated with the conductivity σ_1 (metal) and the light squares with σ_2 (insulator). In the linear approximation the current in the cells is

$$\mathbf{j}=\sigma\mathbf{e}, \tag{1}$$

where σ is a periodic functions with values σ_1 and σ_2. The current and field obey the equations

$$\nabla\cdot\mathbf{j}=0, \quad \nabla\times\mathbf{e}=0 \tag{2}$$

and the conditions at cell boundaries

$$(\mathbf{j}\cdot\mathbf{n})_1=(\mathbf{j}\cdot\mathbf{n})_2, \quad (\mathbf{e}\cdot\boldsymbol{\tau})_1=(\mathbf{e}\cdot\boldsymbol{\tau})_2, \tag{3}$$

where $\mathbf{n}$ and $\boldsymbol{\tau}$ are unit vectors normal and tangent to the medium boundary. The external field and current can be specified either at the surface of the macroscopic specimen or at the cell boundaries. The electric field and current are nonuniform in the cells, so that it is convenient to introduce values of the respective quantities that are averages over the system's surface area or over the surface area of two adjacent cells,

$$\mathbf{E}=\langle\mathbf{e}\rangle, \quad \mathbf{J}=\langle\mathbf{j}\rangle, \tag{4}$$

so that we can define the effective conductivity σ^{eff} as

$$\mathbf{J}=\sigma^{\text{eff}}\mathbf{E}, \tag{5}$$

or

$$\sigma^{\text{eff}}\mathbf{E}^2=\langle\sigma\mathbf{e}^2\rangle. \tag{6}$$

For macroscopic systems the above definitions are equivalent to those of Stroud and Hui.[10]

Since near the texture corners the field and current may have singularities (see below), we must allow for nonlinear effects from the start. The physical mechanisms governing the nonlinearities strongly depend on the material from which the cells are manufactured. If the cells are represented by a ''clean'' metal and a ''dirty'' metal (cell conductivity σ_1 and σ_2, respectively), we can employ the electron temperature approximation.[15] But if, for instance, the cells with conductivity σ_2 are manufactured from a wide-gap insulator or a doped semiconductor, one must allow for the activation mechanism of the field dependence of the current[16] in such cells. In this paper we limit our discussion to the weak nonlinearity approximation. In this case each of the above models reduces to allowing for the cubic term in the field expansion of the current:

$$\mathbf{j}=\sigma\mathbf{e}+\chi\mathbf{e}^2\mathbf{e}. \tag{7}$$

For instance, in the electron temperature approximation,

$$\chi=\frac{\partial\sigma(T)}{\partial T}\frac{\sigma}{\Lambda}, \tag{8}$$

where Λ is the coefficient of heat transfer from electrons to lattice.

To examine periodic lattices of a more general shape, we introduce the concept of effective nonlinear conductivities of such media. Anisotropy in the boundary conditions leads to anisotropy in the solutions of Eqs. (3), and the effective nonlinear characteristics of the media become tensors:

$$J_\alpha=\sigma^{\text{eff}}_{\alpha\beta}E_\beta+\chi^{\text{eff}}_{\alpha\beta\gamma\delta}E_\beta E_\gamma E_\delta. \tag{9}$$

The number of independent components of the nonlinear conductivity tensor is determined by the symmetry of the lattice. For instance, in the case of a square lattice there are four such components: $\chi^{\text{eff}}_{xxxx}=\chi^{\text{eff}}_{yyyy}$, $\chi^{\text{eff}}_{xxyy}=\chi^{\text{eff}}_{yyxx}$, $\chi^{\text{eff}}_{xyyx}=\chi^{\text{eff}}_{yxxy}$, and $\chi^{\text{eff}}_{xyxy}=\chi^{\text{eff}}_{yxyx}$, while for a hexagonal lattice and an isotropic medium there are three such components, since for such media there is the additional condition that $\chi^{\text{eff}}_{xxxx}=\chi^{\text{eff}}_{xxyy}+\chi^{\text{eff}}_{xyxy}+\chi^{\text{eff}}_{xyyx}$.

To find the effective nonlinear conductivity we need to solve Eqs. (2) with conditions (3) and (7) and calculate the components of the average current by integrating the local expression over the pair of adjacent cells. In an isotropic medium, the effective nonlinear conductivity to first order in powers of the nonlinear component conductivity can be determined only from the known linear field in an equivalent linear medium.[9,10] For an anisotropic medium we generalize the Aharony–Stroud–Hui relation[9,10] and show how to obtain some of the components of the effective nonlinear conductivity tensor from this relation. Writing the expression for local energy dissipation and employing Tellegen's theorem (see Ref. 10), we obtain

$$\chi^{\text{eff}}_{\alpha\beta\gamma\delta}E_\alpha E_\beta E_\gamma E_\delta=\langle\chi\mathbf{e}^4\rangle, \tag{10}$$

where the local electric field $\mathbf{e}$ can be found by solving the linear equations. For a square lattice Eq. (10) yields

$$\chi^{\text{eff}}_{xxxx}(E_x^4+E_y^4)+2(\chi^{\text{eff}}_{xxyy}+\chi^{\text{eff}}_{xyxy}+\chi^{\text{eff}}_{xyyx})E_x^2E_y^2=\langle\chi\mathbf{e}^4\rangle, \tag{11}$$

i.e., this approach yields only the component χ^{eff}_{xxxx} and the combination of components comprising the coefficient of $E_x^2E_y^2$.

Let us examine the simple case in which the solution has the maximum possible symmetry (here the external field is directed along the diagonals of the squares, and the solutions in adjacent cells are related through a dual transformation[4]). Let the system of coordinates be oriented along the diagonals of the squares. Then only one component of the nonlinear conductivity tensor remains in (11), and we denote it by $\chi^{\text{eff}}=\chi^{\text{eff}}_{xxxx}$. For this component we have

$$\chi^{\text{eff}}=\frac{\langle\chi\mathbf{e}^4\rangle}{\langle e\rangle^4}. \tag{12}$$

Thus, to calculate the nonlinear conductivity we must solve Eq. (2) with (1) and calculate the correlator (12). Nonlinear effects in the medium are important if the terms in parentheses in

$$\mathbf{J}=(\sigma^{\text{eff}}+\chi^{\text{eff}}\mathbf{E}^2)\mathbf{E} \tag{13}$$

are of the same order. For the external field directed along the diagonals of the squares we define, following Ref. 11, the critical nonlinearity field as

$$E_c=\sqrt{\frac{\sigma^{\text{eff}}}{\chi^{\text{eff}}}}, \tag{14}$$

and the critical current as

$$J_c=\sigma^{\text{eff}}E_c. \tag{15}$$

The critical field and current depend on the effective characteristics of the system, σ^{eff} and χ^{eff}. Note that according to (12) χ^{eff} is determined by two factors. We rewrite (12) as a sum of averages over the components,

$$\chi^{\text{eff}}=\frac{\chi_1\langle\mathbf{e}^4\rangle_1+\chi_2\langle\mathbf{e}^4\rangle_2}{2E^4}. \tag{16}$$

This expression implies that if, for instance, as Eq. (8) shows, the heat transfer coefficient Λ is large, the coefficients χ_1 and χ_2 can be made small. However, the field correlators $\langle\mathbf{e}^4\rangle_1$ and $\langle\mathbf{e}^4\rangle_2$ at certain values of h (or other parameters of the system) may have singularities. The increase in the field correlators is the reason for the anomalous increase in the effective nonlinearities of the structure. The anomalous increase in nonlinear conductivity means that the system has gone into a nonlinear current flow regime. The same result can be obtained by solving the exact equations for the current (Eqs. (2), (3), and (7)).

3. EXACT SOLUTION FOR A PLANAR PERIODIC LATTICE

As noted earlier, χ^{eff} can be expressed in terms of the solution of the linear problem. The linear problem of field and current distributions in a planar periodic checkerboard structure has been studied by Emets.[4] The exact solution was obtained via methods of the theory of boundary-value problems involving analytic functions. By virtue of the two-dimensional nature of the problem, the initial system of equations (1)–(3) allows for a complex representation. If we take into account the periodicity and the inversion symmetry, the problem of finding the current (or field) reduces to the homogeneous Markushevich problem[17] for two analytic functions, $e_1(z)$ and $e_2(z)$ (with $z=x+iy$), in adjacent cells. Mapping the adjacent square via the Weierstrass function into the upper and lower half-planes, we arrive at the problem for two functions that are analytic in, respectively, the upper and lower half-planes with boundary conditions on the real axis. As shown by Emets,[4] for a square lattice such a problem reduces to an equivalent Riemann problem for two pairs of analytic functions, which allows for a complete solution. The expressions for the electric field are:

$$e_1=c_1(h)X(z)+c_2(h)X^{-1}(z), \tag{17}$$

$$e_2=c_3(h)X(z)+c_4(h)X^{-1}(z), \tag{18}$$

where $c_i(h)$ are functions without singularities, and

$$X(z)=\left[\frac{\text{cn}(Kz/L,k)}{\text{sn}(Kz/L,k)\text{dn}(Kz/L,k)}\right]^{2\gamma}, \tag{19}$$

with sn x, cn x, and dn x Jacobi's elliptic functions, K the complete elliptic integral with modulus k (for a square, $k=1/\sqrt{2}$ and $K=1.8541$), and L the length of the side of a square cell. The parameter γ is related to h by the equation

$$\tan\pi\gamma=\frac{1-h}{2\sqrt{h}},\quad 0\leqslant\gamma\leqslant\frac{1}{2}. \tag{20}$$

Let us analyze the general solution. The first and second terms on the right-hand sides of Eqs. (17) and (18) are particular solutions corresponding external fields directed along the diagonals ac and bd (see Fig. 1). Near the vertices a and c the function $X(z)$ exhibits singularities in its behavior:

$$X(z)\propto\frac{1}{z^{2\gamma}}, \tag{21}$$

while at the adjacent corners b and d the function behaves as

$$X(z)\propto z^{2\gamma}. \tag{22}$$

Clearly, for the function $X^{-1}(z)$, which corresponds to the case in which the external field is directed along the diagonal

b and d, the singular behavior described by (21) is associated with the corners b and d. Thus, for an arbitrary direction of the external field, the general solution has singularities at all corners of the lattice. Singularities like (21) in the neighborhood of a corner with a singularity play an important role in the formation of the nonlinear current flow regime.

Qualitatively, the asymptotic behavior (21) and (22) can be explained as follows. Near a corner, at distances much smaller than the cell size, the system has no characteristic spatial scales. Hence the solution of the Laplace equation in this range of scales is given by a power function, with the exponent depending on the charge distribution at the boundary between regions with different conductivities and, hence, on the parameter h.

Note that the exact expressions for the field and current can be found easily when the lattice is located in a magnetic field perpendicular to the lattice surface (see Sec. 4). Only the phases of the functions (17) and (18) change, and the exponent γ is field-dependent. Reduction to the Riemann problem makes it possible to find the exact field distribution in two-dimensional systems belonging to a certain class.[17–19]

4. CALCULATING THE EFFECTIVE NONLINEAR CONDUCTIVITY AND THE NONLINEARITY FIELD

Now let us get down to calculating the above quantities in a planar periodic checkerboard structure as an example. To simplify matters, we limit our discussion to an external field **E** directed along one of the diagonals of the square, e.g., along ac (see Fig. 1). If this is the case, the constants c_2 and c_4 in Eqs. (17) and (18) vanish. The complex fields in adjacent cells are related by a dual transformation:[4]

$$e_2(z)=\frac{i}{\sqrt{h}}e_1^*(z^*). \tag{23}$$

By expressing the constants c_1 and c_3 in terms of the average field **E**, we can write the absolute value of the electric field in the square with conductivity σ_1 in the form

$$|e_1(z)|=\frac{1}{I}\sqrt{\frac{1+h}{2}}\,|X(z)||\mathbf{E}|, \tag{24}$$

where

$$I(\gamma)=\frac{\pi^{3/2}}{2K\cos\pi\gamma}\left[\Gamma\left(\frac{3}{4}+\frac{\gamma}{2}\right)\Gamma\left(\frac{3}{4}-\frac{\gamma}{2}\right)\right]^{-1}, \tag{25}$$

with $\Gamma(x)$ the gamma function. Combining (23), (24), and (16), we arrive at an expression for the nonlinear conductivity in a form convenient for further analysis:

$$\chi^{\text{eff}}=\frac{(h^2\chi_1+\chi_2)(h+1)^2}{8I^4h^2}\langle|X(z)|^4\rangle. \tag{26}$$

As noted in Sec. 3, the function $X(z)$ has a singularity, and at certain values of γ ($\gamma\geqslant 1/4$) the integral with respect to z in the expression for $\langle|X(z)|^4\rangle$ diverges. A general analysis of the expression for χ^{eff} shows (see Appendix A) that nonlinear conductivity grows with decreasing h and, as h tends to a certain threshold value h_c, behaves as

$$\chi^{\text{eff}}=\frac{C}{h-h_c},\quad h>h_c, \tag{27}$$

where

$$C=\frac{\pi^2(h_c^2\chi_1+\chi_2)(h_c+1)}{8I_c^4K^2\sqrt{h_c}},\quad I_c=I(\gamma_c). \tag{28}$$

For a square lattice, Eq. (20) yields $h_c=(\sqrt{2}-1)^2$. When $h\leqslant h_c$, the integral for χ^{eff} diverges. What is the nature of this divergence? A close study of the general expression (16) for χ^{eff} immediately reveals that the anomalous increase in nonlinear conductivity is due to the increase in the correlators $\langle|\mathbf{e}|^4\rangle_{1,2}$, since all the other factors in the formula are finite. The regions near ''singular'' corners, where the solutions behave asymptotically as $z^{-2\gamma}$, provide the main contribution to the integral for $\langle|\mathbf{e}|^4\rangle_{1,2}$.

Hence, to estimate the correlators as $h\to h_c$ we need only consider small neighborhoods of these points or, more precisely, the neighborhood of one singular corner in an elementary cell. For all directions of the external field, the solution in this neighborhood has the form (21). On the basis of this reasoning we conclude that

$$\langle|e|^4\rangle_{1,2}\sim\int\int|z|^{-8\gamma}dxdy\sim\int r^{-8\gamma+1}dr. \tag{29}$$

This integral is divergent for $\gamma\geqslant 1/4$. From (20) we can easily obtain the corresponding interval for h: $h\leqslant h_c$. To determine the strength of this divergence in $h-h_c$, we expand γ in the exponent of the integrand in (29) in powers of the small parameter $h-h_c$. The result is

$$\langle|e|^4\rangle_{1,2}\sim\int r^{-1+\text{const}\times(h-h_c)}dr\sim\frac{1}{h-h_c}. \tag{30}$$

As $h\to h_c$, only the nonlinear conductivity diverges. Hence the formulas for E_c and J_c show that the dependence of the critical nonlinearity field and the critical current near h_c is determined entirely by the behavior of χ^{eff}, i.e.,

$$E_c,\ J_c\propto\sqrt{h-h_c}. \tag{31}$$

Note that in a disordered film the critical field and current vanish only when $h\to 0$, and they have different critical exponents.[13]

Thus, when $h\leqslant h_c$, the system is in a nonlinear current flow regime. We now show that the same result can be obtained by solving the nonlinear equations (2), (3), and (7) directly. We start by calculating the first-order correction $\tilde{\mathbf{e}}$ to the electric field near a singular corner, and we use perturbation-theory techniques to account for the nonlinear term in (7).

Let us send the external current along the diagonal ac. In the linear approximation, the asymptotic expression for the potential in the neighborhood of corner a can be obtained from (17) and (18). In cylindrical coordinates the potential can be expressed as

$$\varphi^{(1)}=r^{\lambda}A_1\sin\lambda\left(\theta-\frac{\pi}{4}\right),$$

$$\varphi^{(2)} = -r^{\lambda} A_2 \cos \lambda\left(\theta + \frac{\pi}{4}\right), \tag{32}$$

where $\varphi^{(1)}$ and $\varphi^{(2)}$ are, respectively, the solutions for the regions $0<\theta<\pi/2$ and $-\pi/2<\theta<0$, $A_1=(E\sqrt{1+h}/\sqrt{2}I\lambda)\times(K/L)^{\lambda-1}$, $A_2=\sqrt{h}A_1$, and E is the external field strength. To first order in the nonlinearity, the equation for the potential $\tilde{\varphi}$ corresponding to the field $\tilde{\mathbf{e}}$ has the form

$$\Delta\tilde{\varphi}(r,\theta) = \frac{\chi}{\sigma}(\mathbf{e}, \operatorname{grad} \mathbf{e}^2), \tag{33}$$

where $\mathbf{e}$ is the field of the linear problem for a square lattice in a small neighborhood of the corner; the potential of this field is given by (32). The right-hand side of Eq. (33) depends on the coordinate r as $r^{-1-6\gamma}$, so we seek a solution for $\tilde{\varphi}$ in the form $\tilde{\varphi}^{(k)}(r,\theta)=r^{1-6\gamma}f_k(\theta)$. The function $f_k(\theta)$ satisfies the equation

$$\frac{d^2 f_k}{d\theta^2} + \lambda_n^2 f_k = C_k \cos\left[\lambda\left(\theta+\frac{\pi}{4}\right)+\theta_k\right], \tag{34}$$

where $\lambda=1-2\gamma$, $\lambda_n=1-6\gamma$, $C_k=4\gamma(\chi_k/\sigma_k)A_k^3\lambda^3$, $\theta_1=\pi\gamma$, $\theta_2=0$, and k labels the regions (for $k=1$, $0<\theta<\pi/2$, and for $k=2$, $-\pi/2<\theta<0$).

The solution of Eqs. (34) for each region can be found in the form of the sum of a particular solution and the general solution of the homogeneous equation corresponding to (34). The unknown coefficients in the homogeneous solution in each region are uniquely determined by the boundary conditions (3) for the field $\tilde{\mathbf{e}}^{(k)}$ and the current $\tilde{\mathbf{j}}^{(k)}$, where $\tilde{\mathbf{j}}^{(k)}=\sigma_k\tilde{\mathbf{e}}^{(k)}+\chi_k(\mathbf{e}^{(k)})^3$. We write the full formula for the solution $\tilde{\varphi}^{(k)}$:

$$\begin{aligned}
\tilde{\varphi}^{(1)} &= \frac{C_1}{32\gamma(1/4-\gamma)} r^{\lambda_n}\left\{B_1 \sin \lambda_n\left(\theta-\frac{\pi}{4}\right)\right. \\
&\quad \left. + \sin \lambda\left(\theta-\frac{\pi}{4}\right)\right\}, \\
\tilde{\varphi}^{(2)} &= \frac{C_2}{32\gamma(1/4-\gamma)} r^{\lambda_n}\left\{B_2 \cos \lambda_n\left(\theta+\frac{\pi}{4}\right)\right. \\
&\quad \left. - \cos \lambda\left(\theta+\frac{\pi}{4}\right)\right\},
\end{aligned} \tag{35}$$

where B_1 and B_2 are constants that exhibit no singularities when $\gamma\to 1/4$. This solution will help us find the region in which current flow is nonlinear.

Comparing the solutions (35) and (32), we notice that because λ_n is less than λ, as we move away from the corner the potential $\tilde{\varphi}$ decreases in comparison to φ, while as we move closer to the corner, $\tilde{\varphi}$ increases. In the region where $\tilde{\varphi}<\varphi$, the fields and current can be assumed approximately linear. But in the region where $\tilde{\varphi}>\varphi$ the field is substantially nonlinear, so that the solution $\tilde{\varphi}$ is insufficient for describing the field. As the solution (35) shows, as $\gamma\to 1/4$, the amplitude of $\tilde{\varphi}$ grows as $1/(1/4-\gamma)$.

We arrive at the same conclusion if we turn to Eq. (34), which resembles the equation for a harmonic oscillator with natural frequency $|\lambda_n|$ and a periodic driving force of frequency λ. In this equation $\gamma=1/4$ corresponds to $|\lambda_n|$ equal to λ, i.e., to resonance. Thus, even by studying only the first-order correction to the field in the nonlinearity in the neighborhood of a singular corner, we conclude that at certain values of the parameter h, more precisely, at $h=h_c$, the region where nonlinear current flow is possible broadens, which in turn may mean that the nonlinear current flow regime has set in in the entire system.

5. GENERALIZATION TO OBLIQUE LATTICES

We now investigate how the nonlinear conductivity χ^{eff} and the critical field E_c change under variations of the geometry of the structure. In calculating χ^{eff} we note that the main contribution to the correlators $\langle\mathbf{e}^4\rangle_{1,2}$ is provided by regions near the singular corners, where the asymptotic behavior of the electric field is given by (21). Hence, of the entire structure we need only small neighborhoods of the points at which several regions with different conductivities meet and in which for a given direction of the external current the correlators $\langle\mathbf{e}^4\rangle_{1,2}$ diverge. Inside each such neighborhood we calculate the field $\mathbf{e}$ and then, using the resultant solution, estimate χ^{eff}. Thus, the problem of calculating χ^{eff} for different lattices reduces to determining the electric field in a small neighborhood of singular sites with different inhomogeneity geometries.

We begin by investigating how the results for χ^{eff} and E_c change in a lattice consisting of rhombi (with angles α) with conductivities σ_1 and σ_2. To determine the field in the neighborhood of a singular site we must solve the following problem. Suppose that we have a plane partitioned into four regions with conductivities σ_1 and σ_2 by two straight lines that intersect at an angle α. We wish to solve for the current flow through such an inhomogeneous structure. The electric field is given by Eqs. (1) and (2) and condition (3) at the boundary of the regions. The solution for the scalar potential $\varphi(r,\theta)$, where r and θ are the cylindrical coordinates and the origin is chosen at the point of intersection of the straight lines, has the general form

$$\varphi^{(k)} = r^{\lambda} A_k \cos(\lambda\theta+\theta_k), \tag{36}$$

where k labels the regions formed by the intersecting straight lines, A_k and θ_k are constants, and $\lambda=1-2\gamma$.

Note that the potential $\varphi(r,\theta)$ has the form (36) even when an arbitrary number of sectors with different conductivities meet at the origin. Calculating the $\langle\mathbf{e}^4\rangle_{1,2}$, we see that the fact that the dependence of φ on the coordinate r is represented by a power function leads to a situation in which for the neighborhood of a corner with an arbitrary inhomogeneity we obtain at values of $\langle\mathbf{e}^4\rangle_{1,2}$ similar to (29), but with another γ. Hence, to estimate the contribution of the given region to χ^{eff}, we must find the corresponding exponent γ.

We now return to an examination of the field near the point where four sectors meet. Plugging (36) into the boundary condition (3), we arrive at a homogeneous system of algebraic equations (see Appendix B), and the condition that this system has a solution yields an expression for λ. In the case at hand there can be two values of λ. The value of λ specified by the equation

$$\tan\frac{\lambda\alpha}{2}\tan\frac{\lambda(\pi-\alpha)}{2}=h, \tag{37}$$

corresponds to the particular solution with the external field directed along the bisector of sectors with high conductivity, with the current ''penetrating'' the bottleneck formed by sectors with low conductivity. Of course, it is this solution with λ from (37) that leads to the divergence of the field $\mathbf{e}$ at the origin. The particular solution with the external field directed along the bisector of sectors with low conductivity corresponds to the value of λ specified by the equation

$$\tan\frac{\lambda\alpha}{2}\tan\frac{\lambda(\pi-\alpha)}{2}=\frac{1}{h}. \tag{38}$$

Using (37), we find the value h_c at which the correlators $\langle\mathbf{e}^4\rangle_{1,2}$ begin to diverge. Plugging $\lambda=1/2$ (corresponding to $\gamma=1/4$) into (37), we get

$$h_c=\tan\left(\frac{\alpha}{4}\right)\frac{1-\tan(\alpha/4)}{1+\tan(\alpha/4)}. \tag{39}$$

Clearly, as α deviates from $\pi/2$, h_c decreases and the linear range characterized by h broadens.

We now examine a site at which $2N$ sectors with conductivities σ_1 and σ_2 meet at an angle $\alpha=\pi/N$. We wish to calculate the value of h_c at which the contribution of the given region to $\langle\mathbf{e}^4\rangle_{1,2}$ leads to divergence. Doing similar calculations, we arrive at an expression determining λ (see Eq. (65) in Appendix B), from which at $\lambda=1/2$ we obtain

$$h_c=S-\sqrt{S^2-1},\quad S=8\cos^2\left(\frac{\pi}{2N}\right)-1. \tag{40}$$

In particular, for six sectors with $\alpha=\pi/3$ and $N=3$, Eq. (40) yields $h_c=5-2\sqrt{6}$. By analyzing (40) we can easily see that as N increases, h_c decreases, approaching the limit $h_c=7-4\sqrt{3}$.

The results for h_c make it possible to draw certain conclusions concerning the dependence of h_c on the geometry of the regions formed by sectors with two types of conductivity. First, among all the patterns formed by sectors with equal angles, the one with $\alpha=\pi/2$ yields the maximum value of h_c (Eq. (40) clearly demonstrates this). Second, for a given number of sectors, the pattern yielding the maximum value of h_c is the one in which the sector angles are equal. This result follows, in particular, from Eq. (39).

6. CRITICAL PROPERTIES OF LATTICES IN A MAGNETIC FIELD

Let us see how a magnetic field influences the nonlinear effects. For the linear case, in the expression for the current we must allow for the Hall term:

$$j_\alpha=\sigma_{\alpha\beta}e_\beta+\sigma^H\varepsilon_{\alpha\beta}e_\beta, \tag{41}$$

where σ^H is the Hall conductivity, and $\varepsilon_{\alpha\beta}$ is the totally antisymmetric symbol. We introduce the Hall parameter $\beta=-\sigma^H/\sigma$, which assumes the values β_1 and β_2 in the first and second cells, respectively. In a magnetic field the nonlinear conductivity becomes a tensor with symmetric and antisymmetric components. The structure of this tensor can easily be established if we use the electron temperature approximation (see Sec. 2). In this model the symmetric part of the nonlinear conductivity tensor is

$$\chi^{\rm sym}_{\alpha\beta\gamma\delta}=\chi_s\delta_{\alpha\beta}\delta_{\gamma\delta}, \tag{42}$$

with χ_s defined in (8). We can easily show that for the dissipative part of the effective nonlinear conductivity we have

$$\chi^{\rm eff}_s=\frac{\langle\chi_s\mathbf{e}^4\rangle}{\langle\mathbf{e}\rangle^4}. \tag{43}$$

Thus, to calculate $\chi^{\rm eff}_s$ we must solve the linear problem for a periodic lattice in a magnetic field. Let us make use of the problem of a square planar lattice, which we examined earlier in detail. The exact solution for the field in the linear approximation with allowance for the Hall term in such a lattice can be obtained by the same method as the one used in the absence of a magnetic field.[19] As noted earlier, calculating E_c only requires examining the asymptotic behavior of the solution near one singular corner. Introduction of a magnetic field does not change this behavior, which is that of a power function (Eqs. (21) and (22)). Here the exponent γ depends on the magnetic field:

$$\tan\pi\gamma=\frac{\sqrt{(1-h)^2+(h\beta_1-\beta_2)^2}}{2\sqrt{h}}. \tag{44}$$

Thus, the dependences of the nonlinear conductivity $\chi^{\rm eff}$ and the critical field E_c on $h-h_c$ maintain their form. The critical value h_c, which corresponds to $\gamma_c=1/4$, now depends on the magnetic field:

$$h_c=\frac{(3+\beta_1\beta_2)-\sqrt{8+4\beta_1\beta_2-(\beta_1-\beta_2)^2}}{1+\beta_1^2}. \tag{45}$$

We now discuss the physical meaning of the above expressions and compare it with what was obtained in a two-dimensional random system. If a two-dimensional disordered system is placed in an external magnetic field, allowance for the field in the linear approximation amounts to the following.[20,21] The amplitudes of the quadratic field and current fluctuations depend on the magnetic field strength B. In the simplest case this dependence is characterized by the parameter $\omega_c\tau$, where ω_c is the cyclotron frequency, and τ is the average time between collisions. In a weak field ($\omega_c\tau\ll1$) a quadratic correction to the fluctuation amplitude appears, while in a strong field ($\omega_c\tau\gg1$) the amplitude increases in proportion to the field strength B.[20,21] Here the magnetic field has no qualitative effect on the asymptotic behavior of the effective linear conductivity and correlators at small values of h: the effective conductivity decreases as $\sqrt{h}$, and the field correlators diverge as $1/\sqrt{h}$ (see Refs. 20 and 21). For a square lattice the effective conductivity and correlation functions coincide with the corresponding quantities for a random system with two equal component concentrations.[22,19]

As we found earlier, the dissipative nonlinear conductivity and the correlators $\langle\mathbf{e}^4\rangle_{1,2}$ diverge at finite values of h_c, whose dependence on the magnetic field is given by (45). In weak fields, h_c increases quadratically with field strength, while in strong fields it tends to unity. Note that at a given

finite value $h=\sigma_2/\sigma_1$, $h>h_c$, there exists a critical value of the magnetic field, B_c, at which the nonlinear conductivity and the correlators $\langle \mathbf{e}^4\rangle_{1,2}$ diverge. For $\sigma_2/\sigma_1 \neq \beta_2/\beta_1$, Eq. (45) implies that the critical value of the magnetic field corresponds to $\omega_c\tau$ being roughly equal to unity at $\beta_1 \sim \beta_2$. This value of the parameter is known to correspond to the transition from a classically weak field to a classically strong field.

The critical field B_c for a specific material can be estimated from the condition $\omega_c\tau \sim 1$. We can easily derive an expression for B_c in the case of an arbitrary angle α formed by the intersection of two straight lines if we solve the current-flow problem for such a geometry with allowance for the magnetic field. Following the line of reasoning described in Appendix B, we find that

$$B_c = \sqrt{\frac{(1+c^2)h-(1+h^2)c}{c(hr_1-r_2)^2}}, \tag{46}$$

where

$$c=\cot\frac{\alpha}{4}\cot\frac{\pi-\alpha}{4}, \quad r_1=\frac{\beta_1}{B}, \quad r_2=\frac{\beta_2}{B}.$$

In deriving (46) we assumed that the parameters r_1 and r_2 are weakly field-dependent.

Summing up, we conclude that there are values of the parameter h and the angle α at which the magnetic field assumes a critical value. In such a field there is a reversible transition to the nonlinear current flow regime.

7. CONCLUSION

A periodic checkerboard lattice is remarkable because in the general case we can establish the nature of the singularities and analyze their role in current flow. As the exact solution shows, the singularities in the higher-order field moments are determined by the asymptotic behavior of the solutions near the corners of the squares. This property makes it possible to generalize the results to lattices of a more general shape.

The important role that singularities play in random media and periodic lattices has been noted by Dreĭzin and Dykhne,[23] Bergman,[24] Dubson *et al.*,[8] and Dykhne *et al.*[25] Balagurov[5] remarked that in a square lattice the effective conductivity has a logarithmic singularity that depends on the component concentrations. As a result of our research we find that singularities of the electric field and current depend on the lattice parameters, the conductivity ratio, and the external magnetic field. The critical exponents of conductivity and second-order correlators of the lattice and of a two-dimensional random medium coincide. This places the systems in a single universal class.

We have discovered a novel property of two-dimensional lattices. We show that the linear current flow regime is unstable for a critical value h_c of the parameter h, i.e., using the linear approximation to study such systems at $h \leqslant h_c$ is unjustified. The effect results from singularities in the field and divergences in the higher-order field moments.[8,24] We have calculated the critical value h_c and the critical values of the angle parameters and the magnetic field for different types of lattices. We believe that another class of exactly solvable two-dimensional lattices exists in which the singularities can be studied exactly. Classification of such singularities in lattices, apparently, has yet to be done.

We did not examine the cases where $h \leqslant h_c$ and $B \geqslant B_c$. Finding the fields in the neighborhoods of singular angles in a structure where the linear current flow regime breaks down requires solving nonlinear equations.

Note that it may be simpler to experimentally study nonlinear effects in disordered and periodic structures as functions of the ratio of the linear conductivities and the magnetic field than to investigate the concentration dependence.[6] Indeed, the parameters h and B can easily be measured using a single specimen. For instance, for the poorly conducting components we could take a semiconductor with high photosensitivity. The parameter h can be varied, since it is proportional to light intensity. It is especially simple to take the system into the nonlinear regime by varying the magnetic field. This, as we have just seen, does not require using ultrahigh fields. Nonlinear effects can easily be observed by generating higher harmonics, which are proportional to the higher-order moments of electric current.[8,26]

Singularities similar to those studied in the present work also show up in other phenomena, e.g., in thermal and noise processes.[25]

One of the authors (A.M.S.) is especially grateful to A. M. Dykhne, A. A. Snarskiĭ, and P. M. Hui for their useful remarks. The work was supported by grants from the State Committee for Higher Education (Grant No. 95-0-7.4-173) and the Russian Fund for Fundamental Research (Grant No. 97-02-16923a).

APPENDIX A

To estimate $\langle |X(z)|^4\rangle$, we write it in the form

$$\langle |X(z)|^4\rangle = \frac{1}{2}\int\int |X(z)|^4 \frac{dz\,dz^*}{L^2}, \tag{47}$$

where integration is over the domain onto which the square $abcd$ is mapped. The analysis of the integral becomes clear-cut if we express $X(z)$ (see Eq. (19)) in terms of the Weierstrass function $\wp$:

$$X(z)=\wp(z)^{-\gamma}[\wp(z)^2-\wp_L^2]^{\gamma}, \tag{48}$$

where $\wp_L=\wp(L)$. In the integral in (47), we can transform to new variables $\wp$ and $\overline{\wp}$. The Jacobian of the transformation is

$$\left|\frac{dz}{d\wp}\right|^2 = \frac{L^2}{4K^2}|\wp(\wp^2-\wp_L^2)|^{-1}. \tag{49}$$

We can then write (47) as

$$\langle |X(z)|^4\rangle = \frac{1}{8K^2}\int\int |\wp|^{-(1+4\gamma)} \times |\wp^2-\wp_L^2|^{-(1-4\gamma)} d\wp\, d\overline{\wp}. \tag{50}$$

The integral in (50) is taken over the entire upper half-plane of the variable $\wp$. Introducing the change of variables $\wp/\wp_L=\sqrt{x}\,\exp(i\theta/2)$, we obtain

$$\langle |X(z)|^4 \rangle = \frac{\wp_L^{4\gamma-1}}{16K^2} \int_0^\infty x^{-(1+4\gamma)/2} f(x)\,dx, \tag{51}$$

where

$$f(x) = \int_0^{2\pi} (1 - 2x\,\cos\,\theta + x^2)^{(4\gamma-1)/2} d\theta. \tag{52}$$

The function $f(x)$ has the property $f(1/x) = x^{1-4\gamma} f(x)$, So we can reduce the integration with respect to x to the interval $[0,1]$, i.e.,

$$\langle |X(z)|^4 \rangle = \frac{\wp_L^{4\gamma-1}}{8K^2} \int_0^1 f(x) x^{-(4\gamma+1)/2} dx. \tag{53}$$

The function $f(x)$ can be expressed in terms of the hypergeometric function,[27] which allows writing (53) as

$$\langle |X(z)|^4 \rangle = \frac{\pi \wp_L^{4\gamma-1}}{4K^2} \int_0^1 x^{-(4\gamma+1)/2} \times F\left(\frac{1-4\gamma}{2}, \frac{1-4\gamma}{2},\ 1,\ x^2 \right) dx. \tag{54}$$

If we expand the hypergeometric function in powers of x, the first term behaves as $x^{-(4\gamma+1)/2}$, and at certain values of the parameter γ the integral in (54) diverges. Clearly, the integral diverges for $\gamma \geqslant \gamma_c$, where $\gamma_c = 1/4$. For $\gamma < \gamma_c$ and $\gamma_c - \gamma \ll 1$ the correction to the divergent term is of order $(\gamma_c - \gamma)^2$. Thus, near γ_c instead of (54) we can write

$$\langle |X(z)|^4 \rangle \simeq \frac{\pi}{8K^2} \left\{ \frac{1}{\gamma_c - \gamma} + 4\zeta(3)(\gamma_c - \gamma)^2 + \cdots \right\}, \tag{55}$$

where $\zeta(x)$ is the Riemann zeta function. The difference $\gamma_c - \gamma$ is related to the difference $h - h_c$ by

$$\gamma_c - \gamma = (h - h_c)\,\frac{h_c + 1}{8\pi} h_c^{-3/2}, \tag{56}$$

which follows from (20). Plugging (56) in (55) and then into (26), we finally arrive at (27).

APPENDIX B

Let us find the electric field in a planar medium separated into M sectors by rays. Suppose that the angle of the kth sector is α_k and the conductivity of that sector is σ_k. It is convenient to seek the electric field in the form

$$e_r = \lambda r^{\lambda-1} f(\theta), \quad e_\theta = r^{\lambda-1} f'(\theta), \tag{57}$$

where the function f_k in each sector obeys the equation

$$f_k'' + \lambda^2 f_k = 0 \tag{58}$$

and the following conditions at the boundary $\theta = \phi_k$ between sectors k and $k+1$:

$$f_k(\phi_k) = f_{k+1}(\phi_k), \quad \sigma_k f_k'(\phi_k) = \sigma_{k+1} f_{k+1}'(\phi_k). \tag{59}$$

If we write the solution of Eq. (58) in the kth sector in the form $f_k = a_k \cos\lambda\theta + b_k \sin\lambda\theta$, the boundary conditions (59) lead to a relationship between the coefficients a_k and b_k in adjacent sectors:

$$\mathbf{x}_{k+1} = \Omega_k \mathbf{x}_k, \quad \Omega_k = T^{-1}(\lambda\phi_k) H(1/h_k) T(\lambda\phi_k), \tag{60}$$

where $\mathbf{x}_k$ is a column vector with elements a_k and b_k, $T(\theta)$ is the matrix of a two-dimensional rotation through the angle θ, $H(h)$ is a diagonal matrix with diagonal elements 1 and h, and $h_k = \sigma_{k+1}/\sigma_k$. Writing the condition $\mathbf{x}_{M+1} = \mathbf{x}_1$ via (60), we arrive at the equation $(\Pi_{i=1}^M \Omega_i - I)\mathbf{x}_1 = 0$, which has a nonzero solution only if

$$\det\left(\prod_{i=1}^M \Omega_i - I \right) = 0, \tag{61}$$

with $\det \Pi_{i=1}^M \Omega_i = 1$. Equation (61) can be used to find λ.

Let us calculate λ for some typical cases. Suppose that N straight lines intersect at a single point (the origin) and partition the plane into $2N$ sectors with $\alpha = \pi/N$ and conductivities of two types, σ_1 and σ_2. For this geometry the product of the matrices Ω_i in (61) can be written

$$\prod_{i=1}^{2N} \Omega_i = U^N, \tag{62}$$

with $U = H(h)T(\lambda\alpha)H(1/h)T(\lambda\alpha)$ and $h = \sigma_2/\sigma_1$. The eigenvalues of the matrix U are $\exp(i\psi)$ and $\exp(-i\psi)$, with ψ obeying the equation

$$\cos\,\psi = \cos^2(\lambda\alpha) - \frac{1+h^2}{2h} \sin^2(\lambda\alpha). \tag{63}$$

From (61) and (62) it follows that λ can be found by noting that one of the eigenvalues of U^N is unity, i.e., from the condition $\pm\psi N = 2\pi$, or

$$\cos\,\psi = \cos\frac{2\pi}{N}. \tag{64}$$

Plugging (64) into (63), we arrive at an equation for finding λ:

$$\tan^2\left(\frac{\pi\lambda}{N} \right) = \frac{2h[1 - \cos(2\pi/N)]}{1 + 2h\,\cos(2\pi/N) + h^2}. \tag{65}$$

The values of λ tend to unity as $h \to 1$. When there are four sectors ($N = 2$), Eq. (65) yields two values of λ, which can be written $1 + 2\gamma$ and $1 - 2\gamma$, with γ defined in (20). When $N > 2$, there is only one value of λ.

Let us now examine a different structure, which consists of four sectors with conductivities σ_1 and σ_2. The sectors are formed by the intersection of two straight lines with an angle α between them. The value of λ can be found in the same way as before, with $N = 2$ in (62) and $U = H(h)T(\lambda(\pi - \alpha))H(1/h)T(\lambda\alpha)$. However, there is a simpler way of calculating λ in this case. The structure has two symmetry axes, which lie on the bisectors of sectors with different conductivities. If an external current flows along one axis, we notice that the electric field potential is symmetric with respect to this axis and antisymmetric with respect to the other axis. Such symmetry makes it possible to determine the phase values in each sector. Moreover, the problem has inversion symmetry, $\mathbf{e}(-\mathbf{r}) = \mathbf{e}(\mathbf{r})$, which ultimately allows for reduction of the complete system of boundary conditions (for an arbitrarily directed external current) to the boundary conditions only between two adjacent sectors. Taking the point of intersection of the straight lines as the coordinate origin and directing the x axis and the

external current along the bisector of the sector with angle α and conductivity σ_1, we obtain the electric potential in the form

$$\varphi^{(1)}=r^{\lambda}A_1\cos\lambda\theta,\quad \varphi^{(2)}=-r^{\lambda}A_2\sin\lambda\left(\theta-\frac{\pi}{2}\right),$$

where $k=1$ corresponds to the region $-\alpha/2<\theta<\alpha/2$, and at $k=2$ the angle θ varies within the interval $\alpha/2<\theta<\pi-\alpha/2$. For an external current flowing in this direction we need only consider the boundary conditions at the boundary between two sectors, e.g., $\theta=\alpha/2$. Plugging $\varphi^{(1)}$ and $\varphi^{(2)}$ into the boundary condition (3), we arrive at a system of algebraic equations,

$$\begin{aligned} &A_1\cos\frac{\lambda\alpha}{2}+A_2\sin\frac{\lambda(\alpha-\pi)}{2}=0,\\ &A_1\sigma_1\sin\frac{\lambda\alpha}{2}-A_2\sigma_2\cos\frac{\lambda(\alpha-\pi)}{2}=0. \end{aligned} \tag{66}$$

A nonzero solution exists only when the determinant of (66) is zero, i.e., when condition (37) holds. Equation (37) gives the value of λ, and one equation in (66) (with λ known) yields the value of A_1/A_2. In the special case $\alpha=\pi/2$, the result agrees with (20) if we put $\lambda=1-2\gamma$. We now find the second particular solution with the external current flowing along the y axis. Reasoning along similar lines, we arrive at the following expressions for the electric potential in the first and second regions:

$$\varphi^{(1)}=-r^{\lambda}B_1\sin\lambda\theta,\quad \varphi^{(2)}=r^{\lambda}B_2\cos\lambda\left(\theta-\frac{\pi}{2}\right),$$

where λ is now determined by Eq. (38).

*)E-mail: satanin@phys.unn.runnet.ru

[1] *Proceedings of the Third International Conference on Electrical Transport and Optical Properties of Inhomogeneous Media*, Physica A **207**, (1994).
[2] A. M. Dykhne, Zh. Éksp. Teor. Fiz. **59**, 110 (1970) [Sov. Phys. JETP **32**, 63 (1971)].
[3] B. Ya. Balagurov, Zh. Éksp. Teor. Fiz. **79**, 1561 (1980) [Sov. Phys. JETP **52**, 787 (1980)].
[4] Yu. P. Emets, Zh. Éksp. Teor. Fiz. **96**, 701 (1989) [Sov. Phys. JETP **69**, 397 (1989)].
[5] B. Ya. Balagurov, Zh. Éksp. Teor. Fiz. **88**, 1664 (1985) [Sov. Phys. JETP **61**, 991 (1985)].
[6] Y. Gefen, W.-H. Shih, R. B. Laibowitz, and J. M. Viggiano, Phys. Rev. Lett. **57**, 3097 (1986).
[7] L. de Arcangelis, S. Redner, and A. Coniglio, Phys. Rev. B **34**, 4656 (1986).
[8] M. A. Dubson, Y. C. Hui, M. B. Weissman, and J. C. Garland, Phys. Rev. B **39**, 6807 (1989).
[9] A. Aharony, Phys. Rev. Lett. **58**, 2726 (1987).
[10] D. Stroud and P. M. Hui, Phys. Rev. B **37**, 8719 (1988).
[11] R. Blumenfeld and D. J. Bergman, Phys. Rev. B **43**, 13 682 (1991).
[12] A. M. Satanin, S. V. Khor'kov, and A. Yu. Ugol'nikov, JETP Lett. **62**, 322 (1995).
[13] A. M. Satanin, Pis'ma Zh. Tekh. Fiz. **21**, No. 8, 44 (1995) [Tech. Phys. Lett. **21**, 652 (1995)].
[14] A. M. Satanin, V. V. Skuzovatkin, and S. V. Khor'kov, JETP Lett. **64**, 538 (1996).
[15] F. G. Bass and Yu. G. Gurevich, *Hot Electrons and Strong Electromagnetic Waves in Semiconductor Plasma and Gas Discharge* [in Russian], Nauka, Moscow (1975).
[16] I. E. Goldman and A. G. Zhdan, Fiz. Tekh. Poluprovodn. **10**, 1839 (1976) [Sov. Phys. Semicond. **10**, 1098 (1976)].
[17] L. V. Gorozhda, Yu. P. Emets, N. I. Zhukov, and É. I. Zverovich, Zh. Tekh. Fiz. **49**, 1577 (1979) [Sov. Phys. Tech. Phys. **24**, 877 (1979)].
[18] Yu. M. Vasetskiĭ, L. V. Gorozhda, and Yu. P. Emets, Zh. Tekh. Fiz. **52**, 601 (1982) [Sov. Phys. Tech. Phys. **27**, 389 (1982)].
[19] Yu. P. Emets, *Electrical Characteristics of Composites with Regular Structure* [in Russian], Naukova Dumka, Kiev (1986).
[20] A. M. Dykhne, Zh. Éksp. Teor. Fiz. **59**, 641 (1970) [Sov. Phys. JETP **32**, 348 (1971)].
[21] B. Ya. Balagurov, Fiz. Tverd. Tela (Leningrad) **20**, 3332 (1978) [Sov. Phys. Solid State **20**, 1922 (1978)].
[22] Yu. M. Vasetskiĭ and Yu. P. Emets, Zh. Prikl. Mekh. Tekh. Fiz. **5**, 117 (1979).
[23] Yu. A. Dreĭzin and A. M. Dykhne, Zh. Éksp. Teor. Fiz. **84**, 1756 (1983) [Sov. Phys. JETP **57**, 1024 (1983)].
[24] D. J. Bergman, Phys. Rev. B **39**, 4598 (1989).
[25] A. M. Dykhne, V. V. Zosimov, and S. A. Rybak, Dokl. Ross. Akad. Nauk **345**, 467 (1995) [Phys. Dokl. **40**, 613 (1995)].
[26] A. A. Snarskiĭ, Pis'ma Zh. Tekh. Fiz. **21**, No. 1, 3 (1995) [Tech. Phys. Lett. **21**, 1 (1995)].
[27] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*, Academic Press, New York (1980).

Translated by Eugene Yankovsky

Optical absorption by small metallic particles

P. M. Tomchuk and B. P. Tomchuk

Institute of Physics, Ukrainian National Academy of Sciences, 252650 Kiev, Ukraine
(Submitted 10 October 1996)
Zh. Éksp. Teor. Fiz. **112**, 661–678 (August 1997)

We study theoretically the dependence of absorption by small metallic particles on particle shape and wave polarization in the IR frequency range. We examine the electric and magnetic absorption by small particles. The particles may be either larger or smaller than the electron mean free path. We show that for asymmetric particles smaller than the mean free path the light-induced conductivity is a tensor. We also show that the total absorption and the electric-to-magnetic absorption ratio are strongly dependent on particle shape and wave polarization. Finally, we construct curves representing the dependence of the ratio of the electric and magnetic contributions to absorption on the degree of particle asymmetry for different wave polarizations. Similar curves are constructed for the ratio of the components of the light-induced conductivity tensor. © *1997 American Institute of Physics.* [S1063-7761(97)01908-2]

1. INTRODUCTION

Small metallic particles are interesting objects for physical studies. Whenever the characteristic size of such a particle becomes comparable to a physical quantity with a dimension of length (the de Broglie wavelength, the mean free path, the electromagnetic wavelength, the skin-layer depth, etc.), the small particles and their ensembles reveal new properties. For instance, when the particle size is comparable to the de Broglie wavelength, quantization of the electron spectrum becomes important. When the particles are smaller than the mean free path, there is an abrupt change in the intensity of electron–lattice energy exchange.[1–3] The absorption of light also changes. The optical properties of particles also vary depending on whether the particle is larger or smaller than the light's wavelength and the depth of the skin layer.

The sharp decrease (by several orders of magnitude[2,3]) in electron–lattice energy exchange in a small metallic particle has especially far-reaching physical consequences.[4] It is precisely because of such a decrease in energy exchange in a particle that we are able to produce, by supplying energy to the electron subsystem, hot electrons in stationary (quasistationary) conditions. Such hot electrons can be generated in islet metallic films deposited on insulator substrates by sending a current through a tunnel-coupled system of small particles[5,6] or by pointing a laser beam at the film.[7,8]

The presence of hot electrons makes it possible to explain the special features of electron and photon emission by islet metallic films illuminated by laser light.[4,9] What these features amount to is that photon emission is observed in the illumination of the specimen by carbon dioxide laser photons whose energy is 30 to 40 times lower that the electron work function for a metallic islet, while among the emitted photons there are those whose energy exceeds that of the primary photons by a factor of several tens.[9] And this all occurs at laser beam intensities at which multiphoton processes still play no role and the film is not destroyed. Nothing of the sort has been observed for the same illumination intensities in solid metallic films or bulk metals.

In this paper we study optical absorption by small metallic particles. These particles are much smaller than the light's wavelength and can be either larger or smaller than the electron mean free path. A characteristic feature of such particles is the strong dependence of absorption in the IR frequency range on the particle shape and wave polarization. For instance, the luminous power absorbed by particles of the same volume but differing in shape can differ by several orders of magnitude.[10] In addition to drawing attention for purely scientific reasons, these features of absorption by small particles are important for practical reasons. In particular, when paint containing small metallic particles is used to cover solid surfaces, the absorption and reflection of light by such surfaces can change drastically. Controlling such processes requires knowing the properties of the particles.

In space, metallic particles pose a serious hazard for spacecraft. Such particles can be destroyed by laser beams, but for this process to be successful we must know the peculiarities of optical absorption.

Studies of optical properties of small particles have a relatively long history, and the basic results are well covered in monographs (see, e.g., Ref. 11). But as for the effect of the shape of small particles on the absorption by such particles (especially particles much smaller than the mean free path), there is still a lot to be studied. We find it convenient to discuss what has been done in this field after we state the problem.

2. STATEMENT OF THE PROBLEM

Suppose that a metallic particle is exposed to an electromagnetic wave

$$\begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix} = \begin{pmatrix} \mathbf{E}^{(0)} \\ \mathbf{H}^{(0)} \end{pmatrix} \exp\{i(\mathbf{k}\cdot\mathbf{r} - \omega t)\}, \tag{1}$$

where $\mathbf{E}$ and $\mathbf{H}$ are the electric and magnetic fields of the wave, ω and $\mathbf{k}$ are the wave's frequency and wave vector, and $\mathbf{r}$ and t denote the spatial coordinates and time, respectively.

 1063-7761/97/080360-10$10.00

We assume that the wavelength $\lambda = 2\pi/k$ is considerably larger than the characteristic size of the particle. In this case the metallic particle is actually in spatially uniform, but time-varying, electric ($\mathbf{E}$) and magnetic ($\mathbf{H}$) fields. The electric field $\mathbf{E}^{(0)}e^{-i\omega t}$ induces a local potential electric field $\mathbf{E}_{\rm loc}$ inside the particle, which in turn gives rise to electric current (with current density $\mathbf{j}_e$). The magnetic field $\mathbf{H}^{(0)}e^{-i\omega t}$ induces in the particle a vortex (eddy) electric field $\mathbf{E}_{\rm ed}$, which gives rise to an eddy electric current $\mathbf{j}_m$ (here $\mathbf{j}_m$ is the corresponding current density).

As a result, the total dissipation of the wave's energy (absorption by particle) is

$$W = W_e + W_m = \frac{1}{2}\,\mathrm{Re}\int_V d\mathbf{r}(\mathbf{j}_e\cdot\mathbf{E}^*_{\rm loc} + \mathbf{j}_m\cdot\mathbf{E}^*_{\rm ed}), \tag{2}$$

where V is the particle volume.

The first term on the right-hand side of Eq. (2) corresponds to electric absorption and the second to magnetic absorption. Thus, to determine the total absorption we must know the potential electric field $\mathbf{E}_{\rm loc}$ and the eddy electric field $\mathbf{E}_{\rm ed}$, and the corresponding current densities $\mathbf{j}_e$ and $\mathbf{j}_m$.

For particles that are either larger or smaller than the electron mean free path, simultaneous electric and magnetic absorptions are known, i.e., appropriate formulas exist, only when the particles are spherical. The total energy absorbed by a spherical particle (see, e.g., Ref. 12) is

$$W = \frac{9}{8\pi}V\omega\varepsilon''\left[\frac{1}{(2+\varepsilon')^2+\varepsilon''^2} + \frac{\omega^2R^2}{90c^2}\right]|\mathbf{E}^{(0)}|^2, \tag{3}$$

where ε' and ε'' are, respectively, the real and imaginary parts of the dielectric constant, R is the particle radius, and c is the speed of light. The first term on the right-hand side of Eq. (3) describes electric absorption and the second term magnetic absorption.

If the particle is larger than the mean free path, i.e., bulk scattering is dominant, the expression for the dielectric constant of the metal has the standard form

$$\varepsilon_0 = \varepsilon' + i\varepsilon'' = 1 - \frac{\omega_p^2}{\omega^2+\nu^2} + i\,\frac{\nu}{\omega}\,\frac{\omega_p^2}{\omega^2+\nu^2}, \tag{4}$$

where ω_p is the plasma frequency, and ν is the collision rate.

We now use Eqs. (3) and (4) to estimate the relative contributions of electric and magnetic scattering to the total absorption by the particle. For example, let us take a gold particle. Then $\omega_p \approx 5\times10^{15}\mathrm{s}^{-1}$ and $\nu \approx 10^{13}\mathrm{s}^{-1}$. We assume that $R = 3\times10^{-6}$cm and that ω is the frequency of a carbon dioxide laser, i.e., $\omega \approx 2\times10^{14}\mathrm{s}^{-1}$. Then Eq. (3) yields $\varepsilon' \approx -600$ and $\varepsilon'' \approx 30$, and the magnetic-to-electric absorption ratio is

$$\frac{W_m}{W_e} = \frac{1}{90}\left(\frac{\omega R}{c}\right)^2|\varepsilon_0|^2 \approx 2. \tag{5}$$

We see that for the given set of parameters magnetic absorption is twice as large as electric. Obviously, for different parameters of the particle and a different frequency range electric absorption can be either larger or smaller than magnetic absorption. Hence, when studying the dependence of optical absorption by a small metallic particle on particle shape, we must allow for both electric and magnetic absorption. However, for asymmetric particles smaller than the mean free path we know of but a single attempt (Ref. 10) to obtain the total absorbed power. The only quantity that was studied was the light-induced conductivity $\sigma(\omega)$, defined by

$$\mathbf{j}_e = \sigma\mathbf{E}_{\rm loc}, \tag{6}$$

for spherical[13] and cylindrical[14] particles, and for particles shaped like parallelepipeds.[15] All these papers use a quantum mechanical approach, and the shape of the particles was chosen in such a way that the potential well corresponding to such a symmetry allowed for an analytic solution of the Schrödinger equation.

For particles larger than the mean free path the light-induced conductivity is known ($\sigma = \omega\varepsilon''/4\pi$), and allowing for the effect of particle shape on absorption reduces to finding $\mathbf{E}_{\rm loc}$ and $\mathbf{E}_{\rm ed}$. The local electric field $\mathbf{E}_{\rm loc}$ in the case of an ellipsoidal particle is independent of the coordinates and can easily be expressed in terms of the components of the depolarization tensor. Many researchers have used this feature to allow for the dependence of electric absorption on the particle shape in the case where bulk scattering is dominant (see, e.g., Ref. 11 and the literature cited therein).

For perfect conductivity ($\varepsilon'' \gg |\varepsilon'|$), magnetic absorption by ellipsoidal metallic particles larger than the electron mean free path was studied by Levin and Muratov.[16]

In most theoretical papers devoted to optical absorption by islet metallic films, attention is focused mainly on the effect of the interaction between the metallic particles (islets) on $\mathbf{E}_{\rm loc}$ and hence on electric absorption (see, e.g., Ref. 17 and the literature cited therein). In some cases, taking this interaction into account can indeed lead to a change in the absorbed power.[2] However, the effect of particle shape on electric and magnetic absorption, and the polarization dependence of the electric-to-magnetic absorption ratio have yet to be studied. But it is well known that, depending on these factors, the absorbed power can vary not just severalfold, but by several orders of magnitude.

3. LOCAL FIELDS

The metallic particles examined below are ellipsoidal. Such an assumption has certain advantages. First, by considering ellipsoids of different oblateness and elongation, we can obtain the majority of real particle shapes (from ''pancake'' to antenna-like). Second, finding the potential ($\mathbf{E}_{\rm loc}$) and eddy ($\mathbf{E}_{\rm ed}$) local fields for such particles is easy.

As is known,[12] for ellipsoidal particles the potential local electric field $\mathbf{E}_{\rm loc}$ induced by a uniform external electric field $\mathbf{E}^{(0)}$ is coordinate-independent. The field $\mathbf{E}_{\rm loc}$ can be linearly expressed in terms of $\mathbf{E}^{(0)}$ by employing the depolarization tensor. In terms of the principal axes of the depolarization tensor, which coincide with the principal axes of the ellipsoid, we have[11]

$$(E_{\rm loc})_j = E_j^{(0)} - 4\pi L_j P_j = E_j^{(0)} - L_j(\varepsilon-1)(E_{\rm loc})_j, \tag{7}$$

where L_j are the principal values of the components of the depolarization tensor, and $\mathbf{P}$ is the polarization vector.

For asymmetric particles smaller than the mean free path, light-induced conductivity becomes a tensor, as we will

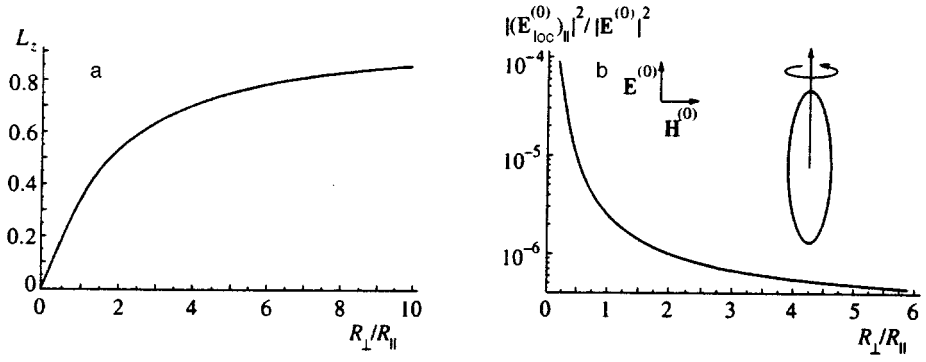


FIG. 1. Dependence of (a) the factor of depolarization along the ellipsoid's axis of revolution on the ellipsoid's semiaxis ratio, and (b) the ratio of the square of the local electric field inside the particle to the square of the wave's electric field on the ellipsoid's semiaxis ratio for the chosen wave polarization.

shortly see, and Eq. (7) needs to be modified. This will be done a little bit later. We also note that in the case of a particle ensemble, the polarization vector in a given particle is induced not only by the field of the external wave, but also by the dipoles induced by this wave in other particles.[4] Here we ignore such effects, but they can easily be incorporated into the picture.[2]

Finding $(E_{\rm loc})_j$ from (7), we get

$$|(E_{\rm loc})_j|^2=\frac{(E_j^{(0)})^2}{[1+L_j(\varepsilon'-1)]^2+(L_j\varepsilon'')^2}. \tag{8}$$

When the light-induced conductivity of the small particles becomes a tensor (about this case later), we must substitute $4\pi\sigma_{jj}/\omega$ for ε'' in (8) (σ_{jj} is the jth diagonal component of the light-induced conductivity tensor).

Now let us find the eddy local field $\mathbf{E}_{\rm ed}$. This field must obey Maxwell's equations

$$\nabla\times\mathbf{E}_{\rm ed}=-i\frac{\omega}{c}\mathbf{H}^{(0)},$$

$$\nabla\cdot\mathbf{E}_{\rm ed}=0, \tag{9}$$

augmented by the boundary condition

$$\mathbf{E}_{\rm ed}\cdot\mathbf{n}_s|_S=0, \tag{10}$$

where $\mathbf{n}_s$ is a unit vector normal to the surface S.

Before we go any further we must make the following remark. On the right-hand side of the first equation in (9), for the magnetic field inside the particle we take the external (spatially uniform) magnetic field $\mathbf{H}^{(0)}$. Such an approximation is justified if the depth δ_H of the skin layer is much greater than the characteristic particle size R:

$$\delta_H\equiv\left(\frac{\omega}{c}\,\mathrm{Im}\,\sqrt{\varepsilon}\right)^{-1}\gg R.$$

For an ellipsoidal particle, R is the semimajor axis of the ellipsoid. Below we assume that the inequality is true. This constitutes the most interesting case since the contribution of eddy currents to absorption is then at its maximum.

Bearing in mind that $\mathbf{H}^{(0)}$ is constant, we can write the solution of the system of equations (9) as

$$(E_{\rm ed})_j=\sum_{k=1}^{3}\alpha_{jk}x_k\quad(x_1=x,x_2=y,x_3=z) \tag{11}$$

and uniquely determine the set of constants α_{ij} from the system (9) and the boundary condition (10). As a result we get

$$(E_{\rm ed})_x=\frac{i\omega}{c}\left(\frac{zH_y^{(0)}}{R_z^2+R_x^2}-\frac{yH_z^{(0)}}{R_x^2+R_y^2}\right)R_x^2. \tag{12}$$

The other components of $\mathbf{E}_{\rm ed}$ can be obtained via cyclic permutations. In Eq. (12), R_y, R_y, and R_z are the semiaxes of the ellipsoid.

Knowing $\mathbf{E}_{\rm loc}$ and $\mathbf{E}_{\rm ed}$ for particles whose characteristic size is greater than the electron mean free path, it is easy, in accordance with (2), to derive a formula for the absorbed power. Note that in this case the currents are related to the field through relationships of type (6).

If to simplify matters we take an ellipsoid of revolution (with the z axis chosen as the axis of revolution), then combining (2), (8), (12), and (6) we get

$$W=V\frac{\omega\varepsilon''}{8\pi}\left\{\sum\frac{(E_j^{(0)})^2}{[1+L_j(\varepsilon'-1)]^2+(L_j\varepsilon'')^2}+\frac{\omega^2R_\perp^2}{10c^2}\right.$$
$$\left.\times(H_\parallel^{(0)})^2+\frac{\omega^2}{5c^2}\frac{R_\perp^2R_\parallel^2}{R_\perp^2+R_\parallel^2}(H_\perp^{(0)})^2\right\}, \tag{13}$$

where

$$H_\parallel^{(0)}=H_z^{(0)},\quad H_\perp^{(0)}=\sqrt{(H_x^{(0)})^2+(H_y^{(0)})^2},$$

$$R_x=R_y=R_\perp,\quad R_z=R_\parallel.$$

Equation (13) is a generalization of (3) to the case of ellipsoidal particles. Note that in (13) $E^{(0)}=H^{(0)}$. In (13) we did not express $H^{(0)}$ in terms of $E^{(0)}$ (as we did in (3)) so that the dependence of the absorbed power on wave polarization might be more graphic. The principal values of the components of the depolarization tensor for particles whose shape is that of ellipsoids of revolution are[12]

$$L_x=L_y=\frac{1}{2}(1-L_z),$$

$$L_z=\begin{cases}\dfrac{1-e_p^2}{2e_p^3}\left[\ln\dfrac{1+e_p}{1-e_p}-2e_p\right], & R_\parallel>R_\perp,\\[2ex] \dfrac{1+e_p^2}{e_p^3}[e_p-\arctan e_p], & R_\parallel<R_\perp,\end{cases} \tag{14}$$

where $e_p^2=|1-R_\perp^2/R_\parallel^2|$.

Figure 1a depicts the dependence of L_z on the ellipsoid semiaxis ratio $R_\perp/R_\parallel$. The components L_j of the depolarization tensor can vary between zero to unity. We see that the denominator in (8) contains L_j^2 as a cofactor of $(\varepsilon'-1)^2$. As the above estimates show, in the frequency range of a carbon dioxide laser, $\varepsilon'^2\approx 4\times 10^4$. Consequently, $|\mathbf{E}_{\rm loc}|^2$ can strongly depend on particle shape. Figure 1b depicts the dependence of $|\mathbf{E}_{\rm loc}|^2/|\mathbf{E}^{(0)}|^2$ on the semiaxis ratio $R_\perp/R_\parallel$ for the case where $\mathbf{E}^{(0)}$ is directed along the ellipsoid's axis of revolution. We see that the ratio of the square of the local field to the square of the external field can vary by several orders of magnitude. According to (13), this means that electric absorption by metallic particles can vary by several orders of magnitude, depending on particle shape and wave polarization.

This fact has important consequences. As we noted earlier (see the Introduction), when an islet metallic film is illuminated by laser light, the electron gas of the metallic particles heats up. This leads to electron and photon emission. In particular, electron emission here is thermionic, which according to Richardson's law is proportional to $\exp\{-\varphi/kT_e\}$, where φ is the electron work function, and T_e is the electron temperature. Electron emission is strongly dependent on the electron temperature T_e, which is determined by the energy absorbed by a particle. The absorbed energy W in turn is also strongly dependent on particle shape and wave polarization. In this situation, introducing average (effective) absorption cross sections of some kind is absolutely out of the question. The particles that absorb the most light are the electrons that emit light. In other words, these phenomena are determined primarily by particles with maximum absorption and not by particles with some average (''effective'') absorption, which constitute the majority in an islet film.

4. ELECTRON DISTRIBUTION FUNCTION

In Sec. 3 we determined the local fields and at the same time derived a general expression for the power absorbed by an ellipsoidal metallic particle in the case of bulk scattering (i.e., for particles larger than the mean free path). Here we focus on the case of particles smaller than the mean free path. Incidentally, the method we now develop can also be applied to particles larger than the mean free path.

Thus, knowing the potential and local eddy electric fields in the particle, to determine the absorbed power we must also derive expressions for the high-frequency currents induced by these fields.

By definition, the current density is

$$\mathbf{j}=\frac{2e}{(2\pi\hbar)^3}\int \mathbf{v}f d^3(mv)=\frac{2m^3e}{(2\pi\hbar)^3}\int \mathbf{v}f(\mathbf{v})d^3v, \quad (15)$$

where $f(\mathbf{v})$ is the electron distribution function over the velocities $\mathbf{v}$, and e and m are electron charge and mass. In the presence of local fields, the distribution function can be represented as a sum of two terms,

$$f(\mathbf{v})=f_0(\varepsilon)+f_1(\mathbf{v}),$$

where $f_0(\varepsilon)$ is the Fermi distribution function, which depends only on the energy $\varepsilon=mv^2/2$, and $f_1(\mathbf{v})$ is a correction generated by the local fields.

We can find $f_1(\mathbf{v})$ by solving the appropriate kinetic equation. In the linear approximation in the electromagnetic field, the kinetic equation has the form

$$(\nu-i\omega)f_1+\mathbf{v}\cdot\frac{\partial f_1}{\partial \mathbf{r}}+e\mathbf{F}\cdot\mathbf{v}\frac{\partial f_0}{\partial\varepsilon}=0. \quad (16)$$

Here we have allowed for the fact that $f_1\propto e^{-i\omega\tau}$; ν is the bulk collision rate, and

$$\mathbf{F}=\mathbf{E}_{\rm loc}^{(0)}+\mathbf{E}_{\rm ed}^{(0)}. \quad (17)$$

Equation (16) must be augmented by boundary conditions for f_1. For these we take, as is often done, the conditions for diffuse electron scattering at the boundary, i.e.,

$$f_1(\mathbf{r},\mathbf{v})|_S=0 \quad \text{for} \quad v_n<0, \quad (18)$$

where v_n is the velocity component normal to the surface S.

Operating on these assumptions, Lesskin *et al.*[18] studied magnetic scattering by a spherical metallic particle. To solve Eq. (16) with the boundary conditions (18) we employ the method of characteristic curves, which demonstrated its effectiveness in Ref. 18. But for ellipsoidal particles the method used in Ref. 18 needs to be modified. The essence of this modification will be clarified later.

Thus, we transform to a deformed system of coordinates in which the original particle ellipsoid,

$$\sum_{i=1}^{3}\frac{x_i^2}{R_i^2}=1 \quad (19)$$

becomes a sphere (of radius R). In other words, we assume that

$$x_i=\frac{x_i'}{\gamma_i}, \quad \gamma_i=\frac{R}{R_i}, \quad R=(R_1R_2R_3)^{1/3}, \quad \gamma_1\gamma_2\gamma_3=1. \quad (20)$$

Under such a deformation the shape of the particle changes, but not its volume. This means that the electron number density remains unchanged, and so does the normalization of the function f.

In the deformed system of coordinates, Eq. (16) and the boundary conditions (18) acquire the form

$$(\nu-i\omega)f_1+\mathbf{v}'\cdot\frac{\partial f_1}{\partial \mathbf{r}'}+e\mathbf{F}(\mathbf{r}')\cdot\mathbf{v}\frac{\partial f_0}{\partial\varepsilon}=0, \quad (21)$$

$$f_1(\mathbf{r}',\mathbf{v}')|_{r'=R}=0 \quad \text{for} \quad \mathbf{r}'\cdot\mathbf{v}'<0. \quad (22)$$

In (21) and (22) we also introduced the ''deformed'' velocity components

$$v_i'=\gamma_i v_i. \quad (23)$$

Equation (21) for the characteristic curves has the form

$$\frac{dx_i'}{v_i'}=-\frac{df_1}{\tilde{\nu}f_1}=dt', \quad \tilde{\nu}\equiv\nu-i\omega, \quad (24)$$

which implies that

$$\mathbf{r}'=\mathbf{v}'t'+\mathbf{R}, \quad (25)$$

where $\mathbf{R}$ is the radius vector whose tip is at the given point of the sphere from which the trajectory begins. Here the parameter t' can formally be considered the ''time'' of motion of the electron along the trajectory.

If in (25) we transfer $\mathbf{v}'t'$ to the left-hand side and square the resulting equation, the solution of this new (scalar) equation can be written

$$t'=\frac{1}{v'^2}[\mathbf{r}'\cdot\mathbf{v}'+\sqrt{(R^2-r'^2)v'^2+(\mathbf{r}'\cdot\mathbf{v}')^2}]. \tag{26}$$

The characteristic curve (26) depends only on the absolute value of R and not on the orientation of $\mathbf{R}$. Such independence of the characteristic curve from the position of a point on the surface was achieved by transforming to the coordinates (20).

From (26) we also see that $t'=0$ at $r'=R$. Bearing this in mind, we can use (24) to find an f_1 that satisfies Eq. (21) and the boundary condition (22):

$$f_1=-\frac{\partial f_0}{\partial\varepsilon}\int_0^t d\tau\,\exp\{-\tilde{\nu}(t'-\tau)\}e\mathbf{v}\cdot\mathbf{F}(\mathbf{r}'-\mathbf{v}'(t'-\tau)). \tag{27}$$

Allowing for the coordinate dependence of $\mathbf{F}$ (see (17) and (12)), from (27) we obtain

$$f_1=-e\frac{\partial f_0}{\partial\varepsilon}\left\{\mathbf{v}\cdot\mathbf{E}_{\rm loc}^{(0)}+\sum_{i,j=1}^{3}\alpha_{ij}v_i\left[\frac{x_j'}{\gamma_j}+v_j\frac{\partial}{\partial\tilde{\nu}}\right]\right\}\times\frac{1-\exp\{-\tilde{\nu}t'\}}{\tilde{\nu}}. \tag{28}$$

If initially the particle is spherical, then $\alpha_{ij}=-\alpha_{ji}$, and the last term in (28) vanishes.

5. ELECTRIC ABSORPTION

Combining (28), (14), and (2), we obtain an expression for the electric absorption:

$$W_e=\frac{e^2m^3R^3}{(2\pi\hbar)^3}\,\mathrm{Re}\left[\frac{1}{\tilde{\nu}}\int\left|\mathbf{v}\cdot\mathbf{E}_{\rm loc}^{(0)}\right|^2\delta(\varepsilon-\mu)\times(1-\exp\{-\tilde{\nu}t'\})d^3r'd^3v\right], \tag{29}$$

where μ is the Fermi energy, and where we allowed for the fact that

$$\frac{\partial f_0}{\partial\varepsilon}\approx-\delta(\varepsilon-\mu).$$

Allowing for the form of t' (according to (26)), it is convenient to integrate with respect to $\mathbf{r}'$ in (29) by directing the z' axis along the vector $\mathbf{v}'$ and introducing two new variables,

$$\zeta=\frac{r'}{R},\quad \eta=\frac{v'}{R}t'. \tag{30}$$

As a result we obtain

$$\begin{aligned}\int(1-\exp\{\tilde{\nu}t'\})d^3r' &= 2\pi\int_0^R dr'r'^2\int_0^\pi d\theta\,\sin\theta\\ &\times(1-\exp\{-\tilde{\nu}t'\})\\ &=R^3\int_0^1 d\zeta\zeta^2\int_{1-\zeta}^{1+\zeta}d\eta\frac{\eta^2-\zeta^2+1}{2\eta^2\zeta}\\ &\times\left(1-\exp\left\{-\frac{\tilde{\nu}R\eta}{v'}\right\}\right)\\ &=\frac{R^3}{2}\int_0^2\frac{d\eta}{\eta^2}\\ &\times\left(1-\exp\left\{-\frac{\tilde{\nu}R\eta}{v'}\right\}\right)\int_{|\eta-1|}^1 d\zeta\zeta\\ &\times(\eta^2-\zeta^2+1).\end{aligned}$$

Further calculation of the integral is easy, and as a result from (29) we obtain

$$W_e=\frac{\pi e^2m^3R^3}{(2\pi\hbar)^3}\,\mathrm{Re}\left[\frac{1}{\tilde{\nu}}\int d^3v\left|\mathbf{v}\cdot\mathbf{E}_{\rm loc}^{(0)}\right|^2\delta(\varepsilon-\mu)\psi(q)\right], \tag{31}$$

where we have introduced

$$\psi(q)=\frac{4}{3}-\frac{2}{q}+\frac{4}{q^3}-\frac{4}{q^2}\left(1-\frac{1}{q}\right)\exp\{-q\},$$

$$q\equiv q_1-iq_2=\frac{2\tilde{\nu}}{v'}R=\frac{2\nu}{v'}R-i\frac{2\omega}{v'}R. \tag{32}$$

Equation (31) gives the general form of the electric absorption by an ellipsoidal metallic particle for an arbitrary bulk-to-surface scattering ratio.

The above case of bulk scattering (the Drude case) follows from (31) when $q\gg1$. Then, according to (32), $\psi(q)\approx4/3$, and from Eq. (31) we obtain for the electric absorption

$$W_e\approx V\frac{e^2n}{m}\frac{\nu}{\nu^2+\omega^2}\frac{|\mathbf{E}_{\rm loc}^{(0)}|^2}{2}\equiv V\sigma(\omega)\frac{|\mathbf{E}_{\rm loc}^{(0)}|^2}{2}, \tag{33}$$

where n is the electron concentration, which can be expressed in terms of the Fermi velocity v_F or the Fermi energy μ,

$$n=\frac{8\pi}{3}\frac{(mv_F)^3}{(2\pi\hbar)^3},\quad v_F=\sqrt{\frac{2\mu}{m}}. \tag{34}$$

Clearly, Eq. (33) corresponds to the first term in (13).

We now analyze the situation in which the particle is smaller than the mean free path, and hence surface scattering is dominant. This corresponds to

$$q_1=\frac{2\nu}{v'}R\ll1. \tag{35}$$

As for $q_2\equiv2\omega R/v'$, when surface scattering is dominant, this parameter can be either larger or smaller than unity. It is of interest then to study the two limits

$$q_2=\frac{2\omega}{v'}R\gg 1, \tag{36a}$$

$$q_2\ll 1. \tag{36b}$$

The case (36a) corresponds to high-frequency surface scattering, and (36b) to low-frequency surface scattering.

If we ignore bulk scattering ($q_1\to 0$) and assume that q_2 is arbitrary, then (32) yields

$$\mathrm{Re}\left[\frac{1}{\tilde{\nu}}\psi(q)\right]\approx\frac{1}{\omega}\left[\frac{2}{q_2}-\frac{4}{q_2^2}\sin q_2+\frac{4}{q^3}(1-\cos q_2)\right]. \tag{37}$$

This expression is present in (31). We see that terms oscillating as functions of particle size have emerged. Such oscillation effects in spherical particles were studied by Austin and Wilkinson[19] for electric absorption, and by Lesskin *et al.*[18] for magnetic absorption. These effects, which are moderate by themselves, are even less important for asymmetric particles. The reason is that the ''deformed'' velocity v', which enters into the expression for q_2, is angle-dependent. In view of this, the integration over angles smooths out the oscillation effects. Furthermore, Eq. (37) implies that these oscillations can exist only when $q_2=2R\omega/v'\approx 1$, i.e., when the electron transit time from wall to wall, $2R/v'$, coincides with the period of the electromagnetic wave. In the limiting cases given by (36), these effects are negligible.

Let us start with the high-frequency case (36a). For $q_2\gg 1$ and $q_1\ll 1$ we have

$$\mathrm{Re}\left[\frac{1}{\tilde{\nu}}\psi(q)\right]\approx\frac{2}{\omega q_2},$$

which in accordance with (31) yields

$$W_e\approx\frac{\pi e^2m^3R^2}{(2\pi\hbar)^3\omega^2}\int d^3v\,v'|\mathbf{v}\cdot\mathbf{E}^{(0)}_{\mathrm{loc}}|^2\delta(\varepsilon-\mu). \tag{38}$$

To study the dependence of absorption on particle shape, we need only consider an ellipsoid of revolution. In this case

$$v'=R\sqrt{\frac{v_\perp^2}{R_\perp^2}+\frac{v_\parallel^2}{R_\parallel^2}}, \tag{39}$$

where $v_\perp$ and $v_\parallel$ are the electron velocity components perpendicular and parallel to the axis of revolution. With allowance for (39), the integral in (38) can easily be calculated:

$$W_e=V\frac{ne^2}{m\omega^2}\frac{v_F}{R_\perp}\frac{9}{16}[\varphi_\perp|E^{(0)}_{\perp,\mathrm{loc}}|^2+\varphi_\parallel|E^{(0)}_{\parallel,\mathrm{loc}}|^2]. \tag{40}$$

Here $\varphi_\perp$ and $\varphi_\parallel$ are functions of the ellipsoid's eccentricity (we note once more that $e_p^2=|1-R_\perp^2/R_\parallel^2|$):

$$\varphi_\perp=\frac{1}{2}\begin{cases}\frac{1}{2}\left(1+\frac{1}{2e_p^2}\right)\sqrt{1-e_p^2}+\frac{1}{e_p}\left(1-\frac{1}{4e_p^2}\right)\\ \quad\times\arcsin e_p, \quad R_\perp<R_\parallel,\\ \frac{1}{2}\left(1-\frac{1}{2e_p^2}\right)\sqrt{1+e_p^2}+\frac{1}{e_p}\left(1+\frac{1}{4e_p^2}\right)\\ \quad\times\ln(\sqrt{1+e_p^2}+e_p), \quad R_\perp>R_\parallel,\end{cases} \tag{41}$$

$$\varphi_\parallel=\begin{cases}\frac{1}{2}\left(1-\frac{1}{2e_p^2}\right)\sqrt{1-e_p^2}+\frac{1}{4e_p^3}\arcsin e_p,\\ \quad R_\perp<R_\parallel,\\ \frac{1}{2}\left(1+\frac{1}{2e_p^2}\right)\sqrt{1+e_p^2}-\frac{1}{4e_p^3}\ln(\sqrt{1+e_p^2}+e_p),\\ \quad R_\perp>R_\parallel.\end{cases} \tag{42}$$

Having the general expression (40) for the electric absorption of an ellipsoidal metallic particle in the case of high-frequency scattering, we can easily find the components of the light-induced conductivity tensor. To this end, we write the expression for electric absorption in terms of the principal values σ_{jj} of the conductivity tensor,

$$W_e=V\frac{1}{2}\sum_{j=1}^{3}\sigma_{jj}|E^{(0)}_{j,\mathrm{loc}}|^2, \tag{43}$$

and compare it with (40). As a result we have

$$\sigma_{xx}=\sigma_{yy}\equiv\sigma_\perp=\frac{ne^2}{m\omega^2}\frac{v_F}{R_\perp}\frac{9}{8}\varphi_\perp,$$

$$\sigma_{zz}=\sigma_\parallel=\frac{ne^2}{m\omega^2}\frac{v_F}{R_\perp}\frac{9}{8}\varphi_\parallel. \tag{44}$$

The case of a spherical particle follows from (44) as $e_p\to 0$. Allowing for (41) and (42), we find that

$$\sigma_\perp=\sigma_\parallel=\frac{ne^2}{m\omega^2}\frac{3}{4}\frac{v_F}{R}. \tag{45}$$

Comparing (44) and (45), we conclude that the light-induced conductivity of metallic particles smaller than the mean free path is a scalar quantity only when the particles are symmetric.

In the general case of asymmetric particles, the light-induced conductivity becomes a tensor whose components depend on particle shape. Figure 2a depicts the dependence

of $\sigma_\perp/\sigma_\parallel$ on the ellipsoid's semiaxis ratio $R_\perp/R_\parallel$. Equation (44) was used to build the curves. We see that the components of the light-induced conductivity tensor differ considerably, depending on the degree of particle asymmetry.

Comparing (45) with the expression for the conductivity that follows from (33) for $\omega \gg \nu$, we see that in the case of a spherical particle, the expression for the electric conductivity dominated by surface scattering can be obtained from a similar expression for the Drude case by formally substituting $3v_F/4R$ for ν. This method is widely used in the literature. However, in the case of asymmetric particles, this procedure leads to incorrect results. The appearance in (40) of the factor $v_F/R_\perp$, which has the formal meaning of ''transit frequency,'' is due to the fact that the particle volume V can be expressed as a separate factor. But when surface scattering dominates, the absorbed power is proportional to the surface area of the particle. This can easily be seen by using (40)–(42) to derive simple analytic expressions for highly elongated and highly flattened ellipsoids:

$$W_e \approx \frac{9\pi}{128} V \frac{ne^2}{m\omega^2} \frac{v_F}{R_\parallel} \left(\frac{3}{2} \left| (E^{(0)}_{\rm loc})_\perp \right|^2 + \left| (E^{(0)}_{\rm loc})_\parallel \right|^2 \right), \quad R_\perp \ll R_\parallel, \tag{46}$$

$$W_e \approx \frac{9}{32} V \frac{ne^2}{m\omega^2} \frac{v_F}{R_\parallel} \left(\frac{1}{2} \left| (E^{(0)}_{\rm loc})_\perp \right|^2 + \left| (E^{(0)}_{\rm loc})_\parallel \right|^2 \right), \quad R_\perp \gg R_\parallel. \tag{47}$$

The factors $v/R_\perp \approx R_\parallel R_\perp$ in (46) and $V/R_\parallel \approx R_\perp^2$ in (47) are simply the surface areas of the corresponding ellipsoids in the specified limits.[1)]

Let us now examine the case (36b) of low-frequency surface scattering, i.e.,

$$q_1 \ll q_2 \ll 1. \tag{48}$$

Here the frequency of the electromagnetic wave is much higher than the bulk collision rate but is much lower than the frequency of transit from wall to wall. If condition (48) is met, we have $\psi(q) \approx q/2$, and from (31) we obtain

$$W_e \approx \frac{\pi e^2 m^3 R^4}{(2\pi\hbar)^3} \int \frac{d^3v}{v'} |\mathbf{v}\cdot\mathbf{E}^{(0)}_{\rm loc}|^2 \delta(\varepsilon-\mu). \tag{49}$$

After evaluating the integral in (49) we can transform the result to (43), where instead of (44) we will have

$$\sigma_\perp = \frac{9}{8} \frac{ne^2}{m} \frac{R_\perp}{v_F} \begin{cases} -\dfrac{1}{2e_P^2} \sqrt{1+e_p^2} + \dfrac{1}{e_p} \left(1 + \dfrac{1}{2e_p^2} \right) \\ \quad \times \ln(\sqrt{1+e_p^2}+e_p), \quad R_\perp > R_\parallel, \\ \dfrac{1}{2e_P^2} \sqrt{1-e_p^2} + \dfrac{1}{e_p} \left(1 - \dfrac{1}{2e_p^2} \right) \arcsin e_p, \\ \quad R_\perp < R_\parallel, \end{cases} \tag{50}$$

$$\sigma_\parallel = \frac{ne^2}{m} \frac{R_\perp}{v_F} \begin{cases} \dfrac{1}{e_p^2} \sqrt{1+e_p^2} - \dfrac{1}{e_p^3} \ln(\sqrt{1+e_p^2}+e_p), \quad R_\perp > R_\parallel, \\ -\dfrac{1}{e_p^2} \sqrt{1-e_p^2} + \dfrac{1}{e_p^2} \arcsin e_p, \qquad R_\perp < R_\parallel. \end{cases} \tag{51}$$

Formula (49) for the absorbed power acquires a simple analytic form for the limits of highly flattened and highly elongated ellipsoids:

$$W_e \approx \frac{9}{16} V \frac{ne^2}{m} \frac{R_\parallel}{v_F} \left\{ \left[\ln\left(2 \frac{R_\perp}{R_\parallel} \right) - \frac{1}{2} \right] \left| E^{(0)}_{\perp,\rm loc} \right|^2 + \left| E^{(0)}_{\parallel,\rm loc} \right|^2 \right\}, \quad R_\perp \gg R_\parallel, \tag{52}$$

$$W_e \approx V \frac{ne^2}{m} \frac{R_\perp}{v_F} \left(\frac{1}{2} \left| E^{(0)}_{\perp,\rm loc} \right|^2 + \left| E^{(0)}_{\parallel,\rm loc} \right|^2 \right), \quad R_\perp \ll R_\parallel, \tag{53}$$

In addition, at $R_\perp = R_\parallel$, Eq. (49) yields the well-known result for a spherical particle:

$$W_e \approx \frac{3}{8} V \frac{ne^2}{m} \frac{R}{v_F} |E^{(0)}_{\rm loc}|^2. \tag{54}$$

Figure 2b depicts the dependence of $\sigma_\perp/\sigma_\parallel$ on the ellipsoid's semiaxis ratio $R_\perp/R_\parallel$ constructed from Eqs. (50) and (51). Comparing Figs. 2a and b, we see that the effect of particle asymmetry on the ratio of the components of the conductivity tensor differs not only quantitatively but also qualitatively in the high- and low-frequency cases (provided that surface scattering is dominant).

6. MAGNETIC ABSORPTION

Magnetic absorption is given by the second term in (2). Combining (11), (15), and (28), we obtain an expression for the magnetic absorption,

$$W_m = \frac{e^2 m^3}{(2\pi\hbar)^3} \,\mathrm{Re} \int d^3r' d^3v \, \delta(\varepsilon-\mu) \times \sum_{(1)}^{(3)} \frac{\alpha^*_{lk}\alpha_{ij}}{\gamma_l\gamma_k\gamma_i\gamma_j} v'_l x'_k v'_i x'_j \frac{1-\exp\{-\tilde{\nu}t'\}}{\tilde{\nu}}, \tag{55}$$

where summation is over all indices from 1 to 3. To calculate the integral with respect to r', we direct the z' axis along the

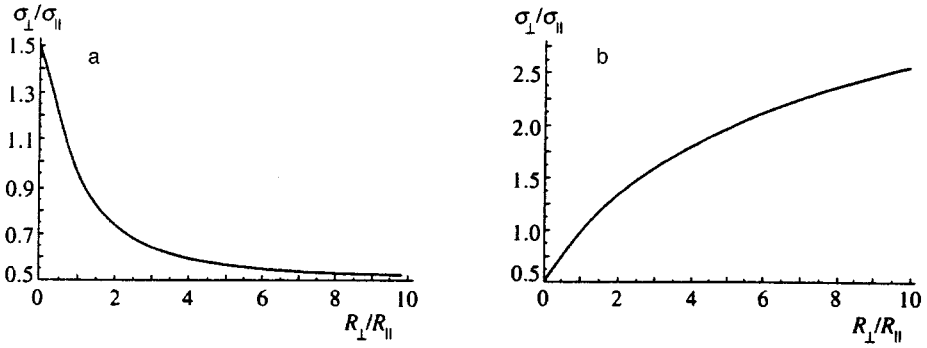


FIG. 2. Ratio of the conductivity perpendicular to the ellipsoid's axis of revolution ($\sigma_\perp$) to the conductivity along the axis ($\sigma_\parallel$) as a function of the ellipsoid's semiaxis ratio: (a) $\omega \gg v_F/R_\perp, v_F/R_\parallel$; (b) $\omega \ll v_F/R_\perp, v_F/R_\parallel$.

vector $\mathbf{r}'$. Then, according to (26), t' is independent of the angle φ' (in the plane perpendicular to $\mathbf{v}'$). Hence in (55) we can first integrate with respect to φ'. It can be shown that

$$\int_0^{2\pi} d\varphi' x_j' x_k' = 2\pi r'^2 \left[\frac{v_j' v_k'}{v'^2} + \frac{\sin^2 \theta'}{2} \left(\delta_{jk} - 3 \frac{v_j' v_k'}{v'^2} \right) \right], \tag{56}$$

where θ' is the angle between $\mathbf{r}'$ and $\mathbf{v}'$.

After (56) is plugged into (55), the calculation becomes similar to the one used in calculating the expression (29) for the electric absorption. As a result, Eq. (55) becomes

$$W_m = \frac{\pi e^2 m^3 R^5}{2(2\pi\hbar)^3} \operatorname{Re}\left\{ \frac{1}{\tilde{\nu}} \int d^3v\, \delta(\varepsilon-\mu) \times \left[\sum_{(1)}^{(3)} \frac{|\alpha_{ij}|^2}{\gamma_i^2 \gamma_j^2} v_i'^2 \psi_1(q) + 2 \times \sum_{(1)}^{(3)} \frac{|\alpha_{ij}+\alpha_{ji}|^2}{\gamma_i^2 \gamma_j^2} \frac{v_i'^2 v_j'^2}{v'^2} \psi_2(q) \right] \right\}, \tag{57}$$

where

$$\psi_1 = \frac{8}{15} - \frac{1}{q} + \frac{4}{q^3} - \frac{24}{q^5} - 8\left(\frac{1}{q^3} + \frac{3}{q^4} + \frac{3}{q^5} \right) \exp\{-q\},$$

$$\psi_2 = \frac{2}{5} - \frac{1}{q} + \frac{8}{3q^2} - \frac{6}{q^3} + \frac{32}{q^5} - 2\left(\frac{1}{q^2} + \frac{5}{q^3} + \frac{16}{q^4} + \frac{16}{q^5} \right) \times \exp\{-q\} - \frac{3}{4}\psi_1(q). \tag{58}$$

Equation (57) determines the magnetic absorption by a particle in general form for an arbitrary ratio of the bulk and surface contributions. For spherical particles, the last term on the right-hand side of Eq. (57) vanishes, since in this case $\alpha_{ij} = -\alpha_{ji}$.

Information about the scattering mechanism is contained in the quantities $q = 2R\tilde{\nu}/v'$ and $\tilde{\nu} \equiv \nu - i\omega$.

From (57) we can derive simple analytic expressions in the limits of pure bulk scattering and pure surface scattering. In the first case ($|q| \gg 1$) we arrive at the known expression for magnetic absorption determined by the second and third terms on the right-hand side of Eq. (13).

When surface scattering is dominant, the high-frequency case is the most interesting, since the eddy field E_{ed} is proportional to the frequency, with the result that the relative role of magnetic absorption grows with frequency. We therefore assume that $q_1 \ll 1$ and $q_2 \gg 1$. Then from (57) it follows that

$$W_m \approx \frac{\pi e^2 m^3 R^4}{4(2\pi\hbar)^3 \omega^2} \int d^3v\, \delta(\varepsilon-\mu) \times \left(\sum_{(1)}^{(3)} \frac{|\alpha_{ij}|^2}{\gamma_i^2} v_i^2 v' + \frac{1}{2} \sum_{(1)}^{(3)} \left| \alpha_{ij} + \alpha_{ji} \right|^2 \frac{v_i^2 v_j^2}{v'} \right). \tag{59}$$

Again, if we limit ourselves to an ellipsoid of revolution, the integrals in (59) can easily be evaluated and we obtain

$$W_m \cong \frac{9}{128} V \frac{e^2 n v_F}{mc^2} R_\perp \left[\varphi_\perp (H_\parallel^{(0)})^2 + \frac{R_\parallel^4}{(R_\perp^2 + R_\parallel^2)^2} \Phi (H_\perp^{(0)})^2 \right]. \tag{60}$$

Here, in addition to the function $\varphi_\perp$ defined in (41), we have introduced a new function Φ:

$$\Phi = \begin{cases} \left(2 - \frac{1}{4e_p^2} - e_p^2 \right) \sqrt{1-e_p^2} + \frac{1}{2e_p} \left(1 + \frac{1}{2e_p^2} \right) \arcsin e_p, \\ \quad R_\perp < R_\parallel, \\ \left(2 + \frac{1}{4e_p^2} + e_p^2 \right) \sqrt{1+e_p^2} + \frac{1}{2e_p} \\ \quad \times \left(1 - \frac{1}{2e_p^2} \right) \ln\left(\sqrt{1+e_p^2} + e_p \right), \quad R_\perp > R_\parallel. \end{cases} \tag{61}$$

For spherical particles (i.e., as $e_p \to 0$), the result obtained by Lesskin *et al.* follows from (60):[18]

$$W_m \approx \frac{3}{64} V \frac{e^2 n v_F}{mc^2} R (H^{(0)})^2. \tag{62}$$

Equation (60) acquires a simple analytic form for the limits of highly elongated and highly flattened ellipsoids:

$$W_m \approx \frac{9}{128} \frac{3\pi}{8} V \frac{e^2 n v_F}{mc^2} R_\perp \left[\frac{1}{2} (H_\parallel^{(0)})^2 + (H_\perp^{(0)})^2 \right], \quad R_\perp \ll R_\parallel, \tag{63}$$

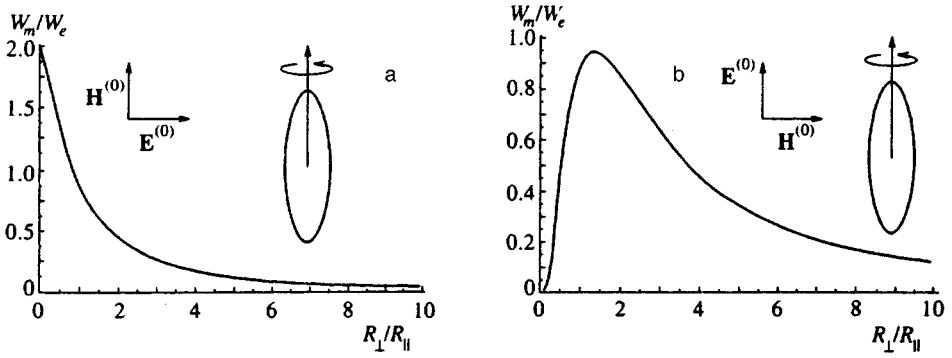


FIG. 3. Dependence of the magnetic-to-electric absorption ratio on the ellipsoid's semiaxis ratio for the chosen wave polarization.

$$W_m \approx \frac{9}{128} V \frac{e^2 n v_F}{mc^2} R_\parallel \left[\left(\frac{R_\perp}{2R_\parallel} \right)^2 (H_\parallel^{(0)})^2 + (H_\perp^{(0)})^2 \right], \quad R_\perp \gg R_\parallel . \tag{64}$$

From (63) it follows that in the case of a highly elongated ellipsoid, magnetic absorption is twice as high when the magnetic field is perpendicular to the axis of revolution than when it is parallel to that axis. The situation is similar for bulk scattering.

Earlier we estimated the relative contribution of the electric and magnetic terms of spherical particles to absorption (see Eq. (5)). Now, having the expressions for the electric (Eq. (40)) and magnetic (Eq. (60)) absorption by asymmetric particles, we can return to that problem.

For an asymmetric particle, the ratio of the electric and magnetic contributions to absorption (at fixed frequency) is strongly dependent on the degree of particle asymmetry and wave polarization. Figures 3a and b depict the dependence of W_m/W_e on the ellipsoid's semiaxis ratio for two different polarizations. Comparing the two figures, we see that these curves differ strongly not only quantitatively but also qualitatively.

7. CONCLUSION

In this paper we have derived, for the first time, expressions for electric and magnetic absorption by nonspherical particles smaller than the electron mean free path. We have found that for small asymmetric particles, their electric and magnetic absorption can vary by several orders of magnitude under particle shape variations with the volume remaining constant. Such drastic variations in absorption can also occur under variations of wave polarization.

Simple analytic formulas have been derived for highly elongated and highly flattened particle shapes.

We have also established that for nonspherical metallic particles smaller than the electron mean free path, the light-induced conductivity is a tensor, in contrast to the Drude case. We have found the components of the conductivity tensor for particles in the form of an ellipsoid of revolution. We have studied the dependence of these components on the degree of particle asymmetry.

In conclusion, we note that most theoretical investigations into the optical properties of islet metallic films are devoted to the mutual effect of the particles on local fields and electric absorption. In recent years the reflection of IR radiation from a layer of small metallic particles has also been studied (see, e.g., Ref. 21). Allowance for the mutual effect of particles can indeed strongly influence the values of the local fields. Estimates have shown[2] that in favorable cases, allowance for this mutual effect can alter the local field inside a given particle severalfold. However, our research has shown that allowing for particle shape together with considering the electric and magnetic absorption can change the total absorption by several orders of magnitude. These factors, therefore, must be taken into account from the outset. These features of absorption are even more important in such phenomena as electron and phonon emission by islet metallic films illuminated by laser light.[8,9] It is in such phenomena that the extremal cross sections of absorption by an ensemble of metallic islets play a much more important role than the average (effective) cross sections.

This work was supported by a grant from the International Science Foundation (Grant No. K6D100).

[1)]The fact that absorption is proportional to the surface area of the small spherical particles was reported by Manykin *et al.*[20]

[1] E. D. Belotskii and P. M. Tomchuk, Int. J. Electron. **69**, 169 (1990).
[2] E. D. Belotskiĭ, S. P. Luk'yanets, and P. M. Tomchuk, Zh. Éksp. Teor. Fiz. **101**, 163 (1992) [Sov. Phys. JETP **74**, 88 (1992)].
[3] S. A. Gorban', S. A. Nepijko, and P. M. Tomchuk, Int. J. Electron. **70**, 485 (1991).
[4] E. D. Belotskii and P. M. Tomchuk, Surf. Sci. **239**, 143 (1990).
[5] P. M. Tomchuk and R. D. Fedorovich, Fiz. Tverd. Tela (Leningrad) **8**, 276 (1966) [Sov. Phys. Solid State **8**, 276 (1996)].
[6] P. M. Tomchuk and R. D. Fedorovich, Fiz. Tverd. Tela (Leningrad) **8**, 3131 (1966) [Sov. Phys. Solid State **8**, 2510 (1996)].
[7] R. D. Fedorovich, A. G. Naumovets, and P. M. Tomchuk, Prog. Surf. Sci. **42**, 189 (1993).
[8] A. A. Benditskiĭ, L. V. Viduta, Yu. A. Kulyupin *et al.*, Izv. Akad. Nauk SSSR, Ser. Fiz. **50**, 1643 (1986).
[9] A. A. Benditskiĭ, L. V. Viduta, V. I. Konev *et al.*, Poverkhnost' No. 10, 48 (1988).
[10] P. M. Tomchuk, Surf. Sci. **330**, 350 (1995).
[11] Yu. I. Petrov, *The Physics of Small Particles* [in Russia], Nauka, Moscow (1982).
[12] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media*, 2nd ed, Pergamon Press, Oxford (1960).
[13] A. Kawabata and R. Kubo, J. Phys. Soc. Jpn. **21**, 1765 (1966).
[14] R. Ruppin and Yaton, Phys. Status Solidi B **74**, 647 (1976).
[15] D. M. Wood and N. W. Ashcroft, Phys. Rev. B **25**, 6255 (1982).
[16] M. L. Levin and R. Z. Muratov, Zh. Tekh. Fiz. **38**, 1623 (1968) [Sov. Phys. Tech. Phys. **13**, 1318 (1968)].

[17] C. Pecharroman and J. E. Iglesias, Phys. Rev. B **49**, 7137 (1994).
[18] A. G. Lesskin, V. E. Pasternak, and A. A. Yushkanov, Zh. Éksp. Teor. Fiz. **83**, 310 (1982) [Sov. Phys. JETP **56**, 710 (1982)].
[19] E. I. Austin and M. Wilkinson, J. Phys.: Condens. Matter **5**, 8451 (1993).
[20] É. A. Manykin, P. P. Poluéktov, and Yu. G. Rubezhnyĭ, Zh. Éksp. Teor. Fiz. **70**, 2117 (1976) [Sov. Phys. JETP **43**, 1105 (1976)].
[21] A. Ya. Blank and A. Ya. Sharshanov, Fiz. Nizk. Temp. **21**, 263 (1995) [Low Temp. Phys. **19**, 263 (1993)].

Translated by Eugene Yankovsky

Inelastic light scattering by electrons and plasmons in metals

L. A. Falkovsky[1)]

L. D. Landau Institute for Theoretical Physics, Russian Academy of Sciences, 117334 Moscow, Russia

S. Klama[2)]

Institute of Molecular Physics, Polish Academy of Sciences Smoluchowskiego 17, 60-179 Poznan, Poland
(Submitted 14 November 1996)
Zh. Éksp. Teor. Fiz. **112**, 679–689 (August 1997)

The cross section of the inelastic light scattering by electron–hole plasma in metals is studied. The Coulomb interaction of electron excitations is taken into account self-consistently. The system of the Boltzmann equation for electronic fluctuations and Maxwell's equations for the interaction field is solved. The Raman spectra consist of the electron–hole background, diffuson and plasmon resonances. The widths of this background and resonance are determined by the electron collision rate as well as by the decay of the incident and scattered radiation in the metal. The line shape depends on the screening of the electron–light interaction, i.e., on the incident radiation frequency. © *1997 American Institute of Physics.* [S1063-7761(97)02008-8]

1. INTRODUCTION

In recent times the inelastic light scattering has attracted considerable attention because of the puzzle of high-temperature superconductivity.[1–5] Also, synchrotron sources of radiation[6,7] with high resolution have led to an advance in experimental investigations of electronic excitations in solids and liquids. Light scattering experiments give access to detailed information on various elementary excitations: phonons,[8] plasmons[9,10] and magnons.[11] The influence of the transition to the superconducting state on the Raman light scattering was first studied theoretically in Ref. 12, and more recently for the synchrotron radiation frequency.[13]

The Raman spectra present a very complex picture and attention to their subtle aspects was paid only recently. In particular, it has been discovered recently that the interaction of the phonon resonances with the electron–hole continuum leads to characteristic changes in the shape of the resonance line.[14–16] The resonance line acquires the Fano resonance line shape, also known as the Breit–Wigner resonance in nuclear physics. The second example, where the resonance peak has a specific shape, is the two-magnon resonance, in which two magnons are involved in the electron transition through a gap.[17]

The inelastic light scattering by normal metals has not yet been studied experimentally in such a detailed manner as that by superconductors. We would like to concentrate on an effect which has already been noticed,[18] but not investigated in detail: the influence of the spatial distribution of the incident and scattered light in a metal on the resonance line shape. Up to now in investigations of Raman scattering, the solid state is treated as a nonabsorbent substance, where the incident and scattered light possess well defined wave vectors. The existence of imaginary parts of the wave vector changes the line shape. This effect should be considered as competitive with the Fano resonance.

In this paper we study how the above-mentioned field decay in a metal affects resonance with the plasmon excitation. The interaction between plasmon resonance and the electron–hole excitations is taken into account. In Fig. 1 the effect of the electron–hole excitations is represented by a loop. The electron–hole excitations are shown in Fig. 1a. This diagram describes a continuum with a width depending on the collision rate. In the ''dirty'' limit the diagram involves a diffusion pole. The diagram 1b, where the electromagnetic interaction is shown by the dotted line, has a narrow plasmon pole at the electron plasma frequency ω_0. It will be shown that the influence of the electron loop leads to asymmetry of the plasmon resonance (the Fano effect). Two contributions shown in Fig. 1 should be integrated with the factor $|U(\mathbf{r},t)|^2=|A^{(i)}A^{(s)}|^2$, whose width is determined by the spatial damping of incident (i) and scattered (s) light.

The Raman cross section is expressed in terms of a linear response of the electron system with the Coulomb interaction to the external field $U(\mathbf{r},t)$. We solve self-consistently Boltzmann's equation for electron fluctuations and Maxwell's equations for the interaction field.

For the applicability of the Boltzmann equation the following conditions should be fulfilled:

$$|k|=|k^{(i)}-k^{(s)}|\ll k_F, \quad |\omega|=|\omega^{(i)}-\omega^{(s)}|\ll\varepsilon_F,$$

where k_F and ε_F are the Fermi momentum and energy.[3)] The first condition allows an analytical expression to be found for the distribution function of charge carriers. We are interested in small values of $|\omega|$ and $|\omega|\simeq\omega_0$. In the latter case the condition $e^2/\hbar v\ll 1$ should be fulfilled, where v is the Fermi velocity. A rigorous method is not known for an evaluation of the response function for an arbitrary $e^2/\hbar v$, but the permittivity calculated from the Boltzmann equation coincides with Lindhard's expression obtained in the limit $|k|\ll k_F$ and for arbitrary values of $|\omega|$. Therefore we use the kinetic equation and write it in the τ-approximation. The τ-approximation can be well justified for elastic scattering processes.[19,20] An attempt to include the inelastic scattering by phonons has been made in Ref. 21, where the response function was obtained by the Green's functions method. Strictly speaking, the collision rate τ for large frequency transfer ($|\omega| \simeq \omega_0$) depends on ω and is determined by

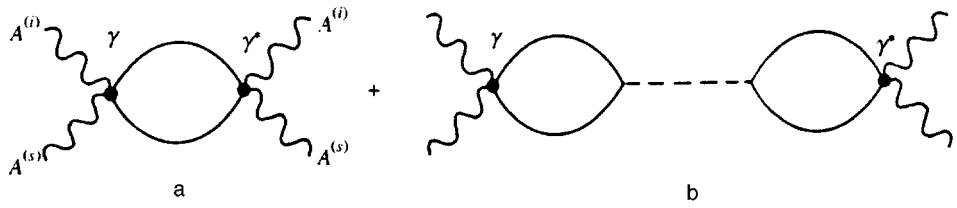


FIG. 1. Feynman diagrams depicting the two contributions to Raman scattering in metals. (a) The electron–hole contribution; (b) the contribution of an electromagnetic excitation (plasmon, the dotted line) in the electron–hole plasma. Here the solid line is the electron Green's function, and the wavy line is the incident and scattered light. The black dots are vertices describing the electron interaction with the incident and scattered light. The Coulomb interaction is represented by an empty vertex.

electron–electron collisions. The long wave part of the Coulomb interaction is taken into account in a self-consistent way (the Vlasov–Landau approximation; for the scattering problem, see Ref. 18); The contributions of the local field will be disregarded.

2. INELASTIC LIGHT SCATTERING AND SUSCEPTIBILITY

We consider light scattering by a metal which occupies the half-space $z>0$. The Raman cross section has the form

$$d\sigma(k_x,\omega)=\left(\frac{8\pi e^2}{mc\hbar\omega^{(i)}}\right)^2\frac{\Sigma(k_x,\omega)}{1-\exp(-\hbar\omega/k_BT)}\times\frac{k_z^{(s)}\omega^{(s)}d\omega^{(s)}d\Omega^{(s)}}{c(2\pi)^3},\tag{1}$$

where the density–density correlation function $\Sigma(k_x,\omega)$ contains the bulk and surface contributions.[18] Here we are interested in the bulk part only. Then we can use the even continuation of all the fields in the semi-space $z<0$ and take the spatial Fourier transform. The correlator $\Sigma(k_x,\omega)$ is expressed in terms of the generalized susceptibility $\chi(\mathbf{k},\omega)$:

$$\Sigma(k_x,\omega)=-\int\frac{dk_z}{2\pi}\,|U(\mathbf{k},\omega)|^2\ \mathrm{Im}\ \chi(\mathbf{k},\omega).\tag{2}$$

The generalized susceptibility $\chi(\mathbf{k},\omega)$ in the field $U(\mathbf{k},\omega)$ is defined as follows

$$\int\frac{2d^3p}{(2\pi)^3}\,\gamma^*(\mathbf{p})f_p(\mathbf{k},\omega)\equiv-\chi(\mathbf{k},\omega)U(\mathbf{k},\omega),\tag{3}$$

where $f_p(\mathbf{r},t)$ is the electron distribution function with a specular boundary condition at $z=0$.

The field $U(\mathbf{r},t)$ considered below as the external force is the product of the vector potentials of the incident light $\mathbf{A}^{(i)}(\mathbf{r},t)$ and scattered light $\mathbf{A}^{(s)}(\mathbf{r},t)$:

$$\mathbf{A}^{(i)}(\mathbf{r},t)\mathbf{A}^{(s)}(\mathbf{r},t)\simeq U(\mathbf{r},t)=U(\mathbf{k},\omega)\exp[i(\mathbf{k}_s,\mathbf{s}-\omega t)],$$

where $\mathbf{k}_s=\mathbf{k}_s^{(i)}-\mathbf{k}_s^{(s)}$. Here the subscript s denotes vector components parallel to the surface. For the unbounded space $\mathbf{e}^{(i)}$, $\mathbf{e}^{(s)}$ are the polarization vectors of the incident and scattered fields, respectively, and they are included in the vertex function $\gamma(\mathbf{p})$. For a half space the polarization vectors $\mathbf{e}^{(i)}$, $\mathbf{e}^{(s)}$ are determined by the solution of the electrodynamic problem (for details, see Refs. 18 and 20. The vertex function $\gamma(\mathbf{p})$ contains resonance denominators appearing in the second order of the perturbation theory with respect to $\mathbf{A}^{(i)}$, $\mathbf{A}^{(s)}$. The Fourier transform of the field $U(\mathbf{r},t)$ is given by

$$U(\mathbf{k},\omega)=\frac{2i\zeta}{\zeta^2-k_z^2},\tag{4}$$

where the complex quantity $\zeta=\zeta_1+i\zeta_2$ is the sum of the normal components of the wave vector of the incident and scattered light in the metal.

The electron distribution function $f_p(\mathbf{r},t)$ is sought in the form

$$f_p(\mathbf{r},t)=f_0(\varepsilon(\mathbf{p},\mathbf{r},t)-\mu)+\frac{df_0}{d\varepsilon_0}\,\delta f_p(\mathbf{r},t),$$

where $f_0(\varepsilon(\mathbf{p},\mathbf{r},t)-\mu)$ is the Fermi–Dirac distribution function depending on the local electron energy

$$\varepsilon(\mathbf{p},\mathbf{r},t)=\varepsilon_0(\mathbf{p})+\gamma(\mathbf{p})U(\mathbf{r},t).$$

The nonequilibrium part of the electron distribution function is governed by the Boltzmann equation:

$$-i(\omega-\mathbf{kv})\,\delta f_p(\mathbf{k},\omega)=[i\omega\gamma(\mathbf{p})U(\mathbf{k},\omega)-e\mathbf{vE}(\mathbf{k},\omega)]-\nu(\delta f_p(\mathbf{k},\omega)-\langle\delta f_p(\mathbf{k},\omega)\rangle),\tag{5}$$

where ν is the collision rate and the angle brackets denote averaging over the Fermi surface,

$$\langle(\cdots)\rangle=\frac{1}{n_0}\int\frac{2dS}{(2\pi)^3v}\,(\cdots),\quad n_0=\int\frac{2dS}{(2\pi)^3v}.$$

Here n_0 is the density of electron states; we assume $T\ll\epsilon_F$.

The electric field $\mathbf{E}(\mathbf{k},\omega)$ describes the electron–electron Coulomb interaction and is determined by the Maxwell equation

$$\mathrm{curl}\ \mathrm{curl}\ \mathbf{E}(\mathbf{r},\omega)-\frac{\omega^2}{c^2}\,\mathbf{D}(\mathbf{r},\omega)=\frac{4\pi i\omega}{c^2}\,\mathbf{j}(\mathbf{r},\omega)\tag{6}$$

with

$$D_\alpha(\mathbf{k},\omega)=\epsilon^0_{\alpha\beta}E_\beta(\mathbf{k},\omega),$$

$$\mathbf{j}(\mathbf{k},\omega)=e\int\frac{2d^3p}{(2\pi)^3}\,\mathbf{v}\delta f_p(\mathbf{k},\omega),\tag{7}$$

where $\epsilon^0_{\alpha\beta}$ is the dielectric constant of the filled bands.

Substituting the solution of Boltzman's equation (5) into Eq. (7) we obtain

$$j_\alpha(\mathbf{k},\omega)=\sigma_{\alpha\beta}(\mathbf{k},\omega)E_\beta(\mathbf{k},\omega)+\Gamma_\alpha^{(\gamma)}(\mathbf{k},\omega)U(\mathbf{k},\omega), \quad (8)$$

where

$$\sigma_{\alpha\beta}(\mathbf{k},\omega)=ie^2n_0\left\langle\frac{v_\alpha\tilde{v}_\beta}{\tilde{\omega}-\mathbf{kv}}\right\rangle,$$

$$\Gamma_\alpha^{(\gamma)}(\mathbf{k},\omega)=e\omega n_0\left\langle\frac{v_\alpha\tilde{\gamma}(\mathbf{p})}{\tilde{\omega}-\mathbf{kv}}\right\rangle,$$

$$\tilde{\mathbf{v}}=\mathbf{v}+\frac{i\nu\langle\mathbf{v}/\tilde{\omega}-\mathbf{kv})\rangle}{1-\langle i\nu/(\tilde{\omega}-\mathbf{kv})\rangle},$$

$$\tilde{\gamma}(\mathbf{p})=\gamma(\mathbf{p})+\frac{i\nu\langle\gamma(\mathbf{p})/(\tilde{\omega}-\mathbf{kv})\rangle}{1-\langle i\nu/(\tilde{\omega}-\mathbf{kv})\rangle},$$

$\tilde{\omega}=\omega+i\nu$.

The solutions of the Maxwell equation (6) and (7) has the form

$$E_\alpha(\mathbf{k},\omega)=\frac{4\pi i\omega}{c^2}\,U(\mathbf{k},\omega)\mathscr{D}_{\alpha\beta}(\mathbf{k},\omega)\Gamma_\beta^{(\gamma)}(\mathbf{k},\omega) \quad (9)$$

with the matrix

$$\mathscr{D}_{\alpha\beta}(\mathbf{k},\omega)=\left(k^2\delta_{\alpha\beta}-k_\alpha k_\beta-\frac{\omega^2}{c^2}\,\epsilon_{\alpha\beta}(\mathbf{k},\omega)\right)^{-1},$$

where $\epsilon_{\alpha\beta}(\mathbf{k},\omega)$ is related to the conductivity tensor

$$\epsilon_{\alpha\beta}(\mathbf{k},\omega)=\epsilon_{\alpha\beta}^0+\frac{4\pi i}{\omega}\,\sigma_{\alpha\beta}(\mathbf{k},\omega). \quad (10)$$

Using the solution of the Boltzmann equation (5) and the Maxwell equation (9) we obtain the generalized susceptibility (3):

$$\chi(\mathbf{k},\omega)=\omega n_0\left\langle\frac{\gamma^*(\mathbf{p})\tilde{\gamma}(\mathbf{p})}{\tilde{\omega}-\mathbf{kv}}\right\rangle-\frac{4\pi}{c^2}\times\tilde{\Gamma}_\alpha^{(\gamma^*)}(\mathbf{k},\omega)\mathscr{D}_{\alpha\beta}(\mathbf{k},\omega)\Gamma_\beta^{(\gamma)}(\mathbf{k}\omega), \quad (11)$$

where

$$\tilde{\Gamma}_\alpha^{(\gamma^*)}(\mathbf{k},\omega)=e\omega n_0\left\langle\frac{\tilde{v}_\alpha\gamma^*(\mathbf{p})}{\tilde{\omega}-\mathbf{kv}}\right\rangle.$$

For normal propagation of the incident and scattered waves ($\mathbf{k}=k_z=k$) the expression on (11) reads

$$\chi(k,\omega)=\omega n_0\left\langle\frac{\gamma^*(\mathbf{p})\tilde{\gamma}(\mathbf{p})}{\tilde{\omega}-kv_z}\right\rangle+\frac{4\pi}{\omega^2\epsilon_{zz}(k,\omega)}\times\tilde{\Gamma}_z^{(\gamma^*)}(k,\omega)\Gamma_z^{(\gamma)}(k,\omega). \quad (12)$$

Substituting (12) into (2) we obtain

$$\Sigma(k_x=0,\omega)=\Sigma_1(0,\omega)+\Sigma_2(0,\omega),$$

where $\Sigma_1(0,\omega)$ is the contribution of the first term in (12), related to the excitation of electron–hole pairs and $\Sigma_2(0,\omega)$ exhibits the plasmon resonance associated with $\epsilon_{zz}(k,\omega)$. These two contributions are shown in Fig. 1; the loops correspond to $\Gamma_z(k,\omega)$ and the dotted line to $\epsilon_{zz}(k,\omega)$.

3. LARGE AND LOW SPATIAL DISPERSION

Let us consider two important limiting cases.

(i) Large-k limit: $kv\gg|\tilde{\omega}|$.

Since we have $\langle v_z/\tilde{\omega}-kv_z\rangle=-1/k+i\pi(\tilde{\omega}/k^2)\langle\delta(\mu)/v\rangle$, with $\mu=\cos(\mathbf{k},\mathbf{v})$, we obtain the susceptibility (12) in the form

$$\chi(k,\omega)=\frac{\pi n_0\omega}{k}\,\langle|\gamma(\mathbf{p})-\langle\gamma(\mathbf{p})\rangle|^2\delta(\mu)/v\rangle. \quad (13)$$

Equation (13) describes the screening of light scattering by the Coulomb electron–electron interaction.[22] This means that in the frequency range $\omega\simeq v|\zeta|<\omega_0$, up to the plasma frequency ω_0 the density fluctuations are screened due to the Coulomb electron–electron interaction.

(ii) Low k-limit: $kv\ll|\tilde{\omega}|$.

We expand all integrands in χ given by (12) in power series with respect to $kv/|\tilde{\omega}|$. The second term in (12) is proportional to k^2. In the leading approximation the first term gives

$$\chi(k\to0,\omega)=\frac{n_0\omega}{\tilde{\omega}}\left(\langle|\gamma(\mathbf{p})|^2\rangle+\frac{i\nu}{\omega-i\nu\langle v_z^2\rangle k^2\tilde{\omega}^{-2}}\,|\langle\gamma(\mathbf{p})\rangle|^2\right). \quad (14)$$

We retain in the denominator the term proportional to k^2 (the diffusion pole) since it is important at low $\omega\sim k^2\langle v_z^2\rangle/\nu\ll\nu$.

If we leave out of consideration the frequency range near the diffuson pole, we obtain

$$\chi(k,\omega)=\chi(k=0,\omega)+\delta\chi(k,\omega),$$

where the first term results only from the electron–hole excitations

$$-\mathrm{Im}\;\chi(k=0,\omega)=n_0\langle|\gamma(\mathbf{p})-\langle\gamma(\mathbf{p})\rangle|^2\rangle\,\frac{\nu\omega}{\omega^2+\nu^2}. \quad (15)$$

This frequency dependence of the cross section has been obtained in Refs. 19 and 23. The first term in Eq. (12) also gives a contribution proportional to k^2:

$$\chi_1(k,\omega)=\chi(0,\omega)+\delta\chi_1(k,\omega),$$

$$\delta\chi_1(k,\omega)=\frac{n_0\omega k^2}{\tilde{\omega}^3}\left[\langle v_z^2|\gamma(\mathbf{p})|^2\rangle-\frac{\nu^2}{\omega^2}\,\langle v_z^2\rangle|\langle\gamma(\mathbf{p})\rangle|^2+\frac{i\nu}{\omega}\,(\langle\gamma^*(\mathbf{p})\rangle\langle\gamma(\mathbf{p})v_z^2\rangle+\text{c.c.})\right]. \quad (16)$$

The second term in Eq. (12) is of resonant form:

$$\chi_2(k,\omega)=\frac{n_0\omega_0^2\varepsilon_{zz}^0k^2}{\langle v_z^2\rangle\tilde{\omega}^4\varepsilon_{zz}(k,\omega)}\left\langle v_z^2\left(\gamma(\mathbf{p})+\frac{i\nu}{\omega}\,\langle\gamma(\mathbf{p})\rangle\right)\right\rangle\times\left\langle v_z^2\left(\gamma^*(\mathbf{p})+\frac{i\nu}{\omega}\,\langle\gamma^*(\mathbf{p})\rangle\right)\right\rangle, \quad (17)$$

where

$$\tilde{\omega}^3\omega\varepsilon_{zz}(k,\omega)=\varepsilon_{zz}^{(0)}(\tilde{\omega}^3\omega-\omega_0^2\tilde{\omega}^2-\omega_0^2u^2k^2-i\nu\omega_0^2\langle v_z^2\rangle k^2/\omega),$$

the electron plasma frequency is given by $\omega_0^2 = 4\pi e^2 n_0 \langle v_z^2 \rangle / \epsilon_{zz}^0$, and the dispersion parameter is $u^2 = \langle v_z^4 \rangle / \langle v_z^2 \rangle$.

The equation $\epsilon_{zz}(k,\omega) = 0$ gives the frequency and damping (including its dispersion) of a plasmon, $\omega = \omega_{\rm pl}(k) - i\Gamma(k)$, where

$$\omega_{\rm pl}^2(k) = \omega_0^2 + u^2 k^2,$$

$$\Gamma(k) = \nu[1 + (u^2 - \langle v_z^2 \rangle) k^2/\omega_0^2]/2.$$

Below we will omit the small dispersion due to damping.

The terms proportional to k^2 in (16) and (17) can be written as follows

$$\delta\chi(k,\omega) = \frac{n_0 |\langle v_z^2 \gamma(\mathbf{p}) \rangle|^2 k^2}{\langle v_z^2 \rangle \tilde{\omega}} \left[\frac{C_1 \omega}{\tilde{\omega}^2} + \frac{\omega\omega_0^2 - \nu^2 \tilde{\omega} C_3 + 2i\nu\omega\tilde{\omega} C_2}{\Delta} \right], \tag{18}$$

where $\Delta = \tilde{\omega}^3 \omega - \tilde{\omega}^2 \omega_0^2 - \omega_0^2 u^2 k^2$,

$$C_1 = \frac{\langle v_z^2 \rangle \langle v_z^2 |\gamma(\mathbf{p})|^2 \rangle}{|\langle v_z^2 \gamma(\mathbf{p}) \rangle|^2},$$

$$C_2 = \frac{\langle v_z^2 \rangle (\langle \gamma^*(\mathbf{p}) \rangle \langle v_z^2 \gamma(\mathbf{p}) \rangle + \text{c.c.})}{2|\langle v_z^2 \gamma(\mathbf{p}) \rangle|^2},$$

$$C_3 = \frac{\langle v_z^2 \rangle^2 |\langle \gamma(\mathbf{p}) \rangle|^2}{|\langle v_z^2 \gamma(\mathbf{p}) \rangle|^2}.$$

Now we are in a position to calculate the cross section.

4. LIGHT SCATTERING CROSS SECTION

In the ''clean'' limit $|\omega + i\nu| \ll v|\zeta|$ one obtains the cross section by substituting Eq. (13) into (2). For a more interesting case, $\zeta_1 \gg \zeta_2$, we get

$$\Sigma(k_x = 0, \omega) = \pi n_0 \omega \langle |\gamma(\mathbf{p}) - \langle \gamma(\mathbf{p}) \rangle|^2 \delta(\mu)/v \rangle \times \left(\frac{4}{\pi \zeta_1^2} \ln \frac{v\zeta_1}{\omega} + \frac{1}{\zeta_1 \zeta_2} \right). \tag{19}$$

In this case $\Sigma(0,\omega)$ has a maximum at $\omega \simeq v\zeta_1$.

In the ''dirty'' limit $\nu \gg v|\zeta|$, there are two regions: the low and large frequency transfers. For the case of low frequency transfer $|\omega| \leqslant \nu$, we obtain the cross section by substituting (14) into (2). If $\zeta_1 \gg \zeta_2$, we can integrate only $|U|^2$, taking the smooth function Im $\chi(k,\omega)$ at $k = \zeta_1$:

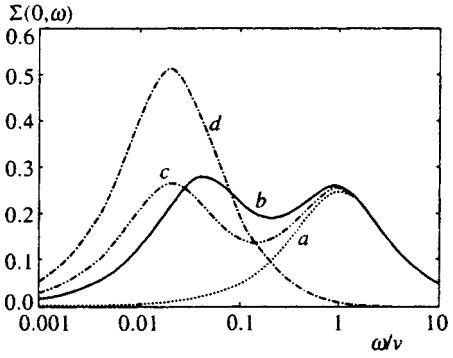


FIG. 2. Diffuson resonance and relaxation maximum for different values of parameters given in the text. In the curve (d) the relaxation maximum is absent because of the complete screening; (b), (c): The screening is partial; the mean free path for (b) is larger than for (c). The curve (a) represents the Raman spectrum without the diffusion resonance (see Eq. (23)).

$$\Sigma(k \to 0, |\omega| \leqslant \nu) = \frac{n_0 \nu \omega}{\zeta_2(\omega^2 + \nu^2)} \left(\langle |\gamma(\mathbf{p})|^2 \rangle - \frac{\omega^2 - \zeta_1^2 \langle v_z^2 \rangle}{\omega^2 + (\zeta_1^2 \langle v_z^2 \rangle/\nu)^2} |\langle \gamma(\mathbf{p}) \rangle|^2 \right). \tag{20}$$

This expression has two maxima (see Fig. 2). The one at $|\omega| \simeq \zeta_1^2 \langle v_z^2 \rangle/\nu \ll \nu$ describes the diffuson excitations. The other at $|\omega| = \nu$ results from relaxation processes in the electron system. The relaxation maximum is absent (curve (d)) if $\langle |\gamma(\mathbf{p})|^2 \rangle = |\langle \gamma(\mathbf{p}) \rangle|^2$. We define this case as the complete-screening limit. For $|\omega| \gg \zeta_1 \langle v_z^2 \rangle^{1/2}$ Eq. (20) transforms to Eq. (23) (see curve (a)) obtained first by Zawadowski and Cardona.[23] In Fig. 2 the following sets of parameters ($w = |\langle \gamma(\mathbf{p}) \rangle|^2 / \langle |\gamma(\mathbf{p})|^2 \rangle$. $v_0 = \langle v_z^2 \rangle \zeta_1^2/\nu^2$) are used: ($a$) $w = 0.5$, $v_0 = 0$; (b) $w = 0.5$, $v_0 = 0.04$; (c) $w = 0.5$, $v_0 = 0.02$; (d) $w = 1.0$, $v_0 = 0.02$. The units of $\Sigma(0,\omega)$ are $n_0 \langle |\gamma(\mathbf{p})|^2 \rangle / \zeta_2$.

For large frequency transfer $|\omega| > \nu$, we find the cross section using (15) and (18). One can obtain a simple form in the limit $u^2 \zeta_1^2 \ll \nu\omega_0$, where the spatial dispersion of a plasmon can be disregarded. Making use of the integrals

$$\int \frac{dk_z}{2\pi} |U(\mathbf{k},\omega)|^2 = \frac{1}{\zeta_2}, \quad \int \frac{dk_z}{2\pi} |U(\mathbf{k},\omega)|^2 k_z^2 = \frac{\zeta_1^2}{\zeta_2}, \tag{21}$$

we obtain

$$\Sigma(0,\omega) = \Sigma_1(0,\omega) + \delta\Sigma(0,\omega), \tag{22}$$

$$\Sigma_1(0,\omega) = n_0 \langle |\gamma(\mathbf{p}) - \langle \gamma(\mathbf{p}) \rangle|^2 \rangle \frac{\nu\omega}{\zeta_2(\omega^2 + \nu^2)}, \tag{23}$$

$$\delta\Sigma(0,\omega) = \frac{n_0 |\langle \gamma(\mathbf{p}) v_z^2 \rangle|^2 \zeta_1^2 \nu\omega}{\langle v_z^2 \rangle \zeta_2 (\omega^2 + \nu^2)^3} \left((3\omega^2 - \nu^2) C_1 + \frac{\omega_0^2 [4\omega^2(\omega^2 - \nu^2) - \omega_0^2(3\omega^2 - \nu^2)]}{(\omega^2 - \omega_0^2)^2 + \omega^2\nu^2} - \frac{(\omega^2 + \nu^2)[\nu^2(3\omega^2 - 2\omega_0^2 - \nu^2) C_3 + 2((\omega^2 - \nu^2)(\omega^2 - \omega_0^2) - 2\omega^2\nu^2)] C_2}{(\omega^2 - \omega_0^2)^2 + \omega^2\nu^2} \right). \tag{24}$$

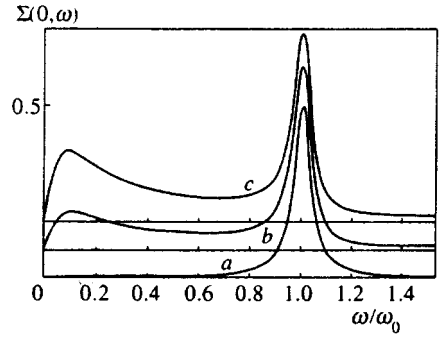


FIG. 3. Theoretical prediction of Raman spectra for normally incident and scattered light. The background is the electron–hole continuum with large collision rate. The resonance corresponds to the excitation of plasmon with low spatial dispersion. The Coulomb screening of electron density fluctuations is complete for the curve (a), and partial for (c). The curves (b) and (c) are shifted along the abscissa. The parameters are defined in the text.

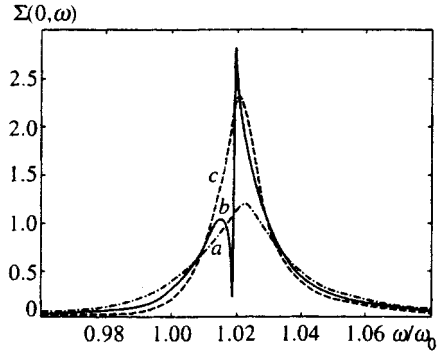


FIG. 4. Plasmon resonance in light scattering for crossover of damping (curve (b) $\nu/\omega_0=0.0115$); (a) plasmon damping is larger than the field damping ($\nu/\omega_0=0.02$); (c) $\nu/\omega_0=0.006$; ν is the electron collision rate, ω_0 is the plasma frequency.

The frequency dependence of the cross section (22) is shown in Fig. 3 for various values of the parameters: $C_1=C_3=1$, $u^2\zeta_1^2/\omega_0^2=0.04$, $\nu/\omega_0=0.08$, and (a) $C_2=1$, $x=0$, (b) $C_2=2$, $x=0.2$, (c) $C_2=2$, $x=0.4$, where $x\equiv u^2\langle|\gamma(\mathbf{p})-\langle\gamma(\mathbf{p})\rangle|^2\rangle\langle v_z^2\rangle/|\langle\gamma(\mathbf{p})v_z^2\rangle|^2$. The units of $\Sigma(0,\omega)$ are $n_0|\langle\gamma(\mathbf{p})v_z^2\rangle|^2/\zeta_2\langle v_z^2\rangle u^2$.

There is a wide background in the range $\nu<|\omega|<\omega_0$, even if the $\mathbf{p}$-dependence of $\gamma(\mathbf{p})$ is weak, when

$$\langle|\gamma(\mathbf{p})-\langle\gamma(\mathbf{p})\rangle|^2\rangle\simeq 0.2|\langle\gamma(\mathbf{p})\rangle|^2.$$

The resonance line shape is asymmetric if the coefficient C_2 differs markedly from unity: the resonance curve drops more rapidly on the side $|\omega|>\omega_0$. We note that the parameters $\gamma(\mathbf{p})$, ζ_1 and ζ_2 depend on frequencies of the incident and scattered light, and ν is a function of ω. All these dependences can be disregarded near the plasmon resonance, but they modify the form of the background.

To take into account the spatial dispersion of the plasmons we calculate the integral

$$\int \frac{dk_z}{2\pi}|U(\mathbf{k},\omega)|^2\frac{k_z^2}{k_z^2-k_p^2}=\frac{\zeta_1}{2\zeta_2}\left(\frac{\zeta_1-k_p\ \mathrm{sign}\ \omega}{\zeta_1^2-k_p^2+2i\zeta_1\zeta_2}+(k_p,\zeta_2\to-k_p,-\zeta_2)\right),$$

where terms of order ζ_2^2/ζ_1^2 are omitted and $k_p^2=(a+ib)/u^2\omega_0^2$,

$$a=(\omega^2-\nu^2)(\omega^2-\omega_0^2)-2\nu^2\omega^2,$$

$$b=\nu\omega[2(\omega^2-\omega_0^2)+\omega^2-\nu^2].$$

Using (2) and (18) we obtain instead of (24) the contribution

$$\delta\Sigma(0,\omega)=\frac{n_0|\langle\gamma(\mathbf{p})v_z^2\rangle|^2\zeta_1}{2\langle v_z^2\rangle\zeta_2}\left[\frac{2\omega\nu(3\omega^2-\nu^2)\zeta_1C_1}{(\omega^2+\nu^2)^3}+\left(\left(\frac{\omega^2\omega_0^2}{\omega^2+\nu^2}-\nu^2C_3\right)\mathscr{L}_2(k_p,\zeta_2)+\nu\omega\left(\frac{\omega_0^2}{\omega^2+\nu^2}-2C_2\right)\mathscr{L}_1(k_p,\zeta_2)+(k_p,\zeta_2\to-k_p,-\zeta_2)\right)\right],\tag{25}$$

where

$$\mathscr{L}_1(k_p,\zeta_2)=[(\zeta_1-k_1\ \mathrm{sign}\ \omega)(a-u^2\omega_0^2\zeta_1^2)-(b-2u^2\omega_0^2\zeta_1\zeta_2)k_2\mathrm{sign}\omega]/\Delta_1,$$

$$\mathscr{L}_2(k_p,\zeta_2)=[(\zeta_1-k_1\mathrm{sign}\ \omega)(b-2u^2\omega_0^2\zeta_1\zeta_2)+(a-u^2\omega_0^2\zeta_1^2)k_2\mathrm{sign}\ \omega]/\Delta_1,$$

$$\Delta_1=(a-u^2\omega_0^2\zeta_1^2)^2+(b-2u^2\omega_0^2\zeta_1\zeta_2)^2,\tag{26}$$

$k_1+ik_2=k_p$ with

$$k_1=(\sqrt{a^2+b^2}+a)^{1/2}/\sqrt{2}u\omega_0,$$

$$k_2=(\sqrt{a^2+b^2}-a)^{1/2}\mathrm{sign}\ \omega/\sqrt{2}u\omega_0.$$

For $u^2\zeta_1^2\ll\nu\omega$ Eq. (25) coincides with (24). Expression (25) is valid if $|\omega+i\nu|\gg v\zeta_1$.

The position of the resonance is given by the condition $a=u^2\omega_0^2\zeta_1^2$. This means that a plasmon is excited with the wave vector equal to twice the wave vector of the incident light (we consider the backscattering geometry). The width of the resonance is determined by competition between the intrinsic plasmon damping, b, and the decay of light, ζ_2. Usually the condition $2u^2\omega_0^2\zeta_1\zeta_2\ll|b|$ holds, where the width of the plasmon resonance is connected with the electron collision rate only (curve (a) in Fig. 4).

The resonance line has a critical behavior ((b) in Fig. 4) when $b\simeq 2u^2\omega_0^2\zeta_1\zeta_2$ (see (26)). In case (c) the resonance width is determined by the decay of the light ζ_2. The parameter x and the units of $\Sigma(0,\omega)$ are defined as for Fig. 3 (see text after (24)), $C_1=C_3=1$, $C_2=2$, $u^2\zeta_1^2/\omega_0^2=0.04$, $x=0.2$, $2\zeta_2/\zeta_1=0.3$.

5. CONCLUSIONS

In this paper we considered the effect of the electron collision rate and incident field decay on inelastic light scattering in metals. Our method is based on the straightforward solution of the Boltzmann equation for the electronic fluctuations and the Maxwell equation for the electron–electron

interaction field. If the incident radiation frequency is not too large, the momentum dependence of the electron–light interaction $\gamma(\mathbf{p})$ has to be taken into consideration. For the scattering geometry A_{1g} (incident and scattered light polarized along the x-axis and propagating along the z-axis)

$$\gamma(\mathbf{p}) = e_x^{(i)} e_x^{(s)} \left[1 + \frac{1}{m} \sum_n \left(\frac{|p_{fn}^x|^2}{\epsilon_f(\mathbf{p}) - \epsilon_n(\mathbf{p}) + \omega^{(i)}} + \frac{|p_{fn}^x|^2}{\epsilon_f(\mathbf{p}) - \epsilon_n(\mathbf{p}) - \omega^{(s)}} \right) \right], \tag{27}$$

where the subscript f denotes the index of the band in which the carriers exist, the transitions take place into some band n, p_{fn}^x is the electron momentum matrix element, m is the electron mass, and for the semi-infinite metal[18] $e_x^{(i)} = (1 + \sqrt{\epsilon_{xx}(\omega^{(i)})})^{-1}$.

The electron–light interaction is screened by the Coulomb interaction as well as by electron–electron collisions. The smooth part of the Raman spectra (background) is determined by the screened electron–light interaction. This part exists due to **p**-dependent second term in bracket (27). The screening is not effective for the diffusion and plasmon resonances. For complete screening there are only the plasmon resonance with symmetric line shape and the asymmetric diffusion maximum. In the intermediate case, a wide electron–hole background appears and the plasmon resonance has the asymmetric line shape. If the collision rate is low, the width of the resonance depends on the decay of incident radiation. If the collision rate is comparable with the decay of incident radiation, the plasmon resonance curve has a nontrivial form. We emphasize, that the diffusion maxima and the plasmon resonance are located in very different parts of the spectrum. The collision rate is $\nu \approx 10^3$ cm^{-1}, according to the estimate given in Ref. 16 for YBaCuO in the normal state and for the incident radiation in the optical range. Then the diffusion maximum is observed at $\omega \approx 10$ cm^{-1}.

In a layered system having a cylindrical Fermi surface with its axis perpendicular to the surface of the sample, the plasmon peak resonance does not appear for incident light normal to the surface, since $v_z = 0$. However, if the incident light falls on the surface an angle different from zero, the plasmon resonance peak should be observed.

In order to observe the phenomenon of plasmon resonance one needs a source of radiation with frequency comparable to the interband electron energy and good resolution.

The authors thank to A. R. Ferchmin for critical reading of the manuscript. One of us (L.A.F.) was supported under INTAS contract 0101-CT93-0023.

[1)]E-mail: falk@landau.ac.ru
[2)]E-mail: klama@ifmpan.poznan.pl
[3)]We put h and Boltzmann's constant equal to 1 except the final expressions.

[1] M. Boekholt, M. Hoffman, and G. Güntherodt, Physica C **175**, 127 (1991).
[2] F. Slakey, M. V. Klein, J. P. Rice, and D. M. Ginsberg, Phys. Rev. B **43**, 3764 (1991).
[3] A. A. Maksimov, A. V. Puchkov, I. I. Tartakovskii, V. B. Timofeev, D. Reznik, and M. V. Klein, Solid State Commun. **81**, 407 (1992).
[4] D. Reznik, M. V. Klein, W. C. Lee, D. M. Ginsberg, and S.-W. Cheong, Phys. Rev. B **46**, 11 725 (1992).
[5] T. P. Devereaux, D. Enzel, B. Stadlober, R. Hackl, D. H. Leach, and J. J. Neumeier, Phys. Rev. Lett. **72**, 396 (1994).
[6] P. M. Platzman, E. D. Isaacs, H. Williams, P. Zschack, and G. E. Ice, Phys. Rev. B **46**, 12943 (1992).
[7] E. D. Isaacs, P. M. Platzman, P. Metcalf, and J. M. Honig, Phys. Rev. Lett. **76**, 4211 (1996).
[8] M. Born and K. Huang, *Dynamical Theory of Crystal Lattice*, Clarendon, Oxford (1954).
[9] D. Pines and P. Nozieres. *The Theory of Quantum Liquids*, Benjamin, New York (1966).
[10] P. M. Platzman and P. A. Wolff, *Waves and Interactions in Solid State Plasmas*, Academic Press, New York (1973).
[11] M. G. Cottam and D. J. Lockwood, *Light Scattering in Magnetic Solids*, Wiley & Sons, New York (1986).
[12] A. A. Abrikosov and L. A. Falkovsky, Zh. Éksp. Teor. Fiz. **40**, 262 (1961) [Sov. Phys. JETP **13**, 179 (1961)].
[13] P. Johansson and M. Altarelli, Phys. Rev. B **53**, 8726 (1996).
[14] B. Friedl, C. Thomsen, and M. Cardona, Phys. Rev. Lett. **65**, 915 (1990).
[15] T. P. Devereaux, A. Virosztek, and A. Zawadowski, Phys. Rev. B **51**, 505 (1995).
[16] L. A. Falkovsky and S. Klama, Physica C **264**, 1 (1996).
[17] G. Blumberg, P. Abbamonte, M. V. Klein, W. C. Lee, D. M. Ginsberg, L. L. Miller, and A. Zibold, Phys. Rev. B **53**, 11930 (1996).
[18] L. A. Falkovsky and S. Klama, Phys. Rev. B **50**, 5666 (1994).
[19] L. A. Falkovsky, Zh. Éksp. Teor. Fiz. **103**, 666 (1993) [JETP **76**, 331 (1993)].
[20] L. A. Falkovsky and E. G. Mishchenko, Phys. Rev. B **51**, 7239 (1995).
[21] V. N. Kostur, Z. Phys. B **89**, 149 (1992).
[22] A. A. Abrikosov and V. M. Genkin, Zh. Éksp. Teor. Fiz. **65**, 842 (1973) [Sov. Phys. JETP **38**, 417 (1974)].
[23] A. Zawadowski and M. Cardona, Phys. Rev. B **42**, 10732 (1990).

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

Positive magnetoresistivity of $Fe_{0.95}Co_{0.05}Ge_2$ single crystals inside the region of a magnetic phase transition

R. I. Zaĭnullina, N. G. Bebenin, and V. V. Ustinov

Institute of Metal Physics, Ural Branch, Russian Academy of Sciences, 620219 Ekaterinburg, Russia
(Submitted 19 November 1996)
Zh. Éksp. Teor. Fiz. **112**, 690–697 (August 1997)

We report on the detection of two narrow peaks of positive magnetoresistivity in the temperature dependence of the magnetoresistivity of a $Fe_{0.95}Co_{0.05}Ge_2$ single crystal in the vicinity of a ''smeared'' first-order phase transition. The position of these peaks correlates with the position of singularities in the temperature dependence of the temperature derivative of the electrical resistivity and magnetic susceptibility. We show that these singularities in the transport and magnetic properties are, probably due to the presence of two percolation transitions with temperature in the magnetic subsystem of the crystal. © *1997 American Institute of Physics.* [S1063-7761(97)02108-2]

1. INTRODUCTION

Usually the features of the electrical conductivity of conducting magnetic materials in the vicinity of a phase transition are due to the rise in fluctuations in the magnetic subsystem of the crystal. Introducing a magnetic field leads to suppression of spin fluctuations and to reduction in the scattering of current carriers and, hence, to a decrease in electrical resistivity.[1–4] Negative magnetoresistivity is also observed in the vicinity of a metal–insulator transition in magnetic semiconductors, since a magnetic field is a contributory factor in the growth of regions of a conducting ferromagnetic phase.[5] In an earlier paper (see Ref. 6) we were the first to report the detection of a sharp peak of positive magnetoresistivity in a $FeGe_2$ single crystal in the vicinity of a first-order phase transition, and suggested that this effect results from the ''smearing'' of the transition. For this reason it would be interesting to study the processes that take place near the transition temperature.

A detailed study of these processes in $FeGe_2$ is made more difficult by the fact that the width of the transition in this material is extremely small (about 1 K), so that it would be more logical to use substances of the same crystalline and magnetic structure but with a ''broader'' transition region. The solid solutions $Fe_{1-x}Co_xGe_2$ can serve as such a substance. In this paper we report on an experimental study of the temperature dependence of the magnetic susceptibility, electrical resistivity, and magnetoresistivity of $Fe_{0.95}Co_{0.05}Ge_2$ single crystals. By analyzing these results we conclude that the ''smearing'' region of the magnetic phase transition probably contains two distinctive percolation transitions (in temperature) in the magnetic subsystem of the crystal.

2. SPECIMENS AND EXPERIMENTAL METHOD

The $Fe_{0.95}Co_{0.05}Ge_2$ single crystals were grown by the Czochralski method at the Department of General Physics of the Ural State Technical University in Ekaterinburg, Russia. The magnetic measurements were made with a vibrating-reed magnetometer. Magnetoresistivity was measured by the constant-current four-contact method in an electromagnet with two mutually opposed magnetic fields and two mutually opposed currents. The specimens for measuring the magnetoresistivity were parallelepipeds that measured $6\times0.5\times0.5$ mm. In magnetoresistivity measurements the temperature stability was no worse than 0.01 K.

All measurements were made in the heating regime.

3. RESULTS OF EXPERIMENT

3.1. Magnetic ordering in $Fe_{1-x}Co_xGe_2$ single crystals

The compound $FeGe_2$ (the $I4/mcm$ space group) is a metallic magnetic material with two magnetic phase transitions, at $T_1\approx260$ K and at $T_2\approx287$ K. A sequence of magnetic structures is realized in this compound: the paramagnetic ($T>T_2$), the incommensurate ($T_1<T<T_2$), and the collinear antiferromagnetic ($T<T_1$). In the magnetically ordered state the magnetic moments lie in the basal plane (001). The transition at $T=T_1$ is first-order and that at $T=T_2$ second-order.[6]

Single crystals of $Fe_{1-x}Co_xGe_2$ solid solutions have been grown only with small values of x not exceeding 0.05 (Ref. 7). We know of no detailed data on the magnetic structures realized in such single crystals. However, when cobalt atoms are substituted for iron atoms in $Fe_{0.95}Co_{0.05}Ge_2$, the magnetic phase transition temperature decreases by 20–30 K (Ref. 8).

The sequence of magnetic structures realized in this solid solution is probably the same as in $FeGe_2$. This has been corroborated by a study of magnetization curves by Vlasov *et al.*[9] They found that at low temperatures, $Fe_{0.95}Co_{0.05}Ge_2$ has a collinear antiferromagnetic structure, with the antiferromagnetic axes directed along the axes of the [110] type, just as in $FeGe_2$.

When cobalt is substituted for iron, the number of defects in the crystal increases. One indication of this is the decrease in the ratio $\rho_{293\,K}/\rho_{4.2\,K}$, where ρ is the electrical resistivity at the given temperature. According to Vlasov *et al.*,[10] as the cobalt content x grows from 0.001 to 0.05, the value of $\rho_{293\,K}/\rho_{4.2\,K}$ drops from 40–50 to 4–6. Furthermore, on the basis of a detailed study of the field dependence of the reversible and irreversible susceptibility of $Fe_{0.95}Co_{0.05}Ge_2$,

 1063-7761/97/080376-05$10.00

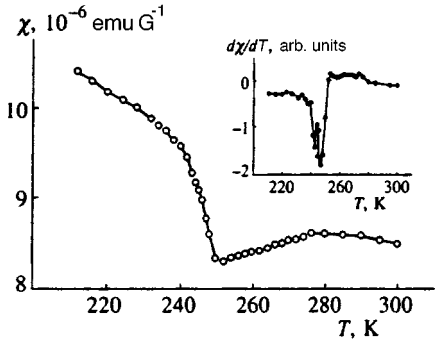


FIG. 1. Temperature dependence of susceptibility of the $Fe_{0.95}Co_{0.05}Ge_2$ single crystal in a field $H=10$ kOe ($\mathbf{H}\|[110]$); the inset depicts the temperature derivative of susceptibility as a function of temperature.

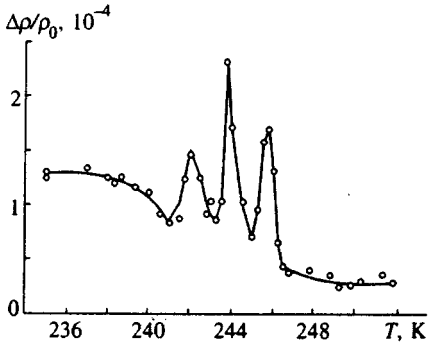


FIG. 3. Temperature dependence of longitudinal magnetoresistivity in the vicinity of a phase transition.

Vlasov *et al.*[9] showed that adding cobalt leads to a rise in the inhomogeneities of the crystal and magnetic structures.

Thus, the solid solution $Fe_{0.95}Co_{0.05}Ge_2$ can be expected to have a first-order phase transition (just as $FeGe_2$ has) whose ''width,'' however, is considerably larger than that in $FeGe_2$.

Figure 1 depicts the temperature dependence of the susceptibility of a $Fe_{0.95}Co_{0.05}Ge_2$ single-crystal specimen in a field $H=10$ kOe applied along the [110] axis. We see that between 240 and 250 K there is a significant drop in susceptibility. As the temperature increases, a mild peak in susceptibility appears at $T=270$ K. From the above it follows that two phase transitions are possible in $Fe_{0.95}Co_{0.05}Ge_2$: a first-order phase transition ''smeared'' in the temperature range from 240 K to 250 K, and a second-order phase transition at $T_2=270$ K.

3.2. Electrical resistivity and magnetoresistivity of $Fe_{0.95}Co_{0.05}Ge_2$

Figure 2 depicts the temperature dependence of the longitudinal ($\mathbf{H}\|\mathbf{j}$) magnetoresistivity

$$\frac{\Delta\rho}{\rho_0}=\frac{\rho_H-\rho_0}{\rho_0}$$

(ρ_H is the resistivity in the magnetic field) in a field $H=15$ kOe when the current density vector $\mathbf{j}$ is parallel to the [110] axis. We see that $\Delta\rho/\rho_0$ gradually decreases as the temperature grows but that in the 240–250 K range the magnetoresistivity behaves irregularly.

In Fig. 3 the temperature dependence of $\Delta\rho/\rho_0$ in this temperature range is depicted in detail. Three peaks of positive magnetoresistivity, at 242, 243.8, and 245.8 K, are clearly visible, the first peak being lower than the other two.

At higher temperatures the magnetoresistivity is extremely low and is essentially independent of the temperature.

The inset in Fig. 4 depicts the temperature dependence of the electrical conductivity ρ_0 for $\mathbf{j}\|[110]$ over the temperature range in which the magnetoresistivity exhibits abnormal behavior. At first glance it appears that $\rho_0(T)$ is an increasing function without singularities. However, at temperatures where the peaks in magnetoresistivity are observed, the $d\rho_0/dT$ vs. T curve depicted in Fig. 4 has its minima, with the left one being the least pronounced.

No appreciable temperature hysteresis of the electrical conductivity has been observed.

4. DISCUSSION

What is remarkable about our results is that in the phase transition region, the presence of a magnetic field leads not to a decrease in resistance due to suppression of fluctuations in the magnetic subsystem of the crystal but, on the contrary, to an increase in resistance, i.e., to greater scattering of current carriers. Below we develop the ideas that explain the

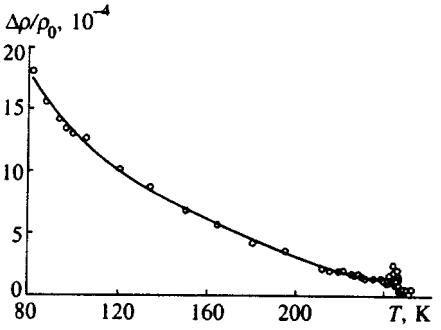


FIG. 2. Temperature dependence of longitudinal magnetoresistivity of the $Fe_{0.95}Co_{0.05}Ge_2$ single crystal in a field $H=15$ kOe ($\mathbf{H}\|\mathbf{j}\|[110]$).

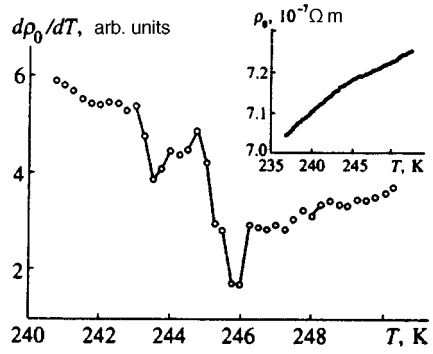


FIG. 4. Temperature dependence of $d\rho_0/dT$ and electrical resistivity ρ_0 (inset) of the $Fe_{0.95}Co_{0.05}Ge_2$ single crystal for $\mathbf{j}\|[110]$.

reasons for the peaks in magnetoresistivity, the number of peaks, and the correlations between the magnetoresistivity and the temperature coefficient of resistance.

Consider the phase transition region. We assume that the specimen consists of microregions, each of which is characterized by its own transition temperature T_c; in other words, we specify a function $T_c(\mathbf{r})$, where the vector $\mathbf{r}$ varies over the specimen. The transition is assumed to be narrow. i.e., the transition width δ_c is small compared to the average transition temperature T_c. We assume that in the phase transition region the electrical resistivity of the specimen, ρ_0, can represented by a sum of two terms, the first (the noncritical) depending only on T, $\rho_0^{nc}=\rho_0^{nc}(T)$, and the second (the critical) being a functional of $T-T_c(\mathbf{r})$, $\rho_0^c=\{F_0\{T-T_c(\mathbf{r})\}$ (for the sake of simplicity we drop the tensor indices). In a magnetic field the resistivity depends on H: first, explicitly (e.g., due to a change in the scattering probability) and, second, because of a shift in values of $T_c(\mathbf{r})$, i.e., $T_c(\mathbf{r},H)=T_c(\mathbf{r})+\Delta T_c(\mathbf{r},H)$. For the critical part of the resistivity we can write $\rho_H^c=F_H\{T-T_c(\mathbf{r})-\Delta T_c(\mathbf{r},H)\}$, where F_H is the corresponding functional. If $\Delta T_c \ll \delta_c$, we can assume that $\Delta T_c(\mathbf{r},H)\approx \Delta \overline{T}_c(H)$, since the variation ΔT_c caused by variation in $\mathbf{r}$ is a quantity of the next order of smallness. Clearly, in this case $\rho_H^c(T)=\widetilde{\rho}_H(T-\Delta \overline{T}_c(H))$, where $\widetilde{\rho}_H=F_H\{T-T_c(\mathbf{r})\}$. Bearing in mind the smallness of $\Delta T_c(H)$, we arrive at an expression for the magnetoresistivity $\Delta\rho/\rho_0$,

$$\frac{\Delta\rho}{\rho_0}=\frac{\Delta\rho^{nc}}{\rho_0}+\frac{\Delta\widetilde{\rho}}{\rho_0}-\frac{1}{\rho_0}\frac{d\widetilde{\rho}_H}{dT}\Delta\overline{T}_c(H), \tag{1}$$

where we have introduced the notation $\Delta\rho^{nc}/\rho_0=(\rho_H^{nc}-\rho_0^{nc})/\rho_0$ and $\Delta\widetilde{\rho}/\rho_0=(\widetilde{\rho}_H-\rho_0^c)/\rho_0$. Since $\rho_0^{nc}(T)$ has no singularities in the transition region and ρ_0^c and $\widetilde{\rho}_H$ depend on the same functional argument $T-T_c(\mathbf{r})$, in weak magnetic fields the singularities of the derivative $d\widetilde{\rho}_H/dT$ (if they exist) must be observed at the same values of T as the singularities of $d\rho_0/dT$.

Note that the correlation between the magnetoresistivity and the temperature resistance coefficients exists only when $\Delta T_c \ll \delta_c$.

We now employ the above ideas to analyze the processes that occur in the first-order phase transition in $Fe_{0.95}Co_{0.05}Ge_2$, which takes place near $T_1=245\,K$ in the 240–250 K temperature range. In $FeGe_2$ the quantity $\Delta T_1(H)$ is positive.[10] This should also be the case for $Fe_{0.95}Co_{0.05}Ge_2$, since in these materials the low-temperature phase has the higher susceptibility (see Fig. 1), and a magnetic field broadens the region of existence of a phase with a higher magnetization (for the same value of H) owing to a larger gain in the Zeeman energy. In $FeGe_2$ in a field $H=15$ kOe, the shift in the transition temperature amounts to about 0.1 K. It should be expected that $\Delta T_1(H)$ in $Fe_{0.95}Co_{0.05}Ge_2$ is of the same order of magnitude. Since the transition width δ_1 in $Fe_{0.95}Co_{0.05}Ge_2$ is of order 10 K (see Fig. 1), the condition $\Delta T_1 \ll \delta_1$ is sure to be met. Hence, the maxima in $\Delta\rho/\rho_0$ coincide with the minima of $d\rho_0/dT$, which was observed in the experiments.

To understand the nature of the singularities in the derivative $\Delta\rho_0/\rho_0$, we examine the processes taking place inside the transition region under temperature variations. We introduce the notation

$$T_{\min}=\min[T_1(\mathbf{r})],\quad T_{\max}=\max[T_1(\mathbf{r})]=T_{\min}+\delta_1.$$

When $T<T_{\min}$, the entire specimen is in the low-temperature phase. As the temperature becomes higher than $T_{\min}$, ''droplets'' of the high-temperature phase begin to form. The relative volume v of the high-temperature phase is given by the integral of the distribution function for the transition temperatures:

$$v=\int_{T_{\min}}^{T} W(T_1)\,dT_1. \tag{2}$$

At $T=T^{(1)}$ the volume of the high temperature phase reaches its critical value v_c, and a simply connected region of the high-temperature phase penetrating the entire specimen is formed, i.e., a distinctive percolation transition takes place. At temperatures $T_{\min}<T^{(1)}<T<T^{(2)}<T_{\max}$ the simply connected regions of low- and high-temperature phases coexist. The decay of the simply connected region of the low-temperature phase into individual clusters (the second percolation transition) occurs at $T=T^{(2)}$. A further increase in temperature leads to a decrease in the volume of the region occupied by the low-temperature phase, and at $T>T_{\max}$ the entire specimen is in the high-temperature phase.

Below we will need information about the temperature dependence of the area S of the interphase boundary surface (per unit volume). The relationship between v and S has apparently never been specially studied, so that we are forced to use simple qualitative ideas. At $T<T_{\min}$ we have $S=0$. As the temperature begins to exceed $T_{\min}$, the surface area of the interphase boundary grows. Near $T=T^{(1)}$ individual droplets begin to fuse, which probably leads to a sharp decrease in the surface area of the interphase boundary. A sharp increase in $S(T)$ due to the decay of an infinite low-temperature phase cluster must occur near the temperature $T^{(2)}$, after which S decreases and vanishes completely at $T=T_{\max}$. Thus, we can expect that the function S vs. T in $T_{\min}<T<T_{\max}$ is represented by a double-humped curve.

The resistivity of the specimen is determined by the resistivity of the homogeneous phases, the cluster distribution in the specimen, and the scattering of current carriers at the cluster boundaries. For simplicity, we assume that for a fixed temperature the resistivities of the phases are the same. Then we can write

$$\rho_0=\rho_h+\rho_{\text{inh}}, \tag{3}$$

where the first term on the right-hand side is the resistivity of the specimen in a magnetically homogeneous state, and ρ_{inh} results from scattering of current carriers at the cluster boundaries. Obviously, in our simplified model the critical resistivity ρ_0^c reduces to ρ_{inh}. As a rough approximation, we can assume that ρ_{inh} is proportional to S. Then the foregoing implies that $\rho_{\text{inh}}(T)$ vanishes outside the interval $(T_{\min}, T_{\max})$, while inside it has two peaks. Obviously, if for a large enough value of $d\rho_{\text{inh}}(T)/dT$ the derivative $d\rho_h/dT$ is an increasing function, $d\rho_0/dT$ has two maxima and two

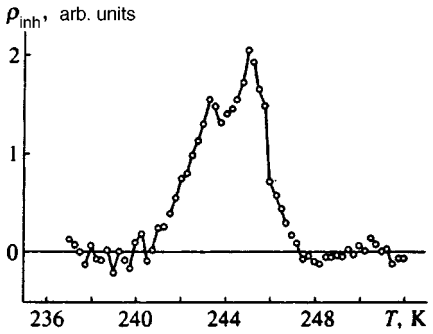


FIG. 5. Temperature dependence of the inhomogeneous part of the electrical resistivity of the $Fe_{0.95}Co_{0.05}Ge_2$ single crystal in the vicinity of a magnetic phase transition.

minima. But if $d\rho_h/dT$ is a decreasing function, $d\rho_0/dT$ has three maxima and three minima: one minimum near $T_{\min}$, another one at T somewhat higher than $T^{(1)}$, then at $T \approx T^{(2)}$ there is a maximum, and higher up on the temperature scale the third minimum. These ideas show that the percolation transition temperatures can be estimated from the position of the extrema of the derivative $d\rho_0/dT$ or from the position of the extrema of the temperature dependence of magnetoresistivity.

Outside the transition region $\rho_h = \rho_0$. Figure 4 shows that $\rho_0(T)$ in $Fe_{0.95}Co_{0.05}Ge_2$ is a monotonically increasing concave downward function, which means that $d\rho_h/dT$ decreases with increasing T, and because of this the derivative $d\rho_0/dT$ must have three minima, as observed.

To get an idea about $\rho_{\text{inh}}(T)$, we approximated the temperature dependence of ρ_0 over two temperature intervals, $237 \leqslant T \leqslant 241$ K and $248 \leqslant T \leqslant 252$ K, by a single fourth-degree polynomial $\rho_h^{\text{fit}}(T)$. Next we assumed that $\rho_h^{\text{fit}}(T)$ is a good approximation to $\rho_h(T)$ at all temperatures T from 237 to 252 K, after which we calculated the difference of $\rho_0(T)$ and ρ_h^{fit}, which approximates $\rho_{\text{inh}}(T)$. The result is depicted in Fig. 5. The values of ρ_{inh} are finite in the interval 241 to 247 K. Two maxima are clearly visible in ρ_{inh} vs. T: at $T = 243.2$ K and at $T = 245$ K.

The resulting curve representing $\rho_{\text{inh}}(T)$ has all the features that follow from the adopted model. Hence, according to the data obtained in electrical measurements, the transition region in the studied specimen of $Fe_{0.95}Co_{0.05}Ge_2$ extends from 241 to 247 K, and for $T^{(1)}$ and $T^{(2)}$ we can take the values corresponding to the positions of the central magnetoresistivity peak and the minimum between the central and right peaks, i.e., 244 K and 245 K, respectively.

The formation and decay of infinite clusters must affect not only the resistivity but also the magnetic susceptibility, since as the small droplets fuse into a large simply connected region extending through the entire specimen, the number of magnetic moments belonging to the interphase boundary surface must change discontinuously. Assuming that the contribution of the interphase boundary surface to the average magnetization m of the specimen is proportional to the area of this surface, we can write

$$m = m_{ht} v + m_{lt}(1 - v) + \alpha S, \tag{4}$$

where m_{ht} is the magnetization of the high-temperature phase, m_{lt} is the magnetization of the low-temperature phase, and α is a coefficient that can be assumed constant inside the transition region, just as m_{ht} and m_{lt} are. Combining Eqs. (2) and (4), we obtain

$$\frac{dm}{dT} = (m_{ht} - m_{lt}) W(T) + \alpha \frac{dS}{dT}. \tag{5}$$

Since $m_{ht} < m_{lt}$ in $Fe_{0.95}Co_{0.05}Ge_2$ and the W vs. T curve has its usual ''dome' shape, we can expect that inside the transition region the temperature dependence of dm/dT has a broad minimum, with the ''comb'' representing the temperature dependence of dS/dT superimposed on this curve.

As the inset in Fig. 1 shows, $d\chi/dT$ vs. T has exactly this shape. Two minima are clearly visible. The temperatures of the first minimum ($T = 243$ K) and the maximum that follows (at $T = 244-245$ K) are almost equal to the temperatures $T^{(1)}$ and $T^{(2)}$ determined from magnetoresistivity data. The minimum at the left edge of the transition region is absent due to the sharp increase in $W(T)$ in this region.

Note that the smearing of the transition calculated from magnetic measurements proves to be almost twice as large as that calculated from electrical measurements. The reason, most likely, is that the specimen's magnetic moment is more sensitive to the presence of small clusters of the other phase than the resistivity is, since the magnetic susceptibilities of the phases in $Fe_{0.95}Co_{0.05}Ge_2$ differ considerably.

5. CONCLUSION

In the vicinity of the ''smeared'' first-order magnetic phase transition in a $Fe_{0.95}Co_{0.05}Ge_2$ single crystal, we have detected three peaks of positive magnetoresistivity in the dependence of $\Delta\rho/\rho_0$ on T, and extrema in the temperature dependence of the temperature derivatives of electrical resistivity and magnetic susceptibility. We found that these features are correlated.

Analysis of the results shows that the presence of peaks of positive magnetoresistivity and the presence of singularities in the temperature dependences of $d\rho_0/dT$ and $d\chi/dT$ are probably a distinctive consequence of the ''smearing'' of the transition. This fact leads to the emergence inside the phase transition region of clusters of low- and high-temperature phases. Under temperature variations the system of clusters undergoes two percolation-like transitions. The area of the cluster boundary in these transitions changes abruptly, which probably explains the presence of the aforementioned singularities in the temperature dependence of $d\rho_0/dT$ and $d\chi/dT$, and the emergence of peaks in the temperature dependence of magnetoresistivity.

Our analysis of the reasons for the changes in resistivity and magnetization in the vicinity of ''smeared'' phase transitions is of a fairly general nature, and is not related to the specific features of the single crystal we studied.

[1] P. G. de Gennes and J. Friedel, J. Phys. Chem. Solids **4**, 71 (1958).
[2] S. Alexander, J. S. Helman, and I. Balberg, Phys. Rev. B **13**, 304 (1976).
[3] I. Balberg and J. S. Helman, Phys. Rev. B **18**, 303 (1978).
[4] K. B. Vlasov, E. A. Rozenberg, A. G. Titova, and Yu. M. Yakovlev, Fiz.

Tverd. Tela (Leningrad) **24**, 1338 (1982) [Sov. Phys. Solid State **24**, 759 (1982)].
[5] É. L. Nagaev, *The Physics of Magnetic Semiconductors* [in Russian], Nauka, Moscow (1979), pp. 344–368.
[6] R. I. Zaĭnullina, K. B. Vlasov, N. G. Bebenin, and V. V. Ustinov, Fiz. Tverd. Tela (St. Petersburg) **38**, 2831 (1996) [Phys. Solid State **38**, 1550 (1996)].
[7] T. I. Papushina and A. A. Frolov, Neorg. Mater. **28**, 608 (1984).
[8] I. I. Piratinskaya, G. P. Zinov'ev, A. V. Mikhel'son, and R. P. Krentsis, *Physics of Metals and Their Compounds* [in Russian], Issue 6, Ural State Univ. Press, Sverdlovsk (1978), p. 67.
[9] K. B. Vlasov, R. I. Zaĭnullina, and M. A. Milyaev, Fiz. Met. Metalloved. **75**, 65 (1993).
[10] K. B. Vlasov, R. I. Zaĭnullina, and V. N. Syromyatnikov, Fiz. Met. Metalloved. **61**, 1219 (1986).

Translated by Eugene Yankovsky

Ionic photoconductivity of $RbAg_4I_5$ superionic crystals

M. V. Bogatyrenko and S. I. Bredikhin

Institute of Solid State Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

(Submitted 6 December 1996)
Zh. Éksp. Teor. Fiz. **112**, 698–706 (August 1997)

A new effect of illumination on ionic conductivity and activation energy of migration of mobile Ag^+ cations in $RbAg_4I_5$ superionic crystals has been detected and studied. Reversible changes in the ionic conductivity due to illumination of superionic crystals are caused by reversible changes in the structure of electronic centers caused by elastic strain around these centers. The effect of elastic deformation on the process of ionic transport and activation energy for diffusion of mobile silver cations has been studied. Photostimulated recovery of the ionic conductivity after its change due to preliminary illumination of a $RbAg_4I_5$ superionic crystal with light of wavelength $\lambda \simeq 430$ nm has been detected. This recovery of the ionic conductivity is due to excitation of centers in complexes generated by previous illumination of tested samples.
 [S1063-7761(97)02208-7]

1. INTRODUCTION

Superionic conductors belong to a specific class of solids one of whose ionic sublattices is structurally disordered at temperatures considerably lower than the melting temperature. Investigation of superionic conductors attract a lot of attention because they have several unusual and even paradoxical properties, primarily an anomalously high ionic conductivity. The distinctive feature of superionic crystals is the presence of two types of charge carriers, namely electrons and ions. Interaction between the electronic and ionic systems results in several fundamentally novel phenomena being observed in superionic materials.[1,2] Nonetheless, no information about feasible causes and mechanisms of the effect of the electronic subsystem on ionic transport is available presently. The reported study is the first step in this direction. Our aim was investigation of one manifestation of interaction between the ionic and electronic subsystems, namely, the effect of excited electronic centers on the ionic transport in $RbAg_4I_5$.

Features of superionic materials show up most clearly in solids called ''materials with a melted cation sublattice.'' $RbAg_4I_5$ crystals, which are the most typical representatives of this group, have the minimal activation energy of mobile cations ($\Delta E_a \simeq 0.1$ eV) and the maximal ionic conductivity ($\sigma_i \simeq 0.32\ (\Omega \cdot \mathrm{cm})^{-1}$ at room temperature).[3] Depending on the temperature, $RbAg_4I_5$ can be in one of three phase states, α, β, or γ. The high-temperature cubic α-phase exists at temperatures above 209 K. As the temperature is lowered, a first-order phase transition to the rhombohedric β-phase occurs at $T_1 = 209$ K. Both α- and β-phases are superionic. At $T_2 = 122$ K a phase transition to the hexagonal, nonsuperionic γ-phase occurs.

Ion migration in superionic materials, just as in classical ionic crystals, can be formally described by the model of simple hops. The ionic conductivity as a function of temperature follows the Arrhenius law and is described by the formula

$$\sigma_i = n_i (Ze)^2 \frac{D_0 \exp(-\Delta E_a / kT)}{kT}, \tag{1}$$

where n_i is the concentration of conducting ions, Ze is the ion charge, and ΔE_a is the activation energy. In the superionic α-phase of $RbAg_4I_5$, all ($n_i = 10^{22}$ cm^{-3}) Ag^+ cations are disordered and contribute to the ionic conductivity.

Thus, the situation in superionic materials should be radically different from semiconductors, in which the primary effect of illumination on the conductivity is the change in the carrier concentration due to generation of nonequilibrium carriers. This means that the only feasible effect of illumination on the ionic conductivity of $RbAg_4I_5$ is a change in the mobility of Ag^+ cations. With a aim of detecting such a property, we have undertaken this study of an illumination effect on the ionic conductivity and activation energy of motion of Ag^+ cations.

2. ILLUMINATION EFFECT ON IONIC CONDUCTIVITY OF $RbAg_4I_5$ SUPERIONIC CRYSTALS

Experiments were conducted on $RbAg_4I_5$ single crystals grown by the zone melting technique. By using special techniques of purification and optimal crystallization conditions, the concentration of heavy-metal impurities was reduced below 10^{-4} percent. Measurements were taken using the four-terminal technique and an alternating current. Small silver contacts were melted into a sample immediately before an experiment. The contact quality was determined by recording conductivity versus frequency. The ionic conductivity was derived from the slope of the current–voltage characteristic at a frequency of $\sim 10^4$ Hz. In order to measure the ionic conductivity as a function of temperature, we placed the sample into an optical thermostat which allowed us to vary the temperature between 77 and 450 K. Samples were exposed to light generated by a DKSSh-120 xenon discharge tube and dispersed by a wide-aperture MDR-12 grating monochromator. An optical system collected light passing

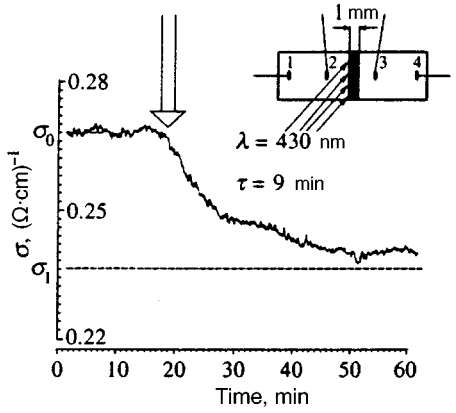


FIG. 1. Effect of light exposure on the ionic conductivity of a $RbAg_4I_5$ crystal (the arrow indicates the moment when light with wavelength $\lambda \simeq 430$ nm is turned on).

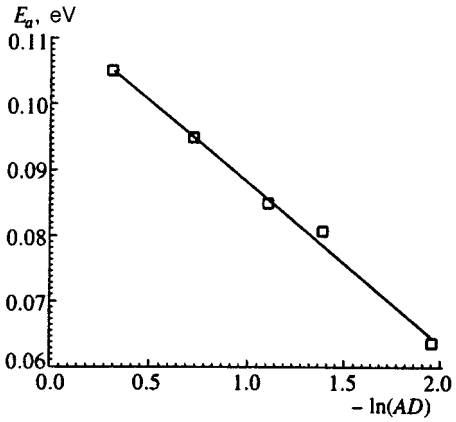


FIG. 3. Activation energy E_a versus logarithm of a preexponential factor, $\ln(AD_0)$, measured after exposing the sample to light in the process of its decolorization.

through the monochromator and focused it into a spot with dimensions of $\sim 0.5 \times 3$ mm in the region between the potential contacts (Fig. 1).

Our measurements indicate that exposure of $RbAg_4I_5$ crystals to light with wavelength $\lambda \simeq 430$ nm exciting electronic centers leads to reversible changes in the ionic conductivity. As an illustration, Fig. 1 shows that the exposure of a $RbAg_4I_5$ sample to light with wavelength $\lambda \simeq 430$ nm at room temperature leads to a decrease in the ionic conductivity. The time constant characterizing the transition to a new conductivity is $\tau \simeq 9$ min. Figure 1 shows that the new value of conductivity (σ_1) is essentially constant with time at a fixed light intensity. When the illumination is turned off, the ionic conductivity returns to the initial value σ_0. Note that the rate of transition in the latter case is slower, and the initial ionic conductivity is restored only after tens of hours.

To study the mechanism of the illumination effect on the activation energy for the diffusion of silver cations, we recorded the conductivity as a function of temperature in the range 280–420 K. Figure 2 shows the logarithm of the ionic conductivity versus reciprocal temperature in the initial state of a $RbAg_4I_5$ sample and after exposure to light with wavelength $\lambda \simeq 430$ nm. It clearly demonstrates that after the exposure the activation energy of the silver cation migration is

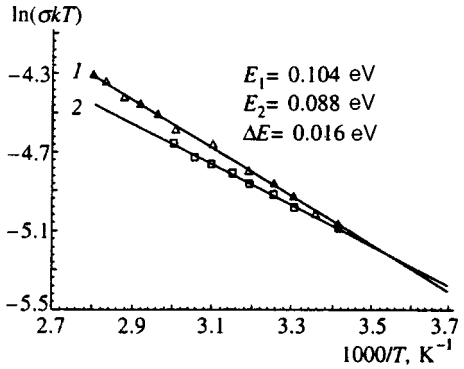


FIG. 2. Logarithm of ionic conductivity as a function of reciprocal temperature for $RbAg_4I_5$ (*1*) in the initial state and (*2*) after exposure to light with wavelength $\lambda \simeq 430$ nm.

lower. After switching off the light, the initial activation energy is restored. Measurements of the activation energy E_a and logarithm of the preexponential factor, $\ln(AD_0)$, during relaxation of the ionic conductivity to its initial value σ_0 indicate that their variations are correlated (Fig. 3) in accordance with the compensation law or the Meyer–Neldel rule[4]:

$$\ln(AD_0) = (43 \pm 3)\Delta E_a - (3.6 \pm 0.01). \tag{2}$$

Let us discuss feasible mechanisms of the light effect on the ionic conductivity of $RbAg_4I_5$ crystals. Previously[2] we detected an effect of photoinduced changes in the local concentration of mobile Ag^+ cations in a region of a $RbAg_4I_5$ superionic crystal exposed to light with $\lambda \simeq 430$ nm, which leads to excitation of electronic centers and generates nonequilibrium electrons in the conductance band. Given the high concentration of mobile Ag^+ cations in the superionic α-phase, they screen electrostatic interaction between ionized centers and photoexcited electrons so diffusion flows of electrons and cations from the exposed region into the sample volume are produced. Generation of silver cation vacancies $[V]^-_{Ag^+}$ and holes in the illuminated region of the crystal results in formation of $h^+[V]^-_{Ag^+}$ centers similar to color centers in additively colored $RbAg_4I_5$ crystals.[2,5,6] Note that the color center $h^+[V]^-_{Ag^+}$ consists of a hole localized at an iodine ion next to a silver cation vacancy. According to the literature, the iodine ion radius is 2.2 Å, whereas the iodine atomic radius is 1.33 Å. Hence, an elastic strain field is generated around such centers, and this field probably changes the ionic conductivity and activation energy for migration of Ag^+ cations in $RbAg_4I_5$ superionic crystals.

Now let us discuss the effect of internal elastic strain fields generated in a superionic crystal by light on ionic transport. Transport properties of ionic crystals are described by the most simple formula in the case of dc conductivity:

$$\sigma_i = \frac{N(Ze)^2 \nu_0 a_0^2}{kT} \exp\left(-\frac{\Delta G}{kT}\right), \tag{3}$$

where Ze is the conducting ion charge, $\Delta G = \Delta E + P\Delta V - T\Delta S$ is the Gibbs energy of defect formation and diffusion, ν_0 is typically comparable to the phonon frequency and

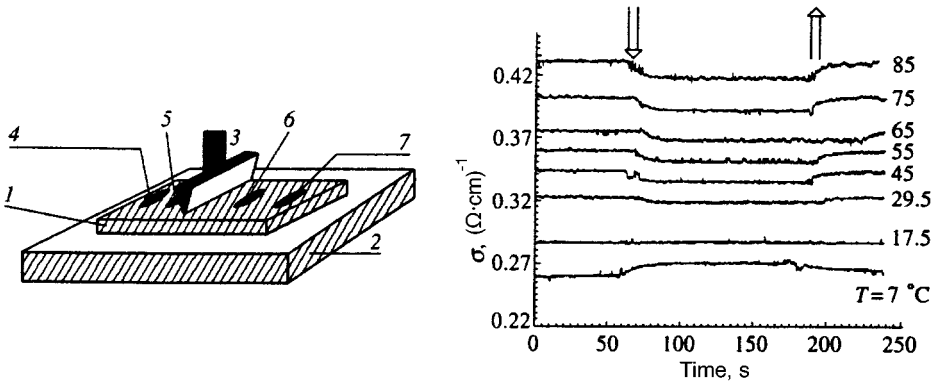


FIG. 4. Effect of external stress $P \simeq 1 \cdot 10^8$ Pa on the ionic conductivity at different temperatures. The moments of stress application and lifting are shown by the arrows.

is usually taken to be equal to the frequency of ion oscillation at a lattice site, and a_0 is the hop length. In superionic crystals, as in ionic crystals with impurity conductivity, the expressions for both ionic conductivity and activation volume ΔV_a do not contain contributions due to defect formation.[7,8] In this case, the expression for the ionic conductivity takes the form

$$\sigma_i = \frac{n_i (Ze)^2 \nu_0 a_0^2}{kT} \exp\left(\frac{S_m}{k}\right) \exp\left(-\frac{\Delta E_m}{kT}\right) \times \exp\left(-\frac{P \Delta V_m}{kT}\right), \tag{4}$$

where ΔE_m is the defect migration energy, and ΔV_m is the activation volume for the defect migration.[7,8] It is clear that in superionic conductors the effect of pressure on conductivity shows up in terms of the activation volume for defect migration.[7,8]

It follows from the expressions (1) and (4) for the ionic conductivity given above that the activation energy ΔE_a measured in experiments includes the migration energy ΔE_m of mobile cations and a fraction due to the activation volume for diffusion:

$$\Delta E_a = \Delta E_m + P \Delta V_m . \tag{5}$$

The activation volume is related to the ionic conductivity as a function of pressure by the following expression:

$$\Delta V_m = -kT \left[\left(\frac{d \ln \sigma}{dP} \right) - \left(\gamma - \frac{2}{3} \right) K \right], \tag{6}$$

where K is the isothermal compressibility constant and γ is the Grüneisen parameter.[7,8] Accurate measurements of activation volumes for diffusion in $RbAg_4I_5$ superionic crystals were conducted by Allen and Lazarus[7] and Samara.[8] Measurements of the activation volume in the low-temperature, nonsuperionic γ-phase yield $\Delta V_a \simeq 9$ cm^3/mol, which is comparable to the total activation volume in AgBr (10.6 cm^3/mol) and AgCl (11.6 cm^3/mol).[7] At the same time, migration volumes ΔV_m measured in $RbAg_4I_5$ superionic α- and β-phases turned out to be negative (-0.4 cm^3/mol and -0.2 cm^3/mol, respectively). The negative values of the activation volume for the diffusion of silver cations in the superionic α- and β-phases in $RbAg_4I_5$ has no unambiguous interpretation at present. One feasible explanation was given by Flygare and Haggins.[9] They demonstrated that the dimension of a conducting Ag^+ cation is slightly smaller than the optimal size of a channel in the three-dimensional lattice of iodine ions, through which silver cations travel.[10] In this case, an increase in pressure results in a smaller channel width and a ratio between the channel width and ion dimension closer to the optimal value, hence the lower activation energy and higher ionic conductivity.[11]

Our analysis given above indicates that photogenerated changes in the ionic conductivity of $RbAg_4I_5$ superionic crystals (Fig. 1) and decrease in the activation energy (Fig. 2) are due to elastic strain fields generated in the sample region exposed to light. As can be seen in Fig. 2, exposure of a $RbAg_4I_5$ superionic crystal to light with wavelength $\lambda \simeq 430$ nm results in an activation energy for migration of silver cations approximately 0.01 eV lower. Simple estimates indicate that the ~ 0.01-eV decrease in the activation energy corresponds to an elastic stress of about $P_{el} \simeq 5 \cdot 10^8$ Pa in the illuminated sample region.

3. EFFECT OF LOCAL ELASTIC STRESS ON IONIC CONDUCTIVITY OF SUPERIONIC CRYSTALS

In order to check the mechanism of the effect of local elastic strain on ionic transport, we have studied the effect of local elastic stress applied to the sample on the ionic conductivity and activation energy for diffusion of mobile silver cations in $RbAg_4I_5$ superionic crystals. The ac conductivity was measured (Fig. 4) by the four-terminal technique on samples shaped as plates *1* with dimensions of $0.5 \times 3 \times 8$ mm^3. The sample was placed on the surface of a sapphire substrate *2* in the optical thermostat. Strain was produced by a top die whose end-piece *3* was a sapphire prism with a 0.1×3 mm^2 pressing surface. External stress was applied to a sample region between the two potential contacts *4–7* with an area of 0.1×3 mm^2. The stress applied in experiments was lower than the $RbAg_4I_5$ microscopic strength $P \simeq 5 \cdot 10^8$ Pa.[12] Note that the experimental configuration in local strain studies (Fig. 4) was identical to that of experiments on the photoinduced ionic conductivity (Fig. 1).

The measurements demonstrated that the local strain in the crystal leads to a reversible change in the ionic conductivity (Fig. 4). When the external stress was lifted, the ionic conductivity recovered. Figure 4 shows measurements of the effect of external stress $P \simeq 1 \cdot 10^8$ Pa on the ionic conductivity at different temperatures. The moments when the stress

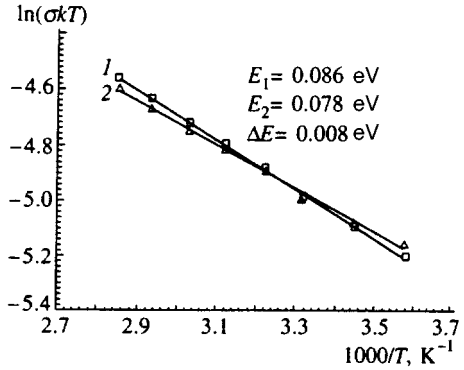


FIG. 5. Logarithm of ionic conductivity as a function of reciprocal temperature for the $RbAg_4I_5$ sample (*1*) in the initial state and (*2*) under local external stress $P \simeq 1 \cdot 10^8$ Pa.

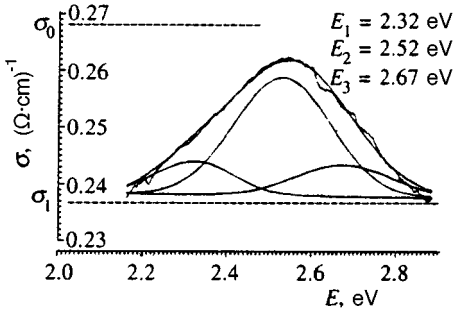


FIG. 6. Spectrum of ionic conductivity recovery (wavy line). The wavelength scanning rate is 0.5 nm/min; σ_0 is the ionic conductivity in the initial state, and σ_1 is the conductivity after an exposure to light with $\lambda \simeq 430$ nm for more than 0.5 h. The energies E_1, E_2, and E_3 are peak positions of the spectral components. The thin, smooth line shows the sum of the three spectral components and is almost identical to the recorded spectrum.

was applied and lifted are marked by jumps in the conductivity. Figure 4 shows that at temperatures higher than room temperature, local strain leads to a reversible decrease in ionic conductivity, and the amplitude $\Delta\sigma/\sigma_0$ drops with decreasing temperature. At a temperature of about 17.5 °C, external stress has little effect on ionic conductivity, and at temperatures below 17 °C it results in an increase in ionic conductivity. Figure 5 shows the logarithm of ionic conductivity versus reciprocal temperature in the initial $RbAg_4I_5$ sample ($E_0 \simeq 0.086$ eV) and under a local stress $P \simeq 1 \cdot 10^8$ Pa ($E_p \simeq 0.078$ eV). It is clear that the applied stress $P \simeq 1 \cdot 10^8$ Pa leads to a similar decrease in activation energy for diffusion of mobile silver cations ($\Delta E_a \simeq 0.008$ eV) as in exposure of the $RbAg_4I_5$ sample to light with wavelength $\lambda \simeq 430$ nm.

Thus, our testing experiments on the effect of external local stress confirm our assumption that the ionic photoconductivity is due to elastic stress P_{el} generated in the sample region exposed to light and changes in the ion diffusion coefficient caused by the change in the activation energy of the diffusion ($\Delta E_a = \Delta E_m + P_{el}\Delta V_m$).

4. EFFECT OF PHOTOSTIMULATED RECOVERY OF IONIC CONDUCTIVITY

Apart from ionic photoconductivity, we have detected in this study an effect of photostimulated recovery of ionic conductivity after its photoinduced change caused by exposure of a $RbAg_4I_5$ superionic crystal to light with wavelength $\lambda \simeq 430$ nm. As was shown above, ionic photoconductivity is due to long-lived centers in the form of donor–acceptor pairs including silver cation vacancies $[V]^-_{Ag^+}$ and hole centers generated in the illuminated region of the sample. Strain fields generated around such centers change the activation energy for diffusion of mobile silver cations. It is natural to expect that internal excitation of donor–acceptor pairs associated with charge transfer within the complex should result in lattice relaxation and changes in elastic strain fields around donor–acceptor pairs. As a consequence, one should detect an opposite change in the ionic conductivity when a $RbAg_4I_5$ sample is exposed to light generating internal excitations in complexes created by the previous illumination of the sample with light of wavelength $\lambda \simeq 430$ nm. Let us recall that illumination of $RbAg_4I_5$ samples at wavelength $\lambda \simeq 430$ nm at room temperature leads to a drop in ionic conductivity after about ten minutes (Fig. 1), whereas the recovery of ionic conductivity to its initial value σ_0 is a much slower process and takes tens of hours. Therefore, the decay of the concentration of donor–acceptor pairs generated by light with wavelength $\lambda \simeq 430$ nm in the tested samples during the first hour can be neglected, and the concentration of such complexes can be considered a constant.

Our experiments have shown that apart from changes in ionic conductivity induced by light of wavelength $\lambda \simeq 430$ nm, subsequent exposure of $RbAg_4I_5$ samples to light with wavelength $\lambda > 430$ nm leads to a partial, reversible recovery of conductivity. Note that the photoinduced recovery time of ionic conductivity to its new value is considerably shorter than the time constant of ionic photoconductivity. A spectrum of ionic conductivity recovery is given in Fig. 6. It shows that the exposure of $RbAg_4I_5$ samples to light with wavelength $\lambda > 430$ leads to a partial recovery of conductivity.

The experimental recovery spectrum was approximated as a sum of three Gaussian curves. One can see that the spectrum contains three bands with peaks at $E_1 = 2.32$ eV, $E_2 = 2.52$ eV, and $E_3 = 2.67$ eV and FWHM $\Delta W_1 = 0.13$ eV, $\Delta W_2 = 0.16$ eV, and $\Delta W_3 = 0.16$ eV, respectively. The peak positions and FWHM of these spectral ranges are in good agreement with absorption bands associated with additive coloring[13] and photoinduced coloring[14] of $RbAg_4I_5$ superionic crystals.

This coincidence of absorption spectra of additively colored and optically colored $RbAg_4I_5$ crystals with the spectrum of ionic conductivity recovery indicates that latter is caused by internal excitation of complexes generated by the preliminary illumination of the samples ($\lambda \simeq 430$ nm). Internal excitation of donor–acceptor pairs leading to charge transfer within the complexes and variations in local strain fields result in changes in the activation energy for diffusion of mobile silver cations and partial recovery of the ionic conductivity altered by previous exposure to light of wavelength $\lambda \simeq 430$ nm.

5. CONCLUSIONS

This paper reports on the first observation and investigation of ionic photoconductivity in superionic crystals. We have determined that ionic photoconductivity is due to generation of long-lived centers in the form of donor–acceptor pairs including silver cation vacancies $[V]^{-}_{Ag^+}$ and hole centers in the region of the $RbAg_4I_5$ sample exposed to light. The elastic strain fields around such complexes change the activation energy for diffusion of mobile silver cations. We have studied the effect of external stress on ionic conductivity and demonstrated that deformation of a small region in a crystal leads to a reversible change in ionic conductivity. The coincidence between measurements of ionic photoconductivity and results of testing experiments with local stress confirms that ionic photoconductivity is due to local elastic stress P_{el} in the region of a superionic crystal exposed to light. We have detected and studied the partial photoinduced recovery of ionic conductivity and investigated its spectral characteristics. We have shown that ionic conductivity recovery is due to internal excitations in complexes generated by previous illumination of tested samples at wavelength $\lambda \simeq 430$ nm.

This work was supported by the Russian Fund for Fundamental Research (Project No. 95-02-06106).

[1] S. Bredikhin, T. Hattori, and M. Ishigame, Phys. Rev. B **50**, 2444 (1994).
[2] S. Bredikhin, N. Kovaleva, T. Hattori, and M. Ishigame, Solid State Ionics **74**, 149 (1994).
[3] B. B. Owencs and G. R. Argue, Science **157**, 308 (1967).
[4] S. I. Bredikhin, N. N. Kovaleva, and N. V. Lichkova, Zh. Éksp. Teor. Fiz. **96**, 735 (1989) [Sov. Phys. JETP **69**, 417 (1989)].
[5] V. N. Andreev and V. G. Goffman, Fiz. Tverd. Tela (Leningrad) **25**, 3480 (1983) [Sov. Phys. Solid State **25**, 2004 (1983)].
[6] S. I. Bredikhin, N. N. Kovaleva, N. V. Lichkova, and I. Sh. Khasanov, Fiz. Tverd. Tela (Leningrad) **30**, 1901 (1988) [Sov. Phys. Solid State **30**, 1099 (1988)].
[7] P. C. Allen and D. Lazarus, Phys. Rev. B **17**, 1913 (1978).
[8] G. A. Samara, Solid State Phys. **38**, 1 (1984).
[9] W. H. Flygare and R. A. Haggins, J. Phys. Chem. Solids **34**, 1199 (1973).
[10] S. Geller, Science **157**, 310 (1967).
[11] Yu Yao and J. T. Kummer, J. Inorg. Nucl. Chem. **29**, 2453 (1967).
[12] Yu. M. Gerbshteĭn, E. I. Nikulin, and F. A. Chudnovskiĭ, Fiz. Tverd. Tela (Leningrad) **25**, 1148 (1983) [Sov. Phys. Solid State **25**, 659 (1983)].
[13] S. I. Bredikhin, N. N. Kovaleva, I. Sh. Khasanov, and N. V. Lichkova, Solid State Ionics **28**, 280 (1988).
[14] N. Kovaleva, A. Boris, S. Bredikhin, and T. Awano, Radiat. Eff. Defects Solids **134**, 457 (1995).

Translation was provided by the Russian Editorial office.

Observation of the dielectric local mode related to divacancies in *p*-silicon

V. A. Voĭtenko and S. E. Mal'khanov

St. Petersburg State Technical University, 195251 St. Petersburg, Russia
(Submitted 7 December 1996)
Zh. Éksp. Teor. Fiz. **112**, 707–713 (August 1997)

A dielectric local mode related to the motion of divacancies in *p*-silicon has been observed for the first time. The mode manifests itself in Fano-resonance signals in the photoconductivity spectra. We explain the behavior of the corresponding segment in these spectra caused by temperature variations. In light of the new results, we examine the entire set of the experimental facts related to positively charged divacancies: the high values of the cross sections of capture and photoionization of a hole at a single divacancy, the difference in the defect environment of electron and hole silicon, and the quadratic dependence of the concentration of defects containing divacancies on the intensity of the electron flux producing these defects.
 [S1063-7761(97)02308-1]

1. INTRODUCTION

The unique photoelectric properties, the abundance in the earth's crust, and the high adaptability to manufacture processes allows silicon to retain its position in micro- and optoelectronics along with gallium arsenide and related compounds. Ge/Si superlattices are grown by applying synchrotron radiation,[1] technologies used in heteroepitaxially growing GaAs-based materials on silicon substrates are developed,[2] and silicon specimens are synthesized, erbium-doped[3] or in the porous modification,[4] in which hot luminescence of the visual and IR ranges has an intensity sufficient for research and application. This makes various flaw-detection studies of doped silicon specimens highly important. Fano-resonance spectroscopy of the photoconductivity of various impurity centers, chalcogens[5,6] and Zn, Sb, and In acceptors,[7] has been developed through the use of uniaxial deformations of both silicon and germanium. Over the years there have also been studies of the photoconductivity[8] and photocapacitance[9] of deep energy levels of *p*-silicon irradiated by high-energy electrons. The photoionization cross-section spectrum contains a resonance band that at peak intensity corresponds to the wavelength $\lambda = 4\ \mu$m (see Refs. 8 and 9). This band was first identified by Vajda and Cheng[8] as a sudden rise in luminescence emission during annealing the specimens in the temperature interval from 150 to 180 K, in the course of which single vacancies begin to migrate, which leads to divacancy formation.

Note that divacancy signals can also be reliably identified in the spectra of *n*-silicon.[10] In *p*-Si, the radiatively induced defect, a divacancy, is in a positively charged state W^+ (see Ref. 9). In their theoretical research, Karpov *et al.*[11,12] associated the corresponding spectral band at $\lambda = 4\ \mu$m with intracenter energy transitions within separate divacancies, which in view of this could be interpreted as micropores ''frozen'' into silicon. However, such an interpretation leads to a contradiction between the experimental data on ionization energies obtained from optical[11] and thermal measurements.

The present paper reports on the results of an experimental study and calculations of the temperature dependence of the divacancy contribution to the photoconductivity with a spectral peak at $\lambda = 4\ \mu$m in the 30–70 K temperature range. We have found for the first time that the divacancies W^+ must be interpreted not as static micropores but as effective microcavities for the optical phonons in silicon, since the peak in light absorption in which they participate ($\lambda = 4\ \mu$m) coincides with the production threshold of optical phonons on these divacancies; these phonons are dielectric local modes.

2. SAMPLE PREPARATION AND EXPERIMENTAL METHOD

The specimens of boron-doped silicon (KDB-10 brand) are p^+pp^+ sandwich plates that measure $5\times2\times0.2$ mm. The outer p^+ regions are heavily doped to metallic state with a boron concentration 10^{17}–10^{18} cm^{-3}, while the central region is lightly doped and contains 1.3×10^{15} boron atoms per cubic centimeter. To create divacancies the specimens were irradiated by 15-MeV electrons with a total dose of 4×10^{16} electrons per square centimeter. An LUÉ-15 linear accelerator was used to produce the necessary electron flux. The interested reader can find a detailed description of this method in Ref. 13. The direction of electron propagation is shown in Fig. 1a by solid arrows. Electrons with such a high energy traveled right through the plate. As a result the acceptor impurities, those in the surface layer and in the bulk, were almost completely balanced by the vacancies uniformly distributed throughout the plate. As a result of formation of positively charged divacancies, static polarization (whose direction is depicted by dashed arrows in Fig. 1a) is imbedded in the semi-isolated inner region. The polarization causes the bands to bend in the direction of the heavily doped outer layers, as shown in Fig. 1b. As a result of such bending the thermal band gap proves to be the smallest. The corresponding hole transitions from the divacancy levels are depicted by the wavy line in Fig. 1b. By measuring the temperature dependence of the plate's resistance between the contacts K, we were able to estimate the value of the bend by the difference of the thermal and optical (dashed line) energies of charge exchange on the divacancy; it turned out to be 100 meV.

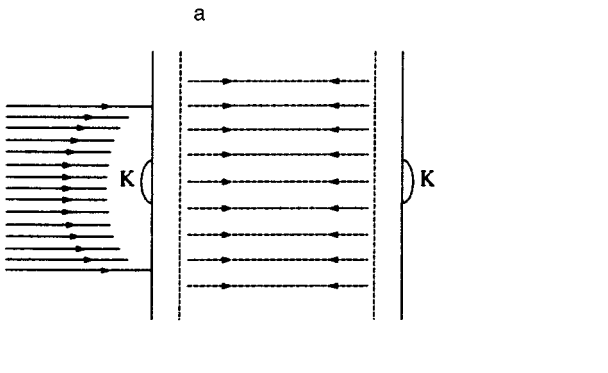


FIG. 1. (a) Structure of a p^+pp^+-Si sandwich plate with imbedded polarization (dashed arrows) and contacts K. The solid arrows show the direction of the high-energy electron flux. (b) Schematic band diagram of the sandwich structure and the hole transitions from the ground state to a divacancy in the band: thermal transitions are designated by a wavy line, and optical transitions by a dashed line.

Photoconductivity measurements were made with an IKS-21 spectrometer with a NaCl prism in the DC or matched-load mode.

3. EXPERIMENTAL RESULTS AND DISCUSSION

Figure 2 depicts the behavior of segments of the photoconductivity spectra under variations of the temperature of a p-Si sample irradiated by electrons. The curve *1* obtained at a specimen temperature $T_1=34$ K has a peak at $\omega_1=0.26$ eV (or $\lambda=5$ μm), corresponding to an edge hole transition denoted by the dashed line in Fig. 1b. With allowance for a small blue shift of 30 meV caused by the renormalization of the edge due to local mode formation, the frequency agrees with the position of the energy level of the holes on the divacancy W^+. According to relaxation measurements,[11] the energy of this level is $E_v=0.23$ eV, as shown in Fig. 1b. As the temperature grows, the shape of this section of the spectrum changes considerably (see curves *2* and *3* in Fig. 2). The former peak at $\omega_1=0.26$ eV is replaced by a stronger band with a peak at a frequency ω_2 shifted to higher values in relation to ω_1 by precisely the frequency of the center-band optical phonon: $\omega_2=\omega_1+\omega_{\rm opt}$. In Fig. 2 this new peak is indicated by downward arrows on curves *2* and *3*, corresponding to the temperatures $T_2=44$ K and $T_3=55$ K. Curve *2* is reconstructed on a linear scale in the inset; so as not to clutter up the figure, the scale grid on the axes has been discarded, but still one can get an idea of the scale from the experimental points.

It is this strong line with $\lambda=4$ μm that is well known and has been discussed in many papers.[8,9,11,12] The line corresponds to a transition with energy $\Delta E=0.32$ eV that takes place deep in the valence band, 90 meV from the edge. If we allow for the renormalization energy of the edge of the valence band, 30 meV (which we found earlier), we find that ω_2 corresponds exactly to the threshold energy of emission of an optical phonon in silicon, $\omega_{\rm opt}=90-30=60$ [meV]. In view of this, the band at $\lambda=4$ μm should be interpreted not as a resonance on a quasidiscrete level, as was done in Ref. 12, but as a one-phonon repetition of a simple charge exchange, depicted in Fig. 1b by a dashed line. The one-phonon nature of the line being discussed is corroborated by the shape of the spectrum, which along the peak can be approximated by the Fano profile (solid curve in inset; see, e.g., Ref. 5):

$$J_{\rm ph}(\omega)=M_\omega^2\frac{\Gamma_1}{\Gamma}\frac{(\varepsilon+q)^2}{\varepsilon^2+1}, \tag{1}$$

where $\varepsilon=(\omega-\omega_{\rm opt}-\omega_1)/\Gamma$ is the energy of a hole transition measured from the bottom of the valence band with allowance for the renormalization discussed earlier, $\Gamma=\omega_{\rm opt}$ Im π, where π is the polarization operator of the optical phonon, M_ω is the transition matrix element, and Γ_1 and q are the Fano profile parameters. The profile factor $q=-2.43$ and the damping factor $\Gamma=33$ meV were determined for curve *2* from the difference of the right and left halfwidths without using fitting methods (see, e.g., Ref. 14).

The experimental points on the temperature dependence of the photocurrent in Fig. 3 are well-approximated at low temperatures by an activation exponential at the optical phonon frequency, which is an indication of a ''freezing'' effect:

$$J_{\rm ph}^{\max}=I_0\ \exp\left(-\frac{\hbar\omega_{\rm opt}}{T}\right). \tag{2}$$

The exponential temperature dependence (solid curve) for the peak photocurrent $J_{\rm ph}^{\max}$ (Eq. (2)) is accompanied by a

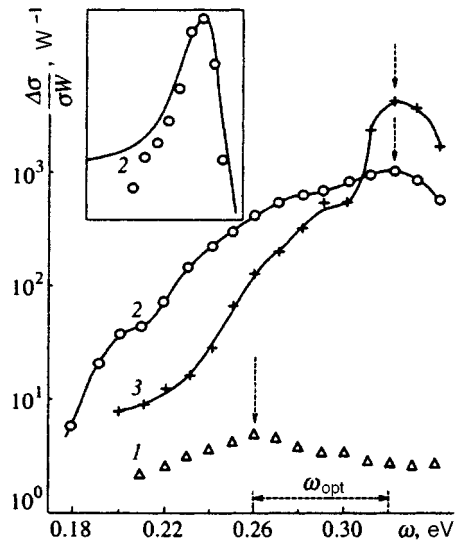


FIG. 2. Photoconductivity spectra at different temperatures: $T_1=34$ K, $T_2=44$ K, and $T_3=55$ K (curves *1*, *2*, and *3*, respectively). The solid curves in the main part of the figure are not approximations. Inset shows the photoconductivity spectrum in the vicinity of the peak for curve *2* on the linear scale (small circles) and its approximation by the Fano profile (solid curve).

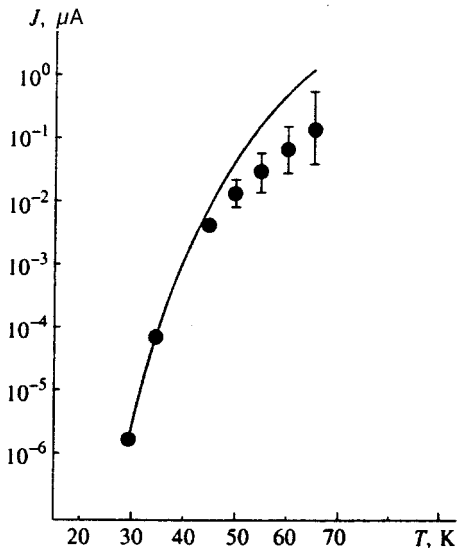


FIG. 3. Temperature dependence of the peak photocurrent according to the spectra depicted in Fig. 2: the solid curve represents the calculated curve, and the heavy dots are the experimental results.

decrease in the phonon damping factor Γ, which leads to a narrowing of the line with increasing temperature. Such characteristic behavior of phonon lines is called temperature or collision narrowing (see, e.g., Ref. 15) and can be caused by either a decrease in the difference of the occupation numbers of the initial and final hole states contributing to the phonon self-energy $\pi(\omega)$, or a decrease in this energy due to the transition of mobile charges into the collisional mode of motion. Upon a further increase in temperature, the variations in the photocurrent become smoother in comparison to (2). The increase in the experimental error in signal measurement visible in Fig. 3 is due to the increase in dark current with specimen temperature.

Another feature of the investigated spectra is that one-photon repetitions *2* and *3* in Fig. 2 are several orders of magnitude more intense than the nonphonon peak (curve *1*). This experimental fact contradicts the ordinary perturbation theory, according to which repetitions and satellites are a reflection of higher-order processes, compared to the initial lines. For instance, in the case of fundamental absorption in a pure CdTe crystal, discussed in Ref. 16, the intensity ratio is reversed: phonon repetition is much weaker. This behavior was partially explained in Ref. 16 by the exciton effect. No such effect is present in the ionization of a divacancy, since the latter remains neutral in the process. A possible explanation of the strengthening of one-photon repetitions of the photoionization of deep divacancy levels is that the positions of the divacancies are correlated in some way, and that this correlation leads to a mutual constructive effect of these divacancies in photon emission.

One of the present authors (S.E.M) in collaboration with other researchers[13] measured the dependence of the concentration of radiation-induced defects containing divacancies on the intensity of electron irradiation. The dependence proved to be quadratic, both for pure divacancies and for C+O+*W* complexes known as *K*-centers. Such a dependence and single-divacancy annealing data[8] indicate that the formation of divacancies occurs mainly because of fusion of the primary close pairs of single divacancies. Since band holes are much more mobile than vacancies, in the time preceding the fusion of a pair of vacancies, τ_v, one of the vacancies is able to capture a hole. This is possible if $\tau_v \gg \tau_p$, where τ_p is the characteristic time of hole capture by a single vacancy. As a result, one of the vacancies in an initial pair in *p*-Si is with an overwhelming probability in a charged state before fusion. Here, according to Keldysh and Proshko,[17] no random spatial distribution of mobile particles (in our case these are vacancies and holes interacting by the Coulomb law) is possible. The particles screen each other.

After immobile divacancies have formed, the Debye–Hückel correlation becomes ''frozen.'' Roughly speaking, this means that the divacancies are separated by approximately the same distances, with the distance being of order of the Debye length r_D. If the phonons emitted by individual divacancies have wave vectors $k \sim r_D^{-1}$, constructive interference may emerge, and this explains the anomalously strong signal at the wavelength $\lambda = 4$ μm and the giant photoionization cross section at the peak, 4×10^{-16} cm^2 (see Ref. 9). At the same time, the large-scale impurity potential generated by charged deep centers leads to ''tails'' in the density of hole states in the forbidden band (see, e.g., Ref. 17), which considerably weaken the resonant amplification of a nonphonon transition. In particular, such suppression of optical resonances leads to a situation in which observing Raman scattering in specimens of heavily or very heavily doped silicon[18] and germanium[19] is entirely out of the question. In view of this, the anomalously high cross section of hole capture by a divacancy, 10^{-14} cm^2 (see Ref. 11), can be explained by the presence of lattice vibration antinodes localized at the divacancies.

The relative intensity of the divacancy-associated contributions to the photocurrent in comparison to other contributions proved to be higher in our measurements (see Fig. 2) than in those conducted by Vajda and Cheng,[8] where the nonphonon peak was even unresolved. The explanation is that in their experiment, defect formation was slowed down by the low temperature at which the specimens were irradiated. The electrostatic polarization in our p^+pp^+ plates, which was absent in Vajda and Cheng's experiment,[8] stimulates the migration of vacancies and the defect-formation process that follows.

One of the possible objections to the proposed ''interference'' mechanism of the redistribution of the intensity of photoconductivity signals among a simple charge exchange in a divacancy and its photon repetition is the lack of such signals from *n*-Si specimens irradiated by electrons. Here the divacancies are studied by the EPR[20] and DLTS[10] methods rather than by measuring the photoconductivity, although the phonons are the very same ones. As an argument against this objection we note that for the conduction electrons in silicon, τ_e can be either shorter or longer than τ_v, since in view of their small effective mass the electrons behave as ultraquantum particles when they are captured by divacancies. Hence, in addition to charged divacancies, irradiated *n*-Si has neutral divacancies occupying totally uncorrelated positions. Suppression of the interference mechanism of amplification of

phonon repetitions makes the observation of electron photoconductivity resonance impossible. The cross section of electron capture by neutral divacancies, $\sigma_n \sim 10^{-15}$ cm^2 (see Ref. 20), is not anomalously high.

To explain the exponential dependence of the photocurrent on temperature, Vajda and Cheng[8] introduced shallow trap levels at an energy $E_v + 0.05$ eV. One must bear in mind, however, that such a temperature dependence of the photocurrent (Eq. (2)) is observed not only in boron-doped silicon, where at low boron concentrations there is indeed an acceptor level near that energy, but also in aluminum-doped silicon, where the acceptor level lies somewhat deeper. In reality, the role of traps probably consists in initiating a more rapid (in comparison to Eq. (1)) decrease in photocurrent as we move down the frequency scale away from the resonance frequency: as the temperature decreases, phonon emission is followed by hole capture in traps.

In conclusion, we note that the frequency of dielectric local modes, not established in our research, can be found from measurements of the low-frequency satellites of one-phonon Raman scattering.[16,21] In the same spectra, in the deep-inelastic region, one can hope to register the charge exchange proper and its phonon repetitions.

[1] Housel Akazawa, J. Appl. Phys. **79**, 9396 (1996).
[2] Maoshang Hao, Chunlin Shao, Tetsuo Soga *et al.*, Jpn. J. Appl. Phys., Part 2 **35**, L960 (1996).
[3] N. T. Bagraev, D. E. Onopko, A. I. Ryskin, and Yu. A. Karpov, Fiz. Tekh. Poluprovodn. **30**, 1855 (1996) [Semiconductors **30**, 967, 970 (1996)].
[4] Yoshihiko Kanemitsu, Phys. Rep. **263**, 1 (1995).
[5] E. Janzen, G. Grossmann, R. Stedman, and H. G. Grimmeiss, Phys. Rev. B **31**, 8000 (1985).
[6] H. G. Grimmeiss, L. Montelius, and K. Larsson, Phys. Rev. B **37**, 6916 (1988).
[7] G. Piao, R. A. Lewis, and P. Fisher, in *Proc. 21st Int. Conf. on the Physics of Semiconductors*, Vol. 2, World Scientific, Singapore (1992), p. 1609.
[8] P. Vajda and L. J. Cheng, J. Appl. Phys. **42**, 2453 (1971).
[9] N. V. Kolesnikov, A. A. Lebedev, S. E. Mal'khanov, Fiz. Tekh. Poluprovodn. **13**, 812 (1979) [Sov. Phys. Semicond. **13**, 479 (1979)].
[10] Zhi-pu You, Xuemei Gao, Xiatong Lin *et al.*, in *Proc. 21st Int. Conf. on the Physics of Semiconductors*, Vol. 2, World Scientific, Singapore (1992), p. 1677.
[11] V. G. Karpov, N. V. Kolesnikov, S. E. Mal'khanov, Fiz. Tekh. Poluprovodn. **16**, 1657 (1982) [Sov. Phys. Semicond. **16**, 1059 (1982)].
[12] V. G. Karpov and N. V. Kolesnikov, Fiz. Tekh. Poluprovodn. **12**, 583 (1978) [Sov. Phys. Semicond. **12**, 338 (1978)].
[13] N. V. Kolesnikov, V. N. Lomasov, S. E. Mal'khanov *et al.*, Fiz. Tekh. Poluprovodn. **18**, 1496 (1984) [Sov. Phys. Semicond. **18**, 936 (1984)].
[14] M. Jouanne, R. Beserman, I. P. Ipatova, and A. V. Subashiev, Solid State Commun. **16**, 1047 (1975).
[15] M. Cardona and I. P. Ipatova, in *Elementary Excitations in Solids*, J. L. Birman, C. Sebenne, and R. F. Wallis (Eds.), Elsevier, Amsterdam (1992), p. 237.
[16] Y. B. Levinson and E. I. Rashba, Rep. Prog. Phys. **36**, 1499 (1973).
[17] L. V. Keldysh and G. P. Proshko, Fiz. Tverd. Tela (Leningrad) **5**, 3378 (1963) [Sov. Phys. Solid State **5**, 2481 (1963)].
[18] A. Compaan, G. Contreras, M. Cardona *et al.*, J. Phys. (Paris) **44**, C5-197 (1983).
[19] A. K. Sood, G. Contreras, and M. Cardona, Phys. Rev. B **31**, 3760 (1985).
[20] J. Bourgoin and M. Lannoo, *Point Defects in Semiconductors II. Experimental Aspects*, Springer-Verlag, Berlin (1983), p. 269.
[21] J. Monecke, W. Cords, G. Irmer *et al.*, Phys. Status Solidi **142**, 237 (1987).

Translated by Eugene Yankovsky

Segregation of impurities and vacancies on phase and antiphase boundaries in alloys

K. D. Belashchenko and V. G. Vaks

Kurchatov Institute Russian Scientific Center, 123182, Moscow, Russia
(Submitted 31 December 1996)
Zh. Éksp. Teor. Fiz. **112**, 714–728 (August 1997)

The equilibrium distribution of low-concentration impurities or vacancies is investigated in the region of a coherent phase boundary or antiphase boundary in a binary alloy. A general expression for the free energy of an inhomogeneous multicomponent alloy, which generalizes the expression previously derived for a binary alloy, is presented. Explicit formulas for the impurity concentration profile $c_{\rm im}(x)$ in terms of the distribution of the principal components of the alloy near a boundary are obtained from this expression in the mean-field and pair-cluster approximations. The shape of this profile is determined by a ''preference potential'' P, which characterizes the attraction of an impurity to one of the alloy components, as well as by the temperature T and the phase transition temperature T_c. At small values of P/T impurities segregate on a phase boundary, and the degree of this segregation, i.e, the height of the maximum of $c_{\rm im}(x)$, in the region of the boundary increases exponentially as the ratio T_c/T increases. For $P \neq 0$ the $c_{\rm im}(x)$ profile near a phase boundary is asymmetric, and as P/T increases, it takes on the form of a ''worn step.'' The maximum on the $c_{\rm im}(x)$ curve then decreases, and at a certain $|P| \gtrsim T_c$ it vanishes. Segregation on an antiphase boundary is investigated in the case of CuZn ordering in a bcc alloy. The form of $c_{\rm im}(x)$ near an antiphase boundary depends significantly both on the form of the potential P and on the stoichiometry of the alloy. At small P impurities segregate on an antiphase boundary, and at fairly large P ''antisegregation,'' i.e., a decrease in the impurity concentration on the antiphase boundary in comparison with the value within the antiphase domains, is also possible. [S1063-7761(97)02408-6]

1. INTRODUCTION

The segregation of impurities on various interfaces is known and is important for many problems in solid-state physics. One well-known example is grain-boundary segregation, i.e., an increase in the impurity concentration on grain boundaries in a polycrystal in comparison with the value within the grains.[1] This type of segregation is usually attributed to the attraction of impurities to a grain boundary, in analogy to the formation of ''Cottrell impurity clouds'' around dislocations. A grain boundary is regarded as a series of edge dislocations. This model is correct and quantitative for small-angle boundaries. Then, if it is assumed that the location of impurities in a lattice is determined mainly by the geometric factor, the region of local contraction near a dislocation core should attract ''small'' impurities, whose radius $r_{\rm im}$ is less than the radius of the main component r_0, and the region of local expansion should attract ''large'' impurities with $r_{\rm im} > r_0$, so that impurities of both these categories should segregate on a boundary. The occurrence of grain-boundary segregation has been confirmed by many experiments and is important for a long list of properties of polycrystals, including their mechanical and corrosion properties.[1]

Segregation phenomena on boundaries not associated with the presence of dislocations or other lattice defects have been studied to a considerably smaller extent. Such boundaries include the phase boundaries appearing during the coherent decomposition of an alloy in a region of different concentration and antiphase boundaries, which coherently join differently ordered regions or antiphase domains in an ordered alloy. The existence of some universal segregation mechanism like the attraction of impurities to dislocations is not evident here. At the same time, the segregation occurring on such boundaries can significantly influence the properties of real alloys undergoing decomposition or ordering.

Vacancies comprise a special case of impurities in alloys. Their possible segregation on phase boundaries or antiphase boundaries has been discussed in a number of papers[2–8] in connection with phase transformations involving the decomposition or ordering of alloys. Interest in this problem arises because the redistribution of atoms among lattice sites during such phase transformations actually takes place specifically according to a vacancy mechanism, in which an atom of one of the principal components of the alloy, for example, A or B in a binary AB substitutional alloy, changes places with a neighboring vacancy. Therefore, the effects of the inhomogeneous distribution of vacancies, particularly their possible segregation on phase boundaries during the growth of new phases when an alloy decomposes or on antiphase boundaries during the growth of antiphase domains upon ordering, can significantly influence the evolution of microstructures. At the same time, these effects are not taken into account in the simple model of the ''direct exchange'' of places by atoms of A and B that has been used in most of the theoretical studies.

However, the segregation of vacancies on phase and antiphase boundaries was discussed in Refs. 2–8 only on a qualitative level and only for certain special models, and the conclusions drawn by different authors regarding the degree

 1063-7761/97/080390-09$10.00

of segregation differ markedly. For example, following the Monte Carlo simulation of spinodal decomposition in a two-dimensional model of an AB alloy with vacancies (an ABv alloy), Yaldram and Binder[2,3] concluded that no appreciable segregation occurs on the phase boundaries at a low vacancy concentration, although the images of the simulated structures that they presented do not rule out such segregation. Fratzl and Penrose[4] noted that for the model of an ABv alloy with nearest-neighbor interactions obeying the relations $V^{AA}=V^{BB}=V^{Av}=V^{Bv}=V^{\mathrm{vv}}=0$ and $V^{AB}>0$ used in Ref. 2, the need for the segregation of vacancies on the phase boundaries follows already from thermodynamic arguments, since some of the ''repulsive'' $A-B$ bonds are then broken. The Monte Carlo simulation in Ref. 4 provided indirect confirmation of the occurrence of such segregation, but without quantitative evaluations. In Refs. 5 and 6 Monte Carlo simulation of the kinetics of $B2$ (CuZn in Ref. 9) ordering was carried out for two-dimensional and three-dimensional models of an alloy with nearest-neighbor interactions conforming to the relations $V^{AA}=V^{BB}=V^{Av}=V^{Bv}=V^{\mathrm{vv}}=0$ and $V^{AB}<0$. The occurrence of segregation on the antiphase boundaries was demonstrated, but its quantitative characteristics were not discussed. Geng and Chen[7,8] considered the segregation of vacancies in the two-dimensional models from Refs. 2 and 5 on the basis of an approximate equation for the evolution of local concentrations previously proposed by Chen.[10] We note that the meaning of this equation is not entirely clear, since its stationary solutions do not translate into the equilibrium results corresponding to the mean-field and pair-cluster approximations used. Geng and Chen[7,8] obtained very large values for the segregation of vacancies on phase and antiphase boundaries, which were an order of magnitude greater than the values obtained in the Monte Carlo simulation in Ref. 2 and in the present work. This is evidently another indication of the unreliability of the equation proposed in Ref. 10.

A systematic approach to the description of the distributions of local concentrations in alloys for an arbitrary deviation from equilibrium was recently developed in Refs. 11 and 12. The application of that approach to investigations of the kinetics of phase transformations led to some new results.[12–15] In the present work this approach is applied to the investigation of the segregation of impurities on phase and antiphase boundaries.

2. FREE ENERGY OF AN INHOMOGENEOUS MULTICOMPONENT ALLOY

Let us first discuss the problem under consideration and the assumptions used in greater detail. We treat a ternary substitutional alloy, in which the concentration of one of the components, say, the vacancies in an ABv alloy, is so small that the interaction of these impurities with one another and their influence on the thermodynamics of the phase transformations can be neglected. Then the equilibrium distribution of such impurities is determined only by their interaction with the principal components of the alloy. Furthermore, if there are several kinds of such impurities, their distributions do not depend on one another, and the expressions presented below can be employed for the impurities of each kind. To simplify the notation we shall use the letter v to denote impurities of any kind, not just vacancies. It is assumed that a boundary between two equilibrium phases appears as a result of a phase transformation that takes place at a temperature greater than the temperature T under consideration. We call this boundary a phase boundary, if it separates unordered phases with different values of the concentration c, and an antiphase boundary, if the regions on opposite sides of the boundary correspond to antiphase domains with different orientations, but the same mean concentration c in the sublattices. For simplicity, the case of a phase boundary between unordered and ordered phases (which is possible for phase transformations involving decomposition with ordering[14]) will not be discussed, although the general formulas in the present section can be applied to that case. The boundary considered is assumed to be stationary, and the distributions of both the principal components A and B and the impurity atoms near it are assumed to be the equilibrium distributions. In the case of moving boundaries this assumption means that the relaxation times of these atomic distributions are presumed to be much shorter than the times for significant displacements of the boundary, as is usually the case for the boundaries of small curvature considered.[13]

To investigate both the problem under discussion and other problems in the theory of inhomogeneous systems it is convenient to consider the free energy F of an inhomogeneous state. This parameter was determined microscopically and investigated in Ref. 11 for binary alloys with an arbitrary deviation from equilibrium. It was shown, in particular, that this free energy has the fundamental property of not increasing as a result of spontaneous processes in the system, in analogy to the nondecreasing Boltzmann entropy. In this section we present the necessary generalizations of the results in Ref. 11 to the case of multicomponent alloys.

Let us consider a substitutional alloy containing atoms of m different kinds $p=p_1,p_2,\ldots,p_m$. Different distributions of the atoms among the lattice sites will be described, as usual, by assigning the set of parameters $\{n_{pi}\}$, where the occupation number operator n_{pi} is equal to unity, if site i is occupied by an atom of kind p, and to zero in other cases. In the substitutional alloy under consideration the n_{pi} for each i are related by the expression $\Sigma_p n_{pi}=1$, so that only $m-1$ of them are independent. It is convenient to label these independent operators by special symbols, for example, Greek letters, and to write $(n_{pi})_{\mathrm{indep}}=n_{\alpha i}$, and the remaining operator, which we shall denote by n_{Bi}, can be expressed in terms of the independent $n_{\alpha i}$:

$$n_{Bi}=1-\sum_{\alpha} n_{\alpha i}. \tag{1}$$

Thus, for an ABv alloy we set $v=\alpha_1$, $\mathrm{A}=\alpha_2$, and $n_{Bi}=1-n_{Ai}-n_{\mathrm{v}i}$.

The energy of the alloy, expressed in terms of all the operators n_{pi}, has the form of a configurational Hamiltonian H of the general form

$$H=\sum_{p}\Phi_i^p n_{pi}+\frac{1}{2!}\sum_{pq,ij}V_{ij}^{pq}n_{pi}n_{qj}$$

$$+\frac{1}{3!}\sum_{pqr,ijk} V_{ijk}^{pqr} n_{pi} n_{qj} n_{rk} + \ldots, \tag{2}$$

where the Φ_i^p are possible external fields (which are present, if not all the lattice sites are equivalent, for example, because of the presence of lattice defects), and the $V_{i\ldots j}^{p\cdots q}$ are the interaction potentials. The effects of fast electron and phonon motions are assumed to be included in the adiabatically averaged values of the configurational potentials $V_{i\ldots j}^{p\cdots q}$.

After elimination of the dependent operators n_{Bi} using relation (1), i.e., in terms of only the independent operators $n_{\alpha i}$, the Hamiltonian (2) takes on the form

$$H = E_0 + \sum_{\alpha i} \varphi_i^\alpha n_{\alpha i} + H_{\text{int}}, \tag{3a}$$

$$H_{\text{int}} = \sum_{\alpha\beta, i>j} v_{ij}^{\alpha\beta} n_{\alpha i} n_{\beta j} + \sum_{\alpha\beta\gamma, i>j>k} v_{ijk}^{\alpha\beta\gamma} n_{\alpha i} n_{\beta j} n_{\gamma k} + \ldots, \tag{3b}$$

where E_0, φ_i^α, and $v_{i\ldots j}^{\alpha\cdots\beta}$ are expressed linearly in terms of the field Φ_i^p and the potentials $V_{i\ldots j}^{p\cdots q}$ in Eq. (2). In particular, for the configurational fields φ_i^α and the pair configuration interactions $v_{ij}^{\alpha\beta}$ we have

$$\varphi_i^\alpha = (\Phi_i^\alpha - \Phi_i^B) + \sum_j (V^{\alpha B} - V^{BB})_{ij} + \sum_{j>k} (V^{\alpha BB} - V^{BBB})_{ijk} + \ldots, \tag{4a}$$

$$v_{ij}^{\alpha\beta} = (V^{\alpha\beta} - V^{\alpha B} - V^{B\beta} + V^{BB})_{ij} + \sum_k (V^{\alpha\beta B} - V^{\alpha BB} - V^{B\beta B} + V^{BBB})_{ijk} + \ldots. \tag{4b}$$

We describe different macroscopic states of the alloy in terms of the distribution function of the atoms among the sites, i.e., a given set of occupation numbers $\{n_{\alpha i}\}$ can be found in terms of the probability $P\{n_{\alpha i}\}$ in this state. Since the $n_{\alpha i}$ are projection operators, i.e., since $n_{\alpha i} n_{\beta i} = \delta_{\alpha\beta} n_{\alpha i}$, the general expression for the distribution function $P\{n_{\alpha i}\}$ can always be written in the form of a ''generalized Gibbs distribution:''[11]

$$P\{n_{\alpha i}\} = \exp\left[\beta\left(\Omega + \sum_{\alpha i} \lambda_i^\alpha n_{\alpha i} - Q\right)\right]. \tag{5}$$

Here $\beta = 1/T$ is the reciprocal temperature, and the correlation term Q is an analog of the interaction Hamiltonian H_{int} in (3b):

$$Q = \sum_{\alpha\beta, i>j} a_{ij}^{\alpha\beta} n_{\alpha i} n_{\beta j} + \sum_{\alpha\beta\gamma, i>j>k} a_{ijk}^{\alpha\beta\gamma} n_{\alpha i} n_{\beta j} n_{\gamma k} + \ldots. \tag{6}$$

In addition, the generalized thermodynamic potential Ω is determined from the normalization condition

$$\Omega = -T \ln \mathrm{Tr} \exp\left[\beta\left(\sum_{\alpha i} \lambda_i^\alpha n_{\alpha i} - Q\right)\right], \tag{7}$$

where $\mathrm{Tr}(\ldots)$ denotes summation over all the configurational states, i.e., over all the sets of occupation numbers $\{n_{\alpha i}\}$.

The mean occupancy $\langle n_{pi} \rangle$ of site i by an atom of kind p (the ''local concentration'') is denoted by c_{pi}, and the average of the operator product $\langle n_{\alpha i} n_{\beta j} \ldots n_{\gamma k} \rangle$ is denoted by $g_{ij\ldots k}^{\alpha\beta\cdots\gamma}$, where the notation $\langle (\ldots) \rangle \equiv \mathrm{Tr}[(\ldots)P]$ denotes averaging over the distribution (5). Then, if the free energy $F = F\{c_{\alpha i}, g_{i\ldots j}^{\alpha\cdots\beta}\}$ is defined by a relation similar to Eq. (23) in Ref. 11,

$$F = \Omega + \left\langle H + \sum_{\alpha i} \lambda_i^\alpha n_{\alpha i} - Q \right\rangle, \tag{8}$$

the derivatives F over its arguments are expressed in terms of φ, λ, v, and a from Eqs. (3)–(6) in the following manner:

$$\partial F / \partial c_{\alpha i} = \varphi_i^\alpha + \lambda_i^\alpha, \tag{9a}$$

$$\partial F / \partial g_{i\ldots j}^{\alpha\cdots\beta} = v_{i\ldots j}^{\alpha\cdots\beta} - a_{i\ldots j}^{\alpha\cdots\beta}. \tag{9b}$$

Following Ref. 11, we can show that the free energy (8) never increases during the spontaneous evolution of the system (the proof is similar to the proof in Ref. 11, being distinguished, for the most part, only by the differences in the notation used for the indices). Therefore, in a stationary, or equilibrium, state the free energy F should have a minimum with respect to its arguments $c_{\alpha i}$ and $g_{i\ldots j}^{\alpha\cdots\beta}$ under the condition of a constant number of particles of each kind $N_\alpha = \Sigma_i c_{\alpha i}$. For the derivative $F_{\alpha i} \equiv \partial F / \partial c_{\alpha i}$ this gives the equation for a conditional minimum

$$F_{\alpha i} = \mu_\alpha = \text{const}, \tag{10}$$

where the Lagrange multiplier μ_α has the meaning of the chemical potential. The conditions for a minimum with respect to $g_{i\ldots j}^{\alpha\cdots\beta}$ do not have restraints, and thus the derivative on the left-hand side of Eq. (9b) should vanish

$$a_{i\ldots j}^{\alpha\cdots\beta} = v_{i\ldots j}^{\alpha\cdots\beta}. \tag{11}$$

Relations (10) and (11) are similar in form to the ordinary equations of Gibbs equilibrium for homogeneous systems. However, in the inhomogeneous alloy under consideration the local concentrations $c_{\alpha i}$ vary with the site label i and must be determined by the system of nonlinear equations (10) and (11) and the boundary conditions.

Real calculations of the free energy can be performed using various approximate methods of statistical physics that are similar to those used for homogeneous systems, such as the mean-field approximation, the cluster-field method[16–18], and the cluster-variation method.[18,19] The simplest and most satisfactory method is the mean-field approximation. The free energy (8) has been written analytically within this approximation,[12] and the derivative $F_{\alpha i}$ in (10) has the form

$$F_{\alpha i} = T \ln \frac{c_{\alpha i}}{c_{Bi}} + \varphi_i^\alpha + \sum_{\beta j} v_{ij}^{\alpha\beta} c_{\beta j}. \tag{12}$$

Here $c_{Bi} = 1 - \Sigma_\alpha c_{\alpha i}$, the equilibrium equations (11) are used, and to simplify the formulas the case in which only pair interactions are present in Hamiltonians (2) and (3) is considered.

Cluster methods are, generally speaking, more exact than the mean-field approximation.[16–20] The cluster-field method (unlike the more complicated cluster-variation method) is also easily generalized to the case of an inhomo-

geneous alloy. In particular, in the simplest version of the cluster-field method, i.e., the pair-cluster method, which coincides with the analogous approximation of the cluster-variation method, the free energy of an inhomogeneous binary alloy $F\{c_{Ai}\}$, as in the mean-field approximation, has been written out analytically and was presented in Refs. 16 and 17. As for a multicomponent alloy, an explicit analytical expression cannot be written within the pair-cluster approximation for $F\{c_{Ai}\}$ in the general case: the algebraic equations obtained are higher than second-order.[16–18] However, for the case considered in the present work of an ABv alloy with a small impurity concentration, $c_{\rm vi} \ll 1$, expressions for F and $F_{\alpha i}$ can also be found analytically in the pair-cluster approximation.

To simplify the notation in the treatment of an ABv alloy below, we shall omit the indices of the principal components A and B and make the replacements

$$c_{Ai} \to c_i, \quad c_{Bi} \simeq 1 - c_i \to c_i',$$
$$v_{ij}^{AA} = V_{ij}^{AA} - 2V_{ij}^{AB} + V_{ij}^{BB} \to v_{ij}. \tag{13}$$

Then, using the methods in Refs. 16–18 for the derivatives $F_{\alpha i}$ in an ABv alloy, for $c_{\rm vi} \ll 1$ we obtain

$$F_{Ai} = \frac{\partial F}{\partial c_i} = \varphi_i^A + T \ln \frac{c_i}{c_i'} + T \sum_{j \neq i} \ln \Big[1 - f_{ij} c_j \times \frac{2}{R_{ij} + 1 + f_{ij}(c_i + c_j)}, \tag{14a}$$

$$F_{\rm vi} = \frac{\partial F}{\partial c_{\rm vi}} = \varphi_i^{\rm v} + T \ln \frac{c_{\rm vi}}{c_i'} - T \sum_{j \neq i} \ln \Big[1 + f_{ij}^{\rm v} c_j \times \frac{2}{R_{ij} + 1 + f_{ij}(c_i - c_j)}. \tag{14b}$$

Here $f_{ij} = f(v_{ij})$ or $f_{ij}^{\rm v} = f(v_{ij}^{A\rm v})$ is the Mayer function for the potential v_{ij} or $v_{ij}^{A\rm v}$:

$$f_{ij} = \exp(-\beta v_{ij}) - 1, \quad f_{ij}^{\rm v} = \exp(-\beta v_{ij}^{A\rm v}) - 1, \tag{15}$$

and the function R_{ij} has the form

$$R_{ij} = [1 + 2f_{ij}(c_i c_j' + c_i' c_j) + f_{ij}^2 (c_i - c_j)^2]^{1/2}. \tag{16}$$

We note that there are no terms with $j = i$ in the sums over j in Eqs. (14). This can be described by setting $v_{ii} = v_{ii}^{A\rm v} = 0$. We also note that the expression (14a) for F_{Ai} coincides with the expression previously derived for an impurity-free AB binary alloy,[17] since the contribution of the impurities to this quantity at small $c_{\rm vi}$ is proportional to $c_{\rm vi}$ and is omitted in the zeroth approximation (14a).

When higher approximations of the cluster-field method, such as the tetrahedral, octahedral, and other similar approximations are used (they must be employed, for example, to obtain a faithful description of the ordering in fcc alloys[17]), the calculation of F and $F_{\alpha i}$ for an inhomogeneous AB or ABv alloy requires the solution of algebraic equations of higher orders, i.e., fourth, sixth, etc., for each lattice site i. However, their solution does not lead to any fundamental difficulties and can be accomplished by numerical methods when specific problems are investigated.

3. GENERAL EXPRESSIONS FOR THE EQUILIBRIUM DISTRIBUTION OF IMPURITIES IN AN ABv ALLOY

The equilibrium distributions of the components in an inhomogeneous alloy can be found from Eq. (10) using any approximation for the free energy $F\{c_{\alpha i}\}$ Below we shall use the mean-field approximation (12) and the pair-cluster approximation (14), under which the distribution of the impurities $c_{\rm vi}$ can be expressed in terms of the distribution of the principal components in a general form.

We first utilize the simpler mean-field approximation. To simplify the formulas, instead of the interaction potential $v_{ij}^{A\rm v}$ from (4b), it is convenient to use the potential P_{ij}, which is equal to the difference between the configurational potentials in the binary alloys Bv and Av:

$$P_{ij} = (V_{ij}^{BB} - 2V_{ij}^{B\rm v} + V_{ij}^{\rm vv}) - (V_{ij}^{AA} - 2V_{ij}^{A\rm v} + V_{ij}^{\rm vv}). \tag{17}$$

When P_{ij} is positive, it is energetically preferable for the impurity to be in a region enriched with atoms of B, and when P_{ij} is negative, it is preferable for it to be in a region enriched with atoms of A. Therefore, we call P_{ij} the ''preference potential'' for the impurity. The potential $v_{ij}^{A\rm v}$ from Eq (4b) is expressed in terms of v_{ij} and P_{ij} as $v_{ij}^{A\rm v} = (v_{ij} + P_{ij})/2$, and the equilibrium equations (10) in the mean-field approximation take the form

$$T \ln \frac{c_i}{c_i'} + \varphi_i^{\rm A} + \sum_j v_{ij} c_j = \mu_{\rm A}, \tag{18a}$$

$$T \ln \frac{c_{\rm vi}}{c_i'} + \varphi_i^{\rm v} + \frac{1}{2} \sum_j (v_{ij} + P_{ij}) c_j = \mu_{\rm v}. \tag{18b}$$

Eliminating the terms with a configuration interaction v_{ij} from Eqs. (18), we obtain an expression for the local impurity concentration $c_{\rm vi}$ in a form that is symmetric relative to the two principal components of the alloy:

$$c_{\rm vi} = A(c_i c_i')^{1/2} \exp\left[-\beta \left(\widetilde{\varphi}_i^{\rm v} + \frac{1}{2} \sum_j P_{ij} c_j \right) \right]. \tag{19}$$

Here the value of the constant $A = \exp(\beta\mu_{\rm v} - \beta\mu_{\rm A}/2)$ is determined from the boundary conditions, for example, from the equilibrium value of $c_{\rm vi}$ far from the phase or antiphase boundary under consideration, and the effective field for the impurities $\widetilde{\varphi}_i^{\rm v}$ is related to the quantities Φ_i^p and V_{ij}^{pq} in (2) by the expression

$$\widetilde{\varphi}_i^{\rm v} = \Phi_i^{\rm v} - \frac{1}{2}(\Phi_i^{\rm A} + \Phi_i^{\rm B}) + \sum_j \left[V_{ij}^{B\rm v} - \frac{1}{2}(V_{ij}^{AB} + V_{ij}^{BB}) \right]. \tag{20}$$

We note that in the inhomogeneous alloy under consideration the potentials V_{ij}^{pq} in (20) can depend not only on the distance between sites $\mathbf{r}_i - \mathbf{r}_j$, but also on the absolute position of a site $\mathbf{r}_i$, so that the last sum in (20) can also depend on $\mathbf{r}_i$.

Let us now discuss the distribution of impurities in the pair-cluster approximation (14). In this case the local concentration $c_{\rm vi}$ depends, generally speaking, not only on c_i, c_i', and P_{ij}, but also on the interactions of the principal components v_{ij}. For convenience in making comparisons

with the result in the mean-field approximation (19), we present the expression for c_{vi} in the pair-cluster approximation in a form analogous to relation (19):

$$c_{vi}=A(c_i c_i')^{1/2}\exp(-\beta\tilde{\varphi}_i^{v}+S_i). \tag{21}$$

Here A and $\tilde{\varphi}_i^{v}$ are the same as in (19), and S_i is given by the expression

$$S_i=\sum_j\left[\frac{1}{2}\ln(1-f_{ij}c_j b_{ij})+\ln(1+f_{ij}^{v}c_j b_{ij}^{v})\right], \tag{22}$$

where b_{ij} and b_{ij}^{v} denote the last cofactor of the second term in the square brackets in Eq. (14a) or (14b), the values of c_i are determined from the equilibrium equation $F_{Ai}=\mu_A=\text{const}$, and F_{Ai} is given by (14a).

The mean-field approximation is applicable, if each of the potentials v_{ij} and v_{ij}^{v} is small in comparison with the temperature:[17] $\beta v_{ij},\beta v_{ij}^{v}\ll 1$. In this case $b_{ij}\simeq b_{ij}^{v}\simeq 1$, $f_{ij}\simeq-\beta v_{ij}$, $f_{ij}^{v}\simeq-\beta(v_{ij}+P_{ij})/2$, and Eq. (21) transforms into (19). If βv_{ij} and βv_{ij}^{v} are not small [for example, when there are strong repulsive interactions with $\beta v_{ij}\gtrsim 1$ (Ref. 16)], the cluster approximations can significantly refine the mean-field approximation.

Relations (18)–(21) can be applied to very different problems in the physics of inhomogeneous alloys, in particular, to the investigation of the aforementioned grain-boundary segregation, where the fields Φ_i^p in Eqs. (2) and (20) include elastic contributions associated with the grain boundaries. We shall apply these relations below to the study of the distribution of impurities near interphase and antiphase boundaries.

4. SEGREGATION ON A PHASE BOUNDARY

Let us first consider the case of a phase boundary between two unordered phases 1 and 2. Far from the boundary, i.e., in a region of homogeneous phase 1 or phase 2, the concentrations c_i have a constant value equal to c_1 or c_2, where c_1 and c_2 are determined from the equations for equilibrium between the phases:

$$\Omega(c_1)=\Omega(c_2),\quad \mu_A(c_1)=\mu_A(c_2). \tag{23}$$

Here $\Omega(c_\nu)=f(c_\nu)-c_\nu\mu_A(c_\nu)$, and $f(c_\nu)$ is the free energy per site of phase ν, where ν is equal to 1 or 2. The contribution of the impurities to Eq. (23) is proportional to their concentration $c_{v\nu}$, i.e., is negligibly small.

In the model with pair interactions under consideration the values of $c_1(T)$ and $c_2(T)$ on the c,T phase diagram are arranged symmetrically relative to the mean value $c=1/2$. Therefore, instead of the local concentration c_i it is convenient to introduce the variable $\xi_i=c_i-1/2$. Then negative, positive, or zero values of the ξ_i correspond to phase 1, phase 2, or the phase boundary, and far from the phase boundary in phase 1 or 2 we have $c_1=1/2-\xi_0$ and $c_2=1/2+\xi_0$, where ξ_0 is determined from Eq. (23).

Let us first utilize the mean-field approximation. In this case the solution of equalities (23) for the phase equilibrium curve (the binodal) can be written in the form

$$\xi_0=\frac{1}{2}\tanh(2\xi_0 T_c/T). \tag{24}$$

Here T_c is the critical temperature, which equals $-v_0/4$ in the mean-field approximation used, where $v_0=\Sigma_j v_{ij}$, and in the homogeneous phases under consideration the potentials v_{ij} and v_0 do not depend on the coordinate of the site $\mathbf{r}_i$.

To simplify the formulas we assume that in the presence of a phase boundary the potentials $v_{ij}^{\alpha\beta}$ depend only on the distances $\mathbf{r}_i-\mathbf{r}_j$ and that the external fields φ_i^{α} are absent. Defining the constant A in (19), for example, in terms of the equilibrium value of the impurities in the first phase $c_v(1)$, we arrive at the following expression for the distribution of the impurities:

$$c_{vi}=c_v(1)\left(\frac{1-4\xi_i^2}{1-4\xi_0^2}\right)^{1/2}\exp\left[-\frac{1}{2}\beta\sum_j P_{ij}(\xi_0+\xi_j)\right]. \tag{25}$$

If the boundary condition in the second phase, rather than the first phase, is used, then $c_v(1)$ in Eq. (25) is replaced by $c_v(2)=c_v(1)\exp(-\beta\xi_0\Sigma_j P_{ij})$ and ξ_0 is replaced by $-\xi_0$.

Let us discuss the form of distribution (25) near a phase boundary. It can be seen, first, that in the absence of the preference potential, i.e., when $P_{ij}=0$ and $c_v(2)=c_v(1)=c_{v0}$ hold, this distribution has a maximum $c_v^{\max}$ at $\xi_i=0$, i.e, on the phase boundary,

$$c_v^{\max}=c_{v0}(1-4\xi_0^2)^{-1/2}, \tag{26}$$

where the value $\xi_0=\xi_0(T)$ is specified by the equation of the binodal (24). At low temperatures ($T\ll T_c$) the difference $\xi_0-1/2$ is exponentially small, and the segregation of impurities on the phase boundary given by Eq. (26) increases exponentially. Actually, however, the equilibrium values of impurity concentrations (25) and (26) may not be achieved at low T because of the slowing of the kinetic processes involving diffusion of the atoms at such T.

For $P=0$ the segregation of impurities occurring on the phase boundary is physically related mainly to entropic, rather than energetic effects. The vanishing of the preference potential (17) implies that the energies corresponding to the attraction of the impurity to both pure components of the alloy are identical and that the expression (19) for c_{vi} contains contributions from only the entropic, but not the energetic, terms in Eqs. (18) when $P=0$. Although the use of Eq. (18a) in the derivation of (19) makes such a separation of the contributions not entirely unequivocal, in any case these remarks point out the importance of the entropic contributions to segregation on a phase boundary. Thus, the arguments advanced in Ref. 4 regarding the ''energetic'' origin of such segregation appear to be at least incomplete.

If the preference potential P is not equal to zero, the distribution of the impurities is asymmetric relative to the phase boundary with different values of $c_v(1)$ and $c_v(2)$ in pure phases 1 and 2, and the maximum of c_{vi} is shifted toward the phase ν which is enriched with impurities. As $|P|$ increases, this shift increases, and the amount by which $c_v^{\max}$ exceeds the bulk value $c_v(\nu)$ decreases. At certain values $|P_{ij}|\sim|v_{ij}|$ this maximum vanishes, and the profile of c_{vi} takes on the form of a worn step (see Figs. 1 and 2).

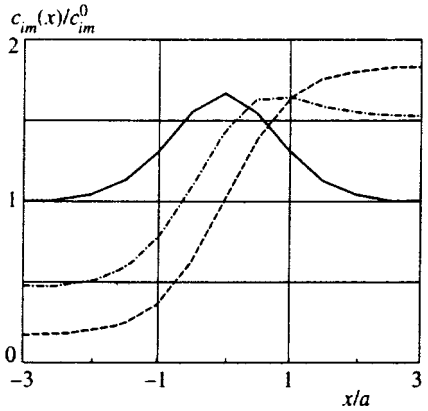


FIG. 1. Impurity concentration profile $c_{\rm im}(x_i)$ near a phase boundary lying in a [1,0,0] plane of a bcc alloy with a lattice constant a calculated in the pair-cluster approximation for $T=0.75T_c$. For clarity, the values of $c_{\rm im}(x_i)$ at the discrete points $x_i=na$ and $x_i=(n\pm 1/2)a$ are joined by straight lines. The configurational potential $v_{ij}<0$ corresponds to the interaction of only the nearest neighbors, and the solid, dot-dashed, and dashed lines correspond to a preference potential P_{ij} equal to 0, $v_{ij}/2$, and v_{ij}.

Let us now discuss the distribution of c_{vi} in the pair-cluster approximation (21). Here, instead of (25) we obtain

$$c_{vi}=c_v(1)\left(\frac{1-4\xi_i^2}{1-4\xi_0^2}\right)^{1/2}\exp[S_i-S(c_1)], \tag{27}$$

where S_i is the same as in (21), and $S(c_1)$ corresponds to the replacement of all the c_i and c_j in (22) by c_1. The values of c_1 and c_2 are specified by the equation of the binodal (23) with the following expressions[16] for $\Omega(c_\nu)$ and $\mu_A(c_\nu)$:

$$\Omega(c_\nu)=T\ln c'_\nu-T\sum_j \ln[1-c_\nu^2 f_{ij}b_{ij}(c_\nu)], \tag{28a}$$

$$\mu_A(c_\nu)=T\ln\frac{c_\nu}{c'_\nu}+T\sum_j \ln[1-c_\nu f_{ij}b_{ij}(c_\nu)], \tag{28b}$$

where the $b_{ij}(c_\nu)$ are the same as in Eq. (22) when $c_i=c_j=c_\nu$.

Numerical calculations based on Eqs. (25)–(28) show that for ordinary models of interactions at not excessively low temperatures ($T\gtrsim T_c/2$) the distributions (25) in the mean-field approximation and (27) in the pair-cluster approximation are similar. The degree of segregation is somewhat greater in the pair-cluster approximation; however, the differences in the concentration profile are usually less than 10%.

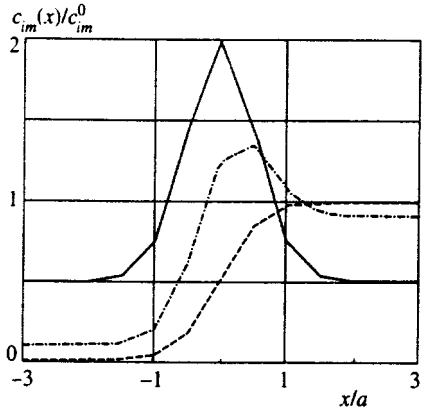


FIG. 2. Same as in Fig. 1, but for $T=0.5T_c$.

Figures 1 and 2 present some results of the calculations of the impurity concentration profile $c_{\rm im}(x_i)=c_{vi}$ near phase boundaries. The model of a bcc alloy with the nearest-neighbor interaction v_{ij} is considered for several values of the preference potential P_{ij}. The phase boundary is assumed to be flat and to lie in a [1,0,0] plane with $x=x_i=0$. The normalizing constant $c_{\rm im}^0$ is chosen equal to the average of the bulk concentrations of the impurity in phases 1 and 2: $c_{\rm im}^0=[c_v(1)+c_v(2)]/2$. The distribution of the c_i of the principal components of the alloy (i.e., the structure of the phase boundary) is found by minimizing the free energy $F\{c_i\}$ with respect to c_i under suitable boundary conditions, which is technically simpler than solving the equation $F_{Ai}=$ const like (18a). After this, $c_{\rm im}(x_i)=c_{vi}$ is calculated from Eq. (25) or (27).

The calculation results presented in Figs. 1 and 2 were obtained using the pair-cluster approximation. These results illustrate the general features discussed here of the segregation of impurities on phase boundaries. It is seen for large $|P_{ij}|\gtrsim|v_{ij}|$, in particular, that the maximum for $c_{\rm im}$ near the phase boundaries vanishes and the profile takes on the form of a worn step. We also note that the degree of segregation in our calculations (both in the pair-cluster approximation and in the mean-field approximation) is similar in order of magnitude to the degree which was visually observed in the Monte Carlo calculations in Ref. 2. At the same time, values of the degree of segregation almost an order of magnitude greater were obtained in Ref. 7. This is another indication of the unreliability of the equations from Ref. 10 that were used in Ref. 7.

5. SEGREGATION ON AN ANTIPHASE BOUNDARY

Let us consider segregation on antiphase boundaries in the case of $B2$ [or CuZn (Ref. 9)] ordering in a bcc structure. In the ordered phase the original bcc lattice decomposes into two cubic sublattices, which are displaced by the translation vector $(a/2,a/2,a/2)$. In one of these sublattices the concentration of the atoms of A is greater than and in the other less than the mean value $c=c_0$ throughout the alloy. To describe inhomogeneous states of such an alloy, particularly, antiphase boundaries, it is convenient to introduce a local mean concentration $\tilde{c}_i$ and an order parameter η_i for each site i and to determine them, for example, using averaging over the nearest neighbors:[13]

$$\tilde{c}_i=\frac{1}{2}\left(c_i+\frac{1}{8}\sum_{j=nn(i)}c_j\right),$$

$$\eta_i=\frac{1}{2}\left(c_i-\frac{1}{8}\sum_{j=nn(i)}c_j\right)\exp(i\mathbf{k}_s\mathbf{r}_i). \tag{29}$$

Here the index $nn(i)$ in the sum denotes summation over all the nearest neighbors of site i, $\mathbf{r}_i$ is the coordinate of that site, and $\mathbf{k}_s=(1,t,1)2\pi/a$ is the superstructure vector. The multi-

plier $\exp(i\mathbf{k}_s\mathbf{r}_i)$ in (29) is equal to 1 or -1 for the sites i lying in the first or second sublattice. The $\tilde{c}_i$ and η_i (unlike the occupancies c_i of the sites, which are equal to $\tilde{c}_i \pm \eta_i$) are smooth functions of the coordinates, and the values in neighboring sites are close. In a homogeneous ordered phase the values of $\tilde{c}_i = c_0$ and $\eta_i = \eta_0$ are constant throughout the alloy, and the order parameter η_0 is determined from the condition of equality between the expressions (10) for the two sublattices. For example, in the mean-field approximation, from Eqs. (18a) we have

$$T \ln \frac{(c_0 - \eta_0)(c_0' - \eta_0)}{(c_0 + \eta_0)(c_0' + \eta_0)} = 2v_0^s \eta_0, \tag{30}$$

where $c_0' = 1 - c_0$ and $v_0^s = \Sigma_j v_{ij} \exp[i\mathbf{k}_s(\mathbf{r}_i - \mathbf{r}_j)]$. The critical ordering temperature T_{c0} equals $(-v_0^s)/4$.

An antiphase boundary is defined as the set of the points $\mathbf{r}_A$ at which the values of the η_i or the values of $\eta(\mathbf{r}_A)$ extrapolated from η_i in the sites i closest to $\mathbf{r}_A$ vanish.[13] For simplicity, below we shall consider the case of a flat antiphase boundary lying in a [1,0,0] plane of the lattice with $x_A = x_i = 0$.

It is convenient to write the distributions of the impurities for the first and second sublattices, c_{vi}^+ and c_{vi}^-, separately. Then the equations for the mean-field approximation (19) for $\tilde{\varphi}_i^v = 0$ take the form

$$c_{vi}^+ = A[(\tilde{c}_i + \eta_i)(\tilde{c}_i' - \eta_i)]^{1/2} \exp\left[-\frac{1}{2}\beta \left(\sum_j P_{ij} \tilde{c}_j + \sum_j P_{ij}^s \eta_j \right) \right], \tag{31a}$$

$$c_{vi}^- = A[(\tilde{c}_i - \eta_i)(\tilde{c}_i' + \eta_i)]^{1/2} \exp\left[-\frac{1}{2}\beta \left(\sum_j P_{ij} \tilde{c}_j - \sum_j P_{ij}^s \eta_j \right) \right], \tag{31b}$$

where $P_{ij}^s = P_{ij} \exp[i\mathbf{k}_s \cdot (\mathbf{r}_i - \mathbf{r}_j)]$. Equations (31) show that the functions c_{vi}^+ and c_{vi}^- vary smoothly, if the values of $\tilde{c}_i$ and η_i vary smoothly in space. However, generally speaking, these functions and, therefore, the values of c_{vi} in neighboring sites differ appreciably.

Let us first discuss the simpler case of an alloy with the stoichiometric composition $c_0 = c_0' = 1/2$. Then the local concentrations $\tilde{c}_i$ do not vary near the antiphase boundary under consideration:[13] $\tilde{c}_i = \text{const} = 1/2$. Let the asymptotic values of η_i to the left and the right of the antiphase boundary (in regions 1 and 2) be equal, respectively, to η_0 and $-\eta_0$. Then distribution (31) takes a form similar to the expression (25) for a phase boundary with the replacement of ξ_i by η_i and P_{ij} by P_{ij}^s:

$$c_{vi}^{\pm} = \tilde{c}_{v0} \left(\frac{1 - 4\eta_i^2}{1 - 4\eta_0^2} \right)^{1/2} \frac{1}{\cosh(\beta P_0^s \eta_0/2)} \times \exp\left(\mp \frac{1}{2}\beta \sum_j P_{ij}^s \eta_j \right), \tag{32}$$

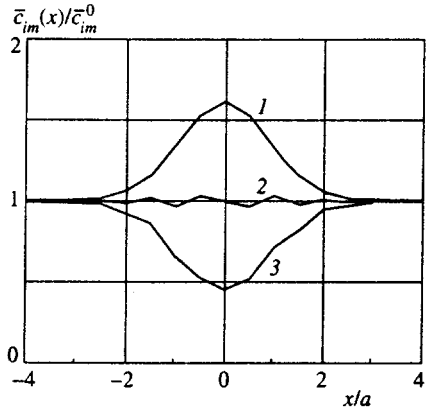


FIG. 3. Locally averaged impurity concentration profile $\tilde{c}_{im}(x_i)$ near an antiphase boundary lying in a [1,0,0] plane with $B2$ ordering calculated in the pair-cluster approximation for a model with a nearest-neighbor interaction $v_{ij} > 0$ at $T = 0.75T_{c0}$ and a mean alloy concentration $c_0 = 1/2$. Curves *1*, *2*, and *3* correspond to a preference potential P_{ij} equal to 0, v_{ij}, and $2v_{ij}$.

where $\tilde{c}_{v0}$ is the mean impurity concentration far from the antiphase boundary, which is identical in both domains, and $P_0^s = \Sigma_j P_{ij}^s$.

Equation (32) shows that at large values of $|P_0^s|$ segregation of the impurities on the antiphase boundary (i.e., in the region $\eta_j \simeq 0$) is suppressed. Taking into account that for the value of c_0 equal to 1/2 considered here the factor $(1 - 4\eta_0^2)^{-1/2}$ is equal to $\cosh(\beta v_0^s \eta_0/2)$ according to Eq. (30), we can approximately estimate the degree of segregation on an antiphase boundary as

$$\frac{c_v(\eta_i = 0)}{\tilde{c}_{v0}} \simeq \frac{\cosh(\beta v_0^s \eta_0/2)}{\cosh(\beta P_0^s \eta_0/2)}. \tag{33}$$

Therefore, when $|P_0^s| \lesssim |v_0^s|$ holds, the antiphase boundary is enriched, and when $|P_0^s| \gtrsim |v_0^s|$ holds, it is depleted of impurities in comparison to the interior of the antiphase domains. All these conclusions remain valid when the pair-cluster approximation (21) is used instead of the mean-field approximation.

Figure. 3 presents some results of calculations of the locally averaged impurity concentration profiles $\tilde{c}_{im}(x)/\tilde{c}_{im}^0 = \tilde{c}_{vi}/\tilde{c}_{v0}$ near an antiphase boundary for a stoichiometric alloy with $c_0 = 1/2$ within the pair-cluster approximation. The values of $\tilde{c}_{vi}$ are determined using averaging over the nearest neighbors similar to that used in Eq. (29):

$$\tilde{c}_{vi} = \frac{1}{2}\left(c_{vi} + \frac{1}{8} \sum_{j = nn(i)} c_{vj} \right). \tag{34}$$

Figure 3 illustrates the general conclusions discussed regarding the segregation of impurities on an antiphase boundary when $c_0 = 1/2$.

For an alloy of nonstoichiometric composition with $c_0 \neq 1/2$ the situation is more complicated. As was discussed in Ref. 13, in this case the values of $\tilde{c}_i$ near antiphase boundaries are no longer constant, and for $\tilde{c}_0 < 1/2$ they have a minimum: $\tilde{c}_i < c_0$. Therefore, the coordinate dependence of

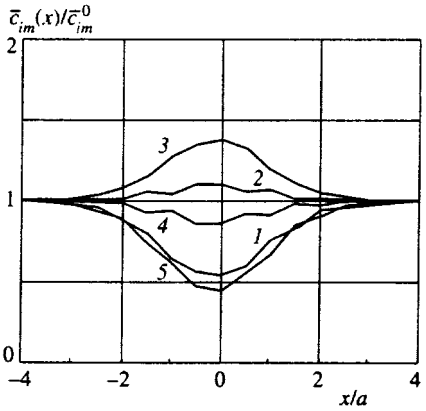


FIG. 4. Same as Fig. 3, but for a model with an interaction v_{ij} in two coordination spheres: $v_1=1$, $v_2=-0.8$ at $T=0.75T_{c0}$ and $c_0=0.425$. Curves *1*, *2*, *3*, *4*, and *5* correspond to $P_{ij}=gv_{ij}$ with g equal to -2, -1, 0, 1, and 2, so that the values of P_0 and P_0^s for these curves are equal, respectively, to (-6.4) and 25.6, (-3.2) and 12.8, 0 and 0, 3.2 and (-12.8), and 6.4 and (-25.6).

the terms with $\tilde{c}_i$ in the exponential functions must also be taken into account in Eq. (31). The influence of these terms can be characterized qualitatively by $P_0=\Sigma_j P_{ij}$. Positive values of P_0 promote an increase in c_{vi} in a region with small $\tilde{c}_i$, i.e, near antiphase boundaries. Conversely, negative values of P_0 decrease this segregation. The potentials P_{ij}^s, which are characterized by P_0^s in Eq. (32), as was noted above, suppress segregation on antiphase boundaries. To illustrate the possible situations, in Figs. 4 and 5 we present the results of calculations of the impurity profiles $\tilde{c}_{\rm im}(x)$ for models with interactions in two coordination spheres. Figure 4 corresponds to models with $|P_0^s|>|P_0|$, and Fig. 5 is for models with $|P_0^s|<|P_0|$. In accordance with the arguments advanced here, an increase in the preference potential leads to the suppression of segregation in the former case and, conversely, to its enhancement in the latter case.

The Monte Carlo calculations of the segregation of vacancies on an antiphase boundary for $B2$ ordering models that are known to us[5,6] correspond to $P_{ij}=0$. The conclusions drawn regarding the occurrence of segregation on an antiphase boundary are consistent with our results. For example, according to Eq. (33), when $P_0^s=0$, the impurity concentration is always greater on an antiphase boundary than in the interior of the domains in a stoichiometric alloy, and curve *3* in Fig. 4 and curve *2* in Fig. 5 show that this is also true for nonstoichiometric alloys. However, the lack of quantitative evaluations of this segregation in Ref. 5 and 6 make a more detailed comparison difficult.

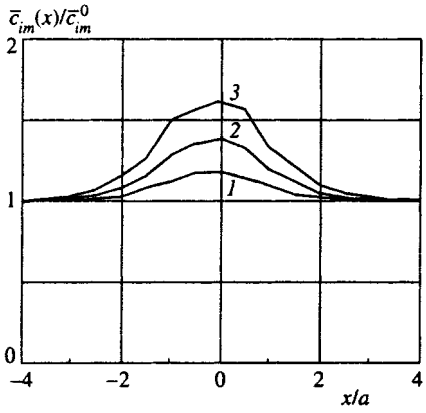


FIG. 5. Same as Fig. 4, but for the preference potentials $P_{ij}=g|v_{ij}|$. Curves *1*, *2*, and *3* correspond to g equal to -2, 0, and 2, so that P_0 and P_0^s for these curves are equal, respectively, to (-25.6) and 6.4, 0 and 0, and 25.6 and (-6.4).

In conclusion, we note that, apart from the calculations of the distribution of impurities described above for ''site-intersecting'' phase or antiphase boundaries, whose centers lie in [1,0,0] atomic planes in a bcc lattice, we also performed similar calculations for ''intersite'' phase or antiphase boundaries, whose centers are halfway between these atomic planes [so that the central plane of such a boundary is obtained from the analogous plane of a site-intersecting boundary as a result of translation by the vector $(a/4,0,0)$]. These intersite boundaries are also stable, and the surface energies of site-intersecting and intersite phase or antiphase boundaries are very close to one another (see, for example, Ref. 21), so that both types of boundaries form with the same frequency in a real alloy. The calculated plots of $c_{\rm im}(x)$ and $\tilde{c}_{\rm im}(x)$ for intersite phase and antiphase boundaries are very similar to the analogous profiles for site-intersecting boundaries presented in Figs. 1–5. The values of $c_{\rm im}(x_i)$ and $\tilde{c}_{\rm im}(x_i)$ for an intersite phase or antiphase boundary can be estimated to good accuracy ($\sim$10%) as the values on the broken curves in Figs. 1–5 for the corresponding points $x_i=(n\pm 1/4)a$ or $x_i=(n\pm 3/4)a$. This similarity between the results for boundaries of different types suggests that the results in Figs. 1–5 are actually fairly general and can be used to estimate the distribution of impurities near phase and antiphase boundaries in different positions and orientations.

6. CONCLUSIONS

In conclusion, let us enumerate the main results of this work. A general expression has been proposed for the free energy of an inhomogeneous multicomponent alloy. It can be used to investigate various problems in the physics of such alloys. In the case of an AB_v alloy with a small concentration of vacancies (v), explicit expressions for the equilibrium distribution of the impurities among the lattice sites c_{vi} in terms of the distribution of the principal components of the alloy have been given in the mean-field and pair-cluster approximations. An investigation of the impurity concentration profiles $c_{\rm im}(x)=c_{\rm vi}$ near a phase boundary has shown that at small values of the preference potential P defined by Eq. (17) the segregation of impurities on the boundary is described by Eq. (26) and that as P increases, the $c_{\rm im}(x)$ profile takes on the form of a worn step. Segregation also occurs near an antiphase boundary between domains with $B2$ ordering when P is small. As P increases, the degree of this segregation in an alloy of stoichiometric composition decreases according to Eq. (33), and antisegregation takes place at large values of P. In nonstoichiometric alloys the influence of the preference potential on the character of the segregation on an antiphase boundary also depends on the form

of this potential, particularly on the ratio between P_0^s and P_0, which were introduced in Sec. 5. Depending on the ratio between P_0^s and P_0, an increase in P can either decrease or increase the degree of segregation on an antiphase boundary.

We thank V. Yu. Dobretsov for his great assistance in this work, as well as D. E. Temkin for his interest in our efforts and some useful remarks. This work was supported by the Russian Fund for Fundamental Research (Project 95-02-06246).

[1] H. Gleiter and B. Chalmers, *High-Angle Grain Boundaries*, Pergamon, Oxford (1972) [Russ. transl., Mir, Moscow (1975, Chap. 3)].
[2] K. Yaldram and K. Binder, Acta Metall. Mater. **39**, 707 (1991); J. Stat. Phys. **62**, 161 (1991).
[3] K. Yaldram and K. Binder, Z. Phys. B **82**, 405 (1991).
[4] P. Fratzl and O. Penrose, Phys. Rev. B **50**, 3477 (1994).
[5] E. Vives and A. Planes, Phys. Rev. B **47**, 2557 (1993).
[6] C. Frontera, E. Vives, and A. Planes, Z. Phys. B **96**, 79 (1994).
[7] C. Geng and L. Q. Chen, Scr. Metall. Mater. **31**, 1507 (1994).
[8] L. Q. Chen, Mater. Res. Soc. Symp. Proc. **319**, 375 (1994).
[9] L. D. Landau and E. M. Lifshitz, *Statistical Physics, Vol. 1*, 3rd. ed., Pergamon Press, Oxford–New York (1980).
[10] L. Q. Chen, Phys. Rev. B **49**, 3791 (1994).
[11] V. G. Vaks, JETP Lett. **63**, 461 (1996).
[12] V. G. Vaks, S. V. Beĭden, and V. Yu. Dobretsov, JETP Lett. **61**, 68 (1995).
[13] V. Yu. Dobretsov, G. Martin, F. Soisson, and V. G. Vaks, Europhys. Lett. **31**, 417 (1995).
[14] V. Yu. Dobretsov, V. G. Vaks, and G. Martin, Phys. Rev. B **54**, 3227 (1996).
[15] K. D. Belashchenko and V. G. Vaks, Phys. Lett. A **222**, 345 (1996).
[16] V. G. Vaks and V. G. Orlov, Fiz. Tverd. Tela (Leningrad) **28**, 3627 (1986) [Sov. Phys. Solid State **28**, 2045 (1986)].
[17] V. G. Vaks, N. E. Zein, and V. V. Kamyshenko, J. Phys. F: Met. Phys. **18**, 1641 (1988).
[18] V. G. Vaks and V. V. Kamyshenko, Izv. Akad. Nauk. SSSR, Met. (**2**), 121 (1990).
[19] J. M. Sanchez, F. Ducastelle, and D. Gratias, Physica A **128**, 334 (1984).
[20] V. G. Vaks, *Introduction to the Microscopic Theory of Ferroelectrics* [in Russian], Nauka, Moscow (1973).
[21] R. Kikuchi and J. W. Cahn, J. Phys. Chem. Solids **20**, 137 (1962); **27**, 1305 (1966).

Translated by P. Shelnitz

Spin liquid in an almost ferromagnetic Kondo lattice

K. A. Kikoin, M. N. Kiselev, and A. S. Mishchenko

Kurchatov Institute, 123182 Moscow, Russia
(Submitted 2 January 1997)
Zh. Éksp. Teor. Fiz. **112**, 729–759 (August 1997)

A theory of stabilization of a spin liquid in a Kondo lattice at temperatures close to the temperature of antiferromagnetic instability has been developed. Kondo exchange scattering of conduction electrons leads to emergence of a state of the spin liquid of the resonating valence bonds (RVB) type at $T > T_K$. Owing to this stabilization, low-energy processes of Kondo scattering with energies below T_K are frozen so that the ''singlet'' state of the Kondo lattice is not realized; instead a strongly correlated spin liquid with developed antiferromagnetic fluctuations occurs. A new version of the Feynman diagram technique has been developed to describe interaction between spin fluctuations and resonant valence bonds in a self-consistent manner. Emergence of a strongly anisotropic RVB spin liquid is discussed. [S1063-7761(97)02508-0]

1. INTRODUCTION

One of the most extraordinary properties of heavy-fermion compounds is the transition of a system of weakly interacting spins, which manifests paramagnetic properties at high temperatures, to a strongly correlated quantum liquid with thermodynamic and magnetic properties typical of Fermi systems at $T < T_{\rm coh} \ll T^*$. This ''dissolution'' of localized spins is usually interpreted in terms of the Kondo lattice model, and the basic mechanism which determines thermal transformation of the spin subsystem is assumed to be Kondo screening of spins by conduction electrons. This screening can be modeled in essentially the same way as a one-impurity lattice, so that the Kondo lattice can be treated as a periodic structure of Kondo impurities coherently scattering conduction electrons.[1,2] The characteristic temperature T^* at which the system switches to another regime is the Kondo temperature T_K, and the ground state in the mean-field approximation is the so-called Kondo singlet.

This simple model, however, ignores spin correlations, whose close relation to heavy fermions is beyond doubt. It is well known that formation of a heavy fermion is in all cases without exception due either to long-range antiferromagnetic order or short-range magnetic correlations. In its interpretation of this relation, the Kondo lattice theory invokes indirect exchange between localized spins via conduction electrons (RKKY exchange), which occurs in the Kondo lattice model in the second order of perturbation theory. Thus, nonlocal spin correlations compete with local effects of spin screening.[3] This naive dichotomy of Doniach's, which takes place in the mean-field approximation, predicts a tendency to antiferromagnetic ordering at small values of the effective coupling constant

$$\alpha = J_{sf}\mathcal{N}(\varepsilon_F)\Omega_0,$$

where J_{sf} is the sf-exchange integral, $\mathcal{N}(\varepsilon_F)$ is the electron density of states at the Fermi surface, Ω_0 is the elementary cell volume. At large α, Kondo screening suppresses the magnetic moment, and the ground state is the Kondo singlet. Then, for α slightly higher than α_c determined by the equation $\alpha_c^2 \exp(1/2\alpha_c) \approx 1$, the RKKY interaction can be calculated by perturbation theory, and magnetic correlations modify properties of the singlet phase.[4,5]

An alternative approach to the problem of competition between the one-site screening and magnetic intersite correlations was suggested by Coleman and Andrei.[6] The two options described by Doniach's simple model were supplemented with a third one, namely, formation of a nonmagnetic spin liquid of the resonant valence bonds (RVB) type with the Fermi statistics of elementary excitation in the spin system (spinons). They demonstrated that the spin liquid state is stabilized by Kondo scattering, but calculated both spin intersite correlations and the single-site sf-exchange between spinons and electrons in the mean-field approximation. Introduction of anomalous one-site averages of the Kondo type is, in reality, equivalent to the assumption that full dynamic spin screening takes place, and the assumption that Kondo singlets are formed at each site owing to multiple ''switching'' of RVB bonds between localized spins and conduction electron spins is equivalent to a translation of electron charge to spin degrees of freedom. Thus, in this scenario, as well as in the mean-field theory of the Kondo lattice,[7] spin degrees of freedom, which manifest themselves at high temperatures as weak paramagnetic, noncharged correlations, have a charge at $T < T_K$ and transform to charged heavy fermions (a critical discussion of this scenario was given in Ref. 8). Naturally, interpretation of the existence of magnetic correlations in the Kondo lattice requires more theoretical effort.[9]

This paper suggests an alternative scenario of formation of a spin liquid in the Kondo lattice described by the Hamiltonian

$$H_{\rm eff} = \sum_{\mathbf{k}\sigma} \varepsilon_k c^+_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma} + J_{sf} \sum_{\mathbf{i}} \left(\mathbf{s_i}\cdot\mathbf{S_i} + \frac{1}{4} \right). \tag{1}$$

Here ε_k is the dispersion relation for conduction electrons, $\mathbf{S_i}$ and $\mathbf{s_i} = (1/2) c^+_{\mathbf{i}\sigma} \hat{\sigma} c_{\mathbf{i}\sigma}$, are operators of a spin localized in the f-shell and of a delocalized conduction electron spin, respectively, and $\hat{\sigma}$ are Pauli matrices. Our approach is based on an understanding that in the critical region, where all

three characteristic temperatures (the Kondo temperature $T_K \sim \varepsilon_F \exp(-1/2\alpha)$, Néel temperature $T_N \sim \varepsilon_F \alpha^2$, and the temperature T^* of the transition to the spin-liquid state) are of the same order, and the Kondo scattering is favorable to the transition to the spin-liquid state so that $T^* > T_N > T_K$, the very existence of spin-liquid correlations impedes the formation of a singlet ground state, since screening of localized spins by the Kondo scattering is, in fact, frozen at temperatures $T \sim T^* > T_K$, and at lower temperatures the system properties are controlled by nonlocal spin-liquid correlations, rather than one-site Kondo scattering. In other words, spin correlations suppress the Kondo effect in both the ordered (magnetic) and disordered (spin-liquid) phases, so Doniach's simple phase diagram should be revised.

Since the spin-liquid state emerges in the critical region $\alpha \sim \alpha_c$ at temperatures close to T_N, the coexistence of heavy fermions and magnetic correlations has a natural interpretation in the proposed model. Moreover, it is obvious that critical spin fluctuations should play an important role in the mechanism of spin-liquid formation. In this study, we have limited our calculations to the range of high temperatures $T > T_K$, where the perturbation theory in the parameter $\alpha \ln(\varepsilon_F/T_K)$ applies. We use the diagram technique for spin operators in the pseudofermion representation[10] in the approximation of noncrossing graphs (or noncrossing approximation, NCA) for the description of the Kondo scattering. The results of high-temperature expansions, which take one-site and intersite correlations into account concurrently, will be extrapolated to the range of temperatures where paramagnetic fluctuations are important. However, when the pseudofermion technique is applied to nonlocal spin-liquid correlations, the problem of nonphysical states arises, and hence the breaking of local spin symmetry.[11–15] With this circumstance in view, we have constructed a Feynman diagram technique for spin Hamiltonians, which allows us, in principle, not only to get rid of nonphysical states, but also to take into account fluctuations of calibration fields.

In Sec. 2 this technique is applied to a spin liquid of the homogeneous RVB phase type[16,17] described in terms of the Heisenberg model; in Sec. 3 the technique is applied to the Kondo lattice, and the mechanism of RVB phase stabilization by Kondo scattering in the mean-field approximation is described.[1)] The mean-field theory for the RVB phase, taking into account critical fluctuations, is generalized in Sec. 4, and Sec. 5 shows how this diagram technique can be used in describing local critical and hydrodynamic fluctuations around the antiferromagnetic instability point.

2. PROJECTION DIAGRAM TECHNIQUE FOR THE HEISENBERG LATTICE

Along with standard perturbation theory techniques developed for Fermi and Bose operators, one can find in the literature a number of diagram techniques for noncommuting operators in terms of which one can write the Hamiltonians of the spin or strongly correlated electrons systems (see, for example, Refs. 19–21 and references therein). Most of these techniques are based on Hubbard's projection operators $X_{\mathbf{j}}^{\lambda\mu} = |\,\mathbf{j}\lambda\rangle\langle\mathbf{j}\mu|$, where $|\mathbf{j}\lambda\rangle$ is a ket vector corresponding to state λ at site j, and the Hamiltonian in the zeroth order of perturbation theory is diagonalized in this approximation.

Diagram techniques for noncommuting operators are harder to handle than the standard Feynman technique. Only in some cases do they allow a self-consistent form of closed equation systems for skeleton diagrams. Goden's procedure factorizing the average of n operators, unlike Wick's procedure, which plays a similar role in the usual diagram technique, is ambiguous, and a successful choice of the hierarchy of couplings largely depends on the theorist's intuition (see, for example, Ref. 22).

For this reason, it is natural to attempt to express Hubbard's operators (and spin operators, which are a special case of these operators) as products of Fermi and Bose operators, and thus restore conditions for using the machinery of the Feynman and Matsubara techniques. Such attempts have been undertaken many times since the 1960s,[10,23–25] up through recent times.[22,26–28] It is clear, however, that these procedures are by no means universal or unambiguous. Moreover, each factorization leads to multiplication and complication of vertices and emergence of local constraints, whose introduction is necessary for the preservation of local gauge invariance, which is a trait of the Hamiltonian in question.

Additional problems arise due to attempts to describe nonlocal spin-liquid RVB excitation. In this case, problems arise on the level of the mean-field approximation for the self-energy part of the one-particle propagator. The usual techniques of self-consistent perturbation theory break the local gauge invariance,[11] and its restoration is quite a difficult physical and mathematical problem.[13,14]

In this section we formulate a version of the diagram technique integrating the two approaches mentioned above, and apply it to Hamiltonians with local SU(2) symmetry. With a view toward using this technique in the description of spin liquid in terms of Hamiltonian (1), we start with the simpler case of the Heisenberg Hamiltonian for spin 1/2 with antiferromagnetic interaction:

$$H = h\sum_{\mathbf{i}} S_{\mathbf{i}}^z + J\sum_{\mathbf{i}} \sum_{\mathbf{j}}^{\langle nn\rangle} \left(\mathbf{S_i}\cdot\mathbf{S_j} - \frac{1}{4}\right); \tag{2}$$

we then pass to a description of the Kondo lattice at high temperatures $T > T_K$, for which the noncrossing approximation (NCA) applies, and the system can be treated as a periodic lattice of independent Kondo scatterers interacting via the RKKY mechanism.[3]

Let us introduce a pseudofermion representation[10] for spin operators:

$$S^+ = f_\uparrow^+ f_\downarrow, \quad S^- = f_\downarrow^+ f_\uparrow, \quad S^z = \frac{1}{2}(f_\uparrow^+ f_\uparrow - f_\downarrow^+ f_\downarrow). \tag{3}$$

These operators satisfy the local constraint condition

$$n = f_\uparrow^+ f_\uparrow + f_\downarrow^+ f_\downarrow = 1 \tag{4}$$

at each site. The first term in Eq. (2) describes Zeeman splitting in an infinitesimal magnetic field $h = g\mu_B H$, and the antiferromagnetic sign of the exchange constant J is taken into account explicitly.

The SU(2) invariance means that the spin operators $\{S^+,S^-,S^z\}$ can be expressed as arbitrary linear combinations of spin-fermions $\{f_\uparrow,f_\downarrow,f_\uparrow^+,f_\downarrow^+\}$:

$$S_{\mathbf{i}}^+=(\cos\,\theta f_{\mathbf{i}\uparrow}^+ +\sin\,\theta f_{\mathbf{i}\downarrow}\,)(\cos\,\theta f_{\mathbf{i}\downarrow}-\sin\,\theta f_{\mathbf{i}\uparrow}^+\,),$$

$$S_{\mathbf{i}}^-=(\cos\,\theta f_{\mathbf{i}\downarrow}^+ -\sin\,\theta f_{\mathbf{i}\uparrow})(\cos\,\theta f_{\mathbf{i}\uparrow}+\sin\,\theta f_{\mathbf{i}\downarrow}^+), \tag{5}$$

$$S^z=\frac{1}{2}(f_{\mathbf{i}\uparrow}^+ f_{\mathbf{i}\uparrow}-f_{\mathbf{i}\downarrow}^+ f_{\mathbf{i}\downarrow}).$$

In particular, for pseudofermion filling factors, we have complete particle–hole symmetry,

$$f_{\mathbf{i}\sigma}^+ f_{\mathbf{i}\sigma}=f_{\mathbf{i}-\sigma}f_{\mathbf{i}-\sigma}^+, \tag{6}$$

which directly follows from condition (4) or from Eqs. (5) for $\theta = 0,\pi/2$. Thus, Hamiltonian (2) can be expressed in the pseudofermion representation as

$$H=H_0+H_{\rm int}=-\frac{h}{2}\sum_{\mathbf{ij}\sigma}\,\sigma f_{\mathbf{i}\sigma}^+ f_{\mathbf{i}\sigma}+\frac{J}{2}\sum_{\mathbf{ij}\sigma}\,f_{\mathbf{i}\sigma_1}^+ f_{\mathbf{j}\sigma_1}f_{\mathbf{j}\sigma_2}^+ f_{\mathbf{i}\sigma_2}. \tag{7}$$

The local constraint places significant limits on the feasibility of using standard diagrammatic techniques, or in any case, makes more difficult practical description of the spin dynamics in the fermion representation, since the functional space in which the spin and fermion operators act has finite dimensionality. One of the most convenient techniques for including the spin kinetamics in the fermion representation was suggested by Abrikosov.[10] $2S+1$ spins (projections) correspond to a localized spin $\mathbf{S_i}$, whereas in its description in terms of pseudofermion operators $2^{(2S+1)}$ orthogonal states emerge, in accordance with the filling numbers (0,1) for all $2S+1$ spin projections. In a specific case of spin $S = 1/2$, there are four fermion states:

$$|0\rangle=|0,0\rangle;\quad |+\rangle=|1,0\rangle;\quad |-\rangle=|0,1\rangle;\quad |2\rangle=|1,1\rangle \tag{8}$$

and only two of them, namely states $|\pm\rangle$, correspond to physical states of the spin operator. Abrikosov suggested ascribing energy $\lambda \gg T$ to each state occupied by a pseudofermion. Then the nonphysical state $|2\rangle$ from set (8) is frozen out after averaging owing to the additional factor $\exp(-\lambda/T)$ in the partition function $\mathcal{Z}(T)$. In order to get rid of the other nonphysical state $|0\rangle$, one must introduce an additional factor $(1/2)\exp(\lambda/T)$ to $\mathcal{Z}(T)$ and take the limit $\lambda/T\to\infty$ in averaging over spin states. As a result, physical states $|\pm\rangle$ become states with the lowest ''energy'', and the final result is independent of λ. Abrikosov's prescription applies to local spin states in the case of a one-impurity Hamiltonian of sf-exchange. In a Kondo lattice it can be used only in the limit of large spin with degeneracy $N\to\infty$, for which NCA becomes an asymptotically exact approximation,[29] but this technique cannot be used in describing spin-liquid correlations.

The starting point of the proposed method is the well known similarity between the Heisenberg and Hubbard Hamiltonians in the limit of strong interaction U in the case of half-filled states. Let us express pseudofermion operators in the form of sums,

$$f_{\mathbf{i}\sigma}^+=f_{\mathbf{i}\sigma}^+(1-n_{\mathbf{i}-\sigma})+f_{\mathbf{i}\sigma}^+ n_{\mathbf{i}-\sigma},$$

$$n_{\mathbf{i}\sigma}=n_{\mathbf{i}\sigma}(1-n_{\mathbf{i}-\sigma})+n_{\mathbf{i}\sigma}n_{\mathbf{i}-\sigma}, \tag{9}$$

and introduce Hubbard's projection operators for pseudofermions, as was done by Hubbard for the real electron operators[30]:

$$X_{\mathbf{i}}^{\sigma 0}=f_{\mathbf{i}\sigma}^+(1-n_{\mathbf{i}-\sigma}),\quad X_{\mathbf{i}}^{2\sigma}=-\sigma f_{\mathbf{i}-\sigma}^+ n_{\mathbf{i}\sigma},$$

$$X_{\mathbf{i}}^{\sigma\sigma}=n_{\mathbf{i}\sigma}(1-n_{\mathbf{i}-\sigma})=n_{\mathbf{i}\sigma}n_{\mathbf{i}-\sigma},\quad X_{\mathbf{i}}^{22}=n_{\mathbf{i}\uparrow}n_{\mathbf{i}\downarrow},$$

$$X_{\mathbf{i}}^{00}=(1-n_{\mathbf{i}\uparrow})(1-n_{\mathbf{i}\downarrow}), \tag{10}$$

$$X_{\mathbf{i}}^{\sigma-\sigma}=f_{\mathbf{i}\sigma}^+ f_{\mathbf{i}-\sigma}=X_{\mathbf{i}}^{\sigma 2}X_{\mathbf{i}}^{2-\sigma}=X_{\mathbf{i}}^{\sigma 0}X_{\mathbf{i}}^{0-\sigma}.$$

These operators form a normalized basis for the group SU(4) with the obvious completeness condition

$$\sum_\mu X_{\mathbf{i}}^{\mu\mu}=1. \tag{11}$$

The second line of Eq. (9) can be rewritten in the form

$$n_{\mathbf{i}\sigma}=X_{\mathbf{i}}^{\sigma 0}X_{\mathbf{i}}^{0\sigma}+X_{\mathbf{i}}^{\sigma 2}X_{\mathbf{i}}^{2\sigma}. \tag{12}$$

As a result, Hamiltonian (2), (7) takes the form

$$H=H_0+H_{\rm int},$$

$$H_0=-\frac{h}{2}\sum_{\mathbf{i}\sigma}\,\sigma X_{\mathbf{i}}^{\sigma\sigma}+\frac{U}{2}\sum_{\mathbf{i}}\,(X_{\mathbf{i}}^{00}+X_{\mathbf{i}}^{22}), \tag{13}$$

$$H_{\rm int}=\frac{J}{2}\sum_{\mathbf{ij}\sigma}\,(X_{\mathbf{i}}^{\sigma-\sigma}X_{\mathbf{j}}^{-\sigma\sigma}-X_{\mathbf{i}}^{\sigma\sigma}X_{\mathbf{j}}^{-\sigma-\sigma})$$

$$=\frac{J}{2}\sum_{\mathbf{ij}\sigma\sigma'}\,\phi_{\mathbf{ij}}^{\sigma}\phi_{\mathbf{ji}}^{\sigma'}.$$

Here

$$\phi_{\mathbf{ij}}^{\sigma}=(\sigma X_{\mathbf{i}}^{2-\sigma}+X_{\mathbf{i}}^{\sigma 0})(\sigma X_{\mathbf{j}}^{-\sigma 2}+X_{\mathbf{j}}^{0\sigma}), \tag{14}$$

and the fictitious Hubbard repulsion parameter U for pseudofermions is introduced so as to preserve the particle–hole symmetry of the Heisenberg Hamiltonian.

Instead of using diagram techniques for the X-operators (see, for example, Refs. 19, 22, 31, and 32), we try to remain within the standard Feynman approach, but use the properties of the projection operators $X_{\mathbf{i}}^{\lambda\mu}$ in explicit form. We take for a basis of the diagram expansion the eigenstates of the Hamiltonian H_0 under the condition $U/T \equiv \beta U\to\infty$. As a result, we have the reduced Hamiltonian

$$\widetilde{H}=\widetilde{H}_0+H_{\rm int},\quad \widetilde{H}_0=-\frac{h}{2}\sum_{\mathbf{i}\sigma}\,\sigma f_{\mathbf{i}\sigma}^+ f_{\mathbf{i}\sigma} \tag{15}$$

with the partition function $\hat{\mathcal{Z}} = \mathrm{Tr}[\exp(-\beta\widetilde{H})]$, which includes only physical states $|\sigma\rangle = |\pm\rangle$ from set (8). The Hamiltonian H_0 reduces to $\widetilde{H}_0$ since the operators $X_{\mathbf{i}}^{\sigma\sigma}$ and $f_{\mathbf{i}\sigma}^+ f_{\mathbf{i}\sigma}$ have identical matrix elements in the reduced (physical) space. Now we can use $\widetilde{H}$ in the form (15) as a zero-approximation Hamiltonian for the Matsubara diagram technique.[2)]

Selection of one of the two forms of Hamiltonian (13) depends on which of the spin system states we are going to describe. Whereas the most convenient representation for the high-temperature paramagnetic phase or a state with long-

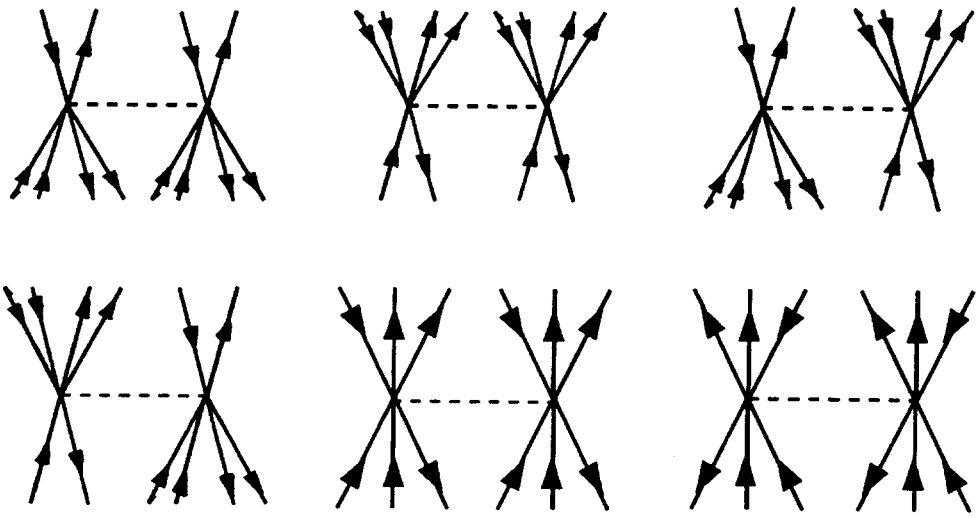


FIG. 1. Bare exchange vertices with due account of projection operators.

range magnetic order is $H_{\rm int}$ in terms of spin operators $X_{\mathbf{i}}^{\sigma\sigma'}$, it is natural to describe the RVB spin-liquid state in terms of operators $\phi_{\mathbf{ij}}^{\sigma}$.

Let us first consider the temperature Green's function

$$\mathscr{K}_{\mathbf{ij}}^{\perp}(\tau)=\langle T_{\tau}S_{\mathbf{i}}^{+}(\tau)S_{\mathbf{j}}^{-}(0)\rangle_{\tilde{H}}, \tag{16}$$

which describes elementary excitations in the standard theory of magnetism ($i\tau$ is the imaginary ''time''). To zeroth-order in the interaction, the function has the form

$$\mathscr{K}_{\mathbf{ij}}^{0}(\tau)=\frac{\delta_{\mathbf{ij}}}{4}\langle T_{\tau}f_{\mathbf{i}\uparrow}^{+}(\tau)f_{\mathbf{i}\downarrow}(\tau)f_{\mathbf{i}\downarrow}^{+}(0)f_{\mathbf{i}\uparrow}(0)\rangle_{\tilde{H}_0}. \tag{17}$$

Averaging is performed with the partition function $\mathscr{Z}=2\cosh(\beta h)$, $\beta=1/kT$. In accordance with Wick's theorem, this average can be presented in the form of a two-fermion loop and reduces to the simple expression

$$\mathscr{K}_{\mathbf{ij}}^{0}(\tau)=\frac{\delta_{\mathbf{ij}}}{4}e^{-h\tau}\begin{cases}\langle n_{\mathbf{i}\uparrow}(1-n_{\mathbf{i}\downarrow})\rangle_{\tilde{H}_0} & (\tau>0)\\ \langle n_{\mathbf{i}\downarrow}(1-n_{\mathbf{i}\uparrow})\rangle_{\tilde{H}_0} & (\tau<0)\end{cases}. \tag{18}$$

One can see that by virtue of Eq. (4), fermion states are generated in pairs, and the emergence of filling factors in the form of averages of projection operators $\langle X^{\sigma\sigma}\rangle_{\tilde{H}_0}$ [see Eq. (10)] shows that spin operators do not drive the system from the space of ''physical'' states $|\pm\rangle$.

Thus, the limit $U\to\infty$ for effective Hamiltonian (13) is equivalent to the limit $\lambda\to\infty$ in Abrikosov's procedure described above, which ''freezes out'' nonphysical pseudofermion states $|0\rangle$ and $|2\rangle$ without breaking the particle–hole symmetry.

The perturbation theory series for the function $\mathscr{K}^{\perp}$ can be constructed in accordance with the usual rules for calculating two-time Green's functions. This procedure leads to Larkin's equation[34]

$$\mathscr{K}=\Sigma+J\Sigma.\mathscr{K}. \tag{19}$$

Here Σ is the irreducible polarization operator, which is not separable with respect to the interaction. In Sec. 5 we will use this version of the diagram technique to calculate the spin diffusion coefficient near the Néel point.

Let us proceed to nonlocal spin-liquid correlations. We consider as an example an RVB homogeneous spin liquid described by the correlator

$$\mathscr{L}_{\mathbf{ij}}=\sum_{\sigma}\langle\phi_{\mathbf{ij}}^{\sigma}\phi_{\mathbf{ji}}^{\sigma}\rangle, \tag{20}$$

i.e., we use the second version of $H_{\rm int}$ in Eq. (13). Thus, the nonphysical states $|0\rangle$ and $|2\rangle$ are eliminated by the Hubbard procedure, since each fermion creation event at each site involves a projection operation in accordance with Eq. (14). This makes exchange vertex (13) more complicated; it can be described in the projection techniques by diagrams with twelve tails, as is shown by Fig. 1.

The role of projectors is to automatically eliminate a state with an opposite projection in creating a fermion with a given spin projection, and this guarantees that the creation operator acts on a state from the physical subspace $|\pm\rangle$. But, although correlator (20) is diagonal in subspace $|\pm\rangle$, the nonphysical states $|0\rangle$ and $|2\rangle$ manifest themselves as intermediate states in any attempt to describe the spin liquid in terms of fermion operators.

In Refs. 11 and 13 it was noted that introduction of a homogeneous RVB state in the mean-field approximation[16] violates the local gauge invariance due to constraint (4), (6), and long-wave fluctuations of gauge fields significantly change the character of RVB excitations in a two-dimensional Heisenberg lattice (see also Refs. 14, 35, and 36). In this paper, we do not consider the problem of long-wave fluctuations in gauge fields. We are interested primarily in nonlocal high-temperature magnetic fluctuations, which are also related, however, to the violation of the constraint.

As was shown in the fundamental study by Baskaran, Zou and Anderson,[16] the description of a uniform RVB state requires ''anomalous'' coupling between pseudofermions at different sites. It is clear that such a procedure drives the system beyond the physical space $|\pm\rangle$. The gauge theory of a spin liquid demonstrates that free propagation of a spinon is impossible. The complex shape of vertices in the projection technique (Fig. 1) indicates the same thing. Nonetheless,

we start construction of our scheme with a demonstration of how far this technique applies in the mean-field approximation; we then consider the possible effect of fluctuations on the mean-field solution.

Let us introduce an anomalous one-particle (one-fermion) temperature Green's function. In order to preserve particle–hole symmetry, let us express it in matrix form:

$$\hat{\mathscr{G}}_{\mathbf{ij}\sigma}(\tau)=-\langle T_\tau \hat{X}_{\mathbf{i}\sigma}(\tau)\hat{X}^+_{\mathbf{j}\sigma}(0)\rangle_{\tilde{H}}, \tag{21}$$

where

$$\hat{X}_{\mathbf{i}\sigma}=\begin{pmatrix} X^{0\sigma}_{\mathbf{i}} & X^{\sigma 0}_{\mathbf{i}} \\ \sigma X^{-\sigma 2}_{\mathbf{i}} & \sigma X^{2-\sigma}_{\mathbf{i}}\end{pmatrix}, \quad \hat{X}^+_{\mathbf{i}\sigma}=\begin{pmatrix} X^{\sigma 0}_{\mathbf{i}} & \sigma X^{2-\sigma}_{\mathbf{i}} \\ X^{0\sigma}_{\mathbf{i}} & \sigma X^{-\sigma 2}_{\mathbf{i}}\end{pmatrix}. \tag{22}$$

This Green's function has the structure

$$\hat{\mathscr{G}}_{\mathbf{ij}\sigma}(\tau)=-\begin{pmatrix} \langle T_\tau(X^{0\sigma}_{\mathbf{i}}(\tau)X^{\sigma 0}_{\mathbf{j}}(0) & \sigma\langle T_\tau(X^{0\sigma}_{\mathbf{i}}(\tau)X^{2-\sigma}_{\mathbf{j}}(0) \\ +X^{\sigma 0}_{\mathbf{i}}(\tau)X^{0\sigma}_{\mathbf{j}}(0))\rangle & +X^{\sigma 0}_{\mathbf{i}}(\tau)X^{-\sigma 2}_{\mathbf{j}}(0)\rangle \\ \sigma\langle T_\tau(X^{-\sigma 2}_{\mathbf{i}}(\tau)X^{\sigma 0}_{\mathbf{j}}(0) & \langle T_\tau(X^{-\sigma 2}_{\mathbf{i}}(\tau)X^{2-\sigma}_{\mathbf{j}}(0) \\ +X^{2-\sigma}_{\mathbf{i}}(\tau)X^{-0\sigma}_{\mathbf{j}}(0)\rangle & +X^{2-\sigma}_{\mathbf{i}}(\tau)X^{-\sigma 2}_{\mathbf{j}}(0)\rangle \end{pmatrix}. \tag{23}$$

The zero (one-site) matrix Green's function

$$\hat{g}_{\mathbf{i}\sigma}(\tau)=-\langle T_\tau \hat{X}_{\mathbf{i}\sigma}(\tau)\hat{X}^+_{\mathbf{i}\sigma}(0)\rangle_{\tilde{H}_0} \tag{24}$$

is diagonal, and its elements are

$$\begin{aligned} g^{(11)}_{\mathbf{i}\sigma}(\tau)&=-\langle T_\tau(X^{0\sigma}_{\mathbf{i}}(\tau)X^{\sigma 0}_{\mathbf{i}}(0)+X^{\sigma 0}_{\mathbf{i}}(\tau)X^{0\sigma}_{\mathbf{i}}(0))\rangle_0, \\ g^{(22)}_{\mathbf{i}\sigma}(\tau)&=-\langle T_\tau(X^{-\sigma 2}_{\mathbf{i}}(\tau)X^{2-\sigma}_{\mathbf{i}}(0) \\ &\quad +X^{2-\sigma}_{\mathbf{i}}(\tau)X^{-\sigma 2}_{\mathbf{i}}(0))\rangle_0. \end{aligned}$$

As in the previous case, the averaging $\langle\ldots\rangle_0\equiv\langle\ldots\rangle_{\tilde{H}_0}$ leaves the one-site Green's function in the physical sector of the Fock space. In particular,

$$\begin{aligned} g^{(11)}_{\mathbf{i}\sigma}(\tau_1-\tau_2)&=-\langle X^{\sigma 0}_{\mathbf{i}}(\tau_1)X^{0\sigma}_{\mathbf{i}}(\tau_2)\rangle_0 \\ &=-\langle X^{\sigma\sigma}_{\mathbf{i}}\rangle_0\ \exp[-i\sigma h(\tau_1-\tau_2)/2] \\ &\quad\times(\tau_1>\tau_2), \\ g^{(11)}_{\mathbf{i}\sigma}(\tau_1-\tau_2)&=\langle X^{\sigma 0}_{\mathbf{i}}(\tau_2)X^{0\sigma}_{\mathbf{i}}(\tau_1)\rangle_0 \\ &=\langle X^{\sigma\sigma}_{\mathbf{i}}\rangle_0\ \exp[-i\sigma h(\tau_1-\tau_2)/2](\tau_2>\tau_1). \end{aligned} \tag{25}$$

Unlike spin Green's functions (17), matrix elements of the function $\hat{g}_{\mathbf{i}\sigma}(\tau)$ formally represent the three-fermion loops containing one particle (spin up) and two hole (spin down) propagators, or one hole and two particle propagators. This function, however, can be simplified using the idempotence property of operator $b^\dagger b$, conditions (4) and (6), and Wick's theorem. By substituting the Hubbard operators in the interaction picture into Eq. (25), we obtain expressions for the elements of the one-site propagator,

$$g^{(11)}_{\mathbf{i}\uparrow}=\frac{1}{2}e^{-h\tau}\begin{cases} -\langle(1-n_{\mathbf{i}\downarrow})\rangle_0 & (\tau>0) \\ \langle n_{\mathbf{i}\downarrow}\rangle_0 & (\tau<0), \end{cases}$$

and a similar expression for the spin-down state.

One can easily check that the Green's function $\mathscr{G}^{(\alpha\alpha)}_{\mathbf{i}\sigma}$ is periodic, $\mathscr{G}_{\mathbf{ii}\sigma}(\tau<0)=-\mathscr{G}_{\mathbf{ii}\sigma}(\tau+1/T)$, so that by introducing the Matsubara frequencies $\omega_n=(2n+1)\pi T$ in the usual manner, we obtain

$$g^{(\alpha\alpha)}_{\mathbf{i}\sigma}(\omega_n)=\frac{1}{2}\,\frac{1}{i\omega_n+(-1)^\alpha\sigma h/2}. \tag{26}$$

The mean-field approximation[16] is based on the introduction of anomalous averages $\langle f^+_{\mathbf{i}\sigma}f_{\mathbf{j}\sigma}\rangle$. For the anomalous matrix Green's function (21), we must introduce four components:

$$\begin{aligned} \Delta^{11}_{ij\sigma}&=\langle X^{\sigma 0}_{\mathbf{i}}(\tau)X^{0\sigma}_{\mathbf{j}}(\tau'\to\tau)\rangle, \\ \Delta^{22}_{ij\sigma}&=\langle X^{-\sigma 2}_{\mathbf{i}}(\tau)X^{2-\sigma}_{\mathbf{j}}(\tau'\to\tau)\rangle, \\ \Delta^{12}_{ij\sigma}&=\langle X^{\sigma 0}_{\mathbf{i}}(\tau)X^{-\sigma 2}_{\mathbf{j}}(\tau'\to\tau)\rangle, \\ \Delta^{21}_{ij\sigma}&=\langle X^{2-\sigma}_{\mathbf{i}}(\tau)X^{0\sigma}_{\mathbf{j}}(\tau'\to\tau)\rangle, \end{aligned} \tag{27}$$

where $\Delta^{11}_{ij\sigma}=\Delta^{22}_{ij\sigma}$. Then one can easily check that the anomalous Green's function also satisfies a periodic condition like that in Eq. (21) on the inverse temperature. Thus, we can use the projection diagram technique in calculating the anomalous average $\Delta=\Sigma_\sigma\langle\phi^\sigma_{\mathbf{ij}}\rangle$, which characterizes a uniform RVB state. This ''order parameter'' can be derived from the relation

$$\Delta=\mathrm{Tr}(\hat{1}+\hat{\tau}_1)\hat{\mathscr{G}}_{\mathbf{ij}}(\tau\to-0), \tag{28}$$

where $\hat{1}$ and $\hat{\tau}_1$ are the Pauli matrices.

Let us rewrite Hubbard operators (10) in the particle–hole representation, $f_{\mathbf{i}\uparrow}\equiv a_{\mathbf{i}}$, $f_{\mathbf{i}\downarrow}\equiv b^+_{\mathbf{i}}$:

$$X^{\uparrow 0}_{\mathbf{i}}=a^+_{\mathbf{i}}b^+_{\mathbf{i}}b_{\mathbf{i}}, \quad X^{2\downarrow}_{\mathbf{i}}=a^+_{\mathbf{i}}b_{\mathbf{i}}b^+_{\mathbf{i}}, \quad X^{\uparrow\downarrow}_{\mathbf{i}}=a^+_{\mathbf{i}}b^+_{\mathbf{i}}\ldots$$

The mean-field approximation (28) corresponds to the following splitting of the interaction Hamiltonian H_{int}:

$$H_{\mathrm{MF}}=J\Delta\sum_{\mathbf{i}}\sum_{\mathbf{j}}^{\langle nn\rangle}(Y^{(h)}_{\mathbf{ij}}+Y^{(p)}_{\mathbf{ij}}), \tag{29}$$

where

$$\begin{aligned} Y^{(p)}_{\mathbf{ij}}&=a^+_{\mathbf{i}}b^+_{\mathbf{i}}b_{\mathbf{i}}b^+_{\mathbf{j}}b_{\mathbf{j}}a_{\mathbf{j}}+a^+_{\mathbf{i}}b_{\mathbf{i}}b^+_{\mathbf{i}}b_{\mathbf{j}}b^+_{\mathbf{j}}a_{\mathbf{j}}+a^+_{\mathbf{i}}b_{\mathbf{i}}b^+_{\mathbf{i}}b^+_{\mathbf{j}}b_{\mathbf{j}}a_{\mathbf{j}} \\ &\quad+a^+_{\mathbf{i}}b^+_{\mathbf{i}}b_{\mathbf{i}}b_{\mathbf{j}}b^+_{\mathbf{j}}a_{\mathbf{j}}, \\ Y^{(h)}_{\mathbf{ij}}&=b_{\mathbf{i}}a_{\mathbf{i}}a^+_{\mathbf{i}}a_{\mathbf{j}}a^+_{\mathbf{j}}b^+_{\mathbf{j}}+b_{\mathbf{i}}a^+_{\mathbf{i}}a_{\mathbf{i}}a^+_{\mathbf{j}}a_{\mathbf{j}}b^+_{\mathbf{j}}+b_{\mathbf{i}}a_{\mathbf{i}}a^+_{\mathbf{i}}a^+_{\mathbf{j}}a_{\mathbf{j}}b^+_{\mathbf{j}} \end{aligned}$$

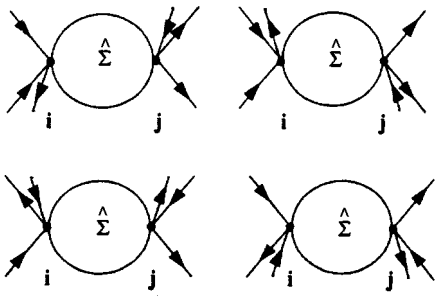


FIG. 2. Components of $\hat{\Sigma}$ matrix for the one-particle Green's function $\mathscr{G}_{\mathbf{ij}}$.

$$+ b_{\mathbf{i}} a_{\mathbf{i}}^{+} a_{\mathbf{i}} a_{\mathbf{j}} a_{\mathbf{j}}^{+} b_{\mathbf{j}}^{+} .$$

In terms of perturbation theory, this approximation can be described by the diagrams for the self-energy component $\hat{\Sigma}_{\mathbf{ij}}$ of Green's function (23) shown in Fig. 2. The four diagrams correspond to the four elements of $\hat{\Sigma}$. The Dyson matrix equation in this approximation is expressed by the diagrams in Fig. 3, in which double lines denote the one-site matrices $g_{\mathbf{i}\sigma}$, the dashed line denotes the Heisenberg exchange constant, and thick lines with two arrows denote the anomalous Green's function $\hat{\mathscr{G}}_{\mathbf{ij}\sigma}$. The Dyson equation

$$\hat{\mathscr{G}}_{\mathbf{ij}\sigma}(\omega_n) = \hat{g}_{\mathbf{i}\sigma}(\omega_n)\left(\delta_{\mathbf{ij}} + \sum_{\mathbf{l}} \hat{\Sigma}_{\mathbf{il}} \hat{\mathscr{G}}_{\mathbf{lj}\sigma}(\omega_n) \right) \tag{30}$$

is Fourier transformed to (as $h \to 0$)

$$2 i\omega_n \mathscr{G}_{\mathbf{k}\sigma}^{(\alpha\beta)}(i\omega_n) = \delta_{\alpha\beta} + J\Delta\varphi(\mathbf{k}) \sum_{\gamma} \mathscr{G}_{\mathbf{k}\sigma}^{(\gamma\beta)}(i\omega_n). \tag{31}$$

A solution of this equation system is

$$\mathscr{G}_{\mathbf{k}\sigma}^{(11)}(i\omega_n) = \frac{1}{2} \frac{i\omega_n - \epsilon_{\mathbf{k}}/2}{i\omega_n (i\omega_n - \epsilon_{\mathbf{k}})},$$

$$\mathscr{G}_{\mathbf{k}\sigma}^{(12)}(i\omega_n) = \frac{1}{2} \frac{\sigma\epsilon_{\mathbf{k}}/2}{i\omega_n (i\omega_n - \epsilon_{\mathbf{k}})}. \tag{32}$$

Here $\epsilon_{\mathbf{k}}$ is the spinon dispersion relation in the mean-field approximation in the form

$$\epsilon_{\mathbf{k}} = J\Delta\varphi(\mathbf{k}) \tag{33}$$

in the case of antiferromagnetic exchange only between nearest neighbors; $\varphi(\mathbf{k})$ is the corresponding form factor:

$$\varphi(\mathbf{k}) = \sum_{\mathbf{l}}^{(nn)} e^{i\mathbf{k}\cdot\mathbf{l}}. \tag{34}$$

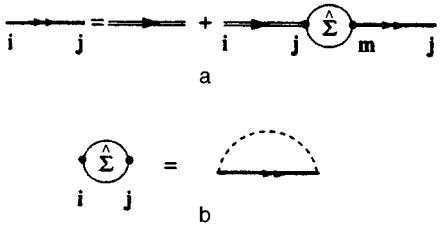


FIG. 3. (a) Dyson equation and (b) self-energy part of the Green's function $\mathscr{G}_{\mathbf{ij}}$ in the mean-field approximation.

By substituting Green's functions (32) into Eq. (28), we obtain a self-consistent equation for Δ:

$$\Delta = (ZN)^{-1} \sum_{\mathbf{k}} \varphi(\mathbf{k}) \tanh \frac{\beta[J\Delta\varphi(\mathbf{k}) - \mu]}{2}, \tag{35}$$

where Z is the coordination number. The chemical potential μ is treated as a Lagrange multiplier when constraint (4) is substituted into the Hamiltonian. This operation corresponds to substitution of $i\omega_n + \mu$ for $i\omega_n$. As usual, the local constraint can be replaced with a global one in the mean-field approximation:

$$N^{-1} T \sum_{\mathbf{k}} \sum_{\omega_n} \mathrm{Tr}(\hat{1} + \hat{\tau}_1) \hat{\mathscr{G}}_{\mathbf{k}}(i\omega_n) = 0. \tag{36}$$

By substituting Green's function (32) into Eq. (36), we obtain another self-consistency condition, which fixes μ at the mid-position of the spinon ''band,'' in accordance with particle–hole symmetry.

The ''phase transition'' temperature T^* at which a nontrivial solution for Δ emerges is given by

$$T^* = \frac{J}{2} (ZN)^{-1} \sum_{\mathbf{k}} \varphi^2(\mathbf{k}), \tag{37}$$

which is usually derived in the mean-field approximation using the functional integration technique (see, for example, Refs. 6, 37, and 38).

Thus, we have found that kinematic constraints on the pseudofermion representation of spin operators taken into account through Hubbard projection operators do not affect the mean-field solution for the RVB state as long as particle–hole symmetry is preserved at each step of the calculation. In this respect, the situation is different from that in which the same problem is solved by the Hubbard operator technique for the $t-J$ model with a finite density of holes,[39] where this symmetry is violated from the outset, since only doubly filled states $|2\rangle$ are excluded. In Ref. 22 another symmetry-based approach to elimination of nonphysical states is suggested, in which the ''fermion'' set $|0\rangle$, $|2\rangle$ is replaced with a unified ''boson'' vacuum $|V\rangle$.

Although the projection technique does not contribute any new features to the mean-field solution for the uniform RVB liquid, it offers, in principle, new opportunities for taking gauge fluctuations into account, which inevitably occur in spinon propagation. Moreover, as will be shown in the next section, in a three-dimensional Kondo lattice, spin liquid is formed in the neighborhood of the antiferromagnetic instability, because magnetic fluctuations are a decisive factor for both the transition temperature to the RVB state and the mechanism of this transition.

3. STABILIZATION OF SPIN LIQUID IN THE KONDO LATTICE AT HIGH TEMPERATURES. MEAN-FIELD APPROXIMATION

It is well known[40] that in the three-dimensional Heisenberg lattice the ground state energy of the RVB phase, E_{SL}, is higher than the antiferromagnetic state energy E_{AFM}. It has also been shown, however, that in the Kondo lattice described by the Hamiltonian (1), spin-flip scattering processes

can lead to stabilization of the RVB phase with respect to the magnetically ordered phase.[6,18] Since antiferromagnetic and spin-liquid correlations in the sf-exchange model are governed by the same coupling constant J_{RKKY}, the temperature at which the spin liquid is formed is close to the point of magnetic instability, $T^*-T_N<T_N$, so that antiferromagnetic correlations can significantly alter the character of a transition to the RVB phase, as compared to the results obtained in the mean-field approximation.

In order to describe formation of spin liquid in the Kondo lattice, we take Hamiltonian (1) in the original form

$$H_{\mathrm{eff}}=\sum_{\mathbf{k}\sigma}\ \varepsilon_k c^+_{\mathbf{k}\sigma}c_{\mathbf{k}\sigma}+\frac{1}{4}J_{sf}\sum_{\mathbf{i}}\ c^+_{\mathbf{i}\sigma}c_{\mathbf{i}\sigma'}f^+_{\mathbf{i}\sigma'}f_{\mathbf{i}\sigma}. \tag{38}$$

As mentioned in the Introduction, we operate in the range of parameters $\alpha\approx\alpha_c$ of Doniach's diagram,[3] in which all characteristic temperatures ($T_K\sim\varepsilon_F\exp(-1/2\alpha)$, $T_{N0}\sim\varepsilon_F\alpha^2$, and T^*, which is to be calculated) are of the same order of magnitude, so that in constructing the real phase diagram one must take into account the mutual effects of all three types of correlation—in particular, the change in the Néel temperature with respect to T_{N0} as given by simple perturbation theory in the parameter α.

As noted above, in this study we limit discussion to the range of high temperatures $T>T_K,T_{N0}$, in which the magnetic subsystem is a lattice of paramagnetic spins immersed in the Fermi sea of conductance electrons, and NCA applies to the one-site paramagnetic sf-scattering, i.e., each spin localized at a lattice site scatters conduction electrons independently of other spins. As the temperature is reduced, both Kondo scattering and correlations among lattice sites due to the indirect RKKY interaction are intensified.

The problem of competition between the indirect exchange among lattice sites and one-site sf-scattering has been discussed in literature many times, largely in terms of the Kondo problem with two impurities. In particular, Varma[41] analyzed the mutual influence of Kondo scattering and RKKY interaction at high temperatures by perturbation theory and concluded that the mutual influence of these two processes is small, at least in the leading logarithmic approximation in $\alpha\ \ln(\varepsilon_F/T)$. In this section, we will show that in the Kondo lattice, the effect of spin-flip scattering on magnetic correlations is a decisive factor for stabilization of the RVB phase in the critical region of Doniach's diagram, $\alpha\sim\alpha_c$.

In describing the intersite magnetic interaction under conditions of Kondo scattering in the noncrossing approximation (NCA), the effective vertex of the RKKY exchange $\tilde{J}_{\mathbf{ij}}(T,\epsilon)$ is determined by the diagram in Fig. 4a. In this diagram, dashed lines denote electron Green's functions, and the ingoing and outgoing lines correspond to pseudofermion operators. The one-site sf-exchange vertices Γ include loops corresponding to the leading logarithmic approximation in $\alpha\ \ln(\varepsilon_F/T)$ for the Kondo problem[3)] (Fig. 5). As a result, the effective interaction is given by

$$\tilde{J}_{\mathbf{ij}}(T,\varepsilon_m)=\Pi(R,\varepsilon_m)\Gamma^2, \tag{39}$$

where $\varepsilon_m=2m\pi T,\ R=|R_{\mathbf{i}}-R_{\mathbf{j}}|$.

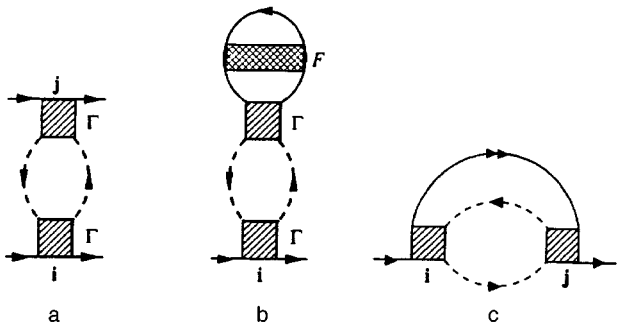


FIG. 4. (a) Effective vertex of renormalized RKKY interaction; self-energy part of the one-particle Green's function in the mean-field approximation (b) for the Néel and (c) RVB phase.

In the spirit of the logarithmic perturbation theory,[10] the argument of vertex Γ should contain only the highest input frequency, which is determined in our case by energies of electronic Green's functions included in the polarization loop $\Pi(R,\varepsilon_m)$ in the integral $\hat{J}_{\mathbf{ij}}(T,\varepsilon_m)$ (Eq. (39)). The polarization operator in the coordinate representation has the form

$$\Pi(\mathbf{R},\varepsilon_m)=T\sum_n\ D(-\mathbf{R},\omega_n+\varepsilon_m)D(\mathbf{R},\omega_n). \tag{40}$$

Since all heavy-fermion systems are characterized by large lattice constants, we use for electronic Green's functions $D(\mathbf{R},\omega_n)$ an expression asymptotic in p_FR:

$$D(R,\omega_n)=-\frac{p_F}{2\pi v_F R}\exp\left(-\frac{|\omega_n|}{2\varepsilon_F}p_FR+ip_FR\ \mathrm{sign}\ \omega_n\right), \tag{41}$$

so that the polarization operator takes the form

$$\Pi(R,\varepsilon_m)=\left(\frac{p_F}{2\pi v_F R}\right)^2 T\sum_{n=-\infty}^{n=\infty}\exp\left(-\frac{|\omega_n|}{v_F}R-\frac{|\omega_n+\varepsilon_m|}{v_F}R+ip_FR[\mathrm{sign}\omega_n+\mathrm{sign}(\omega_n+\varepsilon_m)]\right). \tag{42}$$

In the static limit,

$$\tilde{J}_R(T,0)=T\sum_n\ D^2(R,\omega_n)\Gamma^2(\omega_n,T). \tag{43}$$

The temperature dependence in Eq. (43) is largely determined by one-site vertices, and in the polarization loop one can use the condition $2\pi TR/v_f\ll 1$ and change summation over discrete frequencies to integration (see Appendix I). Then the exchange integral takes the form

FIG. 5. Parquet diagrams for effective vertex Γ.

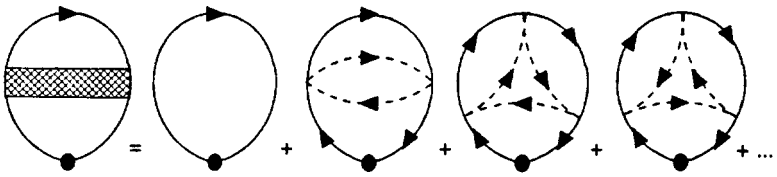


FIG. 6. Single-site diagrams describing Kondo screening of a localized spin.

$$\widetilde{J}_R(T,0)=2\left(\frac{p_F}{2\pi v_F R}\right)^2\cos(2p_FR)\times\int_{T\to 0}^{\infty}\frac{d\varepsilon}{2\pi}\exp\left(-\frac{\varepsilon}{\varepsilon_F}p_FR\right)\Gamma^2(\varepsilon,T). \quad (44)$$

This equation transforms to the standard RKKY exchange integral when modified vertices are replaced with the ''bare'' integrals, $\Gamma\to J_{sf}\Omega_0$, where Ω_0 is the elementary cell volume:

$$J_R^0=(J_{sf}\Omega_0)^2\Pi(R,0)=(J_{sf}\Omega_0)^2\frac{mp_F^4}{\pi^3}\times\left[\frac{\cos(2p_FR)}{(2p_FR)^3}+O\left(\frac{1}{(2p_FR)^4}\right]\right.$$
$$\equiv\left(\frac{J_{sf}^2}{\varepsilon_F}\right)\frac{(p_Fa_0)^6}{2\pi^3}\Phi(2p_FR). \quad (45)$$

Let us substitute into $\widetilde{J}_R(T,0)$ the vertex $\Gamma(\varepsilon,T)$ calculated in the leading logarithmic approximation, in accordance with diagrams given in Fig. 5 with the input frequency ε satisfying the condition $\ln(\varepsilon_F/\overline{\varepsilon})\gg 1$. For the characteristic energy $\overline{\varepsilon}\gg 1$, which determines integral (44) (see Appendix II), we find that the exchange parameter can be approximated by the function

$$\widetilde{J}_R(T,0)\approx\epsilon_F\frac{(p_Fa_0)^6}{2\pi^3}\left(\frac{J_{sf}}{\varepsilon_F}\right)^2\Phi(2p_FR)\times\left(1+2\alpha\,\ln\frac{T}{\varepsilon_F}\right)^{-n}. \quad (46)$$

The exponent n in this function depends on α and the argument of the oscillating function $\Phi(p_FR)$ (see the insert in Fig. 11). Thus, one can see that Kondo scattering has little influence on the form and spatial periodicity of the indirect exchange integral for $T>T_K$.[41] But this integral can be larger, and the larger the separation R between magnetic f-ions, the greater the increase.

In calculating the polarization operator and RKKY integral (46), we assumed that the electron Fermi surface was spherical. Note, however, that the exponent n in Eq. (46) is sensitive to the asymptotic behavior of the function $\Phi(2p_FR)$, so that the role of Kondo processes in intensification of the exchange turns out to be important in the case of a highly anisotropic Fermi surface. In the limiting case of a cylindrical Fermi surface,

$$\Phi(2p_FR)=-\left[\frac{\sin(2p_FR)}{(2p_FR)^2}+O\left(\frac{1}{(2p_FR)^3}\right)\right] \quad (47)$$

(see Appendix I), so that at the same value of p_FR, the integral $\widetilde{J}_R(T,0)$ is larger in the case of a cylindrical Fermi surface than in the case of a spherical surface.

Thus, the spin system can be described at $T>T_K$ by the effective RKKY Hamiltonian with the vertex shown in Fig. 4a in the nearest-neighbor approximation and under the assumption that the RKKY nearest-neighbor coupling has the antiferromagnetic sign. In the mean-field approximation, we treat the problem of stabilization of the spin liquid as a comparison between temperatures of transitions to the RVB state $[T^*(\alpha)]$ and to the antiferromagnetic state $[T_N(\alpha)]$ under conditions of sufficiently strong Kondo scattering, $\alpha\to\alpha_{c0}-0$, and the stabilization criterion is the inequality $T^*(\alpha)>T_N(\alpha)$. The function $T_N(\alpha)$ deviates from the quadratic function prescribed by the bare RKKY vertex. Along with the intensification of one-site vertices described by Eq. (46) and discussed above, there is a dynamic Kondo screening of localized spins, which is the reason for the suppression of antiferromagnetic order as $\alpha\to\alpha_{c0}$.

In the mean-field approximation, the transition temperatures $T_N(\alpha)$ and $T^*(\alpha)$ can be derived from the exchange vertex in Fig. 4a by closing spin-fermion lines, as shown in Figs. 4b and 4c, respectively. The first of these diagrams determines the molecular field for commensurate magnetic ordering characterized by the antiferromagnetic vector $\mathbf{Q}$ such that $\mathbf{Q}\cdot\mathbf{R}_{ij}=\pi$. The suppression of magnetic correlation by Kondo scattering is described by the vertex $F(T)$ in the diagram of Fig. 4b.[42,43] Summation of the set of logarithmic diagrams, the first of which are shown in Fig. 6, yields

$$F(T)=1-2\alpha\,\ln\frac{\varepsilon_F}{T}\Big/\ln\frac{T}{T_K}. \quad (48)$$

Although the function $F(T)$ deviates from this formula as $T\to T_K$,[44] and complete screening occurs only at $T=0$, the suppression of magnetic correlations compensates for the exchange intensification and thus reduces T_N as $\alpha\to\alpha_{c0}$.

The self-energy part of the one-site Green's function $\mathcal{G}_{\mathbf{ii}}$ (Eq. (21)), corresponding to the diagram of Fig. 4b, is

$$\Sigma_N(T)=\lambda\widetilde{J}(R,T)\langle S_z\rangle_T \quad (49)$$

(the factor λ is determined by the lattice configuration). Hence we derive for the mean spin

$$\langle S_z\rangle_T=\frac{1}{2}(\langle a_{\mathbf{i}}^+a_{\mathbf{i}}\rangle+\langle b_{\mathbf{i}}^+b_{\mathbf{i}}\rangle-1)$$

a self-consistent equation

$$\langle S_z\rangle_T=\frac{1}{2}F(T)\tanh\frac{\Sigma_N(T)}{2T}, \quad (50)$$

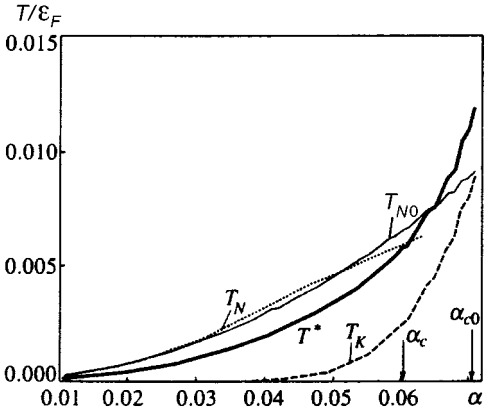


FIG. 7. Generalized Doniach diagram taking the RVB phase into account.

which is, naturally, the standard Brillouin equation for Weiss' molecular field taking Kondo screening into account.

The mean-field equation for Δ (Eq. (28)) is determined by the self-energy part of the anomalous Green's function $\hat{G}_{\mathbf{ij}}(\tau)$ (Eq. (23)) shown in Fig. 4c. This diagram can be substituted into the Dyson equation (Fig. 3), which in this case takes the form

$$\hat{G}(\mathbf{p},\omega_n)=g_0(\omega_n)\Bigg[1-2T\sum_m\sum_{\mathbf{q}}\tilde{J}(\mathbf{p}-\mathbf{q},\omega_n-\omega_m)\times\hat{G}(\mathbf{q},\omega_m)\hat{G}(\mathbf{p},\omega_n)\Bigg]. \tag{51}$$

Here $g_0(\omega_n)$ is the zero one-site Green's function with components (26), and $\tilde{J}(\mathbf{p}-\mathbf{q},\omega_n)$ is a Fourier transform of the indirect exchange integral (39), which in the nearest-neighbor approximation takes the form

$$\tilde{J}(\mathbf{q},\varepsilon_m)=\sum_{\mathbf{l}=0,\langle\mathbf{l}\rangle_{nn}}\tilde{J}_R(\varepsilon_m)e^{-i\mathbf{qr}}=\tilde{J}_0(\varepsilon_m)+\tilde{J}_R(\varepsilon_m)\varphi(\mathbf{q}). \tag{52}$$

The one-site integral $\tilde{J}_0(T,0)$ is estimated as $\alpha^2 T\ln(\varepsilon_F/T)$. Since this integral contains an additional small factor α at $T\sim T^*$, as compared to the intersite integral (46), it can be omitted.

By neglecting, as usual, the frequency dependence of the RKKY interaction, we obtain the mean-field equation (35) for Δ with the coupling constant $J=\tilde{J}_R(T,0)$. As follows from the configuration of the anomalous self-energy part (Fig. 4c), the screening effect responsible for suppression of local magnetic moments does not affect the mean-field parameter Δ, which can be naturally attributed to the singlet nature of the RVB-coupling. The Kondo ''screening'' radius can be estimated by high-temperature perturbation theory to be $\hbar v_F/2T_K$, which is much larger than the correlation radius of the singlet RVB pair, since electron scattering by these pairs is inefficient.

Calculations of the temperatures T^* and T_N by Eqs. (35), (46), (49), and (50) are given in Fig. 7 (see also Ref. 18). This graph shows that as $\alpha\to\alpha_{c0}$, these temperatures become closer, a new critical point α_c emerges in Doniach's diagram, on the right of which the RVB phase is stable with respect to the antiferromagnetic phase, and this stabilization takes place in the logarithmic neighborhood of the Kondo temperature. A calculation of T_N for $\alpha>\alpha_c$ makes no sense, because magnetic ordering in this region should follow another scenario.

Thus, we conclude that stabilization of a homogeneous RVB spin liquid in a three-dimensional Kondo lattice can occur only near the magnetic instability point under conditions of sufficiently strong Kondo screening of localized spins by conduction electrons. This result, obtained in the mean-field approximation, indicates that stabilization of the spin-liquid phase is incompatible with formation of Kondo singlet states characterized by anomalous averages $\langle c_i^+ f_i\rangle$,[6,45] since anomalous Kondo scattering is frozen at $T\approx T^*>T_K$. This resolves Nozières' well-known paradox[46] about the impossibility of screening all spins in the Kondo lattice by electrons from a thin layer of width T_K near the Fermi surface. In the scheme proposed above, the screening vanishes at sufficiently high temperatures above T_K, the Kondo temperature itself is not a singular point of the theory, renormalization of the sf-exchange integral is frozen at about $\tilde{J}(T^*)$, and at $T<T_K$, T^* electrons interact not with localized spins, but with spin-liquid excitations of the spinon type (see also Ref. 47).

In addition to the disadvantages related to violation of local gauge invariance, however, the mean-field approximation in the case of RVB coupling has another flaw, namely, it does not take into account the proximity of the spin system to the antifferomagnetic instability. In the following sections, we discuss possible consequences of this proximity for the RVB state, first in the self-consistent field approximation, then beyond this model.

4. EFFECT OF SPIN FLUCTUATIONS AND MAGNETIC ANISOTROPY ON RVB PHASE STABILIZATION

In the previous section, we determined that antiferromagnetic fluctuations inevitably turn out to be strong in an RVB spin liquid in the three-dimensional Kondo lattice at high temperatures $T\sim T^*$, and can lead, in principle, to magnetic ordering at $T\ll T^*$. Leaving this issue for subsequent studies, let us consider now the effect of spin fluctuations on features of the transition to the spin-liquid state in the mean-field approximation, but using its modification obtained through the projection technique, in which the order parameter is defined by Eq. (28). The diagram technique using Hubbard operators and developed in Sec. 2 allows us to take into consideration long-wave fluctuations of gauge fields due to the U(1) noninvariance of the RVB order parameter. Terms that take the phase of function Δ into account can be introduced into the effective Hamiltonian in standard fashion.[13,14] It is known that long-wave fluctuations in calibration fields do not lead to divergences destabilizing RVB averages in three-dimensional systems. Therefore, the introduction of such fluctuations reduces to the usual Fermi-liquid renormalizations with due account of the particle–hole symmetry condition. In two-dimensional Heisenberg lattices, however, fluctuations are important and must be taken into consideration.[13,14] In what follows, we do not discuss the

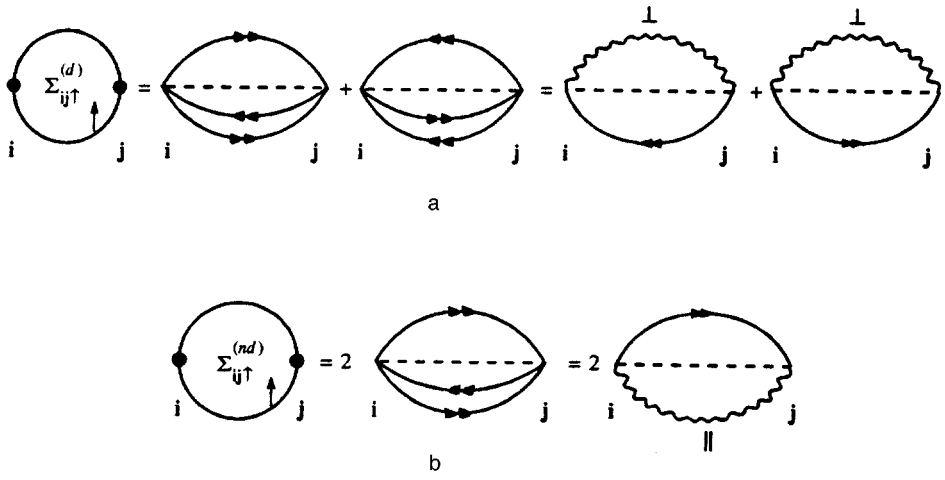


FIG. 8. Self-energy part for the anomalous propagator $\mathscr{G}_{ij}$, including the contribution of critical fluctuations in the mean-field approximation.

issue of long-wave fluctuations in gauge fields, and our analysis is limited to the mean-field approximation in a fixed calibration.

Having expressed the mean-field Hamiltonian in the form of Eq. (29), we considered in the subsequent calculations additional operators in $Y^{(p,h)}_{ij}$ as purely static projection operators, eliminating nonphysical states in thermodynamic averages. We now consider the fluctuation component of this ''kinematic'' interaction by transforming the effective mean-field spinon Hamiltonian for the Kondo lattice as follows:

$$H^{(\mathrm{RKKY})}_{\mathrm{MF}}=\tilde{J}\Delta\sum_{ij\sigma}\phi^{\sigma}_{ij}\equiv\tilde{J}\Delta\sum_{ij}\,(a^{+}_{i}K_{ij}a_{j}-a_{j}K^{+}_{ji}a^{+}_{i}+b_{i}K^{+}_{ij}b^{+}_{j}-b^{+}_{j}K_{ji}b_{i}). \tag{53}$$

Here $\tilde{J}$ is the renormalized constant of the RKKY interaction given by Eq. (46),

$$K_{ij}=S^{-}_{i}S^{+}_{j}-S^{z}_{i}S^{z}_{j}+\frac{1}{4},$$

and $K^{+}_{ij}=K_{ij}$. In the vicinity of the magnetic instability point, it is natural to consider operator K_{ij} as an operator describing critical excitations due to spinon propagation at temperatures close to T_N.

In order to obtain an expression for the spinon Green's function corresponding to this approximation, we reconsider the definition of its self-energy part. In the standard mean-field theory (Fig. 3), projection operators were included in the static approximation. The diagrams in Fig. 8 show how the diagonal and off-diagonal components of the self-energy part of the Green's function $\mathscr{G}_{ij}$, including transverse and longitudinal spin correlators, can be constructed from the vertices shown in Fig. 1. The lines with two arrows in Fig. 8 denote anomalous propagators

$$g^{\uparrow}_{ij}=-\langle T_{\tau}a_{i}(\tau)a^{+}_{j}(\tau')\rangle,$$
$$g^{\downarrow}_{ij}=-\langle T_{\tau}b_{i}(\tau)b^{+}_{j}(\tau')\rangle, \tag{54}$$

and wavy lines denote transverse and longitudinal correlation functions

$$\mathscr{K}^{\pm}_{ij}(\tau\to 0)=\langle T_{\tau}a^{+}_{j}(\tau+0)b^{+}_{j}(\tau+0)b_{i}(0)a_{i}(0)\rangle=\langle T_{\tau}S^{+}_{j}(\tau+0)S^{-}_{i}(0)\rangle,$$
$$\mathscr{K}^{zz}_{ij}(\tau\to 0)=\langle T_{\tau}b^{+}_{j}(\tau+0)b_{j}(\tau+0)b^{+}_{i}(0)b_{i}(0)\rangle=\frac{1}{4}-\langle T_{\tau}S^{z}_{j}(\tau+0)S^{z}_{i}(0)\rangle. \tag{55}$$

Unlike the fully anomalous Green's function (23), the anomalous functions (54) are one-particle propagators, while intersite spin correlators (55) are formed from projection operators. Now the sum of diagonal elements

$$\Sigma^{(d)}_{ij\uparrow}=\Sigma^{(11)}_{ij\uparrow}+\Sigma^{(22)}_{ij\uparrow}$$

in Eq. (28) is determined by the diagrams in Fig. 8a, while the contribution of off-diagonal elements

$$\Sigma^{(nd)}_{ij\uparrow}=\Sigma^{(12)}_{ij\uparrow}+\Sigma^{(21)}_{ij\uparrow}$$

corresponds to the diagrams in Fig. 8b. In deriving these expressions, we have used definition (3) and condition (6). Similar diagrams can be obtained for $\Sigma_{ij\downarrow}$. Summation of all these contributions in the mean-field approximation yields the effective Hamiltonian (53).

In the critical region $T_N<T<T^*$, the main contribution to spin correlators (53) is due to long-wave excitations with $\mathbf{k}\to 0$ and short-wave excitations with $\mathbf{k}\to\mathbf{Q}$ (see, for example, Ref. 48 and Sec. 5). The behavior of the response function $K(\mathbf{k})$ in the long-wave (hydrodynamic) limit $k\to 0$ is determined by fluctuations of the total magnetization of sublattices (which is zero in antiferromagnetic systems) and is diffusive in nature:

$$K^{R}(\mathbf{k},\omega)=K_{0}(\mathbf{k})\frac{iDk^{2}}{\omega+iDk^{2}}, \tag{56}$$

where

$$K_{0}(\mathbf{k})=\mathscr{K}(\mathbf{k},\omega=0)=\frac{\chi_{0}}{\tau+[1-J(\mathbf{k})/J(\mathbf{Q})]}\approx\frac{1}{2}\chi_{0}(T_N),$$

$$J(\mathbf{k})=J\sum_{\langle \mathbf{g}\rangle} e^{i\mathbf{kg}},\quad \chi_0(T)=\frac{S(S+1)}{3T},\quad \tau=\frac{T-T_N}{T_N} \tag{57}$$

(in Eq. (56) we have passed to the retarded Green's function for real frequency ω).

Near the antiferromagnetic vector $\mathbf{Q}$, the response function behavior is relaxation-like:

$$K^R(\mathbf{q},\omega)=\frac{1}{-i\omega/\Gamma\chi_0+K_0^{-1}(\mathbf{q})},\quad \mathbf{q}=\mathbf{k}-\mathbf{Q}, \tag{58}$$

where

$$K_0(\mathbf{q})=\mathscr{K}(q,\omega=0)=\frac{\chi_0}{\tau+(ql_0)^2} \tag{59}$$

is the Ornstein–Zernike static correlation function, and l_0 is the elementary excitation mean free path, which is comparable to the lattice constant.

In the mean-field approximation, we ignore the retardation of the RKKY interaction, and the diagrams in Fig. 8 yield for the self-energy

$$\tilde{\epsilon}_{\mathbf{k}}=\Sigma(\mathbf{k})=2\tilde{J}T^2\sum_{n,m\mathbf{q}}\sum_{s}\varphi(\mathbf{k}-\mathbf{q})g_{\mathbf{k}}(i\omega_n)\mathscr{K}^s_{\mathbf{q}}(i\varepsilon_m)$$

$$\approx\tilde{J}\Delta\left(\frac{\varphi(\mathbf{k})}{2}+2T\sum_{\mathbf{q}}\varphi(\mathbf{k}-\mathbf{q})K_0(\mathbf{q})\right). \tag{60}$$

Here s is the polarization index, while the anomalous Green's function $g_{\mathbf{k}}$ is expressed as $g_{\mathbf{k}}(i\omega_n)=(i\omega_n-\tilde{\epsilon}_{\mathbf{k}})^{-1}$. At high temperatures, we retain only the term with $\varepsilon_m=0$ in the sum over even Matsubara frequencies; then the spin Green's function $\mathscr{K}^s(\mathbf{q},0)$ in Fig. 8 has the same form in both the hydrodynamic and critical regions,[48] so that the main contribution to the spinon spectrum renormalization is due to the static susceptibility $K_0(\mathbf{q})$ (Eq. (59)).

The order parameter Δ defined by Eq. (28) and corresponding to the approximation of Eq. (53) and diagrams of Fig. 8 is given by

$$\Delta=\frac{1}{z}\sum_{\mathbf{pq}}\varphi(\mathbf{p}-\mathbf{q})\left[\frac{1}{2}\delta_{\mathbf{q},0}+2TK_0(\mathbf{q})\right]\tanh\frac{\tilde{\epsilon}_{\mathbf{p}}}{2T}. \tag{61}$$

Self-consistent equations (35) and (61) have been derived for the simplest case of isotropic exchange, which is, generally speaking, never realized in Kondo lattices. Therefore, before analyzing the effect of spin fluctuations on T^*, we generalize the mean-field theory to the case of anisotropic exchange. Let us introduce an exchange integral $J_{\mathbf{ij}}=\{J_\parallel,J_\perp\}$, where $J_\parallel$ and $J_\perp$ are the coupling constants for nearest neighbors in the basal plane and in the perpendicular direction, respectively. The degree of exchange anisotropy is measured by the parameter $\gamma=J_\perp/J_\parallel$. Now, instead of Hamiltonian (29) or (53), we must write the anisotropic mean-field Hamiltonian

$$H_{\mathrm{MF}}=\sum_{\mathbf{i},\rho_\perp}J_\perp\Delta_\perp Y_{\mathbf{i},\mathbf{i}+\rho_\perp}+\sum_{\mathbf{i},\rho_\parallel}J_\parallel\Delta_\parallel Y_{\mathbf{i},\mathbf{i}+\rho_\parallel}. \tag{62}$$

Here the anomalous averages $\langle Y_{\mathbf{i},\mathbf{i}+\rho_u}\rangle$, where $u=\perp,\parallel$, are derived from the equation system

$$\Delta_u=\frac{1}{z_u}\sum_{\mathbf{p}}\phi_u\left(\mathbf{p},\frac{T}{T_N},\gamma\right)\tanh\frac{\tilde{\epsilon}^u_{\mathbf{p}}(T/T_N,\gamma)}{2T} \tag{63}$$

with the dispersion relation

$$\tilde{\epsilon}^u_{\mathbf{p}}(T/T_N,\gamma)=J_u\Delta_u\phi_u(\mathbf{p},T/T_N,\gamma). \tag{64}$$

The structure factor $\phi_u(\mathbf{p},T/T_N,\gamma)$ renormalized by spin fluctuations can be expressed in terms of a structure factor $\varphi_u(\mathbf{p})$ like that in Eq. (33), where summation over nearest neighbors $\mathbf{l}$ is performed only in the basal plane ($\gamma<1$) or in the perpendicular direction ($\tilde{\gamma}=\gamma^{-1}<1$):

$$\phi_u\left(\mathbf{p},\frac{T}{T_N},\gamma\right)=\frac{1}{2}\varphi_u(\mathbf{p})+2T\sum_{\mathbf{q}}\varphi_u(\mathbf{p}-\mathbf{q})K_0(\mathbf{q}). \tag{65}$$

Index γ on the left-hand side of Eq. (65) is due to the anisotropic nature of correlator $K_0(\mathbf{q})$. Thus, the character of the transition to the spin-liquid state is determined by the degree of anisotropy: in the case of $\gamma<1$ spin-liquid correlations emerge first in the basal plane, and if $\gamma>1$ in the z-direction. At lower temperatures, the spin liquid naturally takes on three-dimensional properties, given that $\gamma\neq(0,\infty)$.

The transition temperature to the spin-liquid state, when spin fluctuations are taken into account, is determined by solving the equation

$$T_u^*=\frac{1}{2}\max\{J_\parallel,J_\perp\}\theta_u\left(\frac{T_u^*}{T_N},\gamma\right), \tag{66}$$

where

$$\theta_u\left(\frac{T_u^*}{T_N},\gamma\right)=(z_uN)^{-1}\sum_{\mathbf{p}}\phi_u^2\left(\mathbf{p},\frac{T}{T_N},\gamma\right), \tag{67}$$

$z_\parallel$ is the coordination number in the basal plane, and $z_\perp=2$.

In estimating the role of spin fluctuations for establishing the spin-liquid regime, it is convenient to introduce the temperature

$$T_u^{*(0)}=\frac{1}{2}\max\{J_\parallel,J_\perp\}\theta_u^{(0)} \tag{68}$$

of the transition to the RVB state in the anisotropic lattice without taking spin fluctuations into account. In this case,

$$\theta_u^{(0)}=(z_uN)^{-1}\sum_{\mathbf{p}}\varphi_u^2(\mathbf{p}). \tag{69}$$

Then the condition that the transition occurs by virtue of the spin-fluctuation mechanism is

$$\Upsilon_u(\gamma,T_u^{*(0)}/T_N)=\theta_u(T_u^{*(0)}/T_N,\gamma)/\theta_u^{(0)}>1. \tag{70}$$

The parameter $\Upsilon_u(\gamma)$ ($\Upsilon_u(\tilde{\gamma})$) for a simple cubic lattice is calculated in Appendix III. Critical values of the anisotropy parameters $\gamma_{1,2}$ at which the spin-liquid state stabilizes in almost one-dimensional and almost two-dimensional magnetic lattices are given for the case $T_u^{*(0)}/T_N=1$ in Fig. 9 for different values of τ. It is clear that only in a strongly anisotropic situation, almost one- or two-dimensional (see Eqs. (A.III.7) and (A.III.8)) spin correlations help formation of the spin liquid, and in the anisotropic case, inclusion of antiferromagnetic fluctuations in the mean-field approximation leads to suppression of the spin-liquid phase.

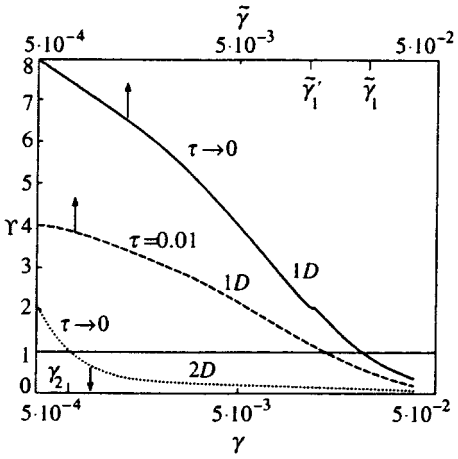


FIG. 9. Parameter Υ describing the effect of critical spin fluctuations on the transition temperature to the RVB phase for the quasi-one-dimensional ($1D$) and quasi-two-dimensional ($2D$) Kondo lattices. Parameter τ characterizes the proximity to the antiferromagnetic instability. The RVB state emerges at $\gamma<\gamma_1$ and $\tilde{\gamma}<\tilde{\gamma}_2$ in the cases of axial and plane magnetic anisotropy, respectively.

The analysis in this section once again indicates that the mean-field approximation is insufficient for the description of the spin liquid. In particular, even the diagrams of Fig. 8 indicate that the static approximation, generally speaking, does not apply to the critical region, since antiferromagnetic fluctuations define their own time and energy scales, which determine the real character of transition from the paramagnetic state to the spin-liquid state.

5. CRITICAL ANTIFERROMAGNETIC FLUCTUATIONS AND SPIN DIFFUSION

As mentioned in Sec. 4, in antiferromagnets critical fluctuations have differing properties in the long-wave ($k\to0$) and short-wave ($k\to Q$) regions, and the spin response function in these regions takes the form of Eqs. (56) and (58), respectively. The critical dynamics of antiferromagnets is usually analyzed using renormalization-group techniques applied to phenomenological models.[49,50] Chubukov[48] calculated the dynamic susceptibility of a two-dimensional antiferromagnet in the diffusion and relaxation regions using a diagram technique in the Schwinger boson representation. We investigate the dynamic susceptibility as a function of frequency and momentum in the thee-dimensional configuration using the pseudofermion technique.

In order to calculate the spin diffusion factor D and the relaxation constant Γ, we need to know, in addition to the spin correlators defined by the Larkin equation (19), the low-frequency behavior of the current correlator:

$$K_{\dot{S}\dot{S}}^{\alpha\beta}(\mathbf{k},\tau)=\delta_{\alpha\beta}\sum_{\mathbf{k}_1\mathbf{k}_2}V(\mathbf{k},\mathbf{p}_1)V(-\mathbf{k},-\mathbf{p}_2)\times\langle T_\tau(S^\mu_{\mathbf{p}_1+\mathbf{k}/2}S^\rho_{-\mathbf{p}_1+\mathbf{k}/2})_\tau(S^\mu_{-\mathbf{p}_2-\mathbf{k}/2}S^\rho_{\mathbf{p}_2-\mathbf{k}/2})_0\rangle, \tag{71}$$

where

$$V(\mathbf{k},\mathbf{p})=J(\mathbf{k}+\mathbf{p}/2)-J(-\mathbf{k}+\mathbf{p}/2).$$

FIG. 10. Self-energy part of the Green's function $\mathscr{G}_{\mathbf{ii}}$ including the contribution of critical fluctuations in the Born approximation.

There is an exact solution for the Fourier transform of correlator $K_{\dot{S}\dot{S}}$ continued to the upper half-plane expressed in terms of irreducible (noncuttable along the interaction line) self-energy parts of the spin and current correlation functions:

$$K^R_{\dot{S}\dot{S}}(\omega)=\Sigma^R_{\dot{S}\dot{S}}+\omega^2\frac{\Sigma^R_{SS}\mathscr{V}\Sigma^R_{SS}}{1-\Sigma^R_{SS}\mathscr{V}}. \tag{72}$$

Here $\mathscr{V}=(\Sigma^R)^{-1}-(K_0)^{-1}$ is the vertex part determined by the static response in the critical region.[19,51]

Using the Kramers–Kronig dispersion relations for retarded and advanced correlation functions, and the analytic properties of irreducible self-energy parts, one can derive from Eqs. (19) and (72) the expression

$$K^R_{SS}(\omega)=K_0\frac{\Gamma_{\mathbf{k},\omega}}{-i\omega+\Gamma_{\mathbf{k},\omega}}, \tag{73}$$

which holds as both $\mathbf{k}\to0$ and $\mathbf{k}\to\mathbf{Q}$.

The spin correlation functions can be expressed in terms of the pseudofermion Green's functions. For example, the expression for the one-site susceptibility has the form

$$\mathscr{K}^\perp_{\mathbf{i}}(\varepsilon_m)=T\sum_m\mathscr{G}_{\mathbf{ii}}(\omega_n+\varepsilon_m)\mathscr{G}_{\mathbf{ii}}(\omega_n), \tag{74}$$

(see Eq. (16)). Here $\mathscr{G}_{\mathbf{ii}}(\omega_n)$ is a Fourier component of the pseudofermion Green's function $\mathscr{G}_{\mathbf{ii}}(\tau)=\langle T_\tau f_{\mathbf{i}}(\tau)f^\dagger_{\mathbf{i}}(0)\rangle$. Since nonphysical states do not appear when calculating single-site averages for $S=1/2$, there is no need to introduce projection operators. As $T\to T_N$, scattering by the relaxation mode contributes a component described by the diagram in Fig. 10 to the self-energy part of the Green's function $\Sigma(\omega_n)$. Unlike the diagram of Fig. 8, here solid lines correspond to one-site propagators $\mathscr{G}_{\mathbf{ii}}$, and points to exchange vertices $\tilde{J}(\mathbf{q})$. The wavy line in this diagram corresponds to the spin Green's function (16) determined by the Larkin equation (19). In the absence of spin-liquid correlations, let us substitute into the self-energy part $\Sigma(\omega_n)$ of the pseudofermion Green's function the ''bare'' function $g_{i\sigma}$ from Eq. (26) and a spin function $\mathscr{K}(\varepsilon_m,\mathbf{q})$ in the form of a relaxator:

$$\Sigma(\omega_n)=\tilde{J}^2T\sum_m N^{-1}\sum_{\mathbf{q}}\varphi(\mathbf{q})^2\frac{1}{i(\varepsilon_m-\omega_n)}\frac{\Gamma\chi_0(T)}{|\varepsilon_m|+b(q)}, \tag{75}$$

where $b(q)=\Gamma[\tau+(ql_0)^2]$, and Γ should be calculated independently using the Dyson and Larkin equations. By calculating the sum over frequencies in Eq. (75) and continuing it analytically to the complex z plane, we obtain the following equation for poles of the pseudofermion Green's function:

$$z-\Sigma(z)=0,$$

$$\Sigma(z)=\sum_{\mathbf{q}}\ \frac{J^2\varphi^2(\mathbf{q})\mathscr{A}}{\pi}\frac{z}{z^2+b_{\mathbf{q}}^2}\left[\psi\left(-\frac{iz}{2\pi T}\right)-\psi\left(\frac{b_{\mathbf{q}}}{2\pi T}\right)-\frac{\pi T}{b_{\mathbf{q}}}+\frac{i\pi T}{z}\right], \tag{76}$$

where $\mathscr{A}=\Gamma\chi_0(T)$, and $\psi(y)$ is the digamma function. Hence, it is clear that the pseudofermion Green's function in this approximation is $\mathscr{G}^R_{\mathbf{ii}}(\omega)\sim[\omega+i\Gamma(T)]^{-1}$. By substituting this into Eq. (74), we find the one-site susceptibility

$$K^R_{\mathbf{i}}=\frac{\overline{\chi_0}}{1-i\omega/\Gamma}, \tag{77}$$

which is, in turn, can be substituted into the Larkin equation (which also includes, generally speaking, vertex corrections[48]), and thus the equation system for Γ and I_0 is closed.

The spin-liquid effects on the behavior of the spin correlation functions in the critical region can be accounted for by introducing anomalous intersite contributions into $\Sigma(\omega)$ (Fig. 10). Nonlocal fermion correlations lead to emergence of a new characteristic length characterizing short-range order, and change the temperature dependence of the static spin susceptibility and dynamic response functions. As a result, we have changes in the scaling behavior and in the frequency and momentum dependence of the spin susceptibility.

The spin diffusion factor is also determined by the self-energy part of the current correlator:[51]

$$D=\lim_{k\to 0,\omega\to 0}\frac{1}{k^2}\frac{\mathrm{Im}\Sigma^R_{\dot{S}\dot{S}}(\mathbf{k},\omega)}{\omega}K_0^{-1}(\mathbf{k}). \tag{78}$$

Since the behavior of the current correlator is fully determined by relaxation processes, effects of nonlocal spin correlations should also change scaling characteristics of the spin susceptibility in the hydrodynamic region.

The calculations described in this section are not considered a complete description of critical phenomena in antiferromagnets. These are instead illustrations given with the following aims: first, to demonstrate applicability of the suggested diagram technique to traditional problems of the theory of magnetic phase transition and, second, to outline feasible methods for taking into account the effect of spin-liquid correlations on antiferromagnetic fluctuations in the critical region.

6. CONCLUSIONS

In this paper, we have demonstrated that the spin-liquid state in the Kondo lattice can be more stable than the Kondo singlet state, owing to the same processes as those responsible for Kondo screening in the case of sufficiently strong antiferromagnetic sf-exchange. This rather paradoxical result can be explained by the fact that strong competition between Kondo scattering and spin-liquid correlations occurs at temperatures near the Néel point. Since all correlation effects at such temperatures have the same order of magnitude, the simple mean-field approximation cannot be used in describing the spin subsystem in a three-dimensional Kondo lattice.

The projection diagram technique suggested in the paper and based on the similarity between the Hubbard Hamiltonian for electrons and Heisenberg Hamiltonian for pseudofermions allows one, in principle, to go beyond the standard mean-field model of the homogeneous RVB phase.[6,16] Attempts to include antiferromagnetic fluctuations in the mean-field approximation (Sec. 4) do not produce any trustworthy results. Preliminary analysis, however, indicates[52] that the diagram technique suggested in the paper may allow one to manage without the mean-field approximation in describing effects which occur in the region of critical antiferromagnetic fluctuations and devise a more realistic scenario of emergence of the spin liquid in the Kondo lattice.

The investigation of spin diffusion near the Néel point reported in Section 5 indicates that the diagram techniques used in describing critical antiferromagnetic correlations at high temperatures may yield new physical results in the hydrodynamic region.

The authors are indebted to Yu. Kagan, N. V. Prokof'ev, G. G. Khaliullin, D. E. Khmel'nitskiĭ, and D. I. Khomskiĭ for helpful discussions and critical remarks. This work was supported by the Russian Fund for Fundamental Research (Project 95-02-04250a), INTAS (Project 93-2834), and Netherlands Organization for Support of Scientific Research (NWO, Project 07-30-002).

APPENDIX I

In calculating the polarization operator $\Pi(\mathbf{R})$ (Eq. (40)), we use the asymptotic form of the Green's function (41). Substituting it into Eq. (42), we obtain the expression for a spherical Fermi surface:

$$\Pi(R,\varepsilon_m)=T\left(\frac{m}{2\pi R}\right)^2\exp\left(-\frac{2|\varepsilon_m|}{v}R\right)\times\frac{\cos(2p_FR+i\varepsilon_mR/v)}{\sinh(2\pi TR/v)}+T\left(\frac{m}{2\pi R}\right)^2\times\exp\left(-\frac{|\varepsilon_m|}{v}R\right)\left[\frac{|\varepsilon_m|}{2\pi T}+\frac{\sinh(|\varepsilon_m|R/v)}{\sinh(2\pi TR/v)}\times\exp\left(-\frac{|\varepsilon_m|}{v}R+2ip_FR\ \mathrm{sign}\ \varepsilon_m\right)\right]. \tag{AI.1}$$

In the static limit, it reduces to

$$\Pi(R,0)=T\left(\frac{m}{2\pi R}\right)^2\frac{\cos(2p_FR)}{\sinh(2\pi TR/v)}=\frac{mp_F^4}{8\pi^3}\frac{\cos(2p_FR)}{(p_FR)^3}\left[1-\frac{\pi^2}{6}\left(\frac{T}{\varepsilon_F}p_FR\right)^2+\dots\right], \tag{AI.2}$$

hence we have Eq. (45) at $T=0$.

In the case of a quasi-two-dimensional cylindrical Fermi surface

$$g(\mathbf{R},z,\omega_n)=\int\frac{d^3p}{(2\pi)^3}\frac{1}{i\omega_n-\xi(\mathbf{p})}\exp(i\mathbf{p}\cdot\mathbf{R}+ip_zz)$$

$$= \int_{-p_{z0}}^{p_{z0}} \frac{dp_z}{2\pi} e^{ip_z z} \int \frac{pdpd\varphi}{(2\pi)^2} \frac{1}{i\omega_n - \xi(\mathbf{p})} e^{i\mathbf{p}\cdot\mathbf{R}},$$

$$g(\mathbf{R},z,\omega_n) = \frac{\sin(p_{z0}z)}{\pi z} G(\mathbf{R},\omega_n). \tag{AI.3}$$

For $p_{z0} \gg p_F$ the effective RKKY interaction is independent of p_{z0},

$$J_{\mathrm{RKKY}}(\mathbf{R}) = \left(\frac{J}{\widetilde{n}_0}\right)^2 \Pi(\mathbf{R},0) = \left(\frac{J}{\widetilde{n}_0}\right)^2 \int \frac{d\omega}{2\pi} g^2(\mathbf{R},\omega)$$

$$= \left(\frac{J}{n_0}\right)^2 \int \frac{d\omega}{2\pi} G^2(\mathbf{R},\omega). \tag{AI.4}$$

Here $\widetilde{n}_0 = 4\pi p_F^2 p_{z0}/(2\pi)^3 = p_{z0}p_F^2/2\pi^2 = p_{z0}n_0/\pi$, $n_0 = p_F^2/2\pi$ is the two-dimensional density of electronic states, and $G(R,\omega_n)$ is the two-dimensional Green's function

$$G(R,\omega_n) = \int \frac{pdpd\varphi}{(2\pi)^2} \frac{1}{i\omega_n - \xi(\mathbf{p})} \exp(ipR\cos\varphi). \tag{AI.5}$$

Let us use the integral representation of the Bessel function

$$J_0(z) = \frac{1}{2\pi} \int_0^{2\pi} d\varphi \exp(iz\cos\varphi) \tag{AI.6}$$

in the asymptotic limit for large $|z|$:

$$J_0(z) \approx \sqrt{\frac{2}{z\pi}} \cos\left(z - \frac{\pi}{4}\right). \tag{AI.7}$$

Then we have

$$G(R,\omega_n) = -i\,\mathrm{sign}\,\omega_n \frac{m}{\sqrt{2\pi p_F R}} \exp\left(-\frac{|\omega_n|}{2\varepsilon_F} p_F R + i\left(p_F R - \frac{\pi}{4}\right)\mathrm{sign}\,\omega_n\right). \tag{AI.8}$$

Substituting this expression into Eq. (AI.4), we obtain

$$\Pi(R,\varepsilon_m) = -T\frac{m^2}{2\pi p_F R} \exp\left(-\frac{2|\varepsilon_m|}{v} R\right) \times \frac{\sin(2p_F R + i\varepsilon_m R/v)}{\sinh(2\pi TR/v)} - T\frac{m^2}{2\pi p_F R} \times \exp\left(-\frac{|\varepsilon_m|}{v} R\right) \left\{ \frac{|\varepsilon_m|}{2\pi T} - \frac{\sinh(|\varepsilon_m|R/v)}{\sinh(2\pi TR/v)} \exp\left[-\frac{|\varepsilon_m|}{v} R + 2i \times \left(p_F R - \frac{\pi}{4}\right)\mathrm{sign}\,\varepsilon_m\right]\right\}. \tag{AI.9}$$

In the low-temperature limit this expression becomes

$$\Pi(R,0) = -T\frac{m^2}{2\pi p_F R} \frac{\sin(2p_F R)}{\sinh(2\pi TR/v)}$$

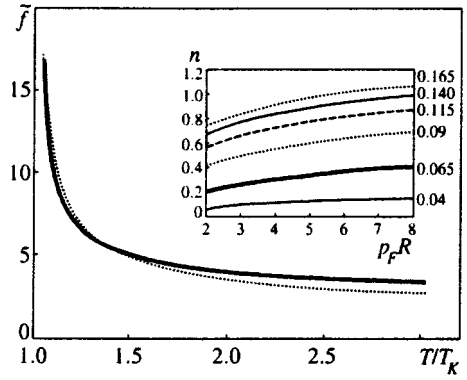


FIG. 11. Numerical values of integral $\widetilde{f}(p_F R)$ (solid line) and of the approximating function $f(p_F R)$ (see text).

$$= -\frac{mp_F^2}{4\pi^2} \frac{\sin(2p_F R)}{(p_F R)^2} \left[1 - \frac{\pi^2}{6}\left(\frac{T}{\varepsilon_F} p_F R\right)^2 + \ldots\right], \tag{AI.10}$$

and at $T=0$ to Eq. (47).

APPENDIX II

Intensification of the RKKY interaction due to Kondo renormalization of the single-site sf-exchange vertex, which is taken into account in the logarithmic approximation,

$$\Gamma(\varepsilon,\alpha) = \frac{1}{(1 + 2\alpha\,\ln(\varepsilon/\varepsilon_F))^2}, \tag{AII.1}$$

is described by the expression

$$f\left(p_F R, \alpha, \frac{T}{\varepsilon_F}\right) = \int_{T/\varepsilon_F}^{\infty} \frac{\exp(-p_F Rx)dx}{(1 + 2\alpha\,\ln(x))^2}. \tag{AII.2}$$

The temperature dependence of this integral is determined by both Doniach's parameter α and the separation between neighboring Kondo centers (parameter $p_F R$).

If we neglect logarithmic renormalization (AII.1), the integral in (A.II.2) equals $1/p_F R$ for $T \ll \varepsilon_F/k_B$, and the integral in (44) reduces to the usual RKKY formula (45). When Kondo processes are taken into account, the function f defined by Eq. (A.II.2) can be approximated in the temperature range of interest, $[T_K, 3T_K]$, by the expression

$$f\left(p_F R, \alpha, \frac{T}{\varepsilon_F}\right) \approx \frac{1}{p_F R} \frac{1}{(1 - 2\alpha\,\ln(T/\varepsilon_F))^{n(p_F R,\alpha)}}, \tag{AII.3}$$

where the exponent $n = n(p_F R, \alpha)$ is independent of temperature. As a result, the high-temperature behavior of the RKKY interaction is determined by the function $\widetilde{f}(p_F R, \alpha, T/\varepsilon_F) = f(p_F R, \alpha, T/\varepsilon_F) p_F R$, which can be approximated as

$$\widetilde{f}\left(p_F R, \alpha, \frac{T}{\varepsilon_F}\right) \approx \frac{1}{(1 + 2\alpha\,\ln(T/\varepsilon_F))^{n(p_F R,\alpha)}}. \tag{AII.4}$$

Figure 11 shows the temperature dependence of the exact function $\widetilde{f}(p_F R = 5.0, \alpha = 0.09)$ calculated numerically (solid line) and the approximate function $\widetilde{f}(p_F R = 0.5, \alpha = 0.09)$ (dotted line) in the temperature range

$T_K<T<3T_K$. The exponent $n=n(p_FR,\alpha)$ of approximate function (AII.4) is shown in the insert as a function of p_FR in the range $2<p_FR<8$ for several values of α in the interval $0.04<\alpha<0.165$. The exponent was determined using the least-square fit in the temperature range $1.2T_K<T<3T_K$.

APPENDIX III

In this Appendix, we calculate the parameter Υ defined by Eq. (70), which characterizes the effect of spin correlations on the transition temperature to the RVB phase for a simple cubic lattice with anisotropic RKKY interaction due, for example, to a nonspherical Fermi surface. Let us introduce $J_\parallel \equiv J_x = J_y$ and $J_\perp \equiv J_z$. Then we must substitute into Eq. (57) for the spin correlator $K_0(\mathbf{q},\gamma)$ the parameter

$$j_q \equiv J_q/|J_Q| = -j_\parallel(\varphi_\parallel + \gamma\varphi_\perp), \tag{AIII.1}$$

where $\varphi_\parallel(q)=2(\cos q_x+\cos q_y)$, $\varphi_\perp(q)=2\cos q_z$, and $j_\parallel = J_\parallel/J_{\mathbf{Q}}$ $(a=1)$. To calculate sums in Eq. (65) like

$$T\sum_{\mathbf{q}} \varphi_u(\mathbf{p}-\mathbf{q})K_0(\mathbf{q},\gamma) = \frac{S(S+1)T}{6T_N j_0}\sum_{\mathbf{q}} \frac{\varphi_u(\mathbf{p}-\mathbf{q})}{T/T_N j_0 - j_{\mathbf{q}}/j_0}, \tag{AIII.2}$$

we use the integral representation for the spin correlator

$$K_0(\mathbf{q},\gamma) = \frac{S(S+1)j_{\mathbf{q}}}{6T_N j_0}\int_0^\infty dt\ \exp\left\{-\left(\frac{T}{T_N j_0}-\frac{j_q}{j_0}\right)t\right\}. \tag{AIII.3}$$

When the interaction in the basal plane is dominant $(\gamma<1)$, the spectrum of spin-liquid excitations has the form

$$\tilde{\epsilon}^{\parallel}_{\mathbf{p}}(T/T_N,\gamma) = \frac{1}{2}J_\parallel\Delta_\parallel\left[1-(2+\gamma)\frac{T}{T_N}A(\gamma,T/T_N)\right]\varphi_\parallel(q), \tag{AIII.4}$$

where the function $A(\gamma,T/T_N)$ can be expressed in terms of integrals of Bessel functions:

$$A(\gamma,\tau) = \int_0^\infty dt\ \exp\{-(2+\gamma)(1+\tau)t\}I_1(t)I_0(t)I_0(\gamma t). \tag{AIII.5}$$

Given that $\theta_\parallel^{(0)}=\theta_\perp^{(0)}=1$ for the simple cubic lattice, we obtain

$$\Upsilon_\parallel(\gamma,T_u^{*(0)}/T_N) = [1-(2+\gamma)(1+\tau)A(\gamma,\tau)]^2/4. \tag{AIII.6}$$

When the interaction perpendicular to the basal plane is dominant $(\tilde{\gamma}<1)$, we have instead of Eqs. (A.III.4)–(A.III.6)

$$\tilde{\epsilon}^{\perp}_{\mathbf{p}}\left(\frac{T}{T_N},\gamma\right) = J_\perp\Delta_\perp\left[1-(1+2\tilde{\gamma})\frac{T}{T_N}\tilde{A}\left(\gamma,\frac{T}{T_N}\right)\right]\cos p_z, \tag{AIII.4$'$}$$

$$\tilde{A}(\tilde{\gamma},\tau) = \int_0^\infty dt\ \exp\{-(1+2\tilde{\gamma})(1+\tau)t\}I_1(t)I_0^2(\tilde{\gamma}t), \tag{AIII.5$'$}$$

$$\Upsilon_\perp(\tilde{\gamma},T_u^{*(0)}/T_N) = [1-(1+2\tilde{\gamma})(1+\tau)\tilde{A}(\tilde{\gamma},\tau)]^2/4. \tag{AIII.6$'$}$$

Using asymptotic expressions for the integrals:

$$A(\gamma,\tau)|_{\gamma,\tau\to 0} \propto -\ln\max(\gamma,\tau),$$

$$\tilde{A}(\tilde{\gamma},\tau)|_{\tilde{\gamma},\tau\to 0} \propto [\max(\tilde{\gamma},\tau)]^{-1/2},$$

we obtain for the neighborhood of T_N in the case of strong anisotropy

$$\Upsilon_\parallel(\gamma,T_u^{*(0)}/T_N) \propto -\ln\max(\gamma,\tau), \tag{AIII.7}$$

$$\Upsilon_\perp(\tilde{\gamma},T_u^{*(0)}/T_N) \propto [\max(\tilde{\gamma},\tau)]^{-1/2}, \tag{AIII.8}$$

and as a result, strong spin fluctuations stabilize the spin liquid.

[1)] Preliminary results of this study were given in the short note.

[2)] A procedure similar to that suggested below was described in Ref. 33 in the cases of the Anderson impurity and Anderson lattice. But since the Anderson Hamiltonian, unlike spin Hamiltonians (1) and (2), does not have local SU(2) symmetry, and the requirement of exact particle-hole symmetry is not imposed, there are many differences between formulations of rules of the diagram techniques.

[3)] Since in the case $S=1/2$ for one-site processes the constraint condition is satisfied automatically,[10] it is unnecessary to introduce projection operators.

[1] G. Zwicknagl, Adv. Phys. **41**, 203 (1992).
[2] A. C. Hewson, *The Kondo Problem to Heavy Fermions*, Cambridge University Press, Cambridge (1993).
[3] S. Doniach, Physica B **91**, 231 (1977).
[4] F. J. Ohkawa, Prog. Theor. Phys. Suppl. No. 108, 209 (1992).
[5] Y. Kuramoto and K. Miyake, Prog. Theor. Phys. Suppl. No. 108, 199 (1992).
[6] P. Coleman and N. Andrei, J. Phys.: Condens. Matter **1**, 4057 (1989).
[7] J. A. Millis and P. A. Lee, Phys. Rev. B **35**, 3394 (1987).
[8] Yu. Kagan, K. A. Kikoin, and N. V. Prokof'ev, Physica B **182**, 201 (1992).
[9] J. Gan, P. Coleman, and N. Andrei, Phys. Lett. **68**, 3476 (1992).
[10] A. A. Abrikosov, Physics **2**, 21 (1965).
[11] I. Affleck and J. B. Marston, Phys. Rev. B **37**, 3774 (1988).
[12] I. Affleck, Z. Zou, T. Hsu, and P. W. Anderson, Phys. Rev. B **38**, 745 (1988).
[13] L. B. Ioffe and A. I. Larkin, Phys. Rev. B **39**, 8988 (1989).
[14] P. A. Lee and N. Nagaosa, Phys. Rev. B **46**, 5621 (1992).
[15] S. Elitzur, Phys. Rev. D **12**, 3978 (1975).
[16] G. Baskaran, Z. Zou, and P. W. Anderson, Solid State Commun. **63**, 973 (1987).
[17] A. Ruckenstein, P. Hirschfeld, and J. Appel, Phys. Rev. B **36**, 857 (1987).
[18] K. A. Kikoin, M. N. Kiselev, and A. S. Mishchenko, JETP Lett. **60**, 358 (1994).
[19] Yu. A. Izyumov and Yu. N. Skryabin, *Statistical Mechanics of Magnetically Ordered Systems* [in Russian], Nauka, Moscow (1987).
[20] V. G. Bar'yakhtar, V. N. Krivoruchko, and D. A. Yablonskii, *Green's Functions in the Theory of Magnetism* [in Russian], Naukova Dumka, Kiev (1984).
[21] H. Keiter and G. Morandi, Phys. Rep. **109**, 227 (1984).
[22] F. Onufrieva and J. Rossat-Mignod, Phys. Rev. B **52**, 7572 (1995).
[23] D. C. Mattis, *The Theory of Magnetism*, Harper and Row, New York (1965).
[24] W. W. Lewis and R. B. Stinchcombe, Proc. Phys. Soc. London **92**, 1002 (1967).
[25] S. E. Barnes, J. Phys. F **6**, 1375 (1976).
[26] G. Kotliar and A. E. Ruckenstein, Phys. Rev. Lett. **57**, 1362 (1986).
[27] P. Coleman, E. Miranda, and A. M. Tsvelik, Phys. Rev. Lett. **70**, 2960 (1993).
[28] X.-G. Wen and P. A. Lee, Phys. Rev. Lett. **76**, 503 (1996).
[29] Y. Ono, T. Matsuura, and Y. Kuroda, Physica C **159**, 878 (1989).
[30] J. Hubbard, Proc. R. Soc. London, Ser. A **285**, 542 (1965).

[31] R. O. Zaitsev, Zh. Éksp Teor. Fiz. **70**, 1100 (1976) [Sov. Phys. JETP **43**, 574 (1976)].
[32] T. Yanagisawa, Phys. Rev. B **40**, 6666 (1989).
[33] J. Brinckmann, Europhys. Lett. **28**, 187 (1994).
[34] A. I. Larkin, Zh. Éksp. Teor. Fiz. **37**, 264 (1960) [Sov. Phys. JETP **10**, 186 (1960)].
[35] K. Kuboku, J. Phys. Soc. Jpn. **62**, 420 (1993).
[36] C. Mudry and E. Fradkin, Phys. Rev. B **49**, 5200 (1994).
[37] D. R. Grempel and M. Lavagna, Solid State Commun. **83**, 595 (1992).
[38] T. Tanamoto, H. Kohno, and H. Fukuyama, J. Phys. Soc. Jpn. **62**, 617 (1993).
[39] I. S. Sandalov and M. Richter, Phys. Rev. B **50**, 12855 (1994).
[40] P. W. Anderson, Mater. Res. Bull. **8**, 153 (1973).
[41] C. M. Varma, in *Theory of Heavy Fermions and Valence Fluctuations*, Springer Series in Solid-State Sciences, T. Kasuya and T. Saso (eds.), Vol. 62, Springer-Verlag, Berlin (1985), p. 277.
[42] K. Yosida and A. Okiji, Prog. Theor. Phys. **34**, 505 (1965).
[43] A. A. Abrikosov and A. A. Migdal, J. Low Temp. Phys. **3**, 519 (1970).
[44] A. M. Tsvelik and P. B. Wiegmann, Adv. Phys. **32**, 453 (1983).
[45] K. Miura, T. Ono, and K. Kuboku, Physica C **179**, 411 (1991).
[46] F. Nozières, Ann. de Phys. **10**, 1 (1985).
[47] K. A. Kikoin, J. Phys.: Condens. Matter **8**, 3601 (1996).
[48] A. Chubukov, Phys. Rev. B **44**, 392 (1991).
[49] B. I. Halperin and P. C. Hohenberg, Rev. Mod. Phys. **49**, 435 (1977).
[50] S. Chakraverty, B. I. Halperin, and D. Nelson, Phys. Rev. B **39**, 2344 (1989).
[51] S. V. Maleyev, Sov. Sci. Rev. A, I. M. Khalatnikov (ed.), Harwood Press, New York (1987), Vol. 8, p. 323.
[52] K. A. Kikoin, M. N. Kiselev, and A. S. Mishchenko, E-prints archive, Cond-Matt/96 08 121.

Translation was provided by the Russian Editorial office.

Magnetic properties of the new organic superconductor λ-$(BETS)_2GaCl_4$

A. É. Primenko and V. D. Kuznetsov

L. D. Mendeleev Chemical-Technological University, 125820 Moscow, Russia

N. D. Kushch and É. B. Yakubskiĭ

N. N. Semyonov Institute of Chemical Technology, Russian Academy of Sciences, 117977 Moscow, Russia
(Submitted 13 February 1997)
Zh. Éksp. Teor. Fiz. **112**, 760–762 (August 1997)

This papr discusses the results of the first investigation of the magnetic properties of the organic superconductor λ-$(BETS)_2GaCl_4$. It is shown that the transition to the superconducting state begins at $T_c \approx 7$ K, which is considerably lower than the value $T_c \approx 10$ K determined from resistive measurements. The estimated value of the critical current density turns out to be two orders of magnitude lower than in superconductors of the family κ-$(ET)_2X$. © *1997 American Institute of Physics.* [S1063-7761(97)02608-5]

1. INTRODUCTION

The current interest in investigating organic superconductors based on cation-radical salts remains strong, due both to the unusual physical properties of these quasi-two-dimensional systems and the fact that the study of quasi-two-dimensional systems is an important route to understanding the physical properties of high-temperature superconductors.

Although the recently synthesized organic superconductor λ-$(BETS)_2GaCl_4$ has been studied intensely, its magnetic properties have not yet been investigated. Hence, measurements of the temperature and field dependence of the magnetic moment of single crystals of this superconductor are of particular interest.

2. EXPERIMENTAL METHODS

Single crystals of the organic superconductor were synthesized by the standard electrochemical techniques described in Ref. 1. The single crystals consisted of thin films with typical dimensions $0.1\times1.0\times0.04$ mm^3 (crystal 1) and $0.7\times1.3\times0.04$ mm^3 (crystal 2). Crystal 1 consisted of a single perfect crystal, while crystal 2 consisted of several single crystals grown together along the plane AC (which we will refer to as multicrystals). This gives us the flexibility of comparing measurements made on a multicrystal and on a single crystal, thereby allowing us to estimate the influence of growth boundaries on the magnetic properties. In light of the extreme brittleness of λ-$(BETS)_2GaCl_4$ single crystals, we used slow cooling from room temperature and fastened the crystals by means of Ramsay paste. The crystals were placed in quartz ampules, which gave a negligibly small contribution to the measured response, and the ampules were then evacuated and sealed off. Then an atmosphere of helium was generated within the ampules by diffusion through the walls; this atmosphere was necessary for heat exchange. In order to measure the magnetic moments we used a SQUID magnetometer.[2] The temperature was stabilized to an accuracy of 0.03 K, and the errors in measuring the magnetic moment were $\approx 4\times10^{-11}$ A·m^2. Measurements of the temperature dependence of the magnetic moment were made in two regimes: cooling in zero field down to a minimum temperature ($T=2$ K) with subsequent application of the field (zero-field cooling) and cooling in a nonzero field (field cooling).

3. RESULTS AND DISCUSSION

The results of the measurements are shown in Figs. 1 and 2. The thermal variation of the magnetic moment obtained in the zero-field cooling and field cooling regimes was similar for both crystals. It is interesting to note that these superconductors have very wide transitions in a magnetic field. For other cation-radical salts this transition is considerably narrower and in the fields used ($B_a \sim 1$–2 mT) is roughly 1–2 K. The temperature at which the superconducting transition begins for this superconductor, determined based on resistive measurements,[3] is ≈11 K, whereas based on the magnetic measurements the transition to the normal state is practically complete after ≈7 K (in a field of 2 mT).

The field dependence of the magnetic moment measured at $T=4.2$ K is practically reversible and similar for both

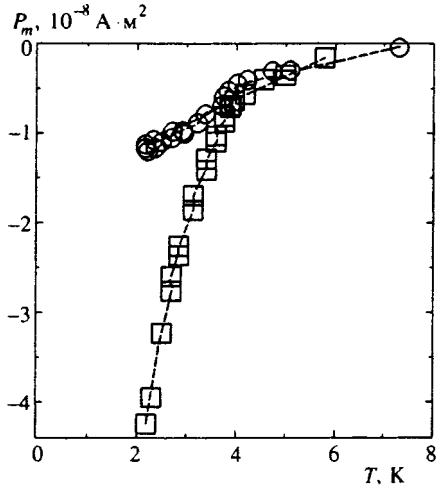


FIG. 1. Temperature dependence of the magnetic moment for crystal 2: in the field-cooled regime (○) and the zero-field-cooled regime (□); $B_a=2$ mT.

 1063-7761/97/080415-02$10.00

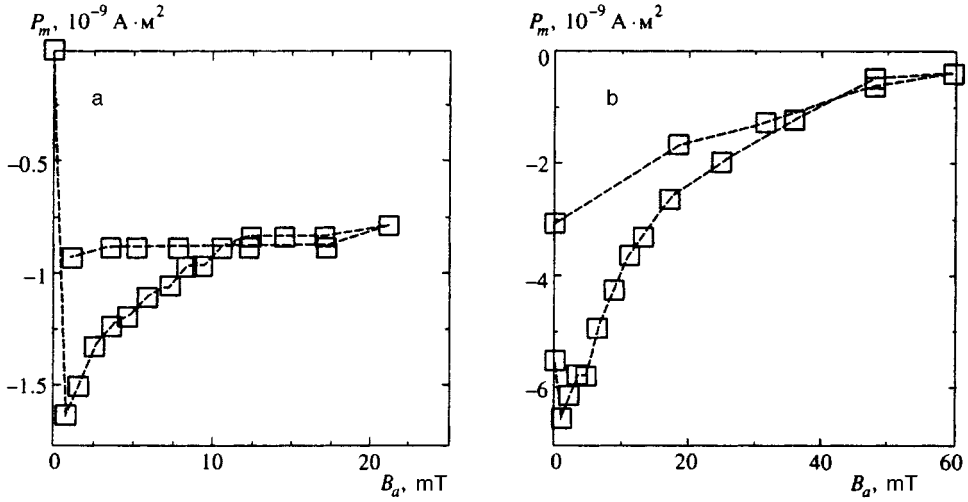


FIG. 2. Magnetization curves for crystals 1 (a) and 2 (b) at $T=4.2$ K.

crystals. Analogous hysteresis curves are observed in Ref. 4 for Y-Ba-Cu-O at temperatures close to T_c (in fact, $T/T_c \approx 0.96$). This unusual form for the magnetization curves is related to the fact that the reversible contribution to the magnetic moment significantly exceeds the irreversible contribution (normally this situation is opposite). The value of the critical current density estimated within the Bean model from the width of the hysteresis loop comes to ≈ 10 A/cm^2. This is two orders of magnitude smaller than in superconductors of the κ-$(ET)_2X$ family. This probably is connected with the large degree of two-dimensionality of the superconductors compared with the κ-$(ET)_2X$.

Because the magnetization curves have practically no linear segment (and also due to the large demagnetizing factor $N \approx 0.9$ in the crystals we used), we were unable to estimate the first critical field. We can only confirm that the field at which penetration begins was at most 0.1 mT for the single crystals.

4. CONCLUSIONS

We have carried out the first studies of the magnetic properties of a new organic superconductor — λ-$(BETS)_2GaCl_4$. We have shown that this superconductor has a very broad phase transition in a magnetic field, which can be related to the large anisotropy of this superconductor. For representatives of the family κ-$(ET)_2X$ this transition is significantly narrower and the transition temperature T_c determined from magnetic measurements is in good agreement with values obtained from resistive measurements.

Measurements on a multicrystal and a single crystal give practically the same results, which may indicate the small influence of intercrystalline boundaries on the magnetic behavior of this system.

The magnetization curves have an almost reversible character with very small hysteresis, corresponding to a critical current density $j_c \approx 10$ A/cm^2 that is two orders of magnitude smaller than for superconductors in the κ-$(ET)_2X$ family.

Overall, we can associate the specific features of the magnetic properties of λ-$(BETS)_2GaCl_4$ compared to κ-$(ET)_2X$ with the more anisotropic crystal structure of the former, which leads to lower stability of the superconducting state compared to its less anisotropic analogs.

The authors take pleasure in acknowledging their gratitude to I. F. Shchegolov for a number of critical remarks regarding the results of this work.

[1] L. K. Montgomery, T. Burgi, and J. C. Huffmann, Physica C **19**, 490 (1994).
[2] V. D. Kuznetsov, Prikl. Tekh. Eksp. **4**, 196 (1985) [in Russian].
[3] H. D. Kush, private communication.
[4] M. Oussena and S. Senoussi, Europhys. Lett. **4**, 625 (1987); S. Senoussi, J. de Phys. III **2**, 1041 (1992).

Translated by Frank J. Crowne