Splitting of a high-energy photon in a strong Coulomb field

R. N. Li, A. I. Mil'shteĭn, and V. M. Strakhovenko

G. I. Budker Nuclear Physics Institute, Russian Academy of Sciences, Siberian Branch, 630090 Novosibirsk, Russia

(Submitted 15 April 1997)
Zh. Éksp. Teor. Fiz. **112**, 1921–1940 (December 1997)

The helical splitting amplitudes of a high-energy photon in an external Coulomb field are found exactly with respect to $Z\alpha$. Both screened and unscreened potentials are investigated. The treatment is carried out within the semiclassical approach, which is valid for small angles between the momenta of all photons. A new representation is used for the semiclassical electronic Green's function. The resulting expressions are analyzed in detail for the case of transverse components of the momenta of the final photons that are large compared with the electron mass. © *1997 American Institute of Physics.* [S1063-7761(97)00112-1]

1. INTRODUCTION

The generation of a virtual electron–positron pair is known to lead to the appearance of such nonlinear quantum electrodynamic effects as Delbrück scattering (coherent photon scattering[1]) and the splitting of a photon into two in an external Coulomb field. Delbrück scattering has been thoroughly studied by theoretical and experimental methods (see the recent review in Ref. 2). At high photon energies ($\omega \gg m$, where m is the electron mass and $\hbar = c = 1$), the photon scattering amplitude was found exactly with respect to $Z\alpha$ ($Z|e|$ is the nuclear charge, $\alpha = e^2/4\pi = 1/137$ is the fine-structure constant, and e is the electron charge). The approach taken depends significantly on the momentum transfer $\Delta = |\mathbf{k}_2 - \mathbf{k}_1|$ ($\mathbf{k}_1$ is the momentum of the initial photon, and $\mathbf{k}_2$ is the momentum of the final photon). The main contribution to the total Delbrück scattering cross section at high photon energy comes from the range of small momentum transfer $\Delta \ll \omega$ (scattering angle $\theta \sim \Delta/\omega \ll 1$). Amplitudes for this case are found in Refs. 3–5 by summing the perturbation diagrams for the interaction with the Coulomb field in a certain approximation, and in Refs. 6 and 7 with the aid of the semiclassical approach. It is found for $\omega \gg m$ and $Z\alpha \sim 1$ that the exact result with respect to $Z\alpha$ differs significantly from the result obtained in lowest-order perturbation theory.

The possibility of using the semiclassical approximation stems from the fact that, in accordance with the uncertainty principle, the characteristic limiting parameter ρ is of the order of $1/\Delta$, and the angular momentum $l \sim \omega\rho \sim \omega/\Delta$ is much greater than unity at small scattering angles. This circumstance is utilized in Refs. 6 and 7, where the semiclassical Green's function is obtained from a convenient integral representation for the exact Green's function of the Dirac equation in a Coulomb field.[8] The semiclassical electronic Green's function for an arbitrary centrosymmetric, decreasing potential is found in Refs. 9 and 10, making it possible to calculate the Delbrück scattering amplitudes in a screened Coulomb potential.

The splitting of a photon in the Coulomb field of a nucleus has yet to be observed, although some events in an experiment performed at DESY[11] were erroneously interpreted as photon splitting. It is shown in Ref. 12 that these events correspond to the creation of an electron–positron pair and a hard photon. A possible experimental scheme is discussed in Ref. 13. Photon splitting is studied theoretically in Refs. 14–18 in lowest-order perturbation theory with respect to $Z\alpha$. The expressions obtained in Refs. 14 and 15 are very cumbersome, and it is difficult to use them for numerical calculations. Nevertheless, such calculations are performed in Refs. 17 and 18. In Ref. 16 the amplitudes of the process are obtained in significantly simpler form, but with logarithmic accuracy (using the equivalent-photon approximation). The comparison made in Ref. 17 of the exact Born cross section with the approximate result in Ref. 16 shows that the maximum difference reaches 20% in the energy range under consideration. The magnitude of the Coulomb corrections in the photon splitting process is still unknown. An experiment is presently being conducted at the G. I. Budker Nuclear Physics Institute of the Siberian Branch of the Russian Academy of Sciences to observe the splitting of a high-energy photon ($\omega \gg m$) in the Coulomb field of heavy nuclei. Therefore, a theoretical investigation of this process would be of unquestionable interest.

In the present work the splitting amplitude of a high-energy photon is calculated exactly with respect to $Z\alpha$ for small angles f_2 and f_3 between the momenta $\mathbf{k}_2$ and $\mathbf{k}_3$ of the final photons and the momentum $\mathbf{k}_1$ of the initial photon. It is in fact small angles that make the main contribution to the total cross section of the process. In addition, small angles and high photon energies enable one to use the semiclassical approach developed in Refs. 6, 7, 9, and 10 to consider Delbrück scattering. We discuss the case of an unscreened Coulomb potential, as well as the influence of screening. The initial expression for the photon splitting amplitude is very complicated, and contains a thirteenfold integral involving the parametrization of the Green's functions in a Coulomb field. The semiclassical approach provides a clear picture of the process, and permits determination of the integration region corresponding to the main contribution to the amplitude. Without it, calculation of the amplitude would probably be impossible.

This paper is organized as follows. In Sec. 2 we transform the exact amplitude, which significantly simplifies further calculations. The kinematics of the process are discussed in Sec. 3. The semiclassical Green's function is derived in

the small-angle approximation in Sec. 4. In Sec. 5 this Green's function is used to calculate the photon splitting amplitude. In Sec. 6 we examine the case of large transverse momentum components of the final photons ($\omega_2 f_2 \geqslant m$, $\omega_3 f_3 \geqslant m$). The asymptotes of the result obtained at small momentum transfers and at $Z\alpha \ll 1$ are discussed in Secs. 7 and 8. In the concluding section the exact Born cross sections are compared with the cross sections obtained using the equivalent-photon approximation (the Weizsäcker–Williams approximation).

2. TRANSFORMATION OF THE AMPLITUDE OF THE PROCESS

According to the Feynman rules, in the Furry representation the photon splitting amplitude has the form

$$M = ie^3 \int d^4x \ \mathrm{Tr}\langle x|\hat{e}_1 \exp(-ik_1x)\mathscr{G}\hat{e}_2^* \times \exp(ik_2x)\mathscr{G}\hat{e}_3^* \ \exp(ik_3x)\mathscr{G}|x\rangle + \{k_2^\mu \leftrightarrow k_3^\mu, e_2^\mu \leftrightarrow e_3^\mu\}. \tag{1}$$

Here e_1^μ and $e_{2,3}^\mu$ are the polarization vectors of the initial and final photons, $\hat{e} = e^\mu \gamma_\mu = -\mathbf{e}\cdot\boldsymbol{\gamma}$, the γ^μ are Dirac matrices, $\mathscr{G} = 1/(\hat{\mathscr{P}} - m + i0)$, and $\mathscr{P}_\mu = i\partial_\mu + g_{\mu 0}(Z\alpha/r)$. Each matrix element of the operator $\mathscr{G}$ is the Green's function of the Dirac equation for an electron in a Coulomb field: $G(x,x') = \langle x|\mathscr{G}|x'\rangle$.

It is convenient to rewrite (1) in a form that contains only the Green's functions of the squared Dirac equation

$$D(x,x') = \langle x|\mathscr{D}|x'\rangle = \langle x|1/(\hat{\mathscr{P}}^2 - m^2 + i0)|x'\rangle.$$

For this purpose, in (1) we represent the left-hand operator $\mathscr{G}$ in the form $\mathscr{G} = \mathscr{D}(\hat{\mathscr{P}} + m)$ and use the commutation relation

$$(\hat{\mathscr{P}} + m)\hat{e} \ \exp(ikx) = \exp(ikx)[-\hat{e}\mathscr{G}^{-1} + \hat{e}\hat{k} - 2\mathbf{e}\cdot\mathbf{p}],$$

$$\mathbf{p} = -i\nabla.$$

We now apply the same transformation to the right-hand operator $\mathscr{G}$ in (1) and take half the sum of the resulting expressions. Taking into account that

$$\int dx \ \mathrm{Tr}\langle x|A_1A_2|x\rangle = \int dx \ \mathrm{Tr}\langle x|A_2A_1|x\rangle,$$

for the arbitrary operators A_1 and A_2, we obtain

$$M = ie^3 \int d^4x \ \mathrm{Tr}\Big\{\Big[\mathbf{e}_1\cdot\mathbf{e}_2^* \ \exp[i(k_2-k_1)x] \times \langle x|\mathscr{D} \exp(ik_3x)(\hat{e}_3^*\hat{k}_3 - 2\mathbf{e}_3^*\cdot\mathbf{p})\mathscr{D}|x\rangle + \mathbf{e}_1\cdot\mathbf{e}_3^* \ \exp[i(k_3-k_1)x]\langle x|\mathscr{D} \exp(ik_2x)(\hat{e}_2^*\hat{k}_2 - 2\mathbf{e}_2^*\cdot\mathbf{p})\mathscr{D}|x\rangle + \mathbf{e}_2^*\cdot\mathbf{e}_3^* \ \exp[-i(k_2+k_3)x] \times \langle x|\mathscr{D} \exp(-ik_1x)(-\hat{e}_1\hat{k}_1 - 2\mathbf{e}_1\cdot\mathbf{p})\mathscr{D}|x\rangle\Big] + \frac{1}{2}[\langle x|\exp(-ik_1x)(-\hat{e}_1\hat{k}_1 - 2\mathbf{e}_1\cdot\mathbf{p})\mathscr{D} \exp(ik_2x) \times (\hat{e}_2^*\hat{k}_2 - 2\mathbf{e}_2^*\cdot\mathbf{p})\mathscr{D} \exp(ik_3x)(\hat{e}_3^*\hat{k}_3 - 2\mathbf{e}_3^*\cdot\mathbf{p}) \times \mathscr{D}|x\rangle + (2\leftrightarrow 3)]\Big\}. \tag{2}$$

Thus, representation (2) is a sum of contributions containing two and three Green's functions: $M = M^{(2)} + M^{(3)}$. The amplitude written in form (1) contains terms of different orders of magnitude. This makes it necessary to allow for cancellation of higher-order terms. The convenience of Eq. (2) stems from the fact that after the trace of the matrix is taken, the expression contains only terms of the required order.

Going from a time-dependent Green's function to an energy-dependent Green's function in (2), taking the integral over time, and omitting the standard factor $2\pi\delta(\omega_1 - \omega_2 - \omega_3)$, for the contribution containing three Green's functions we obtain

$$M^{(3)} = \frac{i}{2}e^3 \int \frac{d\varepsilon}{2\pi} \int d\mathbf{r}_1 \, d\mathbf{r}_2 \, d\mathbf{r}_3 \ \exp[i(\mathbf{k}_1\cdot\mathbf{r}_1 - \mathbf{k}_2\cdot\mathbf{r}_2 - \mathbf{k}_3\cdot\mathbf{r}_3)]\mathrm{Tr}\{[(-\hat{e}_1\hat{k}_1 - 2\mathbf{e}_1\cdot\mathbf{p})D(\mathbf{r}_1,\mathbf{r}_2|\varepsilon-\omega_2)] \times [(\hat{e}_2^*\hat{k}_2 - 2\mathbf{e}_2^*\cdot\mathbf{p})D(\mathbf{r}_2,\mathbf{r}_3|\varepsilon)][(\hat{e}_3^*\hat{k}_3 - 2\mathbf{e}_3^*\cdot\mathbf{p})D(\mathbf{r}_3,\mathbf{r}_1|\varepsilon+\omega_3)]\} + (k_2^\mu \leftrightarrow k_3^\mu, \mathbf{e}_2 \leftrightarrow \mathbf{e}_3). \tag{3}$$

Here $\mathbf{p} = -i\nabla$ differentiates the corresponding Green's function D of the squared Dirac equation with respect to the first argument.

We henceforth call the Green's function $D(\mathbf{r}_1,\mathbf{r}_2|\varepsilon)$ with positive energy ε the electron Green's function, and the function with negative energy the positron Green's function. Let the initial photon propagate along the z axis. Then, in accordance with the semiclassical approach developed in Refs. 6, 7, 9, and 10, at high energies the main contribution to the amplitude is made by the region of integration over the variables z_i such that $z' < z$ for the electron Green's function $D(\mathbf{r},\mathbf{r}'|\varepsilon)$ and $z' > z$ for the positron Green's function. In terms of noncovariant perturbation theory, this corresponds to the fact that over the indicated integration region, the difference between the energy E_n of any intermediate state and the energy of the initial state $E_0 = \omega_1$ is small compared to E_0. In all remaining cases, at least one of the intermediate states has $|E_n - E_0| \sim E_0$, and the corresponding contribution to the amplitude is suppressed.

In addition, there is another constraint on the integration region making the main contribution to the amplitude. It is associated with the properties of the semiclassical Green's function, and has the form

$$z_1 < z_2, z_3, \quad z_1 < 0, \quad \max(z_2,z_3) > 0.$$

All of the indicated inequalities make it possible to represent the main contribution to $M^{(3)}$ in the form of the diagrams shown in Fig. 1. The expressions for the vertices are obvious from Eq. (3). In Fig. 1 the electron Green's functions are represented by arrows running from left to right, and the positron Green's functions are represented by arrows running from right to left. All vertices in these diagrams are spatially ordered. Using this figure, we can easily write the limits for integration over energy and coordinates.

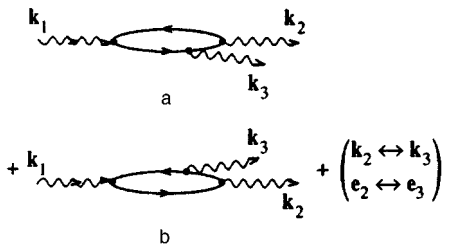


FIG. 1. Perturbation diagrams corresponding to $M^{(3)}$ [Eq. (3)].

As a result, diagrams a and b in Fig. 1 correspond to the following picture. At $\mathbf{r}_1$ a photon with momentum $\mathbf{k}_1$ creates a pair of virtual particles, which transforms into a photon with momentum $\mathbf{k}_2$ at $\mathbf{r}_2$. Between these two events, a photon with momentum $\mathbf{k}_3$ is emitted at $\mathbf{r}_3$ by the electron (a) or positron (b).

Similarly, the expression for the contribution $M^{(2)}$, which contains two Green's functions, has the form

$$
\begin{aligned}
M^{(2)} = ie^3 \int \frac{d\varepsilon}{2\pi} \int d\mathbf{r}_1\, d\mathbf{r}_2\ \mathrm{Tr}\{&\exp[i(\mathbf{k}_1\cdot\mathbf{r}_1 - \mathbf{k}_2\cdot\mathbf{r}_2 \\
&- \mathbf{k}_3\cdot\mathbf{r}_2)]\mathbf{e}_2^*\cdot\mathbf{e}_3^*[(-\hat{e}_1\hat{k}_1 - 2\mathbf{e}_1\cdot\mathbf{p}) \\
&\times D(\mathbf{r}_1,\mathbf{r}_2|\varepsilon-\omega_1)D(\mathbf{r}_2,\mathbf{r}_1|\varepsilon)] \\
&+ [\exp[i(\mathbf{k}_1\cdot\mathbf{r}_1 - \mathbf{k}_2\cdot\mathbf{r}_2 - \mathbf{k}_3\cdot\mathbf{r}_1)]\mathbf{e}_1\cdot\mathbf{e}_3^* \\
&\times D(\mathbf{r}_1,\mathbf{r}_2|\varepsilon-\omega_2)[(\hat{e}_2^*\hat{k}_2 - 2\mathbf{e}_2^*\cdot\mathbf{p})D(\mathbf{r}_2,\mathbf{r}_1|\varepsilon)] \\
&+ (k_2^\mu \leftrightarrow k_3^\mu, \mathbf{e}_2 \leftrightarrow \mathbf{e}_3)]\}. \qquad (4)
\end{aligned}
$$

The diagrams corresponding to the representation (4) of the amplitude are shown in Fig. 2.

3. KINEMATICS OF THE PROCESS

The lifetime of a virtual electron–positron pair (the length of the loop) can be estimated from the uncertainty relation $\tau \sim |\mathbf{r}_2 - \mathbf{r}_1| \sim \omega_1/(m^2+\tilde{\Delta}^2)$, where $\tilde{\Delta} = \max(|\mathbf{k}_{2\perp}|, |\mathbf{k}_{3\perp}|) \ll \omega_1$, and $\mathbf{k}_{2\perp}$ and $\mathbf{k}_{3\perp}$ are the components of the momenta of the final photons perpendicular to $\mathbf{k}_1$. The characteristic transverse distance between the virtual particles can be estimated as $(m^2+\tilde{\Delta}^2)^{-1/2}$. It is clear that the length of the electron–positron loop is much greater than its transverse dimensions. The characteristic impact parameter is $\rho \sim 1/\Delta$, where the momentum transfer $\boldsymbol{\Delta} = \mathbf{k}_2 + \mathbf{k}_3 - \mathbf{k}_1$. For small $\mathbf{k}_{2\perp}$ and $\mathbf{k}_{3\perp}$ ($f_{2,3} \ll 1$) we have

$$
\boldsymbol{\Delta}^2 = (\mathbf{k}_{2\perp} + \mathbf{k}_{3\perp})^2 + \frac{1}{4}\left(\frac{\mathbf{k}_{2\perp}^2}{\omega_2} + \frac{\mathbf{k}_{3\perp}^2}{\omega_3}\right)^2. \qquad (5)
$$

The characteristic orbital angular momentum $l \sim \omega/\Delta$ is much greater than unity, and the semiclassical approximation applies.

Let us consider a screened Coulomb potential. In the Thomas–Fermi model, the screening radius is $r_c \sim (m\alpha)^{-1} Z^{-1/3}$. If $R \ll 1/\Delta \ll r_c$ (R is the radius of the nucleus), screening is negligible, and the amplitude is the same as the amplitude in a Coulomb field. If $1/\Delta \sim r_c$, screening must be taken into account. It is clear that impact parameters $\rho \gg r_c$ do not contribute to the total cross section. Therefore, we henceforth consider only the range of momentum transfer which corresponds to impact parameters $\rho \leqslant r_c$. If $|\boldsymbol{\Delta}_\parallel| = (\mathbf{k}_{2\perp}^2/\omega_2 + \mathbf{k}_{3\perp}^2/\omega_3)/2 \ll r_c^{-1}$, it follows from Eq. (5) that $\rho \leqslant r_c$ only when $|\boldsymbol{\Delta}_\perp| = |\mathbf{k}_{2\perp} + \mathbf{k}_{3\perp}| \geqslant r_c^{-1}$. Thus, the main contribution to the cross section comes from a range of transverse momenta $\boldsymbol{\Delta}_\perp$ bounded from below. In this range $|\boldsymbol{\Delta}_\perp| \gg |\boldsymbol{\Delta}_\parallel|$, i.e., $\Delta \approx \boldsymbol{\Delta}_\perp$. In addition, when $\omega/(m^2+\tilde{\Delta}^2) \gg r_c$, the angles between the vectors $\mathbf{k}_{1,2,3}$ and $\mathbf{r}_{1,2,3}$ are either small or close to π, $|z_i| \gg r_c$, and the appropriate expansion can be used.

According to Furry's theorem, an odd number of quanta are exchanged with the source of the external field during photon splitting, i.e., the amplitude is odd in $Z\alpha$. The range of very small momentum transfer $\Delta \leqslant r_c^{-1}$ is significant only in lowest-order (linear with respect to $Z\alpha$) perturbation theory, owing to the singular behavior of the Coulomb potential in the momentum representation ($-4\pi Z\alpha/\Delta^2$). In this order we use the equivalent-photon approximation, and the corresponding large logarithm appears in the cross section integrated over the angles of one of the final photons.[16]

In higher-order perturbation theory with respect to $Z\alpha$, the integral must be taken over all the momenta corresponding to the external field under the condition that the sum of these momenta equals $\boldsymbol{\Delta}$. Therefore, even if $\Delta \sim r_c^{-1}$, each of the momenta of the external field is not small, and screening can be neglected. In lowest-order perturbation theory screening can be taken into account by multiplying the amplitude by $1 - F(\Delta^2)$, where $F(\Delta^2)$ is the form factor of the atomic electrons. Thus, to find the photon splitting amplitude in a screened Coulomb potential, it is sufficient to solve the problem for the unscreened Coulomb potential.

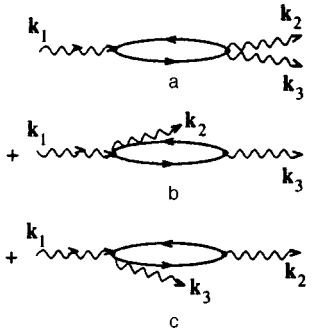


FIG. 2. Perturbation diagrams corresponding to $M^{(2)}$ [Eq. (4)].

4. GREEN'S FUNCTION

We now proceed to a treatment of the Green's function $D(\mathbf{r},\mathbf{r}'|\varepsilon)$ appearing in (3) and (4). A representation of this function for an arbitrary centrosymmetric, decreasing potential was found using the semiclassical approximation in Refs. 9 and 10. In the case of a Coulomb potential and a small angle θ between the vectors $\mathbf{r}$ and $-\mathbf{r}'$, from Eq. (14) in Ref. 9 we obtain

$$D(\mathbf{r},\mathbf{r}'|\varepsilon)=\frac{i\,\exp[i\kappa(r+r')]}{4\pi\kappa rr'}\times\int_0^\infty dl\;l\left[J_0(l\theta)+iZ\alpha\frac{\boldsymbol{\alpha}\cdot(\mathbf{n}+\mathbf{n}')}{l\theta}J_1(l\theta)\right]\times\exp\left[i\frac{l^2(r+r')}{2\kappa rr'}\right]\left(\frac{4\kappa^2rr'}{l^2}\right)^{iZ\alpha\lambda},\qquad(6)$$

where $\boldsymbol{\alpha}=\gamma^0\boldsymbol{\gamma}$, $\kappa^2=\varepsilon^2-m^2$, $\lambda=\varepsilon/\kappa$, $\mathbf{n}=\mathbf{r}/r$, and $\mathbf{n}'=\mathbf{r}'/r'$. Taking into account that

$$\int dl\;lJ_0(l\theta)g(l^2)=\frac{1}{2\pi}\int d\mathbf{q}\,\exp(i\mathbf{q}\cdot\boldsymbol{\theta})g(q^2),$$

$$\frac{\boldsymbol{\theta}}{\theta}J_1(l\theta)=-\frac{1}{l}\frac{\partial}{\partial\boldsymbol{\theta}}J_0(l\theta),$$

where $g(l^2)$ is an arbitrary function and $\mathbf{q}$ is a two-dimensional vector, we can rewrite Eq. (6) in the form

$$D(\mathbf{r},\mathbf{r}'|\varepsilon)=\frac{i\,\exp[i\kappa(r+r')]}{8\pi^2\kappa rr'}\int d\mathbf{q}\times\left(1+Z\alpha\frac{\boldsymbol{\alpha}\cdot\mathbf{q}}{q^2}\right)\exp\left[i\frac{q^2(r+r')}{2\kappa rr'}+i\mathbf{q}\cdot(\boldsymbol{\theta}+\boldsymbol{\theta}')\right]\left(\frac{4\kappa^2rr'}{q^2}\right)^{iZ\alpha\lambda}.\qquad(7)$$

Here $\boldsymbol{\theta}=\mathbf{r}_\perp/r$, and $\boldsymbol{\theta}'=\mathbf{r}'_\perp/r'$. Equation (7) contains only elementary functions, and the angles $\boldsymbol{\theta}$ and $\boldsymbol{\theta}'$ appear only in the factor $\exp\{i\mathbf{q}\cdot(\boldsymbol{\theta}+\boldsymbol{\theta}')\}$. Therefore, the representation (7) for the Green's function is very convenient for calculations.

If the angle between $\mathbf{r}$ and $\mathbf{r}'$ is small, then in a Coulomb field, from Eq. (15) in Ref. 9 we find

$$D(\mathbf{r},\mathbf{r}'|\varepsilon)=-\frac{\exp(i\kappa|\mathbf{r}-\mathbf{r}'|)}{4\pi|\mathbf{r}-\mathbf{r}'|}\left(\frac{r}{r'}\right)^{iZ\alpha\lambda\;\mathrm{sgn}(r-r')}.\qquad(8)$$

It is clear that the Green's function in this case is distinguished from the free function only by the phase factor. It is easy to prove that all phase factors of the form $r^{\pm iZ\alpha}$ in the representations (7) and (8) of the Green's functions cancel when these functions are plugged into the expressions for amplitudes (3) and (4). We also note that κ can be replaced by $|\varepsilon|-m^2/(2|\varepsilon|)$ in (7) and (8) and that to within the required accuracy, the correction $m^2/(2|\varepsilon|)$ is significant only in the phase factor $\exp[i\kappa(r+r')]$.

5. CALCULATION OF THE AMPLITUDES $M^{(3)}$ AND $M^{(2)}$

We now move on to the calculation of diagrams containing three Green's functions (see Fig. 1). It is clear that the contribution of diagram *b* can be obtained from the contribution of diagram *a* by making the replacement $Z\alpha\to-Z\alpha$ and changing the overall sign. This ensures the applicability of Furry's theorem (the sum of the contributions is odd with respect to $Z\alpha$). Therefore, the sum of the contributions of *a* and *b* in Fig. 1 can be obtained from the contribution of *a* by isolating the part that is odd with respect to $Z\alpha$ and multiplying by 2.

To calculate diagram *a*, the integration interval with respect to z_3 is divided in two: $z_3>0$ (photon ahead with momentum $\mathbf{k}_3$) and $z_3<0$ (photon behind). In the former interval the angles between $\mathbf{r}_2$, $\mathbf{r}_3$, and $-\mathbf{r}_1$ are small. In the latter interval the angles between $\mathbf{r}_1$, $\mathbf{r}_3$, and $-\mathbf{r}_2$ are small. We use $M_1^{(3)}$ to denote the contribution of the former interval to diagram *a* and $M_2^{(3)}$ to denote the contribution of the latter interval. We introduce the vectors $\mathbf{f}_2=\mathbf{k}_{2\perp}/\omega_2$ and $\mathbf{f}_3=\mathbf{k}_{3\perp}/\omega_3$ ($|\mathbf{f}_{2,3}|\ll1$), as well as $\boldsymbol{\theta}_i=\mathbf{r}_{i\perp}/r_i=\mathbf{n}_{i\perp}$ ($i=1,2,3$). When the small magnitude of the angles is taken into account, we have $d\mathbf{r}_i=r_i^2dr_i\,d\boldsymbol{\theta}_i$.

It is convenient to perform the calculation for the helical amplitudes $M_{\lambda_1\lambda_2\lambda_3}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)$, for which it is sufficient to find three amplitudes: $M_{+--}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)$, $M_{+++}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)$, and $M_{++-}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)$. All remaining amplitudes can be obtained by replacement of variables.

Note that in the approximation under consideration, there is no need to allow for corrections to the transverse parts of the polarization vectors $\mathbf{e}_{2,3}$, and that the longitudinal parts of $\mathbf{e}_{2,3}$ can be expressed in terms of the transverse parts using the relation $\mathbf{e}\cdot\mathbf{k}=0$: $(\mathbf{e}_{2,3})=-\mathbf{e}_{2,3}\cdot\mathbf{f}_{2,3}$. Thus, we can assume that the transverse part of the polarization vector of the final photon with given helicity is equal to the polarization vector of a photon with the same helicity propagating along the *z* axis. Henceforth we use the notation $\mathbf{e}$ for this polarization vector in the case of positive helicity; the polarization vector with negative helicity will then be $\mathbf{e}^*$. Note that since the arrangement of the vertices in the diagrams is spatially ordered in our approach, the calculation of M_{++-} requires finding two $M^{(3)}$ amplitudes, namely, $M^{(3)}_{++-}$ and $M^{(3)}_{+-+}$.

We substitute (7) and (8) into (3) and perform the obvious expansion at small angles, taking account of terms quadratic in $\mathbf{f}_i$ and $\boldsymbol{\theta}_i$. We introduce the notation $\kappa_2=\omega_2-\varepsilon$ and $\kappa_3=\omega_3+\varepsilon$ and transform to variables

$$\mathbf{q}_2\to\kappa_2\mathbf{q}_2,\quad \mathbf{q}_3\to\kappa_3\mathbf{q}_3,\quad R_1=\frac{\omega_1}{\kappa_2\kappa_3}r_1,$$

$$R_2=\frac{\omega_2}{\varepsilon\kappa_2}r_2,\quad R_3=\frac{\omega_2}{\varepsilon\kappa_3}r_3.$$

After simple integration over $\boldsymbol{\theta}_i$ we obtain

$$M_1^{(3)}=\frac{e^3}{32\pi^3\omega_1\omega_2\omega_3}\int_0^{\omega_2}\varepsilon\kappa_2\kappa_3d\varepsilon\int_0^\infty dR_1\int_0^\infty dR_2\int_0^L\frac{dR_3}{R_1R}\times\int\int d\mathbf{q}_2\,d\mathbf{q}_3\left(\frac{q_2}{q_3}\right)^{2iZ\alpha}\exp(i\Phi)T,\qquad(9)$$

where $L=R_2\omega_3\kappa_2/\omega_2\kappa_3$,

$$T=\frac{1}{4}\,\mathrm{Tr}\left\{\left(1+\frac{Z\alpha\;\boldsymbol{\alpha}\cdot\mathbf{q}_3}{\kappa_3q_3^2}\right)\left(\frac{2}{R_1}\mathbf{e}_1\cdot\mathbf{Q}-\hat{e}_1\hat{k}_1\right)\times\left(1-\frac{Z\alpha\;\boldsymbol{\alpha}\cdot\mathbf{q}_2}{\kappa_2q_2^2}\right)\left[\left(\frac{2}{R}\mathbf{e}_2^*\cdot(\mathbf{Q}+\varepsilon R_3\mathbf{f}_{23})-\hat{e}_2^*\hat{k}_2\right)\times\left(\frac{2}{R}\mathbf{e}_3^*\cdot(\mathbf{Q}+\varepsilon R_2\mathbf{f}_{23})-\hat{e}_3^*\hat{k}_3\right)-\frac{4i}{R}\mathbf{e}_2^*\cdot\mathbf{e}_3^*\right]\right\},\qquad(10)$$

$$\Phi=\left[\left(\frac{1}{R}+\frac{1}{R_1}\right)\frac{\mathbf{Q}^2}{2}+\frac{\varepsilon^2R_2R_3\mathbf{f}_{23}^2}{2R}-\frac{(\kappa_2\mathbf{q}_2-\kappa_3\mathbf{q}_3)\cdot\mathbf{\Delta}}{\omega_1}\right.$$
$$\left.-\frac{\omega_3\kappa_2R_2-\omega_2\kappa_3R_3}{\omega_1R}\mathbf{Q}\cdot\mathbf{f}_{23}-\frac{m^2}{2}(R_1+R)\right],$$
$$R=R_2-R_3,\quad \mathbf{f}_{23}=\mathbf{f}_2-\mathbf{f}_3,\quad \mathbf{Q}=\mathbf{q}_2+\mathbf{q}_3,$$
$$\mathbf{\Delta}=\omega_2\mathbf{f}_2+\omega_3\mathbf{f}_3.$$

For subsequent calculations it is convenient to transform the function T in (10) so that it does not contain $Z\alpha$. For this purpose we utilize the identities

$$Z\alpha\frac{\mathbf{q}_2}{q_2^2}\left(\frac{q_2}{q_3}\right)^{2iZ\alpha}=-\frac{i}{2}\frac{\partial}{\partial\mathbf{q}_2}\left(\frac{q_2}{q_3}\right)^{2iZ\alpha},$$
$$Z\alpha\frac{\mathbf{q}_3}{q_3^2}\left(\frac{q_2}{q_3}\right)^{2iZ\alpha}=\frac{i}{2}\frac{\partial}{\partial\mathbf{q}_3}\left(\frac{q_2}{q_3}\right)^{2iZ\alpha}$$

and integrate by parts over $\mathbf{q}_2$ and $\mathbf{q}_3$ in (9). Thereupon terms appear that contain the variable R_1 in the factors $1/R_1$ and $1/R_1^2$ and do not depend on R_1. Calculating the trace of each matrix and integrating the terms containing $1/R_1^2$ by parts over R_1, for the various polarizations we find

$$T_{+--}=\frac{8}{R_1R^2}(\mathbf{e}\cdot\mathbf{Q})(\mathbf{e}\cdot\mathbf{Q}_2)(\mathbf{e}\cdot\mathbf{Q}_3),$$
$$T_{+++}=-\frac{4}{R_1R^2}\left(\frac{\kappa_2}{\kappa_3}+\frac{\kappa_3}{\kappa_2}\right)(\mathbf{e}\cdot\mathbf{Q})(\mathbf{e}^*\cdot\mathbf{Q}_2)(\mathbf{e}^*\cdot\mathbf{Q}_3)$$
$$-\frac{\omega_1}{\varepsilon R^3}\mathbf{e}^*\cdot\left(\frac{\omega_2}{\kappa_2}\mathbf{Q}_3^2\mathbf{Q}_2-\frac{\omega_3}{\kappa_3}\mathbf{Q}_2^2\mathbf{Q}_3\right)+\frac{2i\omega_1^2}{\kappa_2\kappa_3R^2}\mathbf{e}^*$$
$$\times(\mathbf{Q}_2+\mathbf{Q}_3)+\frac{m^2\omega_1}{\varepsilon R}\mathbf{e}^*\cdot\left(\frac{\omega_2}{\kappa_2}\mathbf{Q}_2-\frac{\omega_3}{\kappa_3}\mathbf{Q}_3\right),\tag{11}$$

$$T_{++-}=-\frac{4}{R_1R^2}\left(\frac{\kappa_2}{\varepsilon}+\frac{\varepsilon}{\kappa_2}\right)(\mathbf{e}\cdot\mathbf{Q})(\mathbf{e}\cdot\mathbf{Q}_2)(\mathbf{e}^*\cdot\mathbf{Q}_3)$$
$$+\frac{\omega_2\omega_3}{\varepsilon\kappa_3R_1R^2}(\mathbf{Q}_2^2+\mathbf{Q}_3^2)(\mathbf{e}\cdot\mathbf{Q})-\frac{\omega_1\omega_2}{\kappa_2\kappa_3R}$$
$$\times\left(\frac{\mathbf{Q}_3^2}{R^2}-m^2\right)(\mathbf{e}\cdot\mathbf{Q}_2)+\frac{4i}{R_1R}\left(\frac{\omega_1\omega_2}{\kappa_2\kappa_3}-2\right)(\mathbf{e}\cdot\mathbf{Q})$$
$$+\frac{2i\omega_1\omega_2}{\kappa_2\kappa_3R^2}\mathbf{e}\cdot(\mathbf{Q}_2+\mathbf{Q}_3),$$

where $\mathbf{Q}_2=\mathbf{Q}+\varepsilon R_2\mathbf{f}_{23}$ and $\mathbf{Q}_3=\mathbf{Q}+\varepsilon R_3\mathbf{f}_{23}$. The function T_{+-+} can be obtained from T_{++-} using the replacements $\omega_2\leftrightarrow\omega_3$, $\kappa_2\leftrightarrow\kappa_3$, $\mathbf{Q}_2\leftrightarrow\mathbf{Q}_3$, and $\varepsilon\to-\varepsilon$.

Similarly, for $M_2^{(3)}$ we obtain

$$M_2^{(3)}=\frac{e^3}{32\pi^3\omega_1\omega_2\omega_3}\int_0^{\omega_2}\varepsilon\kappa_2\kappa_3d\varepsilon\int_0^\infty dR_1\int_0^\infty dR_2$$
$$\times\int_0^{L_1}\frac{dR_3}{rR_2}\int\int d\mathbf{q}_2\,d\mathbf{q}_3\left(\frac{q_2}{q_3}\right)^{2iZ\alpha}\exp(i\widetilde{\Phi})\widetilde{T},\tag{12}$$

where $L_1=R_1\omega_3\kappa_2/\omega_1\varepsilon$, $r=R_1+R_3$,

$$\widetilde{\Phi}=\left[\left(\frac{1}{r}+\frac{1}{R_2}\right)\frac{\mathbf{Q}^2}{2}-\frac{\kappa_3^2R_1R_3\mathbf{f}_3^2}{2r}-\frac{(\kappa_2\mathbf{q}_2-\varepsilon\mathbf{q}_3)\mathbf{\Delta}}{\omega_2}\right]$$
$$+\frac{\omega_3\kappa_2R_1-\varepsilon\omega_1R_3}{\omega_2r}\mathbf{Q}\cdot\mathbf{f}_3-\frac{m^2}{2}(R_2+r),\tag{13}$$

and the function $\widetilde{T}$ for different polarizations is

$$\widetilde{T}_{+--}=-\frac{8}{r^2R_2}(\mathbf{e}\cdot\mathbf{Q})(\mathbf{e}\cdot\mathbf{P}_1)(\mathbf{e}\cdot\mathbf{P}_3);\tag{14}$$
$$\widetilde{T}_{++-}=\frac{4}{r^2R_2}\left(\frac{\kappa_2}{\varepsilon}+\frac{\varepsilon}{\kappa_2}\right)(\mathbf{e}^*\cdot\mathbf{Q})(\mathbf{e}\cdot\mathbf{P}_1)(\mathbf{e}\cdot\mathbf{P}_3)$$
$$+\frac{\omega_2}{\kappa_3r^3}\mathbf{e}\left(\frac{\omega_1}{\kappa_2}\mathbf{P}_3^2\cdot\mathbf{P}_1+\frac{\omega_3}{\varepsilon}\mathbf{P}_1^2\cdot\mathbf{P}_3\right)-\frac{2i\omega_2^2}{\kappa_2\varepsilon r^2}\mathbf{e}$$
$$\times(\mathbf{P}_1+\mathbf{P}_3)-\frac{m^2\omega_2}{\kappa_3r}\mathbf{e}\cdot\left(\frac{\omega_1}{\kappa_2}\mathbf{P}_1+\frac{\omega_3}{\varepsilon}\mathbf{P}_3\right),$$
$$\widetilde{T}_{+++}=\frac{4}{r^2R_2}\left(\frac{\kappa_2}{\kappa_3}+\frac{\kappa_3}{\kappa_2}\right)(\mathbf{e}^*\cdot\mathbf{Q})(\mathbf{e}^*\cdot\mathbf{P}_1)(\mathbf{e}\cdot\mathbf{P}_3)$$
$$+\frac{\omega_1\omega_3}{\varepsilon\kappa_3r^2R_2}(\mathbf{P}_1^2+\mathbf{P}_3^2)(\mathbf{e}^*\cdot\mathbf{Q})+\frac{\omega_1\omega_2}{\kappa_2\varepsilon r}\left(\frac{\mathbf{P}_3^2}{r^2}\right.$$
$$\left.-m^2\right)(\mathbf{e}^*\cdot\mathbf{P}_1)-\frac{4i}{rR_2}\left(\frac{\omega_1\omega_2}{\kappa_2\varepsilon}-2\right)(\mathbf{e}^*\cdot\mathbf{Q})$$
$$-\frac{2i\omega_1\omega_2}{\kappa_2\varepsilon r^2}\mathbf{e}^*\cdot(\mathbf{P}_1+\mathbf{P}_3),$$

where $\mathbf{P}_1=\mathbf{Q}+\kappa_3R_1\mathbf{f}_3$ and $\mathbf{P}_3=\mathbf{Q}-\kappa_3R_3\mathbf{f}_3$. The function $\widetilde{T}_{+-+}$ can be obtained from $\widetilde{T}_{+++}$ by making the replacements $\omega_1\leftrightarrow\omega_3$, $\kappa_2\leftrightarrow-\varepsilon$, $\mathbf{P}_1\leftrightarrow\mathbf{P}_3$, and $\mathbf{e}\leftrightarrow\mathbf{e}^*$. Note that the integrand for the helical amplitudes $M_2^{(3)}$ in (12) and (14) can be obtained from the integrand for $M_1^{(3)}$ in (9) and (11) with the replacements

$$\mathbf{q}_{2,3}\to-\mathbf{q}_{2,3},\quad \omega_1\leftrightarrow\omega_2,\quad \omega_3\to-\omega_3,\quad \kappa_3\leftrightarrow\varepsilon,$$
$$R_1\leftrightarrow R_2,\quad R_3\to-R_3,\quad \mathbf{f}_{23}\leftrightarrow-\mathbf{f}_3,\quad \mathbf{f}_2\to-\mathbf{f}_2.\tag{15}$$

Here $T_{+--}\to\widetilde{T}_{+--}$, $T_{+-+}\to\widetilde{T}_{+-+}$, $T_{+++}\to\widetilde{T}_{++-}\times(\mathbf{e}\leftrightarrow\mathbf{e}^*)$, and $T_{++-}\to\widetilde{T}_{+++}(\mathbf{e}\leftrightarrow\mathbf{e}^*)$.

Similarly, for $M^{(2)}$ we find

$$M_{+--}^{(2)}=0,\quad M_{+++}^{(2)}=\mathbf{e}^*\cdot(\mathbf{M}_{12}+\mathbf{M}_{13}),$$
$$M_{++-}^{(2)}=\mathbf{e}\cdot(\mathbf{M}_{12}+\mathbf{M}_{23}),\quad M_{+-+}^{(2)}=\mathbf{e}\cdot(\mathbf{M}_{13}+\mathbf{M}_{23}),\tag{16}$$
$$\mathbf{M}_{23}=-\frac{ie^3}{16\pi^3}\int_{-\omega_3}^{\omega_2}d\varepsilon\int_0^\infty\frac{dR_1}{R_1^2}\int_0^\infty\frac{dR_2}{R_2^2}\left[R_1\right.$$
$$\left.+\left(\frac{\kappa_2-\kappa_3}{\omega_1}\right)^2R_2\right]\int\int d\mathbf{q}_2\,d\mathbf{q}_3\,\mathbf{Q}\left(\frac{q_2}{q_3}\right)^{2iZ\alpha}$$
$$\times\exp\left\{i\left[\left(\frac{1}{R_1}+\frac{1}{R_2}\right)\frac{\mathbf{Q}^2}{2}+\frac{\omega_2\omega_3\kappa_2\kappa_3}{2\omega_1^2}\mathbf{f}_{23}^2R_2\right.\right.$$
$$\left.\left.-\frac{(\kappa_2\mathbf{q}_2-\kappa_3\mathbf{q}_3)\cdot\mathbf{\Delta}}{\omega_1}-\frac{m^2}{2}(R_1+R_2)\right]\right\},$$

$$\mathbf{M}_{13}=\frac{ie^3}{16\pi^3}\int_0^{\omega_2}d\varepsilon\int_0^\infty\frac{dR_1}{R_1^2}\int_0^\infty\frac{dR_2}{R_2^2}\Bigg[R_2+\left(\frac{\kappa_2-\varepsilon}{\omega_2}\right)^2R_1\Bigg]\int\int d\mathbf{q}_2\,d\mathbf{q}_3\,\mathbf{Q}\left(\frac{q_2}{q_3}\right)^{2iZ\alpha}\times\exp\Bigg\{i\Bigg[\left(\frac{1}{R_1}+\frac{1}{R_2}\right)\frac{\mathbf{Q}^2}{2}-\frac{\omega_1\omega_3\varepsilon\kappa_2}{2\omega_2^2}\mathbf{f}_3^2R_1-\frac{(\kappa_2\mathbf{q}_2-\varepsilon\mathbf{q}_3)\cdot\boldsymbol{\Delta}}{\omega_2}-\frac{m^2}{2}(R_1+R_2)\Bigg]\Bigg\},\tag{17}$$

and M_{12} can be obtained from M_{13} with the replacements $\omega_2\leftrightarrow\omega_3$ and $\mathbf{f}_3\leftrightarrow\mathbf{f}_2$. As we see, a large number of terms in the sum $M^{(2)}+M^{(3)}$ cancel.

In the general case, further transformation of the expressions leads to quadruple integrals of elementary functions, and this problem requires relatively detailed consideration. Below we confine ourselves to a detailed discussion of the case $|\mathbf{k}_{2\perp}|=|\omega_2\mathbf{f}_2|\gg m$, $|\mathbf{k}_{3\perp}|=|\omega_3\mathbf{f}_3|\gg m$, for which the expressions for the amplitudes can be simplified significantly. This range of parameters corresponds to a more virtual electron–positron pair in comparison to the electron mass, which can be neglected in this case. We note that the relation between the momentum transfer $\Delta=|\boldsymbol{\Delta}|$ and the electron mass m can be arbitrary here, since Δ determines the characteristic impact parameter $\rho\sim1/\Delta$, rather than the virtuality of the pair.

6. ZERO-MASS LIMIT

It is not difficult to see that setting $m=0$ in these expressions leads to logarithmic divergences in individual terms (i.e., to the appearance of $\ln m$ for a finite mass). Such logarithms appear, for example, in $M_1^{(3)}$ when the terms in T not containing the multiplier $1/R_1$ are integrated over R_1 [see (11)]. The final result properly contains no logarithms of the mass. However, the cancellation of these logarithms between the individual contributions is quite nontrivial.

When the integral over R_1 is taken in Eq. (9) for $T=T_{+--}$, no logarithms appear, and $M_{+--}^{(2)}=0$. Formulas (11) for T_{+++} and T_{++-} contain terms proportional to $\mathbf{Q}_2^2$. In these terms it is convenient to pass from the variables R_2 and R_3 to R_2 and $y=R_3/R$ and to integrate by parts over y. In the terms containing $\mathbf{Q}_3^2$ we go over to the variables R_3 and $y=R_3/R$ and also integrate by parts over y. As a result, in the double integral over R_2 and R_3, all terms containing the logarithm of the mass cancel, and we can set $m=0$. Thereupon T_{+++} and T_{++-} become

$$T_{+++}=-\frac{4}{R_1R^2}\left(\frac{\kappa_2}{\kappa_3}+\frac{\kappa_3}{\kappa_2}\right)(\mathbf{e}\cdot\mathbf{Q})(\mathbf{e}^*\cdot\mathbf{Q}_2)(\mathbf{e}^*\cdot\mathbf{Q}_3),$$

$$T_{++-}=-\frac{4}{R_1R^2}\left(\frac{\kappa_2}{\varepsilon}+\frac{\varepsilon}{\kappa_2}\right)(\mathbf{e}\cdot\mathbf{Q})\times[(\mathbf{e}\cdot\mathbf{Q}_2)(\mathbf{e}^*\cdot\mathbf{Q}_3)-iR].\tag{18}$$

Moreover, terms appear outside the integral when $y=\omega_3\kappa_2/\omega_1\varepsilon$ (upper limit) and $y=0$ (lower limit). Similar transformations must be performed for $M_2^{(3)}$. Accordingly, $\widetilde{T}_{+++}$ and $\widetilde{T}_{++-}$ become

$$\widetilde{T}_{++-}=\frac{4}{r^2R_2}\left(\frac{\kappa_2}{\varepsilon}+\frac{\varepsilon}{\kappa_2}\right)(\mathbf{e}^*\cdot\mathbf{Q})(\mathbf{e}\cdot\mathbf{P}_1)(\mathbf{e}\cdot\mathbf{P}_3),$$

$$\widetilde{T}_{+++}=\frac{4}{r^2R_2}\left(\frac{\kappa_2}{\kappa_3}+\frac{\kappa_3}{\kappa_2}\right)(\mathbf{e}^*\cdot\mathbf{Q})[(\mathbf{e}^*\cdot\mathbf{P}_1)(\mathbf{e}\cdot\mathbf{P}_3)-iR].\tag{19}$$

At the lower limit, terms outside the integral cancel in the sum of $M_1^{(3)}$ and $M_2^{(3)}$. Recall that to calculate $M^{(3)}$, we must find the sum $M_1^{(3)}+(\mathbf{k}_2\leftrightarrow\mathbf{k}_3,\mathbf{e}_2\leftrightarrow\mathbf{e}_3)$, isolate the part that is odd with respect to $Z\alpha$, and multiply by 2. The contributions of terms outside the integral from $M_1^{(3)}$ at the upper limit for M_{+++} then cancel, while they yield a finite expression when $m=0$ in the sum with $\mathbf{e}\cdot\mathbf{M}_{23}$ for M_{++-} and M_{+-+}.

To cancel out the contributions that are singular in m, we must take advantage of the antisymmetry of the integrands under the replacements $\varepsilon\to\omega_2-\omega_3-\varepsilon$ and $\mathbf{q}_2\leftrightarrow-\mathbf{q}_3$. The contributions of terms outside the integral from $M_2^{(3)}$ at the upper limit vanish for M_{++-}, but they yield an expression that is finite at $m=0$ in the sums with $\mathbf{e}^*\cdot\mathbf{M}_{13}$ and $\mathbf{e}\cdot\mathbf{M}_{13}$ for M_{+++} and M_{+-+}, respectively. Similarly, the amplitudes $\mathbf{e}^*\cdot\mathbf{M}_{12}$ and $\mathbf{e}\cdot\mathbf{M}_{12}$ cancel the singular contributions from $M_2^{(3)}(\mathbf{k}_2\leftrightarrow\mathbf{k}_3)$ for M_{+++} and M_{++-}. As for M_{+-+}, the integrated terms from $M_2^{(3)}(\mathbf{k}_2\leftrightarrow\mathbf{k}_3)$ vanish at the upper limit.

As a result, the sum of the terms outside the integral and $M^{(2)}$ yield additional contributions to the helical amplitudes, which we represent in the form

$$\delta M=-\frac{e^3}{4\pi^3}\int_0^\infty\frac{dR}{R}\int\int\frac{d\mathbf{q}_2\,d\mathbf{q}_3}{\mathbf{Q}^2}\times\left[\left(\frac{q_2}{q_3}\right)^{2iZ\alpha}-\left(\frac{q_3}{q_2}\right)^{2iZ\alpha}\right]F.\tag{20}$$

The function F for the various helicities is

$$F_{+--}=0,\quad F_{+-+}=(\mathbf{e}\cdot\mathbf{Q})\left[\int_{-\omega_3}^{\omega_2}d\varepsilon\,\frac{\kappa_2\kappa_3^2}{\omega_1^2\varepsilon}\exp(i\psi_1)-\int_0^{\omega_2}d\varepsilon\frac{\kappa_2\varepsilon^2}{\omega_2^2\kappa_3}\exp(i\psi_2)\right],$$

$$F_{+++}=(\mathbf{e}^*\cdot\mathbf{Q})\int_0^{\omega_2}d\varepsilon\frac{\varepsilon\kappa_2^2}{\omega_2^2\kappa_3}\exp(i\psi_2)+(\omega_2\leftrightarrow\omega_3,\mathbf{f}_2\leftrightarrow\mathbf{f}_3),\tag{21}$$

$$F_{++-}=F_{+-+}(\omega_2\leftrightarrow\omega_3,\mathbf{f}_2\leftrightarrow\mathbf{f}_3),$$

$$\psi_1=\frac{\mathbf{Q}^2}{2R}+\frac{\omega_2\omega_3\kappa_2\kappa_3}{2\omega_1^2}\mathbf{f}_{23}^2R-\frac{(\kappa_2\mathbf{q}_2-\kappa_3\mathbf{q}_3)\cdot\boldsymbol{\Delta}}{\omega_1}.$$

The phase ψ_2 can be obtained from ψ_1 by replacement (15). To obtain Eq. (20) we took the integral over one of the radii. Note that the singularity of the integrand at $\varepsilon=0$ in (21) is

fictitious and that it cancels out in the complete expression for the amplitude of the process. Therefore, it does not need to be evaluated.

We substitute Eqs. (18) and (19) into (9) and (12), respectively, set $m=0$ in the phases Φ and $\widetilde{\Phi}$, and take the elementary integral over R_1 in (9) and over R_2 in (12). Then we pass from the variables $\mathbf{q}_2$ and $\mathbf{q}_3$ to the variables $\mathbf{Q}=\mathbf{q}_2+\mathbf{q}_3$ and $\mathbf{q}=\mathbf{q}_2-\mathbf{q}_3$. As a result, the integral over $\mathbf{q}$ has the form

$$J=\int \frac{d\mathbf{q}}{\mathbf{Q}^2}\left(\frac{|\mathbf{q}+\mathbf{Q}|}{|\mathbf{q}-\mathbf{Q}|}\right)^{2iZ\alpha}\exp\left(-\frac{i}{2}\mathbf{q}\cdot\mathbf{\Delta}\right). \tag{22}$$

We make the replacement $\mathbf{q}\to|\mathbf{Q}|\mathbf{q}$ and represent $\mathbf{Q}$ and $\mathbf{\Delta}$ in the form $\mathbf{Q}=|\mathbf{Q}|\boldsymbol{\lambda}_1$, $\mathbf{\Delta}=|\mathbf{\Delta}|\boldsymbol{\lambda}_2$. It is not difficult to see that J is a function of $S=\boldsymbol{\lambda}_1\cdot\boldsymbol{\lambda}_2$ and $|\mathbf{Q}||\mathbf{\Delta}|$.

Note that in the two-dimensional case the form $P=\varepsilon_{ij}\lambda_1^i\lambda_2^j$ is also invariant under rotation; however, even powers of P can be expressed in terms of S ($P^2=1-S^2$), and odd powers of P change sign upon reflection relative to an arbitrary axis.

At the same time, the integral J is invariant under reflection. This becomes obvious, if the vector $\mathbf{q}$ is reflected simultaneously with $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$. Thus, J does not vary under the replacement $\mathbf{Q}\leftrightarrow\mathbf{\Delta}$. After this replacement, J takes the form

$$J=\int \frac{d\mathbf{q}}{\mathbf{\Delta}^2}\left(\frac{|\mathbf{q}+\mathbf{\Delta}|}{|\mathbf{q}-\mathbf{\Delta}|}\right)^{2iZ\alpha}\exp\left(-\frac{i}{2}\mathbf{q}\cdot\mathbf{Q}\right). \tag{23}$$

Using this representation, we can easily take the integrals over $\mathbf{Q}$ and over all radii. Summing all the contributions, we ultimately obtain

$$M=\frac{8e^3}{\pi^2\omega_1\omega_2\omega_3\mathbf{\Delta}^2}\int d\mathbf{q}(\mathbf{T}\cdot\boldsymbol{\nabla}_{\mathbf{q}}\,\mathrm{Im})\left(\frac{|\mathbf{q}+\mathbf{\Delta}|}{|\mathbf{q}-\mathbf{\Delta}|}\right)^{2iZ\alpha}, \tag{24}$$

$$\mathbf{T}_{+--}=\omega_3\mathbf{e}\int_0^{\omega_2}d\varepsilon\,\frac{\kappa_2^2}{\mathbf{e}^*\cdot\mathbf{a}}\left[\frac{(\mathbf{e}\cdot\mathbf{b})\kappa_3}{\omega_1\mathscr{D}_1}-\frac{(\mathbf{e}\cdot\mathbf{c})\varepsilon}{\omega_2\mathscr{D}_3}\right]+\left(\begin{matrix}\omega_2\leftrightarrow\omega_3\\ \mathbf{f}_2\leftrightarrow\mathbf{f}_3\end{matrix}\right),$$

$$\mathbf{T}_{+++}=\omega_3\int_0^{\omega_2}d\varepsilon\,\kappa_3\left[\mathbf{e}^*\frac{\varepsilon}{\omega_2\mathscr{D}_3}\left(\frac{\kappa_3-\kappa_2}{2}+\frac{\mathbf{e}^*\cdot\mathbf{f}_3}{\mathbf{e}^*\cdot\mathbf{a}}\times(\kappa_2^2+\kappa_3^2)\right)-\mathbf{e}\,\frac{(\mathbf{e}^*\cdot\mathbf{b})(\kappa_2^2+\kappa_3^2)}{2(\mathbf{e}\cdot\mathbf{a})\omega_1\mathscr{D}_1}\right]+\left(\begin{matrix}\omega_2\leftrightarrow\omega_3\\ \mathbf{f}_2\leftrightarrow\mathbf{f}_3\end{matrix}\right),$$

$$\mathbf{T}_{++-}=\omega_3\int_0^{\omega_2}d\varepsilon\,\kappa_2\left[\mathbf{e}\,\frac{\kappa_3}{\omega_1\mathscr{D}_1}\left(\frac{\kappa_2-\varepsilon}{2}-\frac{\mathbf{e}\cdot\mathbf{f}_{23}}{\mathbf{e}\cdot\mathbf{a}}\times(\kappa_2^2+\varepsilon^2)\right)+\mathbf{e}^*\frac{(\mathbf{e}\cdot\mathbf{c})(\kappa_2^2+\varepsilon^2)}{2(\mathbf{e}^*\cdot\mathbf{a})\omega_2\mathscr{D}_3}\right]+\omega_2\mathbf{e}\int_{-\omega_3}^{0}d\varepsilon\,\kappa_3\left[\frac{\kappa_3(\kappa_2^2+\varepsilon^2)}{\mathbf{e}^*\cdot\mathbf{b}}\left(\frac{\mathbf{e}^*\cdot\mathbf{f}_{23}}{\omega_1\mathscr{D}_1}+\frac{\mathbf{e}^*\cdot\mathbf{f}_2}{\omega_3\mathscr{D}_2}\right)-\frac{\varepsilon\kappa_3+\kappa_2^2}{2\omega_1\mathscr{D}_1}+\frac{\varepsilon^2-\kappa_2\kappa_3}{2\omega_3\mathscr{D}_2}\right].$$

It is clear that $\mathbf{T}_{+-+}$ can be obtained from $\mathbf{T}_{++-}$ via the replacement $\omega_2\leftrightarrow\omega_3$, $\mathbf{f}_2\leftrightarrow\mathbf{f}_3$. In Eq. (24) we introduced the notation

$$\mathscr{D}_1=\frac{\omega_2\kappa_3\mathbf{a}^2-\omega_3\kappa_2\mathbf{b}^2}{\omega_1\varepsilon}-i0,$$

$$\mathscr{D}_2=\frac{\omega_2\kappa_3\widetilde{\mathbf{c}}^2-\omega_1\varepsilon\mathbf{b}^2}{\omega_3\kappa_2},\quad \mathscr{D}_3=\frac{\omega_1\varepsilon\mathbf{a}^2+\omega_3\kappa_2\mathbf{c}^2}{\omega_2\kappa_3},$$

$$\mathbf{a}=\mathbf{q}-\mathbf{\Delta}+2\kappa_2\mathbf{f}_2,\quad \mathbf{b}=\mathbf{q}+\mathbf{\Delta}-2\kappa_3\mathbf{f}_3,$$

$$\mathbf{c}=\mathbf{q}+\mathbf{\Delta}-2\varepsilon\mathbf{f}_{23},\quad \widetilde{\mathbf{c}}=\mathbf{q}-\mathbf{\Delta}+2\varepsilon\mathbf{f}_{23}. \tag{25}$$

In deriving (24) we used the identity

$$\mathbf{Q}\exp\left(-\frac{i}{2}\mathbf{q}\cdot\mathbf{Q}\right)=2i\boldsymbol{\nabla}_{\mathbf{q}}\exp\left(-\frac{i}{2}\mathbf{q}\cdot\mathbf{Q}\right)$$

and integrated by parts over $\mathbf{q}$. Note that in (24) the vectors $\mathbf{e}$ and $\mathbf{e}^*$ turned out to be in the denominators by virtue of the relation $2(\mathbf{e}\cdot\mathbf{a})(\mathbf{e}^*\cdot\mathbf{a})=\mathbf{a}^2$.

7. ASYMPTOTE FOR SMALL Δ

In the small-angle approximation ($|\mathbf{f}_2|,|\mathbf{f}_3|\ll1$) the cross section of the photon splitting process has the form

$$d\sigma=\frac{\omega_1^2}{2^8\pi^5}|M|^2x(1-x)dx\;d\mathbf{f}_2\;d\mathbf{f}_3, \tag{26}$$

where $x=\omega_2/\omega_1$, and $\omega_3=\omega_1(1-x)$. We introduce the notation $\boldsymbol{\rho}=(\omega_2\mathbf{f}_2-\omega_3\mathbf{f}_3)/2$. In the variables $\boldsymbol{\rho}$ and $\mathbf{\Delta}$ the cross section of the process is

$$d\sigma=|M|^2\frac{d\mathbf{\Delta}\;d\boldsymbol{\rho}\;dx}{2^8\pi^5\omega_1^2x(1-x)}. \tag{27}$$

Let us consider the asymptote of the expressions obtained for $|\mathbf{\Delta}|\ll|\boldsymbol{\rho}|$. To obtain it we multiply $\mathbf{T}$ in (24) by

$$1=\vartheta(q_0^2-\mathbf{q}^2)+\vartheta(\mathbf{q}^2-q_0^2),$$

where $|\mathbf{\Delta}|\ll q_0\ll|\boldsymbol{\rho}|$. Then for the term in (24) which is proportional to $\vartheta(q_0^2-\mathbf{q}^2)$ we set $\mathbf{q}=0$ and $\mathbf{\Delta}=0$ in $\mathbf{T}$ and integrate by parts over $\mathbf{q}$. Using the relation

$$\boldsymbol{\nabla}_{\mathbf{q}}\vartheta(q_0^2-\mathbf{q}^2)=-2\mathbf{q}\delta(q_0^2-\mathbf{q}^2),$$

we can easily calculate the integral over $\mathbf{q}$, so that for $|\mathbf{q}|=q_0\gg|\mathbf{\Delta}|$ we have

$$\mathrm{Im}\left(\frac{|\mathbf{q}+\mathbf{\Delta}|}{|\mathbf{q}-\mathbf{\Delta}|}\right)^{2iZ\alpha}\approx4Z\alpha\frac{\mathbf{q}\cdot\mathbf{\Delta}}{\mathbf{q}^2}.$$

As a result, at $|\mathbf{q}|<q_0$ the contribution proportional to $Z\alpha$ does not depend on q_0, and the next terms with respect to $Z\alpha$ are small with respect to $|\Delta|/q_0$.

For the term proportional to $\vartheta(\mathbf{q}^2-q_0^2)$ we have

$$\boldsymbol{\nabla}_{\mathbf{q}}\,\mathrm{Im}\left(\frac{|\mathbf{q}+\mathbf{\Delta}|}{|\mathbf{q}-\mathbf{\Delta}|}\right)^{2iZ\alpha}\approx4Z\alpha\frac{\mathbf{q}^2\mathbf{\Delta}-2\mathbf{q}(\mathbf{q}\cdot\mathbf{\Delta})}{|\mathbf{q}|^4}.$$

Setting $\boldsymbol{\Delta}=0$ in $\mathbf{T}$, we first calculate the integral over the angles of the vector $\mathbf{q}$ and then over its magnitude. As a result, the leading term with respect to $q_0/|\boldsymbol{\rho}|$ does not depend on q_0, and is proportional to $Z\alpha$. We take the sum of the contributions of the two regions for $\mathbf{q}$ and calculate the integral over the energy ε. We ultimately obtain

$$M_{+--}=\frac{4iN(\mathbf{e}\cdot\boldsymbol{\rho})^3}{\boldsymbol{\rho}^4}(\boldsymbol{\Delta}\times\boldsymbol{\rho})_z, \quad N=\frac{4Z\alpha e^3\omega_2\omega_3}{\pi\omega_1\boldsymbol{\Delta}^2\boldsymbol{\rho}^2},$$

$$M_{+++}=N\Bigg[\mathbf{e}^*\cdot\boldsymbol{\Delta}+2(\mathbf{e}\cdot\boldsymbol{\Delta})\frac{(\mathbf{e}^*\cdot\boldsymbol{\rho})^2}{\boldsymbol{\rho}^2}\times\Bigg(1+\frac{\omega_2-\omega_3}{\omega_1}\ln\frac{\omega_3}{\omega_2}+\frac{\omega_2^2+\omega_3^2}{2\omega_1^2}\times\Bigg(\ln^2\frac{\omega_3}{\omega_2}+\pi^2\Bigg)\Bigg)\Bigg],$$

$$M_{++-}=N\Bigg[\mathbf{e}\cdot\boldsymbol{\Delta}+2(\mathbf{e}^*\cdot\boldsymbol{\Delta})\frac{(\mathbf{e}\cdot\boldsymbol{\rho})^2}{\boldsymbol{\rho}^2}\times\Bigg(1+\frac{\omega_1+\omega_3}{\omega_2}\Bigg(\ln\frac{\omega_3}{\omega_1}+i\pi\Bigg)+\frac{\omega_1^2+\omega_3^2}{2\omega_2^2}\Bigg(\ln^2\frac{\omega_3}{\omega_1}+2i\pi\ln\frac{\omega_3}{\omega_1}\Bigg)\Bigg)\Bigg], \tag{28}$$

where A_z is the component of $\mathbf{A}$ along $\mathbf{k}_1$. Substituting (28) into (27) and taking the elementary integrals over the angles of the vectors $\boldsymbol{\Delta}$ and $\boldsymbol{\rho}$, we arrive at the expression

$$d\sigma=\frac{4Z^2\alpha^5}{\pi^2}\frac{d\rho^2\ d\Delta^2\ dx}{\rho^4\Delta^2}g(x), \tag{29}$$

where the functions $g(x)$ for the various polarizations have the form

$$g_{+--}(x)=x(1-x),$$

$$g_{+++}(x)=\frac{1}{2}x(1-x)\Bigg[1+\Bigg(1+(2x-1)\ln\frac{1-x}{x}+\frac{x^2+(1-x)^2}{2}\Bigg(\ln^2\frac{1-x}{x}+\pi^2\Bigg)\Bigg)^2\Bigg],$$

$$g_{++-}(x)=\frac{1}{2}x(1-x)\Bigg[1+\Bigg|1+\Bigg(\frac{2}{x}-1\Bigg)\times(\ln(1-x)+i\pi)+\frac{1+(1-x)^2}{2x^2}(\ln^2(1-x)+2i\pi\ln(1-x))\Bigg|^2\Bigg], \tag{30}$$

$$g_{+-+}(x)=g_{++-}(1-x).$$

Equations (29) and (30) are consistent with the corresponding results in Ref. 16, which were obtained using the equivalent-photon approximation. However, this method does not enable us to obtain the amplitudes (28) themselves. A large logarithm appears in the cross section as a result of the integration of (29) over Δ^2 from $\Delta^2_{\min}$ to ρ^2, where $\Delta_{\min}\sim r_c^{-1}$ for a screened Coulomb potential and $\Delta_{\min}\sim\rho^2/\omega_1$ for an unscreened Coulomb potential.

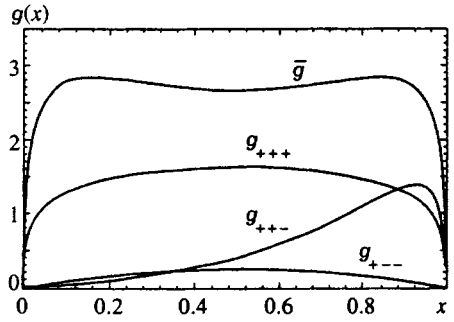


FIG. 3. The functions $g(x)$ (Eq. (30)) for various polarizations, and $\overline{g}(x)$ of Eq. (31).

It is of interest to compare the contributions of the various helical amplitudes to the cross section of the process when $\Delta\to 0$. Figure 3 shows the function $g(x)$ for the various helical states, along with the function

$$\overline{g}(x)=g_{+--}(x)+g_{+++}(x)+g_{++-}(x)+g_{+-+}(x), \tag{31}$$

which results from the summation over the polarizations of the final photons. It is clear that $\overline{g}(x)$ varies weakly over a broad range of x.

When $\Delta\to 0$, the Coulomb corrections to the photon splitting amplitude are small compared to the Born contribution (28). They become significant when $\Delta\sim\rho$, and a separate detailed analysis of their role is required.

8. THE BORN APPROXIMATION

As stated earlier, the photon splitting amplitude in Refs. 14 and 15 in the lowest Born approximation for arbitrary energy and momentum transfer is very complicated for applications. It is therefore of interest to find the term in the expansion of the amplitude (24) that is linear in $Z\alpha$. For this purpose, in (24) we make the substitution

$$(\mathbf{e}\cdot\boldsymbol{\nabla}_{\mathbf{q}})\mathrm{Im}\Bigg(\frac{|\mathbf{q}+\boldsymbol{\Delta}|}{|\mathbf{q}-\boldsymbol{\Delta}|}\Bigg)^{2iZ\alpha}\to Z\alpha\Bigg[\frac{1}{\mathbf{e}^*\cdot(\mathbf{q}+\boldsymbol{\Delta})}-\frac{1}{\mathbf{e}^*\cdot(\mathbf{q}-\boldsymbol{\Delta})}\Bigg]$$

and write the functions $\mathscr{D}_{1,2,3}$ from (25) in the form

$$\mathscr{D}_1=\Bigg(\mathbf{q}+\frac{\kappa_2-\kappa_3}{\omega_1}\boldsymbol{\Delta}\Bigg)^2-\frac{4\omega_2\omega_3\kappa_2\kappa_3}{\omega_1^2}\mathbf{f}_{23}^2-i0,$$

$$\mathscr{D}_2=\Bigg(\mathbf{q}-\frac{\kappa_3+\varepsilon}{\omega_3}\boldsymbol{\Delta}\Bigg)^2-\frac{4\omega_1\omega_2\kappa_3\varepsilon}{\omega_3^2}\mathbf{f}_2^2,$$

$$\mathscr{D}_3=\Bigg(\mathbf{q}+\frac{\kappa_2-\varepsilon}{\omega_2}\boldsymbol{\Delta}\Bigg)^2+\frac{4\omega_1\omega_3\kappa_2\varepsilon}{\omega_2^2}\mathbf{f}_3^2. \tag{32}$$

In each term we shift the integration variable $\mathbf{q}$ so that the functions $\mathscr{D}_{1,2,3}$ cease to depend on the angle ϕ of the vector $\mathbf{q}$. For example, in the terms containing $\mathscr{D}_1$ we make the substitution

$$\mathbf{q}\to\mathbf{q}-\frac{\kappa_2-\kappa_3}{\omega_1}\,\boldsymbol{\Delta}.$$

As a result, passing to the variable $z=\exp(i\phi)$, we easily take the integral over z using residues. Calculating the elementary integrals over $|\mathbf{q}|$ and over ε, for the Born amplitudes we obtain

$$M_{+--}=\frac{2iZ\alpha e^3(\mathbf{f}_2\times\mathbf{f}_3)_z}{\pi\boldsymbol{\Delta}^2(\mathbf{e}^*\cdot\mathbf{f}_2)(\mathbf{e}^*\cdot\mathbf{f}_3)(\mathbf{e}^*\cdot\mathbf{f}_{23})},$$

$$\begin{aligned}M_{+++}=&\frac{2(Z\alpha)e^3\omega_1}{\pi\boldsymbol{\Delta}^2(\mathbf{e}\cdot\mathbf{f}_{23})^2\omega_2\omega_3}\\&\times\Bigg\{(\mathbf{e}\cdot\boldsymbol{\Delta})\Bigg[1+\frac{\mathbf{e}\cdot\mathbf{f}_2+\mathbf{e}\cdot\mathbf{f}_3}{\mathbf{e}\cdot\mathbf{f}_{23}}\ln\frac{a_2}{a_3}\\&+\frac{(\mathbf{e}\cdot\mathbf{f}_2)^2+(\mathbf{e}\cdot\mathbf{f}_3)^2}{(\mathbf{e}\cdot\mathbf{f}_{23})^2}\Bigg(\frac{\pi^2}{6}+\frac{1}{2}\ln^2\frac{a_2}{a_3}\\&+\mathrm{Li}_2(1-a_2)+\mathrm{Li}_2(1-a_3)\Bigg)\Bigg]\\&+\frac{1}{\mathbf{e}\cdot\boldsymbol{\Delta}}\Bigg[\omega_3^2(\mathbf{e}\cdot\mathbf{f}_3)^2\frac{a_2}{1-a_2}\Bigg(1+\frac{a_2\ln a_2}{1-a_2}\Bigg)\\&+\omega_2^2(\mathbf{e}\cdot\mathbf{f}_2)^2\frac{a_3}{1-a_3}\Bigg(1+\frac{a_3\ln a_3}{1-a_3}\Bigg)\Bigg]\\&+\frac{2(\mathbf{e}\cdot\mathbf{f}_2)(\mathbf{e}\cdot\mathbf{f}_3)}{\mathbf{e}\cdot\mathbf{f}_{23}}\Bigg(\omega_3\frac{a_2\ln a_2}{1-a_2}-\omega_2\frac{a_3\ln a_3}{1-a_3}\Bigg)\Bigg\},\end{aligned}$$

$$\begin{aligned}M_{++-}=&\frac{2(Z\alpha)e^3\omega_2}{\pi\boldsymbol{\Delta}^2(\mathbf{e}^*\cdot\mathbf{f}_3)^2\omega_1\omega_3}\\&\times\Bigg\{(\mathbf{e}^*\cdot\boldsymbol{\Delta})\Bigg[1-\frac{(\mathbf{e}^*\cdot\mathbf{f}_2)+(\mathbf{e}^*\cdot\mathbf{f}_{23})}{(\mathbf{e}^*\cdot\mathbf{f}_3)}\ln\frac{-a_1}{a_2}\\&+\frac{(\mathbf{e}^*\cdot\mathbf{f}_2)^2+(\mathbf{e}^*\cdot\mathbf{f}_{23})^2}{(\mathbf{e}^*\cdot\mathbf{f}_3)^2}\Bigg(\frac{\pi^2}{6}+\frac{1}{2}\ln^2\frac{-a_1}{a_2}\\&+\mathrm{Li}_2(1-a_2)+\mathrm{Li}_2(1+a_1)\Bigg)\Bigg]\\&+\frac{1}{\mathbf{e}^*\cdot\boldsymbol{\Delta}}\Bigg[\omega_3^2(\mathbf{e}^*\cdot\mathbf{f}_{23})^2\frac{a_2}{1-a_2}\Bigg(1+\frac{a_2\ln a_2}{1-a_2}\Bigg)\\&-\omega_1^2(\mathbf{e}^*\cdot\mathbf{f}_2)^2\frac{a_1}{1+a_1}\Bigg(1-\frac{a_1\ln(-a_1)}{1+a_1}\Bigg)\Bigg]\\&+\frac{2(\mathbf{e}^*\cdot\mathbf{f}_2)(\mathbf{e}^*\cdot\mathbf{f}_{23})}{(\mathbf{e}^*\cdot\mathbf{f}_3)}\Bigg[\omega_1\frac{a_1\ln(-a_1)}{1+a_1}\\&-\omega_3\frac{a_2\ln a_2}{1-a_2}\Bigg]\Bigg\},\end{aligned}\tag{33}$$

where

$$a_1=\frac{\boldsymbol{\Delta}^2}{\omega_2\omega_3\mathbf{f}_{23}^2},\quad a_2=\frac{\boldsymbol{\Delta}^2}{\omega_1\omega_2\mathbf{f}_2^2},$$

$$a_3=\frac{\boldsymbol{\Delta}^2}{\omega_1\omega_3\mathbf{f}_3^2},\quad \mathrm{Li}_2(x)=-\int_0^x\frac{dt}{t}\ln(1-t).$$

It follows from Eq. (25) that $\ln(-a_1)$ should be understood as $\ln(-a_1+i0)=\ln a_1+i\pi$. In addition,

$$\begin{aligned}\mathrm{Li}_2(1+a_1)&=\mathrm{Li}_2(1+a_1-i0)\\&=\frac{\pi^2}{6}-\ln(1+a_1)[\ln a_1+i\pi]-\mathrm{Li}_2(-a_1).\end{aligned}$$

The result (33) was obtained under the condition $|\boldsymbol{\Delta}_\perp|\gg|\boldsymbol{\Delta}_\parallel|$. It can be shown that it also remains true in the case of $|\boldsymbol{\Delta}_\perp|\sim|\boldsymbol{\Delta}_\parallel|$, if the expression (5) for $\boldsymbol{\Delta}^2$ is used in Eqs. (33). Actually, in (33) the difference between $\boldsymbol{\Delta}^2$ and $\boldsymbol{\Delta}_\perp^2$ is significant only in the outermost factor $1/\boldsymbol{\Delta}^2$. For a screened Coulomb potential the amplitudes (33) should be multiplied by the atomic form factor $1-F(\Delta^2)$. In the case of the Molière potential[19] this factor is

$$1-F(\Delta^2)=\Delta^2\sum_{i=1}^{3}\frac{\alpha_i}{\Delta^2+\beta_i^2},\tag{34}$$

where

$$\alpha_1=0.1,\quad \alpha_2=0.55,\quad \alpha_3=0.35,\quad \beta_i=\beta_0 b_i,$$

$$b_1=6,\quad b_2=1.2,\quad b_3=0.3,\quad \beta_0=mZ^{1/3}/121.\tag{35}$$

Recall that Eq. (33) holds for $|\mathbf{k}_{2\perp}|,|\mathbf{k}_{3\perp}|\gg m$.

9. CROSS SECTION OF THE PROCESS

It was proposed in Ref. 13 that to overcome the difficulties associated with the background in observing photon splitting, events must be recorded with $|\mathbf{f}_{2,3}|\geqslant f_0$, where the angle $f_0\ll 1$ is determined by the experimental conditions. Let us consider the cross section of the process integrated over $\mathbf{f}_3$ in the region $|\mathbf{f}_3|>f_0$. It is interesting to compare the result for this cross section $(d\sigma/dx\,d\mathbf{f}_2)$ obtained from (33) and (26) with the cross section found using the equivalent-photon approximation $(d\sigma_{\mathrm{approx}}/dx\,d\mathbf{f}_2)$. The large logarithm corresponds to the contribution of the region $\Delta\ll\rho=|\omega_2\mathbf{f}_2-\omega_3\mathbf{f}_3)/2|$, where $f_3\approx xf_2/(1-x)$. Taking the integral over Δ^2 in (29) from $\Delta^2_{\min}$ to Δ^2_{eff}, where (see Ref. 16)

$$\Delta^2_{\min}=\Delta^2_\parallel=[\omega_1 f_2^2 x/2(1-x)]^2,\quad \Delta^2_{\mathrm{eff}}=\rho^2=(\omega_1 x f_2)^2,$$

and summing over the polarizations of the final photons, in an unscreened Coulomb potential we obtain

$$\frac{d\sigma_{\mathrm{approx}}}{dx\,d\mathbf{f}_2}=\frac{8Z^2\alpha^5}{\pi^3\omega_1^2}\frac{\overline{g(x)}}{x^2f_2^4}\ln\frac{2(1-x)}{f^2}\,\vartheta\left(\frac{x}{1-x}f_2-f_0\right).\tag{36}$$

The approximate cross section for a screened potential has the form

$$\begin{aligned}\frac{d\sigma_{\mathrm{approx}}}{dx\,d\mathbf{f}_2}=&\frac{4Z^2\alpha^5}{\pi^3\omega_1^2}\frac{\overline{g(x)}}{x^2f_2^4}\left(2\ln\frac{\omega_1 x f_2}{\beta_0}+\gamma\right)\\&\times\vartheta\left(\frac{x}{1-x}f_2-f_0\right).\end{aligned}\tag{37}$$

The function γ in (37) is

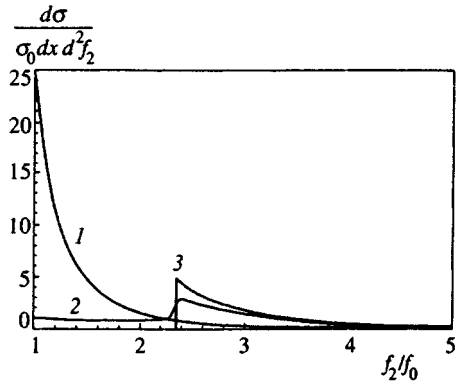


FIG. 4. Dependence of $\sigma_0^{-1} d\sigma/dx\, d\mathbf{f}_2$ on f_2/f_0 for an unscreened potential, $f_0=10^{-3}$, $x=0.7$ (curve *1*), and $x=0.3$ (curve *2*). The quantity σ_0 is defined in the text. Curve *3* corresponds to the equivalent-photon approximation and $x=0.3$.

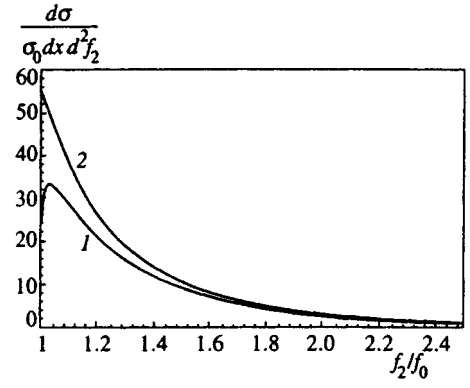


FIG. 5. Same as Fig. 4, but for $x=0.5$ (curve *1*). Curve *2* corresponds to the equivalent-photon approximation.

$$\gamma=1-\sum_{i=1}^{3}\alpha_i^2(\ln a_i+1)-2\sum_{i>j}\alpha_i\alpha_j\frac{a_i\ln a_i-a_j\ln a_j}{a_i-a_j},$$

$$a_i=b_i^2+\frac{\Delta_{\min}^2}{\beta_0^2}, \tag{38}$$

and the coefficients α_i, b_i, and β_0 are defined in (35). If $\Delta_{\min}^2/\beta_0^2 \gg 1$, then $\gamma=-\ln(\Delta_{\min}^2/\beta_0^2)$, and expression (37) transforms into (36). If $\Delta_{\min}^2/\beta_0^2 \ll 1$, then $\gamma=-0.158$.

The dependence of $\sigma_0^{-1} d\sigma/dx\, d\mathbf{f}_2$ on f_2/f_0 for an unscreened potential is shown in Fig. 4 for $f_0=10^{-3}$ and $x=0.7$ (curve *1*), as well as $x=0.3$ (curve *2*), where

$$\sigma_0=\frac{4Z^2\alpha^5\overline{g}(x)}{\pi^3\omega_1^2 f_0^4},$$

and $\overline{g}(x)$ is defined in (31). When $x=0.7$, the curve corresponding to the cross section (36) is essentially identical to curve *1*. When $x=0.3$ the difference between the approximate and exact cross sections is quite significant (in Fig. 4 curve *3* corresponds to the cross section found in the equivalent-photon approximation for $x=0.3$). However, when $x=0.3$, the cross section $d\sigma/dx$ agrees to high accuracy with the cross section found from (36). This is because $d\sigma/dx$ does not vary under the replacement $x\to 1-x$, and when $x=0.7$, as we have already noted, the approximate result (36) agrees well with the exact result.

When $x=0.5$, there is an appreciable difference between the exact and approximate results (see Fig. 5). This is related to the fact that a large logarithm results from the integration over $\mathbf{f}_3$ in the interval $|(1-x)\mathbf{f}_3+x\mathbf{f}_2| \ll x f_2$. After integrating over the azimuthal angle φ between $\mathbf{f}_2$ and $-\mathbf{f}_3$ we should integrate over f_3 from f_0 to $xf_2/(1-x)$ and from $xf_2/(1-x)$ to infinity. If $xf_2/(1-x)\approx f_0$, the contribution of the first interval vanishes, and the cross section becomes approximately half the cross section (36), as can be seen in Fig. 5.

Since the amplitudes (33) were obtained in the zero-mass approximation, it is of interest to estimate the accuracy of this approximation. Numerical results for the cross section $d\sigma/dx\, d\mathbf{f}_2$ were obtained in Ref. 17 (see Table V in Ref. 17) for $Z=79$, $x=0.87$, $\omega_1=1.7$ GeV, 3.4 GeV, and 6.1 GeV, and five values of the angle f_2 in the range 1.2–2.8 mrad. In these calculations the electron mass was taken into account exactly. The data presented in Table I demonstrate the good agreement between our results and the results in Ref. 17. Only at the point corresponding to the smallest value of the transverse component of the momentum $k_{2\perp}=1.77$ MeV does the error reach 7%.

In the case of an unscreened potential, the cross section $d\sigma/dx$ can be approximated to high accuracy by the formula

$$\frac{d\sigma_{\rm Coul}}{dx}=\pi f_0^2\sigma_0\left[\frac{\vartheta(x-1/2)}{x^2}\left(2\ln\frac{2(1-x)}{f_0}-1-F(x)\right)+(x\leftrightarrow 1-x)\right], \tag{39}$$

where

TABLE I. Values of the photon splitting cross section $d\sigma/\omega_1 dx\, d\mathbf{f}_2$ in b/GeV for $Z=79$ and $x=0.87$.

	$\omega_1=1.7$ GeV		$\omega_1=3.4$ GeV		$\omega_1=6.1$ GeV	
f_2 mrad	Present work	Ref. 17	Present work	Ref. 17	Present work	Ref. 17
1.2	22.7	21.1	3.11	3.25	0.57	0.56
1.6	7.5	7.4	0.99	1.03	0.18	0.18
2.0	3.09	3.12	0.40	0.41	0.071	0.072
2.4	1.48	1.51	0.19	0.19	0.034	0.034
2.8	0.79	0.80	0.10	0.10	0.018	0.018

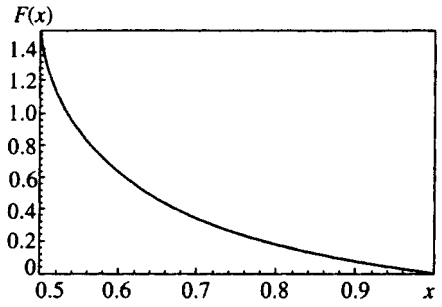


FIG. 6. Plot of $F(x)$ (Eq. (40)).

$$F(x)=\frac{1}{2}+\frac{x}{1-x}+\frac{2x-1}{(1-x)^2}\ln\left(2-\frac{1}{x}\right). \tag{40}$$

If $\omega_1^2 f_0^4/\beta_0^2 \ll 1$, the corresponding expression for the cross section in a screened potential has the form

$$\frac{d\sigma_{\rm scr}}{dx}=\pi f_0^2\sigma_0\left[\frac{\vartheta(x-1/2)}{x^2}\left(2\ln\frac{\omega_1 x f_0}{\beta_0}+0.842-F(x)\right)+(x\leftrightarrow 1-x)\right], \tag{41}$$

where $F(x)$ was defined in (40). The function $F(x)$ characterizes the difference between the exact cross section and the cross section obtained using the equivalent-photon approximation. It is clear from Fig. 6 that this difference becomes significant only for x close to 0.5. It amounts to several percent for total cross sections.

The inequality $\Delta \ll \rho$, which ensures applicability of the equivalent-photon approximation, corresponds to a small angle φ between the vectors $\mathbf{f}_2$ and $-\mathbf{f}_3$ (i.e., the situation in which $\mathbf{f}_2$ and $\mathbf{f}_3$ have almost opposite directions). It is therefore of interest to examine the cross section $d\sigma(\varphi_{\max})/dx$ integrated over φ from $-\varphi_{\max}$ to $\varphi_{\max}$. The dependence of $(\pi f_0^2\sigma_0)^{-1}d\sigma(\varphi_{\max})/dx$ on $\varphi_{\max}$ in an unscreened potential is shown in Fig. 7 for $f_0=10^{-3}$ and various values of x. It can be seen that the cross section approaches its total value at comparatively large values of $\varphi_{\max}$. The same conclusion holds for a screened potential.

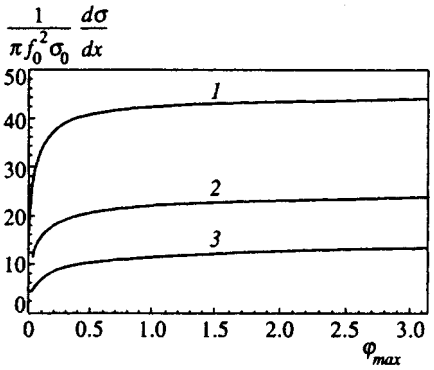


FIG. 7. Dependence of $(\pi f_0^2\sigma_0)^{-1}d\sigma(\varphi_{\max})/dx$ on $\varphi_{\max}$ for an unscreened potential and various values of x: *1*—$x=0.5$, *2*—$x=0.7$, *3*—$x=0.9$; $f_0=10^{-3}$.

One of us (R.N.L.) thanks the International Center for Fundamental Physics in Moscow for supporting this work (INTAS Grant 93-2492-ext).

[1] L. Meitner and H. Kösters, with additional remarks by M. Delbrück, Z. Phys. **84**, 137 (1933).
[2] A. I. Milstein and M. Schumacher, Phys. Rep. **243**, 183 (1994).
[3] M. Cheng and T. T. Wu, Phys. Rev. **182**, 1873 (1969).
[4] M. Cheng and T. T. Wu, Phys. Rev. D **2**, 2444 (1970).
[5] M. Cheng and T. T. Wu, Phys. Rev. D **5**, 3077 (1972).
[6] A. I. Milstein and V. M. Strakhovenko, Phys. Lett. A **95**, 135 (1983).
[7] A. I. Mil'shteĭn and V. M. Strakhovenko, Zh. Éksp. Teor. Fiz. **85**, 14 (1983) [Sov. Phys. JETP **58**, 8 (1983)].
[8] A. I. Milstein and V. M. Strakhovenko, Phys. Lett. A **90**, 447 (1982).
[9] R. N. Lee and A. I. Milstein, Phys. Lett. A **198**, 217 (1995).
[10] R. N. Li and A. I. Mil'shteĭn, Zh. Éksp. Teor. Fiz. **107**, 1393 (1995) [JETP **80**, 777 (1995)].
[11] G. Jarlskog, L. Jönsson, S. Prünster, *et al.*, Phys. Rev. D **8**, 3813 (1973).
[12] R. M. Dzhilkibaev, É. A. Kuraev, V. S. Fadin, and V. A. Khoze, JETP Lett. **19**, 47 (1974).
[13] A. I. Milstein and B. B. Wojttsekhowski, Preprint INP 91-14, Novosibirsk (1991).
[14] Y. Shima, Phys. Rev. **142**, 944 (1966).
[15] V. Constantini, B. De Tollis, and G. Pistoni, Nuovo Cimento A **2**, 733 (1971).
[16] V. N. Baier, V. M. Katkov, E. A. Kuraev, and V. S. Fadin, Phys. Lett. B **49**, 385 (1974).
[17] A. M. Johannessen, K. J. Mork, and I. Overbo, Phys. Rev. D **22**, 1051 (1980).
[18] H.-D. Steinhofer, Z. Phys. C **18**, 139 (1983).
[19] G. Molière, Z. Naturforsch. **2A**, 133 (1947).

Translated by P. Shelnitz

Semiclassical quantization of $SU(3)$ skyrmions

V. B. Kopeliovich

Institute for Nuclear Research, Russian Academy of Sciences, 117312 Moscow, Russia
(Submitted 6 May 1997)
Zh. Éksp. Teor. Fiz. **112**, 1941–1958 (December 1997)

Semiclassical quantization of the $SU(3)$-skyrmion zero modes is performed by means of the collective coordinate method. The quantization condition known for $SU(2)$ solitons quantized with $SU(3)$ collective coordinates is generalized for $SU(3)$ skyrmions with strangeness content different from zero. The quantization of the dipole-type configuration with large strangeness content found recently is considered as an example and the spectrum and the mass splittings of the quantized states are estimated. The energy and baryon number density of $SU(3)$ skyrmions are presented in a form emphasizing their symmetry in different $SU(2)$ subgroups of $SU(3)$, and a lower bound for the static energy of $SU(3)$ skyrmions is derived. © *1997 American Institute of Physics.* [S1063-7761(97)00212-6]

1. INTRODUCTION

The chiral soliton approach first proposed by Skyrme[1] allows one to describe the properties of baryons with fairly good accuracy.[2–4] Considerable progress has been made recently also in understanding the properties of few-nucleon systems.[5–7] Moreover, this approach allows some predictions for the spectrum of states with baryon number $B>1$.[8–13] The quantization of the bound states of skyrmions, primarily their zero modes, is a necessary step towards realization of this approach. Different aspects of this problem have been considered, beginning with Refs. 2, 13, 8, 9, and 14. However, a more general treatment allowing the consideration of arbitrary $SU(3)$ skyrmions was lacking until recently.

In the sector with $B=2$ besides the $SO(3)$ hedgehog with the lowest quantum state interpreted as an H-dibaryon[8,9] the $SU(2)$ torus—a bound $B=2$ state—was discovered 10 years ago.[15] In the flavor-symmetric (FS) case, when all meson masses in the Lagrangian are equal to the pion mass, there are three degenerate tori in the (u,d), (d,s) and (u,s) $SU(2)$ subgroups of $SU(3)$. In the flavor-symmetry-broken (FSB) case the (u,s) and (d,s) tori are degenerate and heavier than the (u,d) torus. Another local minimum with large strangeness content was found recently in the $SU(3)$ extension of the model.[16] This configuration is of molecular type and consists of two interacting $B=1$ skyrmions located in different $SU(2)$ subgroups of $SU(3)$, (u,s) and (d,s). The attraction between two $B=1$ skyrmions in optimal orientation which led to the formation of the torus-like state is not sufficient for this when both skyrmions are located in different $SU(2)$ subgroups of $SU(3)$ and interact due only to one common degree of freedom. To find this configuration a special algorithm was developed allowing for the minimization of an energy functional depending on eight functions of three variables.[16] The position of the known $B=2$ classical configurations representing local minima in $SU(3)$ configuration space is shown on Fig. 1 in the plane with the scalar strangeness content C_S (Ref. 17) as Y-axis and the difference of the U- and D-contents as X-axis. Since the sum of all scalar contents is equal to unity, they are defined uniquely at each point of this plot. The $SO(3)$ hedgehog (1) has all contents equal to 1/3.[12] Intuitively this is clear, since the basis for the $SO(3)$ solitons is formed by the matrices λ_2, $-\lambda_5$, λ_7 and they are located in three $SU(2)$ subgroups of $SU(3)$ on equal footing. The three tori in three different $SU(2)$ subgroups of $SU(3)$ are denoted by the labels (2), (3) and (4), the $u-d$ symmetric state (2) with $C_S=0$ being of special interest. The configurations (3) and (4) can be connected by isorotation in the (u,d) subgroup. The dipole type state (5) found recently[16] has a binding energy about half of that of the torus, i.e., about 0.04 of the mass of the $B=1$ skyrmion.

The zero modes of solitons have been quantized previously in a few cases: for $SU(2)$ solitons rotated in the $SU(2)^2$ as well as in the $SU(3)$ configuration space of collective coordinates,[13,8,14] and also for $SO(3)$ solitons.[8,9] In the case of $SU(2)$ solitons rotated in $SU(3)$ space the quantization condition known as the Guadagnini condition[13] was established; see also Ref. 18.

The quantization of the $SU(2)$ $B=1$ hedgehog yields the spectrum of baryons, mainly the octet and decuplet, and moderate agreement with the data has been achieved.[4] Quantization of the $SU(2)$ torus in the $SU(3)$ space of collective coordinates leads to predictions of a rich spectrum of strange dibaryons.[19,11] Most of them are probably unbound if a natural assumption is made concerning the poorly known Casimir energy of the torus-like solitons; see also the discussion in the last section.

However, these solitons are only particular cases, since other types of solitons exist, e.g., the above-mentioned solitons of dipole type with large strangeness content,[16] point (5) on Fig. 1. In general, one should expect that the map of the local minima in the $SU(3)$ configuration space will become more and more complicated as the baryon number of the configuration increases. In some cases the local minima corresponding to larger strangeness content may have greater binding energy than configurations with small or zero C_S. Therefore, a quantization procedure for arbitrary $SU(3)$ solitons should be developed. This is a subject of the present paper (Ref. 20 contains a preliminary short version).

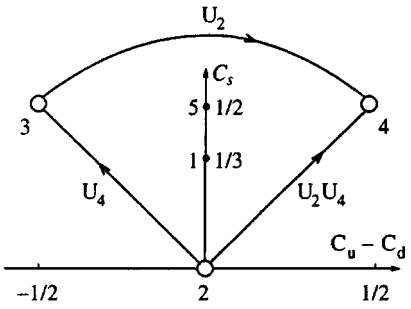


FIG. 1. Map of the different local minima for classical configurations with $B=2$ in the plane (C_u-C_d), C_S. Here C_u, C_d and C_S are the scalar quark contents of the soliton, (1) is the $SO(3)$ hedgehog, (2), (3) and (4) are $SU(2)$ tori in the (u,d), (d,s) and (u,s) subgroups of $SU(3)$, and (5) is the dipole-type configuration (strange skyrmion molecule).

2. THE WESS-ZUMINO-WITTEN TERM

Let us consider the Wess–Zumino (WZ) term in the action which defines the quantum numbers of the system in the quantization procedure. It was written by E. Witten in the elegant form:[21]

$$S^{\rm WZ}=\frac{-iN_c}{240\pi^2}\,\epsilon_{\mu\nu\alpha\beta\gamma}\int_\Omega \mathrm{Tr}\,\tilde{L}_\mu\tilde{L}_\nu\tilde{L}_\alpha\tilde{L}_\beta\tilde{L}_\gamma d^5x', \tag{1}$$

where Ω is the 5-dimensional region with 4-dimensional space–time as its boundary, N_C is the number of colors of the underlying QCD, and $\tilde{L}_\mu=U^\dagger d_\mu U$. As usual, we introduce time-dependent collective coordinates for the quantization of zero modes according to the relation $U(\mathbf{r},t)=A(t)U_0(\mathbf{r})A^\dagger(t)$. Integration by parts is possible then in the expression for the WZ-term in the action, and for the WZ-term contribution to the Lagrangian of the system we obtain

$$L^{\rm WZ}=\frac{-iN_c}{48\pi^2}\,\epsilon_{\alpha\beta\gamma}\int \mathrm{Tr}\,A^\dagger\dot{A}(R_\alpha R_\beta R_\gamma+L_\alpha L_\beta L_\gamma)d^3x, \tag{2}$$

where $L_\alpha=U_0^\dagger d_\alpha U_0=iL_{k,\alpha}\lambda_k$ and $R_\alpha=d_\alpha U_0U_0^\dagger=U_0L_\alpha U_0^\dagger$, or

$$L^{\rm WZ}=\frac{N_c}{24\pi^2}\int\sum_{k=1}^{k=8}\omega_k\mathrm{WZ}_k d^3x=\sum_{k=1}^{k=8}\omega_k L_k^{\rm WZ}, \tag{3}$$

with the angular velocities of rotation in the configuration space defined in the usual way, $A^\dagger\dot{A}=-i\omega_k\lambda_k/2$. Summation over repeated indices is assumed here and below. The functions WZ_k can be expressed through the chiral derivatives $\mathbf{L_k}$:

$$\mathrm{WZ}_i=\mathrm{WZ}_i^R+\mathrm{WZ}_i^L=(R_{ik}(U_0)+\delta_{ik})\mathrm{WZ}_k^L, \tag{4a}$$

$i,k=1,...,8$, and

$$\begin{aligned}
\mathrm{WZ}_1^L&=-(L_1,L_4L_5+L_6L_7)-(L_2L_3L_8)/\sqrt{3}\\
&\quad-2(L_8,L_4L_7-L_5L_6)/\sqrt{3},\\
\mathrm{WZ}_2^L&=-(L_2,L_4L_5+L_6L_7)-(L_3L_1L_8)/\sqrt{3}\\
&\quad-2(L_8,L_4L_6+L_5L_7)/\sqrt{3},\\
\mathrm{WZ}_3^L&=-(L_3,L_4L_5+L_6L_7)-(L_1L_2L_8)/\sqrt{3}\\
&\quad-2(L_8,L_4L_5-L_6L_7)/\sqrt{3},\\
\mathrm{WZ}_4^L&=-(L_4,L_1L_2-L_6L_7)-(L_3L_5L_8)/\sqrt{3}\\
&\quad+2(\tilde{L}_8,L_1L_7+L_2L_6)/\sqrt{3},\\
\mathrm{WZ}_5^L&=-(L_5,L_1L_2-L_6L_7)+(L_3L_4L_8)/\sqrt{3}\\
&\quad-2(\tilde{L}_8,L_1L_6-L_2L_7)/\sqrt{3},\\
\mathrm{WZ}_6^L&=(L_6,L_1L_2+L_4L_5)+(L_3L_7L_8)/\sqrt{3}-2(\tilde{\tilde{L}}_8,L_1L_5\\
&\quad-L_2L_4)/\sqrt{3},\\
\mathrm{WZ}_7^L&=(L_7,L_1L_2+L_4L_5)-(L_3L_6L_8)/\sqrt{3}+2(\tilde{\tilde{L}}_8,L_1L_4\\
&\quad+L_2L_5)/\sqrt{3},\\
\mathrm{WZ}_8^L&=-\sqrt{3}(L_1L_2L_3)+(L_8L_4L_5)+(L_8L_6L_7),
\end{aligned} \tag{5}$$

where $(L_1L_2L_3)$ denotes the mixed product of vectors $\mathbf{L}_1$, $\mathbf{L}_2$, $\mathbf{L}_3$, etc. and

$$\tilde{L}_3=(L_3+\sqrt{3}L_8)/2,\quad \tilde{L}_8=(\sqrt{3}L_3-L_8)/2,$$

$$\tilde{\tilde{L}}_3=(-L_3+\sqrt{3}L_8)/2,\quad \tilde{\tilde{L}}_8=(\sqrt{3}L_3+L_8)/2$$

are the third and eighth components of the chiral derivatives in the (u,s) and (d,s) $SU(2)$-sub-groups. Here $[\tilde{\mathbf{L}}_3\tilde{\mathbf{L}}_8]=-[\mathbf{L}_3,\mathbf{L}_8]$, etc.,

$$R_{ik}(U_0)=\frac{1}{2}\,\mathrm{Tr}\,\lambda_i U_0\lambda_k U_0^\dagger$$

is a real orthogonal matrix, and WZ_i^R are defined by the expressions (5) with the substitution $\mathbf{L}_k\to\mathbf{R}_k$. Relations similar to (5) can be obtained for $\widetilde{\mathrm{WZ}}_3$ and $\widetilde{\mathrm{WZ}}_8$; they are analogs of WZ_3 and WZ_8 for the (u,s) or (d,s) $SU(2)$-subgroups, thus clarifying the symmetry of the WZ-term in the different $SU(2)$ subgroups of $SU(3)$.

The baryon number of the $SU(3)$ skyrmions can be written also in terms of $\mathbf{L}_i$ in a form where its symmetry in the different $SU(2)$ subgroups of $SU(3)$ is obvious:

$$B=-\frac{1}{2\pi^2}\int\left((\mathbf{L}_1\mathbf{L}_2\mathbf{L}_3)+(\mathbf{L}_4\mathbf{L}_5\tilde{\mathbf{L}}_3)+(\mathbf{L}_6\mathbf{L}_7\tilde{\tilde{\mathbf{L}}}_3)+\frac{1}{2}\left[(\mathbf{L}_1,\mathbf{L}_4\mathbf{L}_7-\mathbf{L}_5\mathbf{L}_6)+(\mathbf{L}_2,\mathbf{L}_4\mathbf{L}_6+\mathbf{L}_5\mathbf{L}_7)\right]\right)d^3r. \tag{6}$$

The contributions of the three $SU(2)$ subgroups enter the baryon number on equal footing. In addition, mixed terms corresponding to the interaction of the chiral fields from different subgroups are present also.

It should be noted that the results of calculating the WZ-term according to (5) depend on the orientation of the soliton in the $SU(3)$ configuration space. When solitons are located in the (u,d) $SU(2)$ subgroup of $SU(3)$, only L_1, L_2 and L_3 are different from zero, WZ_8^R and WZ_8^L are both proportional to the B-number density, and the well known quantization condition of Guadagnini,[13] rederived in Ref. 18,

$$Y_R=\frac{2}{\sqrt{3}}\,dL^{\rm WZ}/d\omega_8=N_cB/3, \tag{7}$$

applies, where Y_R is the so-called right hypercharge characterizing the $SU(3)$ irrep under consideration. This relation is generalized to[20]

$$Y_R^{\min}=\frac{2}{\sqrt{3}}\,dL^{\rm WZ}/d\omega_8\approx\frac{1}{3}\,N_cB(1-3C_S), \tag{8}$$

where the scalar strangeness content C_S is defined in terms of the real parts of the diagonal matrix elements of the matrix U:

$$C_S=\frac{\langle 1-\mathrm{Re}\ U_{33}\rangle}{\langle 3-\mathrm{Re}(U_{11}+U_{22}+U_{33})\rangle}, \tag{9}$$

and $\langle\cdots\rangle$ means averaging or integration over the whole 3-dimensional space.[17] This formula was checked in several cases.

a) One can rotate any $SU(2)$ soliton of the (u,d) subgroup by an arbitrary constant $SU(3)$ matrix containing $U_4=\exp(-i\nu\lambda_4)$. In this case $C_S=(1/2)\sin^2\nu$,[17] and both WZ_8^R, WZ_8^L are proportional to $R_{88}=1-(3/2)\sin^2\nu$. As a result, the relation (8) is fulfilled exactly. Solitons (3) and (4) on Fig. 1 can be obtained from the (u,d) soliton (2) by means of U_4 or U_2U_4 rotations and satisfy relation (8). For example, when the skyrmion is located in the (u,s) $SU(2)$ subgroup of $SU(3)$ we have

$$L^{\rm WZ}(u,s)=-\frac{\sqrt{3}N_cB}{12}\,(\omega_8-\sqrt{3}\omega_3). \tag{10a}$$

For skyrmions in the (d,s) $SU(2)$ subgroup

$$L^{\rm WZ}(d,s)=-\frac{\sqrt{3}N_cB}{12}\,(\omega_8+\sqrt{3}\omega_3). \tag{10b}$$

Since we have $C_S=0.5$ in both cases,[17] relation (8) holds. To derive (10a and 10b) we have noted that if the soliton is located in any $SU(2)$ subgroup of $SU(3)$ the two terms in (2) and (4a) give equal contributions.

b) For the $SO(3)$ hedgehog the strangeness content was calculated previously, $C_S=1/3$,[12,11] and $L^{\rm WZ}=0$ according to Ref. 8, at least for periodic $A(t)$.[9] The standard assumption that the angular velocities are constant corresponds to (quasi)periodic behavior of $A(t)$, so relation (8) is satisfied.

c) We obtained the relation (8) numerically for solitons of the form[16]

$$U=U_L(u,s)U(u,d)U_R(d,s), \tag{11}$$

with $U(u,d)=\exp(ia\lambda_2)\exp(ib\lambda_3)$ and $U_L(u,s)$ and $U_R(d,s)$ being deformed interacting $B=1$ $SU(2)$ hedgehogs. For this ansatz we had for the rotated $SU(3)$ Cartan–Maurer currents[16]

$$\begin{aligned}
&L^r_{1i}=s_as_al_{3i}, && L^r_{2i}=d_ia,\\
&L^r_{3i}=(c_{2a}l_{3i}-r_{3i})/2+d_ib, && L^r_{4i}=c_al_{1i},\\
&L^r_{5i}=c_al_{2i}, && L^r_{6i}=s_al_{1i}+r_{1i}(b),\\
&L^r_{7i}=s_al_{2i}+r_{2i}(b), && L^r_{8i}=\sqrt{3}(l_{3i}+r_{3i})/2
\end{aligned} \tag{12}$$

in terms of the $SU(2)$ Cartan–Maurer currents $l_{k,i}$ and $r_{k,i}$ $(i,k=1,2,3)$ and the functions a and b, with

$$r_1(b)=c_br_1-s_br_2,\quad r_2(b)=c_br_2+s_br_1,$$

$$c_b=\cos b,\quad s_b=\sin b,\quad c_a=\cos a,\ \text{etc.}$$

$$i\mathbf{l}_k\tilde{\tau}_k=U_L^\dagger\mathbf{d}U_L,\quad i\mathbf{r}_k\tilde{\tau}_k=\mathbf{d}U_RU_R^\dagger,\quad k=1,2,3.$$

Here

$$\tilde{U}_L(u,s)=f_0+i\tilde{\tau}_kf_k,\quad \tilde{U}_R(d,s)=q_0+i\tilde{\tau}_kq_k,$$

$$k=1,2,3,$$

$\tilde{\tau}$ and $\tilde{\tau}$ are the Pauli matrices corresponding to the (u,s) and (d,s) $SU(2)$-subgroups, and $f_0^2+\ldots+f_3^2=1$, $q_0^2+\ldots+q_3^2=1$.

$$L_i^r=TL_iT^\dagger,\quad U_0=VT,\quad V=U(u,s)\exp(ia\lambda_2),$$

$$T=\exp(ib\lambda_3)U(d,s).$$

The chirally invariant quantities, B-number density (6), and the second-order and Skyrme term contributions to the static energy have the same form in terms of L_{ki} and L^r_{ki}. The formula (4a) should be written then as

$$\mathrm{WZ}_i=[R_{ik}(V)+R_{ik}(T^\dagger)]\mathrm{WZ}_k^L, \tag{4b}$$

with WZ_k^L given in terms of L_n^r according to (5). In the following we shall omit the label «r» everywhere. Relations (10a, and b) can be checked easily with the help of (4), (5) and (12).

Using (12) and (5) we obtain

$$\begin{aligned}
\mathrm{WZ}_8^L=\frac{\sqrt{3}}{2}\,&((\mathbf{l}_1\mathbf{l}_2\mathbf{l}_3)+(\mathbf{r}_1\mathbf{r}_2\mathbf{r}_3)+s_a[(\mathbf{l}_1\mathbf{r}_2-\mathbf{r}_1\mathbf{l}_2,\mathbf{l}_3+\mathbf{r}_3)\\
&-(\mathbf{d}s_a\mathbf{l}_3,\mathbf{r}_3-2\mathbf{d}b)]).
\end{aligned} \tag{13}$$

It follows from (13) and (4) that at large relative distances, for arbitrary but not overlapping solitons, and for $a=0$, we have

$$\begin{aligned}
Y_R^{\min}&=\frac{2}{\sqrt{3}}\,L_8^{\rm WZ}=\frac{1}{2\sqrt{3}\,\pi^2}\int \mathrm{WZ}_8^Ld^3x\\
&=\frac{1}{4\pi^2}\int[(\mathbf{l}_1\mathbf{l}_2\mathbf{l}_3)+(\mathbf{r}_1\mathbf{r}_2\mathbf{r}_3)]d^3x\\
&=-(B_L+B_R)/2,
\end{aligned} \tag{14}$$

where B_L and B_R are the baryon numbers located in the left (u,s) and right (d,s) $SU(2)$ subgroups of $SU(3)$. Relation (8) holds since $C_S=1/2$ for both (u,s) and (d,s) skyrmions. Equation (14) does not hold in the general case for overlapping solitons, since there is no conservation law for the components of the Wess–Zumino term.

For the strange skyrmion molecule[16] we should calculate (3), (5), (8) with $\mathrm{WZ}_8=(R_{8k}(V)+R_{k8}(T))\mathrm{WZ}_k^L$. The contribution $-(B_L+B_R)/2$ also appears with some additional terms which turn out to be small numerically. We obtained $C_S=0.475$ and $Y_R^{\min}=-0.87$ in the FSB case, so relation (8) is satisfied with good accuracy.

It is natural to assume that (8) is valid with good accuracy for any $SU(3)$ skyrmions. However, corrections to this relation are not excluded by our treatment.

3. ROTATION AND STATIC ENERGY

We start with the well known Lagrangian density of the Skyrme model widely used in the literature since Ref. 2. It depends on the parameters $F_\pi = 186$ MeV (experimental value) and the Skyrme parameter e:

$$L_{Sk} = -\frac{F_\pi^2}{16}\,\mathrm{Tr}\;\tilde{L}_\mu\tilde{L}^\mu + \frac{1}{32e^2}\,\mathrm{Tr}(\tilde{L}_\mu\tilde{L}_\nu - \tilde{L}_\nu\tilde{L}_\mu)^2 + L_M \tag{15}$$

We take $e = 4.12$, close to the value suitable for describing, with a bit more complicated Lagrangian, the mass splittings inside the $SU(3)$ multiplets of baryons.[4] The chiral and flavor-symmetry-breaking mass terms L_M in (15) depending on meson masses will be described in detail in Sec. 4.

The expression for the rotation energy density of the system depending on the angular velocities of rotations in the $SU(3)$ collective coordinate space defined in Sec. 2 can be written in more compact form than previously:[20,16]

$$\begin{aligned} L_{\mathrm{rot}} = \frac{F_\pi^2}{32}\,(\tilde{\omega}_1^2 + \tilde{\omega}_2^2 + \ldots + \tilde{\omega}_8^2) + \frac{1}{16e^2}\Big\{ & (\mathbf{s}_{12} + \mathbf{s}_{45})^2 + (\mathbf{s}_{45} \\ & + \mathbf{s}_{67})^2 + (\mathbf{s}_{67} - \mathbf{s}_{12})^2 + \frac{1}{2}\,((2\mathbf{s}_{13} - \mathbf{s}_{46} - \mathbf{s}_{57})^2 + (2\mathbf{s}_{23} \\ & + \mathbf{s}_{47} - \mathbf{s}_{56})^2 + (2\bar{\mathbf{s}}_{34} + \mathbf{s}_{16} - \mathbf{s}_{27})^2 + (2\bar{\mathbf{s}}_{35} + \mathbf{s}_{17} + \mathbf{s}_{26})^2 \\ & + (2\bar{\mathbf{s}}_{36} + \mathbf{s}_{14} + \mathbf{s}_{25})^2 + (2\bar{\mathbf{s}}_{37} + \mathbf{s}_{15} - \mathbf{s}_{24})^2)\Big\}. \end{aligned} \tag{16}$$

Here $\mathbf{s}_{ik} = \tilde{\omega}_i\mathbf{L}_k - \tilde{\omega}_k\mathbf{L}_i$, $i,k = 1,2...8$ are the $SU(3)$ indices, and $\bar{\mathbf{s}}_{34} = (\mathbf{s}_{34} + \sqrt{3}\mathbf{s}_{84})/2$, $\bar{\mathbf{s}}_{35} = (\mathbf{s}_{35} + \sqrt{3}\mathbf{s}_{85})/2$, $\bar{\mathbf{s}}_{36} = (-\mathbf{s}_{36} + \sqrt{3}\mathbf{s}_{86})/2$, $\bar{\mathbf{s}}_{37} = (-\mathbf{s}_{37} + \sqrt{3}\mathbf{s}_{87})/2$, similar to $\tilde{L}_3$ and $\tilde{L}_8$.

To get (16) we used the identity $\mathbf{s}_{ab}\mathbf{s}_{cd} - \mathbf{s}_{ad}\mathbf{s}_{cb} = \mathbf{s}_{ac}\mathbf{s}_{bd}$. The formula (16) possesses remarkable symmetry relative to the different $SU(2)$ subgroups of $SU(3)$. The functions L_8 or $\tilde{L}_8$ do not enter (16) as well as expression (6) for the baryon number density. The functions $\tilde{\omega}_i$ are connected with the body-fixed angular velocities of $SU(3)$ rotations by means of transformation (see (8) above)

$$\tilde{\omega} = V^\dagger\omega V - T\omega T^\dagger, \tag{17a}$$

or

$$\tilde{\omega}_i = (R_{ik}(V^\dagger) - R_{ik}(T))\,\omega_k = R_{ki}\omega_k\,. \tag{17b}$$

$R_{ik}(V^\dagger) = R_{ki}(V)$ and $R_{ik}(T)$ are real orthogonal matrices, $i,k = 1,\cdots,8$, and $\tilde{\omega}_i^2 = 2(\omega_i^2 - R_{kl}(U_0)\omega_k\omega_l)$. Expressions for R_{ik} are given in the Appendix for the general case of the parametrization (11). Relations (17) hold just because we are operating with rotated functions L_{ki}^r in (12).

The expression for static energy can be obtained from (16) by means of the substitution $\tilde{\omega}_i \to 2L_i$ and $\mathbf{s}_{ik} \to 2[\mathbf{L}_i\mathbf{L}_k]$.[16] It can be written in a form which emphasizes quite clearly the lower boundary for the static energy proportional to the winding (baryon) number of the system:

$$\begin{aligned} E_{\mathrm{stat}} = \int \Big\{ \frac{F\pi}{8e}\Big[& (\mathbf{L}_1 - 2\mathbf{n}_{23} - \mathbf{n}_{47} + \mathbf{n}_{56})^2 + (\mathbf{L}_2 - 2\mathbf{n}_{31} \\ & - \mathbf{n}_{46} - \mathbf{n}_{57})^2 + (\mathbf{L}_4 - 2\mathbf{n}_{53} + \mathbf{n}_{17} + \mathbf{n}_{26})^2 + (\mathbf{L}_5 \\ & - 2\mathbf{n}_{34} - \mathbf{n}_{16} + \mathbf{n}_{27})^2 + (\mathbf{L}_6 - 2\mathbf{n}_{73} + \mathbf{n}_{15} - \mathbf{n}_{24})^2 \\ & + (\mathbf{L}_7 - 2\mathbf{n}_{36} - \mathbf{n}_{14} - \mathbf{n}_{25})^2 + \frac{2}{9}\,([\mathbf{L}_3 + \tilde{\mathbf{L}}_3 - 3(\mathbf{n}_{12} \\ & + \mathbf{n}_{45})]^2 + [\tilde{\mathbf{L}}_3 + \tilde{\tilde{\mathbf{L}}}_3 - 3(\mathbf{n}_{45} + \mathbf{n}_{67})]^2 + [\tilde{\tilde{\mathbf{L}}}_3 - \mathbf{L}_3 \\ & - 3(\mathbf{n}_{67} - \mathbf{n}_{12})]^2)\Big] + \mathrm{M.t.} + 3\pi^2\,\frac{F\pi}{e}\,\tilde{B}\Big\}\,d^3\tilde{r}, \end{aligned} \tag{18}$$

where $\tilde{B}$ is the baryon number density given by the integrand in (6), $\tilde{r} = F_\pi e r$ and $\mathbf{n}_{ik} = [\mathbf{L}_i\mathbf{L}_k]$. For $i = 4,5$, $k = 3$ $\tilde{L}_3$ should be taken in $\mathbf{n}_{i3}$. For $i = 6$, $7\tilde{\tilde{L}}_3$ should be taken. In (18) we used relations $\tilde{L}_3 - \tilde{\tilde{L}}_3 = L_3$ and

$$(\mathbf{L}_3 + \tilde{\mathbf{L}}_3)^2 + (\tilde{\mathbf{L}}_3 + \tilde{\tilde{\mathbf{L}}}_3)^2 + (\tilde{\tilde{\mathbf{L}}}_3 - \mathbf{L}_3)^2 = \frac{9}{2}\,(\mathbf{L}_3^2 + \mathbf{L}_8^2).$$

The chiral- and flavor-symmetry-breaking mass term M.t. will be considered in Sec. 4.

From (18) we have the inequality

$$E_{\mathrm{stat}} - \overline{\mathrm{M.t.}} \geqslant 3\pi^2\,\frac{F\pi}{e}\,B. \tag{19}$$

This inequality was obtained first by Skyrme[1] for the $SU(2)$ model and is a particular case of the Bogomol'ny-type bound.

Eight diagonal moments of inertia and 28 off-diagonal ones define the rotation energy, a quadratic form in $\omega_i\omega_k$, according to (16), (17). The analytical expressions for the moments of inertia are too lengthy to be reproduced here. Fortunately, it is possible to perform calculations without explicit analytical formulas, by substituting (17) into (16).

The expression for E_{rot} simplifies considerably when the (u,d) $SU(2)$ soliton is quantized in the $SU(3)$ space of collective coordinates:

$$\begin{aligned} L_{\mathrm{rot}}(SU_2) = \frac{F_\pi^2}{32}\,(\tilde{\omega}_1^2 + \tilde{\omega}_2^2 + \ldots + \tilde{\omega}_7^2) + \frac{1}{8e^2}\Big(& \mathbf{s}_{12}^2 + \mathbf{s}_{23}^2 \\ & + \mathbf{s}_{31}^2 + \frac{1}{4}\,(\tilde{\omega}_4^2 + \ldots + \tilde{\omega}_7^2)(\mathbf{l}_1^2 + \mathbf{l}_2^2 + \mathbf{l}_3^2)\Big), \end{aligned} \tag{20a}$$

or

$$\begin{aligned} L_{\mathrm{rot}}(SU_2) = \frac{F_\pi^2}{8}\Big[& \omega^2\mathbf{f}^2 - (\omega\mathbf{f})^2 + \frac{1-f_0}{2}\,(\omega_4^2 + \omega_5^2 + \omega_6^2 \\ & + \omega_7^2)\Big] + \frac{1}{8e^2}\Big[\tilde{\omega}^2\mathbf{l}_i^2 - (\tilde{\omega}\mathbf{l}_i)^2 + \frac{1-f_0}{2}\,\mathbf{l}_i^2 \\ & \times(\omega_4^2 + \ldots + \omega_7^2)\Big], \end{aligned} \tag{20b}$$

where

$$\tilde{\omega}_i = [R_{ik}(U_0) - \delta_{ik}]\omega_k = 2(f_i f_k - \mathbf{f}^2\delta_{ik} + f_0\epsilon_{ikl}f_l)\omega_k$$

TABLE I. The values of the masses M_{cl}, the mass term M.t. (in MeV), the strangeness content C_S and the moments of inertia (in 10^{-3} MeV^{-1}) for the hedgehog with $B=1$ and the dipole configuration with $B=2$ (Ref. 16) in the flavor-symmetric (FS) and flavor-symmetry-broken (FSB) cases. Here M.t. is included in M_{cl}, $F_\pi=186$ MeV and $e=4.12$. The accuracy of the calculations is at least $\sim 0.5\%$ in the masses and a few percent in other quantities.

	B	M_{cl}	M.t.	C_S	Θ_N	Θ_S	Θ_3	Θ_8	Θ_{38}	$\Theta_{46}=-\Theta_{57}$
FS	1	1702	46	-	5.55	2.04	-	-	-	-
FS	2	3330	87	0.495	4.14	7.13	2.86	8.14	0.01	0.63
FSB	1	1982	199	-	3.24	1.06	-	-	-	-
FSB	2	3885	380	0.475	2.44	4.13	1.70	4.77	0.002	0.24

for $i,k=1,2,3$, and ω and $\tilde{\omega}$ have three components in the (u,d) $SU(2)$ subgroup with $\tilde{\omega}^2=4[\omega^2\mathbf{f}^2-(\tilde{\omega}\mathbf{f})^2]$.

To derive (20b) we used also that

$$\tilde{\omega}_4^2+\ldots+\tilde{\omega}_7^2=2(1-f_0)(\omega_4^2+\ldots+\omega_7^2).$$

Here, $\mathbf{l}_2$ parametrizes the chiral derivatives of U_0: $U_0^\dagger d_k U_0 = i\tau_i l_{i,k}$, and the functions f_0, $\mathbf{f}$ define the matrix U_0 in the usual way

$$\mathbf{l}_i^2=(d_i f_0)^2+\ldots+(d_i f_3)^2.$$

Equation (20b) defines the moments of inertia of arbitrary $SU(2)$ skyrmions rotated in $SU(3)$ configuration space and illustrates well that the $SU(2)$ case is much simpler than general $SU(3)$ case. The analytical expressions for the moments of inertia of axially symmetric $SU(2)$ skyrmions, also rotated in ν-direction, can be found in Refs. 11 and 19.

When the $SU(2)$ hedgehog is quantized in the $SU(3)$ collective coordinates space only two different moments of inertia enter,[8,13,14] $\Theta_1=\Theta_2=\Theta_3$ and $\Theta_4=\Theta_5=\Theta_6=\Theta_7$. For the $SO(3)$ hedgehog the rotation energy also depends on two different moments of inertia, $\Theta_2=\Theta_5=\Theta_7$ and $\Theta_1=\Theta_3=\Theta_4=\Theta_6=\Theta_8$.[8,9] In the case of the strange skyrmion molecule we obtained four different diagonal moments of inertia:[20] $\Theta_1=\Theta_2=\Theta_N$; Θ_3; $\Theta_4=\Theta_5=\Theta_6=\Theta_7=\Theta_S$ and Θ_8. Numerically the difference between Θ_N and Θ_3 is not large while Θ_8 is a bit greater than Θ_S (see Table I below). In view of the symmetry properties of the configuration many off-diagonal moments of inertia are equal to zero. A few of them are nonzero, but at least one order of magnitude smaller than the diagonal moments of inertia, e.g., Θ_{46} and Θ_{57}. For this reason we shall neglect them here in making estimates.

The Lagrangian of the system can be written in terms of the angular velocities of rotation and moments of inertia in the form (in the body-fixed system)

$$L_{\rm rot}=\frac{\Theta_N}{2}(\omega_1^2+\omega_2^2)+\frac{\Theta_3}{2}\omega_3^2+\frac{\Theta_S}{2}(\omega_4^2+\omega_5^2+\omega_6^2+\omega_7^2)+\frac{\Theta_8}{2}\omega_8^2+\Theta_{45}(\omega_4\omega_5-\omega_6\omega_7)+\ldots\,. \tag{21}$$

The above relations between the different moments of inertia of the strange molecule can be obtained in the following way, at large distances between the two $B=1$ hedgehogs.

When the $B=1$ skyrmion is located in the (u,s) $SU(2)$ subgroup of $SU(3)$ we obtain from (12) and (16)

$$L_{\rm rot}(u,s)=\frac{\theta_S}{2}(\omega_1^2+\omega_2^2+\omega_6^2+\omega_7^2)+\frac{\theta_N}{2}\left[\omega_4^2+\omega_5^2+\frac{1}{4}(\omega_3+\sqrt{3}\omega_8)^2\right], \tag{22a}$$

where we have retained the notations used for the (u,d) $B=1$ soliton.

For the $B=1$ skyrmion in the (d,s) subgroup,

$$L_{\rm rot}(d,s)=\frac{\theta_S}{2}(\omega_1^2+\omega_2^2+\omega_4^2+\omega_5^2)+\frac{\theta_N}{2}\left[\omega_6^2+\omega_7^2+\frac{1}{4}(\omega_3-\sqrt{3}\omega_8)^2\right], \tag{22b}$$

with[13,14]

$$\theta_S=\frac{1}{8}\int(1-c_F)\left[F_\pi^2+\frac{1}{e^2}(F'^2+2s_F^2/r^2)\right]d^3r,$$

$$\theta_N=\frac{1}{6}\int s_F^2\left[F_\pi^2+\frac{4}{e^2}(F'^2+s_F^2/r^2)\right]d^3r, \tag{22c}$$

where $F(r)$ is the profile function of the $B=1$ hedgehog and $f_0=\cos F$. Relations (22c) follow immediately from (20b). Note that the combinations of ω_3 and ω_8 which enter the expressions for the rotation energy (22a, and 22b) and the WZW-term (10) are orthogonal to each other, as follows from general arguments.

When two $B=1$ hedgehogs in different subgroups, (d,s) and (u,s), are located at large distances, we should take the sum of the expressions (22a), (22b). Simple relations for the $B=2$ moments of inertia Θ in terms of the $B=1$ inertia θ then appear:

$$\Theta_N=2\theta_S,\quad \Theta_S=\theta_N+\theta_S,\quad \Theta_3=\theta_N/2,$$
$$\Theta_8=3\theta_N/2=3\Theta_3. \tag{23}$$

For interacting hedgehogs in a molecule these relations hold only approximately (see Table I where some numbers are corrected in comparison with Ref. 20).

In the flavor-symmetric (FS) case all meson masses in the Lagrangian are equal to the pion mass, and the distance between centers of the two skyrmions in the molecule equals ~ 1.05 Fm. In the FSB case the kaon mass is included in the Lagrangian (see the next section) and the distance between solitons centers in the molecule is ~ 0.75 Fm.[16,20]

The Hamiltonian of the system can be obtained by the canonical quantization procedure[2,14,8] which we reproduce here for completeness. The components of the body-fixed $SU(3)$ angular momentum J_k^R can be defined as

$$J_k^R=dL/d\omega_k. \tag{24}$$

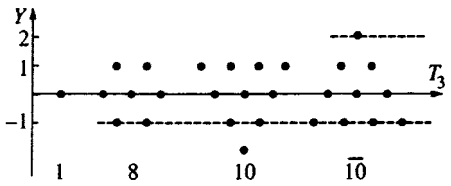


FIG. 2. T_3-Y-diagrams for the lowest $SU(3)$ multiplets allowed for the case of $[SU(2)]^3$ configurations, ansatz (11): singlet $(p,q)=(0,0)$, octet (1,1), decuplet (3,0) and antidecuplet (0,3). The lower dashed line indicates isomultiplets with $Y=-1\simeq Y_R^{\min}$, $T=N$; the upper dashed line shows nonstrange isomultiplets with $Y=B=2$.

This definition coincides identically with another one,

$$J_k^R=\frac{1}{2i}\,\mathrm{Tr}\;A\lambda_k\pi, \tag{25}$$

where $\pi_{\alpha\beta}=dL/d\dot{A}_{\beta\alpha}$. In the canonical quantization procedure the substitution

$$\pi_{\alpha\beta}=\frac{dL}{d\dot{A}_{\beta\alpha}}\to -i\,\frac{d}{dA_{\beta\alpha}} \tag{26}$$

is made. The commutation relations

$$[J_i^R J_k^R]=-if_{ikl}J_l^R \tag{27}$$

then follow immediately, where f_{ikl} are the $SU(3)$ structure constants.

After the standard quantization procedure the Hamiltonian of the system, $H=\omega_i dL/d\omega_i-L$, is a bilinear function of the generators J_i^R. For the states belonging to a definite $SU(3)$ irrep the rotation energy can be written in the simplified form

$$E_{\rm rot}=\frac{C_2(SU_3)-3Y_R^2/4}{2\Theta_S}+\frac{N(N+1)}{2}\left(\frac{1}{\Theta_N}-\frac{1}{\Theta_S}\right)+\frac{3(Y_R-Y_R^{\min})^2}{8\Theta_8}. \tag{28}$$

The second-order Casimir operator of the $SU(3)$ group is $C_2(SU_3)=(1/3)(p^2+q^2+pq)+p+q$, N is the right isospin (see Fig. 2) and p,q are the numbers of the upper and low indices in the tensor describing the $SU(3)$ irrep (p,q). The terms linear in the angular velocities present in the Lagrangian due to the Wess–Zumino–Witten term are cancelled in the Hamiltonian, but they lead to the quantization condition discussed in the previous section. Corrections of order Θ_{45}^2/Θ_S^2 and $(\Theta_N-\Theta_3)/\Theta_N$ have been neglected in (28). Note that $Y_R^{\min}$ can take arbitrary noninteger values because it is a quantity similar to the strangeness content C_S,[17] not a quantum number. Y_R is a quantum number and can take only integer values. The usual spatial angular momentum is $J=0$ here. The correct description of the usual spatial rotations demands the introduction of a second set of collective coordinates, as it was done previously[11] for the case of flavor $SU(2)$. It was shown that the states of the lowest energy have $J=0$.

It is clear from expression (28) that for $\Theta_8\to 0$ the right hypercharge satisfies $Y_R=Y_R^{\min}=2L_8^{\rm WZ}/\sqrt{3}$, otherwise the quantum correction due to ω_8 would be infinite. For solitons located in (u,d) $SU(2)$ we have $\Theta_8=0$ and $Y_R=2L_8^{\rm WZ}/\sqrt{3}=B$, the quantization condition[13,18] with $N_c=3$.

For the skyrmion molecule[16] we have $L_8^{\rm WZ}\approx-\sqrt{3}/2$, or $Y_R^{\min}\approx-1$, as was explained above. The last term in (25) is absent for $Y_R=-1$, and because of the evident constraints

$$\frac{p+2q}{3}\geqslant Y_R\geqslant-\frac{q+2p}{3} \tag{29}$$

the following lowest $SU(3)$ multiplets are possible: octet, $(p,q)=(1,1)$, decuplet (3,0) and antidecuplet (0,3), Fig. 2. The sum of the classical mass of the soliton and rotational energy for the $B=2$ octet, 10 and $\widetilde{10}$ is equal to $\sim$3.53, 3.74 and 3.89 GeV for $Y_R=-1$ (the flavor-symmetric FS-case). The whole FSB mass term described in the following section, $\Delta M+\delta M_{\rm FS}$, should be added to these numbers. When the FSB mass term is included in the classical mass the sum $M_{cl}+E_{\rm rot}$ equals 4.23, 4.59 and 4.84 GeV for the octet, decuplet and $\widetilde{10}$. Only the mass splitting part of the mass term, $\delta M_{\rm FSB}$, should be added to these numbers (see Table II below). The octets with $Y_R=0$ and 1 have $M_{cl}+E_{\rm rot}+\Delta M$ equal to 4.61 and 4.73 GeV according to (28) (the FS scheme of calculation). The $SU(3)$ singlet with $Y_R=0$ has energy equal to $M_S=M_{cl}+3/8\Theta_8$ which, according to Table I

TABLE II. The values of I, $(1/2)\sin^2\nu$, the mass splitting δM (in MeV) and the masses M (in GeV) for the octet, decuplet and antidecuplet of dibaryons in the flavor-symmetric and flavor-symmetry-broken cases. The binding energy of the configuration $\epsilon=(M_1+M_2-M)/(M_1+M_2)$ relative to the final state F.st. is presented. $M_{\rm FS}=M_{cl,\rm FS}+E_{\rm rot,FS}+\Delta M+\delta M_{\rm FS}$, $M_{\rm FSB}=M_{cl,\rm FSB}+E_{\rm rot,FSB}+\delta M_{\rm FSB}$.

$\lvert p,q;Y,T\rangle$	I	$-\langle\sin^2\nu/2\rangle$	$\delta M_{\rm FS}$	$M_{\rm FS}$	$\delta M_{\rm FSB}$	$M_{\rm FSB}$	F.st.	$\epsilon_{\rm FS}$
$\lvert 8,1,1/2\rangle$	−2/10	−4/10	−385	4.16	−124	4.10	ΛN	0.14
$\lvert 8,0,1\rangle$	−1/10	−11/30	−353	4.19	−114	4.11	ΞN	0.15
$\lvert 8,0,0\rangle$	1/10	−3/10	−289	4.26	−93	4.13	$\Lambda\Lambda$	0.14
$\lvert 8,-1,1/2\rangle$	3/10	−7/30	−224	4.32	−73	4.15	$\Lambda\Xi$	0.14
$\lvert 10,1,3/2\rangle$	−1/8	−3/8	−361	4.40	−117	4.47	ΣN	0.11
$\lvert 10,0,1\rangle$	0	−1/3	−320	4.44	−104	4.48	ΞN	0.11
$\lvert 10,-1,1/2\rangle$	1/8	−7/24	−280	4.48	−91	4.50	$\Lambda\Xi$	0.11
$\lvert 10,-2,0\rangle$	1/4	−1/4	−240	4.53	−78	4.52	$\Xi\Xi$	0.10
$\lvert\overline{10},2,0\rangle$	−1/4	−5/12	−401	4.52	−130	4.72	NN	0.04
$\lvert\overline{10},1,1/2\rangle$	−1/8	−3/8	−361	4.55	−117	4.72	ΛN	0.06
$\lvert\overline{10},0,1\rangle$	0	−1/3	−320	4.59	−104	4.74	ΞN	0.07
$\lvert\overline{10},-1,3/2\rangle$	1/8	−7/24	−280	4.63	−91	4.75	$\Sigma\Xi$	0.09

equals $\simeq 3.38$ GeV in the FS case. This can be compared with the $SO(3)$ hedgehog mass, $M_H = 3.272$ GeV for the same values of the parameters.[12]

4. MASS SPLITTING WITHIN $SU(3)$ MULTIPLETS OF DIBARYONS

The mass splittings inside $SU(3)$ multiplets are defined as usual by the FSB part of the mass terms in the Lagrangian density:

$$L_M = \frac{F_\pi^2 m_\pi^2}{16}\,\mathrm{Tr}(U+U^\dagger-2) + \frac{F_K^2 m_K^2 - F_\pi^2 m_\pi^2}{24}\,\mathrm{Tr}(1 - \sqrt{3}\lambda_8)(U+U^\dagger-2). \tag{30}$$

When (u,d) $SU(2)$ solitons are rotated in the «strange» direction by means of the matrix $U_4 = \exp(-i\nu\lambda_4)$, (30) leads to the substitution[4,11]

$$F_\pi^2 m_\pi^2 \to F_\pi^2 m_\pi^2 + \sin^2\,\nu (F_K^2 m_K^2 - F_\pi^2 m_\pi^2).$$

For the ansatz (11), after averaging over all phases in the matrix $A(t)$ except ν, we can rewrite the mass term in the energy density in the following form:

$$\mathrm{M.t.} = \frac{F_\pi^2 m_\pi^2}{8}\,(3 - v_1 - v_2 - v_3) + \frac{F_K^2 m_K^2 - F_\pi^2 m_\pi^2}{4} \times \left[1 - v_3 + (2v_3 - v_1 - v_2)\,\frac{\sin^2\,\nu}{2}\right], \tag{31a}$$

or, for $F_K = F_\pi$

$$\mathrm{M.t.} = \frac{F_\pi^2 m_\pi^2}{4}\left[(3 - v_1 - v_2 - v_3)\left(\frac{1}{2} + \left(\frac{m_K^2}{m_\pi^2} - 1\right) C_S\right) + \left(\frac{m_K^2}{m_\pi^2} - 1\right)(2v_3 - v_1 - v_2)\,\frac{\sin^2\,\nu}{2}\right]. \tag{31b}$$

Here v_1, v_2 and v_3 are real parts of the diagonal matrix elements of the matrix U, depending on the functions f_i and q_i. For the ansatz (11) we have $(b_0 = 0)$

$$\begin{aligned} v_1 &= c_{a_0} c_a (c_b f_0 - s_b f_3) + s_{a_0} s_a c_b, \\ v_2 &= c_{a_0} c_a (c_b q_0 + s_b q_3) + s_{a_0} s_a [c_b (f_0 q_0 - f_3 q_3) \\ &\quad + s_b (f_3 q_0 + f_0 q_3)] - s_{a_0} (f_1 q_1 + f_2 q_2), \\ v_3 &= f_0 q_0 - f_3 q_3 + s_a [s_b (f_1 q_2 - f_2 q_1) - c_b (f_1 q_1 \\ &\quad + f_2 q_2)], \end{aligned} \tag{32a}$$

a_0 and b_0 are the asymptotic values of the functions a, b. For the local minimum found recently,[16] $a_0 = b_0 = 0$. In this case (32a) simplifies to

$$\begin{aligned} v_1 &= c_a (c_b f_0 - s_b f_3), \\ v_2 &= c_a (c_b q_0 + s_b q_3). \end{aligned} \tag{32b}$$

Here v_3 is given by (32a) since it does not depend on a_0, b_0. If a_0, b_0 are different from zero the ansatz (11) should be written

$$U = \exp(-i a_0 \lambda_2) U_L(u,s) U(u,d) U_R(d,s) \times \exp(-i b_0 \lambda_3)$$

to ensure the correct behavior of $U(\mathbf{r})$ at $\mathbf{r} \to \infty$. For example, for $a = a_0 = \pi/2$, $b = b_0 = 0$ then we have $v_1 = 1$, $v_2 = v_3 = f_0 q_0 - f_3 q_3 - f_1 q_1 - f_2 q_2$, i.e., the skyrmion is located in the (d,s) $SU(2)$ subgroup.

In the FS case the part of the mass term

$$\mathrm{M.t.}_{\mathrm{FS}} = F_\pi^2 m_\pi^2 (3 - v_1 - v_2 - v_3)/8 \tag{33}$$

is included in the classical mass M_{cl} which is minimized. In the FSB case the second part,

$$\Delta M = (F_K^2 m_K^2 - F_\pi^2 m_\pi^2)(1 - v_3)/4 \tag{34}$$

also is included in the minimized M_{cl}; see Table I. In the FS case $\Delta M \simeq 1016$ MeV, while in the FSB case it is squeezed by a factor ~ 3.

The mass splitting inside $SU(3)$ multiplets is defined by the term

$$\delta M = -\frac{1}{4}\,(F_K^2 m_K^2 - F_\pi^2 m_\pi^2)(v_1 + v_2 - 2v_3)\left\langle \frac{1}{2}\sin^2\,\nu \right\rangle, \tag{35}$$

which is not included in M_{cl} and is considered as a perturbation in both cases. Here v is the angle of rotation in the «nonstrange» direction. For two undeformed hedgehogs at large relative distances we have $v_1 + v_2 - 2v_3 \to 2(1 - \cos F)$, where F is the profile function of the $B = 1$ hedgehog, and the coefficient of $\sin^2\,\nu$ is the same as for the rotated $B = 1$ (u,d) hedgehog. Note that in the case of a strange skyrmion molecule with strangeness content close to 0.5 the term (35) defining the mass splitting within multiplets is negative—directly opposite to the case when the nonstrange $SU(2)$ solitons are used as starting configurations and are rotated in the «strange» direction. The quantity δM should be added to the sum of $M_{cl} + E_{\mathrm{rot}}$ calculated at the end of Sec. 3, and ΔM should be added in the FS case.

To obtain the mass splitting within $SU(3)$ multiplets we should calculate, as usual, the matrix elements of the function

$$\left\langle \frac{1}{2}\sin^2\,\nu \right\rangle = \frac{1}{3}\,\langle 1 - D_{88}(\nu)\rangle = \frac{1}{3}\,(1 - I)$$

for each component of the $SU(3)$ multiplets described by the $SU(3)$ D-functions. Then the quantity I is equal to

$$I = \sum_\gamma C^{8;(p,q);(p,q)_\gamma}_{0,0,0;Y,T,T_3;Y,T,T_3} C^{8;(p,q);(p,q)_\gamma}_{0,0,0;Y_R,N,M;Y_R,N,M}, \tag{36}$$

expressed through the Clebsch–Gordan coefficients of the $SU(3)$ group.[22] In the case of a strange molecule we have $Y_R = -1$, $N = 1/2$ for the octet and decuplet, $N = 3/2$ for $\widetilde{10}$. The values of $\langle (1/2)\sin^2\,\nu \rangle$ and the mass splittings are shown in Table II.

For the octet the allowed strangeness of the states is -1, -2, -3, for the decuplet it ranges from -1 to -4, and the nonstrange dibaryons appear in the $\widetilde{10}$, 27-plet, etc. (Fig. 2). The masses of the dibaryons calculated according to the FS and FSB schemes differ, but not very much since the in-

crease of the total mass term in the FS case is compensated by the decrease of $E_{\rm rot}$ in comparison with the FSB case. The states $|10,-2,0\rangle$ and $|\widetilde{10},2,0\rangle$ are supposed to have $J=1$ and the corresponding energy is added, roughly estimated according to our previous results.[11]

When FSB mass terms are included in the minimized static energy M_{cl} they are squeezed by a factor ~ 3 due to the smaller dimensions of the kaon cloud in comparison with the pion cloud,[16] so the moments of inertia are greater and $E_{\rm rot}$ is smaller in the FS case (see Table I). The absolute values of the masses are controlled by the Casimir energy,[23–26] which has the order of magnitude ~ -1 GeV for $B=1$[25,26] and ~ -2 GeV for $B=2$ molecules.

For the 27-plet the value of the difference of I for states with maximum and minimum hypercharge is 3/8, just as for decuplet and antidecuplet. The relative binding ϵ is shown in Table II because it is less sensitive to the method of calculation. M_1 and M_2 are the masses of the final baryons available due to strong interactions, calculated within the same approach (theory-to-theory comparison). Inclusion of configuration mixing usually leads to an increase of the mass splitting by $\sim 0.3-0.4$.[27] Since the results for the mass splitting shown in Table II depend on the starting configuration, and both FS and FSB calculation schemes are not consistent by themselves, one should use some interpolating procedure, e.g., similar to the slow-rotator approximation used successfully in Ref. 4 for the description of the hyperon mass splitting.

5. CONCLUSIONS AND DISCUSSION

The quantization scheme for the $SU(3)$ skyrmions has been presented and the quantization condition known previously[13] is generalized for skyrmions with arbitrary strangeness content, which allows one to investigate the consequences of the existence of different local minima in $SU(3)$ configuration space. The quantization condition (8) is valid for all known $B=2$ local minima shown in Fig. 1. It is proved rigorously in several cases; in other cases it was confirmed by numerical calculation. However, some corrections to relation (8) cannot be excluded. The moments of inertia of arbitrary $SU(3)$ skyrmions can be calculated with the help of formulas (16), (17). Both static and rotational energies as well as the baryon number density of $SU(3)$ skyrmions are presented in a form which makes apparent their symmetry in different $SU(2)$ subgroups of $SU(3)$.

For the dipole-type configuration with $C_S=0.5$ our results are in qualitative agreement with those obtained in Ref. 28 for the interaction potential of two strange baryons located at large distances. The absolute values of the masses of both $B=1$ and $B=2$ states are controlled by the Casimir energies, which make a contribution of order N_c^0 to the masses of the configurations.[22–25] However, the dipole-type configuration does not differ much from the $B=2$ configuration within the product ansatz which we used as a starting point in our calculations.[16] For this reason the Casimir energy of the dipole can be close to twice that for the $B=1$ soliton and can cancel in the binding energies of dibaryons. We conclude therefore that a new branch of strange dibaryons in addition to those known previously[8,9,11] is predicted with a small uncertainty in the absolute values of masses due to the Casimir energy, relative to the corresponding $B=1$ states. The values of masses and bindings we obtained here cannot, however, be taken too seriously, not only because the Casimir energy is poorly known but also because the nonzero mode contributions closely connected with the Casimir energy (principally the breathing and vibrational modes) have not been taken into account. These effects not only decrease the binding energies,[6,7] but can make many of the states listed in Table II unbound.

The prediction by chiral soliton models of a rich spectrum of baryonic states with different values of strangeness remains one of the intriguing properties of such models. The comparison with predictions of the quark or quark-bag models[29,30] is of special interest. Some of these models predict the existence of bound strange baryonic states,[30] similar to the chiral soliton approach.

It is difficult to observe these states, especially those which are above the threshold for decay due to strong interactions. The searches for the H-dibaryon predicted at first within the MIT quark-bag model[29] have been undertaken in different experiments, without success till now. It should be noted that observation of the H-dibaryon can be especially difficult by the following reasons. First, its dimensions are small in the framework of the chiral soliton approach,[12,11] $R_H \sim 0.5-0.6$ Fm. Therefore, estimates of the H-dibaryon production cross section based on the assumption that its dimensions are close to the dimensions of the deuteron may be too optimistic.

Second, it is not clear how the transition from an H-dibaryon to two $B=1$ solitons can proceed. Schwesinger proposed a nontrivial parametrization allowing for the transition from the $SO(3)$ $B=2$ hedgehog to the $B=2$ $SU(2)$ torus (described in Ref. 31). Within this parametrization the two configurations are separated by a potential barrier; moreover, the behavior of some function in this parametrization is singular. So, if such a transition is not possible with smooth functions, it would be difficult to find H-dibaryon in coalescence experiments. However, further investigations of the predictions of effective field theories providing a new approach to the description of the fundamental properties of matter are of interest. The near-threshold enhancement in $p\Lambda$ system which was observed many years ago in, e.g., the reaction $pp \to p\Lambda K^+$[32] and confirmed in recent investigations[33] may be a confirmation of soliton model predictions, because within this approach there is no difference between real and virtual levels.

The problem of the H-dibaryon discussed in Ref. 8 is that of parity doubling: the $SO(3)$ soliton has no definite parity, so a special symmetrization procedure should be done.[8] A similar problem exists for the strange molecules also. For the classical configuration of molecular type we have different $B=1$ skyrmions in different parts of space and in different $SU(2)$ subgroups of $SU(3)$. The molecule has no definite parity, but these configurations are invariant under the combined operation of parity transformation and interchange of $SU(2)$ subgroups. The electric dipole momentum of the molecule is different from zero (this was noted by

M. Luty). (Anti)symmetrization should be performed, similar to the H-particle case, providing a state of definite parity and removing the e.d.m. for the quantized state.

I am thankful to Bernd Schwesinger for valuable discussions and suggestions in the initial stages of the work. I am indebted also to G. Holzwarth and H. Walliser for their interest in the problems of $SU(3)$ skyrmions useful discussions during my visits to Siegen University, and to B. E. Stern for help in the numerical computations. I appreciate also the support by Volkswagenstiftung, FRG at the beginning of the present work.

The work supported by the Russian Fund for Fundamental Research, Grant 95-02-03868a.

APPENDIX

Here we sketch the expressions for the matrix elements R_{ik} which connect the rotation angular velocities in body-fixed and rotated coordinate systems:

$$\tilde{\omega}_i = R_{ki}\omega_k, \quad R_{ik} = R_{ik}(V) - R_{ik}(T^+),$$

$$R_{ik}(T^+) = R_{ki}(T) = \frac{1}{2}\,\mathrm{Tr}\,\lambda_i T^+ \lambda_k T,$$

$$V = U_L(u,s)\exp(ia\lambda_2), \quad T = \exp(ib\lambda_3)U_R(d,s),$$

$$U_0 = VT.$$

Definitions of $U_L(u,s)$ and $U_R(d,s)$ in terms of the functions $f_0,\ldots,f_3$ and $q_0,\ldots,q_3$ are given following expression (12).

We use the notations

$$f_{12}^2 = f_1^2 + f_2^2, \quad q_{12}^2 = q_1^2 + q_2^2, \quad F_1^+ = f_0 f_1 + f_2 f_3,$$

$$F_1^- = f_0 f_1 - f_2 f_3,$$

$$F_2^+ = f_0 f_2 + f_1 f_3, \quad F_2^- = f_0 f_2 - f_1 f_3,$$

$$F_3^+ = f_0 f_3 + f_1 f_2, \quad F_3^- = f_0 f_3 - f_1 f_2,$$

$$Q_1^+ = q_0 q_1 + q_2 q_3, \quad Q_1^- = q_0 q_1 - q_2 q_3,$$

$$Q_2^+ = q_0 q_2 + q_1 q_3, \quad Q_2^- = q_0 q_2 - q_1 q_3,$$

$$s_{bb_0} = \sin(b - b_0), \quad Q_c = c_b Q_2^- + s_b Q_1^+,$$

$$Q_s = s_b Q_2^- - c_b Q_1^+,$$

$$\Delta_F = f_0^2 + f_1^2 - f_2^2 - f_3^2, \quad \delta_F = f_0^2 - f_1^2 + f_2^2 - f_3^2,$$

$$C^+ = c_{b+b_0}(q_1^2 - q_2^2) - 2s_{b+b_0} q_1 q_2,$$

$$S^+ = s_{b+b_0}(q_1^2 - q_2^2) + 2c_{b+b_0} q_1 q_2,$$

$$C^- = c_{bb_0}(q_0^2 - q_3^2) + 2s_{bb_0} q_0 q_3,$$

$$S^- = s_{bb_0}(q_0^2 - q_3^2) - 2c_{bb_0} q_0 q_3. \tag{A1}$$

Here a_0 and b_0 are asymptotic values of functions a and b. For the strange molecule[16] we have $a_0 = b_0 = 0$. When, e.g., $a = a_0 = \pi/2$, $b = b_0 = 0$ hold, the matrix U corresponds to solitons located in the (d,s) $SU(2)$ subgroup of $SU(3)$.

$$R_{11} = s_{2a_0} s_{2a}(1 - f_{12}^2/2) + c_{2a_0} c_{2a} f_0 - c_{2b-2b_0} q_0 - s_{2b-2b_0} q_3,$$

$$R_{12} = c_{2a_0} f_3 + s_{2b-2b_0} q_0 - c_{2b-2b_0} q_3,$$

$$R_{13} = s_{2a_0} c_{2a}(1 - f_{12}^2/2) - c_{2a_0} s_{2a} f_0,$$

$$R_{14} = s_{2a_0} c_a F_2^+ - c_{2a_0} s_a f_2 + c_{2b_0-b} q_2 + s_{2b_0-b} q_1,$$

$$R_{15} = -s_{2a_0} c_a F_1^- + c_{2a_0} s_a f_1 - c_{2b_0-b} q_1 + s_{2b_0-b} q_2,$$

$$R_{16} = s_{2a_0} s_a F_2^+ + c_{2a_0} c_a f_2, \quad R_{17} = -s_{2a_0} s_a F_1^- - c_{2a_0} c_a f_1, \quad R_{18} = -\sqrt{3} s_{a_0} c_{a_0} f_{12}^2,$$

$$R_{21} = -c_{2a} f_3 - s_{2b-2b_0} q_0 + c_{2b-2b_0} q_3,$$

$$R_{22} = f_0 - c_{2b-2b_0} q_0 - s_{2b-2b_0} q_3,$$

$$R_{23} = s_{2a} f_3, \quad R_{24} = s_a f_1 + c_{2b_0-b} q_1 - s_{2b_0-b} q_2,$$

$$R_{25} = s_a f_2 + s_{2b_0-b} q_1 = c_{2b_0-b} q_2, \quad R_{26} = -c_a f_1,$$

$$R_{27} = -c_a f_2, \quad R_{28} = 0.$$

$$R_{31} = s_{2a} c_{2a_0}(1 - f_{12}^2/2) - c_{2a} s_{2a_0} f_0, \quad R_{32} = -s_{2a_0} f_3,$$

$$R_{33} = c_{2a} c_{2a_0}(1 - f_{12}^2/2) + s_{2a} s_{2a_0} f_0 - 1 + q_{12}^2/2,$$

$$R_{34} = c_a c_{2a_0} F_2^+ + s_a s_{2a_0} f_2,$$

$$R_{35} = -c_a c_{2a_0} F_1^- - s_a s_{2a_0} f_1,$$

$$R_{36} = s_a c_{2a_0} F_2^+ - c_a s_{2a_0} f_2 - Q_c,$$

$$R_{37} = -s_a c_{2a_0} F_1^- + c_a s_{2a_0} f_1 - Q_s,$$

$$R_{38} = -\frac{\sqrt{3}}{2}(c_{2a_0} f_{12}^2 + q_{12}^2),$$

$$R_{41} = -s_{2a} c_{a_0} F_2^- + s_{a_0} c_{2a} f_2 - s_{2b-b_0} q_1 - c_{2b-b_0} q_2,$$

$$R_{42} = -s_{a_0} f_1 - c_{2b-b_0} q_1 + s_{2b-b_0} q_2,$$

$$R_{43} = -c_{a_0} c_{2a} F_2^- - s_{a_0} s_{2a} f_2,$$

$$R_{44} = c_{a_0} c_a \Delta_F + s_{a_0} s_a f_0 - c_{bb_0} q_0 = s_{bb_0} q_3,$$

$$R_{45} = 2c_{a_0} c_a F_3^+ + s_{a_0} s_a f_3 + s_{bb_0} q_0 + c_{bb_0} f_3,$$

$$R_{46} = c_{a_0} s_a \Delta_F - s_{a_0} c_a f_0,$$

$$R_{47} = 2c_{a_0} s_a F_3^+ - s_{a_0} c_a f_3, \quad R_{48} = \sqrt{3} c_{a_0} F_2^-.$$

$$R_{51} = c_{a_0} s_{2a} F_1^+ - s_{a_0} c_{2a} f_1 + c_{2b-b_0} q_1 - s_{2b-b_0} q_2,$$

$$R_{52} = -s_{a_0} f_2 - s_{2b-b_0} q_1 - c_{2b-b_0} q_2,$$

$$R_{53} = c_{a_0} c_{2a} F_1^+ + s_{a_0} s_{2a} f_1,$$

$$R_{54} = -2c_{a_0} c_a F_3^- - s_{a_0} s_a f_3 - s_{bb_0} q_0 - c_{bb_0} q_3,$$

$$R_{55} = c_{a_0} c_a \delta_F + s_{a_0} s_a f_0 - c_{bb_0} q_0 + s_{bb_0} q_3,$$

$$R_{56}=-2c_{a_0}s_aF_3^-+s_{a_0}c_af_3,\quad R_{57}=c_{a_0}s_a\delta_F-s_{a_0}c_af_0,$$

$$R_{58}=\sqrt{3}c_{a_0}F_1^+,$$

$$R_{61}=-s_{a_0}s_{2a}F_2^--c_{a_0}c_{2a}f_2,\quad R_{62}=c_{a_0}f_1,$$

$$R_{63}=-s_{a_0}c_{2a}F_2^-+c_{a_0}s_{2a}f_2+c_{b_0}Q_2^++s_{b_0}Q_1^-,$$

$$R_{64}=s_{a_0}c_a\Delta_F-c_{a_0}s_af_0,\quad R_{65}=2s_{a_0}c_aF_3^+-c_{a_0}s_af_3,$$

$$R_{66}=s_{a_0}s_a\Delta_F+c_{a_0}c_af_0-C^--C^+,$$

$$R_{67}=2s_{a_0}s_aF_3^++c_{a_0}c_af_3-S^--S^+,$$

$$R_{68}=-\sqrt{3}[s_{a_0}F_2^-+c_{b_0}Q_2^++s_{b_0}Q_1^-].$$

$$R_{71}=s_{a_0}s_{2a}F_1^++c_{a_0}c_{2a}f_1,\quad R_{72}=c_{a_0}f_2,$$

$$R_{73}=s_{a_0}c_{2a}F_1^+-c_{a_0}s_{2a}f_1-c_{b_0}Q_1^-+s_bQ_2^+,$$

$$R_{74}=-2s_{a_0}c_aF_3^-+c_{a_0}c_af_3,\quad R_{75}=s_{a_0}c_a\delta_F-c_{a_0}s_af_0,$$

$$R_{76}=-2s_{a_0}s_aF_3^--c_{a_0}c_af_3+S^--S^+,$$

$$R_{77}=s_{a_0}s_a\delta_F+c_{a_0}c_af_0-C^-+C^+,$$

$$R_{78}=\sqrt{3}[s_{a_0}F_1^+-s_{b_0}Q_2^++c_{b_0}Q_1^-],$$

$$R_{81}=-\frac{\sqrt{3}}{2}s_{2a}f_{12}^2,\quad R_{82}=0,$$

$$R_{83}=-\frac{\sqrt{3}}{2}(c_{2a}f_{12}^2+q_{12}^2),\quad R_{84}=\sqrt{3}c_aF_2^+,$$

$$R_{85}=-\sqrt{3}c_aF_1^-,\quad R_{86}=\sqrt{3}s_a(F_2^++Q_C),$$

$$R_{87}=-\sqrt{3}s_a(F_1^--Q_S),\quad R_{88}=\frac{3}{2}(q_{12}^2-f_{12}^2). \qquad \text{(A2)}$$

The R_{8i} do not depend on a_0,b_0 because the matrices λ_2,λ_3 commute with λ_8. The orthogonality of the real matrices $R(V)$ and $R(T)$ can be checked immediately from these expressions.

[1] T. H. R. Skyrme, Proc. R. Soc. London, Ser. A **260**, 127 (1961); Nucl. Phys. **31**, 556 (1962).
[2] G. Adkins, C. Nappi, and E. Witten, Nucl. Phys. B **228**, 552 (1983); G. Adkins and C. Nappi, Nucl. Phys. B **233**, 109 (1984).
[3] G. Holzwarth and B. Schwesinger, Rep. Prog. Phys. **49**, 825 (1986); I. Zahed and G. E. Brown, Phys. Rep. **142**, 1 (1986).
[4] B. Schwesinger and H. Weigel, Phys. Lett. B **267**, 438 (1991); H. Weigel, Int. J. Mod. Phys. A **11**, 2419 (1996).
[5] L. Carson, Nucl. Phys. A **535**, 479 (1991); T. Walhout, Nucl. Phys. A **547**, 423 (1992); T. Waindzoch and J. Wambach, Phys. Lett. B **226**, 163 (1992).
[6] R. A. Leese, N. S. Manton, and B. J. Schroers, Nucl. Phys. B **442**, 228 (1995).
[7] N. R. Walet, Nucl. Phys. A **586**, 649 (1995); E-print archive, hep-ph/9603273.
[8] A. P. Balachandran, F. Lizzi, V. Rodgers, and A. Stern, Nucl. Phys. B **256**, 525 (1985).
[9] R. L. Jaffe and C. L. Korpa, Nucl. Phys. B **258**, 468 (1985).
[10] J. Kunz and P. J. Mulders, Phys. Lett. B **215**, 449 (1988).
[11] V. B. Kopeliovich, Yad. Fiz. **47**, 1495 (1988) [Sov. J. Nucl. Phys. **47**, 949 (1988); Yad. Fiz. **51**, 241 (1990) [Sov. J. Nucl. Phys. **51**, 151 (1990)]; Phys. Lett. B **259**, 234 (1991); Yad. Fiz. **56**, 160 (1993).
[12] F. G. Scholtz, B. Schwesinger, and H. B. Geyer, Nucl. Phys. A **561**, 542 (1993).
[13] E. Guadagnini, Nucl. Phys. B **236**, 35 (1984).
[14] M. Praszalowicz, Phys. Lett. B **158**, 264 (1985); M. Chemtob, Nucl. Phys. B **256**, 600 (1985).
[15] V. B. Kopeliovich and B. E. Stern, JETP Lett. **45**, 203 (1987); N. S. Manton, Phys. Lett. B **192**, 177 (1987); J. J. M. Verbaarschot, Phys. Lett. B **195**, 235 (1987); E. Braaten and L. Carson, Phys. Rev. D **38**, 3525 (1988).
[16] V. B. Kopeliovich, B. Schwesinger, and B. E. Stern, JETP Lett. **62**, 185 (1995).
[17] J. Donoghue and C. Nappi, Phys. Lett. B **168**, 105 (1986); H. Yabu, Phys. Lett. B **218**, 124 (1989).
[18] D. I. Diakonov and V. Yu. Petrov, LNPI Preprint 967 (1984).
[19] V. B. Kopeliovich, B. Schwesinger, and B. E. Stern, Phys. Lett. B **242**, 145 (1990); Nucl. Phys. A **549**, 485 (1992).
[20] V. B. Kopeliovich, JETP Lett. **64**, 426 (1996).
[21] E. Witten, Nucl. Phys. B **223**, 422,433 (1983).
[22] J. J. de Swart, Rev. Mod. Phys. **35**, 916 (1963).
[23] I. Zahed, A. Wirzba, and U.-G. Meissner, Phys. Rev. D **33**, 830 (1986).
[24] R. V. Konoplich, A. E. Kudryavtsev, R. V. Martemyanov, and S. G. Rubin, Hadronic J. **11**, 271 (1988).
[25] B. Moussalam, Ann. Phys. (N.Y.) **225**, 264 (1993).
[26] G. Holzwarth and H. Walliser, Nucl. Phys. A **587**, 721 (1995); F. Meier and H. Walliser, E-print archive, hep-ph/9602359.
[27] H. Yabu and K. Ando, Nucl. Phys. B **301**, 601 (1988); M. Praszalowicz, *Proceedings of the International Workshop Baryons as Skyrme Solitons*, Siegen, 28–30 September (1992), ed. G. Holzwarth, p. 43–55.
[28] B. Schwesinger, F. G. Scholtz, and H. B. Geyer, Phys. Rev. D **51**, 1228 (1995).
[29] R. L. Jaffe, Phys. Rev. Lett. **38**, 195 (1977).
[30] T. Goldman *et al.*, Phys. Rev. Lett. **59**, 627 (1987); F. Wong, J.-I. Ping, G.-H. Wu, L.-j. Teng, and T. Goldman, Phys. Rev. C **51**, 3411 (1995); E-print archive, nucl-th/9512014.
[31] V. B. Kopeliovich and B. E. Stern, *SU*(3) *skyrmions. Proceedings of the International Conference Quarks 96''* Yaroslavl'. May (1996), E-print archive, hep-th/9612211.
[32] J. T. Reed *et al.*, Phys. Rev. **168**, 1495 (1969); W. G. Hogan *et al.*, Phys. Rev. **166**, 1472 (1968).
[33] *Proceedings of the 25th INS Symposium*, Tokyo, Japan, Dec. 3–7 (1996).

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

The diffraction mechanism of electron–positron pair photoproduction

A. I. Akhiezer and N. P. Merenkov

Kharkov Physicotechnical Institute, National Scientific Center, 310108 Kharkov, Ukraine*
(Submitted 2 June 1997)
Zh. Éksp. Teor. Fiz. **112**, 1959–1965 (December 1997)

We examine a mechanism of electron–positron pair photoproduction near a giant resonance. The mechanism is based on the possibility of photons being absorbed by nuclei. Our calculations set the cross section of this process at roughly 10^{-30} cm^2. We also discuss the feasibility of observing the effect. © *1997 American Institute of Physics.* [S1063-7761(97)00312-0]

1. INTRODUCTION

A free photon cannot produce an electron–positron pair since such a process is forbidden by energy and momentum conservation laws. But if a photon meets an obstacle, say a nucleus that can absorb it, photon diffraction occurs, with the result that the diffracted photon of an appropriate energy can produce an electron–positron pair. In this paper we study such a mechanism of electron–positron pair production, a mechanism based on the possibility of a photon being absorbed by a nucleus.

Before discussing the mechanism of pair production in photoabsorption, let us touch on the phenomenon of photon absorption by nuclei.[1]

For photon energies lower than the energy of nucleon detachment from the nucleus ($\omega<8$ MeV) the photoabsorption cross section $\sigma_\gamma(\omega)$ is zero. Above the nucleon knockout threshold, for photon energies in the 8–15 MeV range the cross section $\sigma_\gamma(\omega)$ is finite but small, $\sigma_\gamma(\omega)\approx 10^{-27}-10^{-26}$ cm^2, and slowly increases with energy in this range. For photon energies in the 15–30 MeV range there is a high and broad maximum in the photoabsorption cross section, with $\sigma_\gamma(\omega)\approx(1-1.5)\times 10^{-25}$ cm^2 and a halfwidth γ equal to 4–7 MeV, which is known as a giant resonance. For photon energies above 30 MeV the photoabsorption cross section is again small (as it is for $\omega<15$ MeV) and decreases with increasing photon energy.

Here we are particularly interested in the giant resonance region, where the prevailing photoabsorption mechanism is that of collective absorption of photons by a nucleus. Such a process leads to specifically collective nuclear movements, where the protons of the nucleus oscillate with respect to the neutrons of the same nucleus due to the variable electric field of the photon (the Migdal–Goldhaber–Teller model).

The giant-resonance region is extremely interesting in studies of the photoproduction of electron–positron pairs. On the one hand, because of the large value of $\sigma_\gamma(\omega)$, the cross section of this process is large (estimates have shown that this value can compete with radiative corrections to the cross section of pair production in the Coulomb field of a nucleus). On the other hand, the energies of the various particles participating in the process (a photon, an electron and a positron) are such that we can employ the formulas of the ultrarelativistic approximation. In this case the matrix element of the pair photoproduction process is expressed in terms of the scattering amplitude of a photon that can be absorbed by the nucleus.[2]

2. THE SCATTERING AMPLITUDE OF A PHOTON THAT CAN BE ABSORBED BY A NUCLEUS

Now let us discuss the scattering of a photon in the presence of a nucleus. Several effects must be distinguished here. First there is the purely quantum-electrodynamic effect of photon scattering in the electrostatic field of the nucleus, or Delbrück scattering. Akhiezer and Pomeranchuk[3] were the first to develop the theory of such scattering. Then there is photon scattering related to the possibility of a photon of an appropriate energy being transformed into an electron–positron pair. The theory of this process was developed by Bethe and Rohrlich.[4] Finally, there is scattering due to purely nuclear effects of photon absorption. Migdal[5] was the first to detect and study the corresponding mechanism. In what follows we deal with the third effect in the giant-resonance region. But first, following Bethe and Rohrlich,[4] we discuss the general theory of photon scattering related to photoabsorption.

When the energies are high and the scattering angles θ are small, the amplitude of coherent photon scattering (i.e., when the photon frequency does not change) can be written as follows:[4]

$$f(\omega,\theta)=k\int bdbJ_0(bk\theta)(\alpha_1(b,\omega)+i\alpha_2(b,\omega)), \qquad (1)$$

where $k\theta=|k_\perp|$, $k=\omega/c$ (below we use a system of units in which Planck's constant and the speed of light are equal to unity), $k_\perp$ is the transverse component of the 4-momentum of the scattered photon in relation to the 4-momentum $\tilde{k}$ of the initial photon, $J_0(x)$ is a Bessel function, and b is the impact parameter. The imaginary part of the scattering amplitude (1) is related to the function $\alpha_2(b,\omega)$, which in turn determines the photoabsorption cross section via the relationship

$$\sigma_\gamma=2\pi\int bdb\ 2\alpha_2(b,\omega), \qquad (2)$$

while the function $\alpha_1(b,\omega)$, which determines the real part of the amplitude, can be found via the simple dispersion relation

$$\alpha_1(b,\omega)=\frac{1}{\pi}P\int_{-\infty}^{\infty}\frac{d\omega'}{\omega'-\omega}\alpha_2(b,\omega'). \tag{3}$$

Allowing for the relationship between diffraction and absorption,[6] from now on we will assume that the function $\alpha_2(b,\omega)$, related to photon absorption by nuclei, is a step function,

$$\alpha_2(b,\omega)=\Theta(R(\omega)-b), \tag{4}$$

which leads to the following simple relationship:

$$\sigma_\gamma(\omega)=2\pi R^2(\omega). \tag{5}$$

In other words, $\sqrt{2}R(\omega)$ acts as the effective radius of the nucleus in relation to the absorption of a photon of frequency ω. Bearing in mind the nature of a giant resonance, for approximate estimates we can write the photoabsorption cross section in the giant-resonance region as

$$\sigma_\gamma(\omega)=\sigma_0\Theta(\omega_0+\gamma-\omega)\Theta(\omega_0-\gamma-\omega), \tag{6}$$

where, according to what was said in the Introduction,

$$\sigma_0\approx 1.5\times 10^{-25}\ \text{cm}^2,\quad \omega_0\approx 25\ \text{MeV},\quad \gamma\approx 5\ \text{MeV}.$$

Inserting (4) into (3), we get

$$\alpha_1(b,\omega)=\frac{\Theta(R_0-b)}{\pi}\ln\left|\frac{\omega_0+\gamma-\omega}{\omega_0-\gamma-\omega}\right|,\quad R_0=\sqrt{\frac{\sigma_0}{2\pi}}. \tag{7}$$

Clearly, for photon energies in the giant-resonance region, the function α_1 is proportional to $(\omega_0-\omega)/\gamma$, while in the high-energy region ($\omega\gg\omega_0$) it decreases in inverse proportion to frequency: $\alpha_1\propto\gamma/\omega$.

Substituting $\alpha_1(b,\omega)$ (Eq. (7)) and $\alpha_2(b,\omega)$ (Eq. (4)) in the right-hand side of Eq. (1), we arrive at the final expression for the photon scattering amplitude:

$$f(\omega,\theta)=\frac{k}{|k_\perp|}(N_0+iN_1), \tag{8}$$

where

$$N_0=\frac{1}{\pi}R_0J_1(R_0|k_\perp|)\ln\left|\frac{\omega_0+\gamma-\omega}{\omega_0-\gamma-\omega}\right|,$$

$$N_1=R(\omega)J_1(|k_\perp|R(\omega)),$$

with $J_1(x)$ a Bessel function.

We see that in the adopted step-function approximation for $\alpha_2(b,\omega)$ and with the photoabsorption cross section in the form (6) (the giant-resonance region), the real part of the photon scattering amplitude near the giant resonance may be of the same order of magnitude as the imaginary part.

Knowing the photon scattering amplitude, we can write an expression for the electromagnetic field 4-vector potential related to the scattered photon:

$$A_\mu(\mathbf{r},t)=\epsilon_\mu e^{-i\omega t}\frac{e^{ikr}}{r}f(\omega,\theta), \tag{9}$$

where ϵ_μ is the polarization 4-vector of the scattered photon, and k is the photon wavenumber.

3. THE AMPLITUDE AND CROSS SECTION OF ELECTRON–POSITRON PAIR PHOTOPRODUCTION RELATED TO PHOTON ABSORPTION

Now that we have an expression for the electromagnetic field 4-vector potential related to the scattered photon, we can find the matrix element reflecting the electron–positron pair production by the photon in the virtual state:

$$M=e\int dtd^3re^{-i\omega t}e^{i(q_++q_-)x}\frac{e^{ikr}}{r}f(\omega,\theta)\epsilon_\mu g_{\mu\nu}J_\nu, \tag{10}$$

where q_- and q_+ are the 4-momenta of the electron and positron, respectively, $g_{\mu\nu}$ is the metric tensor, e is the positron charge, and integration is over the entire four-dimensional space–time.

The electromagnetic current J_ν in (10) is given by the formula

$$J_\nu=\overline{u(q_-)}\gamma_\nu v(q_+),$$

where $u(q_-)$ and $v(q_+)$ are the electron and positron spinor amplitudes.

Integration of the right-hand side of Eq. (10) over space and time yields

$$M=\frac{4\pi e}{\omega^2-|\mathbf{q}_-+\mathbf{q}_+|^2}f(\omega,\theta)2\pi\delta(\omega-\varepsilon_--\varepsilon_+)\tilde{M}, \tag{11}$$

where $\tilde{M}=\epsilon_\mu g_{\mu\nu}J_\nu$, and ε_- and ε_+ are the energies of the electron and positron, respectively.

To calculate $\tilde{M}$ and the quantity $|\tilde{M}|^2$, which enters into the cross section, we proceed as follows. First we write the metric tensor in the form[2]

$$g_{\mu\nu}=g^\perp_{\mu\nu}+\frac{2}{s}q_\nu\tilde{k}_\mu,\quad s=2(\tilde{k}q)\gg m^2, \tag{12}$$

where $g^\perp_{\mu\nu}$ is the transverse part of the metric tensor, $\tilde{k}_\mu$ is the 4-momentum of the initial photon, and m is the electron mass. The 4-vector q_μ determines s, which does not enter into the final formula (as we will see shortly). Obviously,

$$\tilde{M}=(\epsilon J)_\perp+\frac{2}{s}(\epsilon\tilde{k})(Jq).$$

Following Sudakov,[7] we decompose the electron and positron 4-momenta into longitudinal and transverse components with respect to the 4-vectors $\tilde{k}$ and q:

$$q_+=\alpha_+q+(1-x)\tilde{k}+q^\perp_+,$$

$$q_-=\alpha_-q+x\tilde{k}+q^\perp_-,\quad \alpha_+,\alpha_-\ll 1, \tag{13}$$

where the variable x is the fraction of the energy carried away by the positron, and

$$(q^\perp_\pm q)=(q^\perp_\pm\tilde{k})=0.$$

Using the explicit expression for the current J_ν, we can derive a formula for the contribution of the purely traverse polarization to $|\tilde{M}|^2$ after summing over the spin states of the electron–positron pair:

$$(\epsilon J)^2_\perp=2[\boldsymbol{\epsilon}^2k^2+4(\mathbf{q}_+\cdot\boldsymbol{\epsilon})(\mathbf{q}_-\cdot\boldsymbol{\epsilon})], \tag{14}$$

where $k^2=(q_++q_-)^2$, and we have introduced Euclidean vectors: $\boldsymbol{\epsilon}^2=-\boldsymbol{\epsilon}_\perp^2$ and $\mathbf{q}_\pm^2=-(q_\pm^\perp)^2$.

Similarly, the interference of the longitudinal and transverse polarizations yields a contribution

$$\frac{4}{s}(\epsilon\tilde{k})(\epsilon J)_\perp(Jq)=-8[(1-x)(\boldsymbol{\epsilon}\cdot\mathbf{q}_-)+x(\boldsymbol{\epsilon}\cdot\mathbf{q}_+)](\epsilon\tilde{k}), \tag{15}$$

while the contribution of the purely longitudinal polarization of the scattered virtual photon is

$$\frac{4}{s^2}(\epsilon\tilde{k})^2(Jq)^2=8x(1-x)(\epsilon\tilde{k})^2. \tag{16}$$

Now we express the quantity $(\epsilon\tilde{k})$ in (15) and (16) in terms of the transverse component of the polarization vector using the transversality condition for the electromagnetic field 4-vector potential of the scattered photon: $(\epsilon k)=0$. This yields

$$(\epsilon\tilde{k})=\mathbf{k}\cdot\boldsymbol{\epsilon},\quad \mathbf{k}=\mathbf{q}_-+\mathbf{q}_+. \tag{17}$$

Summing the right-hand sides of Eqs. (14)–(16) and taking (17) into account, we get

$$|\widetilde{M}|^2=2\boldsymbol{\epsilon}^2(k^2-2(x\mathbf{k}-\mathbf{q}_-)^2). \tag{18}$$

Allowing for all possible transverse polarizations, we must put $\boldsymbol{\epsilon}^2=2$. Note that Eq. (18) is valid only if $\omega\gg|\mathbf{k}|\gg m^2$.

Now that we have an expression for $|M|^2$, we can find the differential cross section for photoproduction of an electron–positron pair via the mechanism related to photon absorption by a nucleus:

$$d\sigma=\frac{1}{4\omega}|M|^2\frac{d^3q_-d^3q_+}{(2\pi)^5 4\varepsilon_-\varepsilon_+}\delta(\omega-\varepsilon_--\varepsilon_+). \tag{19}$$

Using the parametrization (13) for the electron and positron 4-momenta, we get

$$\frac{d^4\sigma}{dx\,d\varphi\,dq_-^2\,d\mathbf{k}^2}=\frac{\alpha}{4\pi\mathbf{k}^2}\left[\frac{1-2x(1-x)}{m^2+(x\mathbf{k}-\mathbf{q}_-)^2}+\frac{2m^2}{[m^2+(x\mathbf{k}-\mathbf{q}_-)^2]^2}\right](N_0^2+N_1^2), \tag{20}$$

where φ is the angle between the two-dimensional vectors $\mathbf{k}$ and $\mathbf{q}_-$. The reader will recall that the vector $\mathbf{k}$ is the total transverse momentum of the new pair. Integration with respect to the angle φ from 0 to 2π and with respect to the square of the transverse electron momentum from 0 to $x\mathbf{k}^2$ yields

$$\frac{d^2\sigma}{dx\,d\mathbf{k}^2}=\frac{\alpha}{2\mathbf{k}^2}\left\{[1-2x(1-x)]\ln\frac{\mathbf{k}^2x(1-x)}{m^2}+2x(1-x)\right\}(N_0^2+N_1^2). \tag{21}$$

After (21) is integrated with respect to the fraction x of the electron energy from 0 to 1, we arrive at the following formula for the distribution of pairs with respect to the square of their total transverse momentum:

$$\frac{d\sigma}{d\mathbf{k}^2}=\frac{\alpha(N_0^2+N_1^2)}{2\mathbf{k}^2}\left(\frac{2}{3}\ln\frac{\mathbf{k}^2}{m^2}-\frac{10}{9}\right). \tag{22}$$

When the photon energy is in the vicinity of the giant resonance, $R_0|\mathbf{k}|$ is always much smaller than unity, since $|\mathbf{k}|\leqslant\omega$. Hence we can expand the Bessel function present in N_0 and N_1 in a series and keep only the first-order terms: $J_1(x)\approx x/2$, with $x\ll 1$.

After this the differential distribution (22) can be integrated with respect to $\mathbf{k}^2$. The greatest contribution is provided by the region $\mathbf{k}^2\sim\omega^2$, so that the total cross section of electron–positron production by the mechanism considered here is

$$\sigma=\frac{\alpha\omega^2}{8}\left[R^4(\omega)+\frac{1}{\pi^2}R_0^4\ln^2\left|\frac{\omega_0-\gamma-\omega}{\omega_0+\gamma-\omega}\right|\right]\times\left(\frac{2}{3}\ln\frac{\omega^2}{m^2}-\frac{16}{9}\right). \tag{23}$$

In the giant-resonance region at $\omega=\omega_0$ this cross section is approximately 2×10^{-30} cm^2.

Note that for $\mathbf{k}^2\gg m^2$ the differential (in the square of the transverse momentum transfer) cross section of electron–positron pair production in the Coulomb field of the nucleus is given by the formula[8]

$$\frac{d\sigma_k}{d\mathbf{k}^2}=\frac{8\alpha^3Z^2}{3\mathbf{k}^4}\left[(1-2x((1-x))\ln\frac{\mathbf{k}^2}{m^2}+2\right], \tag{24}$$

which is applicable in the point-nucleus approximation and holds as long as

$$\frac{1}{|\mathbf{k}|}>R,\quad R=1.5A^{1/3}\times10^{-13}\text{ cm}^2, \tag{25}$$

where A is the atomic number of the nucleus. The condition (25) corresponds to impact parameters (at which pair production is possible) that are larger than the radius of the nucleus. For nuclei with $A\approx100$ the condition (25) is met up to $|\mathbf{k}|\approx30$ MeV.

Of course, the main contribution to the total cross section σ_k is provided by values of $\mathbf{k}^2$ of order m^2, so that the ratio of the total cross sections corresponding to the Coulomb mechanism and the mechanism considered here is very large: $\sigma_k/\sigma\approx10^3Z^2$. However, for differential (in $\mathbf{k}^2$) cross sections the ratio may be considerably smaller:

$$\frac{d\sigma_k}{d\sigma}\approx\frac{Z^2}{2},\quad \mathbf{k}^2=400\text{ MeV}^2. \tag{26}$$

In these conditions the cross section of electron–positron pair production due to photoabsorption by nuclei is larger than one percentage point (for nuclei with $Z\approx10$), which is roughly equal to the contribution of radiative corrections to Coulomb scattering.[9]

Note that for transverse momenta of the order of, or greater than, 100 MeV, when the photon finds itself inside the nuclei, Eq. (24) requires substantial modification since the nuclei cannot be interpreted as a point object. For instance, if the nucleus of radius R has a constant charge density, the electrostatic potential of such of such a nucleus is

TABLE I.

y	0.1	0.5	1	5	10	50
$C^2(y)$	0.984	0.856	0.504	0.202	0.0441	0.0001

$$U(r)=\begin{cases}\dfrac{Ze}{r}, & r\geqslant R,\\ \dfrac{Zer^2}{R^3}, & r<R.\end{cases} \tag{27}$$

The Fourier transform of the potential $U(r)$ is

$$U(q)=\frac{4\pi Ze}{q^2}C(y),$$

$$C(y)=\frac{3}{y}\left[\left(1-\frac{2}{y^2}\right)\sin y+\frac{2}{y}\cos y\right],\quad y=qR. \tag{28}$$

Clearly, as $R\to 0$, i.e., in the point-nucleus limit, $C(y)$ is equal to unity. In this limit the differential cross section of pair production in a Coulomb field is given by Eq. (24). But if Rq is of the order of, or larger than, unity, Eq. (24) must by modified by multiplying it into $C^2(y)$. Some values of $C^2(y)$ are listed in Table I. Since the correction decreases substantially as y increases, we cannot rule out the possibility that for pair traverse momenta of several hundred MeV both pair production mechanisms will compete as equals.

The authors would like to express their gratitude to A. G. Sitenko and V. K. Tartakovskiĭ for the fruitful discussion concerning photoabsorption by nuclei. The work was partially supported by the International Soros Science Education Program (A.I.A.) and an INTAS Grant No. 93-1867 (N.P.M.).

*E-mail: nsc@kipt.kharkov.ua

[1] J. M. Eisenberg and W. Greiner, *Nuclear Models*, 2nd ed., North-Holland, Amsterdam (1975).
[2] A. B. Arbuzov, E. A. Kuraev, N. P. Merenkov, and L. Trentadue, Nucl. Phys. B **474**, 271 (1996).
[3] A. I. Akhiezer and I. Ya. Pomeranchuk, Zh. Éksp. Teor. Fiz. **7**, 567 (1937).
[4] H. A. Bethe and F. Rohrlich, Phys. Rev. **86**, 10 (1952).
[5] A. Migdal, Zh. Éksp. Teor. Fiz. **15**, 81 (1945).
[6] A. I. Akhiezer and I. Ya. Pomeranchuk, Zh. Éksp. Teor. Fiz. **16**, 396 (1946); G. Placzek and H. Bethe, Phys. Rev. **57**, 1072 (1940).
[7] V. V. Sudakov, Zh. Éksp. Teor. Fiz. **30**, 87 (1956) [Sov. Phys. JETP **3**, 65 (1956)].
[8] V. G. Zima and N. P. Merenkov, Yad. Fiz. **24**, 998 (1976) [Sov. J. Nucl. Phys. **24**, 522 (1976)].
[9] E. A. Vinokurov, É. A. Kuraev, and N. P. Merenkov, Zh. Éksp. Teor. Fiz. **66**, 1916 (1974) [Sov. Phys. JETP **39**, 942 (1974)].

Translated by Eugene Yankovsky

Ionization of hydrogen and helium atoms by highly charged fast ions in collisions with small momentum transfer

A. B. Voĭtkiv

Electronics Institute, Uzbek Academy of Sciences, 700143 Tashkent, Akademgorodok, Uzbek Republic
(Submitted 15 March 1996; resubmitted 17 June 1997)
Zh. Éksp. Teor. Fiz. **112**, 1966–1977 (December 1997)

Ionization of hydrogen and helium atoms is studied for the case of ''soft'' collisions with highly charged fast ions with $v \lesssim Z \ll v^2$ and $v \gg v_0$, where Z is the ion charge, v is the collision velocity, and $v_0 \sim 1$ is the characteristic velocity of the electron in the ground state of the atom. Analytical expressions are derived for the singly and doubly differential cross section for ionization of a hydrogen atom accompanied by the ejection of a slow electron ($v_e \lesssim v_0$, where v_e is the velocity of the ejected electron with respect to the recoil ion). The results are generalized to the case of single ionization of helium. It is shown that soft collisions provide the main contribution to the hydrogen ionization cross section and for all practical purposes determine the cross section for single ionization of helium. The asymmetry in the angular distribution of the ejected slow electrons and the properties of momentum exchange in such collisions are discussed. Finally, a formula for the cross section for single ionization of helium is proposed. © *1997 American Institute of Physics.* [S1063-7761(97)00412-5]

1. INTRODUCTION

The study of collisions of atoms and highly excited fast ions is of interest in many fields of physics (atomic physics, solid-state physics, etc.) and in related fields (e.g., biophysics). Often the charges Z of these ions are so large that they satisfy $Z \gtrsim v$, notwithstanding the high values of the ion velocity v ($v \gg v_0$, with $v_0 = 1$ a.u. $\simeq 2 \times 10^8$ cm s^{-1}; unless stated otherwise, in what follows we use the atomic system of units). There is a great number of experimental and theoretical papers devoted to the study of the total cross sections for single, double, and multiple ionization of atoms in collisions with such ions (see, e.g., Refs. 1–3 and the literature cited therein). Detailed information about the collision of a highly charged fast ion and an atom can be obtained by studying the various differential cross sections. Significant progress in developing the experimental techniques has made it possible to carry out what is known as kinematically complete experiments in studying the collisions of fast charged particles and atoms,[4,5] experiments in which not only the total ionization cross sections are determined but also the angular and energy distributions of the electrons leaving the atom, the momenta and energies of the recoil ions, etc. Such studies of the process of single ionization of helium atoms in collisions with a highly charged fast ion ($Z = 24$ and $v = 12$), a process in which the momentum transferred to the atom is small (in what follows we call such collisions ''soft''), were conducted by Moshammer *et al.*,[5] who investigated the behavior of the slow electrons, whose energies after leaving the atom do not substantially exceed the atom's ionization potential, and the balance of momenta in the system consisting of the highly charged ion, the electron, and the recoil ion (the researchers also calculated the collision process by the classical-trajectory Monte Carlo method). Studying the characteristics of slow electrons is important because such electrons constitute the bulk of electrons ejected by atoms in collisions with highly charged fast ions in the region of the parameters Z and v considered here.

The plan of the paper is as follows. In Sec. 2 we discuss the angular and energy distributions of slow electrons ejected by hydrogen atoms in collisions with highly charged fast ions with the following parameters: $v \lesssim Z \ll v^2$, $v \gg v_0$, and $v_e \lesssim v_0$, where v_e is the velocity of the ejected electron with respect to the recoil ion. Section 3 is devoted to a semiqualitative generalization of these results to the case of single ionization of helium. Finally, there is a brief discussion, by analogy with the photoionization process, of the balance of momenta in soft collisions.

It is a well-known fact (see, e.g., Ref. 6) that classical mechanics cannot be used to describe soft collisions of a charged particle and an atom.[1)] The quantum mechanical approach adopted in this paper is somewhat less rigorous than the various modifications of the distorted wave approximation (see, e.g., Ref. 7) but has one important advantage: it makes it possible to solve the problem analytically, and the final result is fairly simple. The main approximation in this approach is the grouping (in the impact parameter) of collisions into hard collisions, where the energy transferred to the electron is on the average high, $\varepsilon > I$ ($\varepsilon \gg I$), and soft collisions, with $\varepsilon < I$ (I is the ionization potential). Of course, no strict boundary between these regions can be drawn, but the indeterminancy in the ''boundary'' value of the impact parameter is not very important for the final expressions.

2. COLLISIONS WITH HYDROGEN ATOMS

Suppose that initially a hydrogen atom with its nucleus at the origin of coordinates is in its ground state and that a highly charged structureless ion is moving along a classical linear trajectory $\mathbf{R}(t) = \mathbf{b} + \mathbf{v}t$, where $\mathbf{b}$ is the impact param-

 1063-7761/97/121074-07$10.00

eter vector. Let us partition the entire range of the impact parameters $0 \leqslant b < \infty$ into two subregions: (1) $b < Z/v$, and (2) $b > Z/v$.

In collisions in the first subregion the energy transferred to the electron is on the average much higher than the ionization potential of the atom. Indeed, for $b > 1$ this energy can be estimated at $\varepsilon(b) \simeq 2Z^2/b^2v^2$ (see Ref. 8), i.e., it is high already at $b \simeq Z/v > 1$ and rapidly increases as the impact parameter decreases (e.g., for ions with $Z = 24$ and $v = 12$, which were used in the experiment of Moshammer *et al.*,[5] $\varepsilon(b=1) = 8$) for $b < 1$ the average transferred energy is even higher. For this reason (and because of the small size of the region $b < Z/v$) the contribution of hard collisions to the emission of slow electrons is small, so that we will not take such collisions into account in the future (see, e.g., Refs. 6, 7, 9, and 10 and the literature cited therein for discussions of the spectra of electrons ejected in collisions with large momentum transfer).

Calculations in various approximations[11–13] for the region of impact parameters $b > Z/v > 1$ predict a rapid decrease in the ionization probability as b increases, with the result that even at $b \simeq (1.5-2)Z/v$ the ionization probability becomes much smaller than unity. To describe the transitions of atoms in such collisions we use the scattering matrix approach, in which the transition amplitude has the form (see, e.g., Refs. 14 and 15 and the literature cited therein)

$$A_{\mathbf{k}} = -i \int_{-\infty}^{\infty} dt \langle \psi_{\mathbf{k}}^{(-)}(\mathbf{r},t) | W(\mathbf{r},t) | \varphi_0(\mathbf{r},t) \rangle. \tag{1}$$

Here $\varphi_0(\mathbf{r},t) = \varphi_0(\mathbf{r})\exp(-it/2)$, where $\varphi_0(\mathbf{r}) = \pi^{-1/2} \times \exp(-r)$ is the wave function of the ground state of hydrogen, $\Psi_{\mathbf{k}}^{(-)}(\mathbf{r},t)$ is the wave function of the electron in the final state, with both the field of the atomic nucleus and the field of the highly charged ion acting on the electron,

$$W(\mathbf{r},t) = \frac{Z}{|\mathbf{R}(t)|} - \frac{Z}{|\mathbf{R}(t)-\mathbf{r}|}$$

is the interaction between the atom and the field of the highly charged ion, and $\mathbf{r}$ is the electron radius vector.

Let us estimate the effect of the two centers on the electron in the final state $\psi_{\mathbf{k}}^{(-)}(\mathbf{r},t)$ assuming that classical forces act on the electron, i.e., the forces representing the action of the highly charged fast ion (F_i) and of the atomic nucleus (F_a). As a function of time, in the region occupied by the atom the field of the highly charged fast ion has a maximum with its center at the point $t = 0$ and an effective width $I \simeq b/v$, with $b > 1$ (see Refs. 16 and 8). In the region of impact parameters $b > v$ the field of the highly charged ion is not only extremely weak for the atom (for $v^2 \gg Z$) but also varies adiabatically slowly, so that the probability of the atom becoming ionized is exponentially low (see, e.g., Ref. 11). But for collisions with $Z/v < b < v$ (which provide the main contribution to the emission of slow electrons), we have $T < 1$, i.e., at such impact parameters the field of the highly charged ion has a sharp maximum in the times $|t| < T$, when ionization mostly occurs. Then the distance between the proton and the departing electron can be estimated at $\overline{v}t$ ($t > 0$), where $\overline{v}$ is the average velocity at which the electron travels through the region of space $r \sim 1$ and which for slow electrons is, in order of magnitude, equal to $v_0 = 1$. The distance between the highly charged ion and the electron for $t > 0$ can be assumed to be proportional to the difference in their velocities: $|\mathbf{v} - \overline{\mathbf{v}}|t \simeq vt$. Thus, for the ratio of the forces we have[2)]

$$\frac{F_i}{F_a} \sim \frac{Zv_0^2}{v^2} = \frac{Z}{v^2},$$

which implies that for $Z/v^2 \ll 1$ the behavior of the slow electron in the final state is controlled mainly by the field of the atomic nucleus. Hence the effect of the Coulomb interaction between the highly charged fast ion and such an electron can be taken into account approximately. Below we will see that the main effect is that the angular distribution of the slow electrons becomes asymmetric: most electrons are dragged by the Coulomb attraction of the flying highly charged ion in the direction of the ion's motion.

There is one more effect that leads to asymmetry in the angular distribution of the slow electrons. If we expand the potential of the interaction of the atom and the field of the highly charged ion, $z/|\mathbf{R}(t)| - Z/|\mathbf{R}(t)-\mathbf{r}|$, as a function of electron coordinates and time in monochromatic plane waves, we see that the ion field carries a longitudinal (i.e., parallel to the ion velocity) momentum $q_p \sim 1/v$, with the magnitude and direction of the momentum being independent (for $Z \neq 0$) of the magnitude and sign of the charge of the incident particle (what is important is that if we allow only for the dipole term in the expansion of this interaction, the calculated value of q_p vanishes). The absorption of this momentum by an atomic electron leads (as we will see shortly) to additional asymmetry in the angular distribution of the slow electrons.

In the region of impact parameters $b > Z/v > 1$ the interaction $W(\mathbf{r},t)$ can be written as

$$W(\mathbf{r},t) = W_1(\mathbf{r},t) + W_2(\mathbf{r},t), \tag{2}$$

where

$$W_1(\mathbf{r},t) = -\frac{Z(vtz+by)}{R^3(t)} = -\frac{Z\mathbf{R}(t)\cdot\mathbf{r}}{R^3(t)} = -\mathbf{E}(t)\cdot\mathbf{r},$$

$$W_2(\mathbf{r},t) = \frac{Zr^2}{2R^3(t)} - \frac{3Z}{2R^5(t)}(vtz+by)^2. \tag{3}$$

In Eqs. (3), the z axis is directed along the velocity vector of the highly charged ion and the y axis along the impact parameter vector. In accordance with what was said earlier, the expansion in (2) is exact to within quadrupole terms.

The electron behavior in the field state in the field of two centers is determined by the Hamiltonian

$$H = -\frac{\Delta}{2} - \frac{1}{r} - W \simeq -\frac{\Delta}{2} - \frac{1}{r} - W_1 - W_2. \tag{4}$$

In accordance with the above idea of the relative role of the two centers in the final state, the interaction of a slow electron with the atomic nucleus will be taken into account ex-

actly and its interaction with the field of the highly charged fast ion will be treated in the zeroth approximation of the theory of sudden perturbations:[17]

$$\psi_{\mathbf{k}}^{(-)}(\mathbf{r},t)=\varphi_{\mathbf{k}}^{(-)}(\mathbf{r})\exp\left(-\frac{ik^2t}{2}\right)\exp\left[i\int_{-\infty}^{t}dt'\,W(t')\right], \tag{5}$$

where $\varphi_{\mathbf{k}}^{(-)}(\mathbf{r})$ is the Coulomb wave function for electron–proton scattering, which in the limit $r\to\infty$ is a linear combination of the incident plane wave and a convergent spherical wave and is normalized by the condition $\langle\varphi_{\mathbf{k}}^{(-)}|\varphi_{\mathbf{k}'}^{(-)}\rangle=(2\pi)^{-3}\delta(\mathbf{k}-\mathbf{k}')$, where $\mathbf{k}$ is the wave vector of electron motion with respect to the atomic nucleus. Plugging (5) into (1) yields

$$A_{\mathbf{k}}=-i\left\langle\varphi_{\mathbf{k}}^{(-)}(\mathbf{r})\left|\int_{-\infty}^{\infty}dt W(\mathbf{r},t)\right.\right.$$
$$\left.\left.\times\exp\left[i\omega t-i\int_{-\infty}^{t}dt'\,W(t')\right]\right|\varphi_0(\mathbf{r})\right\rangle, \tag{6}$$

where $\omega=(1+k^2)/2$ is the transition frequency.

For $b>Z/v>1$ the term

$$\int_{-\infty}^{t}dt'\,W(t')$$

is small compared to unity for all values of t. Expanding the exponential function (6) in a series and keeping only the leading terms, we get

$$A_{\mathbf{k}}=A_{\mathbf{k}}^{d}+A_{\mathbf{k}}^{q}+A_{\mathbf{k}}^{i}, \tag{7}$$

where

$$A_{\mathbf{k}}^{d}=-i\left\langle\varphi_{\mathbf{k}}^{(-)}(\mathbf{r})\left|\int_{-\infty}^{\infty}dt W_1(\mathbf{r},t)\exp(i\omega t)\right|\varphi_0(\mathbf{r})\right\rangle,$$
$$A_{\mathbf{k}}^{q}=-i\left\langle\varphi_{\mathbf{k}}^{(-)}(\mathbf{r})\left|\int_{-\infty}^{\infty}dt W_2(\mathbf{r},t)\exp(i\omega t)\right|\varphi_0(\mathbf{r})\right\rangle \tag{8}$$

are the dipole and quadrupole transition amplitudes, respectively, and in

$$A_{\mathbf{k}}^{i}=-i\left\langle\varphi_{\mathbf{k}}^{(-)}(\mathbf{r})\left|\int_{-\infty}^{\infty}dt\,\exp(i\omega t)\right.\right.$$
$$\left.\left.\times(\mathbf{q}(t)\cdot\mathbf{r})W_1(t)\right|\varphi_0(\mathbf{r})\right\rangle \tag{9}$$

we have allowed for the leading term in the interaction between the slow electron in the final state and the field of the highly charged fast ion:

$$\mathbf{q}(t)=\int_{-\infty}^{t}dt\mathbf{E}(t).$$

For the probability of an electron going into a state with a definite values of $\mathbf{k}$ as a result of a collision with an impact parameter $\mathbf{b}$ we have the following expression:

$$w(\mathbf{k},\mathbf{b})=|A_{\mathbf{k}}|^2\simeq|A_{\mathbf{k}}^{d}|^2+2\,\mathrm{Re}(A_{\mathbf{k}}^{d})\mathrm{Re}(A_{\mathbf{k}}^{q}+A_{\mathbf{k}}^{i})$$
$$+2\,\mathrm{Im}(A_{\mathbf{k}}^{d})\mathrm{Im}(A_{\mathbf{k}}^{q}+A_{\mathbf{k}}^{i}). \tag{10}$$

Using the explicit expressions for the interactions W_1 and W_2 and the Coulomb wave functions and averaging (in view of the geometry of the problem) the probability (10) over the angles φ of electron departure ($0\leqslant\varphi<2\pi$, with φ is the azimuthal angle in the plane of the impact parameter), we get

$$w(k,\theta,b)=\frac{1}{2\pi}\int_0^{2\pi}d\varphi\,w(\mathbf{k},\mathbf{b})$$
$$=\frac{2Z^2\omega^2}{v^4}\alpha(k)\left\{2K_0^2(\xi)\cos^2\theta+K_1^2(\xi)\sin^2\theta\right.$$
$$+\frac{8k}{v}\cos\theta[K_1^2(\xi)\sin^2\theta+K_0^2(\xi)(3\cos^2\theta$$
$$-1)]+\frac{4Z}{v^2}\frac{\cos\theta}{\omega}\left[K_0(\xi)K_1(\xi)-\frac{\pi k}{2}\right.$$
$$\times\exp(-\xi)(2K_1(\xi)\sin^2\theta+K_0(\xi)$$
$$\left.\left.\times(3\cos^2\theta-1))\right]\right\}, \tag{11}$$

where the angle θ of electron departure ($0\leqslant\theta\leqslant\pi$) is measured from the direction of the ion velocity vector, $\xi=\omega b/v$, K_0 and K_1 are modified Bessel functions,[18] and

$$\alpha(k)=\frac{2^7k^{-1}}{(1+k^2)^5}\exp\left(-\frac{4}{k}\arctan k\right)$$
$$\times\left[1-\exp\left(-\frac{2\pi}{k}\right)\right]^{-1}. \tag{12}$$

For the probability of ionization with a slow electron being ejected in a collision with a fixed impact parameter we have

$$w(b)=\int_0^{k_{\max}}dk\,k^2\int d\Omega\;w(k,\theta,b), \tag{13}$$

where $k_{\max}\simeq v_0=1$ and $d\Omega=2\pi\sin\theta d\theta$.

The doubly differential (in the departure angle and in the energy $E=k^2/2$) ionization cross section is given by the following expression:

$$\frac{d^2\sigma}{dEd\Omega}=2\pi\int_{b_{\min}}^{\infty}db\;b\sqrt{2E}w(k(E),\theta,b), \tag{14}$$

where $b_{\min}=\lambda Z/v\ll v$ (with λ a constant of order unity) is the lower limit of the region of impact parameters in which the ionization probability is substantially smaller than unity. If in (14) we integrate with respect to the impact parameter, we get

$$\frac{d^2\sigma}{dEd\Omega}=2^8\frac{Z^2}{v^2}\frac{1}{(1+2E)^5}\frac{\exp[-(4/\sqrt{2E})\arctan\sqrt{2E}]}{1-\exp(-2\pi/\sqrt{2E})}$$
$$\times\left\{\sin^2\theta\ln\beta+\cos^2\theta-0.5\sin^2\theta\right.$$
$$+\frac{2^{3.5}\sqrt{E}}{v}\cos\theta(\sin^2\theta\ln\beta+\cos 2\theta)$$
$$+\frac{2Z}{v^2}\cos\theta[\ln^2\beta-2^{1.5}\pi\sqrt{E}(\sin^2\theta\ln\beta$$

$$+\cos 2\theta - 0.5\ \sin^2\ \theta)]\Bigg\}, \tag{15}$$

where $\beta = 1.123v[b_{\min}(E+0.5)]^{-1} = (1.123/\lambda)(v^2/Z) \times (E+0.5)^{-1}$.

Since β, determined to within a constant factor of order unity, contains the large cofactor v^2/Z and enters into (15) under the logarithm sign, we can simply set λ to unity in the definition of β. Note that the accuracy of this approach increases with the geometric dimensions of the region of impact parameters $Z/v < b < v$ (the treatment is asymptotically exact in the formal limit $v \to \infty$, $Z \to \infty$, and $v^2/Z \to \infty$, provided that the condition $Z/v \gtrsim 1$ remains valid).

The electron energy distribution is determined by the following differential cross section:

$$\frac{d\sigma}{dE} = \int d\Omega \frac{d^2\sigma}{dE\,d\Omega} = \frac{2^{11}\pi}{3}\frac{Z^2}{v^2}\frac{1}{(1+2E)^5} \times \frac{\exp[-(4/\sqrt{2E})\arctan\sqrt{2E}]}{1-\exp(-2\pi/\sqrt{2E})} \ln\frac{2.25v^2}{Z(1+2E)}. \tag{16}$$

We see that the probability of electron ejection rapidly decreases with increasing electron energy: the majority of the ejected electrons (approximately 90%) have energies less than the ionization potential of the atom, $I_0 = 0.5$. Note that in the adopted approximation only the dipole transitions between the states φ_0 and $\varphi_{\mathbf{k}}^{(-)}$ of the atom contribute to (16); the transitions between φ_0 and $\varphi_{\mathbf{k}}^{(-)}$ due to the interaction $W_2(\mathbf{r},t)$ (which induces quadrupole electron transitions) and the transitions to final states distorted by the field of the highly charged ion and responsible for the asymmetry of the angular distribution of the slow electrons contribute nothing to (16), so that they have no effect (in the given approximation) on the total number of the ejected slow electrons.

To obtain the angular distribution of these electrons we must integrate (15) with respect to the energies E in $0 \leqslant E \lesssim I_0$. However, since the electron departure probability rapidly decreases with increasing E, the upper limit in the integral over energies can be formally set to infinity. This yields

$$\frac{d\sigma}{d\Omega} = \int_0^\infty dE \frac{d^2\sigma}{dE\,d\Omega} = 3\cdot 0.283\frac{Z^2}{v^2}\Bigg\{\sin^2\ \theta \ln\ \beta_1 + \cos^2\ \theta - 0.5\ \sin^2\ \theta + \frac{8\cdot 0.61}{v}\cos\ \theta(\sin^2\ \theta \ln\ \beta_2 + \cos 2\theta) + \frac{2Z}{v^2}\cos\ \theta[\ln^2\ \beta_1 + \langle \ln^2\ \omega\rangle - \ln^2\ \omega_1 - 2\pi\cdot 0.61(\sin^2\ \theta \ln\ \beta_2 + \cos 2\theta - 0.5\ \sin^2\ \theta)]\Bigg\}, \tag{17}$$

where we have employed the following notation:

$$\beta_1 = \frac{1.12v^2}{Z\omega_1}, \quad \beta_2 = \frac{1.12v^2}{Z\omega_2},$$

$$\omega_1 = \exp\left\{\int_0^\infty dk\ k^2\alpha(k)\ln(\omega)\left[\int_0^\infty dk\ k^2\alpha(k)\right]^{-1}\right\} = 0.71,$$

$$\omega_2 = \exp\left\{\int_0^\infty dk\ k^3\alpha(k)\ln(\omega)\left[\int_0^\infty dk\ k^3\alpha(k)\right]^{-1}\right\} = 0.88, \tag{18}$$

$$\langle \ln^2\ \omega\rangle = \int_0^\infty dk\ k^2\alpha(k)\ln^2(\omega)\left[\int_0^\infty dk\ k^2\alpha(k)\right]^{-1} = 0.234,$$

$\ln^2\ \omega_1 = 0.1$.

The terms in (17) proportional to Z^2/v^2 correspond to electron dipole transitions (the term $|A_{\mathbf{k}}^d|^2$ in (10)), and the terms proportional to Z^2/v^3 and Z^3/v^4, which lead to the asymmetry of the cross section $d\sigma/d\Omega$ with respect to the substitution $\theta \to \pi - \theta$, correspond to the terms $2\ \mathrm{Re}(A_{\mathbf{k}}^d)\mathrm{Re}(A_{\mathbf{k}}^q) + 2\ \mathrm{Im}(A_{\mathbf{k}}^d)\mathrm{Im}(A_{\mathbf{k}}^q)$ and $2\ \mathrm{Re}(A_{\mathbf{k}}^d)\mathrm{Re}(A_{\mathbf{k}}^i) + 2\ \mathrm{Im}(A_{\mathbf{k}}^d)\mathrm{Im}(A_{\mathbf{k}}^i)$ in the expansion (10). The asymmetry in electron ejection can be characterized by the quantity

$$\eta = \left(\int_0^{\pi/2} d\theta\ \sin\ \theta\frac{d\sigma}{d\Omega} - \int_{\pi/2}^{\pi} d\theta\ \sin\ \theta\frac{d\sigma}{d\Omega}\right) \times \left(\int_0^\pi d\theta\ \sin\ \theta\frac{d\sigma}{d\Omega}\right)^{-1}. \tag{19}$$

Combining (18) and (19), we find that

$$\eta \approx \frac{1.83}{v} + \frac{Z}{v^2}\left(1.5\ \ln\frac{1.6v^2}{Z} - 2.87 + 2.15\ \ln^{-1}\frac{1.6v^2}{Z}\right) \approx \frac{1.83}{v} + \frac{Z}{v^2}\left(1.5\ \ln\frac{v^2}{Z} - 2.87 + 2.15\ \ln^{-1}\frac{v^2}{Z}\right). \tag{20}$$

From (20) we see that most slow electrons are ejected by the atom (for $Z > 0$) in the direction of motion of the highly charged fast ion. The first term in (20), which is independent of the magnitude and sign of the ion charge, and the second term, which depends on both magnitude and sign (for $Z < 0$ we must take $|Z|$ under the logarithm sign) reflect, respectively, the absorption of longitudinal momentum $q_p \sim 1/v$ by the atom and the dragging of the departing electron by the electric field of the moving highly charged fast ion (we discussed these effects earlier). Note that the simple additivity of these effects in (20) is a consequence of the expansions (7) and (10), where the terms leading to angular asymmetry (but which are not the dominant terms for complete emission) are taken into account in the first nonvanishing approximation.

Combining (16) and (17), we arrive at the following expression for the contribution to the ionization cross section of hydrogen atoms from collisions in which slow electrons are ejected:

$$\Delta\sigma \approx 8\pi\cdot 0.283\frac{Z^2}{v^2}\ln\frac{1.6v^2}{Z} \approx 8\pi\cdot 0.283\frac{Z^2}{v^2}\ln\frac{v^2}{Z}. \tag{21}$$

This expressions differs from the total ionization cross section calculated in Ref. 8 only in the numerical factor under the logarithm sign,[3)]

$$\sigma_i = 8\pi \cdot 0.283 \frac{Z^2}{v^2} \ln \frac{5v^2}{Z}, \tag{22}$$

which gives a good description of the experimental data for $v_0 \ll v \lesssim Z \ll v^2$. The two expressions, (21) and (22), clearly show that for $v_0 \ll v \lesssim Z \ll v^2$ the collisions resulting in the emission of slow electrons contribute the most to the ionization cross section. Note that the cross sections (15), (21), and (22), considered as functions of the charge and velocity of the incident particle, obey scaling, $\sigma/Z = f(v^2/Z)$, which is characteristic of the hydrogen ionization cross section in collisions satisfying $Z/v \gtrsim 1$ and $v \gg v_0$ (see Ref. 1).

So far we assumed that the highly charged ion is structureless. It is obvious, however, that a highly charged ion carrying electrons can be considered a point charge if $b > Z/v \gtrsim 1 \gg r_2$ holds, where $r_2 \sim 1/Z$ is the size of the ion.

3. SINGLE IONIZATION OF HELIUM

As in the above case, we partition the region of impact parameters into two subregions: (1) hard collisions, and (2) of soft collisions. When highly charged fast ions collide with helium atoms and the impact parameters are small, so that large portions of energy are transferred to the atomic electrons, the most probable process is double-electron ionization, while in single-electron ionization the ejected electron has a high energy. Hence the process of single-electron ionization accompanied by the ejection of a slow electron, in this region of impact parameters b ($b < Z/z_{\text{eff}}\, v$, where z_{eff} is the effective charge of the atom core) is even more strongly suppressed than the process of ejection of a slow electron in hard collisions with hydrogen atoms. But when $b > Z/z_{\text{eff}}\, v$ holds, the probability of detachment of each of the electrons rapidly decreases with increasing impact parameter, so that for the probability $P(b)$ of single-electron ionization of helium we can write

$$P(b) \simeq 2w(b), \quad b > \frac{Z}{z_{\text{eff}}\, v}, \tag{23}$$

where $w(b)$ is the detachment probability for each of the electrons. We assume that the function $w(b)$ can be described as the probability of electron detachment from a hydrogenlike ion with a core charge z_{eff}, which we find from the potential of single-electron ionization of helium: $z_{\text{eff}}^2/2 = I = 0.909$, which yields $z_{\text{eff}} = 1.345 \simeq 1.35$. Of course, this approach is quite crude. We note, however, that using it to calculate the cross sections for the ionization of helium by highly charged fast ions yields a fairly good agreement (see, e.g., Refs. 20 and 21) with the experimental data. For this reason, and also because of its simplicity, we employ this method, bearing in mind that the results obtained by this approach should be interpreted as estimates. The results for hydrogen can easily be generalized to the case of single ionization of helium. For instance, if we introduce the Coulomb system of units, based on the charge z_{eff}, the cross sections (15)–(17) multiplied by 2 describe the corresponding cross sections for single-electron ionization of helium. For the angular asymmetry parameter we have

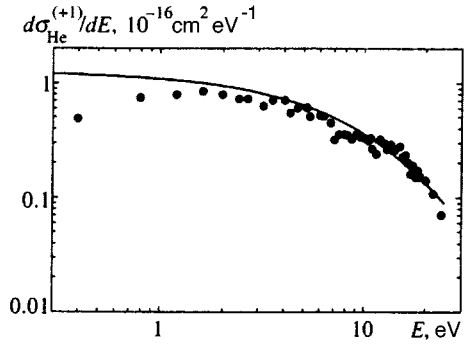


FIG. 1. The differential cross section $d\sigma_{\text{He}}^{(+1)}/dE$ for the ionization of helium by ions with $Z = 24$ and $v = 12$; the dots represent the experimental data of Ref. 5, and the curve represents our estimates.

$$\eta \approx z_{\text{eff}} \left[\frac{1.83}{v} + \frac{Z}{v^2} \left(1.5 \ln \frac{1.6v^2}{Zz_{\text{eff}}} - 2.87 + 2.15 \ln^{-1} \frac{1.6v^2}{Zz_{\text{eff}}} \right) \right]$$
$$\approx z_{\text{eff}} \left[\frac{1.83}{v} + \frac{Z}{v^2} \left(1.5 \ln \frac{v^2}{Z} - 2.87 + 2.15 \ln^{-1} \frac{v^2}{Z} \right) \right] \tag{24}$$

(here we have used atomic units). If we employ (24) to describe the asymmetry in the ionization of helium by ions with $Z = 24$ and $v = 12$ used in the experiment of Moshammer *et al.*,[5] we obtain $\eta \simeq 0.5$, i.e., about 75% of all the slow electrons have a velocity component in the direction of motion of the highly charged fast ion. This value is close to the experimentally determined value of 90% in Ref. 5 (and to the value that follows from classical-trajectory Monte Carlo calculations[5]). Note that in Ref. 5 the classical-trajectory Monte Carlo method was used to study the hypothetical case of single-electron ionization of helium by impact with a highly charged fast anti-ion ($Z = -24$ and $v = 12$). It was found that the asymmetry in the angular distribution of the ejected electron is reversed, i.e., the majority of the ejected electrons move in the direction opposite to that of the anti-ion movement. An estimate of the asymmetry parameters via formula (24) yields $\eta \simeq -0.1$, i.e., Eq. (24) implies that in such collisions (in view of the random selection of the parameters of the problem) the above effects leading to asymmetry balance each other almost perfectly. Thus, in this case a calculation based on classical mechanics and the formula (24) lead to qualitatively different results.

In Fig. 1 we compare the experimentally determined differential (in energy) cross sections and the results of our estimates. In view of the fact that our approach is in no way rigorous, to arrive not only at qualitative agreement but also at a fairly good quantitative agreement we were forced to multiply the calculated cross section $d\sigma_{\text{He}}^{(+1)}/dE$ by a factor $A = 1.4$ (note that such normalization has no effect on the angular distribution of the slow electrons). Since $d\sigma_{\text{He}}^{(+1)}/dE$ rapidly decreases with increasing energy, the value of the normalization constant A was chosen from the condition that the resulting (normalized) expression for the cross section of single-electron ionization accompanied by the ejection of a slow electron,

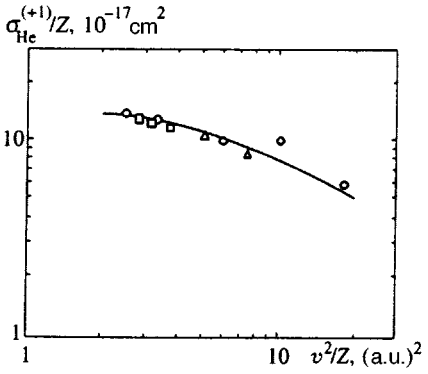


FIG. 2. Single ionization of helium by ions with $8 \leqslant Z \leqslant 54$ and collision energies 1 MeV/amu$\leqslant E_c \leqslant$11.4 MeV/amu: the dots represent the experimental data of Refs. 21–23, and the curve was calculated via (25′).

$$\Delta\sigma_{\mathrm{He}}^{(+1)} \approx A \cdot 16\pi \cdot 0.283 \frac{Z^2}{z_{\mathrm{eff}}^2 v^2} \ln \frac{1.6v^2}{z_{\mathrm{eff}} Z}, \tag{25}$$

reproduce at $Z=24$ and $v=12$ the experimental value of the cross section for single ionization by Au^{24+} ions with a collision energy $E_c=3.6$ MeV/amu (see Refs. 3 and 22), which leads to the following relationship:

$$\Delta\sigma_{\mathrm{He}}^{(+1)} \approx 11.3 \frac{Z^2}{v^2} \ln \frac{1.2v^2}{Z} \tag{25′}$$

$$\approx 11.3 \frac{Z^2}{v^2} \ln \frac{v^2}{Z}. \tag{25″}$$

Note that the experimental value of the cross section $d\sigma_{\mathrm{He}}^{(+1)}/dE$ has a peak at $E \simeq 2$ eV. At the same time, our estimate shows that this cross section monotonically decreases with increasing electron energy and is appreciably larger than the experimental value in the region of very low energies, $E \lesssim 1$ eV. The nature of this discrepancy is unclear (one of the reasons may be that we ignored electron–electron correlations).

In Fig. 2 we compare the results of calculations by formula (25′) and the experimental data of Refs. 3, 21–23 on (total) cross sections of single ionization of helium in a broad range of ion charges and energies ($8 \leqslant Z \leqslant 54$ and 1 MeV/amu$\lesssim E_c \leqslant$11.4 MeV/amu). Finally, in Fig. 3 we give the results of calculations by formula (25′) and the experimental data of Refs. 3 and 22 on (total) cross sections of ionization of helium by ions with Z ranging from 24 to 54 at a fixed collision energy $E_c=3.6$ MeV/amu. If we base our reasoning on the agreement between the results of calculations by formula (25′) and the experimental data corresponding to the parameters $v_0 \ll v \lesssim Z \ll v^2$, we can assume that single ionization occurs for all practical purposes only in soft collisions.

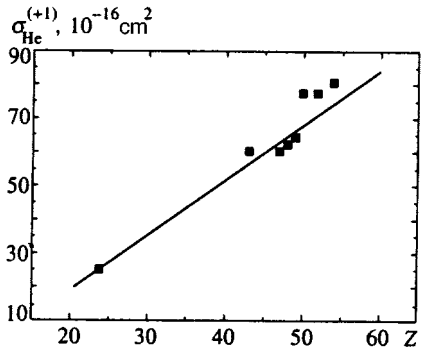


FIG. 3. Single ionization of helium by ions with $24 \leqslant Z \leqslant 54$ at a fixed collision energy $E_c=3.6$ MeV/amu: the dots represent the experimental data of Refs. 3 and 22, and the curve was calculated via (25′).

Formula (25″) yields results whose agreement with the experimental data is slightly worse.

In conclusion, let us touch on the problem of the balance of momenta in the collision of a highly charged fast ion and an atom with $b > Z/v$. In such collisions the average momentum $\mathbf{Q} \simeq Z\mathbf{b}/b^2 v$ transferred from the field of the incident particle to an atomic electron (see, e.g., Ref. 8) is small compared to the characteristic momentum of an electron in the ground state of the atom, $Q_0 \simeq 1$. At the same time, the field of the highly charged fast ion contains characteristic frequencies $\Omega \sim v/b$, which even at $b \simeq v$ are not low compared to the atomic transition frequencies. Hence the ionization of an atom in collisions with $b > Z/v$ resembles[24–26] the ionization of an atom by a light wave, where the atom absorbs a photon whose energy is sufficient for ionization but whose momentum is negligible. In photoionization by a field of moderate frequency, the momentum of the departing electron is balanced by the momentum of the atomic core. Evidently, the same situation occurs in collision ionization with $b > Z/v \gtrsim 1$, which leads to ejection of slow electrons. This, for instance, was detected in the experiments of Moshammer *et al.*,[5,27] who studied ''soft'' collisions with helium atoms.

[1)] In this connection the good agreement between the experimental data and the results obtained by Moshammer *et al.*[5] via the classical-trajectory Monte Carlo method comes as a surprise.

[2)] For electrons moving with velocities $v_0 < v_e < v$ we have $\overline{v} \simeq v_e$ and $F_i/F_a \sim Zv_e^2/v^2 \sim ZE/v^2$. This implies that at electron energies $E \sim z^2/Z$ the effects of the two centers become comparable, while at $E \gg v^2/Z$ the effect of the highly charged ion is predominant.

[3)] Earlier Presnyakov and Uskov[19] found a similar value of the factor under the logarithm sign.

[1] R. K. Janev, L. P. Presnyakov, and V. P. Shevel'ko, *Physics of Highly Charged Ions*, Springer, New York (1985).
[2] C. L. Cocke and R. E. Olson, Phys. Rep. **205**, 205 (1991).
[3] H. Berg, Doctoral Thesis Universität Frankfurt, Frankfurt (1993).
[4] J. Ullrich, R. Dörner, V. Mergel *et al.*, Preprint GSI-94-63 (1994).
[5] R. Moshammer, J. Ullrich, M. Unverzagt, W. Schmidt, P. Jardin, R. E. Olson, R. Mann, R. Dörner, V. Mergel, U. Buck, and H. Schmidt-Böcking, Phys. Rev. Lett. **73**, 3371 (1994).
[6] C. Reinhold and J. Burgdorfer, J. Phys. B **26**, 3101 (1993).
[7] P. D. Fainstein, V. H. Ponce, and R. D. Rivarola, J. Phys. B **24**, 3091 (1991).
[8] A. B. Voĭtkiv and A. V. Koval', Zh. Tekh. Fiz. **64**, No. 3, 188 (1994) [Tech. Phys. **39**, 335 (1994)].
[9] J. J. Macek, in *Ionization of Solids by Heavy Particles*, R. A. Baragiola (ed.), Plenum Press, New York (1988), Vol. 309, p. 1.
[10] N. Stolterfoht, D. Schneider, J. Tanis, *et al.* Europhys. Lett. **4**, 899 (1987); N. Stolterfoht, H. Platten, G. Schiwietz, D. Schneider, L. Gulyás, P. D. Fainstein, and A. Salin, Phys. Rev. A **52**, 3796 (1995).
[11] E. L. Duman, L. I. Men'shikov, B. M. Smirnov, Zh. Éksp. Teor. Fiz. **76**, 516 (1979) [Sov. Phys. JETP **49**, 260 (1979)].
[12] C. O. Reinhold, C. A. Falcon, and J. E. Miraglia, J. Phys. B **20**, 3737 (1987).
[13] V. S. Nikolaev, V. A. Sidorovich, and V. N. Novozhilova, Zh. Éksp. Teor. Fiz. **101**, 1198 (1992) [Sov. Phys. JETP **74**, 640 (1992)].

[14] L. P. Rapoport, B. A. Zon, and N. L. Manakov, *The Theory of Multiphoton Processes in Atoms* [in Russian], Atomizdat, Moscow (1978), p. 135.
[15] S. Basile, F. Trombetta, and G. Ferrante, Kvant. Élektron. (Moscow) **15**, 1149 (1988) [Sov. J. Quantum Electron. **18**, 733 (1988)].
[16] N. Bohr, ''The penetration of atomic particles through matter,'' Kgl. Danske Videnskab. Selskab. Mat.-Fys. Medd. **18**, No. 8 (1948).
[17] A. M. Dykhne and D. L. Yudin, Usp. Fiz. Nauk **125**, 377 (1978) [Sov. Phys. Usp. **21**, 549 (1978)].
[18] *Handbook of Mathematical Functions*, edited by M. Abramowitz and I. A. Stegun, National Bureau of Standards Applied Mathematics Series 55, Washington, D.C. (1964).
[19] L. P. Presnyakov and D. B. Uskov, Zh. Éksp. Teor. Fiz. **86**, 882 (1984) [Sov. Phys. JETP **59**, 515 (1984)].
[20] V. S. Nikolaev and V. A. Sidorovich, Nucl. Instrum. Methods Phys. Res. B **36**, 239 (1989).
[21] J. H. McGuire, A. Müller, B. Schuch, W. Groh, and E. Salzborn, Phys. Rev. A **35**, 2479 (1987).
[22] H. Berg, J. Ullrich, E. Bernstein *et al.*, J. Phys. B **25**, 4075 (1992).
[23] H. Knudsen, L. H. Andersen, P. Hvelplund *et al.* J. Phys. B **17**, 3545 (1984).
[24] E. Fermi, Z. Phys. **29**, 315 (1924).
[25] C. Weizsäcker, Z. Phys. **88**, 612 (1934).
[26] E. Williams, Phys. Rev. **45**, 729 (1934).
[27] R. Moshammer, J. Ullrich, M. Unverzagt *et al.*, Nucl. Instrum. Methods Phys. Res. B **107**, 62 (1996).

Translated by Eugene Yankovsky

Coherence effects accompanying generation of high-order harmonics

D. F. Zaretskiĭ and É. A. Nersesov

Moscow State Engineering Physics Institute, 115409 Moscow, Russia
(Submitted 31 January 1997)
Zh. Éksp. Teor. Fiz. **112**, 1978–1984 (December 1997)

For phase-locked emitters, provided a certain relation exists between the pump wave and atomic beam parameters, a saturation effect is shown to be possible for which the intensity of the high-order harmonics ceases to depend on the atomic density. By means of a simple model that includes variations in the intensity of the pump wave in the plane transverse to the focal axis, an expression is obtained for the optimum atomic density of the medium corresponding to intensity saturation. The dependence of the optimum atomic density on the laser power and harmonic number obtained is found to be in qualitative agreement with recently published experimental data.[8] [S1063-7761(97)00512-X]

1. INTRODUCTION

The generation of high-order harmonics at frequencies corresponding to an odd number of photons of the ionizing laser wave was first examined in the experimental studies Refs. 1–4. One peculiarity of this phenomenon is that the intensity of the harmonic I_s (s is the order of the harmonic) depends on the density n_a of atoms of the medium, established in Ref. 3. In a series of theoretical papers this result was linked with the phenomenon of phase locking during harmonic generation. Thus, the authors of Ref. 5, by numerically solving Maxwell's equations in a nonlinear medium, obtained values of the intensities I_s as functions of the basic parameters of the ionizing wave and the atomic beam.

In Refs. 6 and 7 we examined the generation of high-order harmonics with the aid of the analytical quantum approach which we developed to describe the effects of above-threshold atomic ionization. We established a direct link between harmonic generation and the phenomenon of above-threshold atom ionization in which the high-order harmonics owe their existence to transitions of the atom from the ground state to the continuum with absorption of several pump photons and emission of a photon of the high-order harmonic upon return to the ground state. The expressions derived in Refs. 6 and 7 for the probabilities of spontaneous and induced emission of high-order harmonics under conditions of phase-locking of the emitters provide a satisfactory explanation of the main features of the phenomenon (shape of spectrum, dependence of I_s on the various parameters of the problem such as intensity and size of the focus of the laser wave, density of the atomic beam and its location relative to the focus, etc.).

A recent joint experimental effort of the Lund and Saclay groups,[8] examined the dependence of the intensity of high-order harmonics generation on the density of the atomic medium. They found that in the low-pressure interval (4–14 mbar) the intensity of the harmonics grows approximately quadratically as a function of the atomic density. However, as some optimum concentration is reached the intensity of generation rolls over to a constant level and further growth of the density of the medium is accompanied by a monotonic decrease of the intensity. It was also found that the optimum concentration depends both on the order of the harmonic (it falls with increasing order) and on the intensity of the laser wave (for fixed harmonic order it increases with the pump power).

It would be of interest to seek a qualitative explanation for these experimental results within the framework of the analytical quantum approach. In the present paper we show that under conditions of phase-locking of the emitting atoms for low-to-moderate medium density the dependence of the intensity of the harmonics I_s on the atomic density n_a is quadratic. However, for certain ratios between the pump and beam parameters a saturation effect arises, whereby the intensity I_s ceases to depend on the concentration. The reason for this phenomenon has to do with the substantial phase shift of the emitters arising at distances less than the longitudinal dimension of the interaction volume.

Using a simple model which treats variations in the intensity of the laser wave in the plane transverse to the focal axis, we obtain a relation for the optimum atomic density. The qualitative dependences following from this relation are found to agree well with the data of Ref. 8.

2. PROBABILITY OF SPONTANEOUS EMISSION

We consider the time-dependent problem of transitions from the ground state of the atoms with absorption of several pump photons and the return to the ground state with emission of a high-order-harmonic photon. The probability amplitude for the system to undergo a transition at the time t with emission of photons of the sth harmonic ($\mathbf{K},\Omega=s\omega$) is given by the following expression from Ref. 7 ($\hbar=c=1$):

$$A_\Omega(t)=A_{0s}(eA_{0\Omega})\sum_j \exp[i(s\mathbf{k}-\mathbf{K})\cdot\mathbf{R}_j]\xi^*(s\omega-\Omega)\times\exp[i(s\omega-\Omega-i\lambda)t], \quad (1)$$

where $A_{0\Omega}$ is the amplitude of the vector potential of the emitted wave of frequency Ω; $\mathbf{k}$ and ω are the wave vector and frequency of the ionizing wave; $\mathbf{R}_j$ is the radius vector of the jth atom (the residual ion) and the sum is over all atoms in the interaction volume of the medium with the laser wave; $\xi^*(x)=\mathscr{P}/x+i\pi\delta(x)$; the dimensionless factor A_{0s} is determined by the probability amplitude of multiphoton ionization

of the atom and contains a function describing the envelope of the maxima of the spectrum of above-threshold photoelectrons; the parameter $\lambda \approx +0$ corresponds to adiabatic switching-on of the wave field at $t \to -\infty$.

Equation (1) yields the following formula for the probability per unit time of the transition to the partial final state of the system:

$$\frac{d}{dt}|A_\Omega(t)|^2 = A_{0s}^2 (eA_{0\Omega})^2 \left| \sum_j \exp[i(s\mathbf{k}-\mathbf{K})\cdot\mathbf{R}_j] \right|^2 \times \frac{2\lambda}{(s\omega-\Omega)^2+\lambda^2}. \quad (2)$$

In the case of spontaneous emission of the sth harmonic of frequency ω it is necessary to replace $eA_{0\Omega}$ by the expression $\sqrt{8\pi\alpha/\Omega V}$ in Eq. (2), where V is the normalization volume of the spontaneous emission field, α is the fine-structure constant, and we replace the Lorentzian dependence (the last factor in expression (2)) by a δ-function ensuring conservation of energy in the process. This δ-function is removed by integration over the statistical weight of the emitted photon $\mathbf{K},\Omega$. As a result, the transition probability of the system of atoms from the ground state per unit time is given by

$$w_{sp}^{(s)} = A_{0s}^2 \frac{4\alpha s}{\lambda_0} \int_{(4\pi)} \left| \sum_j \exp[i(s\mathbf{k}-\mathbf{K})\cdot\mathbf{R}_j] \right|^2 d\Omega_K, \quad (3)$$

Here the integration is over directions of emission of the photon $\mathbf{K}$ and $\lambda_0 = 2\pi/\omega$ is the wavelength of the laser radiation.

We will only calculate the probability per unit time of emission of a high-order-harmonic photon. Here for the total transition probability per pump pulse we have the order-of-magnitude estimate $w_{sp}^{(s)}\tau_i$, where τ_i is the duration of the ionizing pulse.

As follows from Eq. (3), the probability $w_{sp}^{(s)}$ can be represented as a product of two factors: the quantity A_{0s}^2, which describes the transition taking place on a single atom and is the same for all atoms, and the sum over j corresponding to the collective response of the medium to the pump waves. The probability A_{0s}^2 pertains to the transition of an atom from the ground state with absorption of several pump photons followed by return to the ground state with emission of one photon of a high-order harmonic. In the case of multiphoton transitions this probability is determined by the component matrix element and involves a multiple sum over intermediate states of the atom, a sum over virtual quasienergy states of the photoelectron, and integration over the continuum.[7] The condition for phase-locking of the atoms radiating during this process is connected with the behavior of the arguments of the summed exponentials in expression (3).

The ensuing calculations are easily carried out in the continuous medium approximation (criterion formulated below), where the sum in expression (3) is replaced by an integral over the volume in which the atoms interact with the pump wave. In this case, Eq. (3) yields

$$w_{sp}^{(s)} = A_{0s}^2 \left(\frac{\pi}{3}\right)^2 \frac{\alpha s}{\lambda_0} (V_{\rm int} n_a)^2 \times \int_{(4\pi)} \left[\frac{2J_1(v)}{v}\right]^2 \frac{\sin^2 u}{u^2} d\Omega_{\mathbf{K}}, \quad (4)$$

where $V_{\rm int} = \pi\rho_0^2 d$ is the volume within which the atomic beam interacts with the ionizing wave, ρ_0 is the radius of the focus of the wave at its center, d is the diameter of the beam of atoms projected transverse to the pump wave, $J_1(v)$ is the Bessel function, the arguments of the diffraction factors are

$$v = s\omega\rho_0(\theta-\theta') \quad \text{and} \quad u = s\omega[\theta^2-(\theta_0^2+\theta'^2)]d/4, \quad (5)$$

where $\theta_0^2 \equiv 2|\Delta n|$ ($\Delta n = n_\omega - n_\Omega$ is the difference in the refractive indices of the medium for waves of the corresponding frequencies), and θ and θ' are the angles between the z axis (the direction of propagation of the ionizing wave) and the vectors $\mathbf{K}$ and $\mathbf{k}$, respectively. The angle θ' is related to the spread in the directions of the wave vector $\mathbf{k}$ resulting from focusing of the laser wave. In the expressions obtained below, it is necessary to average over this angle.

As the numerical calculations of L'Huillier *et al.*[5] have shown, under the conditions of the experiments on high-order harmonic generation using a strong laser wave, where photoionization of the atoms reaches saturation, the main contribution to Δn comes from the photoelectrons and $|\Delta n| = \omega_p^2/2\omega^2$, where $\omega_p = \sqrt{4\pi n_i e^2/m_e}$ is the plasma frequency of the ionized medium ($n_i \approx n_a$ under conditions of saturation).

The continuous medium approximation is valid if the inequality $|s\mathbf{k}-\mathbf{K}|a \ll 1$ is satisfied or, allowing for the parameter Δn,

$$\frac{|\Delta n|a}{\lambda_0/s} = \frac{\theta_0^2 as}{2\lambda_0} \ll 1, \quad (6)$$

where a is the mean distance between the atoms of the medium. It is not hard to see that for a given density of the medium this condition imposes an upper bound on the maximum achievable harmonic orders s.

According to Eqs. (4) and (5) the angular diagram of the intensity of emission of a harmonic consists of two sharp diffraction maxima in the directions $\theta = \theta'$ and $\theta = \sqrt{\theta_0^2+\theta'^2}$, and the resulting intensity depends on both the angular detuning θ_0 and the angular widths $\Delta\theta_\perp$ and $\Delta\theta_\parallel$ of the corresponding diffraction factors.

The estimates in Ref. 7 showed that for parameters typical of the experiments on high-order harmonics generation the total intensity of emission in the direction $\theta = \theta'$ within the limits of the angular width $\Delta\theta_\perp \simeq 2\lambda_0/s\pi\rho_0$ significantly exceeds the analogous quantity in the direction $\theta = \sqrt{\theta_0^2+\theta'^2}$ within the limits of the angular width $\Delta\theta_\parallel \simeq \lambda_0/s\theta_0 d$. For this reason the main contribution to the integral in expression (4) comes from the region of angles θ for which $v \approx 0$. In the case of hard focusing of the laser wave, when the angular divergence at the focus θ_f exceeds the diffraction width $\Delta\theta_\perp$ ($\theta_f > \Delta\theta_\perp$), the integral in expression (4) yields

$$w_{sp}^{(s)}=A_{0s}^2\left(\frac{\pi}{3}\right)^2\frac{\alpha s}{\lambda_0}(V_{\rm int}n_a)^2 2\pi\theta'\Delta\theta_\perp\frac{\sin^2(s\omega\theta_0^2 d/4)}{(s\omega\theta_0^2 d/4)^2}. \tag{7}$$

Assuming for simplicity that the angular density of the intensity of the laser wave is constant within the limits of the focus, after averaging over θ' in expression (7) we obtain

$$w_{sp}^{(s)}=A_{0s}^2\left(\frac{\pi}{3}\right)^2\frac{\alpha s}{\lambda_0}(V_{\rm int}n_a)^2\pi\Delta\theta_\perp\theta_f\frac{\sin^2(s\omega\theta_0^2 d/4)}{(s\omega\theta_0^2 d/4)} \tag{8}$$

for $\theta_f>\Delta\theta_\perp$.

In the soft-focusing limit $\theta_f<\Delta\theta_\perp$, there is no need to average over θ', and the emission probability is given by a formula analogous to (8), where the factor $\pi\Delta\theta_\perp\theta_f$ is replaced by the solid angle $\pi(\Delta\theta_\perp)^2$ of the diffraction spot in the direction of the pump wave.

Wrapping up the discussion of the question of how focusing of the wave affects the intensity of spontaneous emission of harmonics, we derive an expression for the angular divergence θ_f. In the diffraction near zone, where the transverse beam size d is significantly less than the confocal parameter L (this is just this case that is realized in the well-known experiment of Li *et al.*[3]), θ_f is given by

$$\theta_f\simeq\frac{d\lambda_0^2}{(2\pi)^2\rho_0^3}. \tag{9}$$

For the main laser and beam parameters given in the Conclusion it follows that for $\rho_0\simeq10^{-3}$ cm hard focusing takes place, and for $\rho_0\simeq10^{-1}$ cm, soft focusing takes place.

The total intensity of emission of the sth harmonic within the diffraction spot in the direction $\theta=0$ is given by

$$I_s=w_{sp}^{(s)}s\omega/\pi\rho_0^2=A_{0s}^2\frac{(2\pi)^3}{9}\alpha(dn_a)^2\frac{\sin^2(s\omega\theta_0^2 d/4)}{(s\omega\theta_0^2 d/4)^2}. \tag{10}$$

(we restrict the discussion here to the case of large ρ_0, for which $\theta_f<\Delta\theta_\perp$). If the main parameters of the problem are such that the condition $u_0=s\omega\theta_0^2 d/4<1$ holds, the intensity I_s depends quadratically on n_a, and is equal to

$$I_s=A_{0s}^2\frac{(2\pi)^3}{9}\alpha(dn_a)^2. \tag{11}$$

This result obtains under conditions such that all the atoms found inside the volume $V_{\rm int}$ within which the atoms interact with the wave are phase-locked. It is completely clear that expression (11) is applicable for this reason only in sufficiently disperse media. Thus, if we use the parameters of Ref. 3 for our estimates ($\lambda_0=1064$ nm, $d=1$ mm), we obtain the following upper bound on the atomic density:

$$n_a<2.1\times10^{18}/s\ \ {\rm cm}^{-3}. \tag{12}$$

For harmonics with orders $s\simeq30-50$ condition (12) leads to values of $n_a\approx10^{16}$ cm^{-3}.

The intensity (11) groups with n_a for fixed values of the wave parameters until the argument of the diffraction factor $\sin^2 u_0/u_0^2$ in expression (10) becomes significant, $u_0\gg1$. In this case the intensity I_s ceases to depend on n_a and its value is given by

$$I_s=A_{0s}^2\frac{2(2\pi)^3}{9}\frac{m_e^2}{s^2\alpha\lambda_0^2}. \tag{13}$$

Saturation occurs because when the medium is dense, the distance in the pump-wave direction over which a significant phase shift of the coherently radiating atoms accumulates,

$$L_{\rm coh}=\frac{2\lambda_0}{s\theta_0^2}=\frac{2\pi m_e}{s\alpha\lambda_0 n_a}, \tag{14}$$

is less than the longitudinal dimension d region in which the beam interacts with the wave. This effectively replaces the parameter d with the coherence length $L_{\rm coh}$ in expression (10).

The atomic density at which saturation is reached is found from the condition $L_{\rm coh}\simeq d$ and is given by

$$n_a\approx2\times10^{13}/s\lambda_0 d \tag{15}$$

(here λ_0 and d are measured in centimeters, and n_a in cm^{-3}).

The possibility of such a saturation effect under phase-locking conditions for a certain set of values of the parameters of the problem was noted in Ref. 9, where this conclusion was reached through numerical analysis of the results of the experiment of Ref. 3.

In fact, as follows from the results of a recent experimental study,[8] the optimum value of the atomic density at which saturation is reached exceeds the estimate (15) by more than an order of magnitude. This difference may be due to inhomogeneous ionization of the atoms within the focus of the pump wave. This inhomogeneity is a consequence of the laser wave intensity distribution in the plane transverse to the axis of the focal region. To understand this effect qualitatively and estimate it, consider the following model. We divide the focal region into two segments in the transverse direction. We assume that in the inner region abutting the focal axis total ionization of the atoms takes place within the transit time of the pump-wave pulse. On the periphery of the focus the ionization has a partial character. Within this model the diffraction factor $\sin^2 u_0/u_0^2$ in expression (10) changes and the expression for the intensity takes the form

$$I_s\propto\frac{\sin^2(s\omega\theta_0^2 d/4)}{(s\omega\theta_0^2 d/4)^2}\left(\frac{\Delta\rho}{\rho_0}\right)^2+\frac{\sin^2(s\omega\theta'^2_0 d/4)}{(s\omega\theta'^2_0 d/4)^2}\times\left[1-\left(\frac{\Delta\rho}{\rho_0}\right)^2\right]\tilde{\alpha}, \tag{16}$$

where $\Delta\rho$ is the effective radius of the inner region of the focus, $\tilde{\alpha}=|A'_{0s}/A_{0s}|^2$ and θ'_0 and θ_0 are parameters that depend on the degree of ionization of the medium in the outer and inner regions of the focus, respectively.

The transition from a quadratic dependence of the harmonic intensity to saturation as n_a increases takes place when the two terms in expression (16) become of the same order of magnitude, at which point saturation is reached in the inner region ($s\omega\theta_0^2 d/4>1$). Equation (16) then yields

the condition for the density at which the dependence of the harmonic intensity begins to deviate from quadratic:

$$n_a \approx \frac{(\Delta\rho/\rho_0)^2}{[1-(\Delta\rho/\rho_0)^2]\tilde{\alpha}} \frac{m_e\omega}{se^2d}. \quad (17)$$

The value of n_a following from expression (17) is much greater than the estimate (15) if $\Delta\rho/\rho_0 \approx 1$ holds, i.e., almost complete ionization of the atoms over the entire focal region occurs and the ratio of amplitudes is $\tilde{\alpha} < 1$. The experimentally observed dependence[8] of the optimum atomic density for a given harmonic on the pump-wave power (in the range 0.6×10^{15} to 1.5×10^{15} W/cm^2) also follows from condition (17), as well as the harmonic order ($s \simeq 50-70$) for fixed pump-wave intensity.

As follows from condition (17), the optimum density should decrease as a function of the order s, as is experimentally observed. In addition, increasing the pump-wave power for fixed s causes the parameter $\Delta\rho/\rho_0$ to approach unity and, consequently, leads to a growth of n_a, which is also observed in experiment.[8]

3. ESTIMATES; CONCLUSION

In this section we give the pump-wave and atomic-beam parameters, and also numerical estimates of the main quantities for which expressions were obtained in the previous section. In the derivation of the beam and pump parameter values we were guided by two main considerations. The effects of above-threshold ionization of the atoms and generation of high-order harmonics were treated in the multiphoton approximation, in which the Keldysh adiabaticity parameter satisfies $\gamma \geqslant 1$. This condition imposes an upper bound on the intensity of the pump wave in the interaction volume and on its degree of focusing. However, we note that the phase-locking conditions for the emitters implemented in this work do not depend on the mechanism of ionization of an individual atom and are therefore equally applicable in the case of ionization in the tunneling regime, when $\gamma < 1$ holds. As for the properties of the medium, the beam parameters (density, degree of collimation, cross-sectional diameter) are chosen from the optimal conditions for observation of harmonic generation.

We will confine ourselves to a numerical example corresponding to the multiphoton approximation. To check the validity of the formulas obtained in this work, let us turn to the experimental results of Li *et al.*[3] who observed generation of high-order harmonics for hard focusing ($\rho_0 = 18$ μm), a dense medium ($n_a = 5\times10^{17}$ cm^{-3}), high pump-wave intensity at the focus ($I = 3\times10^{13}$ W/cm^2), and comparatively long pulses ($\tau_p = 36$ ps).

Calculating according to our formulas shows that in the case of argon atoms the number of photons with harmonic order $s = 33$ emitted per pulse of the pump wave reaches a value of roughly 10^5 (when the parameter A_{0s} calculated according to the formulas in Ref. 7 is approximately equal to 1.4×10^{-7}). Completely satisfactory agreement is noted between the derived value of N_s and the experimental value.

The results obtained here are based on the supposition that the amplitude A_{0s} [see Eq. (1)] is independent of the coordinates of the atom. This assumption, strictly speaking, is valid only in the case of a homogeneous pump-wave field. In fact, the wave field at the focus is inhomogeneous, which can give rise to an additional phase in the sum in expression (1). However, the characteristic length over which this phase varies significantly is of the order of ρ_0. Since the characteristic width of the transverse diffraction factor is $\Delta\theta_\perp \sim \lambda/s\rho_0 \ll 1$, such a ''sluggish'' dependence of the additional phase does not alter the results.

In conclusion, we may briefly formulate the main results obtained here.

1) We have shown that under the conditions for phase locking of the radiating atoms the dependence of the intensity of harmonic generation on the atomic density of the medium can be different in the limiting cases of sufficiently disperse and dense media (from a quadratic dependence to a constant value);

2) Using a simple model we have allowed for the effect of variation in the pump-wave intensity in the direction transverse to the focal axis. Using this model we have formulated a relation for the optimum atomic density of the medium corresponding to the above-indicated transition. The results obtained are found to be in qualitative agreement with recently published data.[8]

ACKNOWLEDGMENTS

The authors express their gratitude to N. B. Delone for his interest in this work and discussion of its results, and also P. Agostini for making data available from the Lund–Saclay joint experimental study.

[1] A. McPherson, G. Gibson, H. Jara *et al.*, J. Opt. Soc. Am. B **4**, 595 (1987).
[2] M. Ferray, A. L'Huillier, L. A. Lompré *et al.*, J. Phys. B **21**, L31 (1988).
[3] X. Li, A. L'Huillier, M. Ferray *et al.*, Phys. Rev. A **39**, 5751 (1989).
[4] A. L'Huillier, L. A. Lompré, G. Mainfray, and C. Manus, in *Proceedings of the Fifth International Conference on Mulitphoton Processes*, Paris (1990), p. 45.
[5] A. L'Huillier, P. Balcou, S. Candel *et al.*, Phys. Rev. A **46**, 2778 (1992).
[6] E. A. Nersesov and D. F. Zaretsky, Laser Phys. **3**, 1105 (1993).
[7] D. F. Zaretskiĭ and É. A. Nersesov, Zh. Éksp. Teor. Fiz. **109**, 1994 (1996) [JETP **82**, 1073 (1996)].
[8] C. Altucci, T. Starczewski, E. Mevel *et al.*, J. Opt. Soc. Am. B **13**, 148 (1996).
[9] S. C. Rae, K. Burnett, and J. Cooper, Phys. Rev. A **50**, 3438 (1994).

Translated by Paul F. Schippnick

Evolution of polarization in an inhomogeneous isotropic medium

N. I. Petrov

All-Russian Electrical Engineering Institute, 143500 Istra, Moscow Region, Russia
(Submitted 5 March 1997)
Zh. Éksp. Teor. Fiz. **112**, 1985–2000 (December 1997)

The depolarization and rotation of the polarization plane of radiation propagating in a two-dimensional graded-index medium is investigated on the basis of the quantum-mechanical method of coherent states. It is shown that the degree of polarization of both linearly and circularly polarized radiation decreases with increasing distance as a result of interaction between the polarization (spin) and the path (orbital angular momentum) of the beam. The wave nature of the depolarization is emphasized. The depolarization decreases as the radiation wavelength decreases. It is found that the degree of polarization exhibits oscillations of pure diffraction origin during the propagation of light in a single-mode optical fiber. It is shown that the rotation of the polarization plane is nonuniform in character and depends on the offset and the tilt angle of the incident-beam axis relative to the fiber axis. The Berry phase is found to undergo oscillations of a wave nature during the propagation of radiation in an inhomogeneous medium. It is shown that the spread in the angle of rotation of the polarization plane increases with distance and can be determined from measurements of the degree of polarization of the radiation. © *1997 American Institute of Physics.* [S1063-7761(97)00612-4]

1. INTRODUCTION

The polarization properties of fields must be taken into account in many problems involving wave propagation in inhomogeneous media. The polarization of radiation is known to remain constant during propagation in a homogeneous, isotropic, nondispersive medium. Propagation in an inhomogeneous medium, on the other hand, is accompanied by significant changes in the state and degree of polarization. The depolarization of radiation in a randomly inhomogeneous medium has been investigated previously.[1,2] It was shown that two depolarization mechanisms occur: a geometrical mechanism associated with Rytov rotation[3] of the polarization plane and a diffraction mechanism. McLean and Pendry[4] investigate polarization effects in the propagation of electromagnetic waves in two-dimensional random media on the basis of generalized transformation matrices for the description of multiple scattering by inhomogeneities. In optical fibers, as rule, the input polarization changes by virtue of birefringence in the medium.[5–8] However, depolarization also takes place in optical fibers without birefringence. In particular, it has been shown[9] that the degree of linear polarization of radiation in an isotropic parabolic-index fiber decreases with increasing distance. It has also been shown[10] that the degree of linear polarization in a multimode parabolic-index fiber decreases with distance according to an inverse square law as a result of Rytov rotation of the polarization vector, while the degree of circular polarization remains constant. In experimental work,[11] however, preservation of the degree of polarization has not been observed in an isotropic optical fiber.

We know that the polarization vector of a light ray rotates as it propagates along a twisted path in an inhomogeneous medium.[3] This rotation has been observed experimentally in a single-mode fiber wound along a cylinder[12] and has been interpreted as a Berry geometrical phase effect.[13] Rotation of the polarization plane has also been observed in a straight multimode step-index fiber.[14]

In this paper we investigate the evolution of polarization in an isotropic multimode graded-index optical fiber. We show that the depolarization of radiation in an isotropic graded-index medium is the result of diffraction effects, this mechanism prevailing for both linearly polarized and circularly polarized radiation. We analyze the rotation of the polarization vector during propagation in a two-dimensional medium. We show that the rotation of the polarization plane depends on the offset of the axis and the tilt angle of the incident beam relative to the axis of the fiber.

To investigate the evolution of the parameters of the radiation beam, we use quantum-mechanical coherent-state methods, which enable us to calculate averages by the operator approach. The entire system dynamics is transferred to operators in this case. The evolution of the beam parameters can then be investigated by purely algebraic procedures i.e., without having to rely on explicit expressions for the wave functions or having to compute the corresponding integrals.

The choice of the quantum-mechanical formalism in waveguide theory rests on the following considerations. It has been shown[15] that Maxwell's equations for scalar wave beams in the paraxial approximation are very accurately reducible to a parabolic-type equation. This approximation enables us to use well-developed quantum-mechanical methods for the investigation of wave propagation in inhomogeneous media, because the parabolic equation formally coincides with the Schrödinger equation in quantum mechanics for particles moving in a time-dependent potential well. All that is required is to redefine the parameters in the Schrödinger equation. The role of time is now taken by the longitudinal coordinate, and Planck's constant is superseded by the radiation wavelength in vacuum. The potential is defined as a function of the refractive index of the medium.

The intimate relationship between the wave mechanics

of particles and the optics of light beams has been discussed in detail in many papers (see, e.g., Refs. 16–18). Methods developed recently for the solution of time-dependent quantum-mechanical problems,[21] such as the method of coherent states and the density matrix formalism, have been used to investigate the propagation of coherent and partially coherent light in slightly inhomogeneous media without polarization effects.[19,20] The coherent-state method has been used to analyze polarization effects attending light propagation in a graded-index fiber.[22] Completely polarized incident radiation in an inhomogeneous isotropic medium was found to undergo depolarization of diffraction origin.

In principle, the effects discussed below have been obtained on the basis of the classical approach, since the investigated equations are classical and do not contain Planck's constant. However, the results can also be used in the study of quantum systems described by similar Hamiltonians.

2. STATEMENT OF THE PROBLEM

The equation describing the propagation of radiation in an inhomogeneous medium can be obtained from Maxwell's equation for the electric field $\mathbf{E}\exp(-i\nu t)$:

$$\operatorname{curl}(\operatorname{curl}\mathbf{E})=k^2\varepsilon\mathbf{E} \tag{1}$$

or

$$\Delta\mathbf{E}+k^2n^2\mathbf{E}+\nabla(\mathbf{E}\cdot\nabla\ \ln n^2)=0, \tag{2}$$

where $k=2\pi/\lambda$ is the wave number, and $\varepsilon=n^2$ is the dielectric constant of the medium.

We assume below that nonlinear effects and absorption do not occur in the medium.

It is evident from Eq. (2) that the polarization term is small. At relatively large distances, however, small corrections can produce appreciable cumulative effects.

We consider a two-dimensional graded-index optical fiber with a parabolic index profile

$$n^2(x,y)=n_0^2-\omega^2(x^2+y^2), \tag{3}$$

where ω is the gradient parameter, n_0 is the refractive index on the fiber axis, and x,y denote the transverse coordinates in the fiber.

We adopt coherent states, i.e., Gaussian wave packets representing eigenfunctions of the annihilation operators $\hat{a}_1$ and $\hat{a}_2$ (Ref. 23), to serve as functions specifying the spatial distribution of the field in the initial plane:

$$\hat{a}_{1,2}|\alpha_1\alpha_2\rangle=\alpha_{1,2}|\alpha_1\alpha_2\rangle, \tag{4}$$

$$\hat{a}_1=\frac{1}{\sqrt{2}}\left(\sqrt{k\omega}\ \hat{x}+i\sqrt{\frac{k}{\omega}}\ \hat{p}_x\right),$$

$$\hat{a}_2=\frac{1}{\sqrt{2}}\left(\sqrt{k\omega}\ \hat{y}+i\sqrt{\frac{k}{\omega}}\ \hat{p}_y\right),$$

$$\hat{p}_x=-\frac{i}{k}\frac{\partial}{\partial x},\quad \hat{p}_y=-\frac{i}{k}\frac{\partial}{\partial y}.$$

The term *coherent states* was first introduced by Glauber[24] in 1963 as part of a study of oscillator states of an electromagnetic field and its statistical properties. The important consideration here is that these states are analogous to the Gaussian wave packets in coordinate representation formulated and investigated by Schrödinger[25] to establish a relationship between the classical and quantum approaches.

The coherent states $|\alpha_1\alpha_2\rangle$ characterize the spatial distribution of the electric field and have the form

$$|\alpha_1\alpha_2\rangle=\sqrt{\frac{k\omega}{\pi}}\exp\left\{-\frac{k\omega}{2}(x^2+y^2)+\sqrt{2k\omega}(\alpha_1x+\alpha_2y)\right.$$
$$\left.-\frac{1}{2}(\alpha_1^2+\alpha_2^2+|\alpha_1|^2+|\alpha_2|^2)\right\}. \tag{5}$$

The eigenvalues

$$\alpha_1=\frac{1}{\sqrt{2}}\left(\sqrt{k\omega}\ x_0+i\sqrt{\frac{k}{\omega}}p_{x0}\right),$$

$$\alpha_2=\frac{1}{\sqrt{2}}\left(\sqrt{k\omega}\ y_0+i\sqrt{\frac{k}{\omega}}p_{y0}\right)$$

determine the initial coordinates x_0 and y_0 (offset) and the angles p_{x0} and p_{y0} between the ray path and the fiber axis (tilt angles).

For an axial beam ($\alpha_{1,2}=0$) the expression (5) takes the form

$$|00\rangle=\sqrt{\frac{k\omega}{\pi}}\exp\left(-\frac{k\omega}{2}x^2-\frac{k\omega}{2}y^2\right).$$

This is usually the form of the fundamental mode of laser sources. Such a beam also corresponds to the fundamental mode of an optical fiber or the mode of a single-mode fiber. Expressions with $\alpha_{1,2}\neq0$ describe beams having a nonzero offset or a nonzero tilt angle relative to the fiber axis. For example, the equation for a beam with an incident-beam offset x_0 has the form

$$|\alpha_10\rangle=\sqrt{\frac{k\omega}{\pi}}\exp\left(-\frac{k\omega}{2}(x-x_0)^2-\frac{k\omega}{2}y^2\right).$$

The coherent states (5) are localized wave packets, which have minimum width and angular diffraction spreading during propagation in a medium with a quadratic index profile. The centroid of such wave packets moves along the geometrical ray path, i.e., obeys geometrical optics. Moreover, the coherent states (5) form a complete system of functions and are the fiber-mode generating functions. This property can be utilized to expand an arbitrary field $|f\rangle$ in coherent states:

$$|f\rangle=\pi^{-1}\int\int|\alpha\rangle\langle\alpha|f\rangle\ d^2\alpha. \tag{6}$$

In a slightly inhomogeneous medium ($\delta n/n\ll1$ at distances of the order of λ) Maxwell's equations in the paraxial approximation can be reduced to an equivalent Schrödinger equation. A detailed derivation of this equation for a scalar wave field is given in Ref. 17. A similar approach can be used to derive a parabolic equation for a two-component vector wave function.[22] In the paraxial approximation, radiation propagates at small angles relative to the z axis, so that

$$\partial^2 E/\partial z^2 \ll k\partial E/\partial z \approx \partial^2 E/\partial x^2 \approx \partial^2 E/\partial y^2 .$$

Disregarding the term $\partial^2 E/\partial z^2$ in Eq. (2), we obtain the following equation for the two-component wave function:

$$\frac{i}{k}\frac{\partial \psi}{\partial z} = \hat{H}\psi, \tag{7}$$

where

$$\psi = \sqrt{n_0}\,\exp(-ikn_0 z)\begin{pmatrix} E_x \\ E_y \end{pmatrix}, \quad \hat{H} = \hat{H}_0 + \hat{H}_1,$$

$$\hat{H}_0 = \left[-\frac{1}{2k^2 n_0}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) + \frac{1}{2n_0}\,(n_0^2 - n^2) \right] \hat{I}$$

is the Hamiltonian corresponding to the first two terms in Eq. (2), and

$$\hat{H}_1 = -\frac{1}{2k^2 n_0} \begin{pmatrix} \dfrac{\partial}{\partial x}\left(\dfrac{1}{n^2}\dfrac{\partial n^2}{\partial x} \right) & \dfrac{\partial}{\partial x}\left(\dfrac{1}{n^2}\dfrac{\partial n^2}{\partial y} \right) \\ \dfrac{\partial}{\partial y}\left(\dfrac{1}{n^2}\dfrac{\partial n^2}{\partial x} \right) & \dfrac{\partial}{\partial y}\left(\dfrac{1}{n^2}\dfrac{\partial n^2}{\partial y} \right) \end{pmatrix}$$

is the Hamiltonian corresponding to the third term in Eq. (2).

The Hamiltonian $\hat{H}$ can be expressed in terms of the annihilation and creation operators $\hat{a}$ and $\hat{a}^+$ by means of the relations

$$\hat{x} = \frac{1}{\sqrt{2k\omega}}(\hat{a}_1 + \hat{a}_1^+), \quad \hat{y} = \frac{1}{\sqrt{2k\omega}}(\hat{a}_2 + \hat{a}_2^+),$$

$$\frac{\partial}{\partial x} = \left(\frac{k\omega}{2} \right)^{1/2}(\hat{a}_1 - \hat{a}_1^+), \quad \frac{\partial}{\partial y} = \left(\frac{k\omega}{2} \right)^{1/2}(\hat{a}_2 - \hat{a}_2^+).$$

We therefore have

$$\hat{H}_0 = \frac{\omega}{kn_0}\,(\hat{a}_1^+ \hat{a}_1 + \hat{a}_2^+ \hat{a}_2 + 1)\hat{I},$$

$$\hat{H}_1 = \eta(\hat{A}\hat{\sigma}_z + \hat{B}\hat{\sigma}_x + k\hat{M}_z\hat{\sigma}_y + \hat{C}\hat{I} + \hat{I}).$$

Here

$$\eta = \frac{\omega^2}{2k^2 n_0^3}, \quad \hat{A} = \frac{1}{2}(\hat{a}_1^2 - \hat{a}_1^{+2} - \hat{a}_2^2 + \hat{a}_2^{+2}),$$

$$\hat{B} = \hat{a}_1\hat{a}_2 - \hat{a}_1^+\hat{a}_2^+,$$

$$\hat{M}_z = -\frac{i}{k}(\hat{a}_1^+\hat{a}_2 - \hat{a}_1\hat{a}_2^+), \quad \hat{C} = \frac{1}{2}(\hat{a}_1^2 - \hat{a}_1^{+2} + \hat{a}_2^2 - \hat{a}_2^{+2}),$$

$$\hat{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

is the unit matrix, and

$$\hat{\sigma}_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \hat{\sigma}_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \hat{\sigma}_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

are the Pauli spin matrices.

In slightly inhomogeneous media we have $\eta \ll 1$. The Pauli matrices satisfy the relations

$$\hat{\sigma}_x\hat{\sigma}_y = i\hat{\sigma}_z, \quad \hat{\sigma}_y\hat{\sigma}_z = i\hat{\sigma}_x, \quad \hat{\sigma}_z\hat{\sigma}_x = i\hat{\sigma}_y, \tag{8}$$

$$\hat{\sigma}_y\hat{\sigma}_x = -i\hat{\sigma}_z, \quad \hat{\sigma}_z\hat{\sigma}_y = -i\hat{\sigma}_x, \quad \hat{\sigma}_x\hat{\sigma}_z = -i\hat{\sigma}_y .$$

Note that the Hamiltonian $\hat{H}$ is non-Hermitian, i.e., $\hat{H}^+ \neq \hat{H}$. The solution of Eq. (7) can be expressed in terms of the evolution operator $U = \exp(-ik\hat{H}z)$:

$$\Psi(x,y,z) = \hat{U}\Psi(x,y,0).$$

The wave function $\psi(x,y,0)$ describes both the dependence on the space coordinates and the polarization state, which is characterized by the Jones vector in the plane $z=0$. For example, the wave functions describing the linearly polarized state have the form

$$|e_x\rangle = |\alpha_1\alpha_2\rangle\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |e_y\rangle = |\alpha_1\alpha_2\rangle\begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{9}$$

The wave functions of right- and left-circularly polarized states have the forms (respectively)

$$|+1\rangle = \frac{1}{\sqrt{2}}|\alpha_1\alpha_2\rangle\begin{pmatrix} 1 \\ i \end{pmatrix}, \quad |-1\rangle = \frac{1}{\sqrt{2}}|\alpha_1\alpha_2\rangle\begin{pmatrix} 1 \\ -i \end{pmatrix}. \tag{10}$$

or

$$|+1\rangle = \frac{1}{\sqrt{2}}(|e_x\rangle + i|e_y\rangle), \quad |-1\rangle = \frac{1}{\sqrt{2}}(|e_x\rangle - i|e_y\rangle), \tag{11}$$

Polarized radiation is described by the coherency matrix[26]

$$J = \begin{pmatrix} \langle E_x^* E_x \rangle & \langle E_x^* E_y \rangle \\ \langle E_y^* E_x \rangle & \langle E_y^* E_y \rangle \end{pmatrix}. \tag{12}$$

The angle brackets signify statistical ensemble averages. The degree of polarization is given by[26]

$$P = \left(1 - \frac{4\det J}{\mathrm{Tr}^2 J} \right)^{1/2}. \tag{13}$$

3. EVOLUTION OF THE COHERENCY MATRIX

The polarization characteristics of the radiation are described by the elements of the coherency matrix. It would be useful, therefore, to investigate the evolution of the coherency matrix, which is given by the equation

$$\hat{J}(z) = \hat{U}^+\hat{J}(0)\hat{U}$$

or by the equation

$$-\frac{i}{k}\frac{d\hat{J}}{dz} = \hat{H}^+\hat{J} - \hat{J}\hat{H}, \tag{14}$$

where $J(0)$ is the operator of the coherency matrix in the initial plane $z=0$.

Equation (14) is the same as the equation for the correlation function describing the evolution of the spatial coherence of radiation.[20]

3.1. Linear polarization

The wave functions of linearly polarized beams are given by Eqs. (9). The elements of the coherency matrix are

$$J_{i,j} = \langle e_i|\hat{J}(z)|e_j\rangle, \quad i,j = x,y. \tag{15}$$

The operator of the coherency matrix in the plane $z=0$ is determined by the source of the radiation. For example, in the case of radiation linearly polarized along the x axis, we have[26]

$$\hat{J}(0)=I_0\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}=\frac{I_0}{2}(1+\hat{\sigma}_z), \tag{16}$$

where I_0 is the total intensity of the incident beam.

For radiation linearly polarized along the y axis, the operator of the coherency matrix has the form[26]

$$\hat{J}(0)=I_0\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}=\frac{I_0}{2}(1-\hat{\sigma}_z), \tag{17}$$

To solve Eq. (14), we use the relations $[\hat{a}_i,\hat{a}_j^+]=\delta_{i,j}$ $(i,j=1,2)$, Eqs. (8), and the expressions

$$\begin{aligned}&\hat{\sigma}_z|e_z\rangle=|e_x\rangle,\quad \hat{\sigma}_x|e_x\rangle=|e_y\rangle,\quad \hat{\sigma}_y|e_x\rangle=i|e_y\rangle,\\ &\hat{\sigma}_z|e_y\rangle=-|e_y\rangle,\quad \hat{\sigma}_x|e_y\rangle=-|e_x\rangle,\quad \hat{\sigma}_y|e_y\rangle=-i|e_x\rangle\end{aligned} \tag{18}$$

in calculating the matrix elements (15).

Solving Eq. (14) to within small terms proportional to η^2 and substituting the solutions into (13), we obtain the following equation for the quantity representing depolarization:

$$\frac{4\det J}{\mathrm{Tr}^2\,J}=\frac{\omega^2}{2k^2n_0^5}\left[\left(k\omega x_0^2+\frac{k}{\omega}p_{y0}^2\right)(\omega^2z^2+\sin^2(\omega z)-\omega z\,\sin(2\omega z))+2\sin^2(\omega z)\right]. \tag{19}$$

The inclusion of the next two terms in order of smallness, proportional to η^3 and η^4, introduces into (19) terms of higher order in z, which provide a significant contribution at distances greater than a certain z_0. Equation (19) is therefore valid only up to distances $z\leqslant z_0$, where z_0 is dictated by the accuracy of the solution of Eq. (14). In our case this distance is

$$z_0\approx\frac{1}{2k\eta}.$$

Equation (19) takes into account both meridional and sagittal rays with initial coordinates in the plane $z=0$: $x_0\neq0$, $y_0=0$, $p_{x0}=0$, $p_{y0}\neq0$. It is evident from Eq. (19) that the depolarization depends on the wave number and vanishes in the limit $\lambda\to0$. The degree of polarization of pure diffraction origin oscillates in the case of the axial ray ($x_0=0$, $p_{y0}=0$). Inasmuch as the axial ray corresponds to the fundamental mode of the fiber, periodic recovery of the degree of polarization should be observed in a single-mode isotropic fiber.

The dependence of the degree of polarization of linearly polarized radiation on the distance z is represented by the curves in Fig. 1 for various offsets x_0, including meridional rays and radius-preserving ($p_{y0}=x_0\omega$) helically twisted sagittal rays. Clearly, the degree of polarization decreases inversely as the distance squared. The reduction in the degree of polarization becomes more pronounced as the offset or tilt

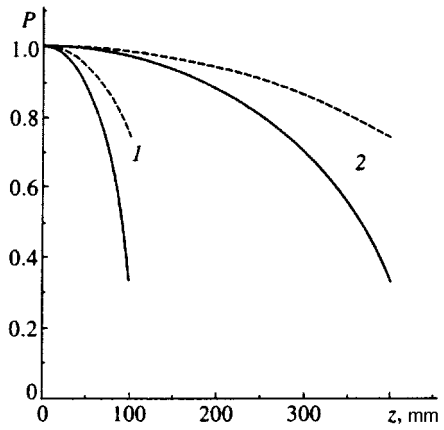


FIG. 1. Degree of linear polarization versus distance for both meridional and sagittal rays (solid curves) and for meridional rays only (dashed curves). *1)* $x_0=20$ μm; *2)* $x_0=5$ μm.

of the beam relative to the fiber axis is increased. The depolarization length of meridional rays is 1.4 times the depolarization length of sagittal rays.

The reduction in the degree of polarization is accompanied by small oscillations. Figure 2a shows the variation of the function $F=10^5(4\det J/\mathrm{Tr}^2\,J)$ with distance z. We see that the rise in this quantity is accompanied by small oscillations.

Substituting Eq. (19) into (13), we obtain

$$P_l\approx1-\frac{1}{4}\frac{\omega}{kn_0^5}(\omega^2x_0^2+p_{y0}^2)(\omega^2z^2+\sin^2(\omega z)-\omega z\,\sin(2\omega z))-\frac{1}{2}\frac{\omega^2}{k^2n_0^5}\sin^2(\omega z). \tag{20}$$

At distances $z\gg\omega^{-1}$, Eq. (20) acquires the form

$$P_l\approx1-\frac{1}{4}\frac{\omega^3}{kn_0^5}(\omega^2x_0^2+p_{y0}^2)z^2-\frac{1}{2}\frac{\omega^2}{k^2n_0^5}\sin^2(\omega z).$$

Consequently, the degree of polarization decreases inversely as the distance squared, and depolarization vanishes in the limit $\lambda\to0$.

For an axial beam ($x_0=0$, $p_{y0}=0$) we encounter oscillations in the degree of polarization of pure diffraction origin:

$$P_l\approx1-\frac{1}{2}\frac{\omega^2}{k^2n_0^5}\sin^2(\omega z). \tag{21}$$

Clearly, the degree of polarization of an axial beam acquires its initial value for $\omega z=n\pi$.

3.2. Circular polarization

The wave functions of right-circularly and left-circularly polarized radiation have the form (10). Consider the evolution of right-circularly polarized radiation. The initial coherency matrix then has the form[26]

$$\hat{J}(0)=\frac{1}{2}I_0\begin{pmatrix}1 & -i\\ i & 1\end{pmatrix}=\frac{I_0}{2}(1+\hat{\sigma}_y). \tag{22}$$

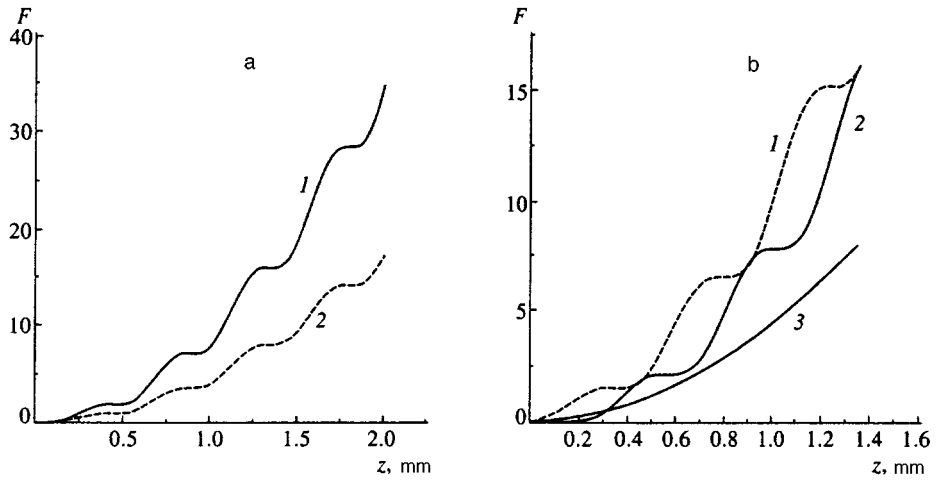


FIG. 2. Function $F=10^5(4\det J/\mathrm{Tr}^2\, J)$ versus distance for an incident-beam offset $x_0=20\ \mu$m. a) Linear polarization: *1*) meridional and sagittal rays; *2*) meridional ray only. b) Circular polarization: *1*) sagittal ray with positive helicity ($p_{y0}=x_0\omega$); *2*) sagittal ray with negative helicity ($p_{y0}=-x_0\omega$); *3*) meridional ray.

The Jones vector of linearly polarized radiation can be expressed in terms of the Jones vector of circularly polarized radiation:

$$E=\begin{bmatrix}E_x\\E_y\end{bmatrix}=\hat{F}\begin{bmatrix}E_+\\E_-\end{bmatrix}, \tag{23}$$

where

$$\hat{F}=\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ -i & i\end{bmatrix}$$

is the transformation matrix.

Accordingly, the evolution of the coherency matrix is given by

$$-\frac{i}{k}\frac{d\hat{J}}{dz}=\hat{H}^{+}\hat{J}-\hat{J}\hat{H}, \tag{24}$$

where $\hat{H}=\hat{F}^{-1}\hat{H}\hat{F}$.

Solving this equation to within small terms in η^2, we obtain an equation for the quantity characterizing the degree of polarization:

$$\frac{4\det J}{\mathrm{Tr}^2\, J}=\frac{\omega^2}{2k^2n_0^5}\left[\left(k\omega x_0^2+\frac{k}{\omega}p_{y0}^2\right)(\omega^2z^2+\sin^2(\omega z))\right.$$
$$\left.\pm 4kx_0p_{y0}\omega z\,\sin^2(\omega z)+2\sin^2(\omega z)\right]. \tag{25}$$

The plus sign corresponds to right-circular polarization, and the minus sign to left-circular polarization. There is clearly an asymmetry with respect to the sign of the twist of sagittal ray paths (Fig. 2b). The depolarization is stronger when the helicities of the ray path and the ''photon'' path have the same sign, and it is weaker when the helicities have opposite signs. The depolarization of meridional rays is weaker than that of sagittal rays.

In the foregoing discussion we have made use of Eqs. (10) and the relations

$$\hat{\sigma}_y|+1\rangle=|+1\rangle,\quad \hat{\sigma}_x|+1\rangle=i|-1\rangle,\quad \hat{\sigma}_z|+1\rangle=|-1\rangle,$$
$$\hat{\sigma}_y|-1\rangle=-|-1\rangle,\quad \hat{\sigma}_x|-1\rangle=-i|+1\rangle,$$
$$\hat{\sigma}_z|-1\rangle=|+1\rangle.$$

Substituting Eq. (25) into (13), we obtain an expression for the degree of polarization

$$P_c\approx 1-\frac{1}{4}\frac{\omega}{kn_0^5}\left[(\omega^2x_0^2+p_{y0}^2)(\omega^2z^2+\sin^2(\omega z))\right.$$
$$\left.\pm 4\omega x_0p_{y0}\omega z\,\sin^2(\omega z)+2\frac{\omega}{k}\sin^2(\omega z)\right]. \tag{26}$$

At distances $z\gg\omega^{-1}$, the dependence of the degree of polarization on the distance has the form

$$P_c\approx 1-\frac{1}{4}\frac{\omega^3}{kn_0^5}(\omega^2x_0^2+p_{y0}^2)z^2-\frac{1}{2}\frac{\omega^2}{k^2n_0^5}\sin^2(\omega z).$$

This expression takes into account the condition $p_{y0}=x_0\omega$ for radius-preserving sagittal rays.

Consequently, like the degree of polarization of linearly polarized radiation, the degree of polarization of circularly polarized radiation decreases inversely as the distance squared. Note that an analogous result is obtained for left-circularly polarized radiation.

For an axial beam ($x_0=0$, $p_{y0}=0$) we have

$$P_c\approx 1-\frac{1}{2}\frac{\omega^2}{k^2n_0^5}\sin^2(\omega z). \tag{27}$$

It is evident that the degree of polarization of an axial beam (fundamental mode) oscillates periodically during propagation, decreasing to a level that depends on the gradient parameter of the medium and the wavelength of the radiation, and then increasing back to its original value.

4. EVOLUTION OF THE BERRY PHASE

The evolution of the polarization vector is described by the elements of the coherency matrix. For example, the rotation of the polarization vector of a beam polarized along the x axis and propagating along a helical path is given by

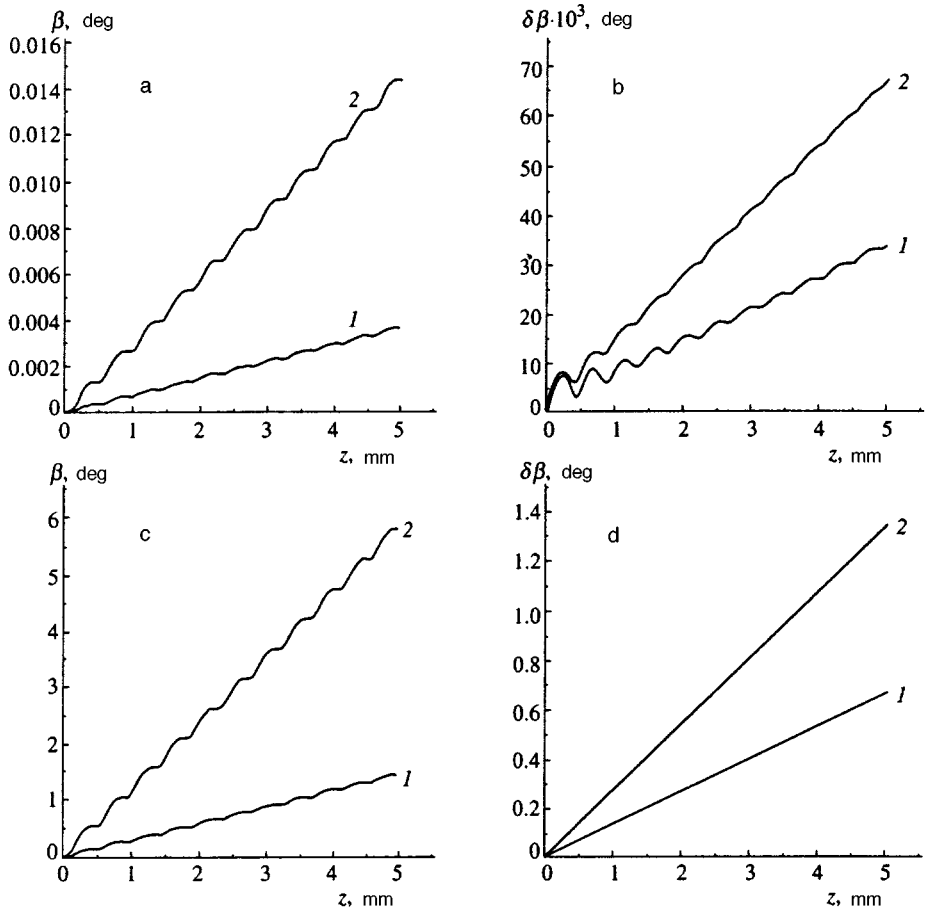


FIG. 3. Angle of rotation β of the polarization vector (a,c) and the spread in the angle of rotation $\delta\beta$ (b,d) versus distance for various incident-beam offsets. a,b: *1)* $x_0=0.5\ \mu$m; *2)* $x_0=1.0\ \mu$m; c,d: *1)* $x_0=10\ \mu$m; *2)* $x_0=20\ \mu$m.

$$\langle\beta\rangle=\frac{1}{2}\arctan\frac{J_{xy}+J_{yx}}{J_{xx}-J_{yy}}$$

$$\approx\frac{1}{2}\arctan\left[\frac{\omega x_0 p_{y0}}{2n_0^3}(2\omega z-\sin(2\omega z))\right]. \tag{28}$$

Equation (28) has been obtained with the inclusion of second-order terms, and holds for sagittal rays with initial coordinates $x_0\neq 0$, $y_0=0$, $p_{x0}=0$, $p_{y0}\neq 0$.

Figures 3a and 3c show the angle of rotation of the polarization plane as a function of distance in an optical fiber with parameters $\omega=7\times10^{-3}\ \mu\text{m}^{-1}$ and $n_0=1.5$ for various offsets. We see that the angle of rotation increases with distance, with sinusoidal oscillations superimposed on a linear trend. Note that the angle of rotation increases linearly with the distance only for comparatively small offsets. Equation (28) then has the form

$$\langle\beta\rangle=\frac{\omega x_0 p_{y0}}{4n_0^3}(2\omega z-\sin(2\omega z)). \tag{29}$$

At distances $z\gg\omega^{-1}$, Eq. (29) assumes the form

$$\langle\beta\rangle=\frac{\omega x_0 p_{y0}}{2n_0^3}\omega z,$$

which coincides with the equation derived in Ref. 27. For $x_0\geqslant\sqrt{2/k\omega}$ (the quantity $\sqrt{2/k\omega}$ specifies the width of the fundamental mode of the fiber), the angle of rotation increases more slowly with increasing distance. However, effects attributable to nonparaxial ray propagation must be taken into account for large offsets of the beam.

The rate of rotation of the polarization vector is given by

$$\nu=\frac{d\langle\beta\rangle}{dz}\approx\frac{\omega^2 x_0 p_{y0}}{2n_0^3}(1-\cos(2\omega z)). \tag{30}$$

Consequently, the rotation of the polarization vector exhibits nonuniform behavior. The rate of rotation is zero where the polarization vector is parallel to the principal normal.

We note that Rytov rotation (or the Berry phase) (Ref. 13) is of pure geometrical origin and does not depend on the wavelength of the radiation. However, the spread in the Berry phase is a wave phenomenon and vanishes in the limit $\lambda\to 0$. Thus, it is given by

$$\langle\delta\beta\rangle^2=\langle\beta^2\rangle-\langle\beta\rangle^2\approx\frac{1}{2}(1-P)$$

$$\approx\frac{\omega^3}{8kn_0^5}(\omega^2x_0^2+p_{y0}^2)z^2+\frac{\omega^2}{4k^2n_0^5}\sin^2(\omega z). \tag{31}$$

The spread in the Berry phase is governed by the degree of polarization of the radiation, and can be determined from measurements of the degree of polarization P: $\langle\delta\beta\rangle\approx(1/2)\sqrt{1-P^2}$. Zel'dovich and Kundinova[14] point out that the depolarization in an optical fiber of length $z\approx 7.5$ cm is 10–30%, which is consistent with the results obtained

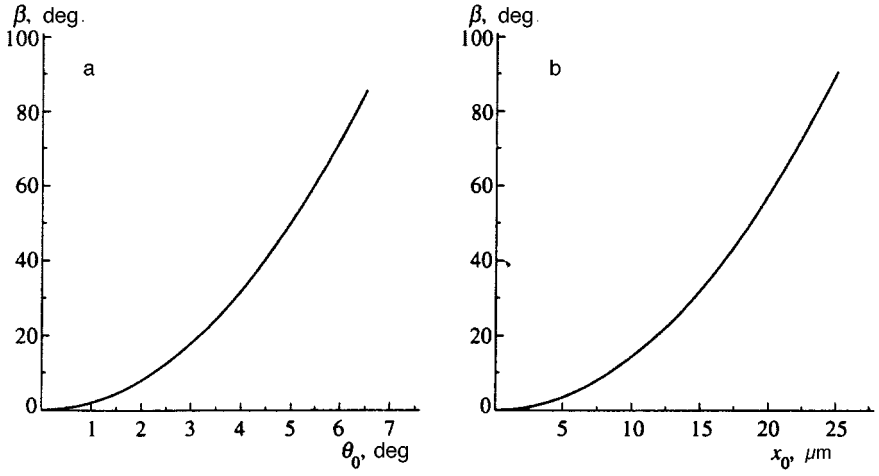


FIG. 4. Angle of rotation of the polarization plane versus a) tilt angle θ_0 of the beam relative to the fiber axis; b) offset x_0 in a graded-index optical fiber with parameters $\omega = 7\times10^{-3}$ μm^{-1}, $n_0 = 1.5$, and length $z = 5$ cm.

here. Clearly, nonvanishing oscillations of the spread in the Berry phase occur in an axial beam, even though the phase itself is equal to zero in this case.

The spread in the Berry phase is a wave phenomenon and vanishes in the geometrical-optics approximation.

Figures 3b and 3d show the spread in the Berry phase as a function of the distance for various offsets of the beam axis. It is evident that the dispersion of the Berry phase increases nonlinearly for relatively small offsets. For an offset $x_0 = 20$ μm, appreciable spreads in the Berry phase, of the order of $\langle \delta\beta \rangle \approx 1°$, occur at distances $z \approx 3.7$ mm. This result corresponds to rotation of the polarization plane through an angle $\langle \beta \rangle \approx 4.4°$.

For radius-preserving ($p_{y0} = x_0\omega$) helically twisted sagittal rays, it follows from Eq. (29) that the rotation angle of the polarization plane increases as the square of the tilt angle θ_0 of the incident beam relative to the fiber axis ($p_{y0} = n_0 \sin\theta_0$) or as the square of the offset x_0 (Figs. 4a and 4b). These results are in good agreement with experimental data. Experiments[14] show that the angle of rotation of the polarization plane of a sagittal ray with a tilt angle in air $\varphi_0 \approx n_0\theta_0 = 12.5°$ relative to the fiber axis ($\Delta n = 0.016$, $n_0 = 1.53$) and length $z = 7.5$ cm is $\beta \approx 115°$. For a radius-preserving, helically twisted sagittal ray with $x_0 = p_{y0}/\omega$ and $\theta_0 \approx 8.19°$ in a fiber of length $z = 7.5$ cm with $\Delta n = 0.016$, it follows from Eq. (29) that the rotation angle of the polarization plane is $\beta \approx 122°$ (Fig. 5). It is evident from Fig. 5 that the theoretical curve describes the experimental data quite well, in spite of the difference in the index profiles of the investigated optical fibers. Note that Liberman and Zel'dovich[28] obtain a cubic dependence of the rotation angle of the polarization plane on the angle between the direction of propagation and the axis of a step-index fiber.

5. DISCUSSION AND CONCLUSION

It is instructive to determine the depolarization length of radiation in real optical fibers. We know[6,7,11] that the propagation of coherent radiation in optical fibers is accompanied by significant changes in the state and degree of polarization. It is evident from Eqs. (20) and (26) that the degree of polarization depends on the gradient parameter ω of the fiber, the incident-beam offset x_0, and the radiation wavelength λ. The depolarization length corresponding to 50% reduction in the degree of polarization of the radiation can be determined from Eqs. (20) and (26):

$$l_d \approx \sqrt{\frac{kn_0}{\omega}}\frac{n_0^2}{\omega^2 x_0}.$$

Substituting the values of the gradient parameter $\omega = 7\times10^{-3}$ μm^{-1} (for a fiber of radius $r_0 = x_0 = 25$ μm and relative difference in the refractive index $\Delta \approx 6.9\times10^{-3}$) and the radiation wavelength $\lambda = 0.63$ μm into the above equation, we obtain $l_d \approx 8.3$ cm. This level of depolarization is observed experimentally in an isotropic multimode optical fiber.[11] The depolarization is intensified when the offset or tilt of the incident beam relative to the fiber axis is increased. Since beams with a large offset x_0 correspond to higher waveguide modes, the depolarization of higher modes is stronger than that of lower modes. Consequently, in coherent communication systems it is better to use single-mode fibers that are not birefringent, where the degree of polarization

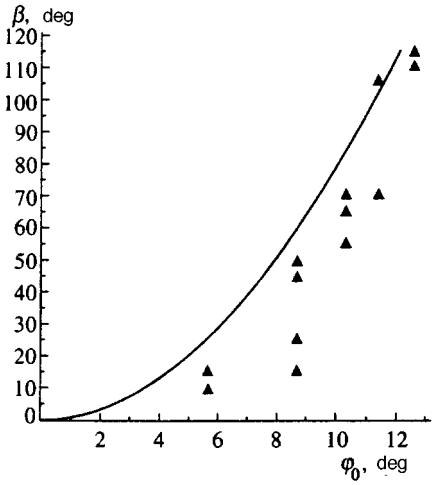


FIG. 5. Angle of rotation of the polarization plane versus tilt angle of the beam relative to the axis of an optical fiber of length 7.5 cm in air; theoretical (solid curve) and experimental[14] values (triangles).

recovers with a period $z=\pi/\omega$ during the propagation of radiation. In graded-index fibers this length is $z\approx 0.5$ mm.

As shown above, inclusion of the ''spin–orbit'' term in Eq. (2) leads to periodic recovery of the polarization of an axial wave beam, which corresponds to the fundamental mode. Note that similar oscillations are exhibited by the width of an axial beam. In the scalar wave problem, the width of an axial beam corresponding to the fundamental mode does not change during propagation, because diffraction spreading of the beam is compensated by the waveguide focusing effect. This shows that a Gaussian wave packet, although an exact solution of the scalar problem, is only an approximate solution of the vector wave equation.

The physical mechanism underlying oscillations in both the degree of polarization and the width of the beam is diffraction. Oscillations of the polarization and the beam width are interdependent in this sense. Most likely, wave packets exist that preserve their width and polarization. It would be interesting, therefore, to look for such wave packets for a given fiber index profile, or to determine index profiles for given beam parameters such that the polarization and beam width are unchanged during propagation. However, this problem requires further investigation.

The depolarization of radiation is a diffractive phenomenon, and the interpretation of Rytov rotation as a ''geometrical'' mechanism of depolarization is incorrect. Rytov rotation merely causes the polarization plane to rotate, but does not alter the degree of polarization. Furthermore, the rotation angle of the polarization plane does not diminish as the radiation wavelength decreases, whereas depolarization vanishes in the limit $\lambda\to 0$. As shown above, depolarization also occurs for meridional rays, which are not subject to Rytov rotation.

Physically, depolarization corresponds to a decrease in the degree of correlation between various components of the field, or to the emergence of a field component that does not correlate with an initial component. In the real world, depolarization is often induced by the scattering of radiation by small-scale inhomogeneities of the medium. The physical cause of depolarization in our case is the scattering of radiation by large-scale inhomogeneities of the medium. Since large-scale inhomogeneities have little influence on the polarization characteristics, they are usually ignored. In waveguide problems, however, small changes accumulate with distance and can produce significant effects.

The depolarization of a circularly polarized laser beam during propagation in the atmosphere is observed experimentally.[29] Atmospheric turbulence is one proposed mechanism to explain the phenomenon. However, it has been shown theoretically[1,2] that atmospheric turbulence cannot account for the experimentally observed depolarization. Depolarization is most likely attributable in this case to the diffraction effects discussed above. A laser beam usually has a nonuniform Gaussian profile. The radiation therefore propagates in a channel with a nonuniform distribution of the refractive index. The gradient parameter ω characterizing this nonuniformity is specified by the beam width $a=\sqrt{2/k\omega}$. For given values of the distance $z=8.045$ km and the radiation wavelength $\lambda=0.6328\ \mu$m and for the measured value of the degree of polarization $P\approx 0.9$ we infer from Eq. (26) that the gradient parameter is $\omega\approx 2.8\times 10^{-5}\ \mu\text{m}^{-1}$. This value corresponds to a diameter of the fundamental mode of the laser beam $a\approx 85\ \mu$m, which is consistent with actual laser beam diameters. Note that depolarization increases as the beam diameter is reduced.

To summarize, the decrease in the degree of polarization in multimode optical fibers has a diffraction origin and can only be explained in wave language. Inasmuch as depolarization vanishes both in the limit $\lambda\to 0$ and in the limit $x_0\to 0$, it can be interpreted as the result of interaction between polarization (spin) and the trajectory (orbital angular momentum). The degree of polarization of both linearly and circularly polarized beams decreases inversely as the distance squared. The depolarization of meridional rays is weaker than that of sagittal rays. Right-circularly polarized radiation and left-circularly polarized radiation are asymmetric with respect to the sign of the twist of sagittal rays. Depolarization intensifies when the offset of the beam axis, the gradient parameter of the fiber, or the wavelength of the radiation increases.

The rotation of the polarization vector in a graded-index optical fiber is nonuniform. The linear growth of the angle of rotation of the polarization plane with increasing distance is accompanied by sinusoidal oscillations. The spread in the Berry phase or the rotation angle of the polarization plane increases with distance, accompanied by fluctuations in the spread of the Berry phase for an axial beam (fundamental mode); these fluctuations are a wave phenomenon. The rotation of the polarization vector increases linearly with distance, and as the squares of the offset and tilt angle of a sagittal ray relative to the fiber axis.

The proposed approach can be used to investigate the evolution of polarization in birefringent, absorbing, or amplifying optical fibers, as well as in fibers with random inhomogeneities.

[1] V. I. Tatarskiĭ, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **10**, 1762 (1967).
[2] Yu. A. Kravtsov, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **13**, 281 (1970).
[3] S. M. Rytov, Dokl. Akad. Nauk SSSR **18**, 2 (1938).
[4] A. S. McLean and J. B. Pendry, J. Mod. Opt. **42**, 339 (1995).
[5] L. G. Cohen, Bell Syst. Tech. J. **50**, 23 (1971).
[6] A. M. Smith, Appl. Opt. **17**, 52 (1978).
[7] P. Kaminov, IEEE J. Quantum Electron. **QE-17**, 15 (1981).
[8] C. R. Manyuk and P. K. Wai, J. Opt. Soc. Am. B **11**, 1288 (1994).
[9] A. D. Shatrov, Radiotekh. Élektron. **26**, 505 (1981).
[10] A. A. Esayan and B. Ya. Zel'dovich, Kvantovaya Élektron. **15**, 235 (1988) [Sov. J. Quantum Electron. **18**, 149 (1988)].
[11] O. I. Kotov, O. L. Marusov, O. L. Nikolaev, and V. N. Filippov, Opt. Spektrosk. **70**, 924 (1991) [Opt. Spectrosc. **70**, 540 (1991)].
[12] A. Tomita and R. Y. Chiao, Phys. Rev. Lett. **57**, 936 (1986).
[13] M. V. Berry, Proc. R. Soc. London, Ser. A **392**, 45 (1984).
[14] B. Ya. Zel'dovich and N. D. Kundinova, Kvantovaya Élektron. **22**, 184 (1995).
[15] M. A. Leontovich and V. A. Fok, Zh. Éksp. Teor. Fiz. **16**, 557 (1946).
[16] D. Marcuse, *Light Transmission Optics*, 2nd ed., Van Nostrand Reinhold, Princeton, N.J. (1982).
[17] J. A. Arnaud, *Beam and Fiber Optics*, Van Nostrand, New York (1976).
[18] G. Eichmann, J. Opt. Soc. Am. **61**, 161 (1971).
[19] S. G. Krivoshlykov and I. N. Sisakyan, Opt. Quantum Electron. **12**, 463 (1980).

[20] S. G. Krivoshlykov, N. I. Petrov, and I. N. Sisakyan, Opt. Quantum Electron. **18**, 253 (1986).
[21] I. A. Malkin and V. I. Man'ko, *Dynamical Symmetries and Coherent States of Quantum Systems* [in Russian], Nauka, Moscow (1979).
[22] N. I. Petrov, J. Mod. Opt. **43**, 2239 (1996).
[23] J. R. Klauder and E. C. G. Sudarshan, *Fundamentals of Quantum Optics*, Benjamin, New York (1968).
[24] R. J. Glauber, Phys. Rev. **131**, 2766 (1963).
[25] E. Schrödinger, Naturwissenschaften **14**, 664 (1926).
[26] R. M. A. Azzam and N. M. Bashara, *Ellipsometry and Polarized Light*, North-Holland Publ. Co., Amsterdam; Elsevier North-Holland, New York (1977).
[27] V. S. Liberman and B. Ya. Zel'dovich, Phys. Rev. A **46**, 5199 (1992).
[28] V. S. Liberman and B. Ya. Zel'dovich, Pure Appl. Opt. **2**, 367 (1993).
[29] D. L. Fried and G. E. Meyers, J. Opt. Soc. Am. **55**, 740 (1965).

Translated by James S. Wood

Nonlinear diffraction in spontaneous three-wave and coherent four-wave light scattering by polaritons

G. Kh. Kitaeva, A. A. Mikhailovskiĭ, and A. N. Penin

M. V. Lomonosov Moscow State University, 119899 Moscow, Russia
(Submitted 24 June 1997)
Zh. Éksp. Teor. Fiz. **112**, 2001–2015 (December 1997)

Spontaneous three-wave and coherent four-wave scattering by polaritons in crystalline media with periodically modulated nonlinear quadratic susceptibility have been studied both theoretically and experimentally. Phase matching conditions and an expression for the scattering spectral line shape with due account of cascade processes in nonlinear diffraction for coherent Stokes polariton light scattering have been derived. Measurements of the light intensity distribution due to the three- and four-wave scattering in a $LiNbO_3$:Nd:Mg crystal with a periodic domain structure are in good agreement with theoretical results. The prospects for using the effects of nonlinear diffraction in spectroscopic studies of media with periodic distributions of nonlinear optical parameters, specifically, in precision measurements of the IR refractive index dispersion and determination of the period and profile of the quadratic susceptibility distribution are discussed. © *1997 American Institute of Physics.* [S1063-7761(97)00712-9]

1. INTRODUCTION

In recent years, there has been steady interest in crystalline structures with periodically modulated optical parameters. The reason is that such materials can be used in frequency converters of electromagnetic radiation (light harmonic converters, frequency doublers, etc.) based on quasi-matched parametric interactions. Before using such materials, one should study the quantitative characteristics of their optical properties over a broad spectral range. Such characteristics include dispersion of the refractive index in both visible and infrared ranges, the period and alignment of a spatial grating, and its profile. These characteristics are measured using various linear and nonlinear optical effects, but only the latter yield the required information about a test sample.

One well-known method of measuring the dispersions of the linear and quadratic susceptibility of crystals in the IR range is by spectroscopy of spontaneous three-wave parametric light scattering.[1] Spontaneous parametric scattering in media with a periodic distribution of quadratic susceptibility demonstrates an effect of nonlinear diffraction resulting in additional spectral features related to the quasi-matched interaction.[2] Measurements of the spectrum and angular distribution of spontaneous parametric scattering in such materials yield dispersion curves in the polariton spectral range, the period of a quadratic susceptibility distribution, and its shape. Notable flaws of this technique, however, are the relatively poor accuracy of refractive index measurements (especially in the range of low-frequency polariton states), the low intensity of signals due to higher harmonics of the quadratic susceptibility space distribution, and poor spatial resolution in measurements of variations in parameters of a nonlinear superlattice generated in a crystal.

Processes of four-wave mixing in media with periodic distribution of nonlinear optical parameters have not been studied in detail by this time, although they can play a decisive role in spectroscopic applications. In this context, nonlinear diffraction is of special interest for measurements of the polariton light scattering. The technique of coherent anti-Stokes Raman scattering (CARS) has been widely used in coherent spectroscopy.[3] Measurements of the Stokes components of scattered light, however, are lacking, although this version of the coherent four-wave light scattering is preferable for studies of crystals without a symmetry center. Although the number of components in the experimental apparatus required for this technique is larger, since all interacting waves should have different frequencies, the optical system is aligned more easily by switching from recording spectra of spontaneous parametric scattering to spectra of coherent light scattering. By using the effect of coherent Stokes Raman scattering (CSRS), one can reduce the uncertainty of measured dispersion curves of the polariton refractive index by one or two orders of magnitude and investigate small samples of materials. The latter circumstance also allows one to study nonperiodic distributions of optical parameters. It is obvious that the effect of nonlinear diffraction in the coherent Stokes scattering should give rise to emergence of additional features in the spectrum and angular distribution of scattered light, and multistep diffraction processes, in which nonlinear diffraction occurs in both polariton state generation and probing light scattering due to this state, should take place. When active spectroscopic methods are employed, signals due to weak spatial harmonics of the nonlinear susceptibility, which do not manifest themselves in spontaneous parametric scattering owing to their small amplitudes, can be detected. Taking into account the measured efficiences of the high-order nonlinear diffraction one can reconstruct the profile of a nonlinear spatial grating with a higher accuracy.

The aim of our study was to investigate spontaneous parametric scattering and coherent Stokes light scattering in materials with a periodic distribution of nonlinear susceptibility using samples of multiply domained $LiNbO_3$:Nd:Mg

crystals and to assess the feasibility of applying these effects to spectroscopic techniques.

2. THEORY

2.1. Three- and four-wave light scattering in a homogeneous noncentrosymmetric medium

In spontaneous parametric scattering, the pump optical wave is scattered due to vacuum fluctuations in a noncentrosymmetric medium.[1] The frequencies of the three waves are related by the temporal phase matching:

$$\omega_L-\omega_P=\omega_S, \tag{1}$$

and the wave vectors $\mathbf{k}_i$ by the condition of spatial phase matching, which is expressed for a homogeneous medium by

$$\mathbf{k}_L-\mathbf{k}_P=\mathbf{k}_S, \tag{2}$$

where the subscripts L, P, and S denote the laser (pumping), polariton (idle), and signal waves, respectively. Condition (2) should be satisfied to within some allowed detuning $\Delta k=|\mathbf{k}_L-\mathbf{k}_P-\mathbf{k}_S|$, which is determined by the absorption and the length of the scattering medium.

Spontaneous parametric scattering is used in measurements of dispersions of the real and imaginary parts of the polariton wave vector. To this end, the pumping and signal frequencies are selected in the transparent range of the material, and the frequency ω_P is in the polariton band of the spectrum. After leaving the test material, the scattered wave is measured to find its distribution in frequency and angle, $I_S(\theta_S,\omega_S)$, where θ_S is the scattering angle in the plane of wave vectors. The detected signal has a maximum when the spatial matching condition (2) is satisfied exactly. The polariton dispersion curve is determined using the tuning curve $\theta_{S0}(\omega_S)$, which connects points of maximum intensity, where

$$\left.\frac{\partial I(\theta_S,\omega_S)}{\partial\theta_S}\right|_{\omega_S=\text{const}}=0.$$

The linear absorption factor for polaritons and the imaginary part of their wave vector are derived from the angular line width of the scattered light spectrum $I(\theta_S,\omega_S=\text{const})$.

Coherent four-wave scattering process in a noncentrossymmetric media can be divided into two types, namely direct and multistep processes.[3] Direct scattering is the parametric interaction of three pumping waves generating a fourth wave whose frequency and wave vector satisfy the matching conditions temporal and spatial. In the case of Stokes scattering, the matching conditions are expressed as follows:

$$\omega_S=\omega_L-\omega_1+\omega_2, \tag{3}$$

$$\Delta\mathbf{k}\equiv\mathbf{k}_S-\mathbf{k}_L+\mathbf{k}_1-\mathbf{k}_2=0, \tag{4}$$

where ω_i are the frequencies of the four interacting waves and $\mathbf{k}_i$ are their wave vectors; the subscripts 1, 2, and L label the incident waves, and the subscript S denotes the scattered wave to be detected. The efficiency of the process is determined by the nonlinear cubic susceptibility $\chi^{(3)}$ of the material.

Multistep scattering is due to the quadratic susceptibility $\chi^{(2)}$ of the material, and it can be treated as two sequential three-wave processes similar to the parametric light scattering. In this case, however, the population of all the interacting modes is high and is determined not by vacuum fluctuations, but by the intensities of all three pumping waves. In the first process, a polariton wave with the difference frequency $\omega_1-\omega_2$ and wave vector $\mathbf{k}_1-\mathbf{k}_2$ resulting from the pump waves with frequencies ω_1 and ω_2 is excited in the material. A rapidly decaying free polariton wave is also excited, its frequency ω_P and wave vector $\mathbf{k}_P$ being related through the dielectric function of the material. The temporal and spatial matching conditions in excitation of polariton waves have the form

$$\omega_P-\omega_1+\omega_2=0, \tag{5}$$

$$\boldsymbol{\tau}\equiv\ \mathbf{k}_P-\mathbf{k}_1+\mathbf{k}_2=0, \tag{6}$$

which are analogous to Eqs. (1) and (2). In the second stage of the multistep process, the test pump wave is scattered by the coherently excited polariton waves. The temporal and spatial matching conditions are similar to Eqs. (5) and (6). As in the case of spontaneous parametric scattering, the matching conditions (4) and (6) should be satisfied to within an uncertainty determined by the absorption and the length of the scattering medium. If the material has both quadratic and cubic nonlinear susceptibilities, both of these processes can occur concurrently, and the necessary condition for detecting multistep light scattering by coherently excited polaritons is that Eqs. (3)–(6) should be valid simultaneously. Multistep coherent scattering was first studied by Coffinet and DeMartini.[4] Its theory was described comprehensively in Refs. 5 and 6, and the case of Stokes scattering was analyzed in our previous publications.[7,8]

2.2. Three- and four-wave light scattering in media with periodic distribution of quadratic nonlinear susceptibility

In the case of a material with a periodic distribution of the quadratic nonlinear susceptibility $\chi^{(2)}$, the Fourier series expansion can be used:

$$\chi^{(2)}(\mathbf{r})=\sum_{-\infty}^{\infty}\chi_m^{(2)}\exp(i\mathbf{q}_m\cdot\mathbf{r}), \tag{7}$$

where the amplitudes of the spatial harmonics are

$$\chi_m^{(2)}=\frac{1}{2\pi}\int_{-d/2}^{d/2}\chi^{(2)}(\mathbf{r})\exp(-i\mathbf{q}_m\cdot\mathbf{r})d\mathbf{r}, \tag{8}$$

$\mathbf{q}_m=\mathbf{q}m$, m is an integer, $\mathbf{q}\equiv(2\pi/d)\mathbf{n}$ is the reciprocal superlattice vector, d is its period, and $\mathbf{n}$ is the unit vector normal to its layers.

2.2.1. Three-wave spontaneous parametric light scattering. Spontaneous parametric light scattering in such materials was discussed in Ref. 2. In the plane wave approximation, the polarization at the frequency of the scattered wave can be expressed as

$$P_S(\omega_S=\omega_L-\omega_P)\sim\sum_{m=-\infty}^{\infty}\chi_m^{(2)}E_LE_P^*$$
$$\times\exp[i(\mathbf{k}_L-\mathbf{k}_P+\mathbf{q}_m)\cdot\mathbf{r}], \quad (9)$$

whence follows the phase matching condition for interacting waves [compare with Eq. (2)]:

$$\mathbf{k}_L-\mathbf{k}_P+\mathbf{q}_m=\mathbf{k}_S. \quad (10)$$

In this case, the intensity distribution over frequency and angle for spontaneous parametric scattering contains additional tuning curves, whose number equals that of the nonzero spatial harmonics $\chi_m^{(2)}$. By measuring the frequencies and angles on any tuning curve with index m and taking into account condition (10), one can determine with a high accuracy the period d of the nonlinear superlattice and alignment of the vector $\mathbf{n}$. The scattered light intensity on each tuning curve of order m is, in turn, proportional to $|\chi_m^{(2)}|^2$. Thus, by measuring the relative distribution of the scattered light intensity in different orders of nonlinear diffraction, one can reconstruct the profile of the quadratic susceptibility spatial distribution in the studied crystal.

2.2.2. Four-wave coherent light scattering by polaritons. In a medium with periodic modulation of its quadratic susceptibility, nonlinear diffraction should take place in all stages of a multistep process. Unlike the spontaneous parametric scattering, the four-wave coherent scattering has not been analyzed theoretically in previous work. Here we describe a consistent calculation of the coherent cascade scattering intensity depending on the relation among wave vectors of waves interacting in a material whose quadratic susceptibility is expressed by Eq. (7). Suppose that three pump waves enter the medium, and their fields are expressed as

$$\mathbf{E}_i=\mathbf{E}_{i0}\exp(-i\omega_it+i\mathbf{k}_i\mathbf{r}),$$

where the subscript i can be 1, 2, or L. In the first stage, the interaction between $\mathbf{E}_1$ and $\mathbf{E}_2$ generates the polarization

$$\mathbf{P}^{(1)}(\mathbf{r},\omega_P=\omega_1-\omega_2)$$
$$=\frac{1}{4\pi}\sum_n\chi_n^{(2)}\exp(in\mathbf{q}\cdot\mathbf{r})\mathbf{E}_{10}\mathbf{E}_{20}^*$$
$$\times\exp[-i(\omega_1-\omega_2)t+i(\mathbf{k}_1-\mathbf{k}_2)\cdot\mathbf{r}]. \quad (11)$$

The spatial Fourier components of the polarization, $\mathbf{P}^{(1)}(\mathbf{r},\omega_P)$ and field $\mathbf{E}(\mathbf{r},\omega_P)$ generated by this polarization, $\mathbf{P}^{(1)}(\mathbf{k},\omega_P)$ and $\mathbf{E}(\mathbf{k},\omega_P)$, are related through the polariton Green's function $\mathbf{G}(\mathbf{k},\omega_P)$[1]:

$$\mathbf{E}(\mathbf{k},\omega_P)=\mathbf{G}(\mathbf{k},\omega_P)\mathbf{P}^{(1)}(\mathbf{k},\omega_P). \quad (12)$$

In this case, the Green's function is assumed to be the same as in a homogeneous medium. A general expression for $\mathbf{G}(\mathbf{k},\omega_P)$ was given in Ref. 5. The Green's function of ordinary polariton waves can be written in the form

$$G(\mathbf{k},\omega_P)=\frac{\omega_P^2}{c^2}\frac{4\pi}{k^2-K_P^2(\omega_P)}, \quad (13)$$

where $K_P(\omega_P)=k_P+i\alpha_P/2$ is the complex wave vector of the polariton state at frequency ω_P, and α_P is the linear absorption of the medium at this frequency. Hence we derive

$$\mathbf{E}(\mathbf{r},\omega_P)=\frac{1}{4\pi}\frac{\mathbf{E}_{10}\mathbf{E}_{20}^*\exp(-i\omega_Pt)}{(2\pi)^3}\int d\mathbf{k}\mathbf{G}(\mathbf{k},\omega_P)$$
$$\times\exp(i\mathbf{k}\cdot\mathbf{r})\sum_n\chi_n^{(2)}\int_{V'}d\mathbf{r}'\exp(i\Delta\mathbf{k}_n^{(1)}\mathbf{r}'), \quad (14)$$

where $\Delta\mathbf{k}_n^{(1)}=\mathbf{k}_1-\mathbf{k}_2-\mathbf{k}+\mathbf{q}n$. Here and in what follows we use the plane-wave approximation, so the amplitudes can be removed from under the integrals over the coordinates. Nonetheless, integration over the coordinates is performed over a limited crystal volume V' exposed simultaneously to the fields $\mathbf{E}_1$ and $\mathbf{E}_2$. The domain of integration over the wave vectors $\mathbf{k}$ contains all possible magnitudes and alignments of $\mathbf{k}$.

In the second stage, the test wave and the field described by Eq. (14) are mixed. As a result, the expression for the polarization oscillating at frequency $\omega_S=\omega_L-\omega_P$ takes the form

$$\mathbf{P}^{(2)}(\mathbf{r},\omega_S=\omega_L-\omega_P)$$
$$=\frac{1}{(4\pi)^2}\frac{\mathbf{E}_{L0}\mathbf{E}_{10}^*\mathbf{E}_{20}\exp(-i\omega_St+i\mathbf{k}_S\cdot\mathbf{r})}{(2\pi)^3}$$
$$\times\sum_{m,n}\chi_m^{(2)}\chi_n^{(2)*}\int d\mathbf{k}\exp(i\Delta\mathbf{k}_m^{(2)}\mathbf{r})\mathbf{G}^*(\mathbf{k},\omega_P)$$
$$\times\int_{V''}d\mathbf{r}''\exp(-i\Delta\mathbf{k}_n^{(1)}\cdot\mathbf{r}''), \quad (15)$$

where $\Delta\mathbf{k}_m^{(2)}=\mathbf{k}_L-\mathbf{k}_S-\mathbf{k}+\mathbf{q}m$. We can easily derive from Eq. (15) the scattered light intensity as a function of the detuning (i.e., the shape of the scattered light spectrum):

$$I_S\propto\left|\sum_{m,n}\chi_m^{(2)}\chi_n^{(2)*}\int d\mathbf{k}\mathbf{G}^*(\mathbf{k},\omega_P)f(\Delta\mathbf{k}_m^{(2)})f^*(\Delta\mathbf{k}_n^{(1)})\right|^2, \quad (16)$$

$$f(\boldsymbol{\xi})=\int_V d\mathbf{r}e^{i\boldsymbol{\xi}\cdot\mathbf{r}}. \quad (17)$$

Here we have assumed for simplicity that the integration domain V' defined for Eq. (14) and the region of interaction V'', which is the nonlinear crystal volume exposed simultaneously to all three incident waves, are identical and denoted by V. In the case of an unlimited interaction region V, the functions $f(\Delta\mathbf{k}_n^{(1)})$ and $f(\Delta\mathbf{k}_m^{(2)})$ are nonvanishing only if the exact conditions $\Delta\mathbf{k}_n^{(1)}=\Delta\mathbf{k}_m^{(2)}=0$ are satisfied, which, in turn imply that the matching conditions for incident waves are fulfilled:

$$\Delta\mathbf{k}_{n-m}\equiv\mathbf{k}_S-\mathbf{k}_L+\mathbf{k}_1-\mathbf{k}_2+(n-m)\mathbf{q}=0. \quad (18)$$

In the case of a limited interaction volume, condition (18) is fulfilled to within some detuning which depends on the characteristic linear dimensions of region V. As follows from Eq. (15), the scattered light intensity as a function of the detun-

ing, $I_S(\tau_1,\tau_2,\ldots)$, measured along the curve defined by the exact matching conditions $\Delta\mathbf{k}_{n-m=k}=0$ (where k is an integer) has the form

$$I_S(\boldsymbol{\tau}_1,\boldsymbol{\tau}_2,\ldots)\propto\left|\sum_n \chi_n^{(2)}\chi_{n+k}^{(2)*}\frac{\omega_P^2}{|\mathbf{k}_P-\boldsymbol{\tau}_n|^2-(k_P+i\alpha_P/2)^2}\right|^2, \tag{19}$$

$$\boldsymbol{\tau}\equiv\mathbf{k}_P-\mathbf{k}_1+\mathbf{k}_2-\mathbf{q}n. \tag{20}$$

The line shape defined by Eq. (19) is the result of interference among Lorentzians with different peak positions determined by the conditions

$$\boldsymbol{\tau}_n=0. \tag{21}$$

This means that peaks in the scattered light spectrum should be detected when conditions (18) and (21) are satisfied simultaneously. These are the matching conditions for coherent multistep four-wave scattering in a medium with the quadratic nonlinear susceptibility $\chi^{(2)}$ defined by a periodic function of coordinates.

If the medium also has a cubic susceptibility, an additional scattered component due to direct four-wave scattering should be detected. In multi-domained crystals of the type under discussion, the signs of the nonlinear susceptibility components of odd orders are the same in all domains. The spatial modulation of $\chi^{(3)}$ due to periodic modulation of the impurity concentration is several orders of magnitude smaller than that of $\chi^{(2)}$ and can be neglected in this model. In crystals shaped as plates with thickness L the total signal intensity is expressed in terms of the effective cubic susceptibility:

$$\chi_{\text{eff}}^{(3)}=\chi^{(3)}+\frac{\omega^2}{c^2}\sum_n \chi_{n+k}^{(2)}\chi_n^{(2)*}\times\frac{1}{|\mathbf{k}_1-\mathbf{k}_2-\mathbf{q}n|^2-(k_P-i\alpha_P/2)^2}, \tag{22}$$

$$I_S\propto|\chi_{\text{eff}}^{(3)}|^2\left[\frac{\sin(\Delta kL/2)}{\Delta kL/2}\right]^2. \tag{23}$$

3. EXPERIMENT

3.1. Samples

Our experiments were performed on two $LiNbO_3$:Nd:Mg crystals with a layered structure determined by the growth procedure. The first sample was a single-domain crystal in which the layered structure manifested itself only as a small periodic modulation of the refractive index ($\Delta n\approx10^{-4}$). The second sample had a 180° domain structure of the ''tail-to-tail, head-to-head'' configuration,[9] and the domains were localized on layers produced in the process of growth. A nonvanishing spontaneous polarization in each domain and periodic alternations in its direction led to periodic alternations in the quadratic susceptibility sign. In the studied samples, the angle between the normal to the domain interfaces (and growth layers) and the crystal c-axis was 57°.

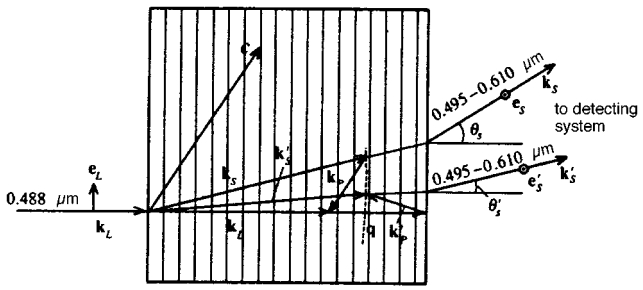


FIG. 1. Configuration of spontaneous parametric scattering and nonlinear diffraction.

3.2. Three-wave spontaneous parametric scattering

In order to record the intensity of light due to spontaneous parametric scattering in crystals with periodic modulation of optical parameters versus scattering angle and wavelength, we used an experimental setup with a photocamera similar to that described in Ref. 10. The light scattering configuration is depicted in Fig. 1. Figures 2a,b show the spectral and angular distributions of light due to spontaneous parametric scattering in the single-domain and multidomain samples, respectively. The comparison between these spectra indicates that the spectrum of the multidomain crystal contains additional components, which are defined more clearly in the region of upper polariton branch ($\omega_P\geqslant900$ cm^{-1}). From the shapes of spectral components corresponding to zero-order diffraction we derived dispersion curves of the ordinary refractive index in the region of the upper polariton branch in single-domain and multidomain samples plotted in Fig. 3. On the base of our measurements and Eqs. (1) and (3), we calculated the domain structure period $d=5.6\pm0.2$ μm. Moreover, we determined that the additional tuning curve on the spectrum of spontaneous parametric scattering in the multidomain sample in the region of the upper polariton branch corresponds to the spatial harmonic $\chi_m^{(2)}$ with $m=-1$.

Analyzing the spontaneous parametric scattering spectrum of the multidomain sample corresponding to scattering by the lower polariton branch in the vicinity of the TM fundamental mode with frequency $\omega_{TO}=580$ cm^{-1}, however, we encountered certain difficulties. As follows from our numerical calculations, the clearly defined component corresponds to two merging curves due to the spatial harmonics $\chi_m^{(2)}$ with $m=\pm1$. A weak tuning curve corresponding to $m=0$ can be seen only at large scattering angles near the phonon component. This indicates that the amplitude of the zeroth spatial harmonic of the quadratic nonlinear susceptibility is small in comparison with the amplitudes of $m=\pm1$ harmonics. Similar relations among nonlinear diffraction efficiencies of different orders for a $LiNbO_3$:Y:Mg crystal with a similar domain structure were given in Ref. 11. Further attempts to detect nonlinear diffraction of higher orders in spectra of spontaneous parametric scattering were futile: at small pumping powers ($P_L\leqslant300$ mW) no additional tuning curves were observed, apparently owing to the small amplitudes of the corresponding spatial harmonics $\chi_m^{(2)}$, and at

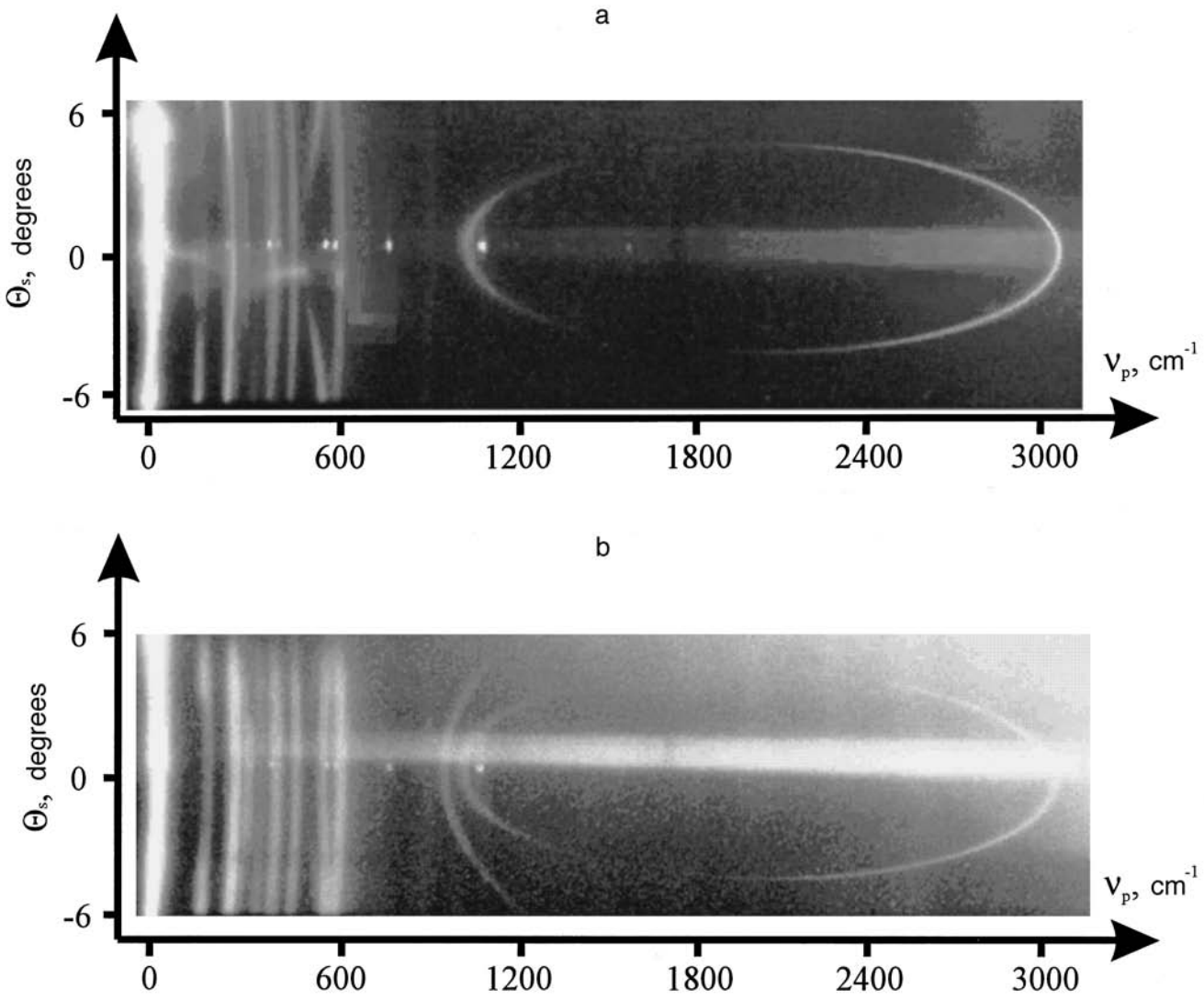


FIG. 2. Intensity of spontaneous parametric light scattering versus scattered light frequency and angle in (a) single-domain and (b) multidomain $LiNbO_3$:Nd:Mg crystals.

larger pumping powers ($P_L \geqslant 1$ W) they were not observable due to the strong effect of photoinduced light scattering.

3.3. Four-wave coherent Stokes scattering by polaritons

In recording these spectra, we used an experimental arrangement described previously.[8] The alignments of pumping light beams with respect to the crystal axes are shown in Fig. 4.

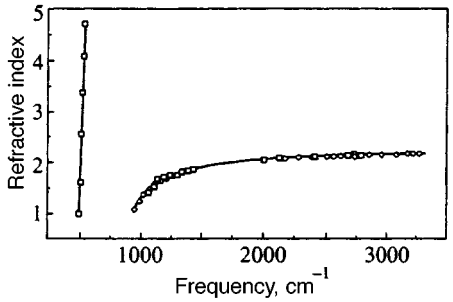


FIG. 3. Dispersion of ordinary refractive index in the polariton branch of the crystal spectrum derived from spontaneous parametric scattering spectra: (squares) single-domain crystal; (circles) multidomain crystal.

Measurements were performed as follows. The frequency difference $\omega_1 - \omega_2 = \omega_P$ was tuned to the polariton state (the wavelength of the first source was constant and equal to 1.064 μm and that of the second source was tuned between 1.08 and 1.22 μm). Then the angles between the incident light beams and crystal axis at which coherent Stokes scattering should occur were approximately calculated from Eqs. (18) and (21) using fairly accurate measurements ($\Delta n_{o,e} \approx 0.0002$) of the refractive index dispersion in the visible range and less accurate data on the dispersion of the ordinary refractive index in the polariton branch measured by means of the spontaneous parametric scattering. In addition, we used the domain superlattice constant derived from spontaneous parametric scattering spectra. The alignment was defined in terms of the angles θ_1 and θ_2 between the pumping IR beams and the probe beam at wavelength $\lambda_L = 532$ nm, and the angle α of the crystal rotation in the plane of the pump wave vectors (Fig. 4). The wave vectors of all interacting beams were in one plane containing the c-axis of the crystal. Next, a set of curves of the scattered light intensity versus the crystal rotation angle at fixed θ_2,

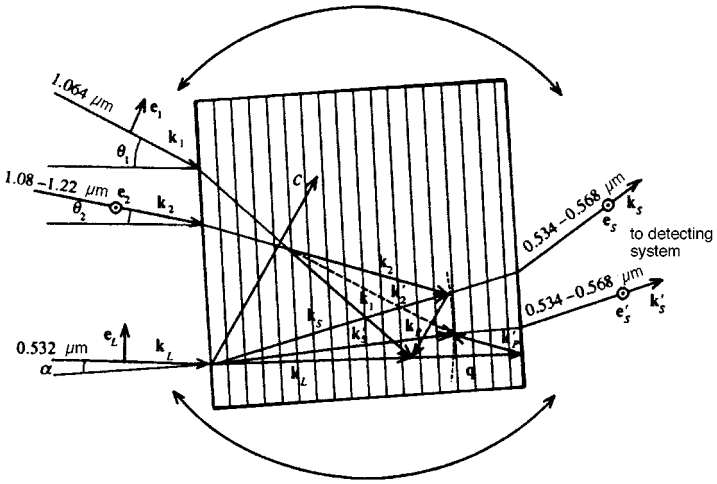


FIG. 4. Configuration of coherent Stokes light scattering by polaritons and nonlinear diffraction.

$I_S(\alpha,\theta_1=\text{const})$, were recorded, each curve of this set being recorded at different θ_1. As a result, we obtained a two-dimensional graph of $I_S(\alpha,\theta_1)$ for fixed frequency ω_P and angle θ_2. It is obvious that the peaks of $I_S(\alpha,\theta_1=\text{const})$ should lie on curves $\alpha=\alpha_{n-m}(\theta_1)$ determined by the condition $\Delta\mathbf{k}_{n-m}=0$ near their crossings with the curves of $\boldsymbol{\tau}_n=0$.

Figure 5 shows numerically calculated curves $\alpha=\alpha_0(\theta_1)$ on which the condition $\Delta\mathbf{k}_{m-n}=0$ holds for both test samples together with measurements of peak positions on the curves of the scattered light intensity $I_S(\alpha,\theta_1=\text{const})$ for frequency $\omega_P=560$ cm^{-1}. Note that, on the scale of this graph, the curves of $\alpha=\alpha_0(\theta_1)$ for the two samples coincide since the difference between the refractive indices of the single-domain and multidomain samples in the visible and near-IR ranges is small, of order 10^{-4}. The curves $\alpha=\alpha_{n-m}(\theta_1)$ for the case of $n-m\neq 0$ were not analyzed, since numerical calculations indicated that in this range the condition $\Delta\mathbf{k}_{m-n\neq 0}=0$ does not hold for all sets of the angles θ_1, θ_2, and α. Figure 5 indicates that the calculations and measurements of the angle $\alpha=\alpha_0(\theta_1)$ are in good agreement. This means that, on one hand, the accuracy of the refractive index dispersion data in the transparency range used in the calculations is fairly good, and, on the other hand, the coherent four-wave interaction model is adequate. The angular widths of the curves $I_S(\alpha,\theta_1=\text{const})$ were $\Delta\alpha=1-4°$ and were determined, in accordance with Eq. (23), by the linear dimensions of the interaction region in the crystal, $L\leqslant 1$ mm.

Equations (22) and (23) imply that the contribution of the direct four-wave scattering due to the cubic susceptibility $\chi^{(3)}$ should be detected along the entire curve $\alpha=\alpha_{n-m\neq 0}(\theta_1)$. Nonetheless, in our experiments the signal was detected only on limited sections of the curve $\alpha=\alpha_{n-m\neq 0}(\theta_1)$, where additional spatial phase matching conditions for multistep scattering are satisfied. This indicates that the background due to direct scattering is small, so $\chi^{(3)}$ is relatively low.

Basic information on the dispersion of the polariton state and modulation of $\chi^{(2)}$ in a crystal can be derived from the curve $I_{\max}(\theta_1)\equiv I(\alpha_0(\theta_1),\theta_1)$. This curve shows the change in I_S on the surface defined by the condition $\Delta\mathbf{k}=0$ and is plotted using measurements of I_S at the peak of each one-parameter curve $I_S(\alpha,\theta_1=\text{const})$ for different θ_1. The curve connecting these peaks yields such parameters of the medium as the refractive index and effective absorption at the polariton frequency, superlattice constant and the ratio of nonlinear diffraction efficiencies of different orders. Figures 6 and 7a yield curves $I_{\max}(\theta_1)$ recorded at one frequency and

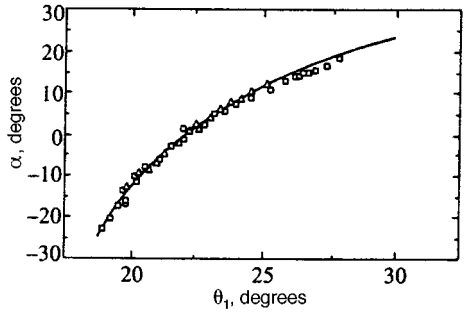


FIG. 5. Curves of $\alpha=\alpha_0(\theta_1)$ at frequency $\omega_P=560$ cm^{-1} and positions of peaks on curve of scattered light intensity $I_S(\alpha,\theta_1=\text{const})$ for (triangles) single-domain and (squares) multiply-domained crystals.

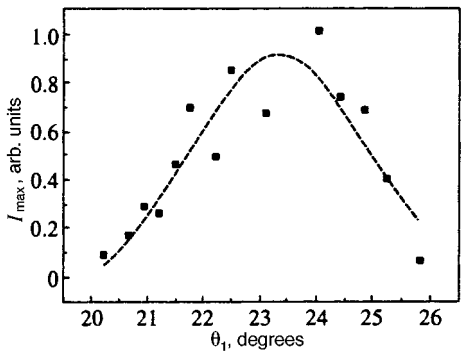


FIG. 6. Curve of $I_{\max}(\theta_1)$ for the single-domain crystal at frequency $\omega_P=560$ cm^{-1} and $\theta_2=0°$: (black squares) experimental data; the dashed line shows calculations.

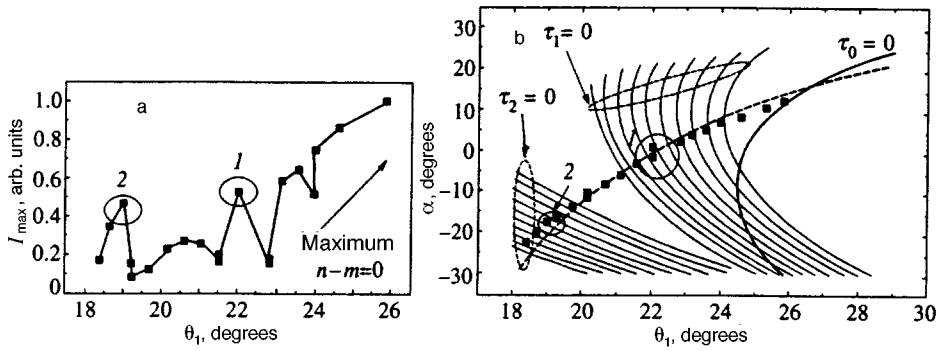


FIG. 7. Angular characteristics of coherent Stokes scattering in the multidomain crystal at frequency $\omega_P = 560$ cm^{-1} and $\theta_2 = 0°$: (a) $I_{max}(\theta_1)$ measured in experiments. Peaks due to the quasi-matching interaction are encircled: *1*—$m-n=0$, $n=1$; *2*—$m-n=0$, $n=2$. (b) Measurements (black squares) of peak positions on curves of the scattered light intensity $I_S(\alpha, \theta_1 = \text{const})$, and (dashed lines) calculated curve of $\alpha = \alpha_0(\theta_1)$, and curves of intensity maxima for the first stage of multistep scattering corresponding to $\tau_n = 0$ ($n = 0,1,2$). The curves of $\tau_n = 0$ for $n = 1,2$ are calculated for several polariton state frequencies which can be excited simultaneously owing to the finite widths of the pump laser lines. Labels *1* and *2* denote positions of peaks on the experimental curve $I_{max}(\theta_1)$ (Fig. 7a).

angle $\theta_2 = 0$ for both tested samples. It is obvious that the peak positions on these curves are defined by the exact matching conditions $\tau_n = 0$ at frequency ω_P. If the polariton frequency and the difference wave vector $\mathbf{k}_1 - \mathbf{k}_2 - \mathbf{q}n = \mathbf{k}_P$ are given, this allows one to calculate the refractive index at the polariton frequency. In this case, the polariton refractive index uncertainty can be about 1%, which is an order of magnitude better than the accuracy of the spontaneous parametric scattering technique in this frequency band. We have measured the refractive index of the tested samples at two frequencies of this band. At $\omega_P = 533$ cm^{-1} we have obtained $n_P = 3.75 \pm 0.05$ (compare to the value derived from the spectra of spontaneous parametric scattering, $n_P = 4.3 \pm 0.6$), and at $\omega_P = 560$ cm^{-1} we have $n_P = 7.04 \pm 0.08$. Note that in the latter case the polariton refractive index could not be derived from three-wave scattering measurements, since in this frequency band the spectral line of the spontaneous parametric scattering strongly overlaps with the Raman line ($\omega_{TO} = 580$ cm^{-1}, $\Delta\omega \approx 20$ cm^{-1}).

In the case of a homogeneous medium, the curve $I_{max}(\theta_1)$ has only one peak, and in the case of a sample with a periodically modulated nonlinear susceptibility it can have additional peaks due to nonlinear diffraction. In principle, nonlinear diffraction can also affect the curves $I_S(\alpha, \theta_1 = \text{const})$. This is possible if condition (17) is satisfied for different $m-n$ at equal θ_1 and θ_2, which was never the case in our experiments. On the curve of $I_{max}(\theta_1)$ plotted in Fig. 7a, the nonlinear diffraction manifests itself in the form of additional peaks (encircled by solid lines). The peak marked by *2* corresponds to $m-n=0$ and $n=2$, and that marked by *1* to $m-n=0$ and $n=1$. The graph shows only one slope of the scattered light intensity peak due to scattering not affected by nonlinear diffraction, since the intensity measurements at larger θ_1 were very inaccurate owing to vignetting by apertures of the light detecting system. For the same reason, we could not adequately interpolate the experimental curve of $I_{max}(\theta_1)$ using a formula like Eq. (19). We derived the domain superlattice constant $d = 5.3 \pm 0.1$ μm from the positions of peaks in the scattered light intensity. Given this value and the refined value of the polariton refractive index, we plotted curves $\tau_n = 0$ and $\alpha = \alpha_0(\theta_1)$ shown in Fig. 7b. The curves $\tau_n = 0$ for $n = 0,1,2$ have the shapes of broad bands, whose widths are determined by the spectrum of polariton states excited simultaneously by pumping laser lines of finite widths.

Similar measurements were performed at the other frequency $\omega_P = 533$ cm^{-1}. Measured and numerically calculated phase matching curves are plotted in Fig. 8a,b. In this case, the curve $I_{max}(\theta_1)$ has two clearly defined peaks of coherent Stokes scattering due to nonlinear diffraction corresponding to $n-m=0$ (marked by *0*) and $n=1$ (labeled by *1*). Figure 8a shows that peak *1* is much stronger than peak *0*, which is in qualitative agreement with the ratio between the respective Fourier components of the quadratic susceptibility in multiply-domained $LiNbO_3$:Y:Mg crystals.[11] The domain superlattice constant is $d = 5.34 \pm 0.05$ μm, which is in good agreement with the value derived from the angular dependence at $\omega_P = 560$ cm^{-1}. In calculating the line shapes using Eq. (22), we took into account angular divergences of pump laser beams and their finite spectral widths. It is noteworthy that the approximation of the experimental curve $I_{max}(\theta_1)$ by Eq. (22) is satisfactory only if the linear absorption at the polariton frequency is assumed to be smaller than that derived from spontaneous parametric scattering spectra[12] and calculations based on the oscillator model.[13] This discrepancy was noted previously[7,8] in studies of homogeneous crystals, but its nature has remained unclear. The small difference between the domain superlattice constants derived from measurements of the spontaneous parametric scattering and coherent Stokes scattering can be ascribed to the inhomogeneous distribution of this parameter over the crystal volume.

4. CONCLUSIONS

In this work we have investigated for the first time the nonlinear diffraction in coherent four-wave light scattering by polaritons in materials with periodically modulated quadratic susceptibility. We have found phase matching condi-

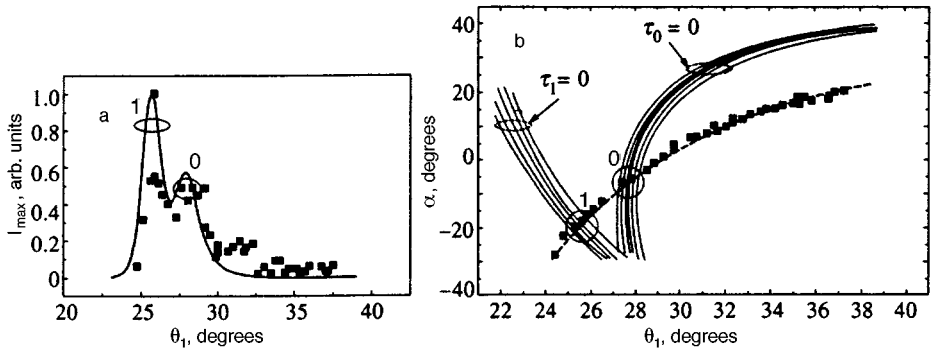


FIG. 8. Angular characteristics of coherent Stokes scattering in the multidomain crystal at frequency $\omega_P = 533$ cm^{-1} and $\theta_2 = 16°$: (a) $I_{max}(\theta_1)$ (black squares) measured in experiments; the solid line shows calculations by Eq. (22) taking into account the angular and spectral distributions of pumping beams and linear absorption $\alpha_P = 500$ cm^{-1} at the polariton frequency. Peaks due to the quasi-matching interaction are encircled: *0*—$m-n=0$, $n=0$; *1*—$m-n=0$, $n=1$. (b) Measurements (filled squares) of peak positions on curves of scattered light intensity $I_S(\alpha, \theta_1 = \text{const})$, (dashed line), calculated curve of $\alpha = \alpha_0(\theta_1)$, and curves of intensity maxima for the first stage of cascade scattering corresponding to $\tau_n = 0$ ($n = 0,1$). The curves of $\tau_n = 0$ for $n = 0,1$ are calculated for several polariton frequencies which can be excited simultaneously owing to finite widths of pumping laser lines. The labels *0* and *1* denote positions of peaks on the experimental curve $I_{max}(\theta_1)$ (Fig. 8a).

tions for this process and an expression for the scattering line shape. The predicted effect has been detected in experiments, and our measurements are in satisfactory agreement with both theoretical predictions and data on the three-wave parametric scattering in the materials studied. We have compared the spectroscopic possibilities of the coherent Stokes scattering technique with those of spontaneous parametric scattering. From our analysis, we have drawn the following conclusions.

1. Cascade diffraction processes are possible in coherent four-wave light scattering by polaritons in a noncentrosymmetric media and with periodically modulated quadratic susceptibility.

2. Coherent Stokes scattering allows one to detect light scattering in higher orders of nonlinear diffraction than the spontaneous parametric scattering, so one should be able to reconstruct, in principle, the profile of a nonlinear diffraction grating with higher accuracy.

3. Active spectroscopy allows one to measure the domain superlattice period and refractive index at the polariton frequency with a higher accuracy than the three-wave spontaneous parametric scattering spectroscopy. Measurements of the linear absorption in the polariton branch derived from coherent Stokes spectra are, however, different from data obtained by other techniques, and the nature of this discrepancy has remained unclear.

4. Intensities of CSRS signals from relatively small regions in crystals are notably higher than in the case of spontaneous parametric scattering. Therefore the four-wave polariton scattering can be used in studies of the spatial distribution of domain system parameters.

Hence it follows that, in coherent polariton light scattering by polaritons, nonlinear diffraction allows one to obtain more accurate measurements of optical materials, but the cost of such an improvement is a more sophisticated experimental facility and a more laborious data-processing technique.

It is a pleasure to express our gratitude to S. P. Kulik, M. V. Chekhova, and A. V. Burlakov for fruitful discussions and information about the test samples. We also thank I. I. Naumova for kindly supplying the $LiNbO_3$:Nd:Mg samples. The work was supported by the Russian Fund for Fundamental Research (Grant No. 96-02-16336a).

[1] D. N. Klyshko, *Photons and Nonlinear Optics* Gordon and Breech, New York (1988).
[2] A. L. Aleksandrovskii, G. Kh. Kitaeva, S. P. Kulik, and A. N. Penin, Zh. Éksp. Teor. Fiz. **90**, 1051 (1986) [Sov. Phys. JETP **63**, 613 (1986)].
[3] S. A. Akhmanov and N. N. Koroteev, *Methods of Nonlinear Optics in Scattered Light Spectroscopy* [in Russian], Nauka, Moscow (1981).
[4] J. P. Coffinet and F. DeMartini, Phys. Rev. Lett. **60**, 22 (1969).
[5] D. N. Klyshko, Kvantovaya Élektronika **2**, 265 (1975) [Sov. J. Quantum Electron. **5**, 149 (1975)].
[6] V. L. Strizhevskii and Yu. N. Yashkir, Kvantovaya Élektronika **2**, 995 (1975) [Sov. J. Quantum Electron. **5**, 541 (1975)].
[7] G. Kh. Kitaeva, P. S. Losevsky, A. A. Mikhailovsky, and A. N. Penin, Opt. Commun. **138**, 242 (1997).
[8] G. Kh. Kitaeva, P. S. Losevskii, A. A. Mikhailovskii, and A. N. Penin, Zh. Éksp. Teor. Fiz. **112**, 441 (1997) [JETP **85**, 241 (1997)].
[9] I. I. Naumova, N. F. Evlanova, O. A. Gliko, and S. V. Lavrishev, Ferroelectrics **190**, 107 (1997).
[10] M. V. Chekhova, S. P. Kulik, and A. N. Penin, Opt. Commun. **114**, 301 (1995).
[11] A. L. Aleksandrovskii, O. A. Gliko, I. I. Naumova, and V. I. Pryalkin, Kvantovaya Élektronika **23**, 657 (1996).
[12] A. V. Burlakov, M. V. Chekhova, S. P. Kulik, and A. N. Penin, in Proceedings Int. Conf. CLEO/Pacific Rim'97, Japan (1997) (to be published).
[13] A. S. Barker and R. Loudon, Phys. Rev. **158**, 433 (1967).

Translation provided by the Russian Editorial office.

Propagation of a beam of gamma rays in the field of a monochromatic laser wave

V. A. Maisheev*)

Institute of High-Energy Physics, 142284 Protvino, Moscow Region, Russia
(Submitted 28 April 1997)
Zh. Éksp. Teor. Fiz. **112**, 2016–2029 (December 1997)

The head-on propagation of a beam of γ rays through the field of a laser wave is investigated. The optical properties of the laser wave (as a medium) are described by the dielectric tensor. The refractive indices are determined, and the polarization characteristics of electromagnetic normal modes capable of propagating in such a medium are investigated. Relations are derived to describe the variation of the initial polarization and intensity of a γ-ray beam as it propagates in a laser field. The influence of laser intensity on the investigated process is discussed. © *1997 American Institute of Physics.* [S1063-7761(97)00812-3]

1. INTRODUCTION

Polarization phenomena[1,2] accompanying the transmission of visible light through anisotropic or gyrotropic media have been known for some time. Analogous phenomena are predicted by theory[3] for γ rays with energies $\geqslant 1$ GeV propagating in single crystals, which are anisotropic media by their very nature. The main process by which γ rays are absorbed in single crystals is the formation of electron–positron pairs. The cross section of the process depends on the direction of linear polarization of the γ rays relative to the crystallographic planes. In general, as a result of interaction with the electric field of the single crystal, a monochromatic, linearly polarized beam of γ rays comprises two electromagnetic waves with different refractive indices, so that linear polarization is transformed into circular polarization or vice versa.

On the other hand, it has been noted[4] that pair production in single crystals is analogous to the process in the field of a linearly polarized electromagnetic wave. The possibility of birefringence of high-energy γ rays moving in a linearly polarized laser wave has been suggested on the basis of this analogy in a study[5] of problems regarding the propagation of γ rays in single crystals, but no actual estimates of the effect are given.

A recent paper[6] shows for the first time that for existing laser parameters, the phenomena of birefringence and rotation of the polarization plane of γ rays with energies in the tens of GeV or higher are quite conspicuous even for comparatively short (a few picoseconds) laser bunches with parameters within the realm of present-day engineering capabilities. The variations of the Stokes parameters and the intensity of γ rays are calculated in the same paper for transmission of the latter through a laser bunch with arbitrary polarization. The known amplitudes for the scattering of light by light[7,8] are used in the calculations.

Here we further investigate the propagation of high-energy γ rays through a laser wave by a more traditional approach[1,2] based on the crystal optics of anisotropic and gyrotropic media. In this approach the anisotropic (gyrotropic) medium is described by the dielectric tensor. Knowing this tensor, we can derive a dispersion relation for the propagation of electromagnetic waves in the given medium and calculate their refractive indices. It should be noted that to calculate the polarization transformation of γ rays moving in the laser field, it is sufficient to know only the difference between the indices, in contrast with processes such as the Čerenkov radiation of charged particles, whose description requires the absolute values of the refractive indices.

The use of the dielectric tensor to describe the properties of the electromagnetic vacuum is a common practice; in Ref. 9, for example, explicit equations are given for the components of this tensor in the case of slowly varying fields and low photon frequencies.

2. DIELECTRIC TENSOR OF A LASER WAVE

We can write the equations for the electromagnetic field in the medium in the form[1,2]

$$\text{curl}\ \mathbf{B}=\frac{1}{c}\frac{\partial \mathbf{D}}{\partial t},\quad \text{div}\ \mathbf{D}=0,$$

$$\text{curl}\ \mathbf{E}=-\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t},\quad \text{div}\ \mathbf{B}=0, \tag{1}$$

where $\mathbf{E}$ is the electric field, $\mathbf{D}$ is the electric displacement, $\mathbf{B}$ is the magnetic induction, t is the time, and c is the speed of light. The properties of the medium (which in our case is the monochromatic electromagnetic field of the laser wave) are manifest in the relationships among the vectors $\mathbf{B}$, $\mathbf{E}$, and $\mathbf{D}$. Choosing to write the equations in this form makes it possible to disregard the magnetic field strength.[1,2] We write the relationship between $\mathbf{D}$ and $\mathbf{E}$,

$$D_i(\omega)=\varepsilon_{ij}E_j(\omega),\quad i,j=1,2,3, \tag{2}$$

where $\varepsilon_{ij}=\varepsilon'_{ij}+i\varepsilon''_{ij}$ is the complex dielectric tensor, and ω is the frequency of the γ ray propagating in the laser wave.

The average energy losses per unit time and per unit volume V of a monochromatic electromagnetic wave (beam of high-energy γ rays) $\mathbf{E}_0\ \exp[i(\mathbf{k}\cdot\mathbf{r}-\omega t)]$ with wave vector $\mathbf{k}$ are[1,2]

$$\tilde{q}=\frac{1}{4\pi V}\int_V \mathbf{E}\cdot\frac{\partial \mathbf{D}}{\partial t}dV=\frac{i\omega}{16\pi}(\varepsilon^*_{ij}-\varepsilon_{ji})E^*_j E_i\,. \tag{3}$$

 1063-7761/97/121102-08$10.00

The energy loss mechanism is the production of electron–positron pairs in the wave field.[10] The process is governed mainly by the transverse part of the dielectric tensor, as its longitudinal components are of higher order in the small interaction constant α (Refs. 4 and 11). Bearing this in mind, we can write Eq. (3) as follows in a coordinate system with one axis directed along the γ-ray wave vector:

$$\tilde{q}=\frac{i\omega J}{4}\{(\varepsilon_{11}^{*}-\varepsilon_{11})(1+\xi_3)+(\varepsilon_{12}^{*}-\varepsilon_{21})(\xi_1-i\xi_2)$$
$$+(\varepsilon_{21}^{*}-\varepsilon_{12})(\xi_1+i\xi_2)+(\varepsilon_{22}^{*}-\varepsilon_{22})(1-\xi_3)\}, \quad (4)$$

where $J=(E_1E_1^{*}+E_2E_2^{*})/8\pi$, and ξ_i denotes the Stokes parameters of the beam of γ rays.

On the other hand, knowing the pair-production cross section $\sigma_{\gamma\gamma}$, we can write

$$\tilde{q}=2n_l\sigma_{\gamma\gamma}(cn_\gamma E_\gamma)=2n_l\sigma_{\gamma\gamma}cJ, \quad (5)$$

where cn_γ is the flux density of γ rays with energy E_γ, and n_l is the number of photons per unit volume of the laser wave. The factor 2 in the equation accounts for the opposing motion of γ photons and laser photons. In our coordinate system with one axis along the γ-ray wave vector, the other two axes are parallel and perpendicular to the linear polarization of the laser wave (if it is zero, the choice of these axes is arbitrary). In this coordinate system pair production is then described by the relation[6]

$$\sigma_{\gamma\gamma}(z)=\sigma_0(z)+\sigma_c(z)\xi_2P_c+\sigma_l(z)\xi_3P_l, \quad 0<z\leqslant 1, \quad (6)$$

where the invariant $z=m^2c^4/E_\gamma E_l$, m is the electron mass, E_l is the laser photon energy, and P_c and P_l are the circular and linear polarizations of the laser photon. Pair production is a threshold process, so that for $E_\gamma E_l\leqslant m^2c^4$ or for $z>1$, the laser wave represents a transparent medium to a beam of γ rays. The functions σ_0, σ_c, and σ_l can be found in Refs. 4, 6, 7, and 10, and enter into Eqs. (17)–(19), (23)–(25) through relations (9)–(11) (see below). Those components of the dielectric tensor associated with the absorption of γ rays can be found by comparing relations (4) and (5).

We use the following dispersion relations in determining the dielectric tensor:[2]

$$\varepsilon_{ij}'-\delta_{ij}=\frac{2}{\pi}P\int_0^\infty\frac{x\varepsilon_{ij}''(x)dx}{x^2-\omega^2}, \quad (7)$$

$$\varepsilon_{ij}''=-\frac{2\omega}{\pi}P\int_0^\infty\frac{(\varepsilon_{ij}'-\delta_{ij})dx}{x^2-\omega^2}, \quad (8)$$

where δ_{ij} is the Kronecker delta. Equating (4) and (5), we obtain

$$\varepsilon_{11}''+\varepsilon_{22}''=4n_lc\sigma_0/\omega, \quad (9)$$

$$\varepsilon_{11}''-\varepsilon_{22}''=4n_lc\sigma_lP_l/\omega, \quad (10)$$

$$\varepsilon_{12}'-\varepsilon_{21}'=4n_lc\sigma_cP_c/\omega. \quad (11)$$

It is readily apparent that $\varepsilon_{12}+\varepsilon_{21}=0$. This result follows from the invariance of the tensor ε_{ij} under rotation of the coordinate system about the wave vector of a circularly polarized laser wave ($P_c\neq 0$, $P_l=0$). The same result can be deduced from the theory of generalized susceptibilities[12] for quantities that change sign upon time reversal, where time corresponds to photon angular momentum in our case. Consequently, the subset of dielectric tensor components associated with absorption is given by

$$\varepsilon_{11}''=2n_lc(\sigma_0+P_l\sigma_l)/\omega, \quad (12)$$

$$\varepsilon_{22}''=2n_lc(\sigma_0-P_l\sigma_l)/\omega, \quad (13)$$

$$\varepsilon_{12}'=-\varepsilon_{21}'=2n_lcP_c\sigma_c/\omega, \quad (14)$$

We can use Eqs. (7) and (8) to find the remaining components of the tensor. The results of calculations of the components of the tensor ε_{ij} are given below for a coordinate system with one axis along the γ-ray wave vector and with the other two axes parallel and perpendicular to the linear polarization of the laser wave P_l. The integrals (7) and (8) reduce to integrals that are evaluated in a recent preprint[5] by the theory of residues:

$$\varepsilon_{11}'-\varepsilon_{22}'=\frac{\alpha}{2\pi}\frac{\langle E^2\rangle}{E_{\rm cr}^2}P_lz^2F_1'(z), \quad (15)$$

$$\varepsilon_{11}'+\varepsilon_{22}'=2+\frac{2\alpha}{\pi}\frac{\langle E^2\rangle}{E_{\rm cr}^2}z^2F_2'(z,1), \quad (16)$$

$$\varepsilon_{12}'=-\varepsilon_{21}'=\frac{\alpha\langle E^2\rangle}{2E_{\rm cr}^2}P_cz^2F_c'(z), \quad (17)$$

$$\varepsilon_{11}''-\varepsilon_{22}''=-\frac{\alpha}{4}\frac{\langle E^2\rangle}{4E_{\rm cr}^2}P_lF_1''(z), \quad (18)$$

$$\varepsilon_{11}''+\varepsilon_{22}''=\frac{\alpha\langle E^2\rangle}{E_{\rm cr}^2}F_2''(z,1), \quad (19)$$

$$\varepsilon_{12}''=-\varepsilon_{21}''=\frac{\alpha\langle E^2\rangle}{\pi E_{\rm cr}^2}P_cF_c''(z), \quad (20)$$

where $\langle E^2\rangle=4\pi n_lE_l$ is the rms electric field of the laser wave, $E_{\rm cr}=m^2c^3/e\hbar$ is a critical field, and the functions F_1', F_2', F_c', F_1'', F_2'', and F_c'' have the form

$$F_1'(z)=\begin{cases}\left[\sqrt{1-z}+\dfrac{z}{2}L_-\right]^2+\left[\sqrt{1+z}-\dfrac{z}{2}L_+\right]^2-\dfrac{\pi^2z^2}{4}, \\ 0<z\leqslant 1, \\ -[\sqrt{z-1}-z\ \arctan\sqrt{z-1}]^2+\left[\sqrt{1+z}-\dfrac{z}{2}L_+\right]^2, \\ z>1,\end{cases} \tag{21}$$

$$F_2'(z,\mu)=\begin{cases}-2-\mu-\left[1+\mu\left(z-\dfrac{z^2}{2}\right)\right]\dfrac{1}{4}L_-^2 \\ \quad-\left[1-\mu\left(z+\dfrac{z^2}{2}\right)\right] \\ \quad\times\dfrac{1}{4}L_+^2+\dfrac{(1+\mu z)\sqrt{1-z}}{2}L_- \\ \quad-\dfrac{(\mu z-1)\sqrt{z+1}}{2}L_++\dfrac{\pi^2}{4}\left[1+\mu\left(z-\dfrac{z^2}{2}\right)\right], \\ 0<z\leqslant 1, \\ -2-\mu^+\left[1+\mu\left(z-\dfrac{z^2}{2}\right)\right] \\ \quad\times\arctan^2(\sqrt{z-1})-\left[1-\mu\left(z+\dfrac{z^2}{2}\right)\right]\dfrac{1}{4}L_+^2 \\ \quad+(1+\mu z)\sqrt{z-1}\arctan\sqrt{z-1} \\ \quad-\dfrac{(\mu z-1)\sqrt{1+z}}{2}L_+, \\ z>1,\end{cases} \tag{22}$$

$$F_c'(z)=\begin{cases}3\sqrt{1-z}-L_-, & 0<z\leqslant 1,\\ 0, & z>1,\end{cases} \tag{23}$$

$$F_1''(z)=\begin{cases}z^4\left(L_-+\dfrac{2\sqrt{1-z}}{z}\right), & 0<z\leqslant 1,\\ 0, & z>1,\end{cases} \tag{24}$$

$$F_2''(z,\mu)=\begin{cases}z^2\left\{\left[1+\mu\left(z-\dfrac{z^2}{2}\right)\right]L_--\sqrt{1-z}(1+\mu z)\right\}, \\ 0<z\leqslant 1, \\ 0, \quad z>1,\end{cases} \tag{25}$$

$$F_c''(z)=\begin{cases}z^2\left(\dfrac{3}{2}\sqrt{1-z}L_--\dfrac{3}{2}\sqrt{1+z}L_+-\dfrac{1}{4}L_-^2\right. \\ \quad\left.+\dfrac{1}{4}L_+^2+\dfrac{\pi^2}{4}\right), \\ 0<z\leqslant 1, \\ z^2\left[3\sqrt{z-1}\arctan\sqrt{z-1}-\dfrac{3}{2}\sqrt{1+z}L_+\right. \\ \quad\left.+\arctan^2(\sqrt{z-1})+\dfrac{1}{4}L_+^2\right], \\ z>1.\end{cases} \tag{26}$$

The functions L_- and L_+ are given by

$$L_-=\ln\frac{1+\sqrt{1-z}}{1-\sqrt{1-z}}, \quad L_+=\ln\frac{\sqrt{1+z}+1}{\sqrt{1+z}-1}.$$

These data completely determine the dielectric tensor for γ rays propagating directly into an oncoming monochromatic laser wave. In a number of problems it is more convenient to use the inverse η_{ij} of the tensor ε_{ij}. When $|\varepsilon_{ij}-\delta_{ij}|\ll 1$, these two tensors are related by

$$\eta_{ij}+\varepsilon_{ij}=2\,\delta_{ij}. \tag{27}$$

Figure 1 shows the components of the dielectric tensor as functions of the invariant z.

3. REFRACTIVE INDICES OF GAMMA RAYS IN A WAVE FIELD

A fundamental problem in the optics of anisotropic and gyrotropic media is the propagation of monochromatic plane waves characterized by definite values of the frequency ω and the wave vector $\mathbf{k}$ in such media. These waves, which satisfy the homogeneous wave equation, are called electromagnetic normal modes,[2] and have the form

$$\mathbf{E}=\mathbf{E}_0\ \exp[i(\mathbf{k}\cdot\mathbf{r}-\omega t)], \quad \mathbf{k}=\omega\tilde{n}\mathbf{s},$$

where $\mathbf{E}_0$ is a complex vector that does not depend on the coordinates r or the time t, $\tilde{n}$ is the complex refractive index, and $\mathbf{s}=\mathbf{k}/|\mathbf{k}|$ is the unit real vector. The vectors $\mathbf{D}$ and $\mathbf{B}$ have the same form.

From Maxwell's equations (1), taking into account the relation between $\mathbf{D}$ and $\mathbf{E}$ in a coordinate system with the x axis along the wave vector, we obtain the wave equation[1,2]

$$\eta_{11}\frac{\partial^2D_1}{\partial x^2}+\eta_{12}\frac{\partial^2D_2}{\partial x^2}-\frac{1}{c^2}\frac{\partial^2D_1}{\partial t^2}=0,$$

$$\eta_{21}\frac{\partial^2D_1}{\partial x^2}+\eta_{22}\frac{\partial^2D_2}{\partial x^2}-\frac{1}{c^2}\frac{\partial^2D_2}{\partial t^2}=0. \tag{28}$$

For a monochromatic plane wave it follows from these equations that

$$(\tilde{n}^{-2}\delta_{ij}-\eta_{ij})D_j=0, \quad i,j=1,2. \tag{29}$$

From the compatibility condition for the two homogeneous equations we find the refractive index for γ rays,

$$\tilde{n}^{-2}=\frac{S}{2}\pm\sqrt{\frac{S^2}{4}-D_\eta}$$

$$=\frac{\eta_{11}+\eta_{22}}{2}\pm\sqrt{\frac{(\eta_{11}-\eta_{22})^2}{4}+\eta_{12}\eta_{21}}, \tag{30}$$

where S and D_η are the trace and determinant of the matrix η_{ij}.

Consequently, a beam of γ rays propagates in the field of a laser wave as a superposition of two electromagnetic waves, which have different refractive indices in general. We note that in our situation the two roots of Eq. (30), with the form $(-1+\text{small quantity})$, are linear and correspond to motion in opposite directions.

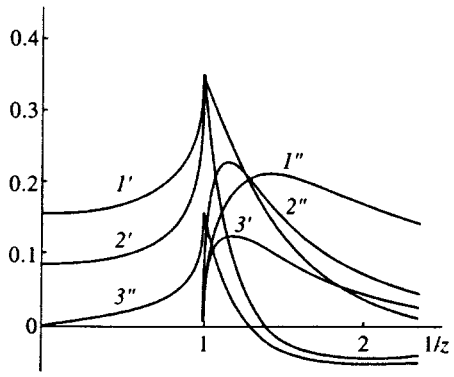


FIG. 1. Components of the dielectric tensor versus the invariant z. $1',1''$) $k(\varepsilon'_{11}+\varepsilon'_{22}-2)/2$, $k(\varepsilon''_{11}+\overline{\varepsilon''_{22}})/2$; $2',2''$) $k(\varepsilon'_{22}-\varepsilon'_{11})/P_l$, $k(\varepsilon''_{22}+\varepsilon''_{11})/P_l$; $3',3''$) $k\varepsilon'_{12}/P_c$, $k\varepsilon''_{21}/P_c$. The components $\varepsilon'_{21}=-\varepsilon'_{12}$, $\varepsilon''_{12}=-\varepsilon''_{21}$, and the factor $k^{-1}=\alpha\langle E^2\rangle/E_{\rm cr}^2$.

The refractive indices are complex quantities in the general case. For $z\geqslant 1$, however, it is evident at once from Eq. (30) that they are real, i.e., the laser wave is a transparent medium in this case. If the laser wave is linearly ($P_c=0$) or circularly ($P_l=0$) polarized, the refractive indices are easily determined, and are equal to $n_1^2=\varepsilon_{11}$ and $n_2^2=\varepsilon_{22}$ in the first case and to $n_{1,2}^2=(\varepsilon_{11}+\varepsilon_{22})/2\pm i\varepsilon_{12}$ in the second case, where the components of the dielectric tensor are written in the previously specified coordinate system, and the relation $\varepsilon_{12}=-\varepsilon_{21}$ is taken into account.

In the limit $z\to\infty$ the refractive indices for linearly and circularly polarized laser waves are respectively

$$n_\perp, n_\parallel = 1+\frac{\alpha}{\pi}\frac{\langle E^2\rangle}{E_0^2}\frac{11\pm 3P_l}{45}, \tag{31}$$

$$n_\Rightarrow,\ n_\Leftrightarrow = 1+\frac{\alpha}{\pi}\frac{\langle E^2\rangle}{E_0^2}\left(\frac{11}{45}\pm\frac{16P_c}{315z}\right). \tag{32}$$

Figures 2 and 3 shows the behavior of these indices as functions of the variable z.

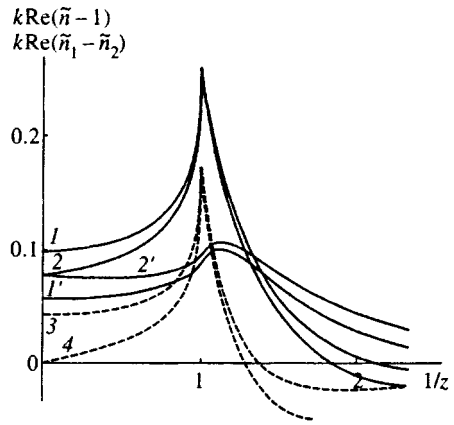


FIG. 2. Real parts of the refractive indices for linearly ($P_l=1$, $P_c=0$,) curves (1 and $1'$) and circularly ($P_l=0$, $P_c=1$, curves 2 and $2'$) polarized laser waves, and the corresponding differences (curves 3 and 4) of these indices versus the invariant z. The factor $k^{-1}=\alpha\langle E^2\rangle/E_{\rm cr}^2$.

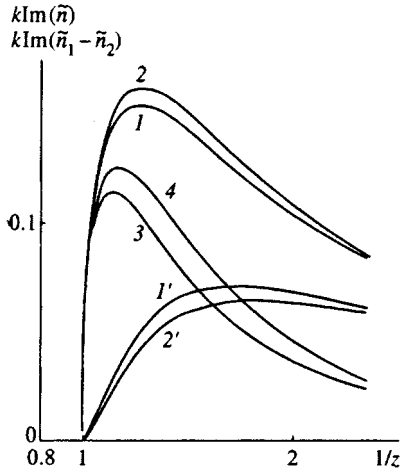


FIG. 3. Imaginary parts of the refractive indices for linearly ($P_l=1$, $P_c=0$,) curves (1 and $1'$) and circularly ($P_l=0$, $P_c=1$,) curves (2 and $2'$) polarized laser waves, and the corresponding differences curves (3 and 4) of these indices versus the invariant z. The factor $k^{-1}=\alpha\langle E^2\rangle/E_{\rm cr}^2$.

4. POLARIZATION CHARACTERISTICS OF GAMMA RAYS PROPAGATING IN A LASER WAVE

From the dispersion relations (29) we find a relation for the components of the vector **D**:

$$\frac{D_1}{D_2}=\kappa=\frac{\tilde{n}^{-2}-\eta_{22}}{\eta_{21}}=\frac{|D_1|}{|D_2|}e^{i\delta}, \tag{33}$$

where δ is the phase shift between D_1 and D_2. Rotating the coordinate system about the γ-ray wave vector, we can reduce this relation to the form $\kappa=0$ or to the form $\kappa=i\rho$ (since $|D_1||D_2|\sin\delta=b_1b_2$, where b_1 and b_2 are the lengths of the semiaxes of an ellipse, and $|\rho|=b_1/b_2$; Ref. 13). The first case corresponds to the propagation of a linearly polarized wave, and the second case to an elliptically polarized wave, where $\rho>0$ ($\rho<0$) corresponds to left (right) polarization of the γ ray.

We first consider the case $z>1$, in which γ rays are not absorbed. If $\eta''_{12}\neq 0$, we have

$$\kappa=i\rho=i\frac{(\eta'_{11}-\eta'_{22})/2\pm\sqrt{(\eta'_{11}-\eta'_{22})^2/4+\eta''^2_{12}}}{\eta''_{12}}. \tag{34}$$

Clearly, normal modes are elliptically polarized in general, where the principal axes of the ellipse are parallel to the axes of the adopted coordinate system. The circular polarization $P_{\rm circ}$ and the linear polarization $P_{\rm line}$ of the wave are determined from the well-known relations

$$P_{\rm circ}=\frac{2\rho}{1+\rho^2}, \tag{35}$$

$$P_{\rm line}=\frac{1-\rho^2}{1+\rho^2}. \tag{36}$$

We see at once that, generally speaking, $P_{\rm circ}\neq P_c$ and $P_{\rm line}\neq P_l$, even if $P_c^2+P_l^2=1$. If $\eta''_{12}=0$, the waves are linearly polarized along the coordinate axes. When absorption is present ($0<z<1$), normal modes propagating in a linearly polarized laser wave ($P_c=0$) are completely linearly

polarized ($P_{\rm line}=\pm 1$) along the coordinate axes. The same behavior is observed for a circularly polarized laser wave ($P_l=0$), where $P_{\rm circ}=\pm 1$.

If $P_c\neq 0$ and $P_l\neq 0$ in the presence of absorption ($z<1$), a beam of γ rays propagating in a laser wave will be a superposition of two elliptically polarized waves, where the principal axes of the polarization ellipses are rotated by an angle φ relative to the previously adopted coordinate system. Indeed, since the refractive indices are complex in this case, the quantity κ is complex as well, and it can be readily established by direct calculation that

$$\kappa^{(1)}\kappa^{(2)}=1, \tag{37}$$

where the parenthesized superscripts correspond to normal modes with refractive indices $\tilde{n}_1$ and $\tilde{n}_2$. Inasmuch as the quantities $\kappa^{(1)}$ and $\kappa^{(2)}$ are related by the simple expression (37), from now on we use just one of them, $\kappa=\kappa^{(1)}$, and drop the superscript.

The following relations can be derived in the present situation:

$$D_1^{(1)}D_1^{(2)}+D_2^{(1)}D_2^{(2)}=2D_2^{(1)}D_2^{(2)}, \tag{38}$$

$$D_1^{(1)}D_1^{*(2)}+D_2^{(1)}D_2^{*(2)}=D_2^{(1)}D_2^{*(2)}\frac{\kappa+\kappa^*}{\kappa^*}, \tag{39}$$

where the parenthesized superscripts refer to normal modes with refractive indices $\tilde{n}_1$ and $\tilde{n}_2$. Consequently, neither the displacement vectors $\mathbf{D}^{(1)}$ and $\mathbf{D}^{(2)}$ nor $\mathbf{D}^{(1)}$ and $\mathbf{D}^{*(2)}$ are orthogonal (see Ref. 2).

We denote the Stokes parameters of the normal mode with index $\tilde{n}_1$ by X_1, X_2, and X_3, and we denote those corresponding to $\tilde{n}_2$ by Y_1, Y_2, and Y_3. We then obtain

$$X_1=\frac{\kappa+\kappa^*}{1+\kappa\kappa^*}, \tag{40}$$

$$X_2=-\frac{i(\kappa^*-\kappa)}{1+\kappa\kappa^*}, \tag{41}$$

$$X_3=\frac{\kappa\kappa^*-1}{1+\kappa\kappa^*}. \tag{42}$$

Moreover, $Y_1=X_1$, $Y_2=-X_2$, and $Y_3=-X_3$. The angle φ can be determined from the relation $\tan 2\varphi=X_1/X_3$ (it is equal to $-\varphi$ for the second wave).

As an illustration, Fig. 4 shows the results of calculations of $|P_{\rm circ}|$, $|P_{\rm lin}|$, and $|\varphi|$ as functions of the variable z for a laser wave with $P_c=P_l=1/\sqrt{2}$.

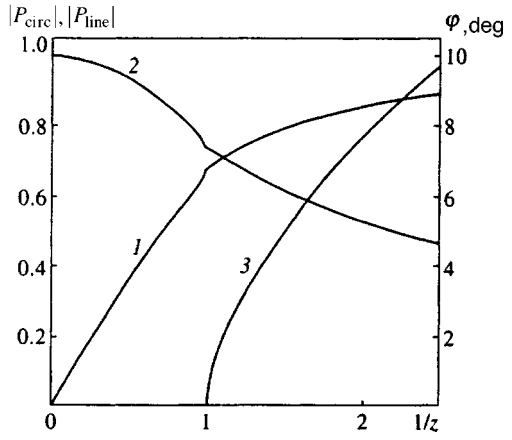


FIG. 4. Variation of the absolute values of $P_{\rm circ}$ (*1*) and $P_{\rm line}$ (*2*) and the angle of rotation φ (in deg) (*3*) of the axes of the polarization ellipse relative to the adopted coordinate system versus the invariant z. Circular and linear polarizations of the laser wave are $P_c=P_l=1/\sqrt{2}$, and for $z>1$ ($1/z<1$), $\varphi=0$.

5. PROPAGATION OF GAMMA RAYS IN A LASER WAVE

The foregoing analysis enables us to find relations describing the variation of the intensity and the polarization of γ rays propagating in a homogeneous ($n_l=$const) laser wave. Representing the beam of γ rays by a superposition of two electromagnetic normal modes with previously determined refractive indices and polarization characteristics, we obtain relations describing the variation of the intensity and the Stokes parameters of the γ-ray beam during its traversal of the laser wave:

$$J_\gamma(x)=J_1(x)+J_2(x)+2J_3(x), \tag{43}$$

$$\xi_1(x)=\frac{(X_1J_1(x)+Y_1J_2(x)+p_1J_3(x))}{J_\gamma(x)}, \tag{44}$$

$$\xi_2(x)=\frac{(X_2J_1(x)+Y_2J_2(x)+p_2J_4(x))}{J_\gamma(x)}, \tag{45}$$

$$\xi_3(x)=\frac{(X_3J_1(x)+Y_3J_2(x)+p_3J_4(x))}{J_\gamma(x)}, \tag{46}$$

where $J_\gamma(x)$ and $\xi_1(x)$, $\xi_2(x)$, $\xi_3(x)$ are the intensity and Stokes parameters of the γ-ray beam at depth x. The partial intensities $J_i(x)$ ($i=1-4$) have the form (the physical significance of these quantities is easily understood by writing out the expression $(\mathbf{D}^{(1)}+\mathbf{D}^{(2)})(\mathbf{D}^{*(1)}+\mathbf{D}^{*(2)}]$

$$J_1(x)=J_1(0)\exp(-2\ \mathrm{Im}(\tilde{n}_1)\ \omega x/c), \tag{47}$$

$$J_2(x)=J_2(0)\exp(-2\ \mathrm{Im}(\tilde{n}_2)\ \omega x/c), \tag{48}$$

$$J_3(x)=\exp(-\mathrm{Im}(\tilde{n}_1+\tilde{n}_2)\ \omega x/c)\{J_3(0)\cos[\mathrm{Re}(\tilde{n}_1-\tilde{n}_2)\ \omega x/c]-J_4(0)\sin[\mathrm{Re}(\tilde{n}_1-\tilde{n}_2)\ \omega x/c]\}, \tag{49}$$

$$J_4(x)=\exp(-\mathrm{Im}(\tilde{n}_1+\tilde{n}_2)\ \omega x/c)\{J_3(0)\sin[\mathrm{Re}(\tilde{n}_1-\tilde{n}_2)\ \omega x/c]+J_4(0)\cos[\mathrm{Re}(\tilde{n}_1-\tilde{n}_2)\ \omega x/c]\}, \tag{50}$$

where the initial partial intensities can be determined from the relations

$$J_1(0)=\frac{\xi_2(0)-f\xi_3(0)}{2(X_2-fX_3)}+\frac{\xi_1(0)-q}{2(X_1-q)}, \tag{51}$$

$$J_2(0)=-\frac{\xi_2(0)-f\xi_3(0)}{2(X_2-fX_3)}+\frac{\xi_1(0)-q}{2(X_1-q)}, \tag{52}$$

$$J_3(0)=\frac{X_1-\xi_1(0)}{2(X_1-q)}, \tag{53}$$

$$J_4(0)=\frac{\xi_3(0)X_2-\xi_2(0)X_3}{p_3(X_2-fX_3)}. \tag{54}$$

In the derivation of these relations we have made use of the relationship between the quantities X_i and Y_i, thereby eliminating them from the equations. We have also assumed that $J_\gamma(0)=1$. The parameters f, q, p_1, p_2, and p_3 have the form

$$f=\frac{i(\kappa\kappa^*-1)}{\kappa-\kappa^*}, \quad q=\frac{1+\kappa\kappa^*}{\kappa+\kappa^*},$$

$$p_1=\frac{2(1+\kappa\kappa^*)}{\kappa+\kappa^*}, \quad p_2=\frac{2(1-\kappa\kappa^*)}{\kappa+\kappa^*},$$

$$p_3=\frac{2i(\kappa-\kappa^*)}{\kappa+\kappa^*},$$

These relations describe the general case of γ-ray propagation $P_l\neq 0$, $P_c\neq 0$, $z<0$. Formally, however, they do not describe cases in which $\kappa+\kappa^*=0$ (e.g., when a beam of photons is completely circularly (or linearly) polarized or when $z>1$ for an arbitrarily polarized beam). Under the stated conditions the propagation of a beam of γ rays can be described by well-known relations[6,14] or by passing to a limit in the above-derived relations, for example, by setting $\kappa=\delta+i\rho$ and letting δ tend to zero. This case is described in the Appendix.

6. INFLUENCE OF THE LASER WAVE INTENSITY ON THE PROPAGATION OF GAMMA RAYS

The influence of the intensity of an electromagnetic wave on electron–positron pair production is investigated in several papers (see Ref. 4 and the literature cited therein). So far we have considered a relatively low-intensity laser wave, whose level can be characterized by the dimensionless parameter

$$\xi^2=\frac{\langle E^2\rangle}{E_{\rm cr}^2}\frac{m^2c^4}{E_l^2}.$$

The relations obtained here for the tensor ε_{ij} hold when $\xi^2\ll 1$. Previously obtained results[4] can be used to write the components of the dielectric tensor with allowance for the series expansion in ξ. A key issue in this treatment of the intensity is the replacement of the variable z by a variable, which we denote by $\tilde{z}$, such that $\tilde{z}=z(1+\xi^2)$. The condition for pair production in the given situation is $\tilde{z}<1$, which implies an increase in the threshold energy of the γ-ray beam at a fixed frequency of the laser wave. On the whole, the corresponding components of the dielectric tensor retain their form, but the variable z is replaced by the variable $\tilde{z}$, and the critical field $E_{\rm cr}$ is replaced by

$$\tilde{E}_{\rm cr}=\frac{m^2c^3(1+\xi^2)}{e\hbar},$$

The functions $F_2'(z,1)$, $F_2''(z,1)$, $F_1'(z)$, $F_1''(z)$, $F_c'(z)$, and $F_c''(z)$ are replaced by $F_2'(\tilde{z},\mu)$, $F_2''(\tilde{z},\mu)$, $F_1'(\tilde{z})$, $F_1''(\tilde{z})$, $F_c'(\tilde{z})$, and $F_c''(\tilde{z})$, where $\mu=1/(1+\xi^2)$. Strictly speaking, the domain of applicability of these refined equations is bounded by the condition $\xi^2\ll 1$, but important information can still be obtained in this case. For example, certain estimates in Ref. 6 are obtained for a laser with $E_l=1.18$ eV and $\xi^2\approx 0.1$. Allowance for the intensity of the laser wave shifts the pair-production threshold from 221 GeV to 247 GeV; in light of the sharp intensity dependence of the refractive indices near threshold, this shift requires appropriate allowance in experiments.

For $\xi^2\gg 1$, pair production in the laser wave becomes similar to the analogous process in a static electromagnetic field. The dielectric tensor for a static field is derived in Ref. 15, and specific results of calculations appear in Ref. 16.

7. DISCUSSION OF THE RESULTS

As mentioned, it was first shown in Ref. 6 that for existing laser parameters it is possible to observe variations of the polarization parameters of γ-ray beams. In that paper, Kotkin and Serbo consider the scattering of light by light, and show that this interaction causes the polarization characteristics of the beam to vary in a way that can be described by means of differential equations for the Stokes parameters. These equations are then solved for a transparent medium, and for circularly ($P_l=0$) and linearly $P_c=0$) polarized γ rays subject to absorption. Despite the sufficiency of such a description of γ-ray transmission through a wave field, several factors are missing from the paper (primarily a description of the laser wave as a well-defined optical medium). For example, the difference between the real refractive indices is determined from the variation of the γ-ray polarization and not as a property of the medium *per se*. This deficit is remedied in the present study of the transmission of γ rays through a laser field.

Here we have obtained the dielectric tensor for a laser field on the basis of γ-ray absorption during the production of electron–positron pairs. This approach to the problem from the standpoint of obtaining final results describing the variation of the beam polarization in a laser bunch is equivalent to the treatment of the scattering of light by light (or, more precisely, the determination of the forward scattering amplitude), since the two processes are related by the optical theorem. On the other hand, according to the principles of macroscopic electrodynamics, an electromagnetic wave (γ rays) becomes polarized as it propagates in a medium, in our case owing to the presence of virtual electron–positron pairs, which introduce lines in the diagrams of both processes in question.[8]

We have determined the refractive indices and polarization characteristics of electromagnetic waves on the basis of dispersion relations. These waves are essentially characteristic solutions of the problem of γ-ray propagation in a medium, and this can be exploited to obtain, in the general case, simple relations describing the variation of the Stokes parameters of a beam of γ rays moving in a laser wave. A comparison of the results of calculations of the variation of the Stokes parameters from Eqs. (43)–(46) with the same results obtained by numerical integration of the differential equations[6] reveals good agreement between the two (to within the numerical errors). The comparison has been made under analogous conditions; in particular, the parameter ξ^2, which takes into account the influence of the laser wave intensity on the investigated process, has been set equal to zero.

The relations that we have derived for the dielectric tensor can be used to include the influence of the intensity of the laser field in the approximation $\xi^2 \ll 1$. Despite the latter restriction, it is important take this factor into account—first of all, in view of the sharp variation of the refractive indices near the pair-production threshold (see Figs. 2 and 3), and second, because values of $\xi^2 = 0.05-0.2$ (and higher) are already fully attainable by technical means.[6]

Let us compare the results of calculations of the dielectric tensor and the refractive indices with published calculations of these quantities for slowly varying fields and low photon frequencies,[8,9,15] We should mention that direct comparison of the components of the tensor is hindered somewhat by the different approaches to the solution. The present study is based on the field equations (1), which do not contain the magnetic field H, whereas the latter is present in Refs. 8 and 9. This situation is discussed in detail in Refs. 1 and 2, and is attributable to the fact that the fields E and B are not completely independent.

However, the refractive indices are observable quantities and must be independent of the particular approach chosen. We carry out the comparison for equal-magnitude, mutually perpendicular, static electric and magnetic fields (representing a kind of electromagnetic wave analog) (Ref. 15), using Eq. (31). As a result, we find that the values of $n_\perp - 1$ and $n_\parallel - 1$ for the laser wave (Eq. (31)) are double the values for the indicated combination of fields calculated according to Ref. 15. This result can be understood by considering the opposing path traveled by γ rays in the laser wave. For massless particles the number of collisions is proportional to $1-\cos\phi$ (ϕ is the angle between the directions of motion of laser and gamma photons).[13] At $\phi = \pi$ this factor is equal to 2, and enters into the refractive indices. As the γ-ray frequency increases, the refractive indices $n_\perp$ and $n_\parallel$ in a static field gradually increase to the values of the invariant parameter $\chi = EE_\gamma / E_{cr} mc^2 \approx 0.7$ and then slowly decrease,[16] i.e., they qualitatively mimic the analogous data for γ rays moving in a laser wave (see Fig. 2), although the peak in a static field is flatter and does not exceed the low-frequency refractive indices by more than 20%.

We illustrate the feasibility of converting a linearly polarized beam of γ rays into a circularly polarized beam using a laser system that radiates 1.18-eV photons and an rms field $\approx 10^{10}$ V/cm (Ref. 6). In this case $\mathrm{Re}(\tilde{n}_\perp - \tilde{n}_\parallel) \approx 8\times 10^{-16}$ near the reaction threshold ($E_\gamma \approx 221$ GeV), corresponding to a quarter-wave plate of thickness ≈ 0.18 cm.

It is important to note that the determination of the absolute values of the refractive indices (and not merely their difference) broadens the category of problems associated with the transmission of particles through a wave; for example, it provides a means of analyzing their Čerenkov radiation[5] in such a medium. For existing laser parameters, roughly speaking, this type of radiation will be emitted by particles with a Lorentz gamma factor $\sim 10^6 - 10^7$, and the particles must be of sufficient mass to ensure low bremsstrahlung losses. The characteristic energies of Čerenkov γ rays are of the order of several hundred GeV.

The polarization of electromagnetic normal modes is no trivial matter. They are elliptically polarized in the general case ($P_c, P_l \neq 0$). For zero absorption the axes of the ellipses are parallel to the linear polarization of the laser wave. Absorption causes the axes to rotate about this direction, and they are no longer mutually orthogonal.

For $\xi^2 \ll 1$, the propagation of γ rays in the field of a linearly polarized laser wave has much in common with the like process in single crystals in the region of coherent pair production.[4] For example, in single crystals the dielectric tensor can be expressed in terms of the same functions F_1', F_2', F_1'', and F_2'' as in a laser medium.[5] A significant difference in single crystals, of course, is the occurrence of at least several equivalent-photon frequencies and the existence of an incoherent component due to thermal fluctuations, which precludes a single crystal from being regarded as a perfectly transparent medium.

It is important to note that no experiments have been performed to date to corroborate the transformation of beam polarization in single crystals, despite the notable lapse of time since the publication of Ref. 3 and the announcement of several proposals to investigate and utilize this phenomenon in modern accelerators.[11,17] The recently disclosed feasibility of experiments in laser beams[6] can be used to verify the basic principles of the theory, not only for a laser medium, but also for single crystals, owing to the similar nature of the effects, and the circular polarization of a laser wave significantly extends the sphere of such experiments.

The author is grateful to G. L. Kotkin and V. G. Serbo for useful advice.

APPENDIX

The relations for the intensity and Stokes parameters of a beam of gamma rays in the case $\kappa + \kappa^* = 0$ are

$$J = J_0 \exp[-(n_1+n_2)][\cosh(n_2-n_1) - a\sinh(n_2-n_1)],$$

$$\xi_1 = \frac{b\sin\Delta + \xi_1^0 \cos\Delta}{\cosh(n_2-n_1) - a\sinh(n_2-n_1)},$$

$$\xi_2=\frac{P_{\rm circ}\,\sinh(n_2-n_1)+aP_{\rm circ}\,\cosh(n_2-n_1)+P_{\rm line}[-b\,\cos\,(\Delta)+\xi_1^0\,\sin\,(\Delta)]}{\cosh(n_2-n_1)-a\,\sinh(n_2-n_1)},$$

$$\xi_3=\frac{P_{\rm line}\,\sinh(n_2-n_1)+aP_{\rm line}\,\cosh(n_2-n_1)+P_{\rm circ}[b\,\cos\,\Delta-\xi_1^0\,\sin\,\Delta]}{\cosh(n_2-n_1)-a\,\sinh(n_2-n_1)},$$

where

$$a=\xi_3^0P_{\rm line}+\xi_2^0P_{\rm circ},\quad b=\xi_3^0P_{\rm circ}+\xi_2^0P_{\rm line},$$

$$n_1=\mathrm{Im}(\tilde{n}_1)\;\omega x/c,\quad n_2=\mathrm{Im}(\tilde{n}_2)\;\omega x/c,$$

$$\Delta=\mathrm{Re}(\tilde{n}_1-\tilde{n}_2)\;\omega x/c,$$

$P_{\rm line}$ and $P_{\rm circ}$ correspond to the polarization of a wave with refractive index $\tilde{n}_1$, ξ_i^0 denotes the initial values of the Stokes parameters, and all other notation is the same as in the main text. In a transparent medium, obviously, $n_1=n_2=0$.

*)E-mail: maisheev@mx.ihep.su

[1] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media*, 2nd ed. (revised and enlarged with L. P. Pitaevskii), Pergamon Press, New York (1984).
[2] V. M. Agranovich and V. L. Ginzburg, *Crystal Optics with Spatial Dispersion, and Excitons*, 2nd ed., Springer-Verlag, Berlin–New York (1984).
[3] N. Cabibbo, G. Da Prato, G. De Franceschi, and U. Mosco, Phys. Rev. Lett. **9**, 435 (1962).
[4] V. N. Baĭer, V. M. Katkov, and V. M. Strakhovenko, *High-Energy Electromagnetic Processes in Oriented Single Crystals* [in Russian], Nauka, Novosibirsk (1989).
[5] V. A. Maisheev, V. L. Mikhalev, and A. M. Frolov, Zh. Éksp. Teor. Fiz. **101**, 1376 (1991) [Sov. Phys. JETP **74**, 740 (1991)]; IHEP Preprint No. 91-142 [in Russian], Institute of High-Energy Physics, Protvino (1991).
[6] G. L. Kotkin and V. G. Serbo, E-Print Archive hep-ph 9611345 (submitted to Phys. Rev. Lett.).
[7] B. De Tollis, Nuovo Cimento **32**, 754 (1964); Nuovo Cimento **35**, 1182 (1965).
[8] V. B. Berestetskiĭ, E. M. Lifshitz, and L. P. Pitaevskiĭ, *Quantum Electrodynamics*, 2nd ed., Pergamon Press, New York (1982).
[9] A. I. Akhiezer and V. B. Berestetskii, *Quantum Electrodynamics*, Wiley, New York (1965) [Nauka, Moscow (1969)].
[10] A. I. Nikishov and V. I. Ritus, Zh. Éksp. Teor. Fiz. **52**, 1707 (1967) [Sov. Phys. JETP **25**, 1135 (1967)].
[11] V. G. Baryshevskiĭ and V. V. Tikhomirov, Usp. Fiz. Nauk **159**, 529 (1989) [Sov. Phys. Usp. **32**, 1013 (1989)].
[12] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, 3rd ed., Pergamon Press, New York (1980).
[13] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, 4th ed., Pergamon Press, New York (1975).
[14] V. A. Maisheev, V. L. Mikhalev, and A. M. Frolov, IHEP Preprint No. 91-30 [in Russian], Institute of High-Energy Physics, Protvino (1991).
[15] V. N. Baĭer, V. M. Katkov, and V. S. Fadin, *Radiation from Relativistic Electrons* [in Russian], Atomizdat, Moscow (1973).
[16] A. M. Frolov, V. A. Maisheev, and V. L. Mikhaljov, Nucl. Instrum. Methods Phys. Res. A **254**, 549 (1987).
[17] K. Piotrzkowski, Nucl. Instrum. Methods Phys. Res. B **119**, 253 (1996).

Translated by James S. Wood

Ordered structures in a nonideal dusty glow-discharge plasma

A. M. Lipaev, V. I. Molotkov, A. P. Nefedov,*) O. F. Petrov, V. M. Torchinskiĭ, V. E. Fortov, A. G. Khrapak, and S. A. Khrapak

Scientific Research Center for the Thermal Physics of Pulsed Effects, Russian Academy of Sciences, 127412 Moscow, Russia

(Submitted 5 May 1997)
Zh. Éksp. Teor. Fiz. **112**, 2030–2044 (December 1997)

The formation of ordered structures of charged macroparticles in a constant-current neon glow-discharge plasma is investigated. Experiments were performed with two types of particles: thin-walled glass spheres 50–63 μm in diameter and particles of Al_2O_3, 3–5 μm in diameter. Formation of quasicrystalline structures is observed in the standing strata and in an artificially created double electric layer. The formation of extended filamentary structures of macroparticles in the absence of visible stratification of the positive column has been observed for the first time. The influence of the discharge parameters on the formation of the ordered structures and their melting is examined. The form of the interaction potential between the charged macroparticles is considered, as well as changes in the conditions for maintaining the discharge in the presence of high concentrations of dust particles. © *1997 American Institute of Physics.* [S1063-7761(97)00912-8]

1. INTRODUCTION

The presence of macroscopic particles can have a substantial effect on the properties of a low-temperature plasma. Particles heated to a sufficiently high temperature can, by emitting electrons and acquiring a positive charge, significantly increase the electron concentration in the plasma. A similar effect can occur under conditions in which the dominant process is photoemission or secondary electron emission. Cold particles, on the other hand, absorb electrons from the plasma, acquire a negative charge, and reduce the free electron concentration in the plasma. Charged particles interact with electric and magnetic fields, and the Coulomb interaction between particles can lead to a highly nonideal plasma.

A dusty plasma was first observed under laboratory conditions by Langmuir in the 1920's.[1] However, its active investigation began only in recent decades in connection with a long list of applications such as the electrophysics and electrodynamics of the combustion products of rocket fuels, the electrophysics of the working body of solid-fuel magnetohydrodynamic generators, and the physics of dust clouds in the atmosphere.[2–6] Dust and dusty plasma are widely distributed in the universe. They have been detected in planetary rings, comet tails, and in interplanetary and interstellar clouds.[7–9]

In the last ten years there has been heightened interest in studying the properties of dusty plasmas in connection with the expanded use of the technology of plasma sputtering and etching in microelectronics and in the production of thin films.[10–13] The presence of particles in plasma not only leads to contamination of the surface of the semiconductor element and thereby to an increased yield of defective components, but also perturbs the plasma in a frequently unpredictable way. The reduction or prevention of these negative effects is impossible without an understanding of the processes of formation and growth of condensed particles in a gas-discharge plasma, their transport mechanism, and their influence on the properties of the discharge.

For surface treatment, low-pressure radio-frequency gas-discharge plasma is usually used.[14] The degree of ionization of such a plasma is low ($\sim 10^{-7}$), the electron energy is a few eV, and the ion energy is near the thermal energy of the atoms (≈ 0.03 eV). A neutral, nonemitting particle incident on such a plasma is buffeted by fluxes of all the particles present in the plasma, including electrons and ions. It is customary to assume that electrons incident on the surface of the particle are absorbed, and the ions raining down upon its surface knock out electrons and recombine with them.

As a consequence of the great difference in masses and temperatures of the electrons and ions, the electron flux exceeds the ion flux by several orders of magnitude, and the particle acquires a negative charge. The negative electrostatic potential on the particle leads to repulsion of electrons and attraction of ions. The charge of the particle changes until the electron and ion fluxes on the particle equalize.

The steady-state charge Z or floating potential φ_p can be estimated within the framework of the orbital motion model, widely used in the theory of plasma probes.[15] This model is valid in the collisionless regime for particles of sufficiently small dimensions

$$R_p \ll \lambda \ll l, \tag{1}$$

where R_p is the particle radius, and λ and l are the typical screening length and the typical mean free path of the electrons or ions, whichever is the smaller. Balance of electron and ion currents leads to the following equation for φ_p:

$$N_e \sqrt{\frac{T_e}{m_e}} \exp\left(\frac{e\varphi_p}{T_e}\right) = N_i \sqrt{\frac{T_i}{m_i}} \exp\left(1 - \frac{e\varphi_p}{T_i}\right), \tag{2}$$

where $T_{e(i)}$ and $m_{e(i)}$ are the temperature and mass of the electrons (ions).

Equation (2) enables one to estimate the potential and charge of an isolated particle in the plasma. The typical charge of a micron-sized particle lies within the range from 10^3 to 10^5 electron charges. The walls of the discharge

chamber and the electrodes acquire a negative potential. All this makes it possible under certain conditions to compensate for the effect of gravitation, and leads to levitation of the particles above the lower electrode or floor of the discharge chamber. Methods for confining a dusty plasma in special traps, which make it possible to reduce contamination of the work surfaces, are based on this effect.

The thermodynamic properties of dusty plasmas are mainly determined by the magnitude of the nonideality parameter Γ, which is equal to the ratio of the Coulomb potential energy to the kinetic energy of thermal motion characterized by the particle temperature T_p:

$$\Gamma = \frac{Z^2 e^2}{a T_p}, \quad a = \left(\frac{3}{4\pi N_p} \right)^{1/3}, \tag{3}$$

where a is the mean distance between particles, and N_p is their concentration. Thanks to the large charge of the particles, nonideality in the interaction between particles can enter significantly earlier than nonideality of the electron–ion subsystem, despite the fact that the particle concentration is usually low in comparison with the electron and ion concentrations.

From the simplest and most widely studied model of a one-component plasma it is well known that for $\Gamma > 1$ short-range order appears in the system, and for $\Gamma \approx 170$ a one-component plasma crystallizes.[16] The one-component model cannot claim to provide a faithful description of the properties of a dusty plasma, above all because it neglects screening effects. Nevertheless, a number of papers, on the basis of qualitative results of the one-component model, express the view that near-range order is possible in a thermal-equilibrium dusty plasma, and even crystallization.[4,5,17]

Recently, ordered structures of dust particles were detected in a thermal plasma at atmospheric pressure at a temperature of around 1700 K (Refs. 18 and 19). Similar considerations led Ikezi[20] to infer the possibility of crystallization of a dust subsystem in a nonequilibrium gas-discharge plasma. Eight years after the publication of this paper, a dust crystal was finally observed experimentally in a high-frequency discharge near the lower electrode at the boundary of the near-cathode region,[21–24] and later in the strata of a stationary glow-discharge.[25,26]

A plasma crystal can have varied crystal structure, with a lattice constant on the order of fractions of a millimeter, which makes it possible to observe it practically with the naked eye. Plasma crystals possess a great many virtues, making them an indispensable instrument both in the study of highly nonideal plasmas and in the study of the fundamental properties of crystals. These include, first and foremost, simplicity of preparation, observation, and parameter control, and their short equilibrium relaxation times and response times to external perturbations.

A dust crystal is not a unique example of the appearance of long-range order in Coulomb systems. In his time, Wigner showed[27] that upon cooling, an electron gas can condense and form an ordered crystalline structure, the so-called ''Wigner crystal.'' Recently, crystallization of a quantum electron liquid with formation of a Wigner crystal was investigated experimentally.[28] Crystal structures, also in one-component systems, were observed in electrostatic vacuum traps on charged macroparticles[29] and in Paul and Penning traps on Mg or Be ions cooled to very low temperatures ($\sim 10^{-3}$ K).[30–32] A Coulomb crystal is also realized in colloidal suspensions.[33] Colloidal crystals consist of almost monodisperse micron-sized particles suspended in an electrolyte, where they become charged to $Z = -(10^3 - 10^4)$ and are screened by ions of both signs in the electrolyte. According to a conjecture of Deryagin and Landau,[34] under certain conditions the Coulomb interaction between the particles makes formation of a crystal structure energetically more favorable. Strong coupling between particles takes place at distances less than the screening radius, which in colloidal suspensions is very small. This leads to the result that for crystallization a rather high particle density ($N_p \sim 10^{12}$ cm^{-3}) is necessary. As a consequence, colloidal crystals are usually opaque, hindering experimental study of their bulk properties. To the drawbacks of colloidal crystals as an instrument of physical study may also be added their long equilibrium relaxation time, amounting to several weeks.

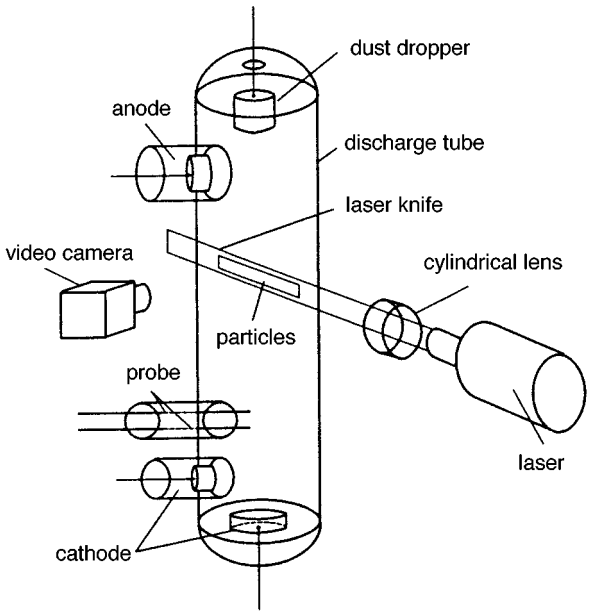


FIG. 1. Schematic of the experimental setup.

The present paper investigates the formation of ordered structures of charged macroparticles of various sizes in a constant-current glow-discharge plasma in neon. The influence of the discharge parameters on the possibility of the existence of quasicrystalline structures of dust particles is examined, along with the conditions for their formation and destruction. We consider the question of the form of the interaction potential between the macroparticles, and also the effect of the macroparticles themselves on the conditions for maintaining the discharge.

2. EXPERIMENTAL SETUP

A glow-discharge was created in neon in a cylindrical, vertically positioned tube with cold electrodes. A schematic of the setup is shown in Fig. 1. The inner diameter of the

tube was 3 cm, the length of the tube was 60 cm, the distance between the electrodes situated in lateral nibs of the tube was 40 cm. A double mobile probe was also placed in the tube. The discharge current was varied within the range from 0.1 to 10 mA, and the neon pressure, from 0.2 to 2 Torr. Discharge regimes with standing strata existed in this range.

Micron-sized particles to be introduced into the plasma were held in a cylindrical container located in the upper part of the gas-discharge tube. The bottom of the container was fashioned from a metal mesh with a spacing of 100 μm. When the container was shaken, particles rained down into the positive column of the discharge. The particles were visualized with the aid of illumination in either the horizontal or vertical planes by a probe laser beam. The beam from an argon laser was formed by a cylindrical lens into a planar converging beam with thickness of the beam waist at the center of the discharge tube equal to 150 μm and width 40 mm.

The horizontal probe beam could be vertically translated along the length of the tube, and the vertical probe beam, both in height and radius. Light scattered by the particles was observed with the aid of a CCD camera at an angle of 60° in the case of the horizontal beam and at an angle of 90° in the case of the vertical beam. The output signal of the camera was recorded on video tape. Note that individual particles and the cloud of particles as a whole are visible in the laser light even to the naked eye.

Oscillations of the discharge connected with repositioning of the cathode spot are observed in a glow-discharge with a cold cathode, which causes fluctuations of the ordered structures. To damp these oscillations, an additional tube with a constriction was introduced. This inset was positioned in the lower part of the discharge tube above the cathode.

3. EXPERIMENTAL RESULTS AND DISCUSSION

3.1. Formation of ordered structures in strata of the positive column

In the experiments we used two types of particles: particles of borosilicate glass ($\rho = 2.3$ g/cm^3) in the form of thin-walled, hollow spheres (microballoons) of diameter 50–63 μm with wall thickness 1–5 μm (the mass of the particles M_p lies in the range from 2×10^{-8} to 10^{-7} g) and Al_2O_3 particles ($\rho = 4$ g/cm^3) with diameter 3–5 μm (M_p lies in the range from 6×10^{-11} to 3×10^{-10} g).

In the presence of standing strata in the positive column of the discharge, dust particles that poured down from the container hovered in the form of a cloud in the center of the luminous part of the stratum. The charged microscopic particles captured into the stratum formed ordered quasicrystalline structures, whose size and shape depended on the discharge parameters. The region of existence of the strata for the given discharge tube lies in the range of neon pressures from 0.1 to 1.7 Torr for the discharge current varying from 0.1 to 10 mA. The length of the luminous part of the stratum was 10 mm at a pressure $p = 1.2$ Torr and grew to 25 mm at $p = 0.2$ Torr. The distance between the luminous parts of the strata depended weakly on the discharge parameters and lay within the limits 35–50 mm. Note that the ordered structures of charged particles formed in the strata have a narrower existence region.

The formation process proceeds as follows: after being shaken out of the container, the particles fall past their equilibrium position and then, over the course of several seconds, rise back up and form into an ordered structure which is preserved for as long as desired provided the discharge parameters are left unchanged. Individual particles can move upward toward the anode and along the periphery of the stratum. In addition, peculiar orbital motions of the particles around the ordered structures are observed. Contrary circular motions of individual particles are also observed. A particle completes one revolution in approximately 10 s.

The simultaneous existence of ordered structures in several neighboring strata was observed. By way of example, Fig. 2a shows an image of structures of charged microballoons of borosilicate glass in two neighboring strata. By varying the discharge parameters (current, pressure) it is possible to obtain coalescence of the structures in neighboring strata into one long, extended formation. Figure 2b displays such a structure.

Figure 3 shows digitized images of structures of Al_2O_3 particles taken in the horizontal plane for two values of the discharge current I_p: 0.4 and 3.9 mA for neon pressure $p = 0.3$ Torr. Figure 4 shows two plots of the conditional particle distribution function $n(r) = n_2(r)/n_1(r)$, obtained by processing the corresponding images in Fig. 3; n_1 and n_2 are the corresponding unary and binary functions.[35,36] The function $n(r)$ is the particle number density at a distance r from some particle. The choice of this function instead of the commonly used pairwise correlation function $g(r)$ is dictated by the inhomogeneity of the investigated structures associated with their relatively small dimensions.

Figure 4 plots the distribution function $n(r)$ for each of the above two discharge regimes for the central part of a horizontal cross section of the structure, and for all of the particles in the given cross section (the central part of the horizontal cross section in our case consists of the particles lying inside a circle whose center coincides with the center of mass of the cross section, and includes around 40% of all the particles observed in the cross section). It can be seen that for $I_p = 0.4$ mA at least three maxima of $n(r)$ can be identified, which suggests the existence of long-range order in the investigated dusty plasma. This corresponds to a crystal-like structure.

It follows from these images and distribution functions that with increasing discharge current there is a tendency toward the destruction of long-range order (''melting''), and at $I_p = 3.9$ mA only short-range order is observed.

Note that the destruction of long-range order accompanying an increase in the discharge current above all entrains the periphery of the structure. The central part of the formation preserves its former order. In addition, with increasing I_p, spontaneous oscillations of individual macroparticles are observed. In this situation, such particles move along circular trajectories whose radii increase with distance from the center of the structure. Circular oscillations of the particles were also observed in a dusty rf discharge plasma when the gas pressure was decreased.[37] As can be seen from Fig. 4, de-

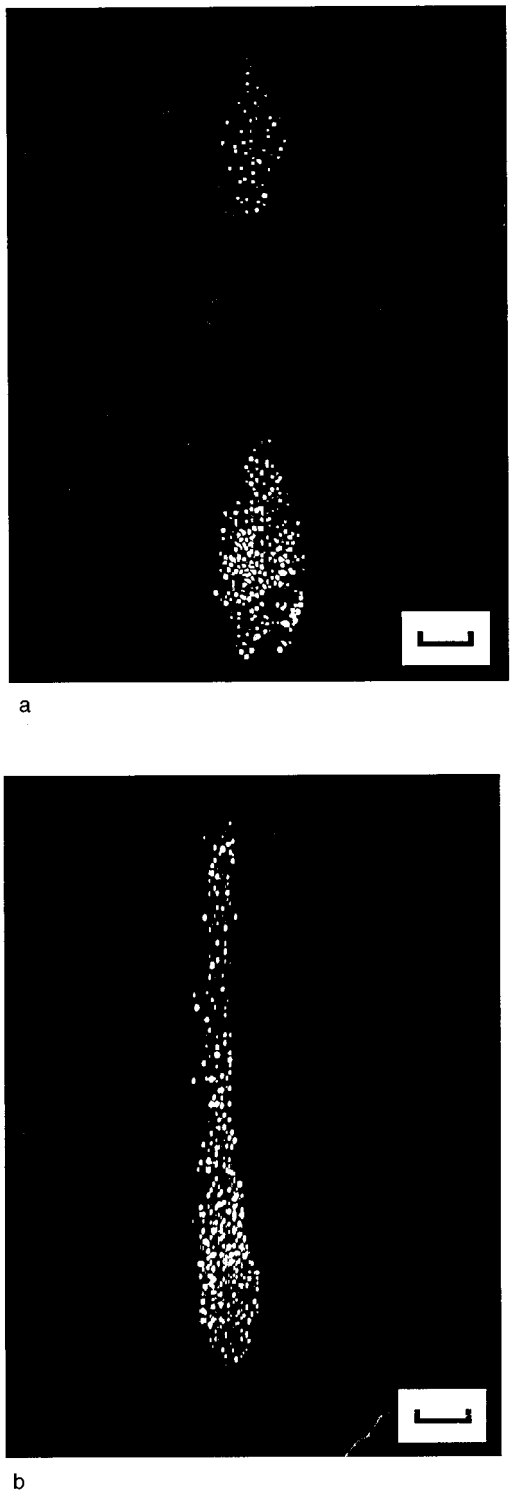


FIG. 2. Video images of structures of charged microballoons of borosilicate glass: a) in two neighboring strata ($p=0.5$ Torr, $I_p=0.5$ mA), b) after their coalescence ($p=0.4$ Torr, $I_p=0.4$ mA). The scale in the figure corresponds to 3 mm.

spite varying I_p by an order of magnitude, the mean interparticle distance a, equal to 250 μm, remains essentially unchanged.

At a higher neon pressure of 0.7 Torr, less ordered structures of charged Al_2O_3 particles are observed in the range of currents from 0.7 to 7 mA. The interparticle distance a in this case is also independent of the discharge current and is equal to 280 μm.

The constancy of the interparticle distance a in the face of a significant change in the discharge current and, consequently, the electron concentration N_e, differs sharply from the appreciable dependence of a on the power fed to the rf discharge (and accordingly on the electron concentration N_e) observed in Ref. 23.

Figure 5 shows a digitized image of the structure of charged microballoons of borosilicate glass in one of the vertical planes. Figure 6 plots the corresponding distribution functions for the central part of the structure and the structure as a whole. These results demonstrate the existence of significant ordering of the particles in the vertical cross sections.

Our experiments revealed the emergence of unusual formations of charged macroparticles: in different discharge regimes, with the disappearance of visible standing strata extended filamentary structures are formed, extending upward from the dark cathode space through the height of the tube. The length of these structures reached 60 mm. Fragments of the filamentary structure are shown in Fig. 7. Near the dark cathode space the number of filaments reaches 7–8, while in the upper part their number is reduced to 1–2. The number of particles per filament (chain) reaches 100–120. Note that the extended filamentary structures are observed for both types of microscopic particles used in the experiments.

Strata in low-pressure discharges have been experimentally investigated in considerable detail.[38–41] In the positive column of the discharge under the conditions of interest, loss of electron energy in elastic collisions is negligibly small and the electron distribution function is formed under the action of the electric field and the inelastic collisions. This can lead to the appearance of strata—a spatial periodicity of the plasma parameters with characteristic scale $\lambda_1=\varepsilon_1/eE_0\approx 4-5$ cm (ε_1 is the first excitation potential, equal to 16.6 eV for neon, and E_0 is the electric field averaged over the length of the stratum).

The electron concentration, their energy distribution, and the electric field are highly irregular along the length of the stratum. The electric field E is relatively strong at the head of the stratum (around 10–15 V/cm at its maximum)—a region occupying 25–30% of the length of the stratum, and weak (around 1 V/cm) outside this region. The maximum value of the electron concentration is displaced relative to the maximum of E in the direction of the anode.[39] The electron energy distribution is substantially bimodal,[39] and the head of the stratum is dominated by the second maximum, whose center lies near the excitation potential ε_1.

Due to the high floating potential of the walls of the discharge tube, the strata have a substantially two-dimensional character: the center–wall potential difference at the head of the stratum reaches 20–30 V, and the change in the potential takes place in a narrow near-wall layer of thickness 2–3 mm (Refs. 40–42). Thus, an electrostatic trap is found at the head of each stratum, which in the case of vertical orientation is capable of confining particles with high enough charge and low enough mass.

The experimental information above was obtained only

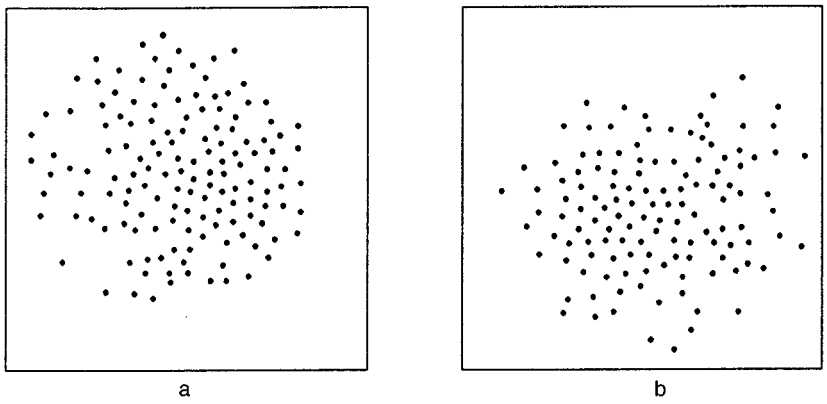


FIG. 3. Digitized video images of structures of Al_2O_3 particles in the horizontal plane for $p=0.5$ Torr: a) $I_p=0.4$ mA, b) $I_p=3.9$ mA. Frame dimensions 5×5 mm^2.

for moving strata; however, there is every reason to believe that for the same discharge parameters, the properties of resting and moving strata are similar.[43] By virtue of the substantially two-dimensional nature of the problem, a theoretical description of the strata in a low-pressure discharge is quite complicated, and at present can lay claim to only a qualitative explanation of the observed effects.[44,42] As a consequence of particles sticking to the probe, efforts under taken in the present work to measure plasma parameters with the aid of a double probe in the presence of macroparticles have not been crowned with success. Nevertheless, the electric field strengths measured in the positive column in a neon discharge in the absence of stratification are in good agreement with the published data. Therefore, in our analysis we will be forced to rely, above all, on the results of an experimental study of strata[40–42] carried out under conditions similar to ours.

The potential of the particle φ_p, and consequently its charge Z, can be estimated with the help of Eq. (2). However, it must be recalled that its derivation employs the Maxwellian electron energy distribution far from the particle. As we have already pointed out, the electron distribution function is bimodal in the strata, and Eq. (2) can be used for estimates only in those regions of the stratum where one of the maxima predominates. Thus, in the region of maximum luminosity, the second maximum, which has energy $T_e\approx\varepsilon_1$, predominates. This makes it possible with the aid of Eq. (2) to estimate the maximum value of the floating potential $\varphi_p\approx-1.65\varepsilon_1\approx-27$ eV. The relation between the potential of a particle and its charge in the linear Debye screening approximation is[45]

$$Z=\varphi_p R_p(1+R_p/\lambda_L), \tag{4}$$

where λ_L is the linearized screening length

$$\lambda_L=\left[4\pi e^2\left(\frac{N_e}{T_e}+\frac{N_i}{T_i}\right)\right]^{-1/2}. \tag{5}$$

This yields a charge $Z\approx-3\times10^4$ for the Al_2O_3 particles, and $Z\approx-7\times10^5$ for the glass particles.

Levitation of macroscopic particles takes place in the region of the stratum where the electrostatic force ZeE is balanced by the gravitational force M_pg. Thus, it is possible to determine the magnitude of the electrostatic field E_m in which particle levitation is possible:

$$E_m=\frac{M_pg}{Ze}. \tag{6}$$

This yields $E_m\approx1$ V/cm for the Al_2O_3 particles and $E_m\approx45$ V/cm for the glass particles. The latter is 3–4 times the fields usually observed at the head of the stratum. However, this disagreement can be partly explained by the separation of particles according to wall thickness, and consequently mass, where this separation takes place directly in

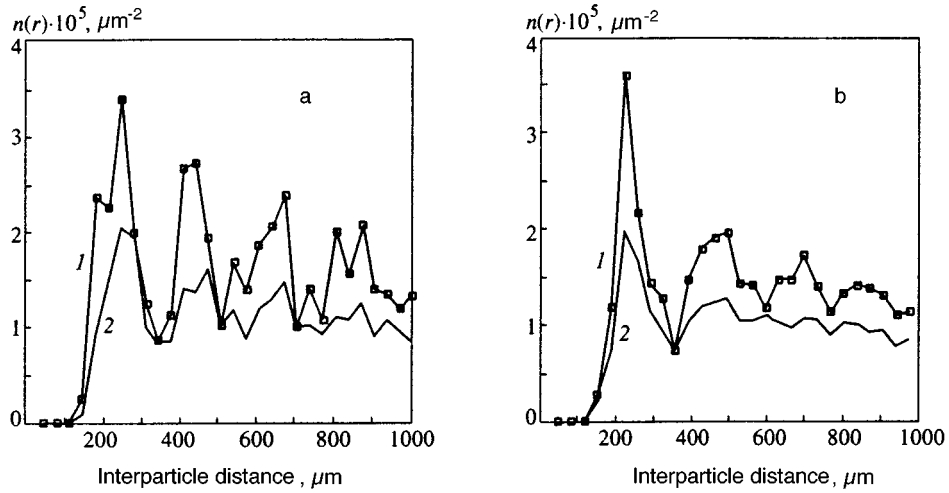


FIG. 4. Distribution functions $n(r)$ for the structures shown in Fig. 3: a) $I_p=0.4$ mA, b) $I_p=3.9$ mA. *1*—$n(r)$ for the central region of the cross section, *2*—for the entire cross section.

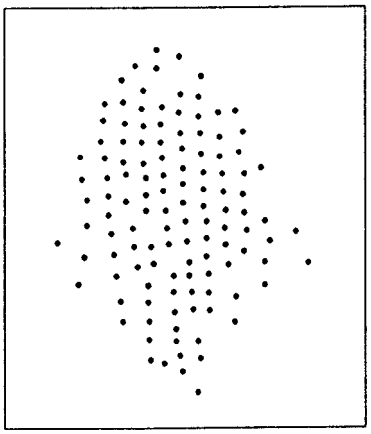


FIG. 5. Digitized video images of structures of charged microballoons of borosilicate glass in the vertical plane for $p=0.2$ Torr and $I_p=0.7$ mA. Frame dimensions 6×7.5 mm^2.

the discharge: particles with minimal M_p/Z are captured, above all the thin-walled ones with wall thickness 1–2 μm. In addition, the very presence of charged particles can lead to an increase in the electric field.

The Coulomb interaction between charged particles is proportional to the product of their charges. Therefore, a large value of Z leads to a strong Coulomb repulsion between particles, and consequently nonideality of the system. For a mean interparticle distance $a=300$ μm at $T=300$ K, the nonideality parameter $\Gamma\sim10^6$ and 10^8 for particles of radius 1.5 and 25 μm, respectively. Note, however, that the particles are screened by the electrons and ions of the plasma, whose concentration varies severalfold along the length of the stratum and lies within the limits $5\times10^7-5\times10^8$ cm^{-3} (Refs. 40–42). The floating potential of the particles is equal in order of magnitude to the electron energy, and significantly exceeds the ion energy. Therefore the screening of the particles is substantially nonlinear. In addition, electron and ion recombination takes place on the surface of the particles, as a result of which there is no reverse charged-particle flux near the particles, and their distribution function is non-Maxwellian. Thus, even the asymptotic behavior of the potential far from the particle surface ceases to be Debye-like, and depends on the distance according to a power law, $\varphi(r)\approx-ZR_P/2r^2$. This effect has long been known, and has been well studied in the theory of spherical electrostatic probes.[15]

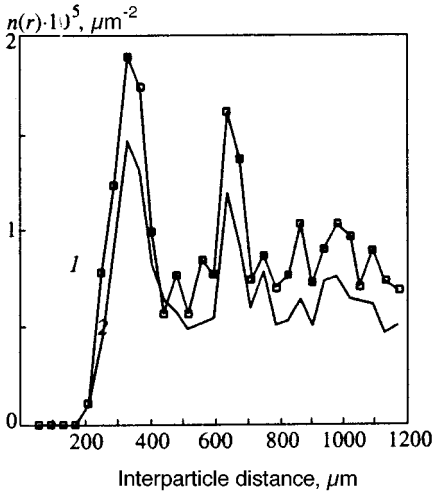


FIG. 6. Distribution functions $n(r)$ for the structure imaged in Fig. 5: *1*—$n(r)$ for the central region of the cross section, *2*—for the entire cross section.

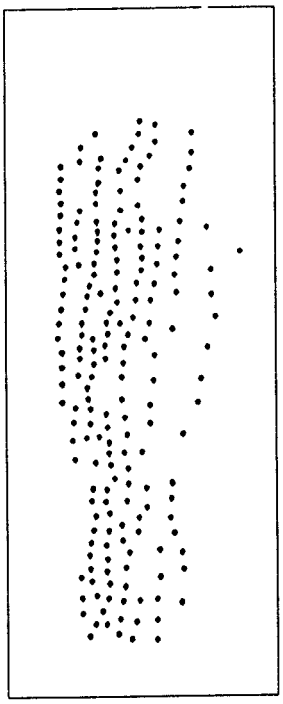


FIG. 7. Digitized video image of a fragment of a filamentary extended structure of charged glass microballoons in the vertical plane for $p=0.4$ Torr and $I_p=0.4$ mA. Frame dimensions 7.5×18 mm^2.

The structure of the screening cloud in the collisionless regime with allowance for the nonlinearity of the Poisson equation and the non-Maxwellian character of the electron and ion energy distribution functions was calculated in Ref. 46 for spherical particles in a helium discharge. It turns out that at small distances the particle potential can be approximated to high accuracy by the Debye–Hückel potential

$$\varphi(r)=\varphi_p\,\frac{R_p}{r}\,\exp\left(-\frac{r-R_p}{\lambda}\right). \tag{7}$$

However $\lambda=\lambda_L$ only in the case of sufficiently small particles ($R_p\ll\lambda_L$), and for large R_p the screening length increases with R_p, reaching and even exceeding the electron screening length $\lambda_e=\sqrt{T_e/4\pi e^2N_e}$.

Figure 8 plots the results of a calculation of $\varphi(r)$ that we performed for a neon discharge, for particles of two sizes, under conditions typical of the mid-region of the strata in neon, and allowing for the non-Maxwellian character of the electron and ion distribution functions due to their absorption by the surface of the macroparticle. It can be seen that the approach to the non-Debye-like dependence $\varphi(r)\propto-1/r^2$ takes place earlier for larger particles, and at distances close to the mean interatomic distance in the levitating ordered structures. This effect can lead to an increase in repulsion at

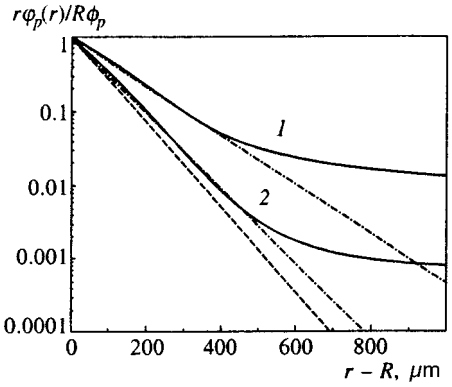


FIG. 8. Dependence of the unit-normalized product $r\varphi_p$ on distance to the surface of the particle in neon ($T_e=50000$ K, $T_i=300$ K, $N_e=5\times10^8$ cm^{-3}) for particles of radius $R=25$ μm (*1*) and $R=1.5$ μm (*2*). The dash–dot curves plot the dependence obtained from Eq. (7) with $\lambda=130$ and 85 μm, respectively. The dashed curve corresponds to Eq. (7) with $\lambda=\lambda_L$.

large interatomic distances, and a slackening of the dependence of the interparticle interaction on the electron and ion concentrations in the screening cloud.

Figure 9 plots the dependence of the screening length λ on the electron concentration N_e. Proper choice of the latter enables one to accurately approximate the calculated potential $\varphi(r)$ near the particle with the help of the Debye–Hückel potential (7). As in the case of the helium discharge,[45] at small radii it is possible to use the linear screening length λ_L, but for $R_p\sim\lambda_L$ it is necessary to allow for nonlinear effects. Note that under the conditions of interest, λ shows essentially no dependence on T_e.

Allowing for screening of the charge of the dust particles and the non-Maxwellian character of the electron and ion velocity distribution functions near the particles, as a nonideality parameter in place of Γ calculated according to Eq. (3), it is possible to use[20]

$$\Gamma_D=\frac{Z^2e^2}{aT_p}\,e^{-a/\lambda} \tag{8}$$

in the case of small Al_2O_3 particles, and

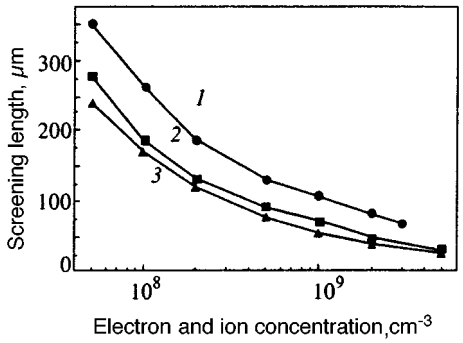


FIG. 9. Dependence of the effective screening length λ on the electron concentration N_e in neon ($T_e=50000$ K, $T_i=300$ K) for particles of radius $R=25$ μm (*1*) and $R=1.5$ μm (*2*). Curve *3* plots the linear Debye–Hückel screening length (5).

$$\Gamma_n=\frac{Z^2e^2}{aT_p}\,\frac{R_p}{2a}. \tag{9}$$

in the case of large glass particles.

In both cases, the degree of nonideality can be reduced by several orders of magnitude in comparison with Γ; however, it is still too large to explain the observed ''melting'' of the ordered structures. One reason for this is that the particles are held back not where the electrostatic force acting on them is maximum (and consequently their charge Z is also maximum), but where it balances the gravitational force. The second and probably the most important reason is that, as a consequence of the fluctuating plasma microfields, particles charged to large values of Z have mean kinetic energy T_p significantly greater than the temperature of the gas. This effect is observed both in a radio-frequency plasma[37,47] and in our experiments.[26] In both cases, with distance from the melting curve, the particle energy reaches a very high value of the order of 50 eV. Such a growth of the energy can explain the observed ''melting'' of dust crystals.

3.2. Formation of ordered structures in an electric double layer

Levitation of macroparticles in the positive column of a constant-current glow-discharge requires an electric field strong enough to balance the force of gravity and is possible not only in strata, but also in a specially organized electric double layer. Toward this end, one can vary the plasma parameters by varying the transverse cross section of the positive column.[48]

However, in contrast to Ref. 48, we use a positive column with a narrow cathode part with radius R_k and a wide anode part with radius R_a ($R_k<R_a$) with an expansion segment in between. In this case, a double layer of space charge separating the two plasma regions with different electron temperatures T_{ek} and T_{ea} and electron concentrations N_{ek} and N_{ea} appears in the mouth of the constriction. The narrow cathode part has higher values of T_{ek} and N_{ek}. The potential jump U at the double layer depends on the gas pressure, the radii R_k and R_a, their ratio R_a/R_k, and the discharge current. U can be estimated from the relation[48]

$$U=T_{ek}-T_{ea}+\frac{T_{ek}+T_{ea}}{2}\ln\frac{N_{ek}}{N_{ea}}. \tag{10}$$

Thus, for the given discharge tube the introduction of a constriction with a radius $R_k=3.5$ mm yields an increase in N_e for $p=0.5$ Torr of almost an order of magnitude, in T of a factor of three, and $U\approx10$ V. Since the longitudinal dimension of the double layer in the mouth of the constriction is not large (~1 cm), the values of the longitudinal electric field in the layer are of the same order of magnitude as at the head of the stratum. Thus, conditions are ensured for the capture and suspension of charged macroparticles, with subsequent formation of an ordered structure.

In experiments with a constriction $R_k=3.5$ mm, we did indeed obtain levitation of particles of both types in the mouth of the narrow part. The constriction was created by introducing an additional cylindrical glass tube with variable

inner diameter into the discharge tube, the wide part of which was placed over the cylindrical cathode.

3.3. Effect of dust particles on discharge parameters

At moderate pressures under stationary steady-state conditions, charge losses in a weakly ionized plasma are associated with ambipolar diffusion toward the walls of the discharge chamber.[38] In this case the electron temperature (and accordingly the longitudinal electric field) in the positive column of the glow-discharge can be determined by equating the ionization rate and the rate of ambipolar diffusion losses.

The presence of dust particles in the discharge can substantially alter this situation. The point is that for a high enough macroparticle concentration, charge loss to their surfaces becomes predominant in comparison with loss to the chamber walls. An idea of the efficiency of the two processes of charged particle loss can be had by comparing their respective rates. The rate of ambipolar losses is given by

$$\nu_d = \frac{D_a}{\Lambda^2}, \tag{11}$$

where D_a is the ambipolar diffusion coefficient, and $\Lambda = R/2.4$ is the characteristic diffusion length for cylindrical discharge geometry[38] (R is the radius of the discharge chamber). The electron loss rate to the surfaces of the dust particles in the orbital motion approximation is given by

$$\nu_p = \pi R_p^2 N_p \sqrt{\frac{8T_e}{\pi m_e}} \exp\left(\frac{e\varphi_p}{T_e}\right). \tag{12}$$

For the ambipolar diffusion coefficient the following estimate applies:[38]

$$D_a \approx D_i \frac{T_e}{T_i} \approx \frac{1}{3N_a\sigma} \sqrt{\frac{8T_i}{\pi m_i}} \frac{T_e}{T_i}. \tag{13}$$

Here σ is the resonant recharging cross section (in the case of motion of the ions in their own gas) and N_a is the atom concentration. Hence we obtain for the ratio of the characteristic rates

$$\frac{\nu_p}{\nu_d} = 3\pi R_p^2 \left(\frac{R}{2.4}\right)^2 N_p N_a \sigma \sqrt{\frac{T_i m_i}{T_e m_e}} \exp\left(\frac{e\varphi_p}{T_e}\right). \tag{14}$$

Let us estimate the macroparticle concentration for which charge loss to the particles is more probable than to the walls. Using the neon discharge parameters $N_a = 3\times 10^{16}$ cm^{-3} ($P = 1$ Torr), $T_e = 50000$ K, $T_i = 300$ K, $R_p = 30$ μm, $\sigma \approx 2\times 10^{-15}$ cm^{-2}, $e\varphi_p/T_e \approx -2$, we find from (14) that the rates coincide for $N_p \approx 250$ cm^{-3}.

Under the conditions of the experiment described here, the concentration of macroparticles of the given size was roughly an order of magnitude greater. Thus, the main channel of charge loss is probably loss to dust particles. Under such conditions, an increase in their density should be accompanied by a rise in the ionization rate needed to maintain the stationary discharge regime. This obviously means a rise in the electron temperature, and consequently the electric field, in the region in which the macroparticles are located. If this field grows to a sufficient magnitude, it will keep the particles from falling as a result of the force of gravity. This mechanism removes the quantitative disagreement between the electric field values required to compensate for the force of gravity and the electric field values usually observed at the head of the stratum in the case of heavy glass particles, and also allows a qualitative explanation of the experimentally observed suspension of an extended ordered structure of dust particles in the absence of visible stratification of the positive column of the glow discharge.

4. CONCLUSION

In the present paper we have investigated the formation of ordered structures of charged macroparticles in the positive column of a neon glow discharge. We observed the formation of quasicrystalline structures whose size and shape depended on the discharge parameters. Formation of the structures took place in a region of high electric field,which balanced the force of gravity, in standing strata and in a double electric layer created by introducing an additional cylindrical glass tube with variable inner diameter into the discharge. The effect of the discharge parameters on the possible existence of quasicrystalline dust-particle structures was investigated. It was found, in particular, that increasing the discharge current leads to the destruction of long-range order (melting of the quasicrystal).

The experiments also revealed the emergence of unusual formations of charged macroparticles. For certain values of the discharge parameters, extended filamentary structures were formed in the absence of stratification of the positive column. The length of such structures reached 6 cm. Note that this is the first reported observation of such structures.

The electrostatic interaction between dust particles, which leads to the formation of ordered structures, was considered in detail with allowance for the nonlinearity of screening with a non-Maxwellian energy distribution of the electrons and ions of the plasma. A difference in the shape of the interaction potential is noted for different sorts of macroparticles (due to a difference in their diameters). An estimate was made of the coupling parameter Γ corresponding to the existence of ordered structures. We point out the necessity of allowing for an additional channel of electron and ion loss in the presence of high dust-particle concentration in the discharge.

ACKNOWLEDGMENTS

We thank Yu. B. Golubovskiĭ, L. D. Tsendin, and V. V. Zhakhovskiĭ for useful discussions. This work was carried out with the partial financial support of the Russian Fund for Fundamental Research (Grant No. 97-02-17565) and a joint grant from the Russian Fund for Fundamental Research and INTAS (Grant No. 95-1335).

*)E-mail: Nefedov@hedric.msk.su

[1] I. Langmuir, G. Found, and A. F. Dittmer, Science **60**, 392 (1924).
[2] S. L. Soo, *Multiphase Fluid Dynamics*, Gower Technical, Brookfield, Vermont (1990).
[3] M. S. Sodha and S. Guha, Adv. Plasma Phys. **4**, 219 (1971).
[4] D. I. Zhukhovitskiĭ, A. G. Khrapak, and I. T. Yakubov, in *Plasma Chemistry*, B. M. Smirnov (ed.), **11**, 130 (1984).

[5] I. T. Yakubov and A. G. Khrapak, Sov. Tech. Rev. B: Thermal Phys. **2**, 269 (1989).
[6] B. M. Smirnov, *Aerosols in Gas and Plasma* [in Russian], IVTAN Press, Moscow (1990).
[7] C. K. Goetz, Rev. Geophys. **27**, 271 (1989).
[8] T. G. Northrop, Phys. Scr. **45**, 475 (1992).
[9] V. N. Tsytovich, Usp. Fiz. Nauk **167**, 57 (1997).
[10] A. Garscadden, B. N. Ganguly, P. D. Haaland, and J. Williams, Plasma Sources Sci. Technol. **3**, 239 (1994).
[11] A. Bouchoule and L. Boufendi, Plasma Sources Sci. Technol. **3**, 292 (1994).
[12] J. Goree, Plasma Sources Sci. Technol. **3**, 400 (1994).
[13] G. S. Selwyn, in *The Physics of Dusty Plasmas*, P. K. Shukla, D. A. Mendis, and V. W. Chow (eds.), World Scientific, Singapore (1996), p. 177.
[14] Yu. P. Raĭzer, M. N. Shneĭder, and N. A. Yatsenko, *High-Frequency Capacitive Discharge* [in Russian], MFTI Press and Nauka, Moscow (1995).
[15] J. E. Allen, Phys. Scr. **45**, 497 (1992).
[16] S. Ichimaru, Rev. Mod. Phys. **54**, 1017 (1982).
[17] V. E. Fortov and I. T. Yakubov, *Nonideal Plasma* [in Russian], Énergoatomizdat, Moscow (1994).
[18] V. E. Fortov, A. P. Nefedov, O. F. Petrov, A. A. Samarian, A. V. Chernyschev, and A. M. Lipaev, JETP Lett. **63**, 187 (1996).
[19] V. E. Fortov, A. P. Nefedov, O. F. Petrov, A. A. Samarian, and A. V. Chernyschev, Phys. Lett. A **219**, 89 (1996).
[20] H. Ikezi, Phys. Fluids **29**, 1764 (1986).
[21] J. H. Chu and L. I, Phys. Rev. Lett. **73**, 652 (1994).
[22] H. Thomas, G. E. Morfill, V. Demmel *et al.*, Phys. Rev. Lett. **73**, 652 (1994).
[23] A. Melzer, T. Trottenberg, and A. Piel, Phys. Lett. A **191**, 301 (1994).
[24] Y. Hayashi and K. Tachibana, Jpn. J. Appl. Phys., Part 2 **33**, L804 (1994).
[25] V. E. Fortov, A. P. Nefedov, V. M. Torchinskiĭ, V. I. Molotkov, A. G. Khrapak, O. F. Petrov, and K. F. Volykhin, JETP Lett. **64**, 92 (1996).
[26] V. E. Fortov, A. P. Nefedov, V. M. Torchinsky, V. I. Molotkov, O. F. Petrov, A. A. Samarian, A. M. Lipaev, and A. G. Khrapak, Phys. Lett. A **229**, 317 (1997).
[27] E. Wigner, Phys. Rev. **46**, 1002 (1934).
[28] H. Jiang, R. L. Willet, H. Stormer *et al.*, Phys. Rev. Lett. **65**, 633 (1990).
[29] R. F. Wuerker, H. Shelton, and R. V. Langmuir, J. Appl. Phys. **30**, 342 (1959).
[30] F. Diedrich, E. Peik, J. M. Chen *et al.*, Phys. Rev. Lett. **59**, 2931 (1987).
[31] S. L. Gilbert, J. J. Bollinger, and D. J. Wineland, Phys. Rev. Lett. **60**, 2022 (1988).
[32] I. Waki, S. Kassner, G. Birkl, and H. Walther, Phys. Rev. Lett. **68**, 2007 (1992).
[33] A. K. Sood, Solid State Phys. **45**, 1 (1991).
[34] B. Deryagin and L. Landau, Acta Phys. Chim. USSR **14**, 633 (1941).
[35] F. M. Kuni, *Statistical Physics and Thermodynamics* [in Russian], Nauka, Moscow (1981).
[36] T. Hill, *Statistical Mechanics: Principles and Selected Applications*, McGraw-Hill, New York (1956).
[37] H. M. Thomas and G. E. Morfill, Nature (London) **379**, 806 (1996).
[38] Yu. P. Raĭzer, *Gas-Discharge Physics* [in Russian], Nauka, Moscow (1987).
[39] Yu. P. Raĭzer, *Principles of the Contemporary Physics of Gas-Discharge Processes* [in Russian], Nauka, Moscow (1980).
[40] Yu. B. Golubovskiĭ, S. U. Nisimov, and I. É. Suleĭmenov, Zh. Tekh. Fiz. **64**, 54 (1994) [Tech. Phys. **39**, 1005 (1994)].
[41] Yu. B. Golubovskiĭ and S. U. Nisimov, Zh. Tekh. Fiz. **65**, 46 (1995) [Tech. Phys. **40**, p. (1995)].
[42] Yu. B. Golubovskiĭ and S. U. Nisimov, Zh. Tekh. Fiz. **66**, 20 (1996) [Tech. Phys. **41**, p. (1996)].
[43] P. S. Landa, N. A. Miskinova, and Yu. V. Ponomarev, Usp. Fiz. Nauk **132**, 601 (1980) [Sov. Phys. Usp. **23**, 813 (1980)].
[44] L. D. Tsendin, Fiz. Plazmy **8**, 400 (1982) [Sov. J. Plasma Phys. **8**, 228 (1982)].
[45] J. E. Daugherty, R. K. Porteous, and D. B. Graves, J. Appl. Phys. **73**, 1617 (1993).
[46] J. E. Daugherty, R. K. Porteous, M. D. Kilgore, and D. B. Graves, J. Appl. Phys. **72**, 3934 (1992).
[47] A. Melzer, A. Homann, and A. Piel, Phys. Rev. E **53**, 2757 (1996).
[48] V. L. Granovskiĭ, *Electrical Current in a Gas. Steady-State Current* [in Russian], Nauka, Moscow (1971).

Translated by Paul F. Schippnick

Surface driven transition in a nematic liquid crystal cell

D. Andrienko, Yu. Kurioz, and Yu. Reznikov*)

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

V. Reshetnyak

Institute of Surface Chemistry, National Academy of Sciences, 252650 Kiev, Ukraine
(Submitted 23 November 1996)
Zh. Éksp. Teor. Fiz. **112**, 2045–2055 (December 1997)

Surface driven reorientation effects in a nematic liquid crystal cell caused by light-induced changes of the anchoring parameters were studied. Theoretical consideration of one-dimensional flat distributions of the director has shown that the director can undergo threshold reorientation between hybrid, homeotropic, and planar alignments as the anchoring energy varies continuously. The threshold reorientation takes place when the reference and light-induced easy axes are perpendicular. In the one-elastic-constant approximation the light-induced transition was found to be of second order as shown by a critical increase of the director thermal fluctuations in the vicinity of the transition point. These effects were experimentally studied in the cells containing 5CB liquid crystal aligned by the photosensitive azo-containing polymer layer. © *1997 American Institute of Physics.* [S1063-7761(97)01012-3]

1. INTRODUCTION

Reorientation effects caused by light acting on the aligning surfaces of liquid crystals have been intensively studied in the past decade[1–7] because of their fascinating expected applications and surprising nonlinear optic effects. The basic idea of this work is to control the orientation of the liquid crystal by the phototransformation of the aligning surface.

Ichimura *et al.*[1,2] in their publications have carefully treated the change of the liquid crystal alignment under the action of light on the azo-containing aligning layers. It was found that trans-cis isomerisation of azo-inits results in the out-of-plane anchoring transition between the planar (the easy orientation axis **e** lies in the plane of the surfaces) and homeotropic (the vector **e** is normal to the cell surfaces) easy axis directions. In turn, the linearly polarized actinic light wave caused in-plane reorientation of the director on the polyvinyl-cinnamate polymer surface due to the cross-linking reaction between cinnamic acid groups.[4] A similar effect was observed under the action of linearly polarized light on azo-containing polymers[5–8] and under the action of light on the bulk of azo-containing nematic liquid crystal.[9,10]

The effects above can be explained by the light-induced changes of the anchoring parameters. In the macroscopic approach and Rapini approximation two basic parameters characterize the anchoring: the easy axis **e** and the anchoring energy W. Both parameters can be induced or changed by the action of light on the photosensitive aligning layer.

As was shown in our previous publications, the in-plane smooth and threshold reorientation effects can be described by changes in the azimuth anchoring energy.[4,8,10] These changes are due to the modification of the orientation distribution function of polymer fragments onto the aligning surface under the action of actinic light, i.e., a change in the order parameter of the fragments.

In accordance with the work of Ichimura *et al.*,[1,2] the out-of-plane reorientation is the result of trans-cis isomerisation of the side chain fragments of the aligning layer. There are two main microscopic scenarios of this effect. The first one is the percolation-like process due to the competition between the planar ability of cis-isomers and homeotropic properties of trans-isomers. This scenario should lead to a discontinuous change in the polar anchoring energy as the cis-isomer concentration increases. The second one is the smooth change of the total aligning ability of the mixture of trans- and cis-isomers, leading to a smooth change in the polar anchoring energy.

The aim of this work is to consider the out-of-plane surface driven reorientation effects in nematic liquid crystal cells caused by smooth changes in the polar anchoring energy and to discuss the microscopic scenario of the process realized under the trans-cis isomerisation of polymer fragments on the aligning layer.

2. THEORY

The system under consideration is a nematic liquid crystal bounded by two plates covered with alignment materials. The free energy of such a sample in the approximation of the surface Rapini potential has the form[11]

$$F=\frac{1}{2}\int (K_1(\operatorname{div}\mathbf{n})+K_2(\mathbf{n}\cdot\operatorname{curl}\ \mathbf{n})^2 + K_3[\mathbf{n}\cdot\operatorname{curl}\ \mathbf{n}]^2)dV-\frac{1}{2}W_0\int(\mathbf{n}\cdot\mathbf{e}_0)^2dS - \frac{1}{2}W_L\int(\mathbf{n}\cdot\mathbf{e}_L)^2dS, \tag{1}$$

where K_1 are the Frank elastic constants, $\mathbf{e}_{0,L}$ are the unit vectors along the easy axes on the surfaces, and $W_{0,L}$ are the anchoring energies which can depend on the light field intensity $I=E^2$.

Let us consider the geometry realized in the experiment described below (Fig. 1). The top plate is the reference one

 1063-7761/97/121119-06$10.00

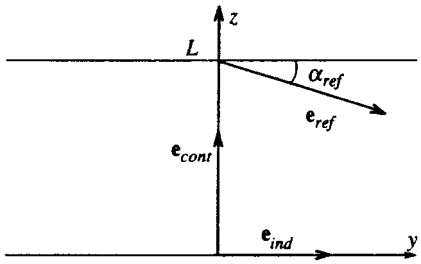


FIG. 1. Geometry of the out-of-plane director reorientation. Bottom surface is the control, upper is the reference one; $\mathbf{e}_{\rm con}$, $\mathbf{e}_{\rm ind}$ and $\mathbf{e}_{\rm ref}$ are the initial, induced and reference easy axes, correspondingly.

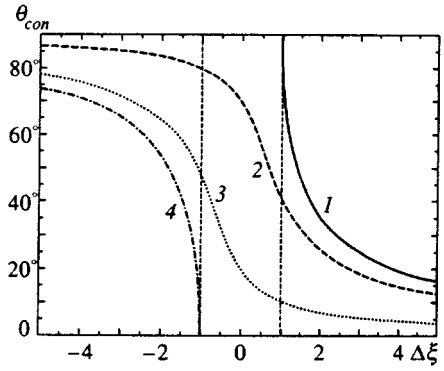


FIG. 2. Calculated dependence of the tilt angle $\theta_{\rm con}$ on the control surface on the anchoring parameter $\Delta\xi$ at the different tilt angles $\alpha_{\rm ref}$ on the reference surface: *1*—$\alpha_{\rm ref}=90°$; *2*—$\alpha_{\rm ref}=70°$; *3*—$\alpha_{\rm ref}=20°$; *4*—$\alpha_{\rm ref}=0°$.

and is characterized by an easy axis direction $\mathbf{e}_{\rm ref}=(0,\cos\alpha_{\rm ref},-\sin\alpha_{\rm ref})$ and constant (large) anchoring energy $W_{\rm ref}$. The bottom photosensitive surface is the control one with the homeotropic easy axis $\mathbf{e}_{\rm con}=(0,0,1)$ and the anchoring energy $W_{\rm con}$. Let us assume that the incident light gives rise to the additional planar easy axis $\mathbf{e}_{\rm ind}=(0,1,0)$ on the control surface with the anchoring energy $W_{\rm ind}$. Both $W_{\rm con}$ and $W_{\rm ind}$ depend on the light intensity I. We also suppose that the director characteristic reorientation times are much shorter than the time over which the anchoring energy varies. In this case the director distribution is determined by the minimum of the total free energy at each instant of time. Further we also assume that the director reorientation occurs in the yz-plane («flat» reorientation). Thus, in the one-elastic-constant approximation, the free energy (1) can be brought into the form

$$\frac{F}{S}=\frac{1}{2}K\int_0^L\theta_z'^2dz-\frac{1}{2}W_{\rm ref}\cos^2(\alpha_{\rm ref}-\theta_{\rm ref})\big|_{z=L}$$

$$-\frac{1}{2}(W_{\rm con}-W_{\rm ind})\sin^2\theta_{\rm con}\big|_{z=0}, \qquad (2)$$

where $\theta(z)$ is an angle between the y-axis and the director, and $\mathbf{L}$ is the cell thickness. Minimization of the free energy (2) leads to the equation for the director distribution

$$\theta_z''=0 \qquad (3)$$

and the boundary conditions

$$L\theta_z'+\frac{1}{2}(\xi_{\rm con}-\xi_{\rm ind})\sin 2\theta\big|_{z=0}=0,$$

$$L\theta_z'+\frac{1}{2}\xi_{\rm ref}\sin 2(\theta-\alpha_{\rm ref})\big|_{z=L}=0, \qquad (4)$$

where $\xi_\mu=W_\mu L/K$ ($\mu=$con,ind,ref) are the anchoring parameters. In the case of infinite anchoring on the reference surface the solution to the equation (3) takes the form

$$\theta=\theta_{\rm con}+(\alpha_{\rm ref}-\theta_{\rm con})z/L, \qquad (5)$$

where $\theta_{\rm con}$ is the tilt angle on the control surface and can be found from the implicit equation

$$\alpha_{\rm ref}-\theta_{\rm con}=\frac{1}{2}(\xi_{\rm ind}-\xi_{\rm con})\sin 2\theta_{\rm con} \qquad (6)$$

From (6) one can obtain the conditions for the stability of the director distribution (5). As a result we find that for $\alpha_{\rm ref}=0$ and $\xi_{\rm ind}-\xi_{\rm con}=\Delta\xi>-1$, only the planar distribution is stable. For $\Delta\xi>1$ the hybrid distribution with tilted orientation of director on the surfaces is realized. In this case decreasing $\xi_{\rm ind}$ causes the tilt angle $\theta_{\rm con}$ to increase monotonically, i.e., smooth reorientation of director toward $\mathbf{e}_{\rm con}$.

For $0<\alpha_{\rm ref}<\pi/2$, increasing $\xi_{\rm ind}$ results in a monotone decrease of the tilt angle $\theta_{\rm con}$ with no critical points for $\Delta\xi$.

For $\alpha_{\rm ref}=\pi/2$ the only $\theta_{\rm con}=\pi/2$ solution is stable for $\Delta\xi<1$. For $\Delta\xi>1$, the director distribution becomes tilted and the angle $\theta_{\rm con}$ decreases toward $\mathbf{e}_{\rm ind}$ as $\xi_{\rm ind}$ increases.

Thus, for $\alpha_{\rm ref}=0$ and $\alpha_{\rm ref}=\pi/2$ the reorientation shows the threshold behavior as the anchoring energy on the control surface varies (Fig. 2). In the one-elastic-constant approximation this reorientation is a second-order orientational transition, so that the director fluctuations are assumed to exhibit a critical behavior in the vicinity of the transition point.

The thermal fluctuations $\delta\mathbf{n}$ of the director in a cell with arbitrary anchoring have been considered in detail in Refs. 12 and 13. In line with the approaches developed there the director fluctuations $\delta\mathbf{n}$ can be regarded as a superposition of the eigenfunctions of the operator $-(K/2)\Delta$ and have the form

$$\delta n_\alpha(\mathbf{r})=\exp(i\mathbf{q}_\perp\cdot\boldsymbol{\rho})$$

$$\times\sum_j\{\delta n_{\alpha j}^+\exp(iq_z^jz)+\delta n_{\alpha j}^-\exp(-iq_z^jz)\}, \qquad (7)$$

where $\alpha=x,y$, $\mathbf{q}_\perp=(q_x,q_y,0)$ is the wave vector of fluctuations in the plane of the cell surfaces, $\boldsymbol{\rho}=(x,y,0)$, and the sum is over the discrete modes of the fluctuation wave vectors q_z^j.

Let us consider the case of $q_x=0$ and the out-of-plane part of the fluctuations (δn_y) which is responsible for the ee-scattering of light and corresponds to the situation realized in the experiment.[11] The boundary conditions (4) linearized with respect to the director fluctuations, which determine the values of q_z^j, have the form

$$L\delta n_{y,z}-(\xi_{\rm con}-\xi_{\rm ind})\delta n_y\big|_{z=0}=0,$$

$$L\delta n_{y,z}+\xi_{\rm ref}\delta n_y\big|_{z=L}=0. \qquad (8)$$

Substituting (7) in the boundary conditions (8), we obtain an equation for the q_z-components of the fluctuation wave vectors:

$$\tan t = \frac{t(\xi_{\text{ref}} + \Delta\xi)}{t^2 - \xi_{\text{ref}}\Delta\xi}, \tag{9}$$

where $t = q_z L$.

Application of the equipartition theorem of classical statistical mechanics gives the fluctuation amplitudes

$$|\delta n_{yj}^{+}|^2 = A\left\{2 + \frac{2\Delta\xi}{\Delta\xi^2 + t^2}(1 - \cos 2t) - \frac{\Delta\xi^2 - t^2}{t(\Delta\xi^2 + t^2)}\sin 2t\right\}^{-1}, \tag{10}$$

where $A = k_B T/KV(q_\perp^2 + q_z^{j2})$. It is seen that the denominator in (10) can be equal to zero for some definite values of $\Delta\xi$ and t. This means that the amplitude of the lowest fluctuation mode experiences an infinite growth, i.e., a second-order structural transition occurs. In the case of a strong anchoring on the reference surface this condition reduces to the existence of the solution to the system of equations

$$\Delta\xi \tan t = t, \quad t^2 - \Delta\xi(1 - \Delta\xi) = 0. \tag{11}$$

The first equation corresponds to Eq. (9) for $\xi_{\text{ref}} = \infty$ and the second one to the vanishing of the denominator in (10). It is apparent that $t = 0$, $\Delta\xi = 1$ is the solution of this system and thus defines the transition point.

The asymptotic behavior of the fluctuation wave vector and amplitude of fluctuation in the limit $\Delta\xi \to 1$ has the form

$$q_z L \sim \sqrt{1 - \Delta\xi}, \quad |\delta n_{y1}|^2 \sim \frac{1}{4VKq_\perp^2(1 - \Delta\xi)}. \tag{12}$$

It is seen that the lowest mode of the splay-bend director fluctuation diverges when $\Delta\xi \to 1$. Moreover, the behavior of δn_{y1} in the vicinity of the point $\Delta\xi = 1$ is typical of the second-order transitions.

This orientational transition is a light-induced anchoring transition and is the surface analog of the well known light-induced Freederiksz effect.[11] In turn, the smooth surface director reorientation for an oblique easy axis direction is the analog of the giant optical nonlinearity widely studied in the past decade.[14] It should be noted that the director can undergo reorientation not only when W changes but also due to the changes of the Frank elastic constants or the cell thickness (the dimensionless parameter determining the director distribution in the cell is $\xi = WL/K$). The orientational transition at a critical cell thickness in a wedge-shaped cell was studied by Barbero *et al.*[15]

In summary, we emphasize the new facts. Firstly, smooth variation of the anchoring can lead to threshold reorientation of the director. Secondly, the threshold reorientation takes place only for a definite cell geometry, namely, when the easy axes are orthogonal on the aligning surfaces. Lastly, if the threshold reorientation occurs, the director thermal fluctuations increase near the transition point. Since the differential cross-section is proportional to the squared amplitude of the fluctuations, the intensity of the ee-scattered light should increase near the transition point.

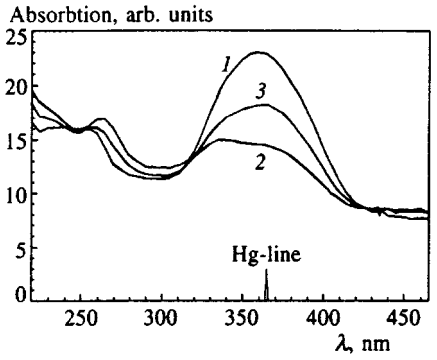


FIG. 3. The initial absorption spectrum of AP and its transformation under UV light irradiation: *1*—initial spectrum; *2*—spectrum obtained after the irradiation with exciting light of Hg-lamp; *3*—spectrum obtained in an hour after the irradiation.

3. MATERIALS AND EXPERIMENTAL PROCEDURE

The experiments were performed with a sandwich cell arrangement filled with the mixture of the liquid crystal 4′ −*n*-pentyl-4-cyanobiphenyl (5CB). The thickness of the cells was 70 μm. One of the aligning plates was the reference. We used three types of reference surfaces: a polydimethyl-siloxane (PS) layer (provides homeotropic alignment, $\alpha_{\text{ref}} = 90°$), an obliquely evaporated layer of In_2O_3 ($\alpha_{\text{ref}} \approx 30°$), and a rubbed polyimide (PI) layer ($\alpha_{\text{ref}} = 4.5°$). The angles α_{ref} were determined by zero magnetic and optical rotation methods[16] in the symmetric cells with a homogeneous director distribution.

The control surface was coated with a layer of photosensitive azo-polymer (AP) (Fig. 3) providing the homeotropic orientation of 5CB. This polymer belongs to a class of photosensitive polymers side fragments of which undergo the trans-cis isomerisation as well as the reorientation of trans-isomers perpendicular to the polarization vector of UV light.[17,18] The initial absorption spectrum of AP and its transformation under UV-irradiation are shown in Fig. 3.

Thus, we used two initial liquid crystal alignments in the cells: the hybrid (PI or In_2O_3 reference surfaces) and homeotropic (PS reference surface). To observe reorientation of the director, the cell C was inserted in the optical circuit normally to the exciting UV beam from a Hg-lamp (Fig. 4). The control surface faced the Hg-lamp. The wavelength spectrum of the exciting light was given by the filter F_1 and is shown in Fig. 3. The polarization $\mathbf{E}_{\text{exc}}$ of the UV light was set by the Glan–Thompson quartz prism Pr. In the case of a hybrid

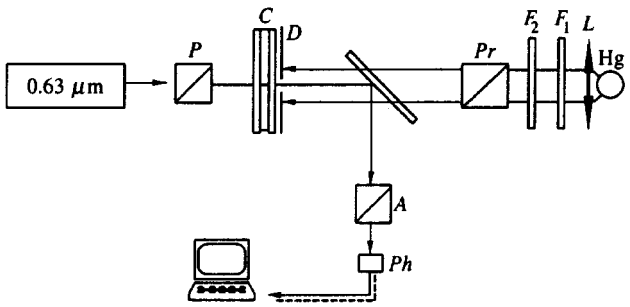


FIG. 4. Experimental setup.

cell with tilted liquid crystal alignment direction, $\mathbf{E}_{exc}$ was parallel to the director. It had an arbitrary direction when the cell with homeotropic orientation were used. The irradiated circle with radius $R=1.5$ mm and homogeneous light intensity distribution was formed by the condenser L and the diaphragm D attached directly to the cell. The UV light intensity I_{exc} was controlled in the range 0.5–50 mW/cm^2 with the filters F_2.

A probe beam from a He–Ne laser ($\lambda_{test}=0.63$ mm, power $P_{test}=0.5$ mW) passed through the polarizer P, cell C, diaphragm D, and analyzer A. The polarizer P and analyzer A were crossed and their axes made a 45° angle with the projection of the director **n** on the aligning surfaces. The intensity I_{test}^{out} of the light behind the analyzer was measured by a photodiode Ph connected to the computer.

The action of the UV light did not affect either the reference aligning layer or the liquid crystal. Therefore, the change of I_{test}^{out} under UV irradiation indicated director reorientation in the cell caused by the phototransformation of the control surface.

The intensity I_{test}^{out} behind the analyzer is given by[19]

$$I_{test}^{out}=I_{test}^{in}\sin^2(\pi\Delta\psi/\lambda), \tag{13}$$

where

$$\Delta\psi=\int_0^L\left(\frac{n_0 n_e}{\sqrt{n_0^2+(n_e^2-n_0^2)n_z^2}}-1\right)dz$$

is the phase difference of the extraordinary and ordinary waves, n_e and n_0 are the refractive indexes of liquid crystal for the extraordinary and ordinary waves, respectively, and $n_z=\sin[\theta_{con}+(\alpha_{ref}-\theta_{con})z/L)$.

To observe the critical behavior of the director fluctuations in the vicinity of the reorientation threshold we used the experimental setup for studying Rayleigh scattering of light described in detail in Ref. 12. The total ee-scattering in the range 2–10 mrad was detected through simultaneous irradiation with the exciting UV light.

4. EXPERIMENTAL RESULTS

The irradiation of cells with the exciting beam caused the test light to appear behind the analyzer, i.e., director reorientation occurred. After the UV light had been cut off the initial state was restored in several minutes.

Observations in a polarizing microscope of the irradiated area in the cells with initially homeotropic liquid crystal alignment (PS reference surface) revealed a typical schlieren structure of the tilted or planarly oriented liquid crystal. Sometimes monodomain orientation of the liquid crystal occurred in this area. If hybrid cells were used (PI and In_2O_3 reference surfaces), tilted director alignment of the irradiated spot was observed. The irradiated area in this case was much less bright than that in the remaining part of the cell, which was attributable to the decrease of the director tilt in the cells under the UV light irradiation.

The dependence of the light intensity behind the analyzer $I_{test}^{out}(t_{exc})$ on the irradiation time t_{exc} is shown in Fig. 5 for different types of reference surfaces. It is seen that there is a smooth director reorientation in the hybrid cell with

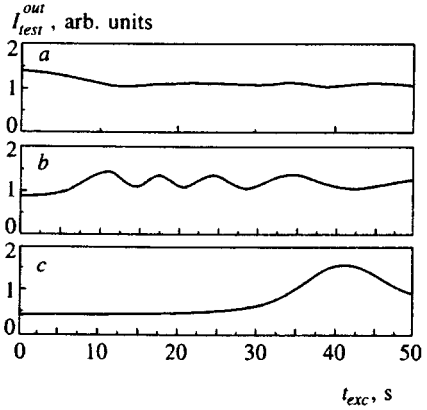


FIG. 5. Light intensity behind the analyzer $I_{test}^{out}(t_{exc})$ as a function of the irradiation time t_{exc}, for different types of reference surfaces: a—$\alpha_{ref}=30°$; b—$\alpha_{ref}=4.5°$; c—$\alpha_{ref}=0°$.

tilted direction of the easy axis $\mathbf{e}_{ref}$ (PI and In_2O_3 reference surfaces). In contrast, we have obvious threshold reorientation in the initially homeotropic cell (PS reference surface). The dependence $\theta_{con}(t_{exc})$ calculated in accordance with Eq. (13) for the threshold reorientation of director is shown in Fig. 6a.

The increase of I_{exc} caused an increase in the number of oscillation periods for all types of cells and is due to the increase in the maximum director deviation angle on the control surface. Within the limits of experimental error the dependence of the stationary value $\theta(I_{exc})$ was linear up to $I_{exc}\simeq 20$ mW/cm^2. In the case of the homeotropic cells the «silent time» (time needed to reach the threshold) was inversely proportional to the intensity I_{exc} (Fig. 7).

The reorientation of the director was accompanied by light scattering caused by fluctuations of the director. The light scattering intensity I_{sc}^{ee} increased with the irradiation time and then went through a maximum. Since the intensity of the scattered light I_{sc}^{ee} is proportional to $\sin 2\beta$, where β is the angle between the director **n** and the wave vector of the incident light,[12,13] this increase can be explained by the change in this angle during reorientation. In the homeotropic cell the scattering only increased with the reorientation. Evidently, it is due to the appearance of schlieren structure in the irradiated area.

It is significant that in the hybrid cells the increase in the light scattering began simultaneously with the director reorientation. At the same time, we found that the intensity of the scattered light increased before the director reorientation in the homeotropic cell. It is seen in Fig. 8 that a definite delay exists between the beginning of the light scattering increase and the director reorientation. Note that the value of this delay depended on the exciting light intensity I_{exc}, i.e., larger intensities give smaller delays.

5. DISCUSSION

The experimental results evidently show the threshold director reorientation in the homeotropic cell and critical behavior of the director fluctuation in the vicinity of the threshold point. There are two possible explanations of this effect.

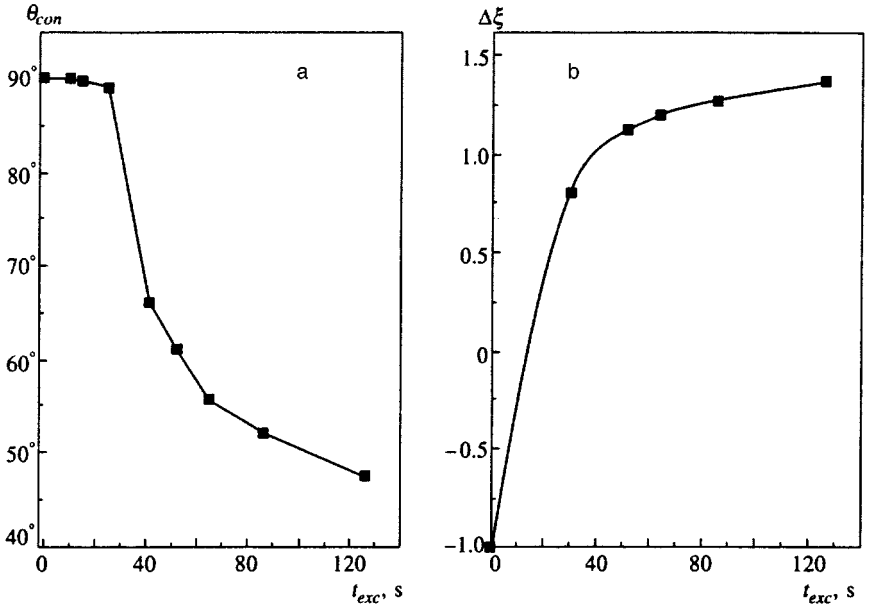


FIG. 6. a) Tilt on the control surface, $\theta_{\rm con}(t_{\rm exc})$, and b) the anchoring difference $\Delta\xi(t_{\rm exc})$, calculated in accordance with Eq. (13).

The first one is the discontinuous change of the anchoring energy on the control surface. Without doubt, if the difference $\Delta W = W_{\rm ind} - W_{\rm con}$ of the anchoring energies as a function of the excitation time has a break, the threshold reorientation of director should occur both in the hybrid and homeotropic cells. The second possibility is a smooth change in the difference ΔW with threshold director reorientation only in the homeotropic cell, which is due to the competition between bulk and surface torque and was discussed above (Sec. 2).

As the experiments have demonstrated, only the homeotropic cell exhibits threshold reorientation of the director. Hence, the model of smooth anchoring energy variation is valid in our case. Comparison of the experimental dependence $\theta_{\rm con}(t_{\rm exc})$, Fig. 6a, and the solution of Eq. (6) allows us to calculate the anchoring difference $\Delta\xi = \Delta WL/K$ at each instant of time. The results obtained for the homeotropic cell are shown in Fig. 6b. The initial value $W_{\rm con}(t_{\rm exc}=0) \approx 4\cdot 10^{-4}$ erg/cm^2 was measured independently in the symmetric cell by the H-field method proposed recently by Lavrentovich *et al.*[16]

It would appear reasonable that the light-induced changes in the anchoring energy $\Delta W = W_{\rm ind} - W_{\rm con}$ are due to the change in the concentrations $C_{\rm trans}$, $C_{\rm cis}$ of trans- and cis-isomers under the action of UV light. In fact, this mechanism is responsible for the planar-homeotropic transitions in the cells with azo-polymer aligning material.[1,2] Moreover, our studies of the absorption spectrum of a thick AP layer in the range of 300–800 nm revealed the typical transformation under trans-cis isomerisation of azo-dye[18] (Fig. 3).

Thus, the difference $\Delta W = W_{\rm ind} - W_{\rm con}$ is a function of $C_{\rm trans}$, $C_{\rm cis}$. At $C_{\rm cis}=0$ the homeotropic orientation of the easy axis is realized, whereas the planar alignment takes place for $C_{\rm trans}=0$. In this model $C_{\rm trans}$, $C_{\rm cis}$ depend on the UV light intensity and give the experimentally obtained increase in the stationary value of the tilt angle $\theta_{\rm con}$ with increasing $I_{\rm exc}$. Moreover, at small irradiation times $C_{\rm trans}(t)$, $C_{\rm cis}(t)$ are linear, and the time needed to achieve their critical values corresponding to the anchoring difference $\Delta\xi = f(C_{\rm cis,trans}) = 1$ is irreversibly proportional $I_{\rm exc}$ as was

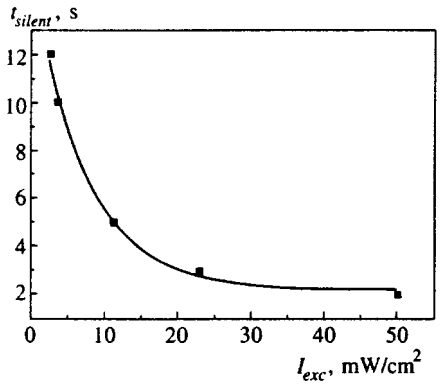


FIG. 7. Dependence of the silent time on the intensity of incident light.

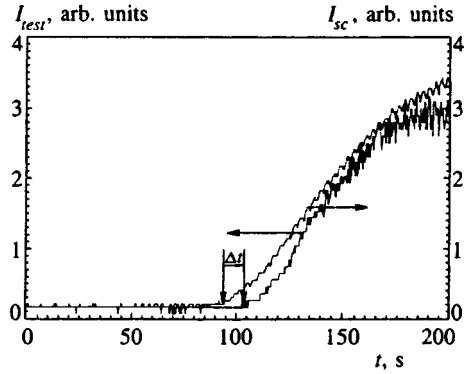


FIG. 8. Critical increase of the light scattering in the vicinity of the reorientation threshold. The increase in the light scattering starts earlier than the director reorientation and is due to the attainment of the critical value for the anchoring parameter.

found experimentally (Fig. 7). The saturation of the light-induced anchoring energy for long exposure times could be caused by the saturation of the concentration of cis-isomers and decimation of trans-isomers.

It should be noted that we observed a difference between the relaxation behavior of AP and liquid crystal in the cell after switching off the UV light. The characteristic time of AP absorption spectrum relaxation was one to two hours for the thick AP layer, while the recovery time of the initial homeotropic orientation in the cells was several seconds. This result requires further study. However, we believe that it reflects the fact that if the relaxation time of the AP spectrum is determined by the lifetime of cis-isomers, $t_{\rm cis}$, in the polymer matrix, the recovery time t_d of the initial director alignment is also determined by the orientational viscosity and elasticity of the liquid crystal and by the interaction between the liquid crystal and the cell walls (boundary conditions). It is well known that the value of $t_{\rm cis}$ strongly depends on the environment of cis-isomers, being hours in a polymer matrix and few seconds in a liquid crystal matrix.[20,21] Since cis-isomers at the cell wall directly interact with liquid crystal molecules at the polymer surface, their lifetime should be comparable with the value of $t_{\rm cis}$ in a liquid crystal matrix, i.e., with the value observed in our experiments. Moreover, it is reasonable to suppose that steady homeotropic alignment occurs for a concentration of cis-isomers at the surface far from its equilibrium value that also makes $t_d \leqslant t_{\rm cis}$. Besides, strong homeotropic anchoring of liquid crystal at the reference surface may also bring about the early recovery of the initial alignment.

The authors thank to A. Iljin, T. Marusii, and S. Slussarenko for useful discussions. We are very grateful to A. Dyadyusha for help in anchoring energy measurements and to A. Tereshchenko who carried out the spectral studies. The research described in this publication was made possible in part by Grants Nos. U5W000, U5W200 from the International Science Foundation, GR/J88111 from EPSRC (UK), and the Ukrainian State Fund for Fundamental Studies Support, grant No. 2.3/406.

*)E-mail: yuri@blue.lci.kent.edu

[1] K. Ichimura, Y. Suzuki, T. Seki, A. Hosoki, and K. Aoki, Langmuir **4**, 1214 (1988).
[2] K. Ichimura, T. Seki, Y. Kawanishi, Y. Suzuki, M. Sakuragi, and T. Tamaki, in *Photo-Reactive Materials for Ultrahigh Density Optical Memory*, Elsevier, Amsterdam (1994), p. 55.
[3] W. Gibbons, P. Shannon, Shao-Tang Sun, and B. Swetlin, Nature (London) **351**, 49 (1991).
[4] A. Dyadyusha, T. Marusii, Yu. Reznikov, V. Reshetnyak, and A. Khizhnyak, JETP Lett. **56**, 17 (1992).
[5] Y. Iimura, J. Kusano, Sh. Kobajashi, Y. Aoyagi, and T. Sugano, Jpn. J. Appl. Phys., Part 1 **32**, 93 (1993).
[6] P. Shannon, W. Gibbons, and Shao-Tang Sun, Nature (London) **368**, 532 (1994).
[7] K. Ichimura, Y. Hauashi, H. Akiama, and T. Ikeda, Appl. Phys. Lett. **63**, 449 (1993).
[8] A. Dyadyusha, V. Reshetnyak, and Yu. Reznikov, Mol. Mater. **5**, 183 (1995).
[9] S. T. San, W. Gibbon, and P. Shannon, Liq. Cryst. **12**, 869 (1992).
[10] D. Voloshchenko, A. Khizhnyak, Yu. Reznikov, and V. Reshetnyak, Jpn. J. Appl. Phys., Part 1 **34**, 566 (1995).
[11] P. G. de Gennes and J. Prost. *The Physics of Liquid Crystals*, Clarendon Press, Oxford (1993).
[12] T. Marusii, Yu. Reznikov, V. Reshetnyak, M. Soskin, and A. Khizhnyak, Zh. Éksp. Teor. Fiz. **91**, 851 (1986) [Sov. Phys. JETP **64**, 502 (1986)].
[13] Yu. Reznikov, V. Reshetnyak, M. Soskin, and A. Khizhnyak, Ukr. Fiz. Zh. **29**, 1269 (1984).
[14] I-C. Khoo and S-T. Wu, *Optics and Nonlinear Optics of Liquid Crystals*, World Scientific, New York (1993).
[15] G. Barbero, N. V. Madhusudana, J. F. Palierne, and G. Durand, Phys. Lett. A **103**, 385 (1984).
[16] D. Subacius, V. Pergamenshchik, and O. Lavrentovich, Appl. Phys. Lett. **67**, 212 (1995).
[17] P. Rochon, J. Gosselin, A. Natansohn, and S. Xie, Appl. Phys. Lett. **60**, 4 (1992).
[18] Ph. Jones, W. Jones, and G. Williams, J. Chem. Soc. Faraday Trans. **86**, 1013 (1990).
[19] B. Zel'dovich and N. Tabirian, Zh. Éksp. Teor. Fiz. **82**, 1126 (1982) [Sov. Phys. JETP **55**, 656 (1982).]
[20] J. Guillet, *Polymer photophysics and photochemistry*, Cambridge University Press (1985).
[21] W. Urbach, H. Hervet, and F. Rondelez, J. Chem. Phys. **83**, 1877 (1985).

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

Magnetohydrodynamics of a viscous spherical layer rotating in a strong potential field

S. V. Starchenko

Borok Geophysical Observatory, Institute of Earth Physics, Russian Academy of Sciences, 152742 Borok, Yaroslavl Region, Russia

(Submitted 15 May 1997)
Zh. Éksp. Teor. Fiz. **112**, 2056–2078 (December 1997)

An analytic solution is given for classical magnetohydrodynamic (MHD) problem of almost rigid-body rotation of a viscous, conducting spherical layer of liquid in an axisymmetric potential magnetic field. Large-scale flows bounded by rigid spheres are described for the first time in a new approximation. Two problems are solved: (1) in which both spheres are insulators and (2) in which the outer sphere is an insulator and the inner sphere a conductor. Axially symmetric flows and azimuthal magnetic fields are maintained by a slightly faster rotation of the inner sphere. The primary regeneration takes place in the boundary and shear MHD layers. The shear layers, described here for the first time, smooth out the large gradients at the boundaries of the MHD structures encompassed by them. There is essentially no azimuthal magnetic field inside these original structures, which are bounded by potential contours tangent to the spheres. An applied constant magnetic field creates a rigid MHD structure outside an axial cylinder tangent to the inner sphere. Inside the cylinder the rotation is faster and the meridional flux depends on height. A magnetic dipole forms a structure tangent to the outer equator. Outside the structure, the rotation is also rigid-body when both spheres are insulators. When a conducting sphere is present, the liquid rotates differentially everywhere, while near the axis and inside the MHD structure, it rotates even faster than the inner sphere. The last example of a general solution is a quadrupole magnetic field. In this case, two equatorially symmetric MHD structures are formed which rotate together with the inner sphere. Outside the structures, as in the most general case, the rotation is differential, the azimuthal magnetic field falls off as the first power of the applied field, and the meridional flux falls off as the square of the field in the first problem, and as the cube in the second.
[S1063-7761(97)01112-8]

1. INTRODUCTION AND STATEMENT OF THE PROBLEM

In the middle of this century, Proudman[1] formulated a problem involving the almost rigid-body rotation of a viscous incompressible liquid between concentric spheres. He linearized the problem, using the smallness of the difference $\Delta\Omega$ between the angular velocities of the rotating spheres, which is characterized by the very small parameter

$$\epsilon \equiv \frac{\Delta\Omega}{\Omega} \ll 1. \tag{1}$$

The large-scale flow was described analytically[1] in the approximation of rapid rotation ($\delta \ll 1$) with a thin viscous Ekman layer, where

$$\delta \equiv \frac{\sqrt{\nu/\Omega}}{L}. \tag{2}$$

Here ν is the kinematic viscosity, Ω is the angular rotation velocity of the outer sphere, and L is the thickness of the spherical layer of liquid illustrated in Fig. 1.

The large-scale flow consists of a differential rotation plus a weak ($O(\delta)$) meridional flow, which depend only on the inside radius s of an axial tangential cylinder that is tangent to the inner sphere. All the liquid outside the tangential cylinder rotates as a rigid body together with the outer sphere. We associate a cylindrical coordinate system (s,φ,z) with the latter, where distances are measured in units of L. The small-scale flows are concentrated in the Ekman boundary layer and in the free shear layer, which ensures a return meridional flow and smoothes out the large gradients at the tangential cylinder. The structure of this layer was determined by Stewartson[2] only ten years after Proudman's paper[1] had been published.

Much later, Ruzmaikin[3,4] pointed out the importance of solving the Proudman-Stewartson problem for the earth's dynamo, and Starchenko[5] laid the foundations for solving the associated, more complicated, magnetohydrodynamic (MHD) problem analytically. Since then this problem has been at the center of attention by researchers, who use numerous models for the hydromagnetic convection, dynamo, and interiors of planets.[6–9] For seven years, repeated but unsuccessful attempts were made to find the most general analytic solution of this MHD problem.[10,11] Only very recently[12] has it become clear that one must begin with a proper statement of the problem.

A first step in this direction had already been taken in Hollerbach's numerical model.[13] He examined the effect of a central axial magnetic dipole on a Proudman–Stewartson system (see Fig. 1) in which a liquid with constant conductivity σ and density ρ is subjected to the strong influence of a weak potential magnetic field. The effect of even a large azimuthal magnetic field on this axially symmetric flow is much weaker, since the azimuthal field cannot cross any of the very thin layers that form the flow.[14–16] Unfortunately,

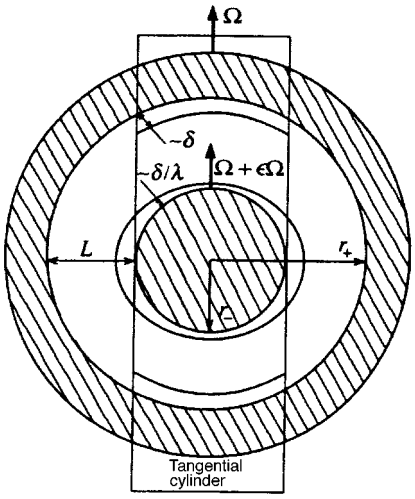


FIG. 1. Rotation of an inner sphere of radius r_- by an amount $\epsilon\Omega$ faster than an outer sphere (r_+) with a rotational velocity of Ω. Shown here are the Ekman ($\sim\delta$) and Hartmann ($\sim\delta/\lambda$) boundary layers. The thickness ($r_+ - r_-$) of the spherical layer equals L.

Hollerbach and the authors of a subsequent important analytic paper[12] did not take the next step in stating the problem: they did not formulate the approximation of a strong potential field, but restricted themselves only to examining the effect of very weak fields. Even these weak fields, which were, therefore, of no practical interest, ended up changing the Proudman–Stewartson flow in a fundamental way. It is therefore logical to formulate the strong-field approximation in this paper in order to solve the MHD problem analytically.

Thus, we assume that the system shown in Fig. 1 lies in a strong potential magnetic field **B** measured in terms of the dimensional magnitude B_p. Accordingly, the principal dimensionless parameter of the problem is very large:

$$\lambda = \frac{B_p}{\sqrt{\rho\Omega/\sigma}} \gg 1. \tag{3}$$

This approach makes it possible in practice to avoid restricting the other basic parameters of the problem, i.e., to assume that we can take arbitrary δ and

$$R_m = \mu\sigma\Omega L^2, \tag{4}$$

where R_m is the magnetic Reynolds number and μ is the magnetic permeability of the liquid, usually taken to be equal to the vacuum permeability μ_0 in SI units.

It is important to note that in contrast to earlier work (including numerical calculations), we do not limit ourselves to the approximation of rapid rotation ($\delta \ll 1$) or to large or small Reynolds numbers R_m. Thus, the results of this paper do not just apply to geophysics and the theory of the MHD dynamo. They can be used in astrophysics to model the magnetoactive interiors of the planets, sun, and stars. It is interesting to compare our results with corresponding numerical simulations. Furthermore, the theory discussed here can be used both in setting up experiments and for comparison with laboratory experiments that have been conducted and are being conducted on extremely diverse liquid systems rotating in magnetic fields.

Let us use their solenoidal and axial symmetry properties to represent the dimensional velocity **v** and magnetic induction **b** vectors in a standard form that takes into account the subsequent linearization of the problem:

$$\mathbf{v} = \epsilon L\Omega([\nabla(\mathbf{e}_\varphi s^{-1}\chi)] + s\zeta\mathbf{e}_\varphi), \tag{5}$$

and

$$\mathbf{b} = B_p(\mathbf{B} + \epsilon R_m[\nabla(\mathbf{e}_\varphi s^{-1}\Psi)] + \epsilon R_m sJ\mathbf{e}_\varphi), \tag{6}$$

where the dimensionless quantities $\mathbf{e}_\varphi$, sJ, and ζ are, respectively, the unit vector in the direction of increasing φ, the azimuthal magnetic field, and the local angular rotation velocity, while Φ and χ are the dimensionless meridional flux functions of the magnetic induction and velocity. The magnetic field (6) and flow (5) vectors obey the complete system of MHD equations linearized in the small parameter ε (for details of the derivation, see Ref. 12):

$$2\frac{\partial\chi}{\partial z} + \delta^2 s\mathscr{D}(s\zeta) + \lambda^2\mathbf{B}\cdot\nabla(s^2 J) \equiv [A] = 0, \tag{7}$$

$$2\frac{\partial\zeta}{\partial z} - \delta^2 s^{-1}\mathscr{D}^2(s^{-1}\chi) + \lambda^2\mathbf{B}\cdot\boldsymbol{\nabla}(s^{-2}\mathbf{B}\cdot\nabla\chi) \equiv [B] = 0, \tag{8}$$

$$s^{-1}\mathscr{D}(sJ) + \mathbf{B}\cdot\nabla\zeta \equiv [C] = 0, \tag{9}$$

and

$$s\mathscr{D}(s^{-1}\Psi) + \mathbf{B}\cdot\nabla\chi \equiv [D] = 0, \tag{10}$$

where the modified Laplacian is given by $\mathscr{D} = \Delta = s^{-2}$.

The boundary conditions for attachment at the rigid boundaries of the inner and outer spheres with radii r_- and r_+, respectively, are

$$\chi = \frac{\partial\chi}{\partial r} = 0 \quad \text{for} \quad r = r_-, r_+, \tag{11}$$

$$\zeta = 1 \quad \text{for} \quad r = r_-, \tag{12}$$

and

$$\zeta = 0 \quad \text{for} \quad r = r_+. \tag{13}$$

The boundary condition corresponding to the absence of an azimuthal magnetic field at an insulating sphere has the form

$$J = 0. \tag{14}$$

At the boundary of a conducting sphere the following continuity conditions must be satisfied:

$$J = j, \quad \frac{\partial(r^2 J)}{\partial r} = H\frac{\partial(r^2 j)}{\partial r}, \tag{15}$$

where H is the ratio of the conductivity of the sphere to the conductivity σ of the liquid, while j is the azimuthal magnetic field inside the rigid conducting sphere. The equation for this field can be derived from Eq. (9), and has the form

$$\mathscr{D}(sj) = 0. \tag{16}$$

We end this section by stating the two very different MHD problems which are of greatest practical interest.

Problem 1: insulator–insulator. Both spheres are insulators. We seek ζ, J, and χ in Eqs. (7)–(9), using the boundary conditions (11)–(14) for both spheres.

Problem 2: insulator–conductor. The inner sphere is an insulator, and the outer a conductor. We seek ζ, J, χ, and j in Eqs. (7)–(9) and (16), using the boundary conditions (11)–(13) for both spheres, and then apply Eq. (14) for the outer sphere and Eq. (15) for the inner sphere.

After the function χ has been found, it is easy to determine the meridional magnetic field Ψ using Eq. (10) by adding the appropriate boundary conditions. Since the actual meridional field $[\nabla(s^{-1}\Psi \mathbf{e}_\varphi)]$ is essentially indistinguishable against the background of the applied potential field $\mathbf{B}$, we shall not study Ψ here.

2. INSULATOR–INSULATOR: ASYMPTOTES AND STRUCTURE

Here we describe the principal asymptotes and the corresponding structures of the general analytic solution to the first MHD problem stated above.

In the approximation of a strong potential field (3), we have a small parameter λ^{-2} in the higher order derivatives in the initial hydromagnetic equations (7)–(9). This makes it possible to reduce the order of the system substantially, taking into account the corresponding viscous effects, which are $O(\delta^2)$, only within the narrow boundary and shear MHD layers. The shear layers are described in the next section. The Hartmann–Ekman boundary layer problem has already been solved by Loper[17,18] and Starchenko[5] in the rapid rotation approximation. In the following we solve the analogous problem in the approximation of a strong applied field (3).

2.1. Boundary layers and asymptotic equations

The radial derivative is perpendicular to the thin boundary layer. Thus, here we can neglect the other derivatives, writing down the original system of Eqs. (7)–(9) in a spherical coordinate system (r,ϑ). Integrating Eq. (9), inside the layer we obtain

$$\frac{\partial J_s}{\partial r} = -B_r(\zeta_s - \zeta). \tag{17}$$

Here and in the following the subscript s denotes small-scale quantities concentrated within narrow layers, and r denotes radial components. We take all quantities without a subscript s to be large-scale quantities that are essentially constant perpendicular to a layer. Integrating Eqs. (7) and (8) perpendicular to the radial layer under consideration and using Eq. (17), we obtain

$$\frac{2\cos\vartheta}{s^2}\frac{\partial \chi_s}{\partial r} + \delta^2 \frac{\partial^2 \zeta_s}{\partial r^2} - \lambda^2 B_r^2(\zeta_s - \zeta) = 0, \tag{18}$$

and

$$2\cos\vartheta s^2(\zeta_s - \zeta) - \delta^2 \frac{\partial^3 \chi_s}{\partial r^3} + \lambda^2 B_r^2 \frac{\partial \chi_s}{\partial r} = 0. \tag{19}$$

The solution of Eqs. (18) and (19) that satisfies the attachment boundary conditions (11)–(13) without the nonpenetration condition $\chi_s = 0$ can be written conveniently in the complex form

$$\zeta_s + \frac{i}{s^2}\frac{\partial \chi_s}{\partial r} = \zeta + (\zeta_\pm - \zeta) \times \exp\left(\pm \frac{r - r_\pm}{\delta} \sqrt{\lambda^2 B_r^2 + 2i\cos\vartheta} \right). \tag{20}$$

Here the upper sign $(+)$ corresponds to the outer sphere $r = r_+$, where $\zeta_+ = 0$, and the lower sign $(-)$, to the inner sphere $r = r_-$, where $\zeta_- = 1$.

Now, integrating Eqs. (20) and (17) and applying the nonpenetration boundary condition together with condition (14), at the outer boundary of the thin MHD layer we obtain

$$\chi = \pm \delta s^2(\zeta_\pm - \zeta)\sqrt{\frac{\sqrt{\lambda^4 B_r^4 + 4\cos^2\vartheta} - \lambda^2 B_r^2}{2\lambda^4 B_r^4 + 8\cos^2\vartheta}}, \tag{21}$$

and

$$J = \pm \delta B_r(\zeta_\pm - \zeta)\sqrt{\frac{\sqrt{\lambda^4 B_r^4 + 4\cos^2\vartheta} + \lambda^2 B_r^2}{2\lambda^4 B_r^4 + 8\cos^2\vartheta}}. \tag{22}$$

This is the boundary condition for the large-scale quantities outside the Hartmann–Ekman boundary layers solved here. Expanding Eqs. (21) and (22) as series in the small parameter $\lambda^{-1} \ll 1$, we obtain the asymptotic conditions at the boundaries of the Hartmann layer:

$$\chi = -\lambda^{-3}\delta s^2 \left| \frac{\cos\vartheta}{B_r^3} \right| \zeta + O(\lambda^{-7}),$$
$$J = -\lambda^{-1}\delta S\zeta + O(\lambda^{-5}) \tag{23}$$

for the outer sphere $r = r_+$ and

$$\chi = \lambda^{-3}\delta(\zeta - 1)s^2 \left| \frac{\cos\vartheta}{B_r^3} \right| + O(\lambda^{-7}),$$
$$J = \lambda^{-1}\delta S(\zeta - 1) + O(\lambda^{-5}) \tag{24}$$

for the inner sphere $r = r_-$. Here $S = \pm 1$ is the sign of the radial component B_r of the applied potential field. Thus, we have obtained a total of four boundary conditions in Eqs. (23) and (24), instead of eight in Eqs. (11)–(14). Accordingly, the order of the system of equations for the large-scale fields should be reduced by a factor of two compared to the original system. This reduction is achieved by dropping both the kinematic viscosity $(\delta^2 s \mathscr{D}(s\zeta)$ in Eq. (7) and $\delta^2 s^{-1}\mathscr{D}^2(\chi/s)$ in Eq. (8)) and the magnetic viscosity $(s^{-1}\mathscr{D}(sJ)$ in Eq. (9)) in the first approximation. The validity of leaving out these terms will be confirmed once again below when we obtain a more general solution of the problem.

The corresponding formal expansion in terms of the small parameter λ^{-1},

$$\zeta = \zeta_0 + \lambda^{-1}\zeta_1 + \ldots, \quad J = \lambda^{-1}J_0 + \lambda^{-2}J_1 + \ldots,$$
$$\chi = \lambda^{-2}\chi_0 + \lambda^{-3}\chi_1 + \ldots, \tag{25}$$

enables us then to find the solution, not only in the first-order (subscript 0), but also in the second-order (subscript 1) approximations. The system of equations for the first approximation is obtained by substituting Eq. (25) into Eqs. (7)–(9) and equating the coefficients for the leading powers of $\lambda \gg 1$:

$$\mathbf{B}\cdot\nabla(s^2 J_0)=0, \quad 2\frac{\partial \zeta_0}{\partial z}+\mathbf{B}\cdot\nabla(s^{-2}\mathbf{B}\cdot\nabla\chi_0)=0,$$

$$\mathbf{B}\cdot\nabla\zeta_0=0. \tag{26}$$

This fourth-order system should satisfy the following boundary conditions:

$$J_0=-\delta S\zeta_0, \quad \chi_0=0 \quad \text{for} \quad r=r_+, \tag{27}$$

and

$$J_0=\delta S(\zeta_0-1), \quad \chi_0=0 \quad \text{for} \quad r=r_-, \tag{28}$$

derived from Eqs. (23) and (24), respectively. The following system of equations for the second approximation (29), together with its boundary conditions (30), can be derived similarly:

$$\delta^2 s\mathscr{D}(s\zeta_0)+\mathbf{B}\cdot\nabla(s^2 J_1)=0, \tag{29a}$$

$$2\frac{\partial \zeta_1}{\partial z}+\mathbf{B}\cdot\nabla(s^{-2}\mathbf{B}_0\cdot\nabla\chi_1)=0, \tag{29b}$$

$$s^{-1}\mathscr{D}(sJ_0)+\mathbf{B}\cdot\nabla\zeta_1=0, \tag{29c}$$

and

$$J_1=-\delta S\zeta_1, \quad \chi_1=-\delta s^2\left|\frac{\cos\vartheta}{B_r^3}\right|\zeta_0 \quad \text{for} \quad r=r_+, \tag{30a}$$

$$J_1=\delta S\zeta_1, \quad \chi_1=\delta s^2\left|\frac{\cos\vartheta}{B_r^3}\right|(\zeta_0-1) \quad \text{for} \quad r=r_-. \tag{30b}$$

The first and second of Eqs. (26) form the simple system

$$\mathbf{B}\cdot\nabla(s^2 J_0)=\mathbf{B}\cdot\nabla\zeta_0=0,$$

which is complete for finding the large-scale quantities J and ζ using just the first boundary conditions in Eqs. (27) and (28). The weak meridional flux $\chi=O(\lambda^{-2})$ can be found later using the remaining boundary conditions in Eqs. (27) and (28). In the degenerate case ($\zeta_0=\text{const}$), the rotation is rigid-body everywhere, and then $\chi=O(\lambda^{-3}\delta)$ can be found using the second approximation Eqs. (26b) with the second boundary conditions from Eqs. (30).

2.2. General solutions inside and outside the ℓ curves

The general solution of the system of Eqs. (26) without the middle equation can conveniently be written in the form

$$\zeta_0=G(V), \quad J_0=s^{-2}F(V), \tag{31}$$

where G and F are free functions determined by the boundary conditions, while a given flux function V determines the applied potential field $\mathbf{B}=[\nabla(V\mathbf{e}_\varphi/s)]$. It follows from this solution that the characteristics of the systems of Eqs. (26) and (30) are parallel to the applied magnetic field $\mathbf{B}$ or, in other words, lie on the $V=\text{const}$ contours. If the contours intersect both spherical boundaries, then different boundary conditions (27) and (28) apply to the different spheres (see points A and B in Fig. 2). Alternatively, either a single boundary condition is applied twice to different points of a single sphere (C and D, E and F in Fig. 2), or the contours do not intersect the boundaries at all (curve ℓ No. 3 in Fig. 2).

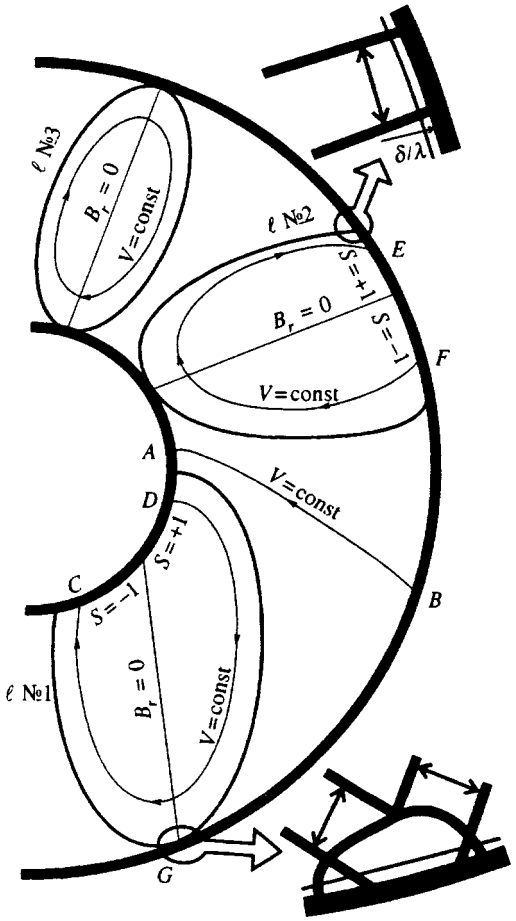


FIG. 2. The three topologically possible positions of the ℓ curves (No. 1, 2, and 3) encompassing the MHD structures in the spherical layer illustrated in Fig. 1. The magnified insets show the structure of the MHD layers: (top) a shear layer of thickness $\sqrt{\delta/\lambda}$ indicated by the arrows, and (bottom) a hilly layer intersecting two shear layers.

Thus, inside the spherical layer there are two types of $V=\text{const}$ contours separated by a special curve that is tangent to only one point on one of the two spheres. Let us call it curve ℓ. In all, three topologically different ℓ curves can exist in our spherical layer, as shown qualitatively in Fig. 2. Let us begin with curve No. 1, which is most widespread in astrophysical applications. This ℓ curve is tangent to the outer sphere in the lower part of the figure. This type of curve intersects two points in the inner sphere, forming a closed region with it. Inside this region, the solution (31), with the first boundary conditions from Eq. (28) applied twice, has the form

$$J_0=0, \quad \zeta_0=1, \tag{32}$$

since the sign $S=\pm 1$ of the function is different at the boundary points C and D joined by the contour $V=\text{const}$ in

the lower part of Fig. 2. Thus, the region inside the curve ℓ rotates together with the inner sphere and ''expels'' the azimuthal field J_0 from there.

Curve ℓ No. 2 in the middle of Fig. 2 is typical of laboratory systems. This curve is tangent to one point on the inner sphere. It intersects two points on the outer sphere, with which it forms another closed region inside which, by analogy with the previous case, we obtain

$$J_0=0, \quad \zeta_0=0. \tag{33}$$

Thus, the region rotates together with the outer sphere and also ''expels'' the field J_0.

The last of the possible ℓ curves, No. 3, is shown in the upper part of Fig. 2. It is tangent to both spheres, so that it is extremely unlikely in any practical applications. This curve is closed and, by analogy with the previous cases, inside it we have

$$J_0=0, \quad \zeta_0=\zeta_3=\text{const}, \tag{34}$$

where the constant $0\leqslant\zeta_3\leqslant 1$ can be determined only after solving the conjugate problem for the original MHD layer along the curve ℓ (see below).

Outside any of the above ℓ curves, generally speaking we obtain nontrivial solutions J_0 and ζ_0, since we apply different boundary conditions, (27) and (28), at the different spheres (see points A and B in Fig. 2). We write these solutions in a form convenient for later practical applications:

$$\zeta_0=[1+S_+(s_+/s_-)^2/S_-]^{-1}, \tag{35}$$

and

$$J_0=-\delta S_+(s_+/s)^2\zeta_0. \tag{36}$$

Here the distances from the rotation axis $s_\pm(V)$ are measured to the corresponding points of intersection of the $V=\text{const}$ profiles with the outer $(+)$ and inner $(-)$ spheres. The functions $S_\pm(V)$ which determine the signs of the component B_r are similarly associated with the spheres.

Having obtained the general solution (32)–(36), let us estimate the range over which it is valid, and thereby confirm the validity of our initial assumptions. To order of magnitude, for the nondegenerate case in (λ,δ) parameter space, Eqs. (25), (35), and (36) imply

$$J=O(\delta/\lambda), \quad \zeta=O(1), \quad \chi=O(1/\lambda^2). \tag{37}$$

Substituting these estimates in the original system of Eqs. (7)–(9), we obtain the inequalities

$$\lambda^{-2}+\delta^2\ll\lambda\delta, \quad \lambda^{-2}\delta^2\ll 1, \quad \lambda^{-1}\delta\ll 1,$$

which are the conditions under which the ''viscous'' terms with the operator $\mathscr{D}$ in the original Eqs. (7)–(9) can be neglected here. All these inequalities are satisfied if for arbitrary δ we have

$$\lambda\gg\max(1,\delta^{-1/3}), \tag{38}$$

which confirms our assumptions in the approximation of a strong applied potential field (3) in the nondegenerate case. In the degenerate case, the angular rotation velocity obtained from Eq. (35) is constant, i.e., $\zeta_0=\text{const}$. Thus, the estimate of χ in Eq. (37) can be replaced, and the range of validity of Eq. (38) expands:

$$\chi=O(\delta\lambda^{-3}), \quad \lambda\gg 1. \tag{39}$$

3. ORIGINAL FREE SHEAR MHD LAYERS

Usually this general nontrivial solution, Eqs. (35) and (36), differs substantially from any of the trivial solutions, Eqs. (32)–(34), inside all the possible ℓ curves (see Fig. 2 and Sec. 4 of this paper); that is, ζ_0 and J_0 are discontinuous on the curve ℓ. These discontinuities are smoothed out by viscosity in the free shear MHD layer along the ℓ curve. As far as we know, no studies of this sort of MHD layer have been undertaken before.

Let us derive the basic equations for this type of original layer moving along a curve ℓ determined by a particular value of the flux function $V(r,\vartheta)=V_\ell=\text{const}$. The layer is best described by an orthogonal coordinate system (V,U) related to the applied potential field

$$\mathbf{B}=[\nabla(V\mathbf{e}_\varphi/s)]=\nabla U. \tag{40}$$

On the ℓ curve distinguished by a layer, the coordinate $V=V_\ell=\text{const}$, while U varies. Outside and inside ℓ, the angular velocity ζ and the transformed azimuthal field

$$P\equiv(\lambda/\delta)s^2J \tag{41}$$

are constants of order unity. Outside the curve ℓ on the outer boundary of our free layer, Eqs. (34) and (35) yield boundary conditions for the small-scale quantities ζ_s and P_s concentrated in the layer:

$$\zeta_s=\zeta(V_l)\equiv\zeta_l=\text{const}, \quad P_s=-s_-^2(\zeta_1-1)=\text{const}. \tag{42}$$

Inside the curve ℓ, on the inner boundary of the layer Eqs. (32)–(34) yield other boundary conditions for the small-scale quantities in the layer:

$$P_s=0, \quad \zeta_s=(1,0,\zeta_3). \tag{43}$$

Comparing Eqs. (42) and (43), we can again convince ourselves of the discontinuity in the large-scale quantities on ℓ. The discontinuity takes place along the coordinate V, so that for describing small-scale variables within the layer, we must add the appropriate ''viscous'' terms from Eqs. (9) and (7) to the asymptotic first-order system (26). As a result, we obtain

$$\frac{\delta}{\lambda}(\nabla V)^2\frac{\partial^2 P_s}{\partial V^2}+s^2(\nabla U)^2\frac{\partial\zeta_s}{\partial U}=0,$$

$$\frac{\delta}{\lambda}(\nabla V)^2\frac{\partial^2\zeta_s}{\partial V^2}+s^{-2}(\nabla U)^2\frac{\partial P_s}{\partial U}=0,$$

where $(\nabla V)^2\partial^2/\partial V^2$ is the leading term in the expansion of the modified Laplacian $\mathscr{D}=\Delta-s^{-2}$, which describes the magnetic and kinematic viscosities, respectively, in the first and second equations. Using a variable ''stretched out'' perpendicular to the layer,

$$x=(V-V_l)\sqrt{\frac{\lambda}{\delta}}, \tag{44}$$

we obtain the final system of equations in the narrow layer from the above equations:

$$\frac{\partial^2 P_s}{\partial x^2}+F_Z\frac{\partial \zeta_s}{\partial U}=0,\quad \frac{\partial^2 \zeta_s}{\partial x^2}+F_P\frac{\partial P_s}{\partial U}=0. \tag{45}$$

Here the functions $F_Z(U)=s^2(\nabla U)^2/(\nabla V)^2$ and $F_P(U)=s^{-2}(\nabla U)^2/(\nabla V)^2$ are given, and depend only on U for $V=V_{\mathscr{l}}=\text{const}$.

The system of Eqs. (45) for the layer must satisfy the boundary conditions (43) and (42) as x approaches $+\infty$ and $-\infty$, respectively. In addition, some boundary conditions must be satisfied at either sphere, i.e., at both ends of the layer in the direction of the coordinate U. To determine these boundary conditions, we compare the characteristic thickness of our free layer that follows from Eqs. (44) and (45),

$$\delta_s=\sqrt{\frac{\delta}{\lambda}}, \tag{46}$$

with the known thickness δ/λ of the Hartmann layer considered above. Obviously $\delta_s \gg \delta/2$, under the conditions considered here, so that we are right to use condition (42) or (43) as a boundary condition where our layer intersects one of the spheres (e.g., in a row along with the points E, F, D, and C in Fig. 2). For the small-scale variables the Hartmann boundary conditions take the form

$$P_s=\begin{cases} s_-^2(\zeta_s-1), & r=r_-,\\ -s_+^2\zeta_s, & r=r_+, \end{cases} \tag{47}$$

and the corresponding structure of the MHD layers is illustrated qualitatively in the magnified inset in the upper part of Fig. 2 next to the point E.

A considerably more complicated situation arises at the edge of the layer, where the curve $\mathscr{l}$ is just tangent to the sphere (see point G of Fig. 2). Here the radial component of the applied field $B_r=0$, so that the Hartmann layer degenerates into a singular hilly layer, which is usually thicker than the Hartmann layer (see Refs. 2 and 12). If this hilly layer is even thinner than the free layer examined above (see Eq. (46)), then we are justified in using Eq. (47) at this end of the layer and complete the statement of the problem for our free layer. Otherwise, we have to solve the problem for a hilly layer from the beginning. The situation in which the hilly layer is thicker than the free layer is illustrated in the lower inset to Fig. 2 near the point G.

Unfortunately, there is no general analytic solution of the system of Eqs. (45) for the layer, although that system is rather simple. It can be solved easily when the applied magnetic field is given. Even then, however, we have to overcome the additional difficulties caused by the hilly layer, as Stewartson[2] did, but only ten years after Proudman's paper.[1] We restrict ourselves to a detailed description of only the large-scale fields, which, as can be seen from the above discussion, can be done independently of the free MHD layers.

4. INSULATOR–CONDUCTOR: ASYMPTOTES AND STRUCTURE

In this section we present the general asymptotic solution of the second problem stated in Sec. 1 and referred to as the ''insulator–conductor'' problem. The search for a solution is analogous in many respects to that in Sec. 2 for the first, or ''insulator–insulator'' problem. In this section, we therefore concentrate our attention only on those features that significantly distinguish the solution of the second problem from that of the first.

4.1. Boundary layers and asymptotic equations

The derivation of the modified boundary conditions is the same as in Sec. 2.1 up to Eq. (21). We then ultimately have the same asymptotic boundary conditions at the outer insulating sphere as in Eq. (23). At an inner conducting sphere only the first modified attachment boundary condition from Eq. (24) is retained unchanged. The magnitude of the large-scale magnetic field J_s at the inner spherical boundary (this is the right-hand side of Eq. (22)) is taken with the upper sign (+). Accordingly, the initial boundary conditions (15) with Eq. (17) take the following form for our large-scale quantities (see Eq. (16) for the azimuthal field j inside this sphere):

$$J+B_r\delta(1-\zeta)\,\mathrm{Re}\,\frac{1}{\sqrt{\lambda^2 B_r^2+2i\,\cos\,\vartheta}}=j,$$

$$2rj-r^2B_r(1-\zeta)=H\,\frac{\partial(r^2 j)}{\partial r}. \tag{48}$$

Based on the obvious order-of-magnitude estimates $j=O(\partial j/\partial r)\ll O(1)=\zeta$ and all the boundary conditions obtained above, we are justified in seeking the large-scale quantities in the form of the expansions

$$\zeta=1+\delta\lambda^{-1}\zeta_0+\ldots,\quad (J,j)=\delta\lambda^{-1}(J_0,j_0)+\ldots,$$

$$\chi=\delta\lambda^{-3}\chi_0+\ldots \tag{49}$$

It is clear that the whole liquid tends to rotate together with the inner sphere: $\zeta\approx 1$. We therefore limit ourselves to the first asymptotic approximation (subscript 0), which is essentially an analog of the second approximation (29) examined in Sec. 2. On deriving the corresponding asymptotic boundary conditions, at the insulating outer sphere ($r=r_+$) we obtain

$$J_0=-\frac{B_r}{|B_r|},\quad \chi_0=-s^2\left|\frac{\cos\,\vartheta}{B_r^3}\right|. \tag{50}$$

At the conducting inner sphere ($r=r_-$) the boundary conditions have the form

$$J_0=j_0,\quad B_r\zeta_0=H\,\frac{\partial j_0}{\partial r}+2r(H-1)j,\quad \chi_0=0. \tag{51}$$

The system of first-order equations differs substantially from the systems for the first problem (see Eqs. (26) and (29)), and it convenient to write it in more detail:

$$\mathbf{B}\cdot\nabla(s^2J_0)=0, \tag{52}$$

$$s^{-1}\mathscr{D}(sJ_0)+\mathbf{B}\cdot\nabla\zeta_0=0, \tag{53}$$

and

$$2\frac{\partial\zeta_0}{\partial z}+\mathbf{B}\cdot\nabla(s^{-2}\mathbf{B}\cdot\nabla\chi_0)=0. \tag{54}$$

4.2. Basic MHD structure and derivation of a solution

For the applied magnetic field (40) being considered here, the general solution of Eq. (52) for the azimuthal field J_0 is taken from Eq. (31). In the nontrivial region, where the contours $V=\text{const}$ intersect each of the spheres once, the arbitrary function F can be found from the first boundary condition in Eq. (50) and, accordingly, the general solution for the azimuthal field has the form

$$J_0=-(s_+/s)^2S_+, \tag{55}$$

where $S_+(V)=B_r/|B_r|$ for $r=r_+$ and $s_+=s(r=r_+,V)$, as in Sec. 2.

Besides the nontrivial region, where J_0 has the form (55), there are also trivial regions with $J_0=0$. They are bounded by part of the spherical boundary and/or the three possible types of ℓ curves. Here we repeat the corresponding results (33) and (34) from Sec. 2 for curves of the second and third types (see Fig. 2). It is merely necessary to replace ζ_0 by ζ in Eqs. (33) and (34) for the second MHD problem.

The significant difference in the first type of ℓ curve lies in the fact that inside the region bounded by it, ζ_0 and χ_0 can be found from Eqs. (53) and (54) with the boundary conditions (52) and with zero azimuthal fields $J_0=0$ and $j(r=r_-,\theta)=0$. The situation in which only ℓ curves of the first type exist is the most widespread in practical applications (see Sec. 6). In this section, we therefore restrict ourselves to finding a general solution only for this situation. In other cases, a solution can be found in similar fashion (see, for example, the following section for curves of the second type).

Thus, in the region outside the type-one ℓ curves, we have already obtained the azimuthal magnetic field (55). This field is zero within the region bounded by an arbitrary curve. In this way we obtain boundary conditions for the field j_0 that are easily found from Eq. (16) and take the form $\mathscr{D}(sj_0)=0$. Now we have everything we need to determine ζ_0 from Eq. (53) and the average boundary condition given by Eq. (51). We conclude the construction of a solution by finding χ_0 using Eq. (54), and employing the last boundary conditions from Eqs. (50) and (51).

We conclude this section by estimating the range of validity of this general solution in the (λ,δ) plane. The above discussion implies that

$$\zeta=1+O(\delta\lambda^{-1}),\quad J=O(\delta\lambda^{-1})=j,\quad \chi=O(\delta\lambda^{-3}), \tag{56}$$

By substituting these estimates in the original equations (7) and (8), we determine the conditions under which it is correct to neglect the terms that have been omitted in our asymptotic equations (52) and (54), respectively (in Eq. (53) all the original terms from Eq. (9) have been retained),

$$\delta\lambda^{-3}+\delta^2\ll\lambda\delta \quad\text{and}\quad \lambda^{-3}\delta^3\ll\lambda^{-1}\delta.$$

Finally, we obtain a range within which the second problem is correct,

$$\lambda\gg\max(1,\delta), \tag{57}$$

which is wider than that for the first problem (see Eq. (38)) and the usual choice $\delta\ll1$.

5. CONSTANT AXIAL MAGNETIC FIELD

The simplest potential magnetic field is directed along the rotation axis z and is constant. Its magnitude varies through the parameter λ, and the field itself has the simple form

$$\mathbf{B}=\mathbf{e}_z, \tag{58}$$

since we have normalized $\mathbf{B}$ to the arbitrary value B_p in Sec. 1. A magnetic field of this sort can be imposed on our system (see Fig. 1) by putting it into a solenoid, as shown in Fig. 3a. The solution is determined by a single ℓ curve of the second type (see Fig. 2) described in Secs. 2 and 4 of this paper for the insulator–insulator and insulator–conductor problems, respectively. The ℓ curve under consideration here is the generator of a tangential cylinder (see Fig. 3b).

5.1. Insulator–insulator

For problem 1 of Sec. 1, the large-scale azimuthal angular velocity ζ and magnetic field sJ can be easily derived from the general trivial (33) and nontrivial (35)–(36) solutions, and for the constant field (58) are given by (see Figs. 3b and 3c)

$$\zeta=\begin{cases}1/2, & s<r_-,\quad\text{and}\quad r>r_-,\\ 0, & s>r_-,\end{cases} \tag{59}$$

and

$$sJ=\begin{cases}-(\delta/2\lambda)s, & s<r_-\quad\text{and}\quad r>r_-,\\ 0, & s>r_-.\end{cases} \tag{60}$$

These expressions (59) and (60) for a hydromagnetic flow are even simpler than for the nonmagnetic flow described by Proudman and Stewartson[1,2] in the approximation of rapid rotation ($\delta\ll1$). Here we have not restricted the value of δ, but nevertheless have simple rigid-body rotation both inside and outside the tangential cylinder.

It is somewhat more difficult to obtain the meridional flow χ. To do this, we must solve the second-order system (29)–(30), since this case is degenerate owing to rigid-body rotation throughout, and

$$\chi_0=0. \tag{61}$$

It is easy to see that in the second approximation, the angular rotation velocity and magnetic field go to zero, i.e.,

$$J_1=0,\quad \zeta_1=0. \tag{62}$$

This greatly simplifies Eq. (29b) and the corresponding boundary conditions from Eq. (30) for the meridional flow:

$$\frac{\partial^2\chi_1}{\partial z^2}=0, \tag{63}$$

and

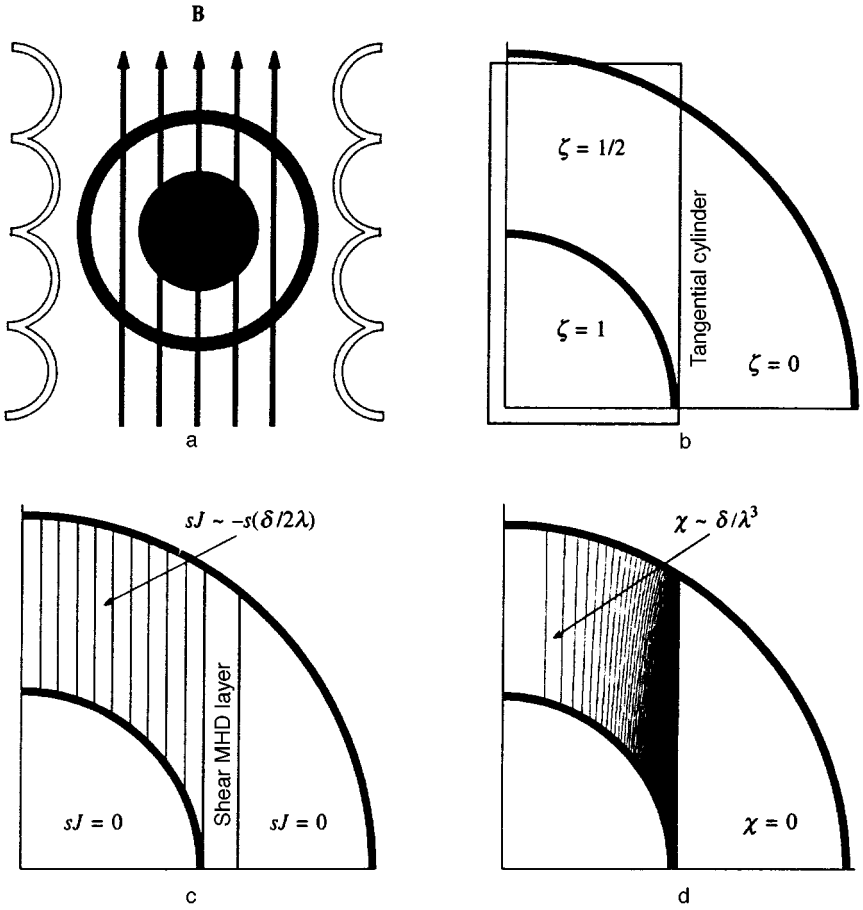


FIG. 3. (a) A constant axial field applied to the system illustrated in Fig. 1, when both spheres are insulators with $r_+=2r_-$. (b) Azimuthal rotation at angular velocity ζ. (c) Azimuthal magnetic field $sJ=-(\delta/\lambda)sJ_0$; the contours are in steps of 0.05 units of sJ_0. (d) Meridional flows $\chi=(\delta/\lambda^3)\chi_1$; contours are in steps of 0.01 units of χ_1.

$$\chi_1=\begin{cases}-\delta s^2\zeta/[1-(s/r_+)^2], & r=r_+,\\ \delta s^2(\zeta-1)/[1-(s/r_-)^2], & r=r_-\end{cases}. \tag{64}$$

Here we have used the fact that $B_r=\cos\vartheta=z/r$ for the constant magnetic field of Eq. (58). The general solution of Eq. (63) depends on two arbitrary functions of the variable s:

$$\chi_1=zF(s)+G(s). \tag{65}$$

The free functions are determined by the boundary conditions (64) for $r=r_-$ with $z=z_-(s)$ and for $r=r_+$ with $z=z_+(s)$, where the heights of the inner $(-)$ and outer $(+)$ boundaries are

$$z_-(s)=r_-\sqrt{1-(s/r_-)^2},\quad z_+(s)=r_+\sqrt{1-(s/r_+)^2}. \tag{66}$$

Substituting Eq. (65) into the boundary conditions (64), we obtain a linear system of equations for F and G:

$$z_+F+G=-\delta s^2r_+^2\zeta/z_+^2,$$
$$z_-F+G=\delta s^2r_-^2(\zeta-1)/z_-^2. \tag{67}$$

Thus, inside the tangential cylinder $(s<r_-)$, the resulting meridional flow (see Eqs. (65)–(67)) can be described by the expression

$$\chi=\frac{\delta s^2}{2\lambda^3}\,\frac{(z_--z)r_+^2z_+^{-2}-(z_+-z)r_-^2z_-^{-2}}{z_+-z_-}, \tag{68}$$

and outside the tangential cylinder $(s>r_-)$ there is no flow in either the first (61) or second approximations and $\chi_1=0$ (see Fig. 3d).

The flow (68) is singular at the tangential cylinder for $s=r_-$, when z_- goes to zero according to Eq. (66). This singularity is smoothed out by viscosity in the MHD layer generated by the layer shown in Fig. 3c. The latter is described above in Sec. 3 and serves to smooth out the discontinuities in the azimuthal rotation and magnetic field (see Eqs. (59) and (60)).

5.2. Conductor–insulator

For the second problem of Sec. 1, the general nontrivial (55) and trivial (33) solutions yield the azimuthal field in the spherical layer:

$$sJ=\begin{cases}-(\delta/\lambda)s, & s<r_-,\quad\text{and}\quad r>r_-,\\ 0, & s>r_-.\end{cases} \tag{69}$$

This field is twice the field (60) for a nonconducting inner sphere.

We now substitute Eq. (69) into the first boundary condition (51) to find $j=(\delta/\lambda)j_0$ inside the conducting sphere. For this we impose the boundary conditions

$$j_0=\begin{cases}-1, & r=r_-,\\ 0, & z=0,\end{cases} \tag{70}$$

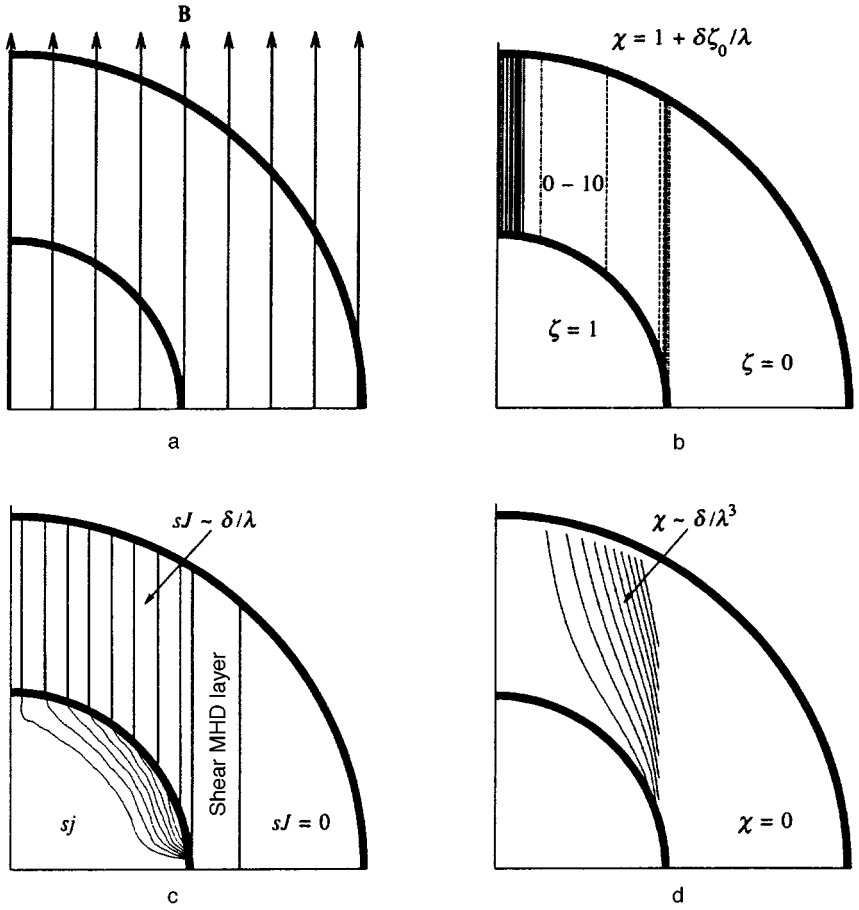


FIG. 4. As in Fig. 3, but with a conducting inner sphere.

on Eq. (16); these are singular at the point ($\theta=\pi/2$, $r=r_-$) because the azimuthal field is antisymmetric. We are unaware of any methods for solving Eq. (16) with singular boundary conditions such as Eq. (70). Thus, we have developed a special numerical solution technique (for details, see Sec. 6.2). The solution obtained by this method is shown in Fig. 4c, together with Eq. (69).

We now substitute the radial gradient of the field j at the inner sphere from the numerical solution of Eq. (16) into the second boundary condition (51), which takes the form

$$\zeta_0=\frac{H}{\cos\vartheta}\frac{\partial j_0}{\partial r} \quad \text{for} \quad r=r_-. \tag{71}$$

An angular rotation velocity $\zeta_0(s)$ which depends only on s is obtained from Eq. (54) and the result is sketched in Fig. 4b. Near the axis the liquid rotates faster than the inner sphere in a superrotation which is somewhat surprising in systems of this type, since the inner sphere is specified to rotate faster than the outer sphere.

Equation (54) for the meridional flow χ_0 is identical to Eq. (63) for χ_1; however, the boundary conditions from Eqs. (51) and (52) are simpler than Eq. (64), and take the form

$$\chi_0=\begin{cases} -s^2/[1-(s/r_+)^2], & r=r_+, \\ 0, & r=r_-. \end{cases} \tag{72}$$

Further calculations analogous to Eqs. (65)–(68), but using the condition (72), yield the resultant meridional flux (Fig. 4d) inside the tangential cylinder:

$$\chi=\frac{\delta}{\lambda^3}\frac{s^2r_+^2(z_--z)}{z_+^2(z_+-z_-)}, \quad s<r_-, \tag{73}$$

while there is no meridional flux ($\chi=0$) outside the cylinder ($s>r_-$). The resulting distributions of the field and velocity are shown in Fig. 4.

6. APPLIED DIPOLE AND QUADRUPOLE FIELDS

In this section we impose dipole and quadrupole fields (40) with sources at the center of the inner sphere on the system sketched in Fig. 1. For the dipole, the corresponding orthogonal system of coordinates (U,V) is expressed in terms of the spherical coordinates (r,ϑ) as

$$V=\frac{\sin^2\vartheta}{r}, \quad U=-\frac{\cos\vartheta}{r^2}, \tag{74}$$

and for the quadrupole, we have

$$V=3\frac{\sin^2\vartheta\cos\vartheta}{r^2}, \quad U=\frac{1-3\cos^2\vartheta}{2r^3}. \tag{75}$$

The ℓ curves generated by both the dipole and the quadrupole are curves of the first topological type (see ℓ curve No.

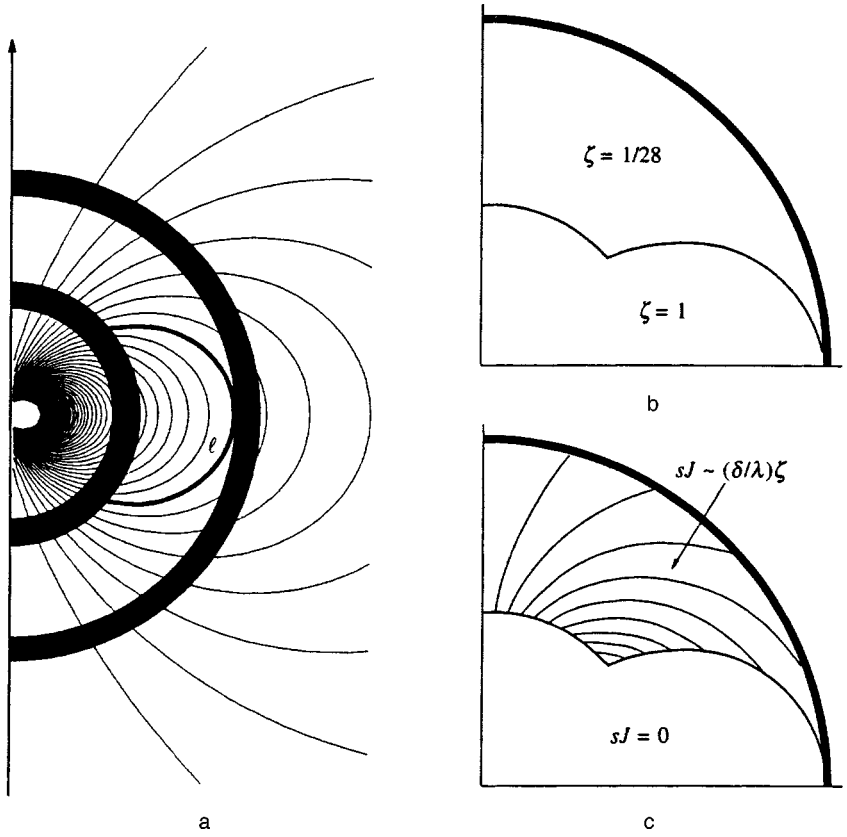


FIG. 5. An applied magnetic dipole (a) for insulating spheres with $r_+ = 3r_-$; azimuthal rotation at angular velocity ζ (b) and azimuthal magnetic field sJ (c).

1 in Fig. 2) described in Secs. 2 and 4. For the dipole, it is a single curve tangent to the equator of the outer sphere that satisfies (see Fig. 5a) the condition

$$V = \frac{\sin^2 \vartheta}{r} = 1/r_+ \equiv V_d, \tag{76}$$

while the quadrupole generates two equatorially symmetric (+ or −) ℓ curves tangent to the outer sphere that satisfy the condition (see Fig. 7a below)

$$V = 3\frac{\sin^2 \vartheta \cos \vartheta}{r^2} = \pm \frac{2}{\sqrt{3} r_+^2} \equiv \pm V_q. \tag{77}$$

6.1. Dipole in the "insulator–insulator" problem (problem 1)

For the "insulator–insulator" problem, the nontrivial solution (34)–(35) outside the dipole curve (76) together with the expansion (25) yield the large-scale quantities

$$\zeta = \frac{1}{1 + (r_+/r_-)^3}, \quad J = -\frac{\delta}{\lambda}\frac{r_-^3}{r^3}\zeta. \tag{78}$$

The azimuthal magnetic field sJ is directed opposite the rigid-body rotation ζ of Eq. (78) outside the curve ℓ (76). In the first approximation, there is no field ($J = 0$) inside ℓ, and the liquid rotates together with the inner sphere ($\zeta = 1$) in accordance with the trivial solution (32).

Therefore, we are again dealing with a degenerate situation, where the liquid rotates as a solid both inside and outside the curve ℓ. As in Sec. 5.1, a meridional flow χ shows up only in the second approximation (χ_1), and is of order δ/λ^3. The corresponding correctness condition $\lambda \gg 1$ is independent of δ. Thus, we can crudely compare our results with Hollerbach's[13] numerical calculations for small δ and $\lambda \leqslant 1$.

Outside ℓ, the angular velocity of rigid-body rotation depends (Eq. (78)) on the ratio of the cubes of the radii, $(r_+/r_-)^3$. For the geophysical situation discussed by Hollerbach, $r_+/r_- \simeq 3$, so that $\zeta \simeq 1/28$, which is consistent with his value of $\zeta \simeq 0.06$ for the maximum value $\lambda \simeq 2$ that he considered. For $\lambda \simeq 0.3$, 0.5, 1, and 2, he obtained characteristic values of the azimuthal magnetic field in the ratio 9:7:4:2, which is in very good agreement with our estimate that $J = O(1/\lambda)$ even for small λ. We are therefore justified in concluding that our asymptotes are good even at the edge, $\lambda \approx 1$, and retain important numerical relationships from the nonlinear problem.[13]

6.2. Dipole in the "insulator–conductor" problem (problem 2)

The general nontrivial (55) and trivial (32) solutions yield an azimuthal magnetic field in the spherical layer (for $r_- < r < r_+$) of

$$sJ = \begin{cases} -s(\delta/\lambda)(r_+/r)^3, & V < 1/r_+, \\ 0, & V = \sin^2 \vartheta / r > 1/r_+. \end{cases} \tag{79}$$

Using the continuity of sJ (see Eq. (51)), we now formulate the boundary conditions at $r = r_-$ for the field sj inside the conducting sphere using Eq. (16):

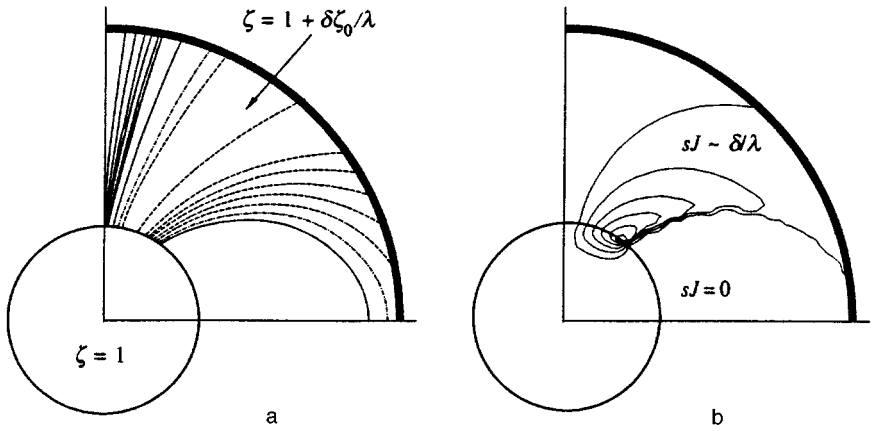


FIG. 6. The case of a magnetic dipole illustrated in Fig. 5a, but with an inner sphere having the same conductivity as the liquid. (a) Azimuthal rotation at velocity $\zeta=1+(\delta/\lambda)\zeta_0$ with contours in steps of 25 units of ζ_0. The dot-dashed curve corresponds to $\zeta_0=0$. Where the contours are discontinuous the liquid rotates more rapidly than the inner sphere (superrotation with $\zeta_0>0$) and where they are dashed, slower ($\zeta_0<0$). (b) Azimuthal magnetic field sJ.

$$sj=\begin{cases}-s(\delta/\lambda)(r_+/r_-)^3, & s^2<r_-^3/r_+,\\ 0, & s^2>r_-^3/r_+.\end{cases} \tag{80}$$

Here a singularity analogous to that in Eq. (70) can be seen. sj is discontinuous at $\vartheta=\arcsin\sqrt{r_-/r_+}$, reaching a maximum to the left and vanishing to the right. The main difficulties in obtaining a corresponding solution of Eq. (16) are connected with this discontinuity. Standard expansions in terms of any polynomials are essentially nonconvergent. In order to overcome the singularity in Eq. (80), it was necessary to use a specially developed finite difference technique. The resulting azimuthal magnetic field is shown in Fig. 6b for the geophysical situation $3r_-=r_+$.

Since the magnetic field (79) in the layer satisfies Eq. (16), Eq. (53) for the angular rotation velocity simplifies and now coincides with the last of Eqs. (26). The corresponding general solution is again Eq. (31) with the boundary conditions from Eq. (51) which for the applied dipole being considered here take the form

$$\zeta_0=\frac{2Hr_-^3}{\cos\vartheta}\frac{\partial j_0}{\partial r}\quad\text{for}\quad r=r_-. \tag{81}$$

Determining $j_r\equiv\partial j_0/\partial r$ using Eq. (80) and the above (see Fig. 6b) numerical solution for $j=\delta j_0/\lambda$, we obtain an azimuthal rotation from Eqs. (31) and (81) which is shown in Fig. 6a. Contours are drawn in this figure in steps of 25 units of ζ_0, which reflects the relatively large magnitude of the differential rotation. The small superrotation regions (see Sec. 5.2) associated with the large positive derivatives of j_r at the inner sphere near the rotation axis and inside the MHD structure are of special interest. Here the liquid rotates much faster than the inner sphere (Fig. 6a, continuous contours). A large part of the layer, on the other hand, rotates more slowly than the inner sphere, which may explain the westward drift of geomagnetic inhomogeneities.

6.3. Quadrupole in the "insulator–insulator" problem (problem 1)

Outside the quadrupole curve $\not\subset$ (77), the trivial solution (34)–(35) and the expansion (25) yield the large-scale quantities

$$\zeta=\frac{1}{1+s_+^2/s_-^2},\quad J=-\frac{\delta}{\lambda}\frac{s_-^2}{s^2}\zeta, \tag{82}$$

where $s_\pm^2(V)$ are the positive real roots of the following equations distinguished by taking the upper $(+)$ or lower $(-)$ signs:

$$(s_\pm^2)^3-r_\pm^2(s_\pm^2)^2+r_\pm^{10}V^2/9=0, \tag{83}$$

where for the quadrupole, $V=3\sin^2\vartheta\cos\vartheta/r^2$ is taken from Eq. (75).

As in the case of the dipole, the azimuthal magnetic field is opposite the rotation and, in the first approximation, inside the quadrupole curve $\not\subset$ there is no field ($J_0=0$) and the liquid rotates together with the inner sphere ($\zeta_0=1$). However, beginning with the quadrupole, the system ceases to be degenerate. As opposed to the dipole, in this kind of nondegenerate system the rotation is differential outside the $\not\subset$ curves and the meridional flow is of order λ^{-2}, as in the more general case (Fig. 7).

7. DISCUSSION AND CONCLUSIONS

The initial assumption in this paper, namely, almost rigid-body rotation ($\epsilon\ll1$), makes it possible to reduce two complicated MHD problems to linear systems of equations in Sec. 1. Then analytic solutions for these problems are found in the limit of a strong potential magnetic field (3). Now these general solutions can be used to estimate the range of validity of the original assumption. To do this we write the initial system of MHD equations in detail, using the left parts $[A]$, $[B]$, $[C]$, and $[D]$ of the system of Eqs. (7)–(10) linearized in ϵ:

$$\epsilon[A]+\epsilon^2\left\{\frac{1}{rs}\frac{\partial(\chi,s^2\zeta)}{\partial(r,\theta)}+\lambda R_m\left[\nabla\left(\mathbf{e}_\varphi\frac{\Psi}{s}\right)\right]\nabla(s^2J)\right\}=0, \tag{84}$$

$$[B]+\epsilon\left[\frac{\partial\zeta^2}{\partial z}-s^{-3}\frac{\partial(\chi,s\mathscr{D}s\zeta)}{r\partial(r,\theta)}+2s^{-3}\frac{\partial\chi}{\partial z}\mathscr{D}\frac{\chi}{s}\right]$$
$$=\epsilon\lambda^2R_m\left\{s\frac{\partial J^2}{\partial z}-\left[\nabla\left(\mathbf{e}_\varphi\frac{\Psi}{s}\right)\right]\nabla\left(s^{-1}\mathscr{D}\frac{\Psi}{s}\right)\right\}, \tag{85}$$

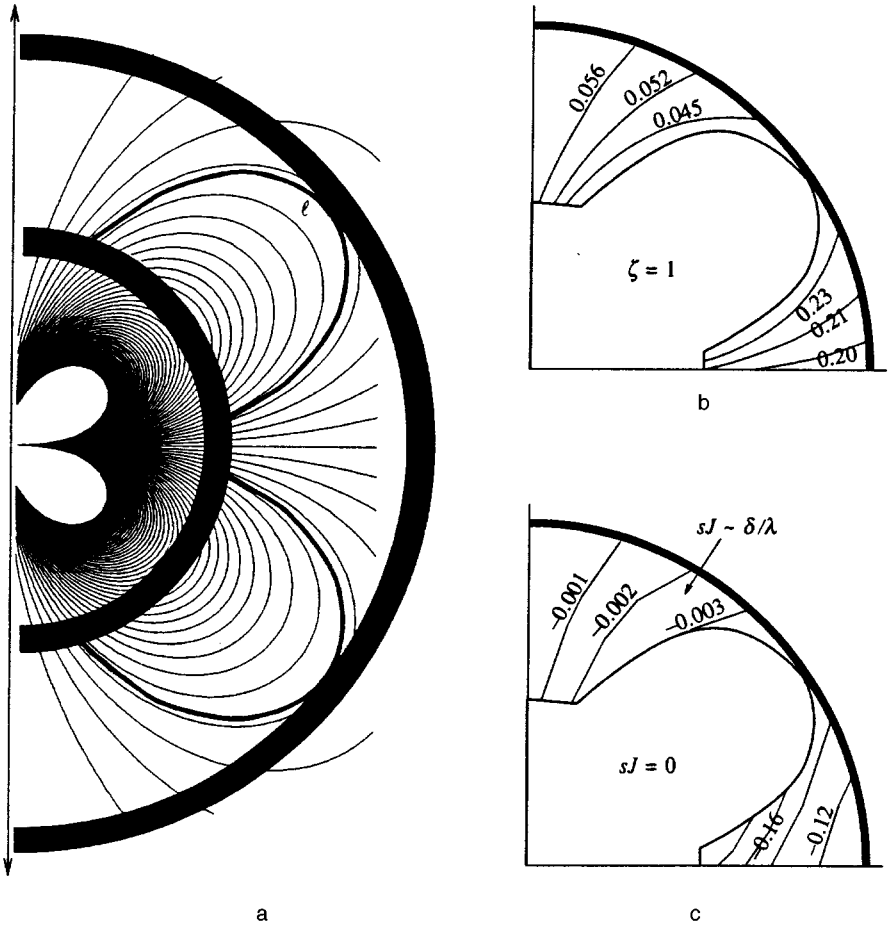


FIG. 7. Magnetic quadrupole (a) with insulating spheres having $r_+ = 2r_-$; azimuthal rotation with angular velocity ζ (b) and azimuthal magnetic field sJ (c).

$$\epsilon[C] + \epsilon^2 R_m \left\{ \left[\nabla \left(\mathbf{e}_\varphi \frac{\Psi}{s} \right) \right] \nabla \zeta - \mathbf{1}_\varphi [(\nabla \chi) \nabla J] \right\} = 0, \quad (86)$$

and

$$\epsilon[D] + \epsilon^2 R_m \left[\nabla \left(\mathbf{e}_\varphi \frac{\Psi}{s} \right) \right] \nabla \chi = 0. \quad (87)$$

For the first, ''insulator–insulator,'' MHD problem (see the solution and estimates of Sec. 2), the linearization is correct if the following inequalities, which replace Eqs. (84)–(87), are satisfied:

$$\lambda \gg \epsilon(\lambda^{-2} + R_m \lambda^{-1}), \quad 1 \gg \epsilon[1 + R_m(\delta^2 + \lambda^{-2})],$$

$$1 \gg \epsilon R_m \lambda^{-2}, \quad \lambda^{-2} \gg \epsilon R_m \lambda^{-4}. \quad (88)$$

With the customary choice of $\delta < 1$ and $R_m > 1$ the correctness condition for problem 1 reduces to

$$\epsilon \ll \min(1, \lambda^2 R_m^{-1}, R_m^{-1} \delta^{-2}). \quad (89)$$

For the second, ''insulator–conductor,'' MHD problem (see the solution and estimates in Sec. 4), the linearization is correct if the following inequalities, which replace Eqs. (84)–(87), are satisfied:

$$\lambda^3 \gg \epsilon(1 + R_m \lambda \delta), \quad 1 \gg \epsilon(1 + R_m \delta \lambda),$$

$$1 \gg \epsilon R_m \lambda^{-3} \delta, \quad 1 \gg \epsilon R_m \lambda^{-3} \delta. \quad (90)$$

The correctness condition for the linearization of problem 2 reduces to

$$\epsilon \ll \min(1, (\lambda R_m \delta)^{-1}). \quad (91)$$

We are therefore justified in applying the general analytic theory in the approximation of Eq. (1) to a wide class of physical MHD systems when conditions (89) or (91) are satisfied, together, respectively, with Eqs. (38) and (39) or Eq. (57).

As an example, let us consider the earth's interior. Recent studies[19] yield a direct estimate of $\epsilon \simeq 10^{-5}$. Reliable estimates have also been obtained[20] of $\lambda \geqslant 10$ and $R_m \simeq 10^8$. The dimensionless thickness of the viscous Ekman layer has been estimated, but less accurately owing to the uncertainty in the viscosity,[21] as $\delta \simeq 10^{-2} - 10^{-7}$. Thus, the strong field conditions (57) are easily satisfied, while the linearization condition (91) is certainly satisfied for $\delta \ll 10^{-4}$. A direct solution of the problem for such narrow MHD layers is essentially impossible, even using the most powerful modern computers (see Refs. 7–9 and 13). Thus, it is logical to use the analytic results obtained here to describe planetary interiors.

In conclusion, we summarize the basic results of this paper as follows:

(1) An analytic solution of the classical MHD problem involving the almost rigid-body rotation of a viscous, con-

ducting spherical layer of liquid in a strong axially symmetric potential magnetic field has been obtained for the first time. The large-scale flows and fields have been described when both spheres are insulators and when the inner sphere is a conductor.

(2) For the first time a description is given of the shear MHD layers that smooth out the large gradients at the boundaries of the MHD structures encompassed by them. The latter original structures ''expel'' the azimuthal magnetic field from regions bounded by potential contours tangent to the spheres.

(3) Applying a constant magnetic field creates a magnetic structure outside the axial cylinder tangent to the inner sphere. Inside the cylinder the rotation is faster, while the meridional flow depends on the height.

(4) An applied dipole field forms an MHD structure tangent to the outer equator. When both spheres are insulating, the liquid inside the structure rotates as a rigid body together with the inner sphere. Outside the structure the rotation is slower, but also rigid-body. When the inner sphere is a conductor, a singularity in the boundary conditions causes the rotation at the axis and inside the MHD structure to be much faster than the rotation of the inner sphere. This sort of superrotation is somewhat surprising, since the given rotation of the inner sphere is faster than that of the outer sphere.

(5) The last example of a general solution is for a quadrupole magnetic field. In this case, two equatorially symmetric structures rotate together with the inner sphere. Outside the structures, as in the most general case, the rotation is differential, the azimuthal magnetic field falls off with the first power of the potential field, and the meridional flow falls off as the square when both spheres are insulators, and as the cube when one sphere is a conductor.

The author thanks his mother for help in the successful completion of this paper. This work was supported by the Russian Fund for Fundamental Research (Grant No. 97-05-64402).

[1] I. Proudman, J. Fluid Mech. **1**, 505 (1956).
[2] K. Stewartson, J. Fluid Mech. **26**, 131 (1966).
[3] A. A. Ruzmaĭkin, Geomagnetizm i Aéronomiya **29**, 299 (1989).
[4] A. Ruzmaikin, in *Theory of Solar and Planetary Dynamos*, M. R. E. Proctor, P. C. Matthews, and A. M. Rucklidge (eds.), Publications of the Newton Institute 1, Cambridge Univ. Press, Cambridge (1993), p. 265.
[5] S. V. Starchenko, Doklady RAN **326**, 983 (1993).
[6] G. A. Glatzmaier and P. H. Roberts, Phys. Earth. Planet. Inter. **91**, 63 (1995).
[7] G. A. Glatzmaier and P. H. Roberts, Nature (London) **377**, 203 (1995).
[8] R. Hollerbach and C. A. Jones, Phys. Earth. Planet. Inter. **75**, 317 (1993).
[9] R. Hollerbach and C. A. Jones, Phys. Earth. Planet. Inter. **87**, 171 (1995).
[10] S. V. Starchenko, J. Geomagn. Geoelectr. **45**, 1563 (1993).
[11] N. Kleeorin, I. Rogachevskii, and A. Ruzmaikin, in *The Cosmic Dynamo*, F. Krause, K.-H. Radler, and G. Rudiger (eds.), Kluwer, Dordrecht (1993), p. 453.
[12] N. Kleeorin, I. Rogachevskii, A. Ruzmaikin, A. M. Soward, and S. V. Starchenko, J. Fluid Mech. **344**, 213 (1997).
[13] R. Hollerbach, Proc. R. Soc. London, Ser. A **444**, 333 (1994).
[14] D. B. Ingham, Phys. Fluids **12**, 389 (1969).
[15] S. Vempaty and D. E. Loper, Phys. Fluids **18**, 1678 (1975).
[16] P. A. Gilman and E. R. Benton, Phys. Fluids **11**, 2397 (1968).
[17] D. E. Loper, Phys. Fluids **13**, 2995 (1970).
[18] D. E. Loper, Phys. Fluids **13**, 2999 (1970).
[19] X. Song and P. G. Richards, Nature (London) **382**, 221 (1996).
[20] S. V. Starchenko, Doklady RAN **348**, 677 (1996).
[21] L. I. Lumb and K. D. Aldridge, J. Geomagn. Geoelectr. **43**, 93 (1991).

Translated by D. H. McNeill

Low field magnetic response of the granular superconductor $La_{1.8}Sr_{0.2}CuO_4$

L. Leylekian, M. Ocio, and L. A. Gurevich*)

Service de Physique de l'Etat Condensé, CEA, CE Saclay, 91191 Gif sur Yvette Cedex, France

M. V. Feigel'man

L. D. Landau Institute for Theoretical Physics, 117940 Moscow, Russia
(Submitted 29 April 1997)
Zh. Éksp. Teor. Fiz. **112**, 2079–2113 (December 1997)

The properties of the low excitation field magnetic response of the granular high-temperature (HT_c) superconductor $La_{1.8}Sr_{0.2}CuO_4$ have been analyzed at low temperatures. The response of the Josephson currents has been derived from the data. It is shown that intergrain current response is fully irreversible, producing a shielding response, but that it does not carry Meissner magnetization. Analysis of the data shows that the system of Josephson currents freezes into a glassy state even in the absence of external magnetic field, which is argued to be a consequence of the *d*-wave nature of superconductivity in $La_{1.8}Sr_{0.2}CuO_4$. The macroscopic diamagnetic response to very weak variations of the magnetic field is shown to be strongly irreversible, but still qualitatively different from any previously known kind of critical-state behavior in superconductors. A phenomenological description of these data is given in terms of a newly proposed fractal model of irreversibility in superconductors. [S1063-7761(97)01212-2]

1. INTRODUCTION

Granular superconductors (SC) are composed of a very large number of small (micron-size) superconductive grains coupled by Josephson tunnelling (or, in some cases, by the proximity effect). These systems are inherently disordered due to randomness in the sizes of grains and their separation. Usually the strength of Josephson coupling between grains is rather weak, so the maximum Josephson energy of the contact between two grains is much below the intragrain superconductive condensation energy. Therefore granular SC can be considered systems with a two-level organization: their short-scale properties are determined by the superconductivity of individual grains, whereas macroscopic SC behavior is governed by weak intergrain coupling.

In studying the latter, one can neglect any internal structure of SC grains and describe them just by the phases ϕ_j of their superconductive order parameters $\Delta_j = |\Delta|_j \exp(i\phi_j)$. As a result, the macroscopic behavior of granular SC can be described by a classical free energy function of the form (see Ref. 1–3):

$$H = \frac{1}{2}\sum_{ij} E_J^{ij} \cos(\phi_i - \phi_j - \alpha_{ij}) + \int d^3r \, \frac{1}{8\pi}(\nabla\times\mathbf{A})^2 - \frac{1}{4\pi}(\nabla\times\mathbf{A})\cdot\mathbf{H}_{\rm ext}), \tag{1.1}$$

where

$$\alpha_{ij} = \frac{2\pi}{\Phi_0}\int_i^j \mathbf{A}\cdot d\mathbf{r}$$

is the phase difference induced by the electromagnetic vector potential $\mathbf{A}$ and $\Phi_0 = \pi\hbar c/e$, whereas the coupling strengths E_J^{ij} are proportional to the maximum Josephson currents: $E_J^{ij} = (\hbar/2e) I_{ij}^c$. The vector potential $\mathbf{A}$ in Eq. (1.1) is the sum of the vector potential $\mathbf{A}_{\rm ext}$ of the external magnetic field $\mathbf{H}_{\rm ext}$ and of the Josephson current-induced vector potential $\mathbf{A}_{\rm ind}$.

In the absence of external magnetic field, the lowest-energy state for the ''Hamiltonian'' (1.1) is clearly a macroscopically superconductive state with all phases ϕ_j equal to one another. Thus, this granular SC system looks like a random XY ferromagnet with randomness in the values of the coupling strengths E_J^{ij}'s (apart from the possible role of the induced vector potential $\mathbf{A}_{\rm ind}$, which will be discussed later); within this analogy, the role of XY ''spin components'' is taken by $S_x = \cos\phi_j$, $S_y = \sin\phi_j$.

The situation becomes much more complicated in the presence of a nonzero external magnetic field $\mathbf{H}_{\rm ext}$, which makes the system randomly frustrated (since magnetic fluxes penetrating plaquettes between neighboring grains are random fractional parts of Φ_0). When the external field is sufficiently strong, $H_{\rm ext} \gg H_0 = \Phi_0/a_0^2$ (here a_0 is the characteristic intergrain distance), the random phases α_{ij} become of the order of π or larger, which implies complete frustration of intergrain coupling—i.e., the system is then expected to resemble an XY spin-glass.

Actually, the random Josephson network in a magnetic field is not exactly identical to an XY spin-glass for the following reasons.[1] First, the effective couplings $\overline{E_J^{ij}} = E_J^{ij}\exp(i\alpha_{ij})$ between ''spins'' $\mathbf{S}_i$ of the frustrated SC network are random complex numbers, whereas in the XY spin-glass model they are real random numbers. Second, the phases α_{ij} generally depend on the total magnetic induction $\mathbf{B} = \mathbf{H}_{\rm ext} + \mathbf{B}_{\rm ind}$, i.e., the effective couplings $\overline{E_J^{ij}}$ depend on the phase variables ϕ_j, which determine the intergrain currents, $I_{ij} = I_{ij}^c \sin(\phi_i - \sigma_j - \alpha_{ij})$. In some cases the effects produced by the self-induced magnetic field $\mathbf{B}_{\rm ind}$ are weak and can be neglected (the quantitative criterion will be discussed later

on), so that phases α_{ij} can be considered to be fixed by the external field.

The model described by the Hamiltonian (1.1) with fixed α_{ij}'s and $H_{\rm ext} \gg H_0$ is usually called the gauge glass model. It is expected on the basis of both analytic[2–5] and numerical[6,7] results that the gauge glass model in 3D space exhibits a true phase transition to a low-temperature glassy superconductive (nonergodic) state. The mean-field theory of such a low-temperature state shows[3,5] that it is characterized by a finite effective penetration depth for the variation of an external field, nonzero macroscopic critical current, and the absence of a macroscopic Meissner effect.

The full model (1.1), with the α's the containing contribution from $\mathbf{B}_{\rm ind}$, is sometimes called a ''gauge glass with screening.''[8] The effect of screening on the existence and properties of the phase transition to a glassy state is not completely clear; some numerical results[8] indicate the absence of a true phase transition in a 3D model with screening. Quantitatively, the strength of screening is determined by the ratio $\beta_L = 2\pi \mathscr{L} I_c / c\Phi_0$ where $\mathscr{L}$ is the characteristic inductance of an elementary intergrain current loop.[9] In ceramics with $\beta_L \ll 1$, screening effects only become important on a long-distance scale $\sim a_0 / \sqrt{\beta_L}$ (i.e., they are similar to strongly type-II superconductors with disorder).

Apart from its relevance to the description of granular superconductors, the gauge glass model with screening is often considered (e.g., Ref. 10) a simplified model describing the large-scale behavior of disordered bulk type-II superconductors in the mixed state (the so-called vortex glass problem). Actually it is unclear *a priori* how these two problems are related; an obvious difference between them is that the basic ingredient of the latter is the vortex lattice, which is clearly an anisotropic object, whereas the former does not contain any preferred direction in 3D space. On the other hand, a granular superconductor in a moderate magnetic field $H_{\rm ext} \leqslant H_0$ can be considered a kind of disordered type-II superconductor, where the notion of a hypervortex (which is a macroscopic analog of an Abrikosov vortex) can be introduced.[2,11] Therefore, the macroscopic properties of a granular network at $H_{\rm ext} \leqslant H_0$ may resemble those of the vortex glass; in such a scenario, a phase transition between vortex glass and gauge glass phases would be expected in a granular superconductive network at $H_{\rm ext} \sim H_0$ (see Ref. 1 for a more detailed discussion of this subject).

Recently, it was noted that granular superconductors may become glassy even in the absence of a external magnetic field, if enough of the Josephson junctions are anomalous, i.e., their minimum Josephson coupling energy corresponds to a phase difference $\Delta\phi = \pi$ instead of 0 (so-called π-junctions). Two completely different origins of π-junctions were proposed: mesoscopic fluctuations in dirty superconductors,[12] and pairing with nonzero momentum.[13,14] Recent experiments revealing the d-wave nature of pairing in high-temperature superconductors[15] suggest the possibility of observing glassy superconductive behavior in HTSC ceramics in virtually zero magnetic field. Note that ceramics with equal concentrations of ordinary and π-junctions are completely equivalent (if screening effects can be neglected) to XY spin-glass. In constrast to the 3D gauge glass model, an XY spin-glass in 3D is expected to have no true thermodynamic phase transition at finite temperature[7]; recently, it has been suggested that the XY spin-glass and d-wave ceramic superconductor might have a new equilibrium ordered phase, the so-called chiral-glass phase.[16] However, these issues are hardly relevant to the measurable response at temperatures much below the ''bare'' glass transition temperature T_g, which we consider in this paper.

Experimental studies of granular superconductors reveal[9,17] the onset of magnetic irreversibility (a difference between Meissner and shielding magnetizations or, in other terms, between field cooled (FC) and zero field cooled (ZFC) magnetizations) below some temperature T_g, which is lower than the SC transition temperature T_c of the grains. However, detailed analysis of the magnetic response in such systems is usually complicated by the mixing of contributions from individual grains and from intergrain currents. The goal of the present paper is to develop a method that makes it possible to extract from the raw d.c. magnetic response data the intergrain contribution, and to compare its behavior with theoretical predictions.

The compound $La_{1.8}Sr_{0.2}CuO_4$ was chosen in this study for experimental convenience, since its critical temperature (≈ 32 K) is within the optimal temperature range of our noise and a.c. susceptibility measurement setup. The sample was fabricated by the standard solid-state reaction of La_2O_3, $SrCO_3$ and CuO.[18] Mixed powder was pressed into pellets, which were sintered in air at 920 °C for 12 hours. The material was then submitted to three cycles of regrinding, sifting to 20 μm, pressing and sintering again at 1100 °C for 12 hours. Samples prepared in two successive runs were used in this study. In the first (sample A), pellets 1 mm thick and 10 mm diameter were obtained, with a density about 80% of the theoretical bulk value. In the second (sample B), cylinders of diameter 6 mm and length 5 to 6 mm were prepared with a density of about 88%. In both preparations, grains sizes were in the range 1–10 μm. Room temperature x-ray powder diffraction patterns divulged a small amount ($<5\%$) of the non-superconductive compound $La_{1-2x}Sr_{2x}Cu_2O_5$.

The rest of this paper is organized as follows. In Sec. 2, a general analysis of the magnetic response data obtained on two different samples (A and B) of $La_{1.8}Sr_{0.2}CuO_4$ ceramic is presented, and the intergrain (Josephson) contribution to the overall response is derived. Section 3 is devoted to a detailed study of the magnetic response of Josephson intergrain network in the low-field range. It is found that the macroscopic critical current is suppressed considerably (by a factor 2) in a magnetic field of only about 2 G. The lower-field d.c.-response to field variations of order 0.05–0.5 G was analyzed for the FC states obtained at $H_{\rm FC} = 0$–10 G and two temperatures, 10 and 20 K. The data at $T = 10$ K and $H_{\rm FC} = 0$ and 0.1 G are shown to be compatible with the Bean critical-state picture,[19] and the low-field critical current value is identified. The rest of the data are in sharp contrast with Bean-model predictions: the screening current grows sublinearly (approximately as the square root) with increasing excitation field strength. Very low-field, low-frequency a.c. measurements are presented, which reveal the strongly irreversible nature of that anomalous response. A new phenom-

enological model is proposed for the treatment of these data. Its first predictions are found to be in a reasonable agreement with the data.

In Sec. 4, a theoretical analysis of our experimental results is given in terms of the existing theories of the gauge-glass state. It is shown that the observed transition temperature to the low-temperature state of the network, and the magnitude of the (low-B,T) critical current, are in sharp contradiction with the (usual) assumption that the zero-field granular network is unfrustrated. In contrast, under the assumption of a strongly frustrated network at $B=0$, all basic measured parameters of the ceramic network are in mutual agreement. We believe that these estimates suggest the existence of a large proportion of π-junctions in $La_{1.8}Sr_{0.2}CuO_4$ ceramics, possibly due to the d-wave nature of superconductivity in cuprates.

Section 5 is devoted to the development of a new model of diamagnetic response in glassy superconductors, which is necessary for the description of the anomalous data described at the end of Sec. 3. This new model (in some sense intermediate between the Bean[19] and the Campbell[20] models) is based on two ideas: the existence of two characteristic ''critical'' currents (J_{c1} and $J_c \gg J_{c1}$), and the fractal nature of free-energy valleys in the ceramic network. Our conclusions are presented in Sec. 6, whereas some technical calculations can be found in the Appendix.

For convenience, the e.m.u. system of units will be used for experimental data, and Gaussian units for theoretical discussions.

2. GENERAL PROPERTIES OF D.C. MAGNETIC RESPONSE

The d.c. magnetization was measured by the classical extraction method. Two SQUID magnetometers were used: one a home-made apparatus used in several previous spin-glass studies,[21] the other a commercial system (Cryogenics S500).

In this section, we describe the static magnetic response of samples A and B and present a preliminary treatment of these data, in order to distinguish between the magnetic response of individual grains and intergrain currents[9,22] (a detailed study of the latter is the subject of the next section). We first present results obtained after cooling the samples in various d.c. fields and applying small field increases. Secondly, we derive from the results the response of the Josephson currents as a function of field and temperature. Finally, we show that the behavior of the field-cooled susceptibility can be satisfactorily accounted for if the system of Josephson currents does not carry Meissner magnetization. We show that the same interpretation accounts fairly well for the FC results which, at first sight, are rather different for samples A and B.

Sample A is a 1 mm-thick pellet with an approximately ellipsoidal shape of 2×6 mm. Its calculated volume is $V \approx 8.5$ mm^3, and the demagnetizing field coefficient for a field parallel to the longitudinal axis is $N \approx 0.06$.[23]

Figure 1 displays the magnetic dipole moment of the sample cooled to 10 K in zero field and submitted to cycles $0 \rightarrow H_{\max} \rightarrow 0$ for several values of $H_{\max}$ up to 2 G. At the lowest increasing fields, the moment increases initially with

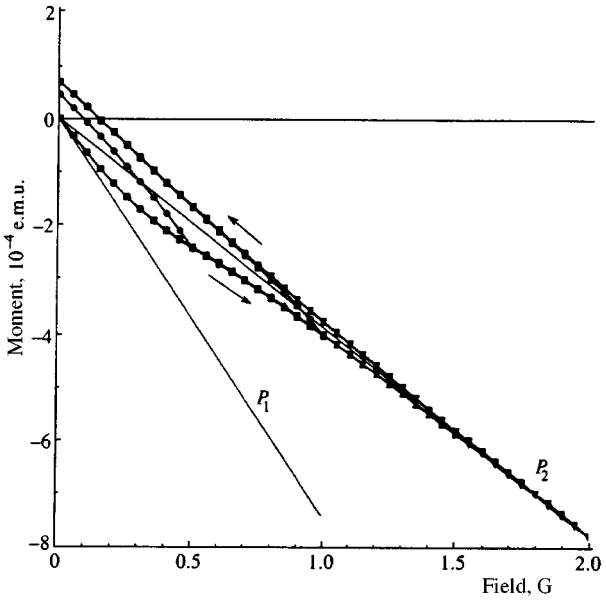


FIG. 1. Magnetic moment of the sample A as a function of field applied in the zero-field cooled state (e.m. units of magnetic moment correspond to cm$^3\cdot$G). $H_{\max}$=0.5 (○), 1 (□), 1.5 (Δ) G. The curves for field loops $0 - H_{\max} - 0$ are superimposed.

slope P_1. Above 1.5 G, it approaches slope P_2. The remanent positive moment saturates for $H_{\max} \geqslant 1$ G. The calculated moment of the sample for perfect volume shielding in a homogeneous field is (e.m.u. system)

$$\mathcal{M} = -\frac{HV}{4\pi(1-N)} = -0.72\times 10^{-3}\cdot H \text{ cm}^3\cdot\text{G}.$$

Owing to uncertainties in the evaluation of the volume, this value is only accurate to $\pm 5\%$. Nevertheless, it is in fair agreement with the slope P_1 in Fig. 1. On the other hand, slope P_2 is about 53%, a rather small value since the density ratio of the sample is about 80%. At such low temperatures (in comparison with $T_c \approx 32$ K), where the lower critical field of the grain's material is above 100 G, one would expect expulsion of the field by the grains with a penetration depth λ. The expected value for the magnetization $M = \mathcal{M}/V$ of the system of uncoupled grains can be calculated as[24]

$$\frac{M}{H} = \frac{1}{4\pi}\,\frac{f}{1-fN-(1-f\ )n}, \tag{2.1}$$

where f is the volume fraction of the superconductive material and n is the demagnetizing-field coefficient for the grains.

As an estimate, we assume the grains to be spherical ($n=1/3$) and, using $M/H \approx 0.53\cdot 1/4\pi$ and $N=0.06$, we find $f \approx 0.41$. This value is considerably below the volume fraction of the sample filled by grains (≈ 0.8); we assume that the difference is due to the intergrain penetration depth λ being comparable to the grain size r, and we estimate an effective value of λ to be

$$f = 0.41 = 0.8\left(1-\frac{\lambda}{r}\right)^3$$

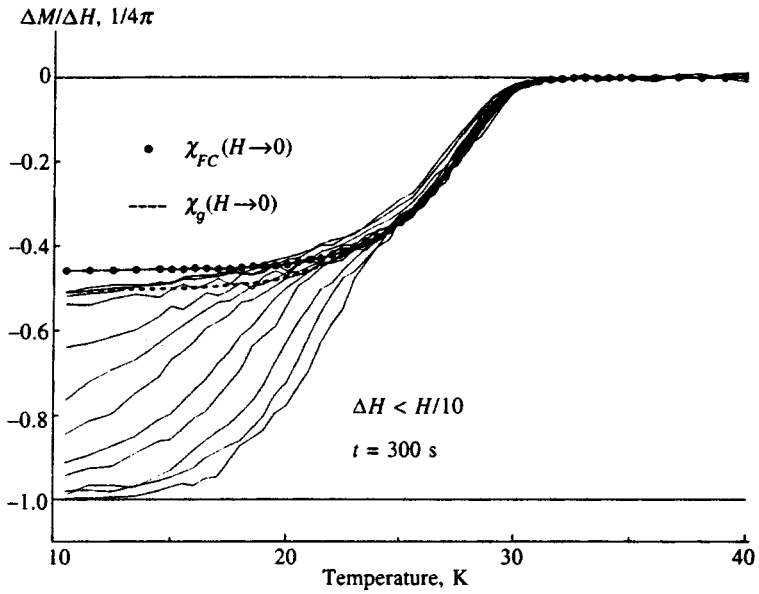


FIG. 2. Shielding susceptibility of the sample *A* as a function of temperature, normalized to the moment for complete shielding. Curves (top to bottom) correspond to $H_{dc}=20.07$, 14.34, 8.60, 5.73, 2.87, 1.99, 1.42, 0.85, 0.57, 0.28, 0.14, 0.06 G.

yielding $\lambda=0.2\,r$. Taking an average size of 5 μm for the grains, we obtain $\lambda\approx 500$ nm. Values reported for the mean penetration depth in $La_{1.8}Sr_{0.2}CuO_4$ are about 200 nm.[25] The value found here is larger than the expected mean value for the homogeneous material, indicating that the grains are not monocrystalline. This will be confirmed below by the results of field-cooling experiments.

The shielding susceptibility is plotted in Fig. 2, as a function of temperature and for several values of the ambient FC field. The measurements were performed according to the following procedure: the sample was cooled in a field H_{dc} down to the working temperature and the moment was measured after waiting 300 sec; then the field was increased by a small amount $\Delta H\leqslant H_{dc}/10$ and the moment was measured again after waiting 300 sec. The figure displays the experimental shielding susceptibility normalized to the value for total shielding, i.e.,

$$\chi_{sh}=\frac{\mathcal{M}(H+\Delta H)-\mathcal{M}(H)}{\Delta H}\,\frac{4\pi(1-N)}{V}.$$

The curves show the double step usually ascribed to the action of both intragrain currents and Josephson intergrain currents.[26] At high temperature, the onset of grain diamagnetism occurs at about 32 K. Above 25 K, the response corresponds to the diamagnetism of the grains. At fixed temperature, it is H_{dc}-independent for $H_{dc}\leqslant 5$ G, and decreases for increasing $H_{dc}>5$ G. Below 25 K the onset of Josephson currents is manifested by a second step of the diamagnetic response. This second step appears at decreasing temperature as H_{dc} increases. At the lowest temperatures, the diamagnetic moment amounts to about 100% of flux expulsion at $H_{dc}=0$, and decreases with increasing H_{dc}. At $H_{dc}>8$ G, the flux expulsion saturates at a value slightly above 50%, which corresponds roughly to the 53% level determined above for the grain response.

The susceptibility in Fig. 2 contains the contributions of grains and Josephson currents. The contributions can be separated along the lines of the work by Dersh and Blatter.[22] The induction in the sample is $B=H+4\pi(M_g+M_j)$, where M_g and M_j stand for the magnetization of grains and of the Josephson currents. It should be noted that the magnetization due to macroscopic circulating currents in a superconductor is sample-size dependent, i.e., the corresponding susceptibility is not a local quantity. At the macroscopic scale of the circulating currents, the magnetization M_g can always be written as $\chi_g H_{local}$, where $\chi_g(H)$ is homogeneous over the sample. In what follows, we consider quantities averaged over the volume of the sample: in that case, M_j is the averaged moment per volume unit due to the currents.

The demagnetizing field effect will be neglected in the calculations. We have verified that, owing to the small value of the demagnetizing factor, this does not modify the essential features of the result while allowing a simpler derivation (the demagnetizing factor will be taken into account when analyzing the data from sample *B*). We get

$$M_g=\chi_g(H+4\pi M_j).$$

Then

$$M=M_g+M_j=\chi_g H+M_j\mu_g,\quad \mu_g=1+4\pi\chi_g,$$

and

$$M_j=\frac{M-\chi_g H}{\mu_g}. \tag{2.2}$$

Equation (2.2) must be considered with care, since χ_g is history- and field-dependent. In fact it is well-adapted to the description of the result of zero (or small) field cooling experiments.

More generally, we must consider the response to field increments δH to obtain $\chi=\delta M/\delta H$. Then, the polarizability[1)] χ_j of the Josephson network reads

$$\chi_j=\frac{\chi-\chi_g}{\mu_g}. \tag{2.3}$$

Note that we can equivalently consider the response of the current system in a homogeneous medium with permeability

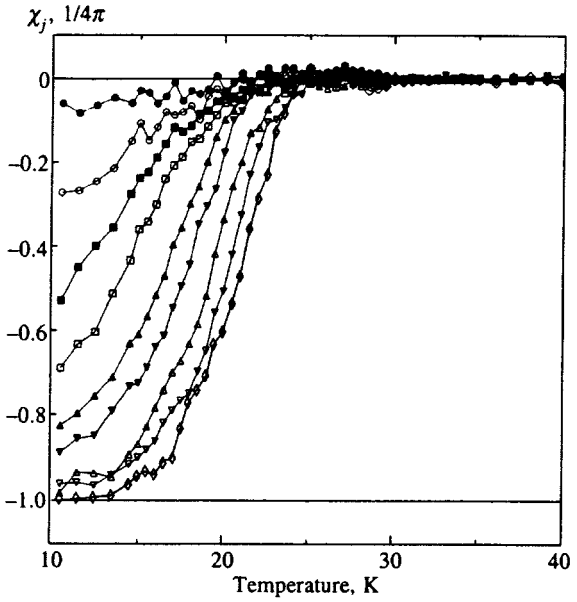


FIG. 3. Josephson currents susceptibility in the sample *A* as a function of temperature. Values have been calculated from data of Fig. 2 and using Eq. (2.3). H_{dc}=5.73 (●), 2.87 (○), 1.99 (■), 1.42 (□), 0.85 (▲), 0.57 (▼), 0.28 (△), 0.14 (▽), 0.06 (◇) G.

μ_g. If the applied field is varied by δH, the Josephson network sees a variation of internal field $\delta H_1=\mu_g \delta H$ and develops a polarization $\delta M_j=\chi_j \delta H_i$. We then recover Eq. (2.3).

The value of χ_g could be determined in principle if we were able to obtain a packing of disconnected grains equivalent to the packing of the sintered sample. In practice this was not possible. Indeed, mechanical grinding resulted in breaking a large number of grains, and thus modifying the characteristics of the material. Nevertheless, it is possible to derive χ_g, at least approximately, from the data of Fig. 2. At high temperatures, above the onset of intergrain currents at $\approx$25 K, the shielding susceptibility χ_{sh} is due to the grains alone, independent of H_{dc} below $\approx$6 G. At low temperatures, for H_{dc} above about 6 G, the χ_{sh} curves coincide, and there is no manifestation of the onset of intergrain currents. Thus, here also, χ_{sh} represents the response of the grains alone.

Hence, the response χ_g of the grains can be reasonably approximated by an interpolation between these two limits. The interpolation curve, obtained by a smoothing procedure between the two curves at H_{dc}=0 G and H_{dc}=20 G, is displayed on Fig. 2 (dashed curve). The values of χ_j derived from Eq. (2.3) are plotted against temperature in Fig. 3, for $H_{dc}<6$ G.

Note that the dependence of χ_j on H_{dc} seen in the figure can be assumed to reflect the behavior of the initial shielding properties of the Josephson network with increasing of H_{dc}. Nevertheless, a nonlinearity in the response due to the correlative increase in the value of ΔH ($\Delta H=H/10$) cannot be ruled out; this issue will be studied in detail in sample *B*. Finally, we point out the similarity between our data and the results of earlier numerical simulations on a gauge glass system.[17]

Above we have discussed the system's responses to the variation of magnetic field at fixed temperature (i.e., shielding responses), and derived from these data the polarizability χ_j of the intergrain system. Now we turn to a description of the results of the field cooling measurements. FC (Meissner) magnetization was measured by the standard procedure between 10 and 40 K for fields from 0.01 to 20 G. The results are plotted against temperature in Fig. 4 and against applied field in Fig. 5. Data are normalized to the value of the moment for 100% shielding.

Even in the weakest field, the flux expulsion rate is no more than 45%, less than the 53% shielding by grains. In weak fields, below 1 G, there is an approximate affinity between the curves of M/H versus T. M/H can be extrapolated linearly to $H\to 0$. The result is plotted in Fig. 2 (solid

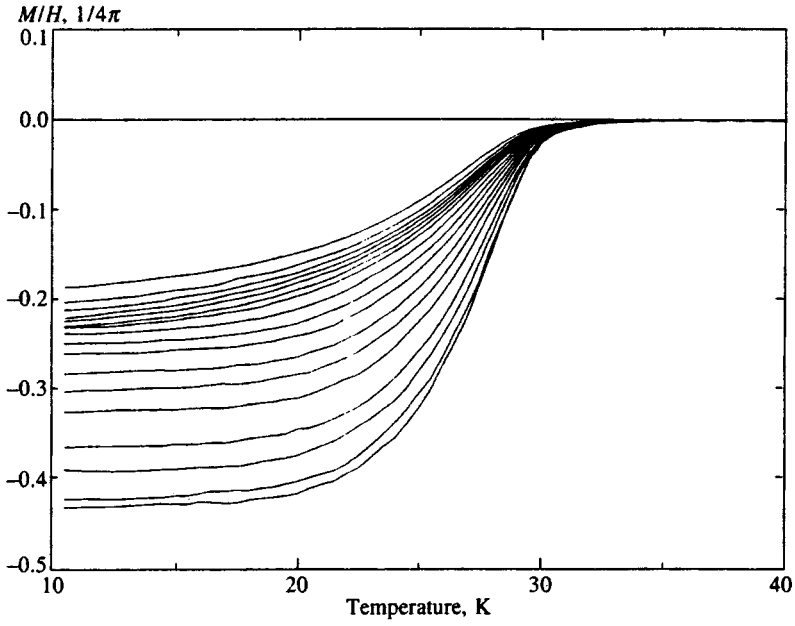


FIG. 4. FC (Meissner) susceptibility of the sample *A* as a function of temperature for fields up to 20 G. Curves (top to bottom) correspond to H_{dc}=20.07, 14.34, 8.60, 5.73, 4.01, 2.87, 1.99, 1.42, 0.85, 0.57, 0.4, 0.28, 0.2, 0.14, 0.08, 0.06, 0.02, 0.01 G.

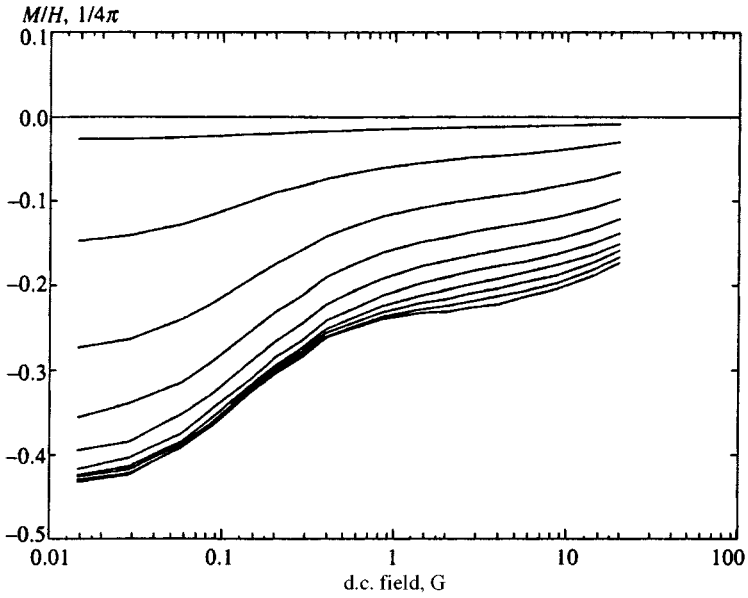


FIG. 5. FC (Meissner) susceptibility of the sample *A* as a function of field for selected temperatures. Curves (top to bottom) $T=36$, 30, 28, 26, 24, 22, 20, 18, 16, 13.5, 10.5 K.

circles): one can see that the extrapolated FC susceptibility coincides exactly with the low d.c. field shielding susceptibility above 25 K. Therefore, at low d.c. fields above 25 K, the response of the grain system is reversible, and it is described well by the low d.c. field shielding curves; this justifies the hypothesis used above for the calculation of χ_j.

On the other hand (see Fig. 5), the behavior of the FC susceptibility as a function of H is nontrivial. M/H decreases with increasing field, and reaches a stable level (about 25% at the lowest temperatures) at roughly 1 G. Whatever the temperature, this decrease is centered at a constant value of the field, about 0.1–0.3 G. Above 5 G, M/H decreases once more with increasing field. Note an essential difference between the FC results presented on Fig. 4 and the shielding results above (Fig. 2): the FC curves do not show any increase in the response M/H with decreasing temperature below 20 K, where the intergrain coupling grows considerably (as can be seen in Fig. 2). This means that the network of intergrain currents does not produce Meissner (FC) magnetization, whereas it does produce *shielding* magnetization.

The behavior of the FC susceptibility $\chi_{FC}=M_{FC}/H$ as a function of applied field H depicted in Fig. 5 shows two nontrivial features: crossover between two plateaus (at low and moderate values of H), which takes place between 0.1 and 1 G independent of temperature, and a value of the low-field χ_{FC} that is appreciably lower than the Meissner response of uncoupled grains (45% versus 53%). These features can be understood in terms of polycrystalline structure of the grains, which can be suspected from the large values of the penetration depth obtained from the results of Fig. 1, and self-shielding (pinning of the magnetic flux) by the Josephson currents when reducing the temperature in an applied field.

We start with the first feature; the curves of FC magnetization in Fig. 5 are rather similar to those measured by Ruppel *et al.*[27] in YBaCuO ceramics, who interpreted their results on the basis of a theory of flux expulsion by strongly anisotropic randomly oriented crystallites derived by Wohllebeen *et al.*[28] We stress that the model is not based on any activated flux creep mechanism. It is thus well-adapted to the analysis of our results; indeed, flux creep effects can scarcely be invoked here, since the temperature has no apparent effect on the characteristic field related to the decrease in magnetization.

The starting point of the model is that, provided the size b of the crystallites is such that $\lambda_\parallel \ll b \ll \lambda_\perp$, the longitudinal magnetization of a crystallite whose c-axis makes an angle α with the field is given by $M=-(H/4\pi)\gamma\cos^2\alpha$, where γ is a factor close to 1 that depends on the ratio $\lambda_\parallel/b$. Averaging over α, one obtains

$$\frac{M}{H}=\frac{\gamma}{3}\frac{1}{4\pi}.$$

It must be stressed that the system of intragrain crystallites is a strongly-coupled system, in contrast to the system of grains that compose the ceramic. Therefore, a grain consists of an ensemble of interconnected Josephson loops surrounding crystallites whose planes are nearly along the field, and are thus transparent to the field. At low fields, this system will expel the flux with a penetration depth that depends on the junction coupling energy. Nevertheless, when the field is such that a loop sees a flux larger than $\sim\Phi_0/2$, the macroscopic magnetization of the Josephson currents vanishes and the system reacts as an ensemble of disconnected crystallites.[29] The characteristic field of this crossover is such that[28]

$$\frac{H_m s_c}{\Phi_0}\approx 0.1. \tag{2.4}$$

Recently determined values for the penetration depth in $La_{1.8}Sr_{0.2}CuO_4$[30] are $\lambda_\parallel=150$ nm and $\lambda_\perp=1500$ nm. Older measurements indicate a higher anisotropy, up to a factor 14.[31] We can thus reasonnably consider that the model can

be applied in our case. Taking $H_m = 0.3$ G, we obtain $s_c = 7.4 \times 10^{-8}$ cm^2. With $s_c \approx \pi b^2$ this leads to a mean diameter $b = 1.5$ μm for the crystallites.

Above H_m, the system acts as an ensemble of crystallites whose average susceptibility is $(\gamma/3)(1/4\pi)$. With the density ratio $f = 0.8$, taking $\gamma = 1$ and assuming spherical crystallites, we obtain from Eq. (2.1) $4\pi M/H = 0.31$, which is above the experimental value (the latter being about 0.25).

Nevertheless, it must be noted that we have neglected here the effect of the factor γ and used a rather unrealistic spherical approximation for the shape of crystallites. Finally, it has been seen that above 5 G, the FC magnetization starts to decrease once more with increasing field, although H_{c1} is larger than 100 G in $La_{1.8}Sr_{0.2}CuO_4$. This might be due to intrinsic pinning inside the crystallites themselves when the applied field is such that the flux in the cross-section of one crystallites is larger than Φ_0. With a mean radius of 0.8 μm for the crystallites, this crossover occurs at about 10 G.

We now turn to the discussion of the second feature mentioned above. At temperatures below 25 K, the Josephson currents become active. Their effect is that at 10 K, the shielding response of the system of grains amounts to about 53%, while the FC susceptibility saturates at about 45%. This difference is enough to be significant, and can be interpreted to be the result of pinning by the Josephson network. In fact, this pinning can be understood as a back shielding effect of the Josephson currents against a decrease in the local internal field, due to the temperature dependence of the grain's system permeability μ_g. We have seen above that the response of the system consists of the two parts: for an applied field H, the internal field due to the grains seen by the integrain currents is $H_i = \mu_g H$, and the intergrain currents system reacts to all variations of H_i with a polarizability χ_j and generates a magnetization $\delta M_j = \chi_j \delta H_i$.

Thus, when the temperature is decreased by dT, the internal field decreases by $H\, d\mu_g/dT$ and the Josephson network tends to screen this variation. Since the intergrain currents give no Meissner effect, we consider their response to be totally irreversible. Thus for a variation dT in the temperature, in a field H, the variation in the induction is

$$dB = (1 + 4\pi\chi_j)\left(\frac{d\mu_g}{dT}\right)_H H dT.$$

On the other hand, $B = (1 + 4\pi\chi_{\rm FC})H$. With $\mu_g = 1 + 4\pi\chi_g$, we finally obtain

$$\chi_{\rm FC} = \chi_g + 4\pi \int_{T_c}^{T} \chi_j \frac{d\chi_g}{dT}\, dT = \chi_g + \chi_j^{\rm FC}. \qquad (2.5)$$

$M_j^{\rm FC} = \chi_j^{\rm FC} H$ is the magnetization produced by the Josephson currents due to variations in μ_g with decreasing temperature. As χ_g is known only in the limit $H_{dc} \to 0$, Eq. (2.5) has been used to calculate $\chi_{\rm FC}$ versus T in the limit of low field.

In order to do so, we started with the values of $\chi_g(H \to 0)$ as derived above; for χ_j, we used the values given in Fig. 3 for the weakest field $H_{\rm dc} = 0.06$ G. The result is plotted in Fig. 6. The agreement of calculated values with

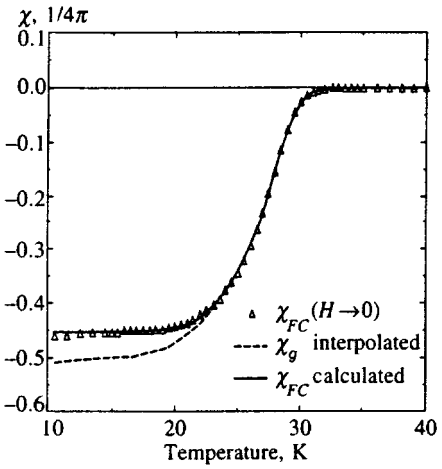


FIG. 6. $\chi_{\rm FC}$ calculated with Eq. (2.5) from the values of χ_g and $\chi_j(0)$ (see the text).

experimental data is satisfactory, although not perfect. This discrepancy is emphasized if we revert Eq. (2.5) to calculate χ_j as a function of χ_g and $\chi_{\rm FC}$.

The reason is that we have used the simplest linear model of back shielding here. In fact, as we will see later, the response of the current system is strongly nonlinear, with the susceptibility χ_j decreasing with increasing ΔH, and this effect becomes stronger as the temperature increases. The result is that the calculated efficiency of back shielding is underestimated, since the value of the experimental susceptibility is determined by applying finite increments ΔH.

2.2. Sample *B*

Sample B was machined from one of the original cylinders, in form of a parallelepiped of dimensions approximately $3 \times 3 \times 6$ mm. Its calculated volume is $V \approx 52.6$ mm^3, and its demagnetizing field coefficient for a longitudinal field is $N \approx 0.19$. In a longitudinal field, its calculated moment for perfect flux expulsion is $\mathcal{M} = 5.1 \pm 0.2 \times 10^{-3} \cdot H$ cm$^3 \cdot$G.

Measurements of the initial magnetization at 10 K are in fair agreement with this value. For $H_{\rm dc}$ above 3 G and up to 30 G, the ratio $\Delta\mathcal{M}/\Delta H$ reaches a stable level at about 3.2×10^{-3} cm^3 which corresponds to the response of the grains alone. With a density ratio of 88% for this sample, using Eq. (2.1) one finds $f = 0.46$, yielding $\lambda = 0.19r$, i.e., the same value as derived for sample A.

The shielding susceptibility was measured in this sample by using a more sophisticated method, in order to reduce the effect of nonlinearity. After cooling the sample to the working temperature in the d.c. field, the field was increased in 5 successive steps ΔH, and $\Delta\mathcal{M}$ was measured. At the lowest fields, $\Delta H = 10$ mG, and (to keep a good signal/noise ratio) $\Delta H = H_{\rm dc}/50$ at the highest ones. The value of $\Delta\mathcal{M}_n/\Sigma_n \Delta H$ was then extrapolated to $\Delta H = 0$ by a least-squares fit.

As in the case of sample A, all curves with $H_{\rm dc} \leqslant 10$ G merge at high temperatures to a common curve, which corresponds to the flux expulsion by the grains. The main difference from sample A is that in sample B, Josephson current shielding sets in at higher temperatures. This is consistent with the fact that sample B is denser, resulting in better cou-

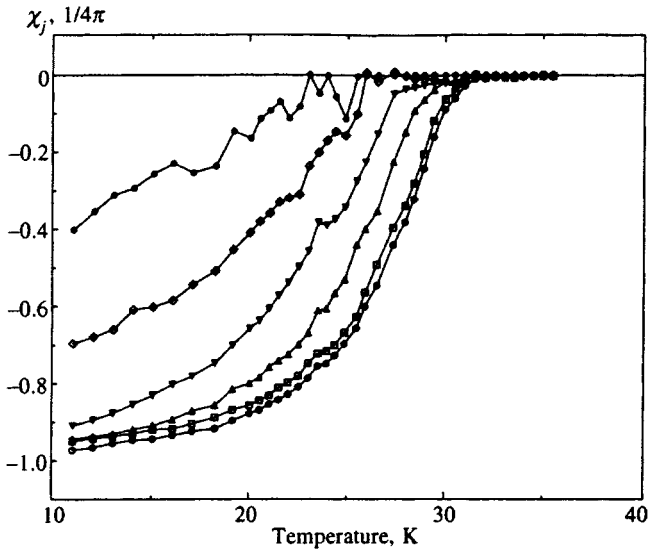


FIG. 7. Josephson current shielding susceptibility as derived from the data and the use of Eq. (2.6). $H_{dc}=0$ (○), 0.1 (□), 0.5 (△), 1 (▽), 2 (◇), 5 (●) G.

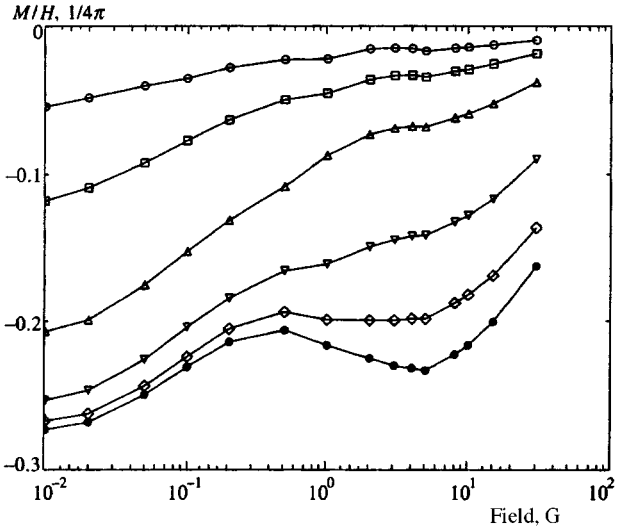


FIG. 8. FC (Meissner) susceptibility of sample *B* normalized to the moment for total flux expulsion, as a function of field. $T=31$ (○), 30 (□), 28 (△), 25 (▽), 20 (◇), 10 (●) K.

pling between grains; moreover, it is larger, which also increases the total shielding magnetization. At low temperature, the magnetization curve at $H_{dc}=20$ G reaches a level slightly above 60%, which corresponds to the low-temperature level for the grains.

The shielding response of the Josephson currents can be obtained with the procedure already used for sample *A*. Here the demagnetizing factor cannot be neglected ($N\approx 0.19$). Two kind of quantities are to be considered. The first involves the responses χ_g and χ of an equivalent sample without demagnetizing field (e.g., an infinitely long cylinder with the same cross section); here χ_g is the response of the system of grains alone, without intergrain currents, and χ is the total response of the system of intragrain plus intergrain currents. The second involves the measured responses $\overline{\chi_g}$ and $\overline{\chi}$; these correspond to the measured moment for each case, normalized to the moment for total flux expulsion from the volume of the sample. The relation between the two kinds of quantities is

$$\frac{M}{H}=\frac{\overline{\chi}}{1-N}=\frac{\chi}{1+4\pi N\chi}.$$

A relation of the same kind holds for χ_g and $\overline{\chi_g}$. With the use of Eq. (2.3), we finally obtain

$$\chi_j=\frac{\overline{\chi}-\overline{\chi_g}}{(1-N\overline{\mu})\overline{\mu_g}}, \tag{2.6}$$

where $\overline{\mu}=1+4\pi\overline{\chi}$, $\overline{\mu_g}=1+4\pi\overline{\chi_g}$.

Similar to the case of sample *A*, an approximate curve has been determined for $\overline{\chi_g}$ by interpolation between the weak H_{dc} curves at high temperatures, and the curve at $H_{dc}=20$ G at low temperatures. The values of χ_j have then been derived from Eq. (2.6) and plotted in Fig. 7. The set of curves is similar to the set for sample *A*, except for the higher onset temperature of the intergrain currents.

Field-cooled magnetization data, normalized to the value of the moment for full flux expulsion, are reported in Fig. 8 as a function of field up to 30 G. At the lowest field and temperature, the FC magnetization does not exceed 28% of its value for full flux expulsion. Furthermore, at low temperatures the curves representing the field dependence present a second maximum at about 5 G. We expect that this complicated behavior is due to the back shielding effect of the intergrain currents, as discussed for sample *A*.

To take them into account, a relation similar to Eq. (2.5) (but with the demagnetizing effect included) should be derived. The internal field is given as usual by $H_i=H-4\pi NM$, and the value of the local field seen by the currents is $H_l=\mu_g H_i$. Thus, under a temperature variation dT,

$$\frac{dH_l}{dT}=\frac{d\mu_g}{dT}(H-4\pi NM)-4\pi N\mu_g\frac{dM}{dT}.$$

With $dB/dT=\mu_j dH_l/dT$, and using the relation

$$M=\frac{\overline{\chi_{FC}}H}{1-N}=\frac{1}{4\pi}\int_{T_c}^{T}\frac{d(B-H_i)}{dT}\,dT,$$

one obtains after integration

$$\overline{\chi_{FC}}=\frac{1-N}{4\pi N}(1-\exp(-4\pi NI)),$$

$$I=\int_{T_c}^{T}\frac{\mu_j}{1-N(1-\mu_g\mu_j)}\frac{d\chi_g}{dT}\,dT. \tag{2.7}$$

Here $\mu_j=1+4\pi\chi_j$, with χ_j plotted in Fig. 7, whereas the value of μ_g was obtained using the relation $\mu_g=(1-N)\overline{\mu_g}/(1-N\overline{\mu_g})$ from the value of $\overline{\chi_g}$ as derived above.

The values of $\overline{\chi_{FC}}$ for $H\to 0$ have been calculated using the values of $\overline{\chi_g}$ as determined above, and the values of χ_j at $H_{dc}=0$. The calculated value of $\overline{\chi_{FC}}$ was found to be about -0.35 at $T=10$ K, whereas its measured value was about -0.28. The discrepancy between measured and calculated values is larger here than in corresponding results for sample

A. We believe that the origin of this discrepancy is the same as in the case of sample A, i.e., it stems from the nonlinear response effect. This effect is numerically larger in sample B, since here the onset of Josephson currents occurs in a range of temperature where χ_g still varies strongly, in contrast to the case of sample A.

The above analysis shows (notwithstanding the aforementioned discrepancy) that the back shielding effect leads to a strong reduction in the field-cooled susceptibility as compared with the susceptibility of the grains alone. It is then easy to understand the complex behavior of $\overline{\chi_{\rm FC}}$ as a function of field: at 10 K, for instance, the onset of back shielding occurs at about 20 G, and its amplitude increases with decreasing field due to the increase in $\overline{\chi_j}$. Starting with the two-step behavior of χ_g expected from the theory of Wohllebeen *et al.*[28] (and seen in the data of sample A, where back shielding is less important), back shielding results in the double maximum shape of the measured curves.

3. DETAILED STUDY OF THE JOSEPHSON NETWORK RESPONSE

3.1. Determination of the global critical current

In this subsection we present the procedure we used to derive the macroscopic critical current in our sample B. This procedure is not quite trivial, since we are interested in the dependence of the critical current on the background d.c. field in the sample, so we need to analyze the magnetization curves, which depend both on the cooling field $H_{\rm dc}$ and the field variation δH.

The magnetization was recorded at 10 and 20 K as a function of increasing ΔH with the smallest possible field steps ($\delta H = 10$ mG), and starting from several FC states. From the $\Delta\mathcal{M}$ data, it is possible to derive the current response $\Delta\mathcal{M}_j$ as a function of ΔH. For this we use Eq. (2.6), which can be written as

$$\Delta\mathcal{M}_j = \frac{\Delta\mathcal{M} - \Delta\mathcal{M}_g}{(1-N\overline{\mu})\overline{\mu_g}}, \tag{3.1}$$

where $\Delta\mathcal{M}_g$ is the magnetization of the grains alone; $\overline{\mu}$ and $\overline{\mu_g}$ are defined in Subsection 2.2. The value of the grain system response is approximately derived in the same section: $\Delta\mathcal{M}_g \approx 3.2\times10^{-3}\cdot H$ cm$^3\cdot$G at 10 K, and $\Delta\mathcal{M}_g \approx 2.9\times10^{-3}\cdot H$ cm$^3\cdot$G at 20 K. Calculated values of $\Delta\mathcal{M}_j$ at 10 K are plotted in Fig. 9. Note the similarity of these results to the magnetization curves of classical type-II superconductors with strong pinning (the difference is that here ΔH plays the role of H).

After cooling the sample at zero d.c. field, its response is obviously symmetric with respect to ΔH. When it is cooled in a finite d.c. field, this is no longer the case, as explained in the previous section. The magnetic moment of the sample just after cooling is $\mathcal{M}_{\rm FC} = \mathcal{M}_g + \mathcal{M}_j^{\rm FC}$ where $\mathcal{M}_j^{\rm FC}$ is the positive moment due to back shielding by the Josephson currents that developed during the cooling process (see Eqs. (2.3) and (2.6)). The total moment produced by the intergrain currents after increasing the field by ΔH is thus $\mathcal{M}_j = \mathcal{M}_j^{\rm FC} + \Delta\mathcal{M}_j$. It is this moment that vanishes when $J_c \to 0$ (at large enough ΔH), and thus $\Delta\mathcal{M}_j$ approaches $-\mathcal{M}_j^{\rm FC}$.

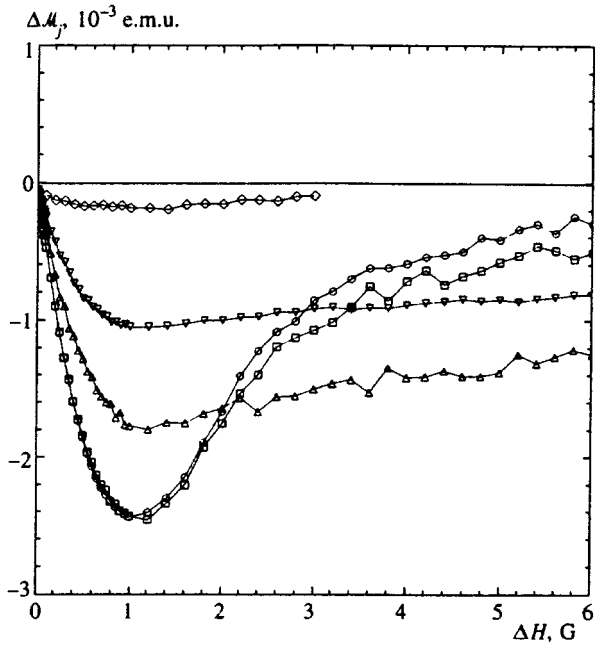


FIG. 9. Shielding moment of the Josephson currents after cooling the sample at 10 K in a d.c. field over the range 0–10 G. $T = 10$ K, $H_{\rm dc} = 0$ (○), 0.1 (□), 2 (△), 4 (▽), 10 (◇) G.

In Fig. 10 we show the data recorded at $T = 10$ K and $H_{\rm dc} = 2$ G. Curves recorded at positive and negative ΔH both converge to the value corresponding to $-\mathcal{M}_j^{\rm FC}$; at 10 K, the value of $-\mathcal{M}_j^{\rm FC}$ is about 1.1×10^{-3} e.m.u.

When $\Delta H > H_{\rm dc}$, it is natural to expect that the response of the Josephson network does not depend on the initial state. A simple illustration can be given by analogy with Bean-like

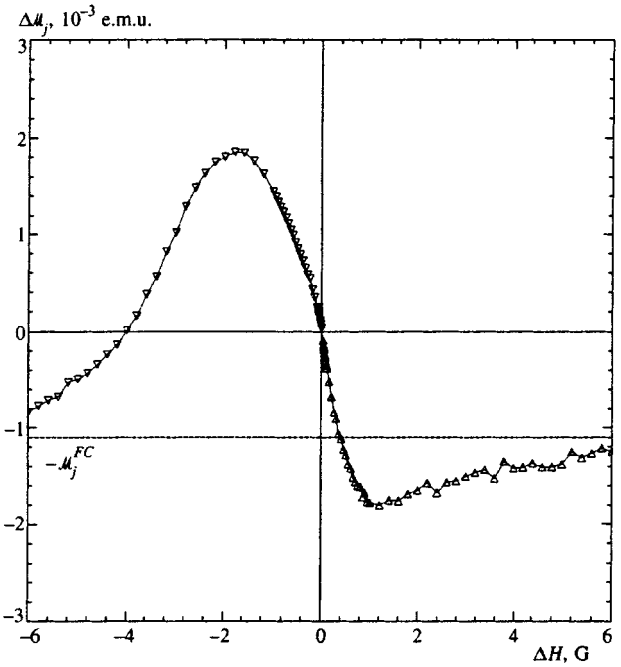


FIG. 10. Shielding moment of the Josephson currents after cooling the sample at 10 K in a d.c. field $H_{\rm dc} = 2$ G. Data are for positive and negative field steps. $T = 10$ K. $H_{\rm dc} = 2$ G, $\Delta H > 0$ (△), $\Delta H < 0$ (▽).

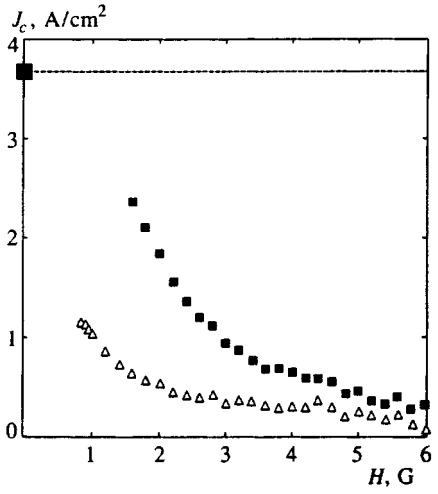


FIG. 11. Calculated values of the average critical current $\overline{J_c}$ as a function of total field for strong field penetration. The large square corresponds to the initial J_c as determined in subsection 3.2. $T=10$ (■), 20 (△) K.

pinning in type-II superconductors.[19] At large ΔH, when the induction profile has penetrated to the center of the sample, the magnetization does not depend on ΔH, but only on J_c. If J_c varies with the induction in the sample, as is the case in real materials, the magnetization depends on the total H, whatever the value of $H_{\rm dc}$ in which the sample was cooled. Indeed, when plotted as a function of the total field $H_{\rm dc}+\Delta H$, the curves giving the total moment of network currents $\mathscr{M}_j^{\rm FC}+\Delta\mathscr{M}_j$ merge in their ''large'' field part (i.e., above their maximum). The values have been calculated, with $-\mathscr{M}_j^{\rm FC}=1.1\times10^{-3}$ emu and 0.7×10^{-3} emu for $H_{\rm dc}=2$ and 4 G respectively. In order to obtain an optimal overlap between the curves, the following values have been used for $\Delta\mathscr{M}_g$: $3.25\times10^{-3}\cdot H$ cm^3 G at $H_{\rm dc}=0$ G, $3.22\times10^{-3}\times H$ cm^3 G at $H_{\rm dc}=2$ and 4 G. The calculated values for $\Delta\mathscr{M}_j$ at large ΔH are extremely sensitive to those for $\Delta\mathscr{M}_g$. This allows us to refine the determination of $\Delta\mathscr{M}_g$. Note that the values quoted above do not differ by more than 1%, which is compatible with experimental accuracy and the possible variations of grain response with $H_{\rm dc}$.

Finally, from the knowledge of the true Josephson shielding response in ''large'' fields, we can now derive a rough evaluation of the critical current. Specifically, above the maximum of $\Delta\mathscr{M}_j$, we calculate the value J_c of the average critical current which would give the value of the measured moment by use of the Bean formula[19] in a cylindrical geometry. For strong penetration, the magnetization is given in e.m.u. by $M=J_cR/3$. With $R=0.15$ cm and the values of the moment measured at 10 and 20 K with $H_{\rm dc}=0$ G, we obtain the data displayed in Fig. 11. Note that the data are limited to fields such that $H\approx H^*=4\pi J_cR$, below which the above approximate evaluation is no longer relevant.

3.2. Low field d.c. response

We can now concentrate on the behavior of the Josephson current moment at small ΔH. For this discussion, the current susceptibility $\Delta\mathscr{M}_j/\Delta H$ is plotted against ΔH at 10 and 20 K in Figs. 12 and 13, respectively. At 10 K after zero

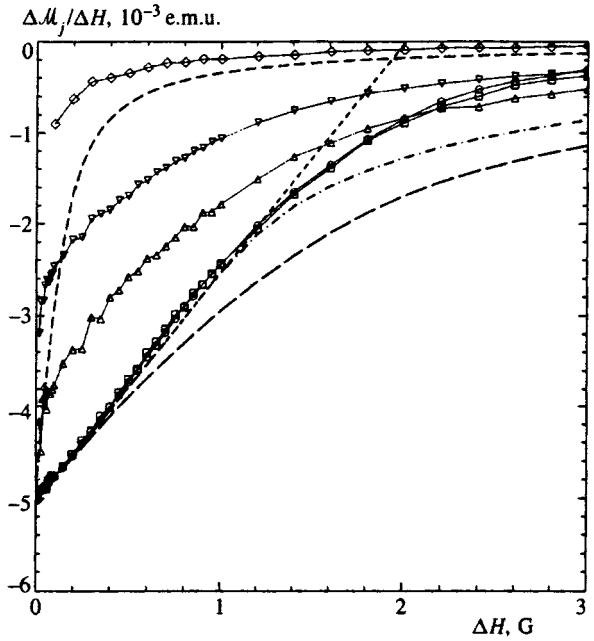


FIG. 12. Josephson current susceptibility at 10 K vs. applied variation ΔH of field, after cooling in d.c. field $H_{\rm dc}=0$ (○), 0.1 (□), 2 (△), 4 (▽), 10 (◇) G. The meaning of dashed and dot-dashed lines is explained in the text; $T=10$ K.

field cooling or cooling in a small field $H_{\rm dc}=0.1$ G, the response varies linearly with ΔH for small values of ΔH up to about 0.5 G. This linear slope of $\Delta\mathscr{M}_j/\Delta H$ is considered to be the result of classical Bean-like pinning with critical current density $J_c=H^*/4\pi R$, where $1/4\pi H^*$ is the initial slope

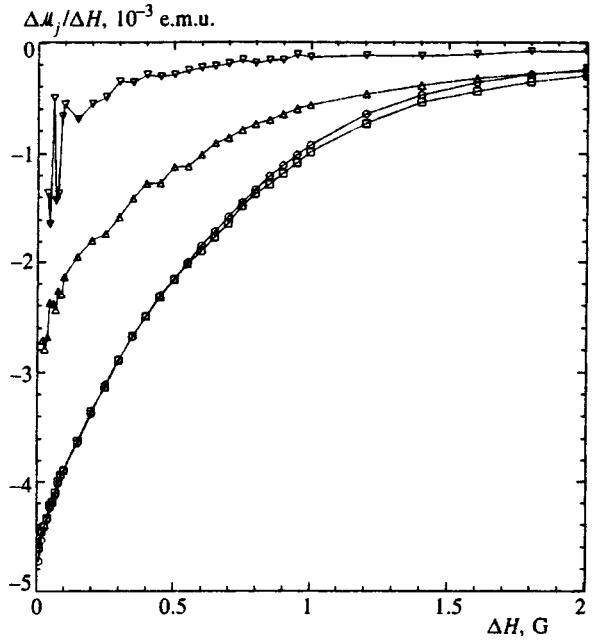


FIG. 13. Josephson currents susceptibility at 20 K vs. applied variation ΔH of field, after cooling in d.c. field $H_{\rm dc}=0$ (○), 0.1 (□), 2 (△), 5 (▽). $T=20$ K.

of the curve.[19] This initial slope is plotted in Fig. 12 as the short-dashed line, which corresponds to $H^*=2$ G, leading to $J_c \approx 3.7$ A/cm^2.

At larger ΔH, the behavior of current susceptibility $\Delta \mathcal{M}_j/\Delta H$ deviates from linear, which is the result of both the magnetic-field dependence of the critical current J_c, an intrinsic effect, and the increasing degree of flux penetration into the sample, a purely size-dependent effect. Usually one uses the Bean model (generally with some B-dependent critical current) in an appropriate geometry to deconvolve these two effects.

However, one should keep in mind that the Bean model is a severe simplification of the problem of constant pinning force, corresponding to the limit $\lambda \to 0$ (i.e., the London penetration depth is assumed to be negligible with respect to the Bean penetration length). For the simplest sample shapes (thin slab or cylinder), this means that the condition $\lambda \ll R$ should be fulfilled, which is usually the case. However the situation is more complicated for samples of square cross-section (like the present one), where the effect of corners may become important even at $\lambda \ll R$. For such a geometry, use of the Bean model leads to exactly the same relation between critical current, external field, and measured magnetization as for cylindrical samples, whereas one expects some difference if finite-λ corrections are taken into account. At the present time, we are not able to evaluate these corrections, nor therefore the value of the magnetization corresponding to our experimental geometry with nonnegligible λ. Nevertheless, we expect that it lies between the curves for two extreme limits. The upper one corresponds to the $\lambda \to 0$ limit, where the magnetization is given simply by Bean's formula for the cylinder:

$$4\pi M/H = -1 + H/H^* - H^2/3H^{*2} \quad \text{for } H < H^*$$

and

$$4\pi M/H = -H^*/3H \quad \text{for } H > H^*.$$

A lower limit (thought rather artificial) consists of the ''double slab'' case in which the variation of magnetization is counted twice (once for each pair of edges):

$$4\pi M/H = -1 + H/H^* \quad \text{for } H < H^*/2$$

and

$$4\pi M/H = -H^*/4H \quad \text{for } H > H^*/2.$$

Both curves are plotted in the Fig. 12 (dot-dashed and long-dashed curves respectively) for $J_c = 3.7$ A/cm^2 and $\Delta \mathcal{M}_j/\Delta H = -5.05 \times 10^{-3}$ cm^3 at $\Delta H \to 0$.

We now discuss the data, starting from that obtained for low d.c., fields, $H_{dc}=0$ and 0.1 G. One can see that after the initial linear part, the absolute value of the measured susceptibility is always less than the calculated value. This corresponds to a decrease in J_c with increasing induction, as classically expected in granular materials, due to the suppression of intergrain critical currents by magnetic field penetration into the Josephson junctions.[9] This ''classical'' behavior for granular superconductors is usually analyzed by considering the volume-averaged Josephson medium as a kind of type-II superconductor in the dirty limit, provided its macroscopic penetration depth λ_J is large compared with the grain size.[11,34]

At $H_{dc} \geqslant 2$ G, the behavior of $\Delta \mathcal{M}_j/\Delta H$ is quite different: there is no initial linear slope, but a monotonic curvature is present down to the smallest ΔH. It is no longer possible to fit a Bean-like curve to the data. For instance, the Bean curve plotted for the lowest ΔH data for $H_{dc}=2$ G is represented in Fig. 12 as a dashed line. It corresponds to a very small critical current of order 0.2 A/cm^2, and it is evident that the effective screening current becomes much larger with increasing ΔH. Here, in contrast to the case of $H_{dc}=0$ G, the absolute value of the measured susceptibility is always larger than the calculated value for a constant shielding current corresponding to the limit $\Delta H \to 0$. This means that whereas at $H_{dc}=0$ the effective screening current density stays constant and then slowly decreases with increasing ΔH (which corresponds to classical Josephson pinning), at $H_{dc} \geqslant 2$ G it increases with ΔH sublinearly (since a linear increase would correspond to a susceptibility independent of ΔH). Such behavior is quite unusual within the commonly accepted picture of screening in superconductors; indeed, we know that for vanishing field excitations, the screening current may be either linear in ΔH and reversible, as in the London (or Campbell[20]) shielding regime, or constant (equal to the initial critical current J_c) and irreversible as in the case of the Bean-type critical state (or of any other known critical model, e.g., Kim model,[32] exponential model,[33] etc).

The above anomalous screening behavior is even more pronounced at 20 K, where even after zero field cooling, no initial linear slope of $\Delta \mathcal{M}_j/\Delta H$ can be seen in the data. All curves show the same anomalous behavior as the data at 10 K in fields from 2 G. This specific behavior is emphasized by plotting the difference between the measured susceptibility $\Delta \mathcal{M}_j/\Delta H$ and its value for total flux expulsion $\Delta \mathcal{M}_j(0)/\Delta H$ against ΔH on a log–log scale. In such a plot, at least in the regime of weak penetration, i.e., where $\Delta \mathcal{M}_j/\Delta H$ is greater than $0.8\Delta \mathcal{M}_j(0)/\Delta H$, sublinear variation of the shielding current density results in a logarithmic slope less than 1 for the curves of $\Delta \mathcal{M}_j/\Delta H$ (for $\Delta \mathcal{M}_j/\Delta H$ less than $0.8\Delta \mathcal{M}_j(0)/\Delta H$, we are in a regime of strong penetration, where it is no longer possible to relate simply the variations of the moment response to those of the shielding current). In Fig. 14, we have plotted the three curves for which data are found in the range above $0.8\Delta \mathcal{M}_j(0)/\Delta H$, i.e., at $T=10$ K, $H_{dc}=0$ and 2 G, and $T=20$ K, $H_{dc}=0$ G.

At 10 K and $H_{dc}=0$ G, the logarithmic slope is about 1 as expected, although in the weakest fields the logarithmic slope becomes closer to 0.5. At 20 K and $H_{dc}=0$ G, the logarithmic slope is about 0.4 at the lowest ΔH. Approximately the same value of the slope characterizes the data obtained at 10 K and $H_{dc}=2$ G, although the dispersion of data points at lowest ΔH makes its accurate determination difficult.

The above anomalous behavior makes it tempting to try a simple Ansatz for the behavior of the response current density of the system versus induction variations. Let us suppose that $J \propto \Delta B^{\alpha}$, with α between 0 and 1. The case with

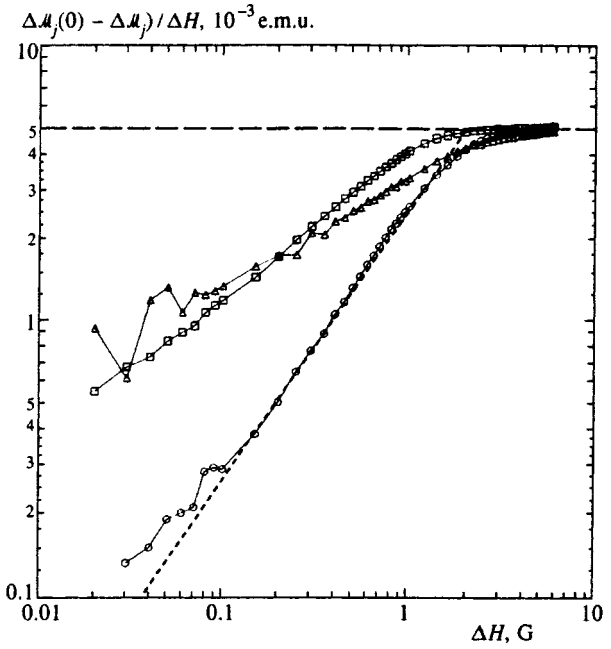


FIG. 14. Difference between the measured susceptibility and its value for perfect shielding for selected data at 10 and 20 K. The short dashed line represents a logarithmic slope 1 expected for a Bean critical state. ○—$T=10$ K, $H=0$ G; △—$T=10$ K, $H=2$ G; □—$T=20$ K, $H=0$ G.

$\alpha=1$ corresponds simply to classical screening with penetration length λ (since $J\propto\Delta B$); the case with $\alpha=0$ corresponds to constant J, i.e., the classical Bean case. Anomalous response arises for noninteger α. For very weak excitation ΔH, the length of induction penetration is small compared with the size of the sample, and we need to consider the effect of the excitation to lowest order in ΔB only. For the purpose of illustration we consider the simplest slab geometry. Then the induction profile is determined by the Maxwell equation

$$\frac{dB}{dx}=-4\pi J_1\left(\frac{\Delta B}{\Delta B_1}\right)^{\alpha}, \tag{3.2}$$

where x is the coordinate perpendicular to the edge of the sample. For an external field ΔH, the induction in the sample is given by

$$\Delta B(x)=\left(\frac{(1-\alpha)4\pi J_1}{\Delta B_1^{\alpha}}\,(x_H-x)\right)^{1/(1-\alpha)}, \tag{3.3}$$

where x_H is the coordinate of penetration and J_1 and ΔB_1 are normalizing factors; $\Delta B=\Delta H$ for $x=0$, i.e.,

$$x_H=\frac{\Delta B_1^{\alpha}}{4\pi J_1(1-\alpha)}\,\Delta H^{1-\alpha}.$$

Then, integrating the field profile (3.3) over x, we get

$$\frac{4\pi\Delta\overline{M}+\Delta H}{\Delta H}\propto\Delta H^{1-\alpha}, \tag{3.4}$$

where $\Delta\overline{M}=\Delta\mathcal{M}/V$ is the mean magnetization variation due to the field variation ΔH.

If we now compare the result (3.4) with the data shown in Fig. 14, we find values of α in the range 0.4–0.5 at both 10 and 20 K.

Thus, a simple choice for the relation between the screening current J_c and the induction variation ΔB allows us to imitate the experimental results for the simplest protocol of a weak monotonic ΔH variation on top of a homogeneous state of the network. Nevertheless, it is evident that ΔB has no clear meaning if the variation of H is nonmonotonic, or if the initial state is obtained by nonzero field cooling. Indeed, in the latter case, induction in the sample varies during cooling due to the variation of μ_g with T, giving the response $\Delta\mathcal{M}_j^{\mathrm{FC}}$ as seen before. Furthermore, we will see below that the response is irreversible, even for extremely low excitation fields.

3.3. Irreversibility: very low field, low frequency a.c. response

Problems of sensitivity limit the range of small excitations which can be used in d.c. experiments. The preceding results clearly show the sublinear nature of the low field response, but they do not allow its precise determination. In order to extend the range of our lower excitations investigation by several orders of magnitude, we have been led to perform a.c. susceptibility measurements. The use of a.c. response measurements is always questionable when equilibrium (or quasiequilibrium) properties are under investigation, since the results can be affected by the time-dependent part of the response function. It has been shown that the latter is the response of a very good conductor with complex conductivity.[35,36] Hence, it is necessary to work at low frequency, in a range where the susceptibility is roughly frequency-independent.

We present here preliminary results obtained on a long cylinder produced by stacking several of the original sample B cylinders. Measurements were made at 20 K, at a working frequency of 1.7 Hz in the equipment used for noise experiments.[36] The sample was simply shifted into the upper half of the third-order gradiometer. At this temperature and frequency, we have verified that the in-phase susceptibility is almost frequency-independent, which ensures that the results are mainly dependent on the (quasi)static part of the response. The susceptibility was recorded using the classical method of SQUID magnetometry. We used a.c. excitation fields in the range 3×10^{-2}–30 mG, and the sample was cooled in d.c. fields from 0 to 8.8 G. From the data, the values of the Josephson network susceptibility were derived using the method developed in Sec. 2, with the susceptibilities in Eq. (2.6) being complex quantities. The susceptibility measured at 4.2 K at the lowest a.c amplitude was taken to be the level for perfect diamagnetism.

Figure 15 displays a log–log plot of the out-of-phase susceptibility χ_j'' versus the amplitude of the a.c. field, and for several values of the FC static field. The response is irreversible down to the lowest a.c. amplitudes, and the irreversibility increases with the superimposed d.c. field. All curves follow a power law, with the same exponent close to 0.5. Going towards the smallest a.c. excitations, they show some downward bend, which could be related to the ap-

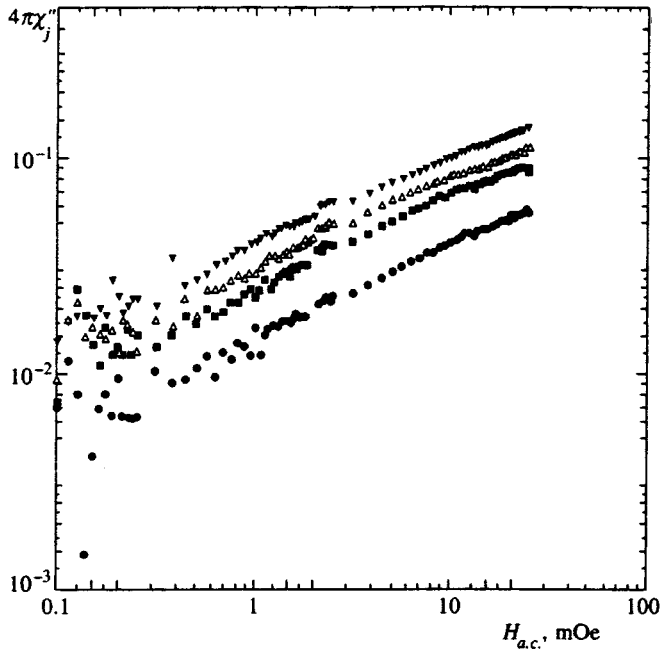


FIG. 15. Out-of-phase susceptibility at 1.7 Hz as a function of a.c. field amplitude. $T=20$ K, $H_{dc}=0$ (●), 2.2 (■), 4.4 (△), 8.8 (▼) Oe.

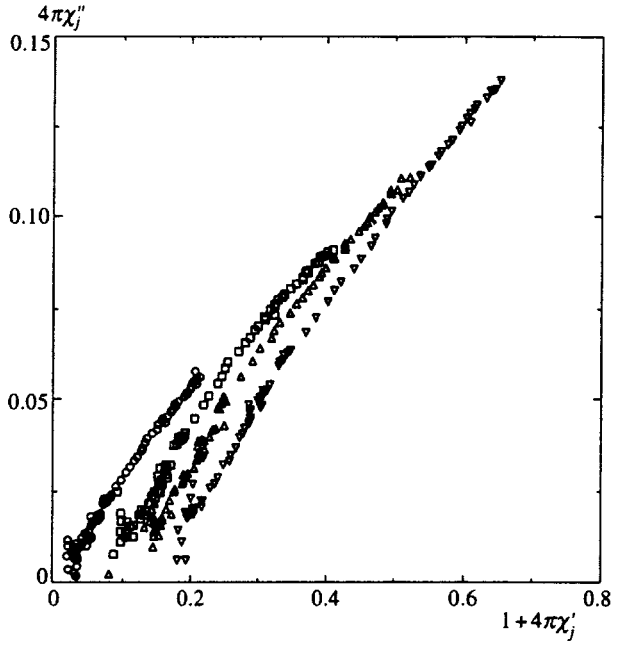


FIG. 17. Plot of $4\pi\chi_j''$ as a function of $1+4\pi\chi_j'$ for $H_{dc}=0$ (○), 2.2 (□), 4.4 (△), 8.8 Oe, $T=20$ K, $f=1.7$ Hz.

proach to a linear regime (with $\chi_j''=0$) below 0.1 mG, although the variance of the data is too great to draw any conclusions.

The in-phase susceptibility χ_j' is plotted as a function of $H_{ac}^{0.5}$ in Fig. 16. Here as well, the anomalous nature of the response is clearly seen. $4\pi\chi_j'$ behaves like $(-1+\delta+\gamma H_{ac}^{0.5})$, where both the constant δ and the slope γ increase with increasing superimposed static field H_{dc}.

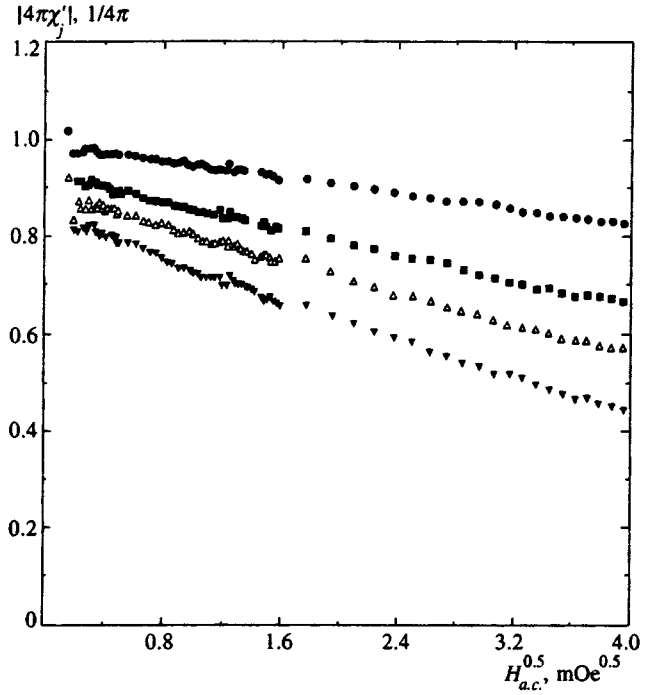


FIG. 16. In-phase susceptibility at 1.7 Hz as a function of the a.c. field amplitude $H_{ac}^{0.5}$. $T=20$ K, $H_{ac}=0$ (●), 2.2 (■), 4.4 (▲), 8.8 (▼) Oe.

The dependence of the harmonic susceptibility on the a.c. field amplitude provides genuine proof of the existence of static irreversibility in the response. This is not surprising by itself, but these results stress the anomalous aspect of this irreversibility. For instance, in the classical Bean case with weak penetration, it is known that $1+4\pi\chi_j'$ and χ_j'' are proportional to H_{ac}, whereas Figs. 15 and 16 clearly show the proportionality to $H_{ac}^{0.5}$. Further evidence is provided by plotting χ_j'' against $1+4\pi\chi_j'$, as displayed in Fig. 17. It can be shown that if the a.c. response is driven only by static irreversibility, the two will be proportional. In the Bean case, the coefficient of proportionality is $4/3\pi$. In Fig. 17, data that fall within the range in 20% variation in χ_j' (where the relations for slab geometry are approximately valid) show that χ_j'' is indeed proportional to $1+4\pi\chi_j'$, but with a bit smaller coefficient $\approx 0.28\pm 0.03$.

In order to understand the meaning of the above results, we generalize the crude *ad hoc* model of Sec. 3.2 to the irreversible case. In order to do so, we generalize the protocol of the Bean model. Specifically, in the Bean model, the current is given by a step function of the variation of induction, $J=J_c\,\mathrm{sign}(\Delta B)$ according to the sign of ΔB, as long as the induction variation is monotonic. If the sign of variation of B is reversed, J also changes sign; this can be written in terms of the variation of the current density (with respect to the initial current distribution obtained after monotonic variation of the field, J_{init}) $\Delta J=-2J_{init}\Theta(-\Delta B_{new})$ where $\Theta(x)=(1/2)(1+\mathrm{sign}\ x)$ and $\Delta B_{new}=B-B_{init}$. Such a representation (which is not needed in the analysis of the Bean model itself) will allow us to construct the necessary generalization of the relation between current and variation of the field used in Eq. (3.2).

Actually our goal here is rather limited: we wish to find a consistent description of the simplest hysteresis cycle,

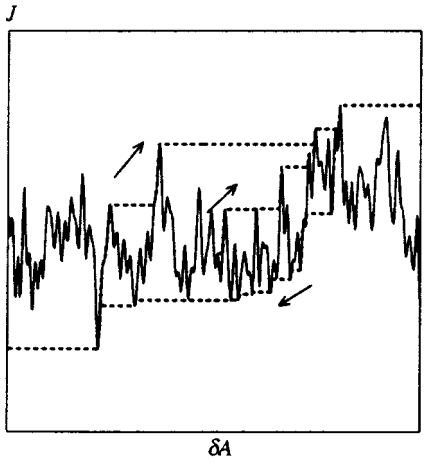


FIG. 18. Picture of a fractal $J(\delta A)$ landscape. An example of a hysteresis loop is shown.

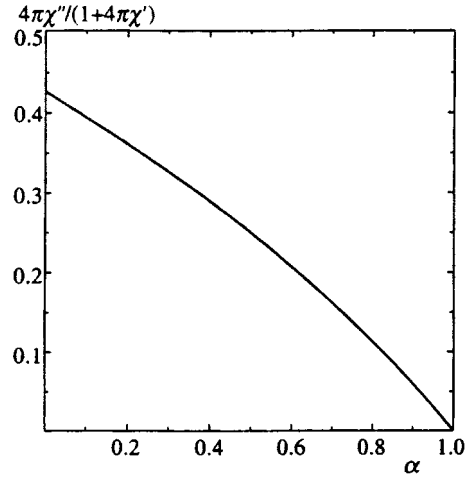


FIG. 19. Values of $4\pi\chi''/(1+4\pi\chi')$ as a function of exponent α.

which consists of an initial increase in ΔB from zero to $\Delta B_{\rm init}$, reversal of the sign of the field variation until the value of $\Delta B=-\Delta B_{\rm init}$ is reached, and another reversal of dB/dt, finishing at $\Delta B_{\rm final}=\Delta B_{\rm init}$. The description of this cycle will be consistent if we find that the value of the current density at the end-point, $J_{\rm final}$, coincides with the value after the original increase in the field $\Delta B_{\rm init}\, J_{\rm init}$. This simply means that the hysteresis loop is closed. It is easy to check that the above condition will be fulfilled by the following choice of the $\Delta J(\Delta B_{\rm new})$ dependence:

$$\Delta J=-\,{\rm sign}(J_{\rm init})2^{1-\alpha}J_1\left(\frac{\Delta B_{\rm new}}{\Delta B_1}\right)^{\alpha}\Theta(-\Delta B_{\rm new}), \quad (3.5)$$

where J_1 and ΔB_1 have the same meaning as in Eq. (3.2).

Actually, the only difference between the Eq. (3.5) and the original used in Eq. (3.2) is the coefficient $2^{1-\alpha}$. The Bean model limit then corresponds to $\alpha\to 0$, so the above coefficient approaches 2, as it should. Then instead of Eq. (3.2) we obtain

$$d\Delta B/dx=\pm 2^{1-\alpha}\mathscr{A}\,\Delta B^{\alpha}, \quad (3.6)$$

where $\mathscr{A}=4\pi J_1/\Delta B_1^{\alpha}$.

The induction profile, induced magnetization and harmonic response are calculated in the Appendix. The main conclusions are that the fundamental components $1+4\pi\chi_j'$ and χ_j'' are both proportional to $h_0^{1-\alpha}$, and that their ratio $R=\chi_j''/(1+4\pi\chi_j')$ decreases from $4/3\pi$ to 0 when α goes from 0 to 1. For $\alpha=0.5$, we get (see Fig. 19) $R\approx 0.25$, a value which is in good agreement with the data presented on Fig. 17. Note that the degree of irreversibility (measured by this ratio) is similar (although a bit lower) to that of the Bean model.

It should be emphasized that the numerical coefficient in Eq. (3.5) was ''fit'' in order to obtain a consistent (i.e., closed) hysteresis loop. One can expect that an analogous equation describing current variation after some more complicated history of field variations will contain a different (history-dependent) numerical coefficient instead of $2^{1-\alpha}$.

It can be seen from Figs. 15 and 16 that $\mathscr{A}=4\pi J_1/\Delta B_1^{\alpha}$ increases with increasing ambient d.c. field. It is natural to expect a decrease in J_1 with increasing $H_{\rm dc}$. The increase in $\mathscr{A}$ with $H_{\rm dc}$ means that ΔB_1^{α} decreases more quickly than J_1 when $H_{\rm dc}$ increases.

The presence and behavior of the constant δ cannot be predicted on the basis of the above simple model. In fact, the latter neglects the possibility of elastic displacement of flux lines under the action of the external applied field. Such an effect would result in a response analogous to the Campbell response due to the elastic displacement of vortices in their pinning potential in type-II superconductors.[20] Campbell's response is linear, so it would result in an a.c. field-independent positive contribution to χ', whose amplitude would be inversely proportional to the strength of the restoring force of the pinning potential wells. It is natural to expect that the pinning force decreases with increasing ambient d.c. field in our granular system, due to the reduction in junction critical currents. Hence, such an effect would make a positive contribution to χ' that would increase with increasing d.c. field. This corresponds rather well to the behavior of the offset δ seen in the data.

4. COMPARISON WITH AN EXISTING THEORY OF GAUGE GLASS: FRUSTRATION AT $H=0$

In this section we compare the experimental results described above with the theoretical results available for randomly frustrated Josephson networks. We start with a simple estimate of the mean energy $E_J=\hbar I_c/2e$ using the experimental value of the low-temperature ZFC ($T=10$ K, $H_{\rm dc}=0$ G) critical current density $J_c\approx 3.7$ A/cm^2. Using the estimate $a_0\approx 5\ \mu$m for the mean size of the grains, one could naively obtain $I_c\approx J_c a_0^2\approx 1\ \mu$A and the corresponding low-temperature Josephson energy $E_J^{\rm naive}\approx 20$ K (this value was derived from J_c measured at $T=10$ K, but we do not expect much difference in the intrinsic Josephson energies at $T=10$ K and at $T\to 0$, since the bulk transition temperature in $La_{1.8}Sr_{0.2}CuO_4$ is $T_c\approx 32$ K).

However, such an estimate is in conflict with the measured value of the glass transition temperature $T_g\approx 29$ K. Indeed, let us assume that the mean coordination number (number of ''interacting neighbors'') Z in the ceramics is

around 6, as for a simple cubic lattice. Then for an estimate of the relation between E_J and T_g, one can use simulation data,[6,7] which give $T_g \approx 0.5E_J(T_g) = 0.5E_{J_0}(1-T_g/T_c)$, where we took into account the linear dependence of E_J on $T_c - T$ close to the bulk transition temperature. As a result, one gets

$$\frac{E_{J_0}}{k_B} \approx \frac{2T_g}{1-T_g/T_c} \approx 600 \text{ K}, \tag{4.1}$$

i.e., 30 times the naive estimate above. However we now show that this discrepancy can be resolved if we assume that the current network producing the measured critical current density J_c was actually strongly frustrated in spite of the lack of a background d.c. field in this measurement.

The macroscopic critical current density J_c for a strongly frustrated Josephson network is calculated in Ref. 5 using the mean-field approach (we are not aware of any calculations of this kind beyond the scope of mean-field theory). It is shown that frustration significantly reduces J_c as compared to its value J_0 for an unfrustrated system:

$$\frac{J_c}{J_0} = \frac{3\sqrt{3}\gamma}{8}(1-T/T_g)^{5/2},$$

where the factor $\gamma \approx 0.065$ was obtained by numerical solution of the slow cooling equations (Refs. 3, 37, 4, p. 183) describing the evolution of the glassy state under slow variations of temperature and magnetic field. In the low-temperature limit, this relation amounts to a factor of 25 reduction in the J_c value with respect to J_0. Correspondingly, the characteristic value of the critical current for an individual junction will be $I_c \approx 25 J_c a_0^2 \approx 25\ \mu$A, which results in a Josephson coupling energy $E_{J_0} \approx 500$ K, in fairly good agreement with the above estimate (4.1).

The above estimates show that the network of Josephson junctions in $La_{1.8}Sr_{0.2}CuO_4$ is frustrated even in the absence of an external magnetic field. A careful reader might question this conclusion, since we have used some results from mean-field theory that may be a poor approximation for a 3D gauge-glass. We believe, however, that the qualitative result of the above estimates is robust, because a significant reduction in J_c with respect to J_0 should be a general feature of a glassy network, so that inaccuracy due to mean-field approximation cannot compensate for a huge discrepancy obtained between E_J^{naive} and the estimate (4.1).

Additional evidence of the glassy nature of our system is provided by the similarity of the low-ΔH diamagnetic response at $T = 20$ K for zero and nonzero H_{dc}, as described in Sec. 3 above, as well as the low-frequency noise data obtained in Ref. 35 on the same type of ceramics.

What might the origin of that frustration be? We believe that most probably it is the result of the *d*-wave nature of superconductivity in cuprates[15] and randomness of the crystalline orientations in ceramics.[13,14] It was shown there that the form of the effective phase-dependent Hamiltonian for such ceramics is of the same form as in (1.1) except for the fact that the random phases α_{ij} at $B = 0$ are just 0 or π, depending on the mutual orientation of grains i and j. Therefore, such a system is equivalent at $B = 0$ to an XY spin-glass, with the low-temperature state characterized by a completely random orientation of phases ϕ_i, as in the gauge-glass model with uniformly random distribution of α_{ij}'s.

Therefore, the low-temperature state is characterized by the presence of randomly distributed intergrain currents and, therefore, of the magnetic field generated by those currents. It means that the actual phases α_{ij} will contain contributions due to the self-induced magnetic field. Its relative importance is characterized by the ratio of the corresponding magnetic flux penetrating elementary loops of the ceramics Φ_{sf} to the flux quantum Φ_0, i.e., just by the parameter $\beta_L = 2\pi \mathscr{L} I_c / c\Phi_0$, where $\mathscr{L}$ is the characteristic inductance of an elementary loop.[9] Estimating the elementary inductance as $\mathscr{L} \approx 2\pi a_0 \mu_g$ and using Eq. (4.1) to estimate I_c, we obtain

$$\beta_L \approx \frac{4\pi^2 a_0 \mu_g I_c}{c\Phi_0} = \frac{8\pi^3 \mu_g a_0 E_J}{\Phi_0^2} \approx 0.1, \tag{4.2}$$

so that self-field (screening) effects are relatively weak, though perhaps not always negligible.

It is also of interest to estimate the effective penetration depth λ_{cer} into the ceramics of a very weak magnetic field perturbation δH. Roughly, the value of λ_{cer} can be estimated to be $a_0/\sqrt{\beta_L} \sim 15\ \mu$m. Another (hopefully more accurate) estimate can be obtained using mean-field results,[5] which enable one to express λ_{cer} via the critical current density J_c:

$$\lambda_{\text{cer}} = \left(\frac{\gamma}{8\pi^2} \frac{c\Phi_0}{J_c \xi_0 \mu_g} \right)^{1/2} \approx 25\ \mu\text{m}, \tag{4.3}$$

where we inserted (as compared with Ref. 5) $\mu_g \approx 0.35$, and approximated the random nearest-neighbor network by a cubic lattice with coordination number $Z = 6$, which amounts to the relation $\xi_0^2 = a_0^2/6$ between the effective interaction range ξ_0 and the intergrain distance a_0.

The characteristic magnetic field variation producing the critical current density J_c at the boundary can be estimated to be $\Delta H_c \sim 4\pi \lambda_{\text{cer}} J_c / c \approx 15$ mG, whereas the numerical solution[5] yields

$$\Delta H_c = \frac{v}{2c\gamma} 4\pi \lambda_{\text{cer}} J_c \approx 30 \text{ mG}. \tag{4.4}$$

In the theoretical approach of Ref. 5, ΔH_c marks a crossover between reversible (although still nonlinear at $\Delta H \leqslant \Delta H_c$) and irreversible penetration of the magnetic field into the intergrain network. The value of ΔH_c obtained in Eq. (4.4) is at the lower end of the range of field variations used to measure our d.c. magnetization curves, so we might just conclude that we always have $\delta H \gg \Delta H_c$, and thus are producing the Bean-like critical state. Indeed, the data at $H_{\text{dc}} = 0$, $T = 10$ K look compatible with such an interpretation (see Fig. 14), where some deviations from the logarithmic slope 1 (which is characteristic of a Bean state) are seen at lowest $\Delta H \leqslant 50$ (mG).

However, as far as the data obtained at 10 K with d.c. fields $H_{\text{dc}} \simeq 2$ G are concerned, or all data at higher temperature ($T = 20$ K), including d.c. and a.c. results at zero-H_{dc} (see Figs. 12–16), the low-field magnetization response is drastically different from Bean-type predictions, as explained at the end of Sec. 3. Qualitatively, the most surprising feature

of these data is the existence of a very broad range of ΔH within which the response is nonlinear but still not like the critical-state response. We are unaware of any microscopic theory that predicts fractional-power behavior of the shielding susceptibility over such a broad range of ΔH variations. It cannot be ruled out *a priori* that such a behavior is related to a very wide range of intergrain critical currents, which might exist in ceramics (we have thus far neglected inhomogeneity of intergrain coupling strengths in our theoretical discussion). Moreover, we may expect that the relative importance of such inhomogeneities increases with the field and/or temperature (see Ref. 38).

In Sec. 5, we attempt to formulate a new phenomenological model appropriate for the understanding of our data (leaving its theoretical justification for a future study); this model will be seen to be an interpolation between Campbell's and Bean's regimes of flux penetration into hard superconductors.

5. FRACTAL MODEL OF DIAMAGNETIC RESPONSE

We showed at the end of Sec. 3.3 that a simple generalization, Eq. (3.6), of Bean's relation between variation of the applied magnetic induction ΔB and current J results in reasonably good agreement with our data. However, in contrast to the original Bean relation, the new one was not based on any physical model—it was just a convenient description of the data. In this Section we propose a phenomenological model that provides a qualitative understanding of the irreversible diamagnetic behavior mimicked by Eq. (3.6).

We begin with the nonlinear response of the current $\mathbf{J}$ to a variation of the vector potential $\delta\mathbf{A}$ derived in Ref. 5 via the mean-field approximation and presented in Fig. 2 of that paper. Here the current induced by a variation of $\delta\mathbf{A}$ is linear at very small $\delta\mathbf{A}$, then grows sublinearly, and finally reaches its maximum value J_c at the critical δA_c such that the differential response $(dJ/dA)_{\delta A_c} \to 0$. For $\delta A > \delta A_c$ the numerical instability of the slow cooling equations was detected and interpreted as an indication of the absence of any solution that would interpolate smoothly between zero and large values of δA (i.e., much greater than δA_c). In other words, some kind of ''phase slip'' was expected to occur in the model,[5] leading to a new metastable state with lower (free) energy at the new value of the vector potential $\mathbf{A}' = \mathbf{A} + \delta\mathbf{A}$ (in other words, a state similar to the one obtained by the FC procedure at constant $\mathbf{A}'$, which does not carry macroscopic current). A further increase in $\delta\mathbf{A}' = \mathbf{A} - \mathbf{A}'$ again induces a macroscopic current until it reaches the maximum value J_c at $\delta A = \delta A_c$, and so on.

Thus, the whole $J(\delta A)$ dependence emerging from the mean-field solution[5] is periodic; it leads immediately to the irreversibility of the response, since the inverse function $\delta A(J)$ is multivalued: different vector potential values may correspond to the same value of current. Of course, such a periodic $J(\delta A)$ dependence does not correspond to the usual Campbell–Bean picture, which would be better represented by

$$J_{CB}(\delta A) = J(\delta A)\,\theta(\delta A_c - \delta A) + J_c\,\theta(\delta A - \delta A_c). \quad (5.1)$$

It is important to note that the $J(\delta A)$ dependence in Ref. 5 was obtained from the space-independent solution for the glassy correlation function $Q_{jj}(t,t') = \langle \cos(\phi_j(t) - \phi_j(t')) \rangle$. Such an approximation, which is reasonable for the description of smooth ''adiabatic'' transformations in a system with long-range interactions, will probably break down when a jump occurs from one metastable state to another. In other words, the aforementioned ''phase slip'' should have something to do with spatially inhomogeneous processes like vortex penetration in hard type-II superconductors.

The problem of the solution of the general history- and space-dependent system of integral equations (which can be derived following the method of Ref. 5) is formidable, and the method to solve it is still unknown. Therefore we can only speculate on possible properties of its solution. The simplest idea would be that the macroscopic $J(\delta A)$ response becomes similar (after averaging over inhomogeneities of the space-dependent solution) to the Campbell–Bean type of response (5.1). Indeed, our analysis of the low-field diamagnetic response at $T = 10$ K and $H_{\rm ext} = 0$ (Sec. 3.2), developed in Sec. 4 on the basis of such an assumption, is in reasonable agreement with the data. However other sets of data (for higher temperature and/or lower field) are described by the completely different Ansatz (3.5).

We now propose a (phenomenological) generalization of the $J(\delta A)$ relation compatible with Eq. (3.5). The relation we are looking for should be an intrinsic (i.e., independent of the sample geometry) and general (i.e., usable for an arbitrary magnetic history of the sample) relation between the current and variation of the vector potential. Recall that Eq. (3.5) was written for the simplest nonmonotonic variation of ΔB, and that it relates the true vector $\mathbf{J}$ and the pseudovector $\Delta\mathbf{B}$. Thus, in writing this equation, some additional information on the geometry of the sample has been used (we use the simplest slab geometry). A natural basic equation should therefore relate the current density $\mathbf{J}$ and the variation of the vector potential $\delta\mathbf{A}$.

In a generalized model, the diamagnetic current response must possess two major properties: it must scale as some fractional power $\alpha \approx 0.5$ of the amplitude of the excitation field δB, and it must be strongly irreversible (as follows from an analysis of the ratio $4\pi\chi''/(1 - 4\pi\chi') \approx 0.28$). We consider these two conditions in turn.

The first condition is rather easy to fulfill: it is enough to suppose that the differential response of the current to the variation of the vector potential δA is given by a nonlinear generalization of the London relation

$$\frac{d\mathbf{J}}{d\mathbf{A}} = -\frac{c}{4\pi}\,\lambda_{\rm eff}^{-2}(J), \quad (5.2)$$

where the current-dependent ''effective penetration depth'' is given by

$$\lambda_{\rm eff} = \lambda_1 |J/J_1|^{\kappa}. \quad (5.3)$$

In the case of a monotonic field variation applied to an initially uniform induction distribution, Eqs. (5.2) and (5.3) lead to the simple relation $J \propto \Delta B^{\alpha}$ with $\alpha = (1 + \kappa)^{-1}$. Indeed, with $dA = \Delta B\, dx$ and approximating $d\Delta B/dx$ by

$\Delta B/\lambda_{\rm eff}$, one obtains $J\propto\Delta B^{1/(1+\kappa)}$. Thus we need to choose $\kappa\approx 1$ in order to reproduce the observed scaling with $\alpha \approx 0.5$.

However, the set of equations (5.2), (5.3) does not satisfy the second condition above: the corresponding solutions are reversible, as follows from the existence of a single-valued function $\delta A(J)\propto J^{1+2\kappa}$, which follows from Eqs. (5.2), (5.3). In other words, the system described by Eqs. (5.2), (5.3) would exhibit nonlinearity and harmonic generation, but would not show finite $\chi''(\omega)$ in the limit $\omega\to 0$. In order to avoid this inconsistency, we need to formulate a model with the same kind of scaling between δA and J as in Eqs. (5.2), (5.3), but with a nonmonotonic $J(\delta A)$ dependence that allows for irreversible behavior.

A model obeying very similar properties was formulated and studied in Sec. 3.2 of Ref. 4 in a different physical context (one-dimensional spin-glass). The low-energy spin configurations in this model are described by a phase variable $\varphi(x)\in(-\pi,\pi)$ such that two such configurations (local energy minima) that differ by a phase shift $\delta\varphi(x_0)=\Phi$ in a region around some point x_0 have a characteristic energy difference $E(\Phi)\propto\Phi^{5/3}$ and a characteristic spatial extent of phase deformation $X(\Phi)\propto\Phi^{1/3}$. This scaling holds for the intermediate range of phase deformations $\varphi_0\ll\Phi\ll\pi$; at smaller $\Phi\leqslant\varphi_0\ll 1$, the energy cost of deformation is $\propto\Phi^2$, whereas at $\Phi\sim\pi$ the energy growth obviously saturates due to 2π periodicity.

The above $E(\Phi)$ scaling leads to a sublinear growth of the characteristic ''force'' $f(\Phi)=dE/d\Phi\propto\Phi^{2/3}$ with Φ over the same intermediate range. The leading contribution to the second derivative $d^2E/d\Phi^2$ (curvature of the energy valleys) comes from the smallest scale $\Phi\sim\varphi_0$, i.e., from the curvature of individual local minima. It was explained in Sec. 3.2 of Ref. 4 that such scaling implies fractal structure of the energy minima as a function of φ with fractal dimension $D_f=1/3$. This means that the number of energy minima discernible on a scale φ grows as $\mathcal{N}\propto\varphi^{-1/3}$ at finer scales; new minima appear primarily due to the splitting of older (broader) ones. This picture emerged in Sec. 3.2 of Ref. 4 from a microscopic analysis of the original Hamiltonian for a one-dimensional spin-glass model formulated in Ref. 39. We can borrow the qualitative features of this construction for our present purpose (leaving for future studies the problem of its microscopic justification for the case of superconductive glasses).

Suppose that the free energy $F(\delta A)$ of the Josephson network behaves (as a function of vector potential variations with respect to a ''virgin'' state with a homogeneous induction) in a way similar to $E(\varphi)$ at $\varphi\ll\pi$. Specifically, suppose that the free energy is parabolic, $\delta F\propto(\delta A)^2$, at very small variations of vector potential $\delta A\leqslant\delta A_{c1}$, but on a larger scale, $\delta A\gg\delta A_{c1}$, it contains many local minima whose characteristic free energies scale (with respect to the lowest state with $\delta A=0$) as

$$F(\delta A)\propto(\delta A)^{\theta+1} \tag{5.4}$$

for $\delta A_{c1}\ll\delta A\ll\delta A_c$, with exponent $\theta\in(0,1)$ (see the definition of δA_c below). Then the characteristic value of the current $J=(1/c)\partial F/\partial A$ scales as

$$J_{\rm char}(\delta A)\approx J_{c1}\left(\frac{\delta A}{\delta A_{c1}}\right)^{\theta} \tag{5.5}$$

over the same interval of δA. At larger variations $\delta A \geqslant\delta A_c$, the growth of the induced current should saturate at the true critical current value J_c, so we can estimate

$$\delta A_c\sim\delta A_{c1}(J_c/J_{c1})^{1/\theta}. \tag{5.6}$$

On the other hand, weak $\delta A\ll\delta A_{c1}$ leads to the usual linear London (or Campbell) response with an effective penetration depth λ_1; matching at $\delta A\sim\delta A_{c1}$ leads to the following estimate:

$$\delta A_{c1}\sim\frac{4\pi}{c}J_{c1}\lambda_1^2. \tag{5.7}$$

The estimate (5.5) looks very much like the previous version defined by (5.2), (5.3), so one can find a relationship between the exponents:

$$\theta=1/(1+2\kappa)=\alpha/(2-\alpha)\approx 0.3. \tag{5.8}$$

However the whole picture is substantially altered: the current is now assumed to be an (irregularly) oscillating function of δA (see Fig. 18), so only its envelope $J_{\rm char}(\delta A)$ defined on a scale δA follows the scaling relation (5.5). As a result, the inverse function $\delta A(J)$ is multivalued, and irreversibility of the response is ensured. Similar to the spin-glass model of Ref. 4, Sec. 3.2, the fractal dimension D_f of the low-energy valleys can be defined; it is now given by $D_f=1-\theta\approx 0.7$.

The proposed picture is based on the existence of two substantially different scales of currents, J_{c1} and J_c, and corresponding vector potential variations δA_{c1} and δA_c. It can thus be compared with the usual Campbell–Bean picture of critical currents in the same way as the thermodynamics of type-II superconductors is compared with that of type-I.

To describe quantitatively the diamagnetic response in the ''fractal'' range (5.4), we need to determine the distribution function $\mathcal{P}[J(\delta A)]$ (which would lead, in particular, to the estimate (5.5) for $J_{\rm char}(\delta A)$). Moreover, in general, a relation of the type of (5.5) might be nonlocal (i.e., the current depends on the $\delta A(x)$ distribution over some region of space, whose size may depend on δA itself; see again Ref. 4, Sec. 3.2). We leave this complicated problem for future studies, and only note here that the mere existence of relation (5.5) is sufficient for the existence of some ''natural'' properties of the response (like the existence of a closed hysteresis loop, as assumed in Sec. 3.3).

6. SUMMARY AND CONCLUSIONS

In this paper, we have presented experimental results on the low temperature (10 and 20 K) response of the granular HT_c superconductor $La_{1.8}Sr_{0.2}CuO_4$ to small field excitations. The general properties of the magnetic response were investigated in two samples (A and B) differing by the strength of the coupling between grains. By cooling the samples in vari-

ous d.c. fields up to 20 G and applying small field increases, we were able to measure the shielding response of the material and to derive a method, inspired by the work of Dersh and Blatter,[22] to deduce from the data the polarizability of the intergrain current system. The field cooled magnetization was measured in fields up to 20 G. Analysis of the results leads to the conclusion that a) the structure of the grains is polycrystalline, resulting in a step decrease of the FC magnetization with increasing field, which can be interpreted on the basis of the model by Wohllebeen *et al.*;[28] b) self-shielding (pinning) by intergrain currents when lowering the temperature strongly reduces the value of the FC magnetization; c) there is no macroscopic Meissner magnetization due to the system of intergrain currents.

Further detailed study of the response of the Josephson network was performed in sample B. It was shown that the response is asymmetric with respect to the sign of variation of the applied field after field cooling; this is due to the shielding currents pinned during cooling. The macroscopic critical current is found to be strongly reduced by moderate values of the external d.c. field, about 2 G.

Very low-field magnetization measurements were performed by applying field steps of 10 mG or low-frequency a.c. fields in the range 50 μG to 30 mG after cooling in d.c. fields up to 8.8 G. The results show that the response is highly nonlinear, and the shielding current growing sublinearly with increasing applied field. Furthermore, the a.c. results show that the response is strongly irreversible down to the smallest excitations used. It is shown that a nonlinear relation between the shielding current and the induction, $J \propto \Delta B^{\alpha}$ with $\alpha \approx 0.5$, together with a natural assumption about the existence of a closed hysteresis loop, yield predictions in reasonable agreement with the data.

Theoretical analysis of our experimental results was developed in the context of existing ''gauge-glass'' theories. It was shown that the extremely low value of the low-temperature, zero-field critical current density ($J_c \approx 3.7$ A/cm^2 at 10 K) together with the rather high temperature of the transition to the low-temperature glassy state, can be coherently interpreted only under the assumption that the Josephson network is strongly frustrated even at zero applied field. This contradicts the usual assumption that frustration in the interactions results solely from local magnetic induction, but supports the hypothesis of the existence of a large proportion of π-junctions in the granular system. These π-junctions are possibly due to the d-wave nature of the pairing, combined with the randomness of grain orientations in $La_{1.8}Sr_{0.2}CuO_4$ ceramics.

Finally, a new model of diamagnetic response in the glassy state of granular superconductors was developed in order to describe the anomalous (fractional-power) behavior of the shielding current response. This model, based on the idea of fractal structure of the free energy landscape in the granular network, can qualitatively account for the main features of the anomalous response. Its further development will be the subject of future studies.

We are grateful to L. B. Ioffe for many important discussions, which helped to clarify a number of issues considered in this paper. Research of M. V. F. was supported by the DGA Grant No. 94-1189, by joint Grant No. M6M300 from the International Science Foundation and the Russian Government, and by Grant No. 95-02-05720 from the Russian Fund for Fundamental Research.

APPENDIX

The hysteretic behavior of the current as a function of the induction variations is represented by the relation

$$\Delta J = \pm 2^{(1-\alpha)\nu} J_c \left(\frac{|\Delta B|}{B_c} \right)^{\alpha}, \tag{A1}$$

with $\nu = 0$ when starting from a zero-induction state, and $\nu = 1$ otherwise. $\Delta J = J - J_0$ and $\Delta B = B - B_0$, where J_0 and B_0 are the (old) values just before the last reversal of the sign of variation of B. The Ansatz (A1) ensures that we have a stable closed hysteresis loop, and that there is no hysteresis for $\alpha = 1$, which describes the London case.

The induction profile is determined by the Maxwell equation, which leads, for the case of weak penetration, to

$$\frac{d\Delta B}{dx} = \pm 2^{(1-\alpha)\nu} \mathscr{A} |\Delta B|^{\alpha}, \tag{A2}$$

where $\mathscr{A} = 4\pi J_c / B_c^{\alpha}$; x is the distance from the edge of the sample. After increasing the applied field from 0 to h_0, starting from a zero-induction state, the induction profile is given by $B^{-\alpha} dB = -\mathscr{A} dx$, leading to

$$x = -\frac{1}{\mathscr{A}} \int_{h_0}^{B} \xi^{-\alpha} \, d\xi = -\frac{B^{1-\alpha} - h_0^{1-\alpha}}{(1-\alpha).\mathscr{A}},$$

where

$$B = (h_0^{1-\alpha} - (1-\alpha).\mathscr{A} x)^{1/(1-\alpha)}. \tag{A3}$$

The field penetrates until $x = x_{h_0} = h_0^{1-\alpha}/(1-\alpha).\mathscr{A}$.

When h decreases from h_0, we obtain $(B_0 - B)^{-\alpha} dB = -2^{1-\alpha} dx$. Hence

$$x = -\frac{1}{2^{1-\alpha}.\mathscr{A}} \int_{h_0 - h}^{B_0 - B} \xi^{-\alpha} \, d\xi$$

$$= -\frac{1}{2^{1-\alpha}.\mathscr{A}(1-\alpha)} \left((B_0 - B)^{1-\alpha} - (h_0 - h)^{1-\alpha} \right).$$

Modification of induction relative to B_0 extends up to $x_h = (h_0 - h)^{1-\alpha}/2^{1-\alpha}.\mathscr{A}$. For $0 < x < x_h$,

$$B = B_0 - 2\left(\left(\frac{h_0 - h}{2} \right)^{1-\alpha} - (1-\alpha).\mathscr{A} x \right)^{1/(1-\alpha)}, \tag{A4}$$

where B_0 is given by Eq. (A3). When $h = -h_0$ is reached, Eq. (A4) gives simply $B = -B_0$. After reversing the sign of variation of h once more, the profiles are simply symmetrical to those given by Eq. (A4).

The average induction can now be derived. After some algebra, one obtains

$$\langle B \rangle = \frac{\mathscr{A}}{2-\alpha} h_0^{2-\alpha} \left[1 - 2\left(\frac{1 - h/h_0}{2} \right)^{2-\alpha} \right] \quad \text{for } h_0 > 0, \tag{A5a}$$

and

$$\langle B\rangle = \frac{\mathscr{A}}{2-\alpha}\, h_0^{2-\alpha}\left[-1+2\left(\frac{1+h/h_0}{2}\right)^{2-\alpha}\right] \quad \text{for } h_0<0. \tag{A5b}$$

For a sinusoidal excitation $h=h_0\cos\omega t$, one gets

$$\frac{\langle B\rangle}{h_0} = \frac{\mathscr{A}h_0^{1-\alpha}}{2-\alpha}\left[1-2\left(\frac{1-\cos\,\omega t}{2}\right)^{2-\alpha}\right]$$

$$\text{for } 2n\pi<\omega t<(2n+1)\pi, \tag{A6a}$$

$$\frac{\langle B\rangle}{h_0} = \frac{\mathscr{A}h_0^{1-\alpha}}{2-\alpha}\left[-1+2\left(\frac{1+\cos\,\omega t}{2}\right)^{2-\alpha}\right]$$

$$\text{for } (2n-1)\pi<\omega t<2n\pi. \tag{A6b}$$

Since $\langle B\rangle/h_0=1+\langle M\rangle/h_0$, the Fourier transform of Eqs. (A6a), (A6b) yields the values of $1+4\pi\chi'$ and $4\pi\chi''$. This can be done numerically. Figure 19 displays the ratio $4\pi\chi''/1+4\pi\chi'$ as a function of α.

*Permanent address: Institute of Solid State Physics, 142432 Chernogolovka, Moscow Region, Russia. Current address: Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands.

[1)]The susceptibility, which in the usual sense is a local quantity representing $(B-H)/H$, has no meaning in the case of circulating currents in a conductor. We speak rather of a polarizability χ_j, which represents the average value $(B-H)/H$, and which describes the global effect of the currents over the whole volume of the sample.

[1] G. Blatter, M. V. Feigel'man, V. B. Geshkenbein, A. I. Larkin, and V. M. Vinokur, Rev. Mod. Phys. **66**, 1125 (1994).
[2] S. John and T. C. Lubensky, Phys. Rev. B **34**, 4815 (1986).
[3] V. M. Vinokur, L. B. Ioffe, A. I. Larkin, and M. V. Feigel'man, Zh. Éksp. Teor. Fiz. **93**, 343 (1987) [Sov. Phys. JETP **66**, 198 (1987)].
[4] V. S. Dotsenko, M. V. Feigel'man, and L. B. Ioffe, in *Spin Glasses and Related Problems*, I. M. Khalatnikov (ed.), Soviet Scientific Review, Section A, Chapter IV [*sic*].
[5] M. V. Feigel'man and L. B. Ioffe, Phys. Rev. Lett. **74**, 3447 (1995).
[6] J. D. Reger, T. A. Tokuyasu, A. P. Young, and M. P. A. Fisher, Phys. Rev. B **44**, 7147 (1991).
[7] D. Huse and H. Seung, Phys. Rev. B **42**, 1059 (1990).
[8] H. S. Bokil and A. P. Young, Phys. Rev. Lett. **74**, 3021 (1995).
[9] M. Borik, M. Chernikov, V. Velasago, and V. Stepankin, J. Low Temp. Phys. **85**, 283 (1991).
[10] D. S. Fisher, M. P. A. Fisher, and D. A. Huse, Phys. Rev. B **43**, 130 (1991).
[11] E. B. Sonin, JETP Lett. **47**, 496 (1988); E. B. Sonin and A. K. Tagancev, Zh. Éksp. Teor. Fiz. **95**, 994 (1989) [Sov. Phys. JETP **68**, 527 (1989)].
[12] B. Z. Spivak and A. Yu. Zyuzin, JETP Lett. **47**, 267 (1988); B. Z. Spivak and S. Kivelson, Phys. Rev. B **43**, 3740 (1991).
[13] V. B. Geshkenbein, A. I. Larkin, and A. Barone, Phys. Rev. B **36**, 235 (1987).
[14] M. Sigrist and T. M. Rice, Rev. Mod. Phys. **67**, 503 (1995).
[15] J. R. Kirtley, C. C. Tsuei, J. Z. Sun, C. C. Chi, L. S. Yu-Jahnes, A. Gupta, M. Rupp, and M. B. Ketchen, Nature (London) **373**, 225 (1995).
[16] H. Kawamura, Phys. Ref. B **51**, 12398 (1995); J. Phys. Soc. Jpn. **64**, 711 (1995).
[17] I. Morgenstern, K. A. Müller, and J. G. Bednorz, Z. Phys. B **69**, 33 (1987).
[18] P. Imbert and G. Jehanno, Hyperfine Interact. **47–48**, 233 (1989).
[19] C. P. Bean, Rev. Mod. Phys. **36**, 31 (1964).
[20] A. M. Campbell, J. Phys. C **2**, 3186 (1969).
[21] M. Ocio, M. Alba, and J. Hammann, J. de Phys. Lett. 46, 1101 (1985).
[22] H. Dersh and G. Blatter, Phys. Rev. B **38**, 11391 (1988).
[23] J. A. Osborn, Phys. Rev. **67**, 351 (1945).
[24] U. Yaron, Y. Kornyushin, and I. Felner, Phys. Rev. B **46**, 14823 (1992).
[25] C. P. Poole, T. Datte, and H. A. Farach, *Copper Oxide Superconductors*, John Wiley & Sons, New York (1988).
[26] R. Renker, I. Apfelstedt, H. Küpfer, C. Politis, H. Rietschel, W. Shauer, W. Wüll, U. Gottwick, H. Kneissel, U. Rauchschwalbe, H. Spille, and F. Steglich, Z. Phys. B **67**, 1 (1987).
[27] S. Ruppel, G. Michels, H. Geus, J. Kalenborn, W. Schlabitz, B. Roden, and D. Wohllebeen, Physica C **174**, 233 (1991).
[28] D. Wohllebeen, G. Michels, and S. Ruppel, Physica C **174**, 242 (1991).
[29] C. Ebner and D. Stroud, Phys. Rev. B **31**, 165 (1985).
[30] T. Shibauchi, H. Kitano, K. Uchinokura, A. Maeda, T. Kimura, and K. Kishio, Phys. Rev. Lett. **72**, 2263 (1994).
[31] M. Suzuki, Physica C **185–189**, 2243 (1991).
[32] Y. B. Kim, C. F. Hempstead, and A. R. Strnad, Phys. Rev. Lett. **9**, 306 (1962).
[33] S. Senoussi, M. Ousséna, G. Collin, and I. A. Campbell, Phys. Rev. B **37**, 9772 (1988).
[34] J. R. Clem, Physica C **153–155**, 50 (1988).
[35] L. Leylekian, M. Ocio, and J. Hammann, Physica B **194–196**, 1865 (1994).
[36] L. Leylekian, M. Ocio, and J. Hammann, Physica C **185–189**, 2243 (1991).
[37] L. B. Ioffe, Phys. Rev. B **38**, 5181 (1988).
[38] L. Glazman, A. Koshelev, and A. Lebed', Zh. Éksp. Teor. Fiz. **94**(6), 259 (1988) [Sov. Phys. JETP **67**, 1235 (1988)].
[39] M. V. Feigel'man and L. B. Ioffe, Z. Phys. B **51**, 237 (1983).

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

Characteristic time of avalanche mixing of granular materials

S. N. Dorogovtsev*)

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted 8 May 1997)
Zh. Éksp. Teor. Fiz. **112**, 2114–2123 (December 1997)

Fractions of a granular material are mixed in a cylinder rotating slowing about its horizontally oriented longitudinal axis. The cylinder is not completely full, and at any instant, mixing occurs only on the free surface of the material. The complete dependence of the characteristic time of such so-called avalanche mixing on the fill level of the cylinder is constructed within a simple geometric approach. This dependence faithfully describes the actual experimental curve. Near the critical point of half-filling, at which mixing to a homogeneous state does not occur, the reciprocal of the characteristic mixing time is proportional to $\delta^2 \ln(|\delta|^{-1})$, where δ is the deviation of the fill level from half-filling. © *1997 American Institute of Physics.* [S1063-7761(97)01312-7]

1. INTRODUCTION

The experiment reported in Ref. 1 demonstrates the essence of so-called avalanche mixing so graphically and effectively that one of the vivid illustrations of this phenomenon appeared on the cover of the March 1995 issue of Nature. Fractions of a granular material are poured into a cylinder rotating slowly about is horizontally oriented longitudinal axis. The cylinder is not completely full, so that a free surface remains at the top. It is assumed that granules can flow only when they lie on the free surface. Everywhere else in the material they do not mix with one another, and they rotate along with the cylinder (see Fig. 1a).

As a result, as the cylinder rotates, avalanches, in which the fractions mix, occur continuously over the free surface of the material. Such mixing on a surface is called avalanche mixing.[1] The drum in which mixing is investigated in an experiment is flat, so that it is actually a disk or, as it is still called, a two-dimensional drum.[2–6] Therefore, there is no need to discuss the question of the mixing of granules along the rotation axis of the cylinder. We ask then how such mixing unfolds.

It has been found that the dynamics of avalanche mixing can in fact be understood based on a simple approach involving essentially geometric arguments.[7–9] Thus, there is no need at all to utilize the ideas of the theory of self-organized criticality[10–12] (we note, however, that the general interest in the problem of the surface flow of granules in a rotating drum[13–18] and in the surface flow of granular materials in general[19] has recently been displayed specifically in connection with problems of self-organized criticality). Unfortunately, the dependences already known from experiments were not constructed in our previous papers.[7,8] Now, however, we can obtain response functions that enable us to directly compare our theory with the available experimental data. We show here that our theory accurately describes the fastest of the possible mixing regimes. Comparing our response functions with the experiment in Ref. 1, we see that the actual experimental situation is very close to this regime.

We assume that the cylinder rotates so slowly that at any instant, the free surface is at an angle of repose to the horizontal (fluctuations of this angle with time are small,[1] and they can be disregarded).

Since the angle of repose, the radius of the cylinder, and its rotation rate do not appear in the response functions, we can henceforth depict the free surface as being horizontal, as in Fig. 1a, set the radius of the cylinder to unity, and assume that the time t is simply the rotation angle of the cylinder. We assume that granules of different fractions differ only in color. For simplicity, as in the experiments, let only two fractions, viz., white and black, mix. (The results can easily be generalized to an arbitrary number of fractions). Their initial arrangement is not important for the quantities that we seek, and we do not specify it. The granules are small, and we describe the state at any point in the material by ρ, which is the concentration of the black material at that site ($\rho=1$ at sites where all the material is black, and $\rho=0$ at sites where it is all white). Let the cylinder rotate counterclockwise.

Obviously, if the cylinder is more than half full, the central portion of the material (in Fig. 1 it is encircled by a dashed line) never mixes, and rotates along with the drum. When it is less than half full, all the material must mix.

Under what circumstances does avalanche mixing develop most rapidly? Obviously, the fastest of all possible regimes is realized if the material mixes completely, i.e., to a homogeneous state, when granules flow over the free surface from its right-hand half to its left-hand half (see Fig. 1a). To clarify, there can be an inhomogeneous distribution of the fractions on the right-hand half of the free surface, while the concentration ρ must be the same everywhere on its left-hand half. We introduce a circle with radius equal to the smallest distance from the center to the free surface (denoted by the dashed line in Fig. 1), and we discuss how the first fraction is distributed on tangents drawn to this circle. It is clear in the situation under consideration that after a long enough time (certainly after the first complete turn of the cylinder) the values of ρ along different tangents of the type denoted by CD in Fig. 1b can differ, but the concentration ρ will be the same everywhere along each individual tangent (Fig. 1b). Then, if any tangent is characterized by the angle φ between

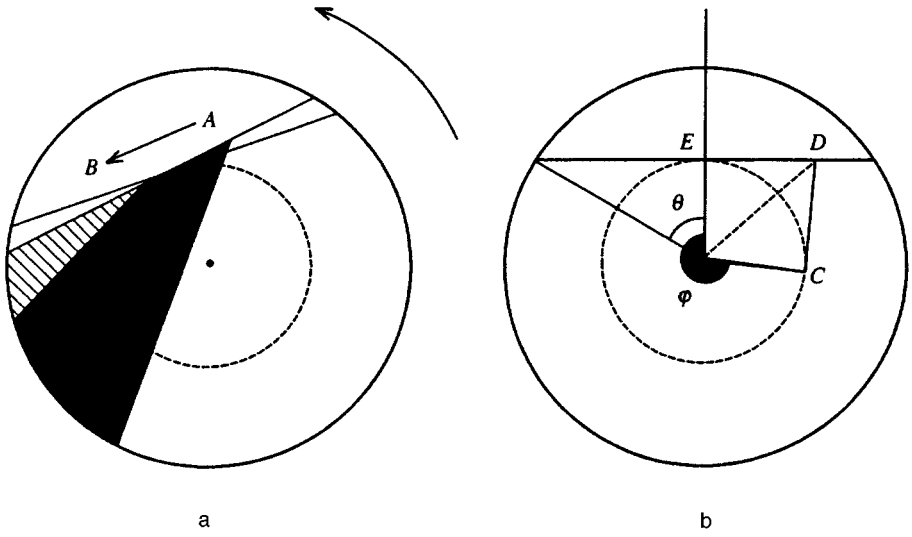


FIG. 1. a) Diagram of avalanche mixing. Granules of different fractions from sector A mix as they flow into sector B in response to an infinitesimal turn of the drum. The free surface of the granular material is at a repose angle to the horizontal. The different fractions are indicated by black and white shading. The regions with mixed material are denoted by hatching. Different degrees of mixing are not shown here or in the figures below. b) Derivation of Eq. (3). The cylinder is more than half full. Since the repose angle does not appear in the response functions, the free surface is drawn horizontal. The angle θ characterizes the relative volume of the empty space. The position of tangent CD is characterized by the angle φ. In the regime of maximally fast mixing, the concentration of the black fraction equals $\rho(\varphi,t)$ at all points on such a tangent at sufficiently large times. The tangent intersects the free surface of the material if $\zeta<2\theta$, where $\zeta\equiv 2\pi-\varphi$.

the corresponding radius vector and a normal to the free surface, the state of the system at time t can be described by the quantity $\rho(\varphi,t)$.

Thus, consideration of the fastest mixing regime, or introduction of the equivalent assumption of complete mixing in the avalanches, drastically simplifies the problem.[7] It is then possible not to explore the subtleties associated with the specific structure of the granules, their cohesion, chaining, etc.

How close is the avalanche mixing observed in the experiment in Ref. 1 to the fastest regime, which can be described analytically? The answer to this question is the main subject of the present work. The principal experimental result that enables us to make such a comparison is the dependence of the characteristic mixing time on the fill level of the rotating drum. Thus, we must construct this dependence for all fill levels.

In general, a characteristic mixing time can be introduced for the following reason. As we shall see below, at long times, the concentration of the black fraction on the left-hand part of the free surface $\rho(\varphi=0,t)$ depends on the time in the following manner (a similar dependence, of course, can be written for other points in the material, and the same dependence also follows from the experiment in Ref. 1):

$$\rho(0,t)=\rho_\infty+C_0\,\exp\{-t/\tau\}\cos\{2\pi t/T+\varphi_0\}. \qquad (1)$$

Here ρ_∞ is the concentration of the black fraction in the material mixed to a homogeneous state, which, like the constants C_0 and φ_0, depends on the initial conditions, i.e., on the size and arrangement of the pure fractions at $t=0$; τ is the desired characteristic mixing time; and T is the oscillation period. We also need to construct plots of τ and T as functions of the fill level.

In the next section we show how equations describing the dynamics of mixing can be obtained in a very simple manner, if the assumption made above is used. The problem actually reduces to an investigation of the dynamics of a zero-dimensional linear system.

In Sec. 3 we use these equations to construct curves that describe the dependence of the parameters τ and T appearing in (1), i.e., the characteristic mixing time and the oscillation period, on the fill level of the cylinder. We see that the experimental situation in Ref. 1 is very close to the fastest mixing regime, and that our theory faithfully describes the experiment.

As it turns out, if the cylinder is exactly half full, the granular material never mixes to a homogeneous state. In Sec. 4 we describe the highly nontrivial, singular behavior of the characteristic mixing time in the vicinity of this special point.

Thus, we see that the dependence of the mixing time on the fill level of the cylinder can be described surprisingly well within an approach based on the assumption that the granules mix completely when they flow over the free surface of the material. In the concluding portion of this paper, we briefly discuss how the mixing process might be described when this very strong assumption is discarded, i.e., if the granules of different fractions are not distributed perfectly homogeneously after they flow over the free surface. It is found that the problem reduces to a description of the dynamics of a one-dimensional linear system. Another char-

acteristic time, which is related to the physical characteristics of the granules, appears in this situation.

2. EQUATIONS DESCRIBING THE DYNAMICS OF AVALANCHE MIXING

Let us describe the kinetics of avalanche mixing in the case of complete mixing of the granules when they flow over the free surface of the material, i.e., the fastest mixing regime. We have already obtained the corresponding equations in Ref. 7. Now, however, we have learned to derive them so simply that we can allow ourselves to present this brief derivation here.

We turn to the case of a more than half full drum. It is convenient to characterize the fraction of empty space by the opening angle θ (see Fig. 1b). We wish to find an asymptote like (1). Therefore, we focus only on fairly long times, at which, as was shown in the Introduction, the concentration of black granules is everywhere identical along each individual tangent like CD. (In Refs. 7 and 8 we also show how avalanche mixing develops at lesser times.) These tangents are characterized by the angle φ, and the concentration of the black fraction at points along the corresponding tangent at time t is $\rho(\varphi,t)$.

We turn the cylinder through the small angle dt at time t. Then the quantity of black material flowing from right to left (from sector A to sector B in Fig. 1) is:

$$\begin{aligned} dt\ \frac{\sin^2\theta}{2}\rho(0,t) &= dt\int_0^{\sin\theta} dr\ r\rho_{\text{right}}(r,t) \\ &= dt\int_0^{2\theta} d\zeta\ \frac{dr(\zeta)}{d\zeta}r(\zeta)\rho_{\text{right}}(r(\zeta),t) \\ &= dt\ \frac{\cos^2\theta}{2}\int_0^{2\theta} d\zeta\ \frac{d\tan^2(\zeta/2)}{d\zeta} \\ &\quad\times\rho(2\pi-\zeta,t). \end{aligned} \tag{2}$$

Here the first integration $\int_0^{\sin\theta} dr$ is carried out from the middle of the free surface to the wall of the cylinder; $r(\zeta)$ is the coordinate of the point of intersection of the tangent characterized by the angle $\varphi=2\pi-\zeta$ with the free surface, i.e., segment ED in Fig. 1b. In addition, $r(\zeta)=\cos\theta\tan(\zeta/2)$, so that $r(2\theta)=\sin\theta$. Thus, the concentration of the black fraction at different points on the right-hand half of the free surface is $\rho_{\text{right}}(r(\zeta),t)=\rho(\zeta,t)$. We now take into account that, as follows from the absence of mixing of the granules at points not belonging to the free surface, $\rho(\varphi,t)=\rho(0,t-\varphi)$ for $t\geqslant\varphi$ and $\varphi\geqslant 0$, and from (2) we quickly obtain the desired equation for $\rho(0,t)$:

$$\rho(0,t)=\cot^2\theta\int_0^{2\theta} d\zeta\ \rho(0,t-2\pi+\zeta)\frac{\sin(\zeta/2)}{\cos^3(\zeta/2)}. \tag{3}$$

Proceeding in the same manner, we can also easily obtain the corresponding equation for a less than half full cylinder (here it is convenient to introduce the angle $\vartheta\equiv\pi-\theta$):

$$\rho(0,t)=\cot^2\vartheta\int_0^{2\vartheta} d\zeta\ \rho(0,t-\zeta)\frac{\sin(\zeta/2)}{\cos^3(\zeta/2)}. \tag{4}$$

As a result, a complex dynamical problem has been reduced to an investigation of the dynamics of a zero-dimensional linear system. We note that Eqs. (3) and (4) do not transform into one another following the formal replacement of θ by $\pi-\vartheta$. The fact is that the half-full point $\theta=\vartheta=\pi/2$ is, as we shall see below, a point where analyticity is lost. At half-filling, it follows at once from Eq. (3), as well as from (4), that $\rho(0,t)=\rho(0,t-\pi)$, and that mixing to a homogeneous state never occurs in the system. Thus, the mixing time goes to infinity at that point.

To find the characteristic mixing time and the oscillation period of the asymptotes of the solutions of Eqs. (3) and (4) (see Eq. (1)), it is sufficient to substitute $\rho(0,t)=e^{zt}$ into these equations. As a result, at more than half-filling, from (3) we obtain

$$e^{2\pi z}=2\cot^2\theta\int_0^{\theta} d\xi\ e^{2z\xi}\frac{\sin\xi}{\cos^3\xi}. \tag{5}$$

At less than half-filling, Eq. (4) gives

$$1=2\cot^2\vartheta\int_0^{\vartheta} d\xi\ e^{-2z\xi}\frac{\sin\xi}{\cos^3\xi}. \tag{6}$$

Thus, we obtain two transcendental equations for z. To determine the characteristic mixing time, we must find the roots closest to zero. As can be seen from the asymptote (1), each of these two roots is associated with a characteristic mixing time and an oscillation period:

$$z=-\tau^{-1}\pm i2\pi/T. \tag{7}$$

We note at once that τ and T depend on only one parameter, i.e., on the fill level of the drum! Now we need only find the roots of Eqs. (5) and (6).

3. DEPENDENCE OF THE MIXING TIME ON THE FILL LEVEL OF THE CYLINDER

The integrals on the right-hand sides of Eqs. (5) and (6) can be expressed in terms of special functions. As usual, although the solution of a transcendental equation cannot be represented in analytic form, it can be found to as high a level of accuracy as desired. Since these special functions are quite exotic, it is simpler to calculate the integrals in (5) and (6) directly. It turns out to be most convenient to perform a single integration by parts and then iterate the following relations:

$$z=\frac{1}{2\pi}\ln\left\{e^{2\theta z}-2z\cot^2\theta\int_0^{\theta} d\xi e^{2z\xi}\tan^2\xi\right\} \tag{8}$$

for a more than half full cylinder ($\theta<\pi/2$), and

$$z=\frac{1}{2\vartheta}\ln\left\{1+2z\cot^2\vartheta e^{2z\vartheta}\int_0^{\vartheta} d\xi e^{-2z\xi}\tan^2\xi\right\} \tag{9}$$

for less than half-filling ($\vartheta<\pi/2$).

These iterations converge very rapidly. As a result, using (7), we obtain the desired dependence of the reciprocal of the mixing time τ^{-1} and the oscillation period T on the fill level of the drum (see Figs. 2 and 3), i.e., our main results. Now it is more instructive to express τ^{-1} and T in terms of $v=(\vartheta-\sin 2\vartheta/2)/\pi$, i.e., the relative fraction of the volume

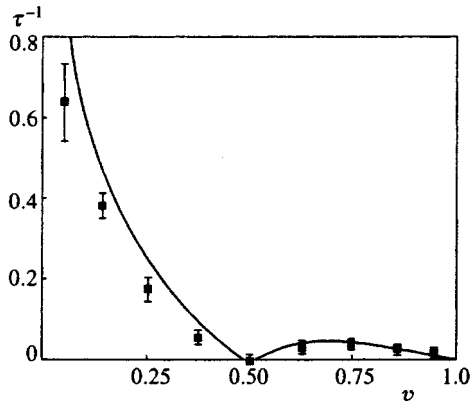


FIG. 2. Dependence of the reciprocal of the characteristic mixing time on the fraction of the volume of the cylinder occupied by granular material. Experimental points from Ref. 1 are shown for comparison.

of the drum occupied by granules, rather than the angles θ and ϑ. For comparison, Fig. 2 also presents the experimental points from Ref. 1 with their errors.

Each value of $\tau(v)$ found corresponds to the shortest of all possible times for avalanche mixing at the respective value of v. Thus, plots of $\tau^{-1}(v)$ for certain specific granular materials should not extend above the limiting curve constructed in Fig. 2. In fact, although the experimental points shown in Fig. 2 are close to the limiting dependence, they are still below it. Thus, the actual experiment[1] is very close to the limiting regime of fastest mixing, and is therefore described excellently by our theory, which does not contain any parameters of the granular material being mixed.

Following Ref. 1, we can also introduce another parameter, viz., the mixing ''rate'' v/τ. In our theory its maximum $v/\tau=0.0683\ldots$ is achieved at $v=0.177\ldots$, i.e., the material mixes most rapidly at that fill level of the drum. These values are also very close to the experimental values: $v/\tau=0.056\pm0.006$ when $v=0.17\pm0.015$.[1]

In Refs. 7 and 8 we found analytic expressions for τ^{-1} and T in the limiting cases of low and high fill levels: at a low drum fill level, $(\vartheta\ll1)\,\tau^{-1}\approx1.392/2\vartheta$ and $T\approx4\pi\vartheta/7.553$, and at a high fill level, $(\theta\ll1)\,\tau^{-1}\approx(1/9)\times(1/2\pi)\theta^2$ and $T\approx2\pi-4\theta/3$. These analytic dependences

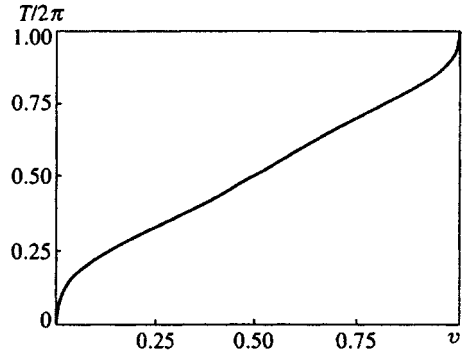


FIG. 3. Dependence of the oscillation period of the concentration of either of the fractions (see Eq. (1)) at long times on the fraction of the volume of the cylinder occupied by granular material.

describe the corresponding portions of the curves in Figs. 2 and 3. We now show how τ and T behave near the special point, i.e, the half-full point.

4. CRITICAL BEHAVIOR NEAR THE HALF-FULL POINT

As was noted in Sec. 2, when the drum is half full, it follows from Eq. (3), as well as from (4), that the concentration of the black fraction $\rho(0,t)$ is a periodic function with period π (half of the rotation period of the drum), and complete mixing is never achieved. Therefore, taking Eq. (7) into account, we can represent the desired root near the half full point in the form $z=2i+s$, where s is a small correction (of course, the conjugate root $z=-2i+s$ could be found with equal success).

Consider, for example, a more than half full cylinder. After substituting the latter relation into Eq. (3), we have

$$e^{2\pi s}=2\cot^2\theta\int_0^\theta d\xi\, e^{2s\xi}\{4\sin 2\xi+\tan^3\xi-7\tan\xi$$
$$-4i(\cos 2\xi+\tan^2\xi-1)\}. \tag{10}$$

We now introduce a small deviation δ from half-filling: $\theta=\pi/2-\delta$, whereupon we can expand $e^{-2s\xi}$, taking the smallness of s into account. It is now simplest to proceed in the following manner: integrating by parts, we can easily isolate the terms that are singular when $\delta\to0$. The remaining integrals, which converge as $\delta\to0$, can also be evaluated without difficulty. After this we need only compare the coefficients of terms with equal powers of δ.

Here we omit these routine but extremely cumbersome calculations, and merely present the expression for the root closest to zero:

$$z=2i-\frac{8\delta}{\pi}i-\frac{16\delta^2}{\pi}\left[\ln\frac{1}{|\delta|}-5/2\right]+\ldots \tag{11}$$

(We have not written out a term $O(\delta^2)$ in the imaginary part of z.) It follows directly from (11) with consideration of (7) that at sufficiently long times ($t\gg\tau$), the mixing process develops according to (1), in which the characteristic mixing time τ and the oscillation period T can be expressed in terms of the deviation of θ from $\pi/2$ in the following manner:

$$\tau^{-1}=\frac{16\delta^2}{\pi}\left[\ln\frac{1}{|\delta|}-\frac{5}{2}\right] \tag{12}$$

and

$$T=\pi+4\delta. \tag{13}$$

Of course, τ and T can also be expressed in terms of the deviation of the relative fraction of the volume filled by the granular material from one-half, noting that $v-1/2\approx2\delta/\pi$.

Similar calculations can be performed with equal success in the case of less than half-filling ($\vartheta<\pi/2$). The final expressions have precisely the same form as (12) and (13) (but here $\delta=\vartheta-\pi/2<0$). Of course, such agreement is observed only in low orders. We note that in the theory of ordinary critical phenomena, the coefficients of the singularities on the two sides of a critical point differ (see, for example, Refs. 20 and 21).

A comparison of the analytic expression (12) with the complete dependence of $\tau(v)$ (see Fig. 2) calculated in the preceding section shows that (12) is suitable only in the very narrow range $|v-1/2|<0.02$ near half-filling.

5. CONCLUSIONS

We have found that in an actual experiment[1] (using colored crystal salt granules), mixing occurs in a regime that is very close to the fastest of all possible regimes for avalanche mixing. (In the fastest mixing regime, granules should mix completely to a homogeneous state as they slide over the free surface of the material.) In this case the description of avalanche mixing becomes extremely simple, i.e., the problem reduces to an investigation of the dynamics of a zero-dimensional linear system, and can, in fact, be solved exactly. It has been found that the simple system studied here has unusual critical behavior near the half-full point of the cylinder, at which the material never mixes to a homogeneous state.

Remarkably, the main characteristic of the problem, viz., the characteristic mixing time τ, which we specifically sought, was found to be extremely insensitive to deviations from a homogeneous distribution of the granules after they flowed over the free surface. (Note that we also neglected fluctuations of the repose angle with time. Such fluctuations can be significant in an experiment at a low fill level of the cylinder.) These deviations are, of course, visible experimentally in the actual patterns of the distribution of the fractions at individual moments in time, and their extent is determined by which specific granular material is being mixed (see Refs. 16, 22, and 23). Consideration of these deviations reduces the problem to the investigation of the dynamics of a one-dimensional linear system.

In fact, since the granules of a fraction can be irregularly distributed even after they have flowed over the left-hand half of the free surface in this case, the quantity $\rho(r,\ \varphi=0,\ t)$, where r is the coordinate of a point on the left-hand half of the free surface measured from its center, will appear in the equation describing the mixing dynamics. The specific form of the map expressing $\rho(r,\ \varphi=0,\ t)$ in terms of the distribution of the black fraction at that moment in time on the right-hand half of the free surface is determined by the characteristics of the granular material. Nevertheless, it is not difficult to see that the map must be linear, if the granules of different fractions differ only in color; if the material is homogeneously mixed on the right-hand half of the free surface, a homogeneous distribution should also be obtained after the material flows over the left-hand half; the quantity of the black fraction before flow in sector A in Fig. 1a should be conserved after flow into sector B. These arguments are sufficient, in principle, to predict the form of such a map. It can be selected, for example, on the basis of semiphenomenological arguments after describing the dependence on the nature of the granular material by a single parameter.

We shall not present the final equations for $\rho(r,\ \varphi=0,\ t)$ here, since they are far from as elegant as Eqs. (3) and (4). Of course, Eqs. (3) and (4) can be obtained from these general equations as the simplest special case. An analytic treatment of the general case can be performed without difficulty for half-filling. It can be seen that one more characteristic time, which is associated with the characteristics of the granules, i.e., with the specific form of the map introduced, then appears in the problem. Still, such a problem can be solved analytically only in the simplest situations. For just this reason, a faithful description of the real experiment in Ref. 1 using a very simple, essentially geometric, theory, which does not contain any characteristics of the granules, is quite impressive.

We thank E. N. Antonov, S. A. Ktitorov, E. K. Kudinov, A. M. Monakhov, A. N. Samukhin, B. N. Shalaev, and Yu. A. Firsov for numerous useful discussions. Some of the results presented were obtained when the author participated in the *Problems in the Physics of Interfaces* workshop in Turin, which was financed by INTAS. This work was partially supported by the Russian Fund for Fundamental Research.

*)E-mail: sn@dor.ioffe.rssi.ru

[1] G. Metcalfe, T. Shinbrot, J. J. McCarthy, and J. M. Ottino, Nature (London) **374**, 39 (1995).
[2] V. Buchholtz, Th. Pöschel, and H.-J. Tillemans, Physica A **216**, 199 (1995).
[3] Th. Pöschel and V. Buchholtz, Phys. Rev. Lett. **71**, 3963 (1993).
[4] F. Cantelaube, Y. Limon-Duparemeur, D. Bideau, and G. H. Ristow, J. Phys. I **5**, 581 (1995).
[5] G. A. Kohring, J. Phys. I **5**, 1551 (1995).
[6] G. H. Ristow, Europhys. Lett. **28**, 97 (1994).
[7] S. N. Dorogovtsev, JETP Lett. **62**, 262 (1995).
[8] S. N. Dorogovtsev, Zh. Éksp. Teor. Fiz. **112**, 257 (1997) [JETP **85**, 141 (1997)].
[9] S. N. Dorogovtsev, submitted to Europhys. Lett.
[10] P. Bak, C. Tang, and K. Wiesenfeld, Phys. Rev. Lett. **59**, 381 (1987).
[11] P. Bak, C. Tang, and K. Wiesenfeld, Phys. Rev. A **38**, 364 (1988).
[12] C. Tang and P. Bak, Phys. Rev. Lett. **60**, 2347 (1988).
[13] J. Rajchenbach, Phys. Rev. Lett. **65**, 2221 (1990).
[14] G. Baumann, I. M. Jánosi, and D. E. Wolf, Europhys. Lett. **27**, 203 (1994).
[15] G. Baumann, I. M. Jánosi, and D. E. Wolf, Phys. Rev. E **51**, 1879 (1995).
[16] G. Baumann and D. E. Wolf, Phys. Rev. E **54**, R4504 (1996).
[17] O. Zik, D. Levine, S. G. Lipson, S. Shrikman, and J. Stavans, Phys. Rev. Lett. **73**, 644 (1994).
[18] K. M. Hill and J. Kakalios, Phys. Rev. E **49**, R3610 (1994).
[19] H. M. Jaeger and S. R. Nagel, Science **255**, 1523 (1992).
[20] D. J. Amit, *Field Theory, Renormalization Group and Critical Phenomena*, World Scientific, Singapore (1984).
[21] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*, Clarendon, New York (1996).
[22] V. Frette, K. Christensen, A. Malthe-Sørenssen, J. Feder, T. Jøssang, and P. Meakin, Nature (London) **379**, 49 (1996).
[23] K. Christensen, A. Corral, V. Frette, J. Feder, and T. Jøssang, Phys. Rev. Lett. **77**, 107 (1996).

Translated by P. Shelnitz

The Ginzburg–Landau expansion and the slope of the upper critical field in superconductors with anisotropic normal-impurity scattering

A. I. Posazhennikova and M. V. Sadovskiĭ*)

Institute of Electrophysics, Ural Branch of the Russian Academy of Sciences, 620049 Ekaterinburg, Russia
(Submitted 16 May 1997)
Zh. Éksp. Teor. Fiz. **112**, 2124–2133 (December 1997)

We carry out the Ginzburg–Landau expansion for superconductors with anisotropic *s* and *d* pairing in the presence of anisotropic normal-impurity scattering, which enhances the stability of *d* pairing with respect to disordering. We find that the slope of the curve of the upper critical field, $|dH_{c2}/dT|_{T_c}$, in superconductors with *d* pairing behaves nonlinearly as disorder grows: at low scattering anisotropy the slope rapidly decreases with increasing impurity concentration, then gradually but nonlinearly increases with concentration, reaches its maximum, and then rapidly decreases, vanishing at the critical impurity concentration. In superconductors with anisotropic *s* pairing, $|dH_{c2}/dT|_{T_c}$ always increases with impurity concentration, finally reaching the familiar asymptotic value characteristic of the isotropic case, irrespective of whether there is anisotropic impurity scattering. © *1997 American Institute of Physics.* [S1063-7761(97)01412-1]

1. INTRODUCTION

The problem of determining the type of Cooper pairing is still occupying center stage in high-T_c superconductivity studies. Most experiments and a number of theoretical models[1] indicate that the majority of high-T_c oxides have $d_{x^2-y^2}$-anisotropic pairing with the zeros of the gap function lying on the Fermi surface. Other variants of anisotropic pairing have also been proposed, e.g., the so-called anisotropic *s* pairing,[2,3] which also gives rise to zeros in the gap function (but without a change in sign in the order parameter) or to minima in the gap function at the Fermi surface in the same directions in the Brillouin zone as in the case of *d* pairing (here, too, there are experimental indications that verify this fact). Borkovski and Hirschfeld[4] and Fehrenbacher and Norman[5] pointed out that controlled injection of normal impurities (disordering) can serve as an effective method of experimentally distinguishing the above types of anisotropic pairing, since it would lead to a markedly different behavior of the density of states in these types of superconductor. In our previous paper (see Ref. 6) we found that measuring the evolution of the slope of the curve of the upper critical field, $|dH_{c2}/dT|_{T_c}$, as the degree of disorder changes, at least in principle, serve the same purpose: in superconductors with *d* pairing the magnitude of this slope rapidly decreases with increasing disorder, while in the case of anisotropic *s* pairing the slope of the field increases with disorder, which is similar to the behavior in the isotropic case.

Recently, Harań and Nagi[7] examined an interesting model with anisotropic impurity scattering. They found that when the *d*-type scattering anisotropy is strong, the breaking of *d*-type Cooper pairs decreases substantially because of normal-impurity scattering, which in the isotropic case is described by the well-known Abrikosov–Gor'kov formula for magnetic impurities in a isotropic superconductor.[4–6] Thus, by allowing for anisotropic impurity scattering we can, at least in principle, resolve one of the main problems of high-T_c superconductor physics: the contradiction between the clear indications that *d* pairing exists in high-T_c superconductors and the relative stability of such superconductors with respect to disordering.[8] This explanation of the remarkable stability of high-T_c superconducting cuprates with respect to disordering, if there is indeed *d* pairing in such cuprates, is not the only one (see, e.g., the explanation given in Ref. 9), but the simplicity of the model of Ref. 7 is appealing and stimulates calculations of other characteristics of superconductors with ''exotic'' types of pairing, with allowance for the possible role of anisotropic normal-impurity scattering. The present paper is a direct generalization of Ref. 6 to this case. It will be shown that allowance for anisotropic impurity scattering leads (in the case of *d* pairing) to striking anomalies in the behavior of the slope of the curve of the upper critical field as a function of the degree of disorder (impurity concentration). As in Ref. 6, we base our reasoning on a microscopic derivation of the Ginzburg–Landau expansion in the impurity system.

2. THE GINZBURG–LANDAU EXPANSION

Let us consider a two-dimensional electron system with an isotropic Fermi surface and a separable Cooper-pairing potential of the form[4,5]

$$V(\mathbf{p},\mathbf{p}')\equiv V(\phi,\phi')=-Ve(\phi)e(\phi'), \tag{1}$$

where ϕ is the polar angle determining the direction of the electron momentum $\mathbf{p}$ in the highly conducting plane, and for $e(\phi)$ we adopt the simplest model:

$$e(\phi)=\begin{cases}\sqrt{2}\,\cos(2\phi) & (d \text{ pairing}),\\ \sqrt{2}|\cos(2\phi)| & (\text{anisotropic } s \text{ pairing}).\end{cases} \tag{2}$$

We assume, as usual, that the attraction constant V is finite in a layer of thickness $2\omega_c$ in the vicinity of the Fermi level (ω_c is the characteristic frequency of the photons that ensure the attraction of the electrons). In this case the superconducting gap (the order parameter) has the form

$$\Delta(\mathbf{p})\equiv\Delta(\phi)=\Delta e(\phi), \tag{3}$$

 1063-7761/97/121162-06$10.00

where the positions of the zeros of the gap function at the Fermi surface for the s and d cases coincide.

We take a superconductor containing ''normal'' (nonmagnetic) impurities that are distributed at random in space with a concentration ρ. Following Harań and Nagi,[7] we assume that the square of the impurity scattering amplitude can be written as

$$|V_{\rm imp}(\mathbf{p},\mathbf{p}')|^2 \equiv |V_{\rm imp}(\phi,\phi')|^2 = |V_0|^2 + |V_1|^2 f(\phi) f(\phi'), \tag{4}$$

where V_0 is the amplitude of isotropic point scattering, V_1 is the amplitude of anisotropic scattering, and the model function $f(\phi)$ (depending on the same polar angle that defines the direction of the electron momentum) determines the nature of the anisotropic impurity scattering. We assume that the scattering is ''essentially'' isotropic and introduce the following constraints:[7]

$$|V_1|^2 \leqslant |V_0|^2, \quad \langle f\rangle = 0, \quad \langle f^2\rangle = 1, \tag{5}$$

where the angle brackets stand for averaging over the directions of momentum at the Fermi surface (the angle ϕ). Accordingly, the second term in (4) describes the deviations from isotropic scattering.

The normal and anomalous Green's functions in such a superconductor are[10]

$$G(\omega,\mathbf{p}) = -\frac{i\tilde{\omega} + \xi_{\mathbf{p}}}{\tilde{\omega}^2 + \xi_{\mathbf{p}}^2 + |\tilde{\Delta}(\mathbf{p})|^2},$$

$$F(\omega,\mathbf{p}) = \frac{\tilde{\Delta}^*(\mathbf{p})}{\tilde{\omega}^2 + \xi_{\mathbf{p}}^2 + |\tilde{\Delta}(\mathbf{p})|^2}, \tag{6}$$

where $\omega = (2n+1)\pi T$,

$$\tilde{\omega}(\mathbf{p}) = \omega + i\rho \int \frac{d\mathbf{p}'}{(2\pi)^2} |V_{\rm imp}(\mathbf{p}-\mathbf{p}')|^2 G(\omega,\mathbf{p}'),$$

$$\tilde{\Delta}(\mathbf{p}) = \Delta(\mathbf{p}) + \rho \int \frac{d\mathbf{p}'}{(2\pi)^2} |V_{\rm imp}(\mathbf{p}-\mathbf{p}')|^2 F(\omega,\mathbf{p}'), \tag{7}$$

and ξ is the electron energy measured from the Fermi level.

To find the transition, or critical, temperature T_c we can restrict ourselves to the linear approximation in Δ in Eqs. (7):

$$\tilde{\omega} = \omega + i\rho \frac{N(0)}{2\pi} \times \int d\xi \int_0^{2\pi} d\phi \{|V_0|^2 + |V_1|^2 f(\phi) f(\phi')\} \frac{\tilde{\omega}}{\tilde{\omega}^2 + \xi^2},$$

$$\tilde{\Delta} = \Delta + \rho \frac{N(0)}{2\pi} \times \int d\xi \int_0^{2\pi} d\phi \{|V_0|^2 + |V_1|^2 f(\phi) f(\phi')\} \frac{\tilde{\Delta}}{\tilde{\omega}^2 + \xi^2}. \tag{8}$$

The linearized equation for the gap function, which determines the transition temperature T_c, is

$$\Delta(\mathbf{p}) = -T_c \sum_{\omega} \int \frac{d\mathbf{p}'}{(2\pi)^2} V(\mathbf{p},\mathbf{p}') \frac{\tilde{\Delta}(\mathbf{p}')}{\tilde{\omega}^2 + \xi_{\mathbf{p}'}^2}. \tag{9}$$

Applying standard methods to Eq. (9) and the renormalization equations (8), we arrive at an equation for the transition temperature T_c in general form:

$$\ln \frac{T_c}{T_{c0}} = (\langle e\rangle^2 + \langle ef\rangle^2 - 1)\left[\Psi\left(\frac{1}{2} + \frac{\gamma_0}{2\pi T_c}\right) - \Psi\left(\frac{1}{2}\right)\right] + \langle ef\rangle^2 \left[\Psi\left(\frac{1}{2}\right) - \Psi\left(\frac{1}{2} + \frac{\gamma_0}{2\pi T_c}\left(1 - \frac{\gamma_1}{\gamma_0}\right)\right)\right], \tag{10}$$

where T_{c0} is the transition temperature in the absence of impurities, $\Psi(x)$ is the logarithmic derivative of the gamma function, $\gamma_0 = \pi\rho V_0^2 N(0)$ and $\gamma_1 = \pi\rho V_1^2 N(0)$ are, respectively, the isotropic and anisotropic scattering frequencies, and $\langle ef\rangle^2$ determines the overlap of the functions $e(\mathbf{p})$ and $f(\mathbf{p})$.

For simplicity we select the function $f(\mathbf{p})$ in a form similar to (2):

$$f(\mathbf{p}) \equiv f(\phi) = \sqrt{2}\cos(2\phi), \tag{11}$$

which ensures a maximum overlap in the d case. A more general approach can be found in Ref. 7. We can now write the renormalization equations (8) as follows:

$$\tilde{\omega} = \omega + i\frac{\gamma_0}{\pi} \int d\xi \frac{\tilde{\omega}}{\tilde{\omega}^2 + \xi^2} + i\frac{\gamma_1}{\pi^2}\cos(2\phi) \int d\xi \int d\phi' \cos(2\phi') \frac{\tilde{\omega}}{\tilde{\omega}^2 + \xi^2},$$

$$\tilde{\Delta} = \Delta + i\frac{\gamma_0}{\pi} \int d\xi \frac{\tilde{\Delta}}{\tilde{\omega}^2 + \xi^2} + i\frac{\gamma_1}{\pi^2}\cos(2\phi) \int d\xi \int d\phi' \cos(2\phi') \frac{\tilde{\Delta}}{\tilde{\omega}^2 + \xi^2}. \tag{12}$$

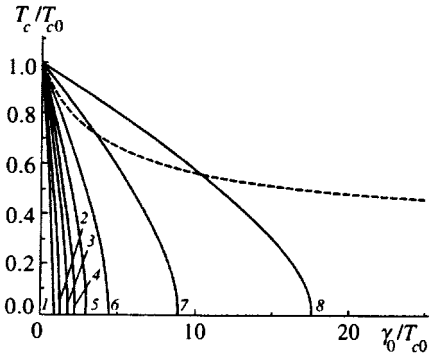


FIG. 1. T_c as a function of the disorder parameter γ_0/T_{c0}. The dashed curve corresponds to the case of s pairing and the solid curves, to the case of anisotropic d pairing for the following values of γ_1/γ_0: curve *1*, 0.0; curve *2*, 0.3; curve *3*, 0.5; curve *4*, 0.6; curve *5*, 0.7; curve *6*, 0.8; curve *7*, 0.9; and curve *8*, 0.95.

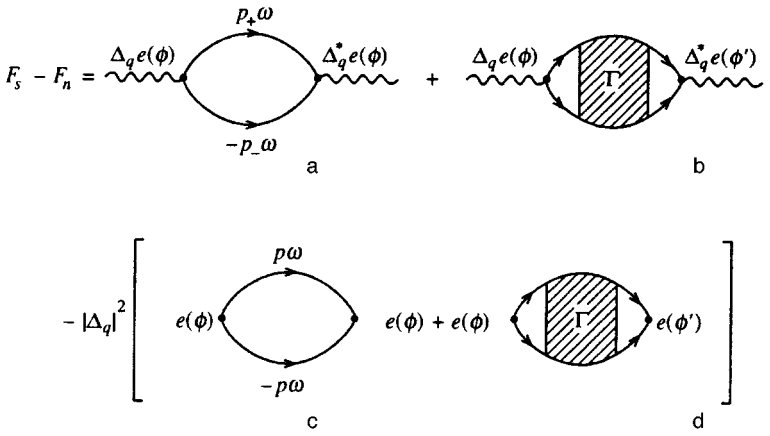


FIG. 2. The diagrammatic representation of the Ginzburg–Landau expansion. The electron lines are ''dressed'' by impurity scattering; Γ is the vertex part of the impurity scattering calculated in the ladder approximation. The diagrams (c) and (d) are calculated for $q=0$ and $T=T_c$, and $p_\pm = p \pm q/2$.

This yields the well-known expression for the renormalized frequency in both cases of interest:

$$\tilde{\omega} = \omega + \gamma_0 \ \mathrm{sgn}\ \omega. \tag{13}$$

In the case of d pairing the symmetry of the gap function in the presence of impurities does not change:

$$\tilde{\Delta} = \Delta \frac{|\tilde{\omega}|}{|\tilde{\omega}| - \gamma_1}. \tag{14}$$

When there is s pairing, the gap is renormalized by a constant independent of the angle ϕ and the frequency γ_1:

$$\tilde{\Delta} = \Delta + \Delta_0 \frac{2\sqrt{2}\gamma_0}{\pi|\omega|}. \tag{15}$$

As a result, the equation for T_c in a superconducting with d pairing becomes

$$\ln \frac{T_c}{T_{c0}} = \Psi\left(\frac{1}{2}\right) - \Psi\left(\frac{1}{2} + \left(1 - \frac{\gamma_1}{\gamma_0}\right)\frac{\gamma_0}{2\pi T_c}\right). \tag{16}$$

For a superconductor with anisotropic s pairing we have

$$\ln \frac{T_c}{T_{c0}} = \left(1 - \frac{8}{\pi^2}\right)\left[\Psi\left(\frac{1}{2} + \frac{\gamma_0}{2\pi T_c}\right) - \Psi\left(\frac{1}{2}\right)\right]. \tag{17}$$

Note that in Eq. (17) there is no dependence on the anisotropic scattering rate.

Figure 1 plots T_c vs. γ_0/T_{c0} for the case of d pairing for different values of γ_1/γ_0. For an s-type superconductor the transition temperature T_c becomes weakly suppressed as γ_0/T_{c0} increases. For a d-type superconductor the transition temperature T_c at small values of γ_1 becomes suppressed very rapidly, but as γ_1/γ_0 increases, the critical value γ_{0c}/T_{c0} at which superconductivity disappears rapidly increases.

As usual, for the order parameter in which the free energy is expanded we take the gap function. Here we assume that the amplitude $\Delta(T)$ is a slowly varying function of position. In momentum space we have the following Fourier transfer of the order parameter:

$$\Delta(\phi, q) = \Delta_q(T) e(\phi). \tag{18}$$

The Ginzburg–Landau expansion for the difference of free energies of the superconducting and normal states has the following form (accurate to within terms quadratic in Δ in the region of small values of q):

$$F_s - F_n = A|\Delta_q|^2 + q^2 C|\Delta_q|^2; \tag{19}$$

it is determined by the diagrams (see Fig. 2) of the loop expansion for the electron free energy in the order-parameter

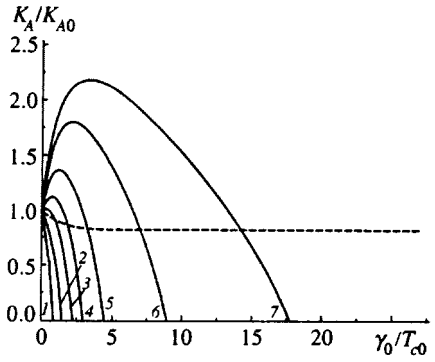


FIG. 3. The dimensionless coefficient K_A/K_{A0} as a function of the disorder parameter γ_0/T_{c0}. The dashed curve corresponds to the case of s pairing and the solid curves to the case of anisotropic d pairing, for the following values of γ_1/γ_0: curve *1*, 0.0; curve *2*, 0.4; curve *3*, 0.6; curve *4*, 0.7; curve *5*, 0.8; curve *6*, 0.9; and curve *7*, 0.95.

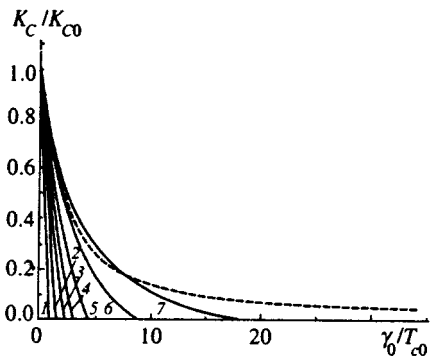


FIG. 4. The dimensionless coefficient K_C/K_{C0} as a function of the disorder parameter γ_0/T_{c0}. The dashed curve corresponds to the case of s pairing and the solid curves to the case of anisotropic d pairing, for the following values of γ_1/γ_0: curve *1*, 0.0; curve *2*, 0.4; curve *3*, 0.6; curve *4*, 0.7; curve *5*, 0.8; curve *6*, 0.9; and curve *7*, 0.95.

fluctuation field with a small wave vector **q**. Subtraction of the diagrams (c) and (d) guarantees that the coefficient A vanishes at the transition point $T=T_c$. In Appendices A and B we give the details of calculations of, respectively, the vertex part $\Gamma_{\mathbf{pp}'}$ and the Ginzburg–Landau coefficients for a d-type superconductor. Note that for d-type superconductors the ''diffusion'' renormalization due to the diagrams of type (b) and (d) is zero to within terms quadratic in q, provided that the anisotropy of impurity scattering is ignored. For an s-type superconductor the calculations are similar, but here there is no dependence on the anisotropic component of the scattering.

As a result, the Ginzburg–Landau coefficients can be written as

$$A=A_0K_A, \quad C=C_0K_C, \tag{20}$$

where A_0 and C_0 are the usual expressions for isotropic s pairing:[11]

$$A_0=N(0)\frac{T-T_c}{T_c}, \quad C_0=N(0)\frac{7\zeta(3)}{48\pi^2}\frac{v_F^2}{T_c^2}; \tag{21}$$

here v_F and $N(0)$ are the electron velocity and the density of states at the Fermi surface. All properties of the models are contained in the dimensionless coefficients K_A and K_C. In the absence of impurities we have $K_A^0=1$ and $K_C^0=3/2$ in both models. For a system with impurities we have the following.

(A) d pairing:

$$K_A=\frac{\gamma_0}{4\pi T_c}\int_{-\omega_c}^{\omega_c}\frac{d\xi}{\xi}\times\int_{-\infty}^{\infty}d\omega\,\frac{\omega+\xi}{(\omega^2+\gamma_0^2)\cosh^2\left(\dfrac{\omega+\xi}{2T_c}\right)}+\frac{\gamma_1(2\gamma_0+\gamma_1)}{4T_c}\times\int_{-\infty}^{\infty}d\omega\,\frac{\omega^2}{(\omega^2+\gamma_0^2)(\omega^2+(\gamma_0-\gamma_1)^2)\cosh^2\left(\dfrac{\omega}{2T_c}\right)}, \tag{22}$$

$$K_C=\frac{3\pi T_c}{7\zeta(3)\gamma_1}\left\{\frac{2\pi T_c}{\gamma_1}\left[\Psi\left(\frac{1}{2}+\frac{\gamma_0-\gamma_1}{2\pi T_c}\right)-\Psi\left(\frac{1}{2}+\frac{\gamma_0}{2\pi T_c}\right)\right]+\Psi'\left(\frac{1}{2}+\frac{\gamma_0-\gamma_1}{2\pi T_c}\right)\right\}; \tag{23}$$

(B) anisotropic s pairing:

$$K_A=\frac{\gamma_0}{\pi T_c}\left\{\frac{1}{4}\int_{-\omega_c}^{\omega_c}\frac{d\xi}{\xi}\times\int_{-\infty}^{\infty}d\omega\,\frac{\omega+\xi}{(\omega^2+\gamma_0^2)\cosh^2\left(\dfrac{\omega+\xi}{2T_c}\right)}+\frac{2\gamma_0}{\pi}\times\int_{-\infty}^{\infty}d\omega\,\frac{1}{(\omega^2+\gamma_0^2)\cosh^2\left(\dfrac{\omega}{2T_c}\right)}\right\}, \tag{24}$$

$$K_C=-\frac{3(\pi^2-8)}{28\pi^2\zeta(3)}\Psi''\left(\frac{1}{2}+\frac{\gamma_0}{2\pi T_c}\right)+\frac{24\pi^2}{7\zeta(3)\gamma_0^2}\frac{T_c^2}{(\pi^2-8)\gamma^2}\ln\frac{T_c}{T_{c0}}+\frac{6\pi}{7\zeta(3)}\frac{T_c}{\gamma_0}. \tag{25}$$

The results of numerical calculations of the dimensionless coefficients as functions of the parameter γ_0/T_{c0} in the case of d pairing for different values of γ_1/γ_0 are depicted in Figs. 3 and 4.

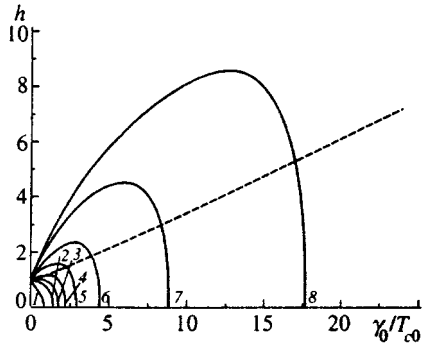


FIG. 5. The normalized slope of the curve of the upper critical field, $h=|dH_{c2}/dT|_{T_c}/|dH_{c2}/dT|_{T_{c0}}$, as a function of the disorder parameter γ_0/T_{c0}. The dashed curve corresponds to the case of s pairing and the solid curves to the case of anisotropic d pairing for the following values of γ_1/γ_0: curve *1*, 0.0; curve *2*, 0.4; curve *3*, 0.5; curve *4*, 0.6; curve *5*, 0.7; curve *6*, 0.8; curve *7*, 0.9; and curve *8*, 0.95.

3. THE UPPER CRITICAL FIELD

As is well known,[11] the behavior of the Ginzburg–Landau coefficients A and C determines the temperature dependence of the upper critical magnetic field near T_c:

$$H_{c2}=\frac{\phi_0}{2\pi\xi^2(T)}=-\frac{\phi_0}{2\pi}\frac{A}{C}, \tag{26}$$

where $\phi_0=c\pi/e$ is the magnetic flux quantum, and $\xi(T)$ is the temperature-dependent coherence length. From this we can easily find the slope of the curve representing the temperature dependence of H_{c2} near T_c, i.e., the temperature derivative of the field:

$$\left|\frac{dH_{c2}}{dT}\right|_{T_c}=\frac{24\pi\phi_0}{7\zeta(3)v_F^2}T_c\frac{K_A}{K_C}. \tag{27}$$

For an s-type superconductor the slope of the curve of the upper critical field is independent of the anisotropic scattering. Figure 5 depicts the dependence of the dimensionless parameter $h=|dH_{c2}/dT|_{T_c}/|dH_{c2}/dT|_{T_{c0}}$ on the disorder parameter γ_0/T_{c0} in the case of d pairing for different values of γ_1/γ_0. In the case of anisotropic s pairing, the slope of the curve of the upper critical field increases with disorder

as usual,[6] and in the strong scattering limit, $\gamma_0 \gg T_c$, the dependence of h on γ_0 becomes linear and the slope is given by the well-known Gor'kov formula[12]

$$\frac{\sigma}{N(0)}\left|\frac{dH_{c2}}{dT}\right|_{T_c}=\frac{8e^2}{\pi^2}\phi_0, \tag{28}$$

where $\sigma=N(0)e^2v_F^2/3\gamma_0$ is the electron conductivity in the normal phase with isotropic s pairing, characteristic of ordinary dirty superconductors. Hence strong disorder suppresses the anisotropy of the gap, and we pass to the usual limit of a dirty superconductor.

In the case of d pairing, the slope of the curve of the field H_{c2} for small values of γ_1/γ_0 rapidly drops to zero on the scale $\gamma_0 \sim T_{c0}$. In the interval $0.5 \leqslant \gamma_1/\gamma_0 \leqslant 0.6$ the behavior of the slope changes dramatically: first h increases slowly but nonlinearly with γ_0/T_{c0}, then it passes through a maximum, and then rapidly drops. The length of the section where the slope grows rapidly increases as $\gamma_1 \to \gamma_0$. We believe that such strong anomalies in the way the slope of the curve of the upper critical field depends on the disorder parameter can be used to determine the pairing type and the possible role of anisotropic scattering in unusual superconductors. Unfortunately, in the case of high-T_c superconducting systems the situation is complicated by the well-known nonlinearity of the temperature dependence of H_{c2}, a feature observed in a broad temperature range starting at T_c, and by a certain indeterminancy in the experimental methods used to determine H_{c2}.

The present work was partially supported by the Russian Fund for Fundamental Research (Project 96-02-16065), by the Statistical Physics State Program (Project IX.1), and by the High-T_c Superconductivity State Program of the Russian Ministry of Science (Project 96-051).

APPENDIX A: CALCULATION OF THE VERTEX PART $\Gamma_{\mathbf{pp}'}$ IN THE LADDER APPROXIMATION

The Bethe–Salpeter equation for the vertex part is

$$\Gamma_{\mathbf{pp}'}=U(\mathbf{p},\mathbf{p}')+\sum_{\mathbf{p}''} U(\mathbf{p},\mathbf{p}'')G^R(\mathbf{p}'')G^A(\mathbf{p}'')\Gamma_{\mathbf{p}''\mathbf{p}'}, \tag{A1}$$

where $U(\mathbf{p},\mathbf{p}')$ is the irreducible vertex part. We take $U(\mathbf{p},\mathbf{p}')$ in the form (the ladder approximation)

$$U(\mathbf{p},\mathbf{p}')=\rho V_0^2+\rho V_1^2 f(\mathbf{p})f(\mathbf{p}'). \tag{A2}$$

Then Eq. (A1) can be written as follows:

$$\Gamma_{\mathbf{pp}'}=\rho V_0^2+\rho V_1^2 f(\mathbf{p})f(\mathbf{p}')+\rho V_0^2\Psi(\mathbf{p}') +\rho V_1^2 f(\mathbf{p})\Phi(\mathbf{p}'), \tag{A3}$$

where

$$\Psi(\mathbf{p}')=\sum_{\mathbf{p}''} G^R(\mathbf{p}'')G^A(\mathbf{p}'')\Gamma_{\mathbf{p}''\mathbf{p}'},$$

$$\Phi(\mathbf{p}')=\sum_{\mathbf{p}''} f(\mathbf{p}'')G^R(\mathbf{p}'')G^A(\mathbf{p}'')\Gamma_{\mathbf{p}''\mathbf{p}'}. \tag{A4}$$

From (A3) we can obtain a self-consisted system of equation for the functions $\Psi(\mathbf{p}')$ and $\Phi(\mathbf{p}')$:

$$\begin{cases}\Psi(\mathbf{p}')=\rho V_0^2 I_1+\rho V_1^2 f(\mathbf{p}')I_2+\rho V_0^2 I_1\Psi(\mathbf{p}')+\rho V_1^2 I_2\Phi(\mathbf{p}'),\\ \Phi(\mathbf{p}')=\rho V_0^2 I_2+\rho V_1^2 f(\mathbf{p}')I_3+\rho V_0^2 I_2\Psi(\mathbf{p}')+\rho V_1^2 I_3\Phi(\mathbf{p}'),\end{cases} \tag{A5}$$

where

$$I_1=\sum_{\mathbf{p}} G^R(\mathbf{p})G^A(\mathbf{p}),$$

$$I_2=\sum_{\mathbf{p}} f(\mathbf{p})G^R(\mathbf{p})G^A(\mathbf{p}),$$

$$I_3=\sum_{\mathbf{p}} f^2(\mathbf{p})G^R(\mathbf{p})G^A(\mathbf{p}). \tag{A6}$$

Solving the system of equations (A5), we arrive at expressions for $\Psi(\mathbf{p}')$ and $\Phi(\mathbf{p}')$ and hence for the vertex part:

$$\Gamma_{\mathbf{pp}'}=\frac{\rho V_0^2(1-\rho V_1^2 I_3+\rho V_1^2 f(\mathbf{p}')I_2)+\rho V_1^2(f(\mathbf{p})f(\mathbf{p}')(1-\rho V_0^2 I_1)+\rho V_0^2 f(\mathbf{p})I_2)}{(1-\rho V_0^2 I_1)(1-\rho V_1^2 I_3)-\rho V_0^2\rho V_1^2 I_2^2}. \tag{A7}$$

APPENDIX B: THE GINZBURG–LANDAU COEFFICIENTS

The diagram (a) in Fig. 2 corresponds to

$$-\frac{T}{(2\pi)^2}\Delta_q^2\sum_{\omega}\int d\mathbf{p}\, 2\cos^2(2\phi)G_{\omega}(\mathbf{p}_+)G_{-\omega}(\mathbf{p}_-)= -\Delta_q^2 TN(0)\sum_{\omega}\int\frac{d\xi}{\widetilde{\omega}^2+\xi^2} +\Delta_q^2 q^2\frac{N(0)\pi v_F^2 T_c}{8}\sum_{\omega}\frac{1}{|\widetilde{\omega}|^3}. \tag{B1}$$

The diagram (c) in Fig. 2 corresponds to

$$-\frac{T}{(2\pi)^2}\Delta_q^2\sum_{\omega}\int d\mathbf{p}\, 2\cos^2(2\phi)G_{\omega}(\mathbf{p})G_{-\omega}(\mathbf{p})=$$

$$-\Delta_q^2 T_c N(0) \sum_{\omega} \int \frac{d\xi}{\widetilde{\omega}^2+\xi^2}. \tag{B2}$$

The diagram with a ''diffuson'' (Fig. 2b) yields

$$-T\sum_{\omega} \sum_{\mathbf{pp}'} \sqrt{2}\cos(2\phi) G^R(\mathbf{p}_+) G^A(\mathbf{p}_-) \Gamma_{\mathbf{pp}'} \times \sqrt{2}\cos(2\phi') G^R(\mathbf{p}'_+) G^A(\mathbf{p}'_-). \tag{B3}$$

If we take (A6) and (A7) into account, Eq. (B3) becomes

$$-TN(0)\pi\gamma_1 \sum_{\omega} \left[\frac{1}{|\widetilde{\omega}|(|\widetilde{\omega}|-\gamma_1)} - \frac{v_F^2(2|\widetilde{\omega}|-\gamma_1)q^2}{8|\widetilde{\omega}|^3(|\widetilde{\omega}|-\gamma_1)^2} \right]. \tag{B4}$$

Note that when there is no anisotropic scattering component, for d-type superconductors the diffusion renormalization due to the diagrams of the type depicted in Fig. 2c is zero to within terms quadratic in q.

Reasoning along similar lines, we arrive at an expression corresponding to the diagram (d) in Fig. 2:

$$-TN(0)\pi\gamma_1 \sum_{\omega} \frac{1}{|\widetilde{\omega}|(|\widetilde{\omega}|-\gamma_1)}. \tag{B5}$$

Writing the expression for $F_s - F_n$ and separating out the coefficients of q raised to the zeroth power and of q^2, we can obtain the expressions for the corresponding Ginzburg–Landau coefficients.

*)E-mail: sadovski@ief.intec.ru

[1] D. Pines, Physica C **235–240**, 113 (1994).
[2] S. Chakravarty, A. Subdø, P. W. Anderson, and S. Strong, Science **261**, 337 (1993).
[3] A. I. Liechtenstein, I. I. Mazin, and O. K. Andersen, Phys. Rev. Lett. **74**, 2303 (1995).
[4] L. S. Borkovski and P. J. Hirschfeld, Phys. Rev. B **49**, 15 404 (1994).
[5] R. Fehrenbacher and M. R. Norman, Phys. Rev. B **50**, 3495 (1994).
[6] A. I. Posazhennikova and M. V. Sadovskiĭ, JETP Lett. **63**, 358 (1996).
[7] G. Harań and A. D. S. Nagi, Phys. Rev. B **54**, 15 463 (1996).
[8] M. V. Sadovskiĭ, Sverkhprovodimost': Fiz. Khim. Tekhnol. **8**, 337 (1995); submitted to Phys. Rep. (1997).
[9] M. V. Sadovskiĭ and A. I. Posazhennikova, JETP Lett. **65**, 270 (1997).
[10] A. A. Abrikosov, L. P. Gor'kov, and I. Ye. Dzyaloshinski, *Quantum Field Theoretical Methods in Statistical Physics*, Pergamon Press, New York (1965).
[11] P. G. de Gennes, *Superconductivity of Metals and Alloys*, W. A. Benjamin, New York (1966).
[12] L. P. Gor'kov, Zh. Éksp. Teor. Fiz. **37**, 1407 (1959) [Sov. Phys. JETP **10**, 998 (1960)].

Translated by Eugene Yankovsky

Investigation of a magnetic phase transition in fcc iron–nickel alloys

S. V. Grigor'ev,* S. A. Klimko, S. V. Maleev, A. I. Okorokov, V. V. Runov, and D. Yu. Chernyshov

St. Petersburg Institute of Nuclear Physics, 188350 Gatchina, Leningrad Region, Russia
(Submitted 21 May 1997)
Zh. Éksp. Teor. Fiz. **112**, 2134–2155 (December 1997)

A magnetic phase transition in carbon-doped (0.1 and 0.7 at. %) $Fe_{70}Ni_{30}$ Invar alloys was investigated by the method of depolarization of a transmitted neutron beam and by small-angle scattering of polarized neutrons. It is shown that for both alloys, two characteristic length scales of magnetic correlations coexist above T_c. Small-angle scattering by critical correlations with radius R_c is described well by the Ornstein–Zernike (OZ) expression. The longer-scale (second) correlations, whose size can be estimated from depolarization data, are not described by the OZ expression, and hypothetically can be modeled by a squared OZ expression, which in coordinate space corresponds to the relation $\langle M(r)M(0)\rangle \propto \exp(-r/R_d)$, where R_d is the correlation length of the second scale. The temperature dependence of the correlation radius R_c was obtained: $R_c \propto ((T-T_c)/T_c)^{-\nu}$, where $\nu \approx 2/3$ is the critical exponent for ferromagnets, over a wide temperature range up to $T_c^{\exp}$, at which the correlation radius becomes constant and equals its maximum value $R_c(T_c)=R_c^{\max}$. The maximum correlation radius established ($R_c^{\max}=140$ Å and 230 Å for the first and second alloys, respectively) characterizes the length-scale of the fluctuation for which the appearance of critical correlations first results in the formation of a ferromagnetic phase, and the phenomenon itself exhibits a ''disruption'' of the second-order phase transition at $T=T_c^{\exp}$, as a result of which a first-order transition arises. Temperature hysteresis was also detected in the measured polarization of the transmitted beam and intensity of small-angle neutron scattering in the alloy above T_c, confirming the character of this magnetic transition as a first-order transition close to a second-order transition.
 [S1063-7761(97)01512-6]

1. INTRODUCTION. TWO LENGTH SCALE PROBLEM

The two length scale problem for phase transitions above T_c has attracted a great deal of attention in recent years. Discussion of this problem commenced with the paper by Andrews[1] on the two-component line detected in measurements of x-ray scattering at a structural phase transition in $SrTiO_3$. A narrow peak superposed on a preexisting wide peak due to conventional critical scattering appeared at temperatures close to the transition temperature T_c.

In the opinion of the author, this indicates the existence of a second, longer scale together with the conventional characteristic scale of critical fluctuations. The existence of a two-component line in the critical scattering of neutrons and x-rays has now been observed in $SrTiO_3$,[1,2] Ho,[3] $RbCaF_3$,[4] $KMnF_3$,[5] Tb,[6] and UPd_2Al_3.[7] It is shown in Ref. 6 that inhomogeneities in Tb that are much larger than the conventional correlation radius are spatially localized within a 0.2 mm thick layer near the surface of the crystal. At the same time, conventional fluctuations are distributed uniformly throughout the entire volume. It is conjectured that the appearance of large clusters in the surface layer is due to the influence of surface defects and elastic deformations on the order parameter.

A corresponding attempt at a theoretical interpretation of these experiments was recently made in Ref. 8, where it is conjectured that the presence of defects engendering large-scale tensile stresses results in the formation of a fluctuation scale that is longer than the conventional fluctuation scale, while the character of the defects (point, linear) changes the critical exponent of the correlation length from $\nu \approx 2/3$ to $\nu_s > \nu$ ($\nu_s = 1,2$ and so on). In Ref. 8 it is also conjectured that the two correlation functions are described by ''Lorentzians'' with substantially different correlation radii. However, in Ref. 7, for example, the existence of an additional, second length scale was observed as a volume effect, and it was shown that the correlation function of the ''second length scale'' is described by a ''squared Lorentzian,'' which, as will be demonstrated below, agrees with the results of our work.

In a recent review,[9] it is concluded on the basis of a set of experiments[1–7] that the interaction of the elastic deformations and order parameter results in the appearance of a second correlation length in critical phenomena above the transition temperature. The question of the form of the correlation function of the ''second length scale'' is solved in Ref. 9 in favor of the ''squared Lorentzian.''

Our objective in the present work is, on the one hand, to demonstrate the coexistence of two length scales of magnetic correlations in Invar FeNi alloys above T_c and, on the other, to develop a method for investigating two length scales of magnetic correlations in ferromagnets.

Magnetic inhomogeneities with a wide spectrum of sizes can be investigated by means of small-angle scattering of polarized neutrons in ferromagnets and to measure at the same time the polarization of the transmitted beam. On the one hand, the typical resolution of small-angle neutron scattering experiments is of the order of $10-10^3$ Å. On the other, magnetic inhomogeneities of magnitude at least about 10^3 Å can be estimated from experiments measuring the depolarization of a transmitted beam.

Investigation of ferromagnetic phase transitions by the method of small-angle neutron scattering yields the correlation radius R_c of critical fluctuations, which are described in the momentum space by an Ornstein–Zernike (OZ) expression. Maleev and Ruban show (in Refs. 10 and 11) that critical correlations at temperatures $T>T_c$ do not yield appreciable depolarization of a neutron beam transmitted through the sample, since the induction is high only at the center of correlation and decreases substantially at $r\sim R_c$. They also establish that depolarization results from the appearance of large magnetic clusters (quasidomains) with a uniform distribution of magnetization within a cluster at $T<T_c$.

Investigations of critical phenomena in a nickel single crystal[12] and in an iron polycrystal[13] confirm the validity of such a description of small-angle scattering of polarized neutrons in complete agreement with Refs. 10 and 11. In contrast, we observe[14,15] an intensification of depolarization of a transmitted neutron beam in a sample of a carbon-doped (0.7 at. %) Invar alloy $Fe_{75}Ni_{25}$ in the paramagnetic temperature range, i.e., in contrast to depolarization in Fe and Ni, the observed depolarization above T_c in the alloy studied was found to be much greater than the polarization computed from the small-angle scattering parameters.

The observation of strong depolarization at temperatures $T>T_c$ can be explained by large-scale magnetic correlations, which are present in the sample together with the conventional critical fluctuations. The investigations established that two characteristic length scales of magnetic inhomogeneities coexist above the Curie temperature T_c. In the present case, T_c is defined to be the temperature at which basic physical quantities, for example, the susceptibility and correlation radius of the critical fluctuations, diverge. Analysis of the polarization of a neutron beam transmitted through a sample showed that magnetic inhomogeneities with characteristic length scale $R_d \geqslant 1000$ Å are formed over a wide temperature range $T>T_c$ in the alloy. Analysis of the magnetic small-angle scattering intensity at the same temperatures yields a different characteristic length scale—the radius of critical fluctuations, which becomes distinguishable within the limits of error ($R_c=40\pm 20$ Å) at $T\approx T_c+25$ K and increases to $R_c=120\pm 10$ Å as T_c is approached. It was also found that the forms of the spin correlation functions corresponding to different length scales of inhomogeneities are different. The correlation function of critical fluctuations is described by the OZ expression, while the long-range correlation can hypothetically be described by a squared OZ expression.

In the present work, we investigated iron–nickel alloys of a different composition ($Fe_{70}Ni_{30}$) and with various admixtures of carbon (0.1 and 0.7 at. %). As will be shown below, the coexistence of two characteristic length scales of magnetic inhomogeneities above T_c was established for both alloys; this shows that the ''two length scale'' situation above T_c is a general situation for fcc FeNi alloys with high iron content.

In the course of the investigations, for both alloys, temperature hysteresis was observed in the measured polarization of the transmitted beam and the small-angle neutron scattering intensity above T_c. Note that in contrast to the materials investigated in Refs. 12 and 13, Fe and Ni alloys with the Invar composition are fundamentally disordered systems.[16] The anomalous behavior of the linear thermal expansion coefficient and Young's modulus in the temperature range from 0 to T_c shows that the magnetism and bulk characteristics of Invar alloys are strongly linked. Disorder in the magnetic subsystem of FeNi Invar alloys resulting from the interaction of the order parameter and elastic deformations in the system near the transition temperature is probably responsible for the emergence of large-scale magnetic correlations at temperatures $T>T_c$, as well as the existence of temperature hysteresis in the magnetic properties near the transition temperature.

This paper is organized as follows. Section 2 is devoted to the theoretical aspects of the method that we employed to investigate two length scales of magnetic correlations in ferromagnets above T_c. Section 3 gives a brief description of the experimental samples and the experimental apparatus on which the measurements were performed. Preliminary results are presented in Sec. 4. In Sec. 5 the Curie temperature of the alloys is determined and the character of the transition is discussed. Sections 6 and 7 are devoted to a demonstration of the existence of two length scales of magnetic correlations above T_c. Finally, the results and the conclusions drawn from the investigations are discussed in Sec. 8.

2. THEORETICAL ASPECTS OF THE METHOD

In the present section we show that measurements of small-angle neutron scattering together with a simultaneous analysis of neutron depolarization make it possible to demonstrate the existence of two length scales of magnetic correlations in ferromagnets above T_c.

It is well known that critical scattering data yield the spin-correlation radius in the system and its temperature dependence, which can be described well on the basis of similarity theory.[17,18] The crux of these ideas about second-order phase transitions is as follows. Magnetization fluctuations appear above the Curie temperature and grow as T_c is approached. The correlation radius R_c, which characterizes long-scale fluctuations, is the only scale that exists in the system above T_c. As T_c is approached, it increases as $R_c \propto \tau^{-\nu}$, where $\tau=(T-T_c)/T_c$ is the relative temperature. At the phase transition ($\tau=0$), the correlation radius becomes infinite. The small-angle neutron elastic scattering cross section above T_c can be described by the OZ expression to high accuracy and increases as T_c is approached[19]:

$$\frac{d\sigma}{d\Omega}=\frac{2}{3}\,r_0^2\gamma^2 S(S+1)\,\frac{1}{r^2}\,\frac{1}{R_c^{-2}+q^2}, \tag{1}$$

where $\mathbf{q}=\mathbf{k}_f-\mathbf{k}_i$, $q\approx|k|\theta$ is the momentum transfer, θ is the scattering angle, $\mathbf{k}_i$ and $\mathbf{k}_f$ are, respectively, the wave vectors before and after scattering, r_0 is the classical electron radius, γ is the gyromagnetic ratio of the neutron, S is the spin of the atom, and r is a constant of the order of the interatomic distance. Thus, the q dependence of small-angle scattering $d\sigma/d\Omega$ makes it possible to obtain the parameter R_c, which is the critical correlation radius at a given temperature.

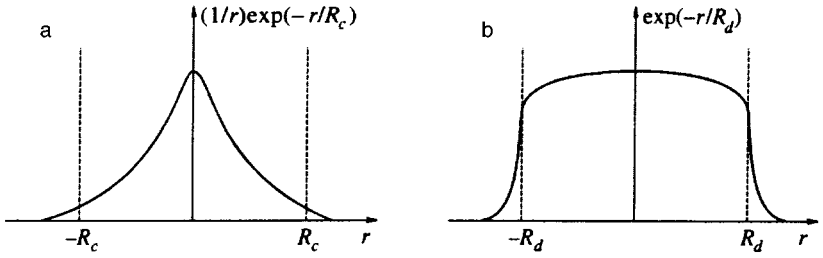


FIG. 1. Schematic curves illustrating the variation of the induction in real space in the case of the critical fluctuations $(1/r)\exp(-r/R_c)$ (a) and in the case of a magnetic cluster $\exp(-r/R_d)$ (b). The neutron scattering cross section is proportional to the Fourier transform of the expressions presented, and corresponds in the first case to the OZ formula (1), and in the second case to the squared OZ expression.

In Refs. 10 and 11 it is shown that the depolarization of neutrons that have passed through a sample can be regarded as resulting from neutron scattering within the angular width of the transmitted beam. Depolarization is determined by the total cross section σ_Ψ (per magnetic atom), in the central detector with angular aperture Ψ. The relation between the directions of the polarization vector $\mathbf{P}_0$ and the neutron wave vector $\mathbf{k}$ is also important for the degree of polarization of the transmitted beam. Thus, in the case $\mathbf{P}\perp\mathbf{k}$ the polarization is given by

$$P=P_0\ \exp[(-3/2)\sigma_\Psi N_0 L], \tag{2}$$

where N_0 is the density of magnetic atoms and L is the thickness of the sample. The contribution of critical fluctuations in the system above T_c to depolarization of the transmitted neutron beam is estimated in Ref. 10. In the case of critical fluctuations, the total scattering cross section σ_Ψ is easily calculated by integrating the expression (1) over the aperture of the detector $[0,\Psi]$. In the case at hand, the total cross section σ_Ψ is small and depends on τ logarithmically:

$$\sigma_\Psi=\frac{2}{3}\ \pi r_0^2\gamma^2 S(S+1)\ \frac{1}{(kr)^2}\ \ln\left[1+\left(\frac{k\Psi}{\kappa}\right)^2\right], \tag{3}$$

where $\kappa=R_c^{-1}$ is the reciprocal of the correlation radius. In real experiments, temperature stabilization and temperature gradients in the sample ordinarily limit how closely T_c can be approached: $\tau_{\min}\sim 10^{-4}$, i.e. measurements of R_c are limited to the temperature range $|\tau|\geqslant\tau_{\min}$. Substituting this minimum value of τ into the expression for σ_Ψ with $\lambda\approx 9$ Å yields the cross section $\sigma_\Psi\leqslant 0.2$ b. For L of the order of several millimeters and $N_0\approx 10^{23}$, depolarization is of the order of several percent. For example, in the case of a neutron beam with wavelength 3 Å passing through a 0.5 cm thick Ni single crystal, depolarization does not exceed 1% even with $\tau=10^{-4}$.[12] Over the entire paramagnetic range ($\tau>\tau_{\min}$), therefore, the polarization P of the transmitted beam is virtually identical to the initial polarization P_0.

We note here that the expression (2) is valid in the Born approximation, the criterion of applicability of which to scattering by magnetic inhomogeneities (diffraction limit) can be formulated as the condition that the angle of precession of the polarization vector in the magnetic field $\mathbf{B}$ of an inhomogeneity R_d be small:

$$\gamma B R_d\lambda/\beta\ll 2\pi. \tag{4}$$

Here λ is the wavelength [Å], R_d is the average correlation radius [cm], and $\beta=3.958\cdot 10^5$ [Å cm$\cdot$s^{-1}] is a constant relating the velocity and wavelength of the neutron. This approximation is reasonable for studying depolarization, since in the case when the polarization vector turns in the field of the inhomogeneity by an angle greater than approximately π, complete effective depolarization of the transmitted beam occurs.

At the same time, as shown in Ref. 20, in the approximation (4), the polarization of a neutron beam transmitted through the sample decreases with wavelength as $\exp(-\alpha\lambda^2)$, where α is a constant. When the opposite inequality holds, the depolarization of the neutrons is independent of wavelength. Therefore, the experimentally measured λ^2 dependence of the polarization should indicate whether or not condition (4) is satisfied and the diffraction limit is applicable.

We note here that with the exception of the dependence of the degree of depolarization on the angular width of the beam, any phenomena associated with depolarization can be formulated in classical terms. We assume that magnetic inhomogeneities are large enough that all neutrons are scattered into the angular aperture of the detector, which renders the quantum and classical approaches equivalent.

Figure 1 displays schematically the variation of the induction in real space in the case of a critical fluctuation ($(1/r)\exp(-r/R_c)$, panel *a* and in the case of a magnetic cluster ($\exp(-r/R_d)$, panel *b*). The neutron scattering cross section is proportional to the Fourier transform of the expressions presented, and corresponds in the first case to the OZ formula (1), and in the second case the squared OZ expression.

It was shown above that critical fluctuations (Fig. 1a) cannot make a large contribution to beam depolarization, and strong beam depolarization requires the presence of large magnetic inhomogeneities with sizes $R_d>R_c$, within which the induction distribution is comparatively uniform (Fig. 1b). Indeed, as shown in Ref. 11, the small-angle magnetic elastic scattering cross section is

$$\sigma_\Psi=\frac{2}{3}\ r_0^2\gamma^2 S^2\int_\Psi d\Omega N_0\int d\mathbf{r}\ \exp(i\mathbf{q}\cdot\mathbf{r})\langle\mathbf{m}(\mathbf{r})\cdot\mathbf{m}(0)\rangle, \tag{5}$$

where $S\mathbf{m}(\mathbf{r})$ is the average spin of a magnetic atom at temperature T, and is related to the magnetization density $M(T,\mathbf{r})$ by

$$g\mu_0 N_0 S\mathbf{m}(\mathbf{r})=M(T,\mathbf{r})=M_0(T)\mathbf{m}(\mathbf{r}),\quad \mathbf{m}^2(\mathbf{r})\leqslant 1. \tag{6}$$

Here $M_0(T)$ is the maximum magnetization density of a cluster. If the classical limit is correct, then the integral over

$d\Omega$ in the expression (5) can be easily calculated and the exponent $\sigma_\Psi N_0 L$ in Eq. (2) can be rewritten in the form

$$\sigma_\Psi N_0 L = \frac{1}{3}\left(\frac{\gamma B(T)}{v}\right)^2 R_d L, \tag{7}$$

where $B=4\pi M$ is the induction inside a cluster, v is the neutron velocity, and R_d is the effective size of a magnetic cluster,

$$R_d = \frac{1}{2}\int_{-\infty}^{\infty} dz \langle \mathbf{m}(z,0)\cdot\mathbf{m}(0,0)\rangle, \tag{8}$$

where the integration extends along the neutron path through the cluster, and the average is taken over all clusters in the sample.

It is clear in Eq. (8) that the effective size R_d of an inhomogeneity depends strongly not only on the characteristic size of the correlation but also on the form of the correlation function $\langle \mathbf{m}(z,0)\cdot\mathbf{m}(0,0)\rangle$. In the case of critical fluctuations, the effective size of a correlation as defined in Eq. (8), $R_d \sim a\ \ln(1/\kappa a) \ll 1/\kappa = R_c$, is small for any reasonably attainable values of the correlation radius. Since in fact the small effective size R_d, and not R_c, appears as a parameter in the expression (7) and determines the degree of depolarization, the onset of critical fluctuations results in only a relatively small depolarization of the neutron beam.

In the case of an inhomogeneity with slowly decreasing induction, the effective size R_d approximately equals the correlation radius of the inhomogeneity, and therefore expression (2) can be rewritten in the form

$$\frac{P}{P_0} = \exp\left[-\frac{1}{2}\left(\frac{\gamma B(T)}{v}\right)^2 R_d L\right], \tag{9}$$

which is identical, to within a factor of 2/3 in the exponent, to the Halperin–Holstein formula.[20] It is shown in a recent paper[21] that the expression (9) also obtains when allowance is made for demagnetization fields in the classical analysis of this problem.

Thus, measurements of the depolarization of the transmitted beam make it possible to determine the scattering cross section σ_Ψ for $q<k\Psi=q_{\min}$, i.e., they yield an integral representation of long-scale magnetic clusters with characteristic size $R>1/q_{\min}$ and a uniform magnetization distribution in a cluster. On the other hand, for small-angle neutron scattering, the characteristic momentum transfer lies in the range $q>k\Psi$, which corresponds to inhomogeneities with characteristic scale $R<1/q_{\min}$. The shape of the correlation function from which scattering takes place can be determined from the form of the q-dependence of the scattering cross section.

3. EXPERIMENT

In the present work, we investigated polycrystalline samples of a carbon-doped (0.1 and 0.7 at. %) fcc iron–nickel $Fe_{70}Ni_{30}$ alloy, which we refer to below as the ''first'' (I) and ''second'' (II) samples, respectively. The samples were produced in an argon atmosphere in an induction furnace, homogenized at 1200 °C for 4 h, and then quenched in water. The samples were certified for fcc structure by the neutron-diffraction method.

Measurements of the intensity of small-angle scattering of polarized neutrons as well as the polarization of the transmitted neutron beam were performed on the samples over a wide temperature range near the transition temperature T_c in 1 K steps. The experiments were performed with the ''Vektor'' multidetector small-angle polarized-neutron scattering setup at the VVR-M reactor of the St. Petersburg Institute of Nuclear Physics. The setup is described in Ref. 22. The scattered neutrons were detected in the momentum-transfer range $3\cdot10^{-3}\leqslant q\leqslant 3\cdot10^{-2}$ Å^{-1} with step $\Delta q=3\cdot10^{-3}$ Å^{-1}. The spectrum-averaged wavelength was $\lambda=9.2$ Å ($\delta\lambda/\lambda=20\%$).

The magnetic scattering intensity I_s analyzed in the present work was defined as the excess above the nuclear scattering, which was taken to be the scattering by the fcc lattice of the alloy near the paramagnetic temperature T_p ($T_p=370$ K for alloy I and $T_p=470$ K for alloy II):

$$I_s(T,q) = I(T,q) - I(T_p,q)R(T), \tag{10}$$

where $I(T,q)=I^{\uparrow}(T,q)+I^{\downarrow}(T,q)$ is the sum of the intensities of the transmitted beam with spin in the state $I^{\uparrow}(T,0)$ parallel to the magnetic field, and in the state $I^{\downarrow}(T,0)$ antiparallel to the magnetic field at temperature T, and $R(T)=I(T,0)/I(T_p,0)$ is the beam attenuation due to the magnetic phase transition. The factor $R(T)$ (10) is close to 1, and yields the second-order correction to the scattering intensity I_s. The polarization $P(T)$ of the transmitted neutron beam was defined in the standard manner:

$$P(T) = \frac{I^{\uparrow}(T,0) - I^{\downarrow}(T,0)}{I^{\uparrow}(T,0) + I^{\downarrow}(T,0)}. \tag{11}$$

We also measured the wavelength dependence of the neutron depolarization for the ''first'' sample at several temperatures near T_c. The measurements were performed with the SPN-1 spectrometer of the pulsed IBR-2 reactor (Dubna). The neutron polarization was analyzed in the range from 1 to 10 Å. In the course of the experiments, the polarization vector of the neutron beam was directed perpendicular to the neutron velocity vector. The guiding field was equal to 1 Oe. The temperature measurements were performed in a helium-filled furnace and temperature stabilization to within 0.1 K.

Structural investigations of the samples were performed with a Mini–SFINKS powder diffractometer (St. Petersburg Institute of Nuclear Physics, Gatchina).[23] The samples were positioned perpendicular to the incident neutron beam. The spectra for both samples were identified unequivocally as corresponding to fcc structure ($Fd3m$, space group 225). No additional reflections indicating the presence of additional phases or ordering of the metal atoms were observed. The following structural parameters were found from the spectra by the Rietveld method (the MRIA package[24] was used): lattice constant 3.58549(2) Å for alloy I and 3.59440(3) Å for sample II; Debye–Waller factor $U_{\rm iso}=0.072(14)$ Å^2 for alloy I and $U_{\rm iso}=0.243(14)$ Å^2 for alloy II.

The lattice constant and Debye–Waller factor ($U_{\rm iso}$) increase with carbon content. It is natural to attribute the in-

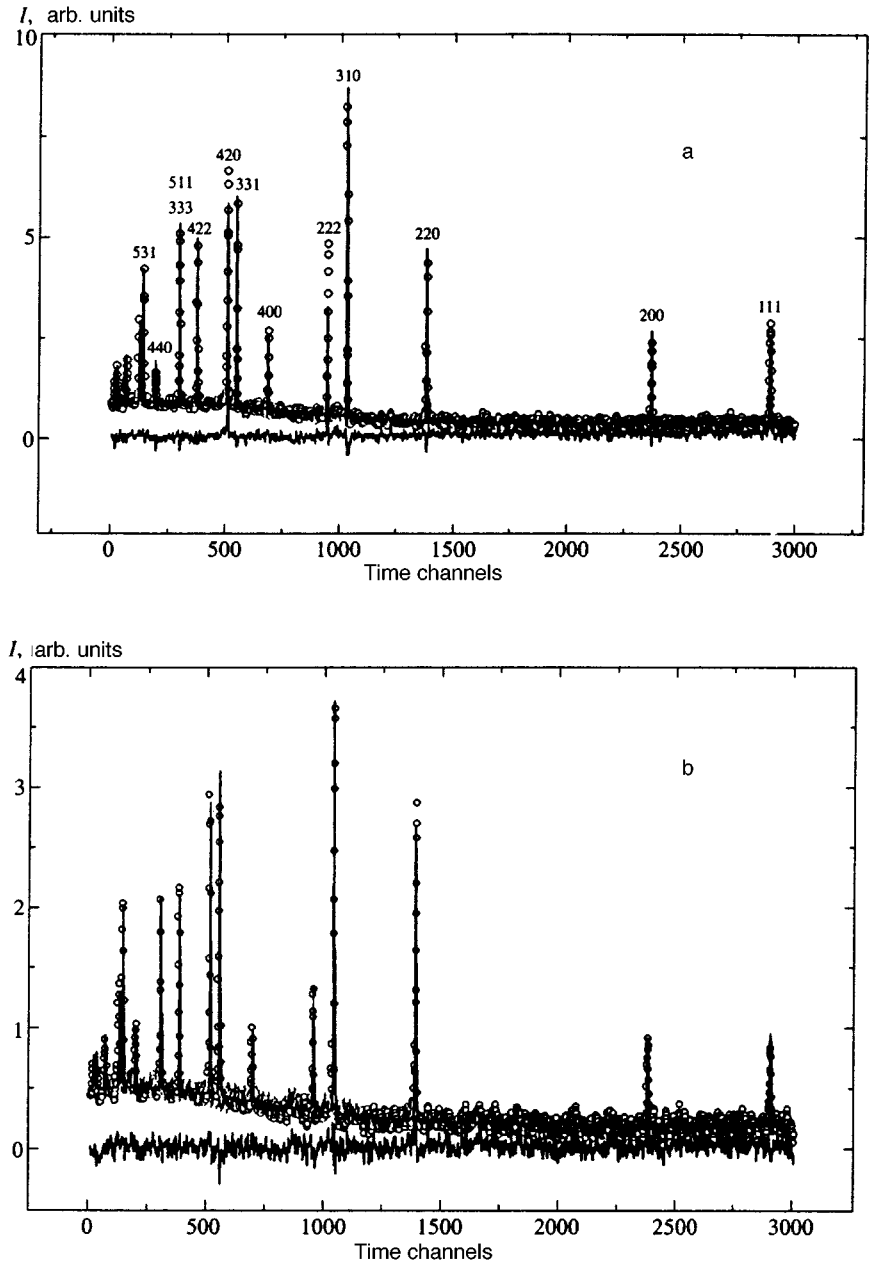


FIG. 2. Diffraction patterns of samples I (a) and II (b) and difference curves of the theoretical and experimental spectra. The circles show the experimental points and the solid lines show the theoretical spectra corresponding to the computed parameters.

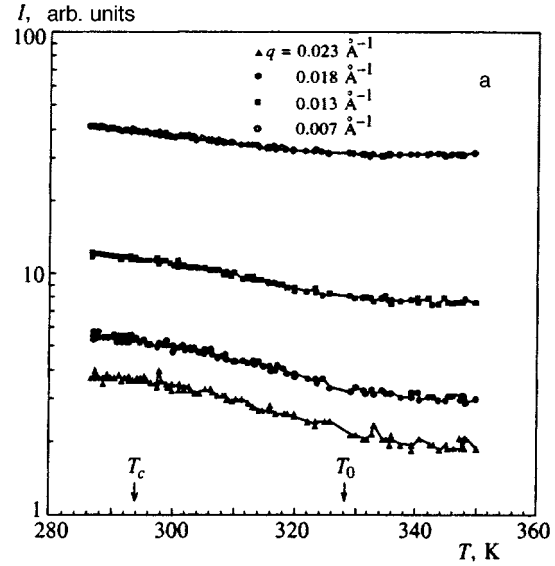
crease in the lattice constant to an increase in the cell volume,[16] and the increase in U_{iso} to static displacements of the metal atoms induced by the insertion of carbon. The shape and sizes of these samples, which are atypical for powder measurements, did not make it possible to properly take a number of systematic corrections into account. Thus, the structural parameters obtained at this stage of the investigations are effective parameters,[25] and have only comparative significance.

The difference curves presented in Fig. 2 (a—for alloy I, b—for alloy II) together with the theoretical and experimental spectra show that the intensities of the reflections (420, 222, 111) for the composition $Fe_{70}Ni_{30}$ (0.1% C) and (220) for the composition $Fe_{70}Ni_{30}$ (0.7% C) are poorly described. One reason for the discrepancies between measured and computed the intensities of separate reflections might be texture, which could be determined only via additional measurements.

4. RESULTS

The overall features of the behavior of the Invar FeNi alloy near T_c were partially studied in Refs. 14 and 15 and mentioned in the introduction. The pattern of small-angle scattering and depolarization for the alloys studied in the present work is qualitatively identical to that described in Refs. 14 and 15 for $Fe_{75}Ni_{25}$ doped with 0.7 at. % carbon.

The temperature dependence of scattering intensity $I(T)$ for different values of momentum transfer q (in the temperature intervals 285–350 K and 380–450 K for alloys I (a) and II (b), respectively) are displayed in Fig. 3. At a phase transition with increasing temperature (forward direction), the



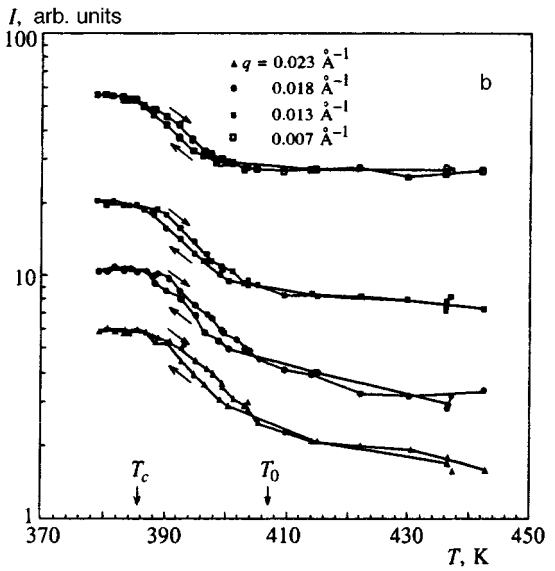
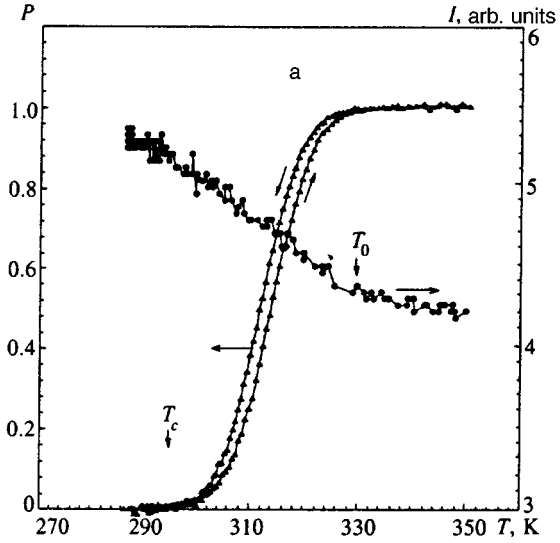


FIG. 3. Small-angle scattering intensity versus temperature for various momentum transfers q for alloys I (a) and II (b). The arrows show the sequence of temperature variations in the measurements.



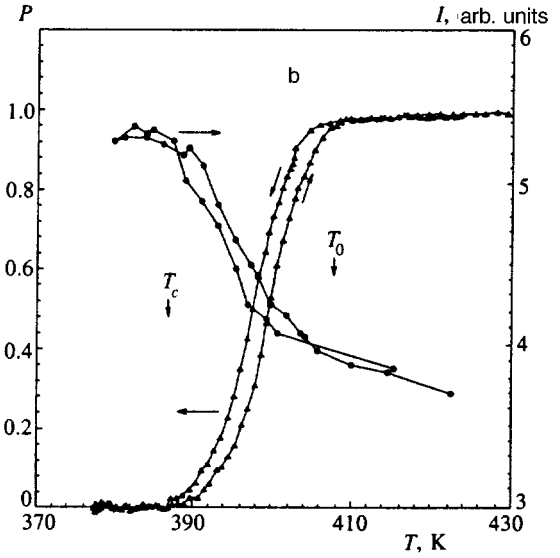


FIG. 4. Temperature dependence of the polarization of a neutron beam transmitted through the sample ($q<0.003$ Å^{-1}) and small-angle scattering intensity for $q=0.018$ Å^{-1} for the alloys I (a) and II (b).

small-angle neutron scattering intensity decreases to the nuclear scattering intensity. In this temperature range, small-angle magnetic scattering can be neglected compared with nuclear scattering. A subsequent decrease in the temperature of alloy II (reverse direction) resulted in an appreciable displacement of the temperature dependence $I(q,T)$ in the direction of low temperatures, and the formation of a hysteresis curve (Fig. 3b). The arrows in the figure show the direction of variation of the temperature in the course of the measurements. We note that on account of the long duration of the measurements, and therefore their high cost, measurements of small-angle neutron scattering intensity with decreasing temperature were not performed for alloy I. However, as we show below (Fig. 4a), hysteresis of the polarization of the transmitted beam occurred in alloy I, just like alloy II (Fig. 4b).

Figure 4 displays the temperature dependence $P(T)$ ($\lambda=9.2$ Å) of the polarization of the transmitted neutron beam and the small-angle scattering intensity for $q=0.018$ Å^{-1} for both alloys. At high temperatures ($T\gg T_c$), the polarization of the transmitted beam equals the polarization P_0 of the incident beam. However, contrary to the conventional notion[12,13] that the polarization decreases rapidly (10%/0.1 K) at a paramagnet–ferromagnet transition, in the present case the beam polarization starts to decrease continuously (at a rate of the order of 10%/3 K) at some temperature that we denote by T_0 ($T_0\approx 330$ K for alloy I and ≈ 410 K for alloy II). The small-angle neutron scattering intensity (Fig. 4) at temperatures $T<T_0$ increases, demonstrating the existence of critical fluctuations in the sample.

As will be shown below, we determine T_c from the small-angle scattering data to be the temperature at which the radius of critical correlations takes on its maximum value. T_c is then tens of degrees less than T_0 (the temperature at which the polarization starts to decrease), and at the Curie tempera-

ture the neutron beam is completely depolarized ($P=0$). The fact that the polarization starts to decrease at a temperature some tens of degrees higher than the Curie temperature would be unusual for standard magnets.

As noted in the introduction, critical fluctuations near T_c depolarize the beam very little. This is explained by the fact that the effective magnetization of a fluctuation that affects the neutron spin is weak, because the magnetization decreases rapidly (as $1/r$) at a distance of the order of the size of the fluctuation (Fig. 1a). Thus, the presence of depolarization in this temperature range attests to the emergence of magnetic inhomogeneities with a large radius and slowly decreasing correlation function (Fig. 1b).

Weak but clearly distinguishable temperature hysteresis in the polarization with increasing and decreasing temperature can be seen in Fig. 4. We note that no relaxational phenomena were observed in the measurements. The polarization measured at the temperature of maximum hysteresis ($P=50\%$, $\Delta P\approx 15\%$) remained constant to within 0.5% for 10 h. The polarization difference $\Delta P=P'-P''$, where P' is the polarization at the 50% level with temperature decreasing from high values, and P'' is the polarization at the temperature corresponding to P' but with temperature increasing from low values, was taken as a measure of the hysteresis. The temperature ranges where hysteresis is observed in the small-angle neutron scattering intensity and in the neutron depolarization are identical, as is the temperature T_0 at which the scattering intensity increases and the polarization decreases with decreasing temperature. We also note that as the magnetic field applied to the sample increases, the magnitude of the temperature hysteresis decreases until it vanishes completely at some tens of oersteds.[26]

5. CURIE TEMPERATURE. ANALYSIS OF THE q DEPENDENCE OF THE SCATTERING INTENSITY

An important question for the present investigation is the question of the Curie temperature. For Invar alloys, the concept of a Curie point is not as clearcut as in the case of pure ferromagnetic materials. For example, the spontaneous magnetization M_s as a function of temperature for low values near the transition temperature does not cut off abruptly at $T=T_c$ but decreases smoothly with increasing T.[16] Strictly speaking, no definite value of T_c can be associated with such a temperature dependence $M_s(T)$. Various methods for determining T_c were examined by Belov,[27] and they yield a discrepancy of 10 K or more for Invar alloys, while for Ni, for example, the values of T_c found by various methods agree to within 1 K.

The Curie temperature can be accurately determined from small-angle neutron scattering experiments, since in the limit $q\ll\kappa$ the small-angle scattering cross section is proportional to the static magnetic susceptibility of the system near the transition.[17] Therefore, the maximum of the temperature dependence of the scattering intensity (scattering cross section) corresponds to the maximum of the susceptibility, and indicates the Curie temperature T_c quite accurately on the temperature scale. At the same time, as pointed out in Sec. 2, the radius of the critical correlations for standard magnets must approach infinity at $T=T_c$. In what follows, we obtain T_c for the experimental alloys as the temperature at which R_c assumes its maximum value, since for the disordered systems investigated here, the correlation radius increases with decreasing temperature to some T' and then becomes constant. We call this temperature the Curie temperature of the given alloy, $T_c=T'$. We note that this method of determining T_c is preferable to other methods, since the measurements can be performed in essentially a vanishing field.

To start, we verify that the correlation functions for spins at which small-angle scattering occurs, i.e., scattering in the range $q\in[3\times10^{-3}, 3\times10^{-2}]$ Å^{-1}, can be described by the OZ expression, i.e., it corresponds to the correlation function of the critical fluctuations. Theoretical and experimental estimates of the contribution of inelastic scattering compared with elastic scattering in the range $q\geqslant\kappa$ have shown that inelastic scattering can be neglected.

Figure 5 shows the reciprocal of the small-angle scattering intensity I^{-1} versus the squared momentum transfer q^2 for different temperatures $T>T_c$ for alloys I (a) and II (b). The curves are linear to within the statistical errors. This confirms the correctness of using an OZ-type relation

$$I(q,\kappa)=\frac{A}{q^2+\kappa^2}, \tag{12}$$

where $\kappa=R_c^{-1}$ is the reciprocal of the correlation radius and A is a free parameter, to analyze the small-angle scattering data $I(q)$. It is evident from the figure that at $T_c\approx 293$ K and $T_c\approx 386$ K for alloys I and II, respectively, the linear dependence $I^{-1}(q^2)$ extrapolated to $q^2\to 0$ does not vanish. This means that near the Curie temperature, $\kappa=R_c^{-1}$ does not vanish but instead approaches a constant.

To obtain the final results, the function $I(q,\kappa)$ was convolved with the instrumental resolution function, which is approximated by a Gaussian function with standard deviation 3.3×10^{-4} Å. The experimental data on $I(q)$ were analyzed in the range $q\in[0.01,0.025]$ Å^{-1} by the least-squares method. The two independent parameters A and κ^2 were determined from the analysis for each value of the temperature T.

The temperature dependence of the squared reciprocal of the correlation radius $\kappa^2(T)$ for both alloys, which decreases with decreasing temperature in the range $T>T_c$ ($\tau\in[0.02,0.2]$) and then becomes constant and different from zero at $T\approx T_c^{\rm exp}$, is displayed in Fig. 6. We note that the increase in the correlation radius $R_c(T)$ is bounded by the value $R_c^{\max}(T_c^{\rm exp})=140\pm10$ Å for alloy I and $R_c^{\max}(T_c^{\rm exp})=230\pm10$ Å for alloy II. The difference between $R_c^{\max}(T_c^{\max})$ for the two alloys indicates that this limit is not determined by the resolution of the apparatus, which sets an upper limit on the experimental range of $R\approx 300$ Å, but is instead a characteristic of the magnetic transition in the samples.

Thus, we are observing a ''disruption'' of the transition at the temperature T_c at which the critical correlation radius reaches the value $R_c^{\max}$. The temperature dependence κ^2 is described by a power law

$$\kappa^2=\kappa_1^2+\kappa_0^2\tau^{2\nu}, \tag{13}$$

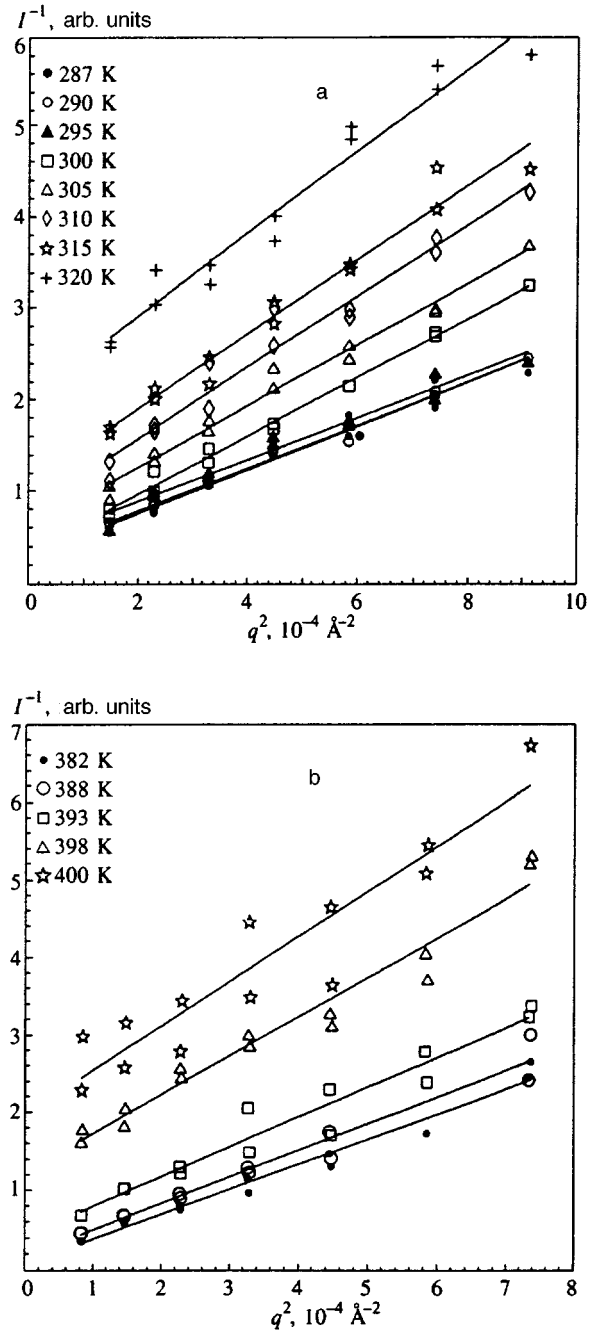


FIG. 5. Reciprocal of the small-angle scattering intensity I^{-1} versus the squared momentum transfer q^2 for different temperatures $T>T_c$ for alloys I (a) and II (b).

TABLE I. Parameters obtained by a least-squares fit of the temperature dependence of the reciprocal of the correlation radius: $\kappa^2=\kappa_1^2+\kappa_0^2\tau^{2\nu}$, where $\tau=(T-T_c)/T_c$.

Alloy	κ_0^2, Å^{-2}	κ_1^2, Å^{-2}	T_c, K	2ν
$Fe_{70}Ni_{30}$ 0.1% C, heating	0.006±0.002	4.5×10^{-5}	293±1	1.3±0.1
$Fe_{70}Ni_{30}$ 0.7% C, heating	0.03±0.01	2×10^{-5}	388±0.5	1.4±0.2
$Fe_{70}Ni_{30}$ 0.7% C, cooling	0.03±0.007	2×10^{-5}	385.5±0.3	1.37±0.06

where ν is the critical exponent, $\tau=(T-T_c^{\exp})/T_c^{\exp}$ is the reduced temperature, $\kappa_1^2=(R_c^{\max})^{-2}$, and $T_c^{\exp}$ is the temperature at which $R_c(T)$ assumes its maximum value. The parameters κ_0^2, κ_1^2, $T_c^{\exp}$, and ν were determined by least-squares analysis.

The resulting values are presented in Table I. The expression (13) with the values $\kappa_1^2=0$ and κ_0^2 determined in the preceding analysis was also fit to the experimental temperature dependence $\kappa^2(T)$. As a result, two free parameters of the least-squares fit were obtained: T_c^{theor} and ν. The quantity ν remained equal to its previous value, to within the limits of error, and the values obtained for T_c^{theor} were 286±1 K for alloy I, and 386.4±0.5 K with heating and 384.3±0.3 K with cooling for alloy II.

The curves $\kappa^2(T)$ obtained with the least-squares parameters are displayed in Fig. 6 (solid curves). They are essentially identical for both types of fits. Especially interesting is the fact that for alloy II, the curves obtained with increasing and decreasing temperature are shifted by 2–3 K, and in consequence the temperatures at which the correlation radius $R_c(T)$ assumes its maximum value differ by 2.5 K, i.e., approximately by the width of the hysteresis, for increasing and decreasing temperature. Thus, on account of hysteresis, the real Curie temperature depends on the history of the sample. Nonetheless, the Curie temperature can be regarded as established to within the hysteresis width (2–3 K) (see Table I).

The parameter κ_0^2 is identical, to within the limits of error, for increasing and decreasing temperature for alloy II, and is quite different from the parameter κ_0^2 for alloy I (Table I). The parameter κ_0, being a coefficient in the power-law relation $R_c(\tau)=\kappa_0^{-1}\tau^{-\nu}$, is effectively the inverse correlation radius at $\tau\sim1$, i.e., in the paramagnetic region for $T\gg T_c$. The values $\kappa_0^{-1}\approx12$ Å and ≈6 Å obtained for alloys I and II, respectively, indicate the presence of critical fluctuations of this scale in the experimental alloys in the far paramagnetic region.

6. ANALYSIS OF NEUTRON POLARIZATION

The polarization $P(\lambda)$ of the transmitted neutron beam as a function of the wavelength was measured at various temperatures near the transition for sample I: $T=289$ K, 303.6 K, and 311.4 K. The resulting λ dependence of the polarization of the transmitted beam (normalized to the polarization of the transmitted beam in the paramagnetic region at $T=370$ K) is described well by an exponential function $P(\lambda)/P_0(\lambda)=\exp(-\alpha\lambda^2)$. To verify this, we constructed a plot of $f=-\ln(P/P_0)$ as a function of λ^2 (Fig. 7). The data demonstrate a linear dependence of f on λ^2 in the wavelength range 1–4 Å. The lack of oscillatory behavior in the resulting curves suggests that the direction of the induction inside the magnetic inhomogeneities is completely disordered. As already mentioned in Sec. 2, the variation of polarization with wavelength according to the law $\exp(-\alpha\lambda^2)$ demonstrates the existence of magnetic correlations in which the polarization vector of the neutrons precesses by a ''small'' angle. In other words, the condition (4) ought to be satisfied.

In the approximation (4), depolarization can be regarded

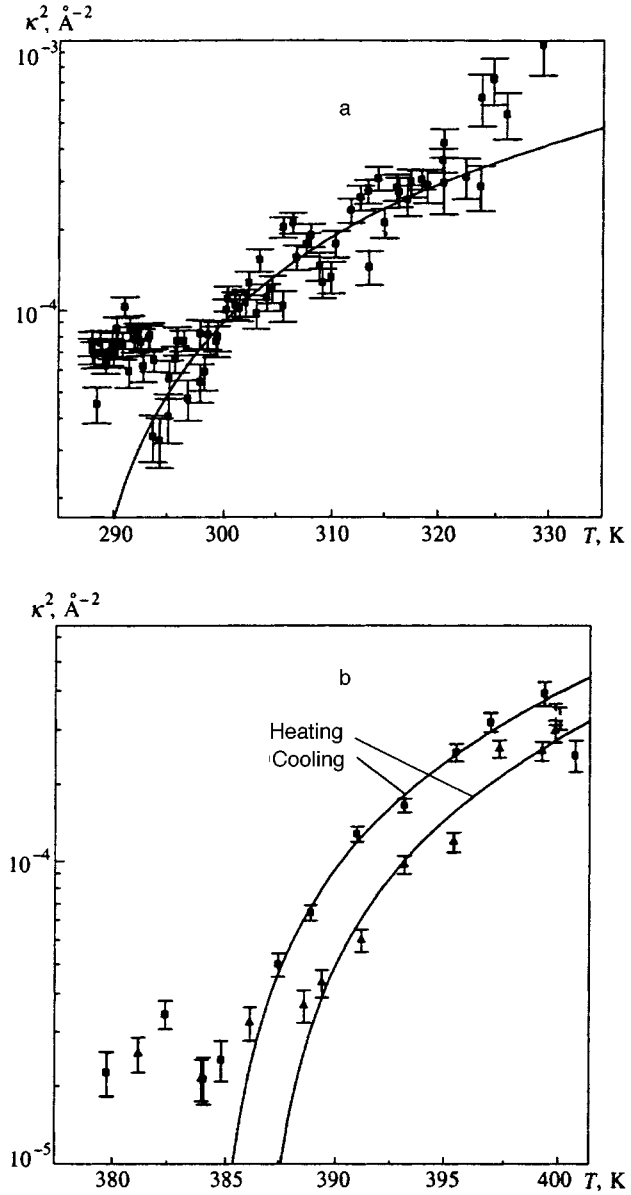


FIG. 6. Squared reciprocal κ^2 of the correlation radius versus temperature T for alloys I (a) and II (b) (increasing and decreasing temperature). The solid curves show the temperature dependence $\kappa^2(T)$ computed using the least-squares parameters.

as being due to uncorrelated small rotations of the polarization vector in the magnetic field of inhomogeneities of size R_d. Equation (9) is valid in this case.

The scale of the magnetic inhomogeneity on which depolarization occurs can be estimated roughly by means of the expression (9) and the temperature dependence of the polarization $P(T)$ at neutron wavelength $\lambda = 9.2$ Å (Fig. 4). Measurement of the magnetization by the method of three-dimensional polarization analysis of the transmitted neutron beam shows that the induction B near T_c does not exceed 0.3 kG for sample I, and 0.5 kG for sample II.[26] In our case $L = 0.13$ cm. Substituting these values of the induction B into the expression (9) and assuming its distribution to be uniform within a magnetic cluster, we obtain from Eq. (9) an estimate of the lower limit on the cluster size R_d. The values

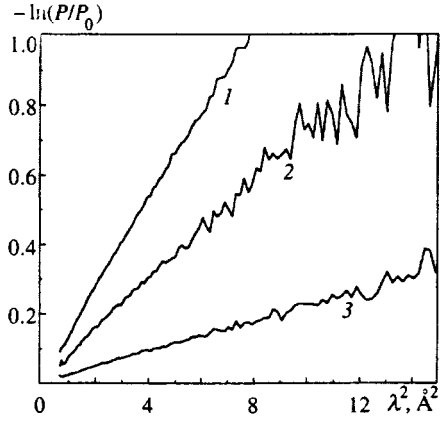


FIG. 7. The quantity $-\ln(P/P_0)$ versus λ^2 at various temperatures near the transition for alloy I: *1*—$T = 289$ K, *2*—303.6 K, *3*—311.4 K.

obtained for the sizes $R_d(T)$ of the magnetic inhomogeneities are of the order of $10^3 - 10^5$ Å, which is hundreds of times greater than R_c over the entire temperature range $T \geqslant T_c$.

7. NEUTRON MEAN FREE PATH

The coexistence of two length scales of magnetic correlations in the system can be demonstrated by using the concept of a neutron mean free path.

As noted in Sec. 2 (with citations to Refs. 10 and 11), the depolarization of the neutrons transmitted through the sample can be regarded as being due to neutron scattering by magnetic inhomogeneities within the angular width of the transmitted beam. Then the depolarization is determined by the total cross section σ_Ψ, normalized to one magnetic atom, in the central detector with angular width Ψ. Thus, the polarization in the case $\mathbf{P} \perp \mathbf{k}$ is given by the relation (2).

At the same time, attenuation of the neutron beam I/I_0 can be regarded as being due to absorption and scattering of neutrons outside the angular aperture of the central detector. Here I and I_0 are, respectively, the intensities of the transmitted and incident beams. A temperature-dependent part $I(T)/I(T_p)$, where $I(T_p)$ is the intensity of the transmitted beam in the paramagnetic region, can be separated out of the total attenuation of the beam. The temperature-dependent part of the beam attenuation characterizes the magnetic scattering cross section $\sigma^m_{1-\Psi}(T)$ near the phase transition, normalized to one magnetic atom, outside the central detector with angular width Ψ. The attenuation of the neutron beam $t(T) = I(T)/I(T_p)$ due to the magnetic scattering at the phase transition is

$$I(T)/I(T_p) = \exp[-\sigma^m_{1-\Psi}(T) N_0 L], \tag{14}$$

where N_0 and L are the same quantities as in the case of the polarization given by the expression (2).

Note that the cross section $\sigma^m_{1-\Psi}(T)$ neglects Bragg scattering, since the neutron wavelength ($\lambda = 9.2$ Å) is greater than the lattice constant of the alloy ($a \approx 3.6$ Å) and therefore the conditions for Bragg reflection are not satisfied. As mentioned above, in the region $q \geqslant \kappa$ the contribution of inelastic scattering is small compared to elastic scattering.

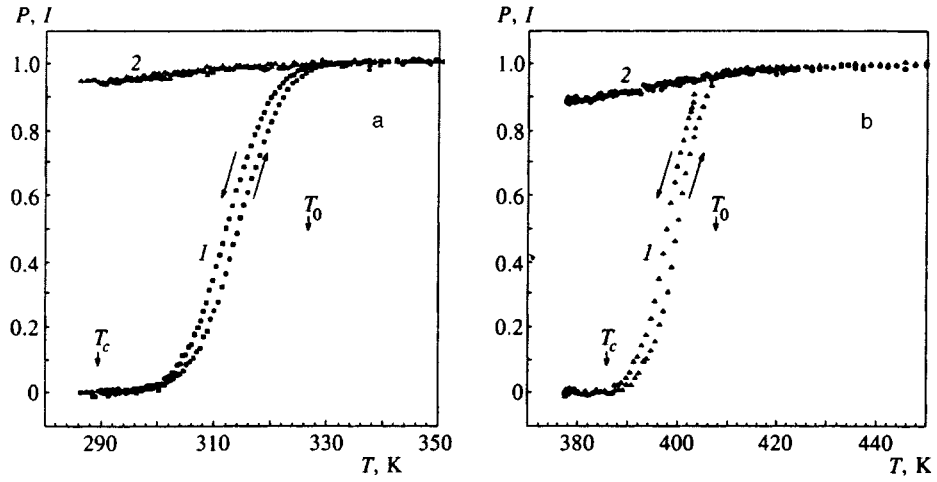


FIG. 8. Polarization $P(T)$ *(1)* and attenuation $I(T,q>0.002\ \text{Å}^{-1})$ *(2)* of the beam versus temperature for alloys I (a) and II (b).

We therefore regard $\sigma^m_{1-\Psi}(T)$ as the elastic scattering cross section of magnetic correlations near the transition.

Thus, the scattering cross section σ_Ψ for $q<k\Psi=q_{\min}$ can be determined by measuring the depolarization of the transmitted beam, i.e., it yields an integral representation of the long-scale magnetic correlations with characteristic size $R\to 1/q_{\min}$. On the other hand, the beam attenuation, which characterizes the removal of neutrons from the transmitted beam at the phase transition, is determined by the total scattering cross section in the region $1>q>k\Psi$, which corresponds to inhomogeneities with characteristic scale $1<R<1/q_{\min}$. In our case $q_{\min}=0.003\ \text{Å}^{-1}$.

The quantities $L_l=[(3/2)\sigma_\Psi(T)N_0]^{-1}$ and $L_s=[\sigma^m_{1-\Psi}(T)N_0]^{-1}$ are the neutron mean free path for long- and short-scale inhomogeneities, respectively. The ratios L/L_l and L/L_s represent the average number of neutron scattering events N_l and N_s per unit length of the sample, and are proportional to the neutron scattering cross sections of inhomogeneities of different scales.

Figure 8 displays the polarization $P(T)/P(T_p)$ and beam attenuation $I(T)/I(T_p)$ versus temperature, and Fig. 9 displays N_l and N_s for alloys I and II. It is clear from the figures that the dependence of N_l on $T\in[T_c,T_0]$ is different from that of N_s. Furthermore, over a wide temperature range N_l is large compared with N_s or, in other words, the mean free path in the case of scattering by long-scale correlations is much shorter than in the case of scattering by conventional critical fluctuations. We also note that for $N=1$ a kink is observed in the function $N_l(T)$. This kink characterizes the transition from single to multiple scattering. It is obvious that polarization for $N>1$ can no longer characterize the cross section σ_Ψ given by the expression (2).

Analysis of the scattering intensity as a function of $q\in[3\times10^{-3},3\times10^{-2}]$ Å shows that the spin correlation function corresponding to this range of q is described by the usual OZ expression for critical correlations. In the case of critical fluctuations, the total cross section $\sigma_{\theta_1,\theta_2}$ in the angular range $\theta_1-\theta_2$ $(\theta_2>\theta_1)$ can be obtained by integrating the expression (1) over the corresponding range of angles. The cross section depends on κ logarithmically:

$$\sigma_{\theta_1,\theta_2}=\frac{2}{3}\,\pi r_0^2\gamma^2 S(S+1)\,\frac{1}{(kr)^2}\,\ln\frac{\theta_2^2+(\kappa/k)^2}{\theta_1^2+(\kappa/k)^2}. \qquad (15)$$

Here the same notation as in Eq. (1) is used; k is the wave vector and κ is the reciprocal of the correlation radius.

Then the total scattering cross section in a small angle Ψ within the width of the transmitted beam has the form (3),

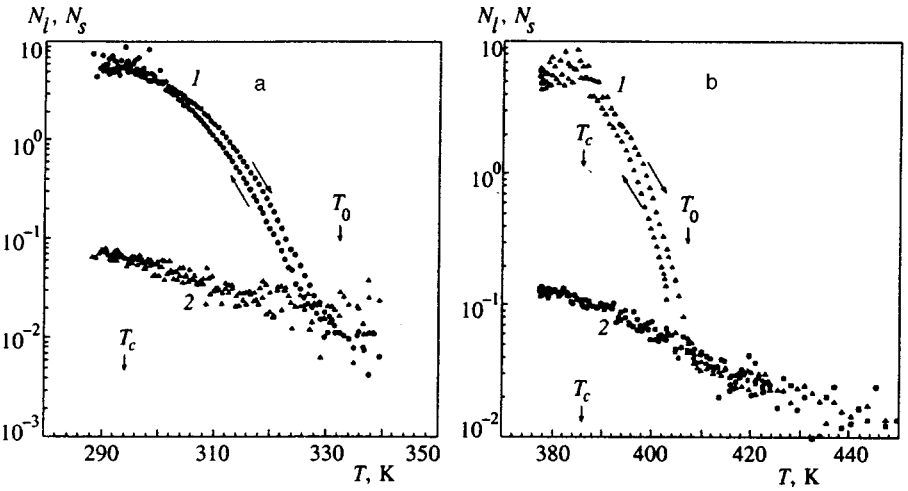


FIG. 9. Reciprocals N_l $(q<0.003\ \text{Å}^{-1})$ *(1)* and N_s $(q>0.003\ \text{Å}^{-1})$ *(2)* of the neutron mean free path versus temperature T for alloys I (a) and II (b).

and the total scattering cross section over the range of angles $\Psi<\theta<1$ is

$$\sigma^m_{1-\Psi}=\frac{2}{3}\,\pi r_0^2\gamma^2 S(S+1)\,\frac{1}{(kr)^2}\,\ln\frac{1+(\kappa/k)^2}{\Psi^2+(\kappa/k)^2}$$

$$\propto \ln\frac{(k/\kappa)^2}{1+(k\Psi/\kappa)^2}. \qquad (16)$$

Since $k\Psi/\kappa\sim 1$ and $k/\kappa\gg 1$, the relation $\sigma^m_{1-\Psi}\gg\sigma_\Psi$ should hold for cross sections associated with critical correlations. As shown in Ref. 10, the cross section σ_Ψ is small, and has little effect on the polarization at essentially any temperature $T>T_c$. However, the values N_s and N_l obtained experimentally above, which are proportional to the cross sections $\sigma^m_{1-\Psi}$ and σ_Ψ, show that $\sigma^m_{1-\Psi}\ll\sigma_\Psi$ (Fig. 9).

The enormous value and temperature dependence of the scattering cross section for small angles Ψ suggest that large magnetic inhomogeneities with sizes $R_d\gg R_c$ are present in the sample. Since the observed inhomogeneities are capable of strongly depolarizing the neutron beam, the depolarization indicates that the induction decreases slowly over their characteristic scale (Fig. 1b). On the other hand, it is clear that the cross section obtained from the beam attenuation is determined by the scattering by critical fluctuations, as confirmed by the fact that the cross section grows very slowly with decreasing temperature.

The depolarization data can thus be explained by the existence in the present system of a second length scale $R_d(T)$ larger than the characteristic size $R_c(T)$ of critical fluctuations at temperatures $T\geqslant T_c$. At the same time, as was found in Sec. 2, the characteristic forms of the inhomogeneities associated with different length scales differ from one another. Inhomogeneities with characteristic size R_c have the form in real space that is typical of critical fluctuations, $(1/r)\exp(r/R_c)$. The form of inhomogeneities with the other scale $R_d(T)$ should correspond to a more uniform distribution of magnetic induction in them.[10,11] The mechanisms leading to the emergence of these inhomogeneities are probably varied.

8. DISCUSSION

In the present work we investigated iron–nickel alloys with the composition $Fe_{70}Ni_{30}$ with admixtures of carbon (0.1 and 0.7 at. %). For these alloys, just as for the alloy ($Fe_{75}Ni_{25}$ with addition of 0.7 at. % carbon) that we investigated previously,[17,18] it was established that two characteristic length scales of magnetic inhomogeneities coexist above T_c. This characterizes the ''two length scale'' situation above T_c as a general situation for iron-enriched FeNi alloys.

A similar experiment, indicating the existence of strong depolarization above the Curie temperature, was performed on the alloy $Fe_{65}Ni_{35}$ and discussed in Ref. 28. However, the authors conclude from the fact that depolarization of the neutron beam is present far above the Curie temperature (T_c+100 K), that depolarization is due to critical fluctuations of the magnetization. Their conclusion is based on the experimentally established fact that above T_c, depolarization depends on the neutron wavelength, and below T_c it is independent of the neutron wavelength. As a result, in their opinion, neutron depolarization above and below T_c should be determined by different mechanisms of the interaction of neutrons with the magnetic system of the sample.

We cannot agree with this interpretation of the data, since, first, as noted above, with reference to the work Maleev and Ruban,[13,14] critical fluctuations near T_c weakly depolarize the neutron beam, and in order for strong depolarization to occur, large magnetic inhomogeneities with sizes $R_d>R_c$ and a comparatively uniform distribution of the induction within their characteristic scale must exist in the sample. Second, the presence or absence of a λ dependence in the polarization of the transmitted beam is determined by the satisfaction or nonsatisfaction, respectively, of the condition (4). As the temperature decreases below T_c, both the induction B and the size R_d of the inhomogeneities increase and the condition (4) ceases to be satisfied, as a result of which the λ dependence in the polarization vanishes. In consequence of this, such a difference cannot be explained by a difference in the mechanisms of interaction of neutrons with the magnetic system, and it simply indicates an increase in the induction B and size R_d of inhomogeneities in the sample. In conclusion, we emphasize once again that the observation of depolarization can be explained by large-scale magnetic inhomogeneities, which exist in the sample together with the conventional critical fluctuations at temperatures $T>T_c$.

These investigations have established the Curie temperature T_c, defined as that point on the temperature scale at which the correlation radius of the critical fluctuations assumes its maximum value. It was shown that the correlation radius does not become infinite as T_c is approached, but instead becomes constant at $T_c^{\exp}$. In our opinion, this demonstrates a disruption of the second-order phase transition at this temperature when the value $R_c(T_c^{\exp})=R_c^{\max}$ typical of the system is reached, and a transformation of the transition into a first-order transition that is close to a second-transition.

In the course of the investigations, temperature hysteresis was also observed in the measured values of the polarization of the transmitted beam and the small-angle neutron scattering intensity at temperatures $T>T_c$. As a result of the existence of hysteresis phenomena, the real Curie temperature depends on the sample history, and is established to within the width of the hysteresis (2–3 K). It was found that two characteristic length scales of magnetic inhomogeneities coexist over a wide temperature range $T>T_c$. It was also established that the forms of the spin correlation functions are different for the two scales. The short-scale correlation characterizing critical fluctuations can be described by the Ornstein–Zernike expression, while the long-scale correlations can be tentatively described by a squared Ornstein–Zernike expression. These features confirm the character of the magnetic phase transition in FeNi Invar alloys as a first-order transition. As far as we know, this is the first experimental confirmation of the theoretical predictions of the existence of a first-order phase transition in this system.[29–31]

We thank V. G. Gavrilyuk and A. L. Sozinov (Institute of Metal Physics, Ukrainian Academy of Sciences) for pro-

viding samples for the investigations, M. K. Runova (St. Petersburg Institute of Nuclear Physics) for assisting in the data analysis, A. L. Malyshev (St. Petersburg Institute of Nuclear Physics) for performing the neutron diffraction measurements, G. P. Kopitsa, O. V. Radionov (St. Petersburg Institute of Nuclear Physics), and S. V. Kozhevnikov and Yu. V. Nikitenko (Laboratory of Neutron Physics, Joint Institute of Nuclear Research) for assisting in the measurements.

This work was performed at the VVR-M reactor with financial support from the Ministry of Science of the Russian Federation (Reg. N01-48). The work was supported by the Russian Fund for Fundamental Research (Project L-EN-96-15-96775) and the State Science and Technology Program ''Neutron Investigations of Condensed Media.''

[a)]e-mail: grigor@rvv.lnpi.spb.su

[1] S. R. Andrews, J. Phys. C **19**, 3721 (1986).
[2] G. Shirane, R. A. Cowley, M. Matsuda, and S. M. Shapiro, Phys. Rev. B **48**, 15595 (1993).
[3] T. R. Tharston, G. Helgensen, D. Gibbs, J. P. Hill, B. D. Gaulin, and G. Shirane, Phys. Rev. Lett. **70**, 3151 (1993).
[4] T. W. Ryan, R. J. Nelmes, K. A. Cowley, and A. Gibaud, Phys. Rev. Lett. **56**, 2704 (1986).
[5] U. J. Nicholls and R. A. Cowley, J. Phys. C **20**, 3417 (1987).
[6] K. Hirota, G. Shirane, P. M. Gehring, and C. F. Majkrzak, Phys. Rev. B **49**, 11967 (1994).
[7] N. Sato, N. Aso, K. Hirota *et al.*, Phys. Rev. B **53**, 14043 (1996).
[8] M. Altarelli, M. D. Nunez-Regueiro, and M. Papular, Phys. Rev. Lett. **74**, 3840 (1995).
[9] R. A. Cowley, Phys. Scr. **T66**, 24 (1996).
[10] S. V. Maleev and V. A. Ruban, Zh. Éksp. Teor. Fiz. **62**, 415 (1972) [Sov. Phys. JETP **35**, 222 (1972)].
[11] S. V. Maleyev, J. Phys. C **7**, 23 (1982).
[12] G. M. Drabkin, E. I. Zabilarov, Ya. A. Kasman, and A. I. Okorokov, Zh. Éksp. Teor. Fiz. **56**, 478 (1969) [Sov. Phys. JETP **29**, 261 (1969)].
[13] A. I. Okorokov, V. V. Runov, A. G. Gukasov, and G. M. Drabkin, Izv. Akad. Nauk SSSR, Ser. Fiz. **42**, 1770 (1978).
[14] S. V. Grigor'ev, V. V. Runov, and A. I. Okorokov, Preprint No. 2112 [in Russian], St. Petersburg Institute of Nuclear Physics (1996).
[15] S. V. Grigoriev, S. V. Maleyev, V. V. Runov, and A. I. Okorokov, Physica B (1997), in press.
[16] V. L. Sedov, *Antiferromagnetism of Gamma Iron. The Invar Problem* [in Russian], Nauka, Moscow (1987).
[17] M. F. Collins, V. J. Minkievicz, R. R. Nathans, L. Passal, and G. Shirane, Phys. Rev. **179**, 417 (1969).
[18] V. J. Minkievicz, M. F. Collins, R. Nathans, L. Passal, and G. Shirane, Phys. Rev. **182**, 624 (1969).
[19] A. Z. Patashinskiĭ and V. L. Pokrovskiĭ, *Fluctuation Theory of Phase Transitions*, 1st ed., Pergamon Press, New York (1979) [Russian original, Nauka, Moscow (1982)].
[20] O. Halperin and T. Holstein, Phys. Rev. **59**, 960 (1941).
[21] R. Rosman and M. Th. Rekveldt, Phys. Rev. B **43**, 8437 (1991).
[22] S. V. Grigor'ev, V. V. Runov, A. I. Okorokov, A. D. Tret'yakov, O. A. Gubin, and G. P. Kopitsa, Preprint No. 2028, St. Petersburg Institute of Nuclear Physics (1995).
[23] O. K. Antson, A. P. Bulkin, P. E. Hiismaki *et al.*, Physica B **156–157**, 567 (1989).
[24] V. B. Zlokazov and V. V. Chernyshev, J. Appl. Crystallogr. **25**, 447 (1992).
[25] K. N. Trueblood and J. D. Dunits, Acta Cryst. B **39**, 120 (1983).
[26] S. V. Grigoriev, J. Magn. Magn. Mat. (1997), in press.
[27] K. P. Belov, *Magnetic Transitions*, Consultants Bureau, New York (1961) [Russian original, Fizmatgiz, Moscow (1959)].
[28] S. Mitsuda and Y. Endoh, J. Phys. Soc. Jpn. **54**, 1570 (1985).
[29] A. Z. Menshikov, Physica B **161**, 1 (1989).
[30] M. Shimizu and S. Hirooka, Phys. Lett. A **28**, 530 (1968).
[31] M. Shimizu, J. Magn. Magn. Mater. **10**, 231 (1979).

Translated by M. E. Alferieff

Effect of electroelastic anisotropy of DNA-like molecules on their tertiary structure

V. L. Golo

Mechanico-Mathematical Department of M. V. Lomonosov Moscow State University, 119899 Moscow, Russia

Yu. M. Evdokimov

Institute of Molecular Biology, Russian Academy of Sciences, 117984 Moscow, Russia

E. I. Kats

L. D. Landau Institute of Theoretical Physics, Russian Academy of Sciences, 117940 Moscow, Russia

(Submitted 26 May 1997)
Zh. Éksp. Teor. Fiz. **112**, 2156–2168 (December 1997)

Under certain conditions, mechanical forces can cause an anisotropic molecule like DNA to assume a toroidal spatial structure. A simple model describing such a behavior is suggested. The model incorporates anisotropic elastic energy and external electrical forces. The steady-state structures formed by a macromolecule have been studied numerically using this model. There exist ranges of model parameters, namely, the anisotropy of the elastic tensor, magnitude and orientation of forces, and modulation periods, where molecules have toroidal, spherical, or extended structures. Estimates of parameters characteristic of these structures are consistent with experimental data. In particular, the toroidal structure dimension corresponds to experimental dimensions of toroidal globules produced as a result of so-called PSI condensation of DNA molecules. © *1997 American Institute of Physics.* [S1063-7761(97)01612-0]

1. INTRODUCTION

It is known that, depending on external conditions, isolated two-chain DNA macromolecules can be condensed in fairly compact globules of various shapes, densities, and structures (see, for example, Refs. 1, 2, and references therein). Globular structures in DNA were detected about twenty-five years ago,[3] and their toroidal configuration was determined.[4] Later both experimental[5,6] and theoretical investigations of processes shaping toroidal particles were published. These include, for example, studies based on macroscopic concepts[7,8] originating from the work of I. M. Lifshits[9] concerning the coil-to-globule transition and studies based on the microscopic description of DNA conformations.[10,11] Investigations of these processes have continued to the present time.[12,13] This fact alone indicates that some questions in this field have not been completely answered.

As a result of these studies, it has become clear that there are many factors affecting the globule structure. Usually formation of toroidal structures is induced by adding a neutral polymer and a mono- or polyvalent salt to an aqueous solution of DNA, from which derives the term applied to this phenomenon, namely PSI (Polymer- and Salt-Induced) condensation.

According to the established concept of PSI condensation,[9,12] the neutral polymer generates osmotic pressure, which makes the spatial molecular configuration more compact, and the salt partly neutralizes negative charges of DNA phosphate groups. In addition to these main factors, an essential role in the PSI condensation is played by intrinsic DNA parameters (contour and persistent lengths), intermonomer interaction, etc. It is hardly possible or advisable to take into account all these factors in an exact form, since many physical parameters that are essential for description of these effects are unknown. Therefore the existing theories of the PSI condensation which is responsible for formation of toroidal structures, are approximate to a greater or lesser extent. Some treatments[8,13] ignore electrostatic forces, while others[12] neglect effects of excluded volume. On the other hand, it turns out that, irrespective of the approximations used in these theories, DNA structures predicted by different models have similar qualitative characteristics, and they are described by similar functions of the macromolecule parameters.

Therefore it seems probable that some underlying causes lead to the formation of toroidal structures independent of specific approximations in the models describing them. This assumption is corroborated by investigations[9,13] based on a very general macroscopic approach and demonstrating that minimization of both bulk and surface free energies expressed with due account of well known scaling relations for hard polymers[14] yields results for toroidal configurations that are in qualitative agreement with experimental data.

A microscopic approach to conformations of DNA molecules, which is based, in a sense, on quite opposite assumptions, was developed by Manning.[10,11,15–17] He considered a DNA molecule as an elastic rod extended by electrostatic repulsive forces acting between the negative charges of the phosphate groups. These forces are balanced by elastic forces compressing the rod. When phosphate charges are partly neutralized by oppositely charged salt ions,[18] compressing forces can lead under certain conditions to the well known buckling instability of a straight rod.[19] Thus we obtain the instability criterion for a straight molecule conformation due to axial compressive forces:

$$4\xi_b \frac{T}{F_a} < \pi^2 R^4.$$

Here ξ_b is the bending persistent length, F_a is the axial compressing force (numerically equal to the electrostatic extending force), T is the temperature, and R is the DNA molecule radius. The buckling instability should lead to a toroidal molecule in this model.

Strictly speaking, the macroscopic polymer approach applies only on a macroscopic scale, i.e., on a scale larger than the persistent length, which is about 500 Å. The microscopic approach can be used on scales smaller than this length. On the other hand, the characteristic dimensions of the experimentally detected globules fall between these two ranges, i.e., they are comparable to the persistent length. However, these are the scales (they can be called mesoscopic scales) on which the elastic energy of DNA molecules plays an important role. Electrical forces also should not be ignored on these scales (as will be shown below, elastic and electric forces are comparable). Moreover, the condensation of oppositely charged ions on an anisotropic and nonuniformly charged DNA molecule should also be nonuniform and anisotropic.

The aim of this work was to investigate a ''minimal'' model of DNA-like molecules taking into account only the elastic and nonuniform electric forces, and to analyze the spatial structures which appear in this model. Here we stress once again that in our model the anisotropic structure and properties of the DNA molecule manifest themselves both in the anisotropy of elastic moduli (i.e., $a_{ij} \neq \delta_{ij}$) and in the existence of a nonuniform electroelastic modulus. If the system is treated as isotropic, both these effects vanish, and toroidal structures cannot form.

2. MODEL DESCRIPTION

Following the scheme described in Refs. 20 and 21, we consider the axial line of a DNA molecule as an elastic string or rod. As was noted above, the conformation of this string is largely determined on intermediate (mesoscopic) scales by its elastic energy. The latter can be expanded, as usual, in terms of the strain tensor. This expansion is conveniently expressed in the form

$$E_0 = \int_0^L ds \left(\frac{1}{2} a_{ij} \omega_i \omega_j + b_i \omega_i \right), \tag{1}$$

which is a generalization of the classical Kirchhoff problem of the equilibrium of an elastic rod.[22]

In Eq. (1) L is the elastic string length, s is the coordinate along the axial line, a_{ik} is the matrix of elastic moduli (the anisotropy in a_{ik} reflects the existence of two nonidentical helices on the microscopic scale), and the vector **b** describes spontaneous deformation of a DNA molecule resulting in formation of the superhelical structure. The physical cause of spontaneous deformation can be, for example, adsorption of a DNA molecule on nucleosomes. In this paper, we are interested only in PSI condensed structures, so we neglect the spontaneous moment **b** due to adsorption.

The occurrence in (1) of the vector $\boldsymbol{\omega}$ needs explanation. In order to describe the conformation of the DNA axial line, we should introduce a Frenet local reference frame $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, where the vector $\mathbf{v}_1$ is tangent to the axial line, and the vectors $\mathbf{v}_2$ and $\mathbf{v}_3$ are aligned with the main axes of the rod deformation,[19] i.e.,

$$\mathbf{v}_1 = \frac{d\mathbf{r}}{ds}, \quad \mathbf{v}_i \cdot \mathbf{v}_j = \delta_{ij}. \tag{2}$$

The vector $\boldsymbol{\omega}$ in expression (1) for the energy describes changes in the Frenet local reference parallel to the string:

$$\frac{d}{ds}\mathbf{v}_j = [\boldsymbol{\omega} \times \mathbf{v}_j]. \tag{3}$$

Equation (3) means that all permissible deformations of the curve are described by rotations of the Frenet reference frame, i.e., the elastic string is not extensible (which is a fairly good approximation for DNA molecules, whose torsion and bending moduli are several orders of magnitude smaller than the stretching modulus).

For the problems of interest, the presence of internal anisotropic elasticity of the molecule, i.e., the fact that $a_{ij} \neq \delta_{ij}$, is essential. This anisotropy of the elastic moduli is introduced in a natural way, since in replacing the double helix with a single elastic string (axial line) we must take into account the anisotropic DNA structure. In terms of mechanics, polynucleotide chains forming the double helix can be modeled by helical grooves on the elastic rod discussed above, one of these grooves being narrow and another being wide. This fact leads to a consequence that is very important for the PSI condensation. Let us recall that a DNA molecule is charged and carries a charge of $-2e$ (where e is the elementary charge) per pair of bases (i.e., for a contour length of 3.4 Å). The negative DNA charge is partly neutralized by condensation of oppositely charged ions in the solution. The problem of a charged string (polyion) in a solution containing oppositely charged ions has been treated by many authors.[16,18,23] They demonstrated that, when the polyion charge exceeds a certain critical value, oppositely charged ions ''condense'' on the polyion and reduce its charge density to the critical value.

The distinguishing feature of our model is, as was noted above, the inherent anisotropy of the macromolecule or polyion due to different charge densities in the narrow and wide grooves.[24,25] This means, in particular, that neutralization of the molecular charge by oppositely charged ions is nonuniform and anisotropic over the molecule length, for example, owing to the inhomogeneity of the field generated by negative charges of phosphate groups in the narrow or wide groove.

An exact solution to the appropriate electrostatic problem would obviously be beyond the accuracy of our simplified model of DNA molecules composed of two chains. In our opinion, however, the only important effect is the presence of an inhomogeneous field generating a nonuniform torque

$$\mathbf{M}_{ee} = [\mathbf{F} \times \mathbf{v}_1]. \tag{4}$$

Here **F** is the force due to partial neutralization of the phosphate group negative charge. Let us reiterate that in our model this force is nonuniform along the molecule axis.

The electroelastic moment given by Eq. (4) affects the molecule conformation defined by the vector $\boldsymbol{\omega}$. The vector **F** is related in a complicated way to the adsorption of oppositely charged ions and the molecule conformation. Generally speaking, the problem of macromolecule conformation should be solved with due account of the osmotic pressure produced by the solvent, the excluded DNA volume, and the conformation entropy (the latter two should be considered in a self-consistent manner, like the force **F**). It is clear that a problem with such a general statement can hardly be solved. In the next section, we will show that a considerably simplified model incorporating only the macromolecule elasticity and nonuniform electroelastic moment yields a qualitatively adequate description of most spatial structures formed by DNA molecules. Moreover, under some fairly natural assumptions concerning the values of parameters in our model, reasonable quantitative agreement is also obtained.

3. RESULTS

We consider the DNA molecule as an anisotropic elastic rod acted upon by an electro-elastic moment $\mathbf{M}_{ee}$ that varies over the rod axis. In this section, our aim is to determine all possible steady-state conformations of such a rod. The solution is obtained in two stages. On the first stage we determine purely mechanical characteristics of the rod. The equilibrium condition in the laboratory frame is expressed in the well-known form:[19]

$$\frac{d\mathbf{M}_u}{ds}=[\mathbf{F}_u\times(\mathbf{v}_1)_u], \tag{5}$$

where $\mathbf{M}_u$ is the torsional moment due to the force **F**. Here the subscript u distinguishes physical quantities measured in the fixed (laboratory) reference frame.

It is convenient to convert this equilibrium condition to the moving (local) reference frame characterized by the Frenet coordinate system introduced above. To this end, the operator $\hat{R}$ locally transforming the laboratory reference frame to the moving one should be applied to Eq. (5). As a result, we have the following equation in the local reference frame:

$$\frac{d\mathbf{M}}{ds}=-[\boldsymbol{\omega}\times\mathbf{M}]+[\mathbf{F}\times\mathbf{v}_1], \tag{6}$$

where the vector tangent to the axial line in the moving reference frame is, by definition, $\mathbf{v}_1\equiv(1,0,0)$. By solving Eq. (6) we obtain $\boldsymbol{\omega}$ (recall that, by definition,[19] $\mathbf{M}=\delta E/\delta\boldsymbol{\omega}$).

Given $\boldsymbol{\omega}$, we reconstruct the tangent vector in the laboratory frame:

$$\frac{d(\mathbf{v}_1)_u}{ds}=[\boldsymbol{\omega}\times(\mathbf{v}_1)_u]. \tag{7}$$

This is the first stage of the solution, which reduces to the calculation of the vector field $\mathbf{v}_1$. The second stage is determination of the steady-state configuration $\mathbf{r}(s)$ based on the Eq. (2). This procedure cannot be performed analytically, so we have solved the equation system (2), (6), and (7) numerically. We have employed the fourth-order Adams scheme. The calculation accuracy depended on the specific configuration under investigation, but even in the worst case the relative error was within 10^{-3}, which seems reasonable for a qualitative assessment of configurations.

In solving the problem numerically, one should first transform it to dimensionless variables. The variables in Eqs. (2), (6), and (7) have the following dimensionalities:

$$[\omega]=\frac{1}{\text{cm}},\quad [a_{ij}]=\text{erg}\cdot\text{cm},\quad [F]=\frac{\text{erg}}{\text{cm}}.$$

It seems convenient to introduce a characteristic energy ε and length l. They can be estimated as follows. The many DNA elastic moduli are known to be of order 10^{-19} erg·cm.[1,17] In our numerical analysis, we assumed that the numerical factors in front of this parameters are of the order of unity, hence

$$\varepsilon l\simeq 10^{-19}\ \text{erg}\cdot\text{cm}. \tag{8}$$

Unfortunately, to the best of our knowledge, no experimental measurements of the electroelastic moment $\mathbf{M}_{ee}$ (or the corresponding force **F**) are available. Therefore we have to use approximate estimates of the Coulomb energy. The latter essentially depends on the unneutralized charge of DNA phosphate groups. If this charge is assumed to be approximately the elementary charge on a characteristic scale of about the DNA molecule diameter ($d\simeq 10^{-7}$ cm), the estimate for the force F is 10^{-5} dyn. If the degree of neutralizations is considerable (the charge per characteristic length is $0.1e$), the force is about 10^{-7} dyn. Note that the threshold of the buckling instability calculated by Manning[15] corresponds to a force that is one or two orders of magnitude smaller. However, Manning[15] considered the onset of instability, whereas we are discussing conditions for formation of a toroidal structure. In addition, the compressive axial force is only important for the buckling instability and the torsional moment appears later owing to the bending of the elastic rod. In the case under discussion, a torsional moment appears even in an initially straight configuration.

Assuming that electric forces are of the same order as elastic forces, i.e., that the dimensionless electroelastic moment, like the dimensionless elastic moduli, is of order unity, we can determine the ratio between the characteristic parameters ε and l:

$$\frac{l}{\varepsilon}\simeq 10^{7}\frac{\text{cm}}{\text{erg}}. \tag{9}$$

Thus, we can derive from Eqs. (8) and (9) the two characteristic parameters:

$$\varepsilon\simeq 10^{-13}\ \text{erg},\quad l\simeq 10^{-6}\ \text{cm}.$$

Note that the characteristic energy needed for formation of a compact three-dimensional structure is one order of magnitude higher than the heat energy (which is about 10^{-14} erg), and this justifies our mechanical model, which neglects thermal fluctuations. The length l acts in our model as an effective persistent length, over which Coulomb forces are comparable to elastic ones. In accordance with the above

reasoning concerning the value of ε, the effective persistent length (as expected) is an order of magnitude smaller than the thermal persistent length. In our opinion, these two facts are quite natural, since the compact DNA configuration arose, for example, under conditions of PSI condensation by osmotic pressure of about 10^5 dyn/cm^2,[1,12] and the respective forces are stronger than thermal fluctuations.

Before proceeding to the results of numerical calculations based on Eqs. (2), (6), and (7), there is another important note. As follows from the expression for $\mathbf{M}_{ee}$ and definition of the energy density, there is no bulk contribution to the energy if $\mathbf{F}=\text{const}$. Therefore, any nontrivial effects due to the electroelastic moment can be described only by taking into account the function $\mathbf{F}(s)$. Given that the double helix is a three-dimensional periodic structure and considering only the most important features of conformations, we analyzed numerically only simple periodic functions $\mathbf{F}(s)$.

The force $\mathbf{F}(s)$ is a vector, so its amplitude and direction in the local reference can both be modulated, while the local reference is determined by the molecule conformation, which, in turn, depends on the force. In our numerical analysis, we included both types of modulation, and it turned out that the results essentially depend on both the magnitude and direction of the force.

Specifically, we specified in our numerical solution the force $\mathbf{F}$ in the form

$$\mathbf{F}=A\sin(s/k_a)\hat{R}(\mathbf{n},s/k_o)\mathbf{f},$$

where A is the force amplitude, k_a is the period of the amplitude modulation, and $\hat{R}(\mathbf{n},s/k_o)\mathbf{f}$ is the operator rotating the unit vector $\mathbf{f}$ aligned with the force with respect to the local normal $\mathbf{n}$ to the curve, and k_o is the modulation period of the force orientation. Since the model described above implies nonuniform (i.e., depending on the coordinate along the axial line of the molecule) and anisotropic (i.e., depending on the alignment of the molecule section in space) charge neutralization in the narrow groove of the DNA molecule, it is natural to assume that the period k_o corresponds to the helix pitch, i.e., its period is ten times the distance between base pairs (34 Å), and this value ($k_o=0.34$ in our dimensionless units) was used in calculations, unless stated otherwise.

Numerical solutions of the system of equations (2), (6), and (7) are presented in the form of graphs that show steady-state configurations of the elastic string, and these configurations were compared with DNA shapes observed, for example, under conditions of PSI condensation.

Figure 1 shows a toroidal structure obtained with the following numerical parameters:

(1) the matrix of elastic moduli

$$\begin{pmatrix} 1.7 & 0 & 0 \\ 0 & 1.1 & 0 \\ 0 & 0 & 1.3 \end{pmatrix};$$

(2) force amplitude $A=5.6$;
(3) amplitude modulation period $k_a=0.17$;
(4) projection of the force unit vector on the local normal $f_3=0.01$;
(5) orientational modulation period $k_o=0.34$.

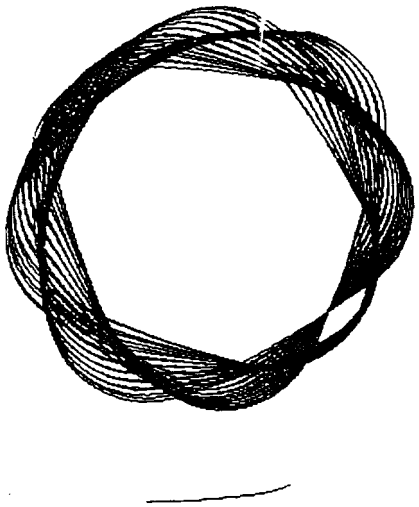


FIG. 1. Toroidal globule (the parameters are given in the text). At the bottom the effective persistent length is shown.

For convenience the effective persistent length l, which is approximately 1/5 of the overall structure dimension, is also shown in the graph.

It is interesting to compare Fig. 1 with an electron micrograph (Fig. 2) of a toroidal globule.[26]

According to some experimental data,[3,4,13] toroidal globules have typical radii of about 500 Å, so the above estimate $l\simeq 10^{-6}$ is in agreement with these data. Note also that the estimate of the characteristic energy needed to form a toroidal configuration, $\varepsilon\simeq 10^{-13}$ erg is also in agreement with characteristic energies of toroidal globules given in Ref. 27. This agreement can be considered as a consistency test of the suggested model.

The parameters listed above have considerable effect on the three-dimensional configuration formed by a macromolecule. The model parameters of the configuration shown in Fig. 3 are the same as in Fig. 1, except the matrix of elastic moduli, which is modified as follows:

$$\begin{pmatrix} 1.7 & 0 & 0 \\ 0 & 1.1 & 0.2 \\ 0 & 0.2 & 1.05 \end{pmatrix}.$$

The resulting configuration is spherical. Note that this change in the configuration is due to a small change in the anisotropy of the elastic tensor (we refer not to the orders of magnitude but the values because our results belong to the parameter range specified, as was mentioned above, by the scales $\varepsilon\simeq 10^{-13}$ erg and $l\simeq 10^{-6}$ cm).

The force alignment is also essential. For example, if we take $f_3=0.03$ instead of $f_3=0.01$, which is the case in the structures shown by Figs. 2 and 3, a particle shaped as a torus becomes unstable against a transformation to an extended structure. This is illustrated by Fig. 4 (in order to obtain a more compact structure, we used in this case the unit matrix of elastic constants).

There are also certain limitations on the permissible contour length L of a macromolecule. Figures 1, 3, and 4 correspond to $L=300$. It is clear from the geometrical viewpoint

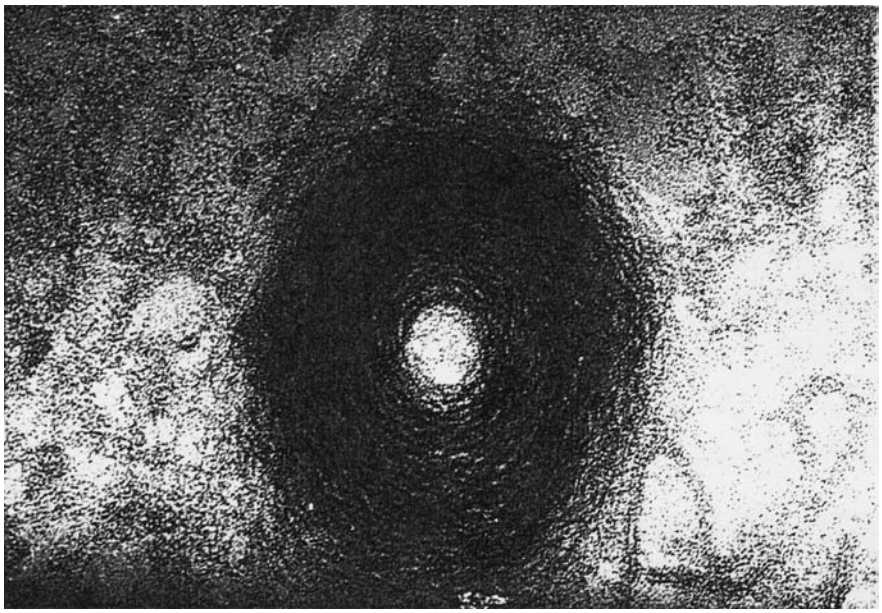


FIG. 2. Electron micrograph of a toroidal globule.[26]

that formation of toroidal structures is impossible at very small L, i.e., for the specific parameters used in our calculations, for $L<30$. If the contour length, however, is too large, toroidal configurations are also unstable. Figure 5a shows as an illustration a configuration corresponding to $L=3000$. It was obtained for the following values of model parameters:

(1) the matrix of elastic constants

$$\begin{pmatrix} 1.4 & 0 & 0 \\ 0 & 1.1 & 0 \\ 0 & 0 & 0.9 \end{pmatrix};$$

(2) force amplitude $A=22.4$;
(3) amplitude of the modulation period $k_a=0.034$;
(4) projection of the force unit vector on the local normal, $f_3=0.01$;
(5) orientational modulation period $k_o=0.34$.

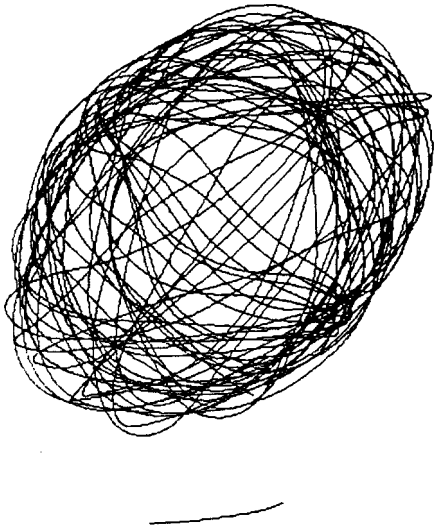


FIG. 3. Spherical globule (the parameters are given in the text).

Figure 5b shows a so-called Poincaré section of the resulting configuration. It is plotted as follows. An arbitrary point is selected on the trajectory (in this specific case, $s=0.07$), and a plane normal to the tangent vector at this point is defined. Then all sequential crossing points of the trajectory with this plane are marked. It is clearly seen in Fig. 5b that the resulting shell is spherical. Note that similar structures were also predicted by the macroscopic approach.[9] Such changes in configurations of three-dimensional structures with the molecule length are in agreement with experimental data, and they seem quite natural from the intuitive viewpoint regarding the behavior of anisotropic elastic rods acted upon by nonuniform external forces.

4. CONCLUSIONS

Let us summarize our results. We have demonstrated that there are purely mechanical forces that can shape an anisotropic DNA molecule in a toroidal three-dimensional structure. In order to give an adequate description of these mechanical forces, we have discussed a simple model incorporating Kirchhoff's anisotropic energy[22] (see also Ref. 19) and external electric forces. Actually, this model applies to intermediate (mesoscopic) length scales, where the configuration is determined by the mechanical energy. The full description of the system should, of course, also include microscopic aspects[15] and macroscopic polymer effects.[9,13] Our

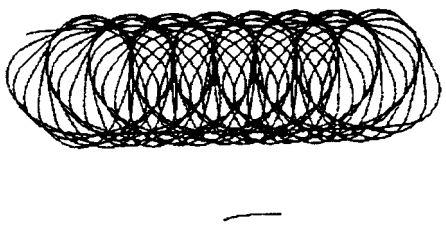


FIG. 4. Extended globule.

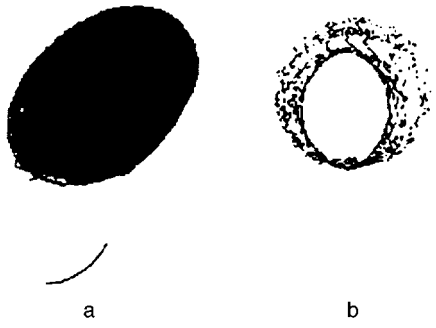


FIG. 5. (a) Spherical shell. (b) Poincaré section of the spherical shell.

estimates and qualitative agreement between numerical calculations by the mechanical model and experimental data indicate, however, that at least some features of PSI condensation are adequately described by our model.

In this connection, we would like to note the following. The issue of so-called kinks or sharp bends in structures has been discussed in the literature.[15] Direct observations of such kinks have also been reported.[28,29] Some authors, however, suppose that there are no kinks (see the discussion in Ref. 1). In our model, we also obtain configurations with sections looking like kinks (although no kinks are possible in elasticity theory). One of such configurations is shown in Fig. 6 as an illustration. The model parameters in this graph are the following:

(1) matrix of elastic constants

$$\begin{pmatrix} 1.1 & 0 & 0 \\ 0 & 1.1 & 0 \\ 0 & 0 & 1.1 \end{pmatrix};$$

(2) force amplitude $A=5.6$;
(3) amplitude modulation period $k_a=0.34$;
(4) projection of the force unit vector on the local normal, $f_3=0.01$;
(5) orientational modulation period $k_o=0.34$;
(6) $L=30$.

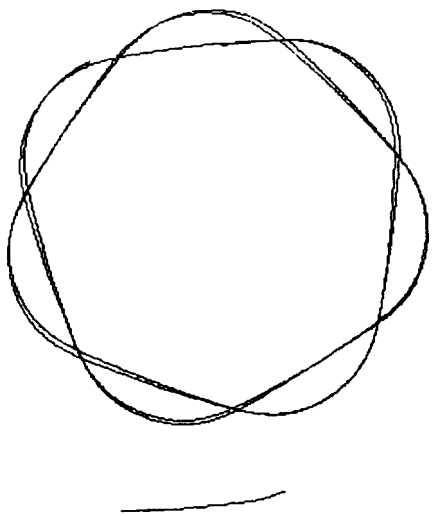


FIG. 6. Quasi-kink structure.

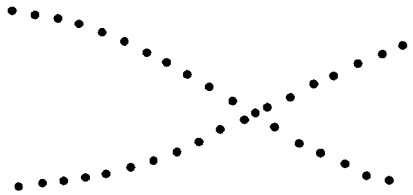


FIG. 7. Poincare section of a toroidal structure.

In the quasikink configuration obtained with these parameters, the quasikink angle is 70°, which is in agreement with experimental results.[28,29]

An important issue related to the biophysical functions of compact DNA structures (and promising applications of such objects as biological sensors) is the liquid-crystal ordering (in particular, the cholesteric one) of DNA segments in compact globules. Our simple model does not contain factors that would lead to such ordering (namely the steric or long-range interaction). Purely mechanical factors, however, can give rise to nematic ordering of neighboring macromolecule sections. As an illustration of this fact, in Fig. 7 we present a Poincaré section of a toroidal structure shown in Fig. 1. A certain type of ordering in the pattern of the crossing points is obvious, and it indicates that a sort of liquid-crystal ordering is present in the structure. Unfortunately, quantitative estimates of the order parameter and even its classification are impossible within our approach to the problem, moreover, attempts of such estimates are futile for the reasons discussed above.

We conclude from our analysis that there are five basic factors that control the macromolecule configuration, namely, the matrix of elastic constants a_{ij}, the amplitude A and alignment $\mathbf{f}$ of the Coulomb force, the periods of both the amplitude and orientational modulation, and the total contour length L. In our simple model, these parameters are controlled by various physical factors. For example, the force amplitude depends on the concentration and valence of salts in the solution, whereas the force modulation and anisotropy of elastic moduli can be regulated, in principle, by introducing (intercalating) specific chemical agents into the DNA structure.

It is feasible that toroidal configurations can be destroyed by introducing a sufficient quantity of dye molecules into the two DNA chains. Impurities or dopants that affect adsorption of charges due to salts in the solution (or even replacement of a monovalent salt with a polyvalent one) should change the force modulation period and essentially modify the formation conditions and shapes of DNA three-dimensional configurations.

This work is part of the Statistical Physics program sponsored by Ministry of Science and Technology of Russia. It was also partly supported by INTAS (Grant No. 94-40-78) and the Russian Fund for Fundamental Research.

[1] W. Saenger, *Principles of Nucleic Acid Structure*, Springer, New York (1984).
[2] A. Yu. Grosberg and A. R. Khokhlov, *Statistical Physics of Macromolecules*, AIP Press, New York (1994).
[3] L. S. Lerman, Proc. Natl. Acad. Sci. USA **68**, 1886 (1971).

[4] Y. M. Evdokimov, A. L. Platonov, A. S. Tikhonenko, and Y. M. Varshavsky, Fed. Europ. Biochem. Soc. Lett. **23**, 180 (1972); Yu. M. Evdokimov, N. M. Akimenko, N. E. Glukhova, A. S. Tikhonenko, and Ya. M. Varshavskii, Mol. Biol. **7**, 151 (1973).
[5] U. Laemmli, Proc. Natl. Acad. Sci. USA **72**, 4288 (1975).
[6] M. Aosule and J. A. Schellman, J. Mol. Biol. **121**, 311 (1978).
[7] A. Yu. Grosberg, Biofizika **24**, 32 (1979).
[8] A. Yu. Grosberg and A. V. Zhestkov, Biofizika **30**, 698 (1985).
[9] I. M. Lifshits, Zh. Éksp. Teor. Fiz. **55**, 2408 (1968) [Sov. Phys. JETP **28**, 1280 (1969)].
[10] G. S. Manning, Biopolymers **20**, 1261 (1981).
[11] G. S. Manning, Comments Mol. Cell. Biophys. **5**, 311 (1982).
[12] J. Ubbink and T. Odijk, Biophys. J. **68**, 54 (1995).
[13] V. V. Vasilevskaya, A. R. Khokhlov, S. Kidoaki, and K. Yoshikawa, Biopolymers **41**, 51 (1997).
[14] P. de Gennes, *Scaling Concepts in Polymer Physics*, Cornell University Press, Ithaca, USA (1979).
[15] G. S. Manning, in *Theoretical Chemistry and Molecular Biophysics*, ed. by D. L. Beveridge and R. Lavery, Adenine Press, New York (1990), p. 191.
[16] M. O. Fenley, G. S. Manning, and W. K. Olson, Biopolymers **30**, 1191 (1990).
[17] G. S. Manning, Biopolymers **31**, 1543 (1991).
[18] G. S. Manning, J. Biomolec. Structure and Dynamics **7**, 41 (1989).
[19] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity*, Pergamon Press, Oxford (1980).
[20] J. F. Marko and E. D. Siggia, Macromolecules **27**, 981 (1994).
[21] V. L. Golo and E. I. Kats, JETP Lett. **60**, 679 (1994).
[22] G. Kirchhoff, *Mechanik*, Tenbnir, Berlin (1897).
[23] M. Lebret, J. Chem. Phys. **76**, 6243 (1982).
[24] K. Zakrewska and B. Pullman, Nucl. Acids Research **11**, 8841 (1983).
[25] K. Zakrewska, R. Lavery, and B. Pullman, Nucl. Acids Research **11**, 8825 (1983).
[26] Yu. M. Evdokimov, T. L. Pyatigorskaya, O. F. Polyvtsev *et al.*, Nucl. Acids Research **3**, 2353 (1976).
[27] S. C. Riemer and V. A. Bloomfield, Biopolymers **17**, 785 (1978).
[28] T. J. Richmond, J. T. Finch, B. Rushton, D. Rhodes, and A. Klug, Nature (London) **311**, 532 (1984).
[29] K. K. Ebralidse, S. A. Grachev, and A. D. Mirzabekov, Nature (London) **331**, 365 (1988).

Translation provided by the Russian Editorial office.

Drift of vertical Bloch lines in a domain wall with a noninterchangeable spectrum in perpendicularly magnetized films

G. E. Khodenkov

Joint Commercial Magneto-optoelectronics Laboratory of the Institute of General Physics at the N. P. Ogaryov Mordovian State University, Russian Academy of Sciences, 430000 Saransk, Russia
(Submitted 24 October 1996)
Zh. Éksp. Teor. Fiz. **112**, 2169–2177 (December 1997)

The dynamics of vertical Bloch lines in variable external magnetic fields is examined with allowance for the magnetostatic noninterchangeability of the spectrum of the domain wall. The drift velocity of the translational motion of vertical Bloch lines is calculated and is found to be nonzero in second order in the weak oscillating field. © *1997 American Institute of Physics.* [S1063-7761(97)01712-5]

1. INTRODUCTION

When we speak of the drift of Bloch lines we mean their translational motion along a domain wall initiated by an oscillating external magnetic field of only one definite polarization. This effect was first detected in experiments and studied in in-plane magnetized yttrium-iron garnet films.[1–4] In addition to the drift of Bloch lines in the domain walls, magnetic materials of various types exhibit a drift of the domain walls proper. The presence of drift presupposes the existence of nonlinear mechanisms leading to the occurrence of even terms in the dependence of the velocities of domain walls or Bloch lines on the amplitude of the external field.

For domain walls in uniaxial ferromagnets this mechanism was studied, in particular, by Bar'yakhtar *et al.*[5] Here the drift of the domain wall is proportional to a quadratic combination of the magnetic fields of different polarizations. Such a drift can lead, at least in principle, to a drift of vertical Bloch lines, but here such mechanisms are not considered.

A general theory explaining the drift of Bloch lines was proposed by Iordanskiĭ and Marchenko.[6] For lines of different topological types with allowance for the general form of magnetostatic energy and the presence of a combination of uniaxial and cubic anisotropy, the frequency dependence of the drift has been derived in an approximation that is quadratic in the external fields of arbitrary polarizations. Laws that govern the transformation of the force leading to drift have also been established. Grishin and Martynovich[7] proposed another mechanism for the drift of vertical Bloch lines in perpendicularly magnetized films with magnetic-bubble domains (magnetic-bubble films). However, for the mechanism to become real the magnetic parameters of the film's material must be spatially inhomogeneous on both sides of the domain wall.

2. STATEMENT OF THE PROBLEM

In this paper we will study a particular mechanism for the drift of vertical Bloch lines in domain walls of uniaxial ferromagnets with a large anisotropy constant. The mechanism is based on the long-known noninterchangeability of the spectrum of magnons localized at domain walls.[8,9] In 1986, V. M. Chetverikov pointed out that such an effect can exist, and the corresponding static effects were also numerically studied by other researchers (see the relevant literature cited in Ref. 10, where the effect of magnetostatic noninterchangeability on the motion of vertical Bloch lines in constant fields was studied). The noninterchangeability of the spectrum modifies the dissipation processes in vertical Bloch lines in such a way that the asymmetry in the velocity of the motion of a Bloch line in the forward and reverse directions along a domain wall emerges already in the first nonlinear (quadratic) approximation in the external field strength. The usual approach to studying the translational dynamics of vertical Bloch lines in magnetic-bubble films is to employ sufficiently long unipolar magnetic-field pulses. Here, in contrast to Refs. 1–4, the possible noninterchangeability effect appears superposed on the large (larger than the effect) main contribution among the odd (in field strength) contributions.

The theory proposed in Ref. 10 referred to constant external fields. The goal of the present investigation is to derive the equations of motion of vertical Bloch lines in variable magnetic fields of a definite polarization, when the velocity asymmetry effect, due in this case to the noninterchangeability of the spectrum, is clearly visible in pure form, i.e., in the form of translational drift of the lines. We will limit ourselves to perpendicularly magnetized very thin films, where the twist of the domain wall is small and can be ignored (as can its effect on the vertical Bloch lines).

Suppose that we have a uniaxial ferromagnet whose easy-magnetization axis is collinear with the z axis and whose Q-factor is high, $Q=H_a/4\pi M\gg 1$, where H_a is the anisotropy field, and M is the absolute value of the magnetization vector. The unperturbed 180-degree domain wall in this ferromagnet lies in the xz plane. To describe the domain wall we use the Slonczewski equations for the variables $y=q(x,t)$ and $\psi(x,t)$, i.e., we use the equations describing the behavior of the domain-wall surface and the value of the azimuthal angle of the magnetization vector at this surface, with both being dependent on the x-coordinate (the x axis lies in the plane of the domain wall at right angles to the easy-magnetization axis) and time t:

$$\eta_D\dot{\psi}+\alpha\dot{q}-\eta_D h_z=q''-\kappa^2 q-\frac{(\sin 2\psi)'}{2\sqrt{Q}}, \tag{1a}$$

 1063-7761/97/121187-05$10.00

$$\eta_D\dot{q}-\alpha\dot{\psi}-h_x\sin\psi+h_y\cos\psi=-\psi''+\sin\psi\cos\psi-\frac{q'\cos 2\psi}{\sqrt{Q}}. \tag{1b}$$

Equations (1) are expressed in terms of the following dimensionless variables:

$$x\to\frac{x}{\Lambda},\quad q\to\frac{q}{\Lambda},\quad t\to t(4\pi\gamma M),$$

$$h_{x,y}\to\frac{h_{x,y}}{8M},\quad h_z\to\frac{h_z}{4\pi M}, \tag{2}$$

with the dots and primes on the variables designating derivatives with respect to t and x. The x-coordinate along the domain wall is measured in units of the width of the Bloch line, $\Lambda=\Delta/\sqrt{Q}$, where Δ is the width of the domain wall; the time variable t is measured in units of $1/4\pi\gamma M$, where γ is the magnetomechanical ratio; and the position of the domain wall $q(x,t)$ is measured in units of Δ. Weak external magnetic fields $\mathbf{h}(t)$ are measured in the units specified in (2), which are proportional to M. The equations contain the following small parameters: $\alpha<1$ is the Gilbert or viscous damping parameter, and $\kappa<1$ is the restoring force constant of the domain wall. The equations also contain a topological charge $\eta_D=\pm 1$, with the ''plus'' corresponding to a domain wall with the directions of magnetization in the domains $M_z(y\to\pm\infty)=\mp M$ and the ''minus,'' to magnetizations with opposite orientations.

The contribution to Eq. (1) of the ''noninterchangeable'' magnetostatic energy $\propto\sin^2(\psi-q'/\sqrt{Q})$ is expanded in powers of the small parameter $1/\sqrt{Q}$. Note that in the linear approximation Eqs. (1) lead to a noninterchangeable spectrum of the wall magnons that coincides with the exact spectrum.[9]

One of the main assumptions of the present work is that noninterchangeable effects are taken into account everywhere to within $1/\sqrt{Q}$. The other main assumptions are

$$|h_{x,y,z}|<1,\quad \alpha<1,\quad \omega<\kappa<1. \tag{3}$$

The importance of the smallness of the amplitudes of the external fields and the damping parameter has been pointed out earlier, and the third inequality in (3) presupposes that the frequencies ω of external fields (measured in units of $4\pi\gamma M$) are below the frequency $\omega_0=\kappa$ of the homogeneous resonance of the domain wall, so that free wall magnons are not excited.

Equations (1) correspond to the local field momentum $\eta_D q\psi'$ which, in view of (1) (and this can be verified), satisfies the continuity equation

$$-\frac{d(\eta_D q\psi')}{dt}+\alpha(\dot{\psi}\psi'-q\dot{q}')+\frac{\partial}{\partial x}\left\{-h_x\cos\psi-h_y\times\sin\psi-\frac{\psi'^2}{2}+q''q-\frac{q'^2}{2}-\kappa^2\frac{q^2}{2}-\frac{q(\sin 2\psi)'}{2\sqrt{Q}}+\frac{(q'\sin 2\psi)'}{2\sqrt{Q}}-\frac{\cos 2\psi}{4}-\sin 2\psi\frac{q'}{\sqrt{Q}}\right\}=0. \tag{4}$$

According to (4), the variation of the local momentum with time is balanced by the damping of the momentum due to viscous processes and the difference of the momentum flux (the second and third terms on the left-hand side). Note that in the stationary case considered in Ref. 10 an energy balance equation was used to simplify the calculations. One of the advantages of using (4) in comparison to an energy balance equation is that in the time-dependent case Eq. (4) does not contain time derivatives of the external fields.

3. PERTURBATION THEORY

To derive the equations of motion of vertical Bloch lines, we begin with the perturbation theory formulated in Ref. 11. The differences consist (a) in the presence of noninterchangeable terms in Eqs. (1) and (b) in the use of an equation of momentum flux conservation, which simplifies the selection of necessary terms in the expansions and the calculations considerably.

Consider Eqs. (1), where in view of the inequalities (2) we have dropped the dissipative terms and the contributions of external fields. Their effect will be taken into account later via Eq. (4). We look for the solutions of Eqs. (1) in the form of series:

$$\psi(x,t)=\psi^{(0)}(u)+\psi^{(1)}(u,t)+\cdots, \tag{5a}$$

$$q(x,t)=q^{(0)}(u)+q^{(1)}(u,t)+\cdots. \tag{5b}$$

Here the local variable is $u=x-X(t)$, where $X(t)$ is the coordinate of the vertical Bloch line on the x axis, and it is the equation of this vertical line that we are seeking. In the present treatment it is sufficient to determine only the first two terms in the expansions (5a) and (5b).

The zeroth-order equations are

$$q^{(0)\prime\prime}-\kappa^2 q^{(0)}=\frac{(\sin\psi^{(0)}\cos\psi^{(0)})'}{\sqrt{Q}}, \tag{6a}$$

$$-\psi^{(0)\prime\prime}+\sin\psi^{(0)}\cos\psi^{(0)}=\frac{\cos 2\psi^{(0)}q^{(0)\prime}}{\sqrt{Q}}. \tag{6b}$$

Since according to (6a) we have $q^{(0)}\propto 1/\sqrt{Q}$, the right-hand side of Eq. (6b) is of order $1/Q$ and can be dropped. Among the solutions of Eq. (6b) with a zero right-hand side (i.e., without allowance for noninterchangeability) we select the following two, both corresponding to a 180-degree vertical Bloch line:

$$\cos\psi^{(0)}=-\tanh u,\quad \psi^{(0)\prime}=\frac{1}{\cosh u},\quad \eta_L=1, \tag{7a}$$

$$\cos\psi^{(0)}=\tanh u,\quad \psi^{(0)\prime}=-\frac{1}{\cosh u},\quad \eta_L=-1. \tag{7b}$$

To the vertical Bloch line of type (7a) we assign a positive topological charge $\eta_L=1$ and to the vertical Bloch line of type (7b), a negative topological charge $\eta_L=-1$. Both solutions have the same directions of magnetizations in the domain wall as $x\to\pm\infty$, i.e., $\psi(x\to-\infty)\to 0$ and $\psi(x\to\infty)\to\eta_L\pi$.

After this, in the $\kappa<1$ approximation we can write (see Ref. 10) the following approximate solution of Eq. (6a), which determines the deformation of the domain wall in the static state due to the noninterchangeable part of the magnetostatic energy;

$$q^{(0)}=\frac{\eta_L}{\sqrt{Q}}\left[\frac{1}{\cosh u}-\frac{\pi\kappa}{2}\exp(-\kappa|u|)\right]. \tag{8}$$

Note that the sign of the deformation of the domain wall depends only on the sign of the topological charge of the Bloch line, η_L, and is independent of the sign of the topological charge of the domain wall, η_D. Also note that it is precisely the deformation of the surface of the domain wall, Eq. (8), that leads to noninterchangeable effects in the dynamics of a vertical Bloch line.

The first-order equations are

$$q^{(1)\prime\prime}-\kappa^2 q^{(1)}=\eta_D\dot{\psi}^{(0)}+\frac{(\cos 2\psi^{(0)}\psi^{(1)})'}{\sqrt{Q}}, \tag{9a}$$

$$\hat{L}\psi^{(1)}=\eta_D\dot{q}^{(0)}+\cos 2\psi^{(0)}\frac{q^{(1)\prime}}{\sqrt{Q}}-2\sin 2\psi^{(0)}\frac{\psi^{(1)}q^{(0)\prime}}{\sqrt{Q}}, \tag{9b}$$

where

$$\hat{L}=-\frac{d^2}{du^2}-\cos 2\psi^{(0)}. \tag{9c}$$

Note that the time derivatives on the right-hand sides of Eqs. (9a) and (9b) are $\partial/\partial t=-\dot{X}\partial/\partial u$.

Since (9b) and (8) imply $\psi^{(1)}\propto 1/\sqrt{Q}$, the second term on the right-hand side of Eq. (9a) is of order $1/Q$ and can be dropped. After this, in the $\kappa<1$ approximation, the solution of Eq. (9a) can be written as

$$q^{(1)}=\frac{\eta_L\eta_D\dot{X}}{2\kappa}\exp(-\kappa|u|). \tag{10}$$

Here we note the important fact that the sum of the deformations (8) and (10) of the surface of the domain wall has no definite symmetry with respect to reversal of the sign of the velocity $\dot{X}$ of the vertical Bloch line.

In this approximation, Eq. (9b) reduces to

$$\hat{L}\psi^{(1)}=-\dot{X}\eta_D q^{(0)\prime}+\cos 2\psi^{(0)}\frac{q^{(1)\prime}}{\sqrt{Q}}. \tag{11}$$

This equation is meaningful since the kernel of operator $\hat{L}$, which is the symmetric function $1/\cosh u$, is automatically orthogonal to the antisymmetric right-hand side of Eq. (11). An approximate solution of Eq. (11) has the form[10]

$$\psi^{(1)}=\frac{\eta_L\eta_D\dot{X}}{2\sqrt{Q}}\left[\frac{u}{\cosh u}-\pi\ \mathrm{sgn}(u)\left(1-\frac{1}{\cosh u}\right)\exp(-\kappa|u|)\right], \tag{12}$$

where $\mathrm{sgn}(u)$ is the sign function. In the limit $\kappa=0$ this solution satisfies the initial equation exactly. A numerical check has shown that this approximate solution differs only slightly from the exact solution at the points of extrema of the exact solution.

We also note that in calculating (12) we ignored (and this is also done below) terms proportional to κ^2, since in calculating noninterchangeable effects we use, where possible, the limit $\kappa=0$. Also, the sum of (7) and (12) has no definite symmetry with respect to reversal of the sign of the velocity $\dot{X}$ of the vertical Bloch line (just as the sum of the deformations of the domain wall, (8) and (10) has no such symmetry). These facts are decisive in the mechanism of drift of a vertical Bloch line.

4. EQUATIONS OF MOTION AND THE DRIFT OF A VERTICAL BLOCH LINE

The zeroth- and first-order approximations (7), (8), (10), and (12) are sufficient for setting up the equations that describe the motion of a vertical Bloch line with allowance for noninterchangeable effects. If we integrate Eq. (4) for the momentum density with respect to x within infinite limits and assume that the moving vertical Bloch line emits no spin waves, we arrive at the integral form of this equation, which actually is the equation of motion of a vertical Bloch line:

$$-\frac{dP}{dt}+F_g+F_d+F_e=0, \tag{13}$$

where

$$P=\eta_D\langle q\psi'\rangle, \tag{14a}$$

$$F_g=-\pi\eta_L\eta_D\dot{\bar{q}}(t), \tag{14b}$$

$$F_d=\alpha\langle\dot{\psi}\psi'-q\dot{q}'\rangle, \tag{14c}$$

$$F_e=2h_x(t). \tag{14d}$$

According to (13), the variation of the total momentum (14a) of the vertical Bloch line with time is balanced by the gyrotropic force (14b), the viscous force (14c), and the external force (14d) generated by the field $h_x(t)$. Here and in what follows the angle brackets indicate that the expression inside the brackets is integrated with respect to u over infinite limits. The partial differential equations (1a) and (1b) have therefore been reduced to the ordinary differential equation (13), which coincides in form with the equation of motion of a material point.

The important thing is that the action of the field $h_z(t)$, which does not enter explicitly into the continuity equation (4), shifts the domain wall as a whole, which in turn shifts the vertical Bloch line due to the gyrotropic effect. Besides, the field $h_z(t)$ changes the asymptotic behavior of the desired solutions of Eqs. (1a) and (1b). Actually, we must add the term $h_z(t)$ to the right-hand side of the first-order approximation equation (9a). As a result, the solution (10) undergoes a change: to its right-hand side we must add the expression

$$\bar{q}=\frac{\eta_D h_z(t)}{\kappa^2}.$$

It is the allowance for this additional contribution in the expression for the local field momentum $\eta_D q\psi'$ that leads to the emergence of the gyrotropic force F_g (Eq. (14b)). The quantity $\overline{q(t)}$ together with the corresponding angle $\overline{\psi(t)}$ can be found from the linearized variant of Eqs. (1), where $h_z(t)$ is taken into account explicitly and the dependence on the coordinate x is dropped. Justification of a more general, nonlinear, asymptotic representation can be found in Ref. 11.

We also note that in writing the expressions (14b) and (14d) we ignored all nonlinearities in the weak external fields. According to Ref. 5, under certain conditions these nonlinearities can lead to a drift of the domain wall and hence to a drift of a vertical Bloch lines in the wall. This effect is not connected to the effects of noninterchangeability of the spectrum, requires using magnetic fields of two different polarizations, and hence is not considered here.

Next, to derive the equations of motion of a vertical Bloch line, we only need to calculate via (7), (8), (10), and (12) the momentum P (see Eq. (14a)) and the viscous force (14c) that enter into Eq. (13). In calculating the various integrals, in addition to allowing for the inequalities (3), we must leave only the first-order terms in $1/\sqrt{Q}$, where the velocity $\dot{X}$ of a vertical Bloch line (measured in units of $4\pi M\gamma\Lambda$) is also assumed small.

The momentum of a vertical Bloch line has the well-known form[11]

$$P \approx \langle \eta_D q^{(1)} \psi^{(0)\prime} \rangle \approx \frac{\pi \dot{X}}{2\kappa} \left\langle \frac{\exp(-\kappa|u|)}{\cosh u} \right\rangle \approx \frac{\pi^2 \dot{X}}{2\kappa}. \quad (15)$$

Thanks to the large factor $1/\kappa$ in (15), we can ignore all other corrections to the momentum.

The viscous force (14c) consists of two partial contributions: one is quadratic in the angle ψ, and the other is quadratic in the coordinate q of the domain wall. The first contribution can be written as follows:

$$-\alpha\dot{X}\langle(\psi^{(0)\prime})^2 + 2\psi^{(0)\prime}\psi^{(1)\prime}\rangle \approx -2\alpha\dot{X} - \alpha\frac{\eta_D\dot{X}^2}{\sqrt{Q}}\left\langle \frac{1}{\cosh u} \times\left(\frac{1}{\cosh u} - \frac{u\sinh u}{\cosh^2 u} - \pi\,\mathrm{sgn}(u)\frac{\sinh u}{\cosh^2 u}\right)\right\rangle = -2\alpha\dot{X} - \alpha\eta_D\dot{X}^2\frac{1-\pi}{\sqrt{Q}}. \quad (16a)$$

In calculating the integrals we allowed for the fact that their values are determined by rapidly decreasing exponential functions with exponents equal to unity, so that in the integrands we can ignore the exponential factors with exponents $\kappa<1$. We also took into account the well-known relationships $|u|' = \mathrm{sgn}(u)$ and $[\mathrm{sgn}(u)]' = 2\delta(u)$, where $\delta(u)$ is the Dirac delta function. Within the same approximations, for the second partial contribution to the viscous force we have

$$\alpha\langle q'\dot{q}\rangle = -\alpha\dot{X}\langle(q')^2\rangle \approx -2\alpha\dot{X}\langle q^{(0)\prime}q^{(1)\prime}\rangle \approx \frac{-\alpha\eta_D\dot{X}^2}{\sqrt{Q}}\left\langle \mathrm{sgn}(u)\frac{\sinh u}{\cosh^2 u}\right\rangle = -\frac{2\alpha\pi\eta_D\dot{X}^2}{\sqrt{Q}}. \quad (16b)$$

Thus, with allowance for noninterchangeable effects, the total value of the viscous force acting on a vertical Bloch line is

$$F_d = -2\alpha\dot{X} - \frac{\alpha\eta_D(1+\pi)\dot{X}^2}{\sqrt{Q}}. \quad (17)$$

Note that this force contains terms that are quadratic in the velocity, do not depend on the topological charge of the vertical Bloch line, and are due to the noninterchangeability of the magnetostatic energy. The viscous force has no definite symmetry with respect to reversal of the sign of $\dot{X}$, since neither the deformation of the surface of the domain wall (the sum of (8) and (10)) nor the sum of the azimuthal angles (7) and (12) has such symmetry. The direction of motion of a vertical Bloch line in which the dissipative force is at its maximum can be determined by a simple rule. The dissipative force is at its maximum if the sign of the amplitude of the dynamical (gyrotropic) bending of the domain wall, Eq. (10), coincides with the sign of the amplitude of static deformation of the domain wall, Eq. (8).

Collecting all the results, (14), (15), and (17), and inserting them in Eq. (13), we arrive at the equation of motion of a vertical Bloch line. Note that calculating the higher-order terms in the series (5a) and (5b) via the integral equation reflecting the conservation of momentum (in our approach this equation acts as the solvability condition for the problem used in Ref. 11) makes it possible to derive terms that are cubic in the velocity, which coincide exactly with the those calculated in Ref. 11.

Here we limit our discussion to the situation in which the dissipative force (17) contains cubic terms together with the quadratic terms calculated earlier. The effective equation of motion of a vertical Bloch line is

$$\frac{\pi^2}{2\kappa}\frac{d(\dot{X}+\dot{X}^3)}{dt} + 2\alpha\left(\dot{X} + \frac{\pi^2\dot{X}^3}{8\kappa}\right) + \frac{\alpha\eta_D(1+\pi)}{\sqrt{Q}}\dot{X}^2 + \pi\eta_D\eta_L\dot{\overline{q}} - 2h_x(t) = 0, \quad (18a)$$

where $\dot{\overline{q}}$ (needed if we wish to allow for the field $h_z(t)$) can be found by solving the following linear equation:

$$\ddot{\overline{q}} + \alpha\dot{\overline{q}} + \kappa^2\overline{q} = \eta_D h_z. \quad (18b)$$

Let us now study some of the simple corollaries of Eqs. (18a) and (18b). Suppose that the external field h_x is constant. Then Eq. (18a) implies that

$$\dot{X} = \frac{h_z}{\alpha}\left(1 - \frac{\eta_D a h_x}{\sqrt{Q}\alpha}\right),$$

where $a=(1+\pi)/2$ (in Ref. 10 the coefficient was erroneously written as $(1+2\pi)/2$). Analysis of this expression shows that the velocity of a vertical Bloch line is largest in absolute value when the static (8) and dynamic (10) amplitudes of deformation of the domain wall (i.e., the total deformation is at its minimum).

Let us now consider a variable external field $h_x=h\cos\omega t$ and calculate the drift velocity of a vertical Bloch line. Assuming that the noninterchangeable term is small and solving Eq. (18a) by the method of successive approximations, we find the mean velocity of drift motion of a vertical Bloch line:

$$\overline{\dot{X}}=4\pi M\gamma\Lambda\,\frac{-\eta_D(1+\pi)/\sqrt{Q}}{4\alpha^2+(\pi^2/2\kappa)^2(\omega/4\pi M\gamma)^2}\left(\frac{h}{8M}\right)^2. \tag{19}$$

Note that calculations show that in the case of a variable field $h_z(t)=h\cos\omega t$ the factor $(h/8M)^2$ in (19) must be replaced by

$$\left(\frac{\pi}{2}\right)^2\left(\frac{\omega}{4\pi\gamma M}\right)^2\left(\frac{h_z}{4\pi\gamma M\kappa^2}\right)^2.$$

Of course, in addition to the drift (19) the vertical Bloch line experiences vibrational motion. It must be said at this point that the field and frequency dependence (19) coincides with the corresponding results of Iordanskiĭ and Marchenko[6] (formula (27) in Ref. 6) and that for $Q>1$ Eq. (19) determines the constants introduced in Ref. 6 within the noninterchangeability mechanism.

The direction of drift is determined by the same line of reasoning as in the case of uniform motion of a vertical Bloch line in a constant field.[10] With the adopted mechanism for vertical Bloch lines of type (7) the direction of drift depends only on the sign of the topological charge of the domain wall, η_D. Note that in addition to vertical Bloch lines of the form (7) there are two more types with $\eta_L=\pm1$ but with other asymptotic values of the azimuthal angle: $\psi(x\to-\infty)=\pi$ and $\psi(x\to\infty)=2\pi$ $(\eta_L=1)$ or 0 $(\eta_L=-1)$. It can be verified that for such vertical Bloch lines the sign in front of the external force (14d) and the sign in front of the term with $h_x(t)$ in Eq. (18a) change, but the sign in Eq. (19) for the drift velocity remains unchanged.

In conclusion we list the main approximations for which the equation of motion of a vertical Bloch line, Eq. (18), which allows for the noninterchangeability of the spectrum of the domain wall, is valid. The specimens were thin films of perpendicularly magnetized uniaxial ferromagnets with a large Q-factor, $Q\gg1$, in which the twist of the domain wall can be ignored. Note that this class incorporates magnetic-bubble films of ferrite garnets near the magnetic moment compensation point, where formally the width parameter of a vertical Bloch line, $\Lambda(\to\infty)$ exceeds the film thickness. The Gilbert damping parameter α is also assumed small. Although the restoring force constant of the domain wall, κ, is small, it must exceed a threshold value above which flexural instability of the domain wall develops. The amplitudes of external exciting fields are small, and the frequencies of these fields are below the frequency of homogeneous resonance of the domain wall.

The work was partially supported by the Russian Fund for Fundamental Research (Grant No. 95-02-05498a).

[1] L. M. Dedukh, V. S. Gornakov, and V. I. Nikitenko, Phys. Status Solidi A **75**, K117 (1983).
[2] V. S. Gornakov, L. M. Dedukh, and V. I. Nikitenko, Zh. Éksp. Teor. Fiz. **86**, 1505 (1984) [Sov. Phys. JETP **59**, 881 (1984)].
[3] L. M. Dedukh, V. I. Nikitenko, and V. T. Synogach, Fiz. Tverd. Tela (Leningrad) **26**, 3463 (1984) [Sov. Phys. Solid State **26**, 2083 (1984)].
[4] V. S. Gornakov, L. M. Dedukh, and V. I. Nikitenko, Zh. Éksp. Teor. Fiz. **94**, No. 3, 245 (1988) [Sov. Phys. JETP **67**, 570 (1988)].
[5] V. G. Bar'yakhtar, Yu. I. Gorobets, and S. I. Denisov, Zh. Éksp. Teor. Fiz. **98**, 1345 (1990) [Sov. Phys. JETP **71**, 751 (1990)].
[6] S. V. Iordanskiĭ and V. I. Marchenko, Zh. Éksp. Teor. Fiz. **91**, 1867 (1986) [Sov. Phys. JETP **64**, 1105 (1986)].
[7] A. M. Grishin and A. Yu. Martynovich, Zh. Tekh. Fiz. **60**, No. 3, 118 (1990) [Sov. Phys. Tech. Phys. **35**, 343 (1990)].
[8] J. F. Janak, Phys. Rev. A **134**, 441 (1964).
[9] I. A. Gilinskiĭ, Zh. Éksp. Teor. Fiz. **68**, 1032 (1975) [Sov. Phys. JETP **41**, 511 (1975)].
[10] G. E. Khodenkov, Fiz. Tverd. Tela (St. Petersburg) **38**, 1149 (1996) [Phys. Solid State **38**, 636 (1996)].
[11] A. K. Zvezdin and A. F. Popkov, Zh. Éksp. Teor. Fiz. **91**, 1789 (1986) [Sov. Phys. JETP **64**, 1059 (1986)].

Translated by Eugene Yankovsky

Galvanomagnetic properties of the new organic metal $ET_2TlHg(S_{0.45}Se_{0.55}CN)_4$

S. I. Pesotskii

Institute of Chemical Physics in Chernogolovka Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia; Walther-Meissner-Institut, D-85748, Garching, Germany; International Laboratory of High Magnetic Fields and Low Temperatures, 53-529 Wroclaw, Poland

A. E. Kovalev

Institute of Solid State Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia; International Laboratory of High Magnetic Fields and Low Temperatures, 53-529 Wroclaw, Poland

P. Boehm, W. Biberacher and K. Andres

Walther-Meissner-Institut, D-85748, Garching, Germany

M. V. Kartsovnik

Walther-Meissner-Institut, D-85748, Garching, Germany; Institute of Solid State Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

N. D. Kushch and E. B. Yagubskii

Institute of Chemical Physics in Chernogolovka Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

K. V. Van

Institute of Experimental Mineralogy, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

(Submitted 4 February 1997)
Zh. Éksp. Teor. Fiz. **112**, 2178–2183 (December 1997)

Magnetoresistance and magnetic torque have been studied in the single crystals of the mixed complex $(ET)_2TlHg(S_{0.45}Se_{0.55}CN)_4$. The quantum oscillations (Shubnikov–de Haas and de Haas–van Alphen) have been observed and investigated in details in this compound. The behavior of these oscillations is close to that in the complex $(ET)_2TlHg(SeCN)_4$ but is characterized by a stronger carrier scattering induced by the disordered positions of the S and Se atoms in the anion. [S1063-7761(97)01812-X]

1. INTRODUCTION

In recent years the isostructural organic metals of the α-$ET_2XHg(YCN)_4$ series, where ET stands for bis(ethylenedithio)tetrathiafulvalene ($C_{10}H_8S_8$), X=K, Tl, Rb, NH_4, and Y=S, Se, have been a subject of intensive studies.[1] These compounds have revealed a variety of surprising properties in high magnetic fields, arising from the specifics of their electronic band structure. The Fermi surface FS,[2] essentially the same in all these metals, uniquely combines features characteristic of quasi-one-dimensional (Q1D) and quasi-two-dimensional (Q2D) systems being represented by the coexistence of a slightly warped cylinder and a pair of open sheets. The series can be subdivided into two groups: One of them includes the salts with X=K, Tl, Rb and Y=S which undergo a phase transition into an antiferromagnetic groundstate which is believed to be driven by a Peierls-like instability of the Q1D FS part. However, unlike in most other Q1D metals, the transition does not lead to an insulating state[3–5] although it brings about an essential rearrangement of the FS.[6] In contrast, no transition is observed in the salts of the other group with $X=NH_4$, Y=S and X=Tl, K, Y=Se and the FS retains its initial shape at low temperatures.[7–9] As a result, the metals of the two groups exhibit very different properties in magnetic fields.

In particular, the salt $ET_2TlHg(SCN)_4$ belonging to the first group exhibits the following properties. (1) A phase transition into an antiferromagnetic groundstate at $T_p \approx 10$ K (Ref 5); (2) Strong angle-dependent magnetoresistance oscillations (AMRO) which are presumed to originate from the electron motion along new open sheets of the reconstructed FS (Ref. 6); (3) Shubnikov–de Haas (SdH) and de Haas–van Alphen (dHvA) oscillations with fundamental frequency $F_0=670$ T at $\mathbf{H}\|\mathbf{b}^*$ (where $\mathbf{b}^*$ is the normal to the crystal highly-conducting ac-plane) demonstrating enormously high second-harmonic contribution at low temperature, in fields above 10 T (Ref. 10). On the other hand, the selenium-containing complex, $ET_2TlHg(SeCN)_4$, is characterized by (1) the absence of the phase transition; (2) AMRO associated with the electron orbits on the cylindrical FS; and (3) strong SdH and dHvA oscillations ($F_0=650$ T) with a normal harmonic content described by the standard Lifshitz–Kosevich formula in the field up to 15 T (Ref. 8).

In order to obtain more information on the effect of subtle structural changes in the electronic band structure andthe groundstate properties induced by the substitution of the S atoms by the Se ones in the anion, we have carried out

magnetic-field studies of a new organic conductor with the mixed composition, $ET_2TlHg(S_{0.45}Se_{0.55}CN)_4$.[11]

2. EXPERIMENT AND SAMPLES

In the experiment, single crystals of the mixed salt $ET_2TlHg(S_{0.45}Se_{0.55}CN)_4$ having a typical size of $\approx 1\times 1\times 0.2\ mm^3$ were used. The samples were prepared by electrocrystallization in benzonitrile as described by Rousseau *et al.*[11] or in trichlorethane with the addition of absolute ethanol (10 vol. %) at 20° at constant current, $I=0.5\ \mu A$. In the latter case the electrolyte was a mixture of cyclic 18-crown-6 ether and TlSCN and $Hg(SeCN)_2$ salts in the ratio of 1:3:1. The samples obtained from different batches had different quality. Nevertheless, the microprobe analysis has shown the same chemical composition for all of them, giving a stable S:Se ratio of 8:1, in agreement with the chemical formula of the complex. A x-ray photomethod test has confirmed that these samples are single crystals isostructural to the other metals of the α-phase series.

The magnetoresistance and torque measurements were carried out as functions of the direction and strength of the magnetic field and the temperature. The magnetoresistance was measured by the standard a.c. technique at 330 Hz with a current $I=100\ \mu A$ directed perpendicular to the highly-conducting layers. This geometry provides the best conditions for observation of AMRO.[6,12] The magnetic field up to 15 T was generated by a superconducting solenoid. The measurements of magnetoresistance were performed at temperatures down 1.5 K. The dHvA effect was studied by a cantilever torquemeter as described in Ref. 13. These measurements were carried out in magnetic fields up to 14 T and temperatures down to 0.5 K.

3. RESULTS AND DISCUSSION

As in the case of the selenium salt,[8] the samples of the mixed complex do not show any sign of the phase transition characteristic of its sulfur analog.[5]

The SdH and dHvA oscillations in the samples of the mixed salt are clearly exhibited in the field dependence of the resistance and magnetic torque in fields $\geqslant 10$ T (see Figs. 1 and 2). FFT analysis yields the oscillation fundamental frequency of 650 ± 10 T for $\mathbf{H}\|\mathbf{b}^*$. This value is close to those obtained for both the pure sulfur and selenium salts.[8,10] Neither SdH nor dHvA oscillations show any significant contribution of frequencies other than the fundamental within the field and temperature range studied. The angle dependence of the oscillation frequency is well described by the expression $F(\theta)=F_0/\cos\theta$, where θ is the angle between the field and $\mathbf{b}^*$ directions (Fig. 3), in accordance with the cylindrical shape of the corresponding FS part. Basically, the described behavior is very similar to that found for the selenium salt. We only note that the value of the background magnetoresistance and the SdH amplitude in the mixed complex are, respectively, a factor of three and a factor of six lower than in the selenium analog.[8]

From the temperature dependence of the SdH and dHvA amplitudes we estimate the cyclotron mass of the carriers (m^*) on the cylindrical part of the FS. In the conducting plane it satisfies $m_0^*=(2.0\pm 0.1)m_e$, where m_e is the free

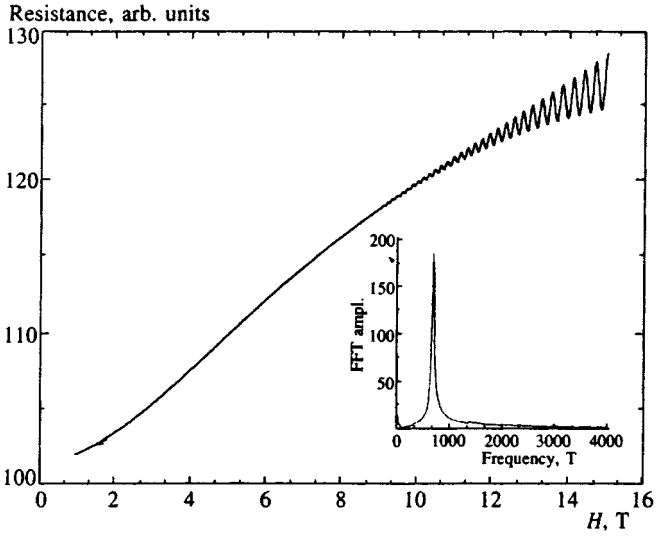


FIG. 1. Shubnikov–de Haas oscillations in $ET_2TlHg(S_{0.45}Se_{0.55}CN)_4$; $\mathbf{H}\|\mathbf{b}^*$, $T=1.5$ K, $\mathbf{I}\|\mathbf{b}^*$. Inset: FFT of the oscillations.

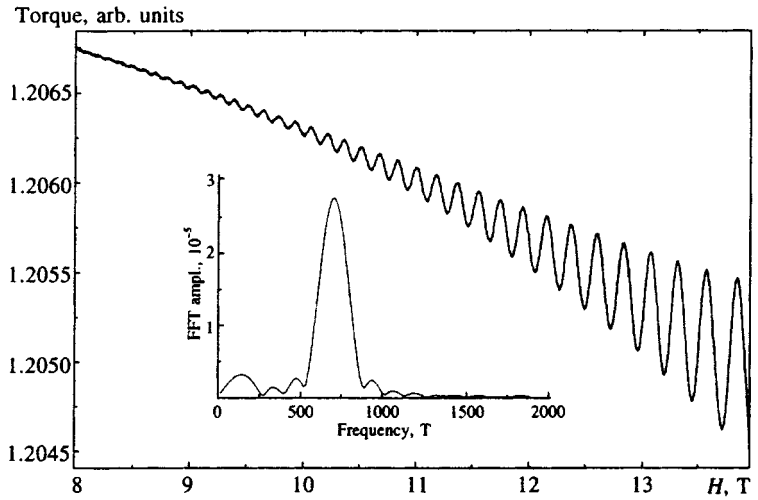


FIG. 2. De Haas–van Alphen oscillations in $ET_2TlHg(S_{0.45}Se_{0.55}CN)_4$. The magnetic field is tilted by the angle $-21.6°$ with respect to the $\mathbf{b}^*$ direction; $T=0.5$ K. Inset: FFT of the oscillations.

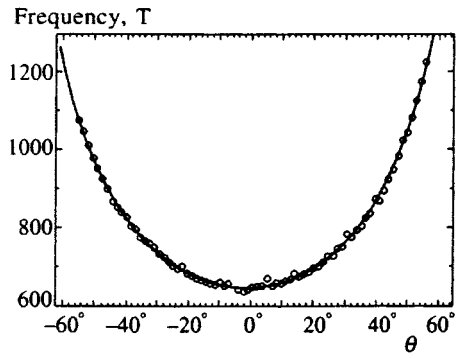


FIG. 3. Angular dependence of dHvA frequency. The solid line corresponds to the relation $F(\theta)=650T/\cos\theta$.

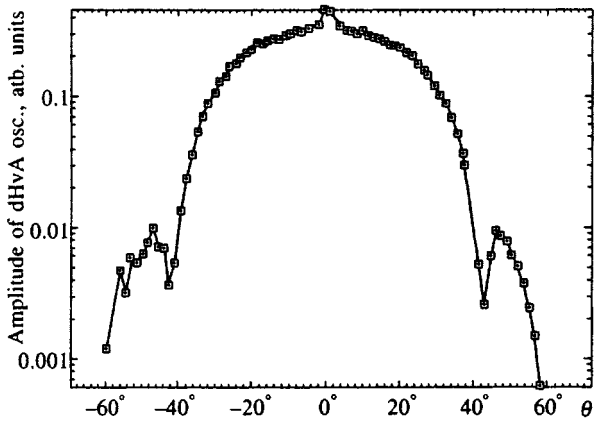


FIG. 5. Angular dependence of dHvA amplitude. $T=0.5$ K.

electron mass. The mass of the carriers in the selenium salt has the same value.[8] The dependence of the cyclotron mass on the field direction for the dHvA oscillations (Fig. 4) obeys the relation $m^*(\theta)=m_0^*/\cos\theta$, within experimental error, in agreement with the cylindrical shape of the FS.

The angular dependence of the dHvA amplitude in the mixed complex is shown in Fig. 5. There are two pairs of angles, $\theta\approx\pm 42°$ and $\theta\approx\pm 58°$, at which the amplitude vanishes due to the spin-splitting effect.[14] Using the experimentally obtained cyclotron mass, we evaluate the g-factor from the condition for the spin-splitting zeros,[14] taking into account the cylindrical shape of the FS.

$$\cos\left(\frac{\pi}{2}\,pg\,\frac{m_0^*}{m_e\cos\theta}\right)=0,$$

where $p=1$ corresponds to the fundamental harmonic, $g\approx 1.7$. This value is very close to that found for the selenium analog.[8] As in the latter case, we attribute the deviation from the free electron g-factor to the substantial effect of the electron-phonon interaction.

The Dingle temperature obtained from the field dependence of the oscillation amplitude in various samples of the mixed complex varies from 1.4 to 2.2 K, which is much higher than the value $T_D\approx 0.6$ K reported for the selenium salt.[9] The high Dingle temperature explains the relatively low magnetoresistance and the SdH amplitude in the mixed complex as compared to the selenium salt.

These results suggest that the mixed complex, $ET_2TIHg(S_{0.45}Se_{0.55}CN)_4$, is a ''dirty'' version of the pure selenium salt $ET_2TIHg(SeCN)_4$. This conclusion, in principle, is consistent with the absence of the AMRO like those found in the pure selenium salt.[8] Indeed, the amplitude of the AMRO is expected to depend on magnetic field and the sample quality as $\sim(\omega\tau)^2$ where ω is the cyclotron frequency, $\omega=eH/m^*c$, and τ is the relaxation time. Using the Dingle temperature to estimate the relaxation time, one obtains an $\omega\tau$ value in the mixed salt an order of magnitude lower than that obtained for the pure selenium analog at the same magnetic field. Therefore, the AMRO amplitude is expected to be an order of magnitude lower in the compound studied. Figure 6, which illustrates the angular dependence of the magnetoresistance at an arbitrary azimuthal angle in ac-plane shows only one shallow maximum at around $-40°$ which could be attributed to the AMRO effect, in agreement with the above consideration. (A picture like Fig. 6 was observed at the different azimuthal angles.) Note, however, that a different (in comparison with the selenium salt) shape or degree of distortion of the FS cannot be completely ruled out as a possible reason for the weakness of AMRO in the present case. Thus, the reason for the absence of AMRO in

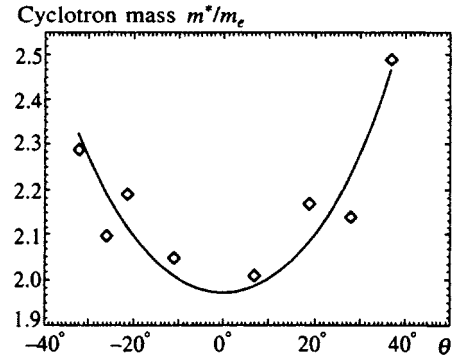


FIG. 4. Angular dependence of the cyclotron mass. The solid line corresponds to the relation $m^*\theta=2.0m_e/\cos\theta$.

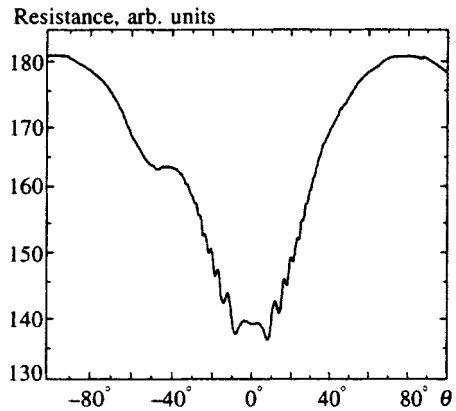


FIG. 6. Angular dependence of the magnetoresistance. $H=15$ T, $T=1.5$ K, $\mathbf{I}\|\mathbf{b}^*$.

the mixed complex is still obscure and a more detailed study is needed for its clarification.

Summarizing, we have studied the magnetoresistance and magnetic torque of the mixed complex $ET_2TIHg(S_{0.45}Se_{0.55}CN)_4$. The complex is found to closely resemble its pure selenium analog, $ET_2TIHg(SeCN)_4$ but is likely characterized by a stronger carrier scattering induced by the disordered positions of the S and Se atoms in the anion. The disorder may be a reason for the suppression of the density-wave instability at low temperatures. On the other hand, unrelated to the disorder, the low-temperature properties are determined by the details of the inherent electronic band structure of the material. The latters could be essentially different from those of $ET_2TIHg(SCN)_4$ due to the presence of the Se atoms. Further studies are necessary to understand the crucial factors responsible for the low-temperature ground state in the system.

The present work is supported by the Volkswagen Stiftung Grant No. 1/70 206, by the Russian Fund for Fundamental Research, Grant No. 96-02-17475, and by a grant from the Russian State Scientific Technical Program on ''The physics of quantum and wave processes.'' S.I.P., A.E.K., M.V.K. and N.D.K. also aknowledge partial support under grant INTAS-93-2400-EXT.

[1] J. Wosnitza, *Fermi Surface of Low-Dimensional Organic Metals and Superconductors*, Springer, Berlin, Heidelberg (1996). J. S. Brooks *et al.*, Phys. Rev. B **52**, 14457 (1995). M. V. Kartsovnik and V. N. Laukhin, J. de Phys. I **6**, 1753 (1996).
[2] H. Mori, S. Tanaka, K. Oshima *et al.*, Bull. Chem. Soc. Jpn. **63**, 2183 (1990).
[3] T. Sasaki, N. Toyota, M. Tokumoto, N. Kinoshita, and H. Anzai, Solid State Commun. **75**, 93 (1990).
[4] N. Kinoshita, M. Tokumoto, and H. Anzai, J. Phys. Soc. Jpn. **60**, 2131 (1991).
[5] N. D. Kushch, L. I. Buravov, M. V. Kartsovnik *et al.*, Synth. Met. **46**, 271 (1992).
[6] M. V. Kartsovnik, A. E. Kovalev, and N. D. Kushch, J. de Phys. I **3**, 1187 (1993).
[7] J. Wosnitza, G. W. Crabtree, H. H. Wang *et al.*, Phys. Rev. B **45**, 3018 (1992).
[8] A. E. Kovalev, M. V. Kartsovnik, and N. D. Kushch, Solid State Commun. **87**, 705 (1993).
[9] T. Sasaki, O. Azawa, H. Mori, S. Tanaka, T. Fukasa, and N. Toyota, J. Phys. Soc. Jpn. **65**, 213 (1996).
[10] M. V. Kartsovnik, A. E. Kovalev, V. N. Laukhin, S. I. Pesotskii, and N. D. Kushch, JETP Lett. **55**, 339 (1992).
[11] R. Rousseau, M.-L. Doublet, E. Canadell *et al.*, J. de Phys. I **6**, 1527 (1996).
[12] M. V. Kartsovnik, V. N. Laukhin, S. I. Pesotskii, I. F. Schegolev, and V. M. Yakovenko, J. de Phys. I **2**, 89 (1992).
[13] P. Christ, W. Biberacher, H. Müller, and K. Andres, Solid State Commun. **91**, 451 (1994).
[14] D. Schoenberg, *Magnetic Oscillations in Metals*, Cambridge University Press, Cambridge (1984).

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

Modeling a dimer state in $CuGeO_3$ in the two-dimensional anisotropic Heisenberg model with alternated exchange interaction

S. S. Aplesnin

L. V. Kirenskiĭ Institute of Physics, Siberian Branch of the Russian Academy of Sciences, 660036 Krasnoyarsk, Russia
(Submitted 28 March 1997)
Zh. Éksp. Teor. Fiz. **112**, 2184–2197 (December 1997)

The two-dimensional Heisenberg spin-1/2 model with alternated exchange interaction along the c axis and an anisotropic distribution of the exchange interaction in the lattice, $J_b/J_c=0.1$, is examined. A quantum Monte Carlo method is used to calculate the phase diagrams of the antiferromagnet, the dimer state in a plane, the value of the alternation δ of the exchange interaction, and the anisotropy $\Delta=1-J^{xy}/J^z$ of the exchange interaction, $\Delta\sim\delta^{0.58(6)}$. The following characteristics are calculated for $\Delta=0.25$: the dependence of the temperature of the dimer-state–paramagnet transition on the alternation of the exchange interaction, $T_c(\delta)=0.55(4)(\delta-0.082(6))^{0.50(3)}$, the singlet–triplet energy gap, and the dependence of the magnetization on the external field for some values of δ. The value of the exchange interaction, $J_c=127$ K, the alternation of the exchange interaction, $\delta=0.11J_c$, and the correlation radius along the c axis, $\xi_c\approx 28c$, are determined. Finally, it is found that the temperature dependence of the susceptibility and the specific heat are in good agreement with the experimental data. © *1997 American Institute of Physics.* [S1063-7761(97)01912-4]

1. INTRODUCTION

At liquid-helium temperatures the compound $CuGeO_3$ is in a nonmagnetic state separated from the excited state by an energy gap E. The value of this gap in the spectrum of magnetic excitations is known from inelastic neutron scattering experiments ($E_N=23.5$ K; see Refs. 1 and 2), from electron spin resonance (EPR) experiments ($E_{\rm ESR}=22\pm 5$ K; see Ref. 3), and from the temperature dependence of the magnetic specific heat ($E_C=30$ K; see Ref. 4). The magnetic susceptibility rapidly tends to zero along all three axes when $T<T_c=14$ K holds (see Refs. 5 and 6), and so does the integral spin-resonance intensity.[7,8] Measurements of the magnetization in strong magnetic fields reveal a jump in magnetization, $\Delta M\approx 0.025\mu_B$, in critical fields H_c along the a, b, and c axes at $H_{ca}=12.9$ T, $H_{cb}=12.4$ T, and $H_{cc}=13.6$ T (see Ref. 8), which also points to the presence of an energy gap.

Most researchers (see, e.g., Refs. 6 and 8) relate the formation of a dimer state in $CuGeO_3$ to a spin–Peierls transition. Crystal structure studies reveal a doubling of the crystal cell along the c and a axes when the lattice constants along these axes change by approximately 0.2% at temperatures lower than $T_c=14$ K (Ref. 9). Measurements of lattice constants and the energy gap at high pressures (up to 4 GPa)[10,11] do not suggest that a spin–Peierls transition takes place in $CuGeO_3$, since the lattice constant c is practically pressure-independent, while the gap increases by 20%. Calorimetric studies in the vicinity of T_c reveal a similarity in the jump in the specific heat, $\Delta C\approx 0.12$ J mol^{-1} K^{-1} (Ref. 4), which roughly amounts to 7% of the specific heat at $T=14$ K. According to the model of a spin–Peierls transition, the jump in specific heat at T_c in the Hartree–Fock approximation is $\Delta C=1.43\gamma T_c$ ($\gamma=2k_B^2/3J$) (Ref. 12), and for $CuGeO_3$ the theoretical estimate of $\Delta C\approx 5$ J mol^{-1} K^{-1} exceeds the experimental value by more than a factor of ten.

In another model[13,14] it is assumed that there is a competing antiferromagnetic interaction J_2 along the c axis that involves next-nearest neighbors. In this model there is an energy gap in the absence of dimerization of the lattice, and the susceptibility is well described in the high-temperature region with the parameters $\alpha=J_2/J_1=0.36$ and $J_1=160$ K (Ref. 13). However, the calculated transition temperature proved to be higher than the experimentally observed value, $T>14$ K. To resolve this contradiction, Castilla *et al.*[14] proposed considering a spin–Peierls interaction in an antiferromagnet with $\alpha<\alpha_c=0.25$ (Ref. 15), with $\alpha=0.24$, and $J_1=150$ K, i.e., as if the gap was caused by weak dimerization of the lattice and the temperature behavior of $\chi(T)$ was due to the competition of exchange interaction. This model, however, does not describe the magnetization curve $M(H)$ in strong pulsed fields,[16] and most neutron-diffraction studies[17–19] do not reveal this fairly strong interaction of about 37 K. Besides, the introduction of impurities within this model does not facilitate the formation of antiferromagnetic order. In Ref. 20 the present author suggested allowing for four-spin exchange, which forms a singlet–triplet gap and describes fairly well the temperature behavior of $\chi(T)$ in $CuGeO_3$; however, the four-spin exchange constant has an unrealistic value if the anisotropy of exchange is taken into account.

In all the models mentioned above it is assumed that the spin system is one-dimensional. Nishi *et al.*[18] used neutrons to find the three exchange parameters $J_c=120$ K, $J_b/J_c=0.1$, and $J_a/J_c=-0.01$ along the three directions. No interaction J_2 along the c axis was detected. Using the crystal structure parameters, Khomskii *et al.*[21] calculated the exchange interaction parameters, $J_n^c=11.6$ meV, $J_b/J_c\approx 0.06$, $J_a/J_c\approx -3\times 10^{-5}$, $J_{nn}^c/J_n^c=0.23-0.3$, and estimated the spin–Peierls transition temperature in the

three-dimensional case, $T_{SP}\sim\lambda' J_c/\ln(\lambda' J_c/zJ_b)$, where z is the number of nearest chains, and λ' is the spin–phonon coupling constant. For $CuGeO_3$ we have $\lambda'=0.2J_c$ (Ref. 22), and the calculated value of the spin–Peierls transition temperature $T_{SP}\sim 38$ K exceeds the experimental value $T_c^{\rm expt}=14$ K severalfold. All the models ignore the anisotropy of susceptibility, which in the vicinity of the maximum susceptibility amounts to roughly 30%; the anisotropy of the paramagnetic Curie point $\theta_b/\theta_c\approx 0.75$ (Ref. 5), and the anisotropy of the critical fields $H_{ca}/H_{ba}=1.055$ (Ref. 8). Thus, analysis of the experimental data shows that $CuGeO_3$ is a quasi-two-dimensional magnetic material with anisotropic exchange interactions, has a gap in the excitation spectrum, and has a zero spin average at a site.

As in the one-dimensional model, we consider the alternation of the exchange interaction between the nearest neighbors along the c axis. The alternation of this interaction can be caused not only by nonlinear displacements of atoms related to a structural phase transition but also by nonlinear interactions between the spin and phonon subsystems, when the frequencies of the elastic lattice vibrations are close to magnon frequencies. In $CuGeO_3$, Nishi *et al.*[18] detected a softening of the phonon mode from 34 cm^{-3} at 300 K to 30 cm^{-1} at 5 K. Several questions arise in connection with this model: Can a dimer (singlet) state form in the two-dimensional Heisenberg model? What effect will the anisotropy of the exchange interaction have on this state? What is the dependence of the energy gap between the ground and excited states on the size of alternation of the exchange interaction? If we establish the magnetic structure, we can conjecture about the nature of the elementary excitations and the possible effects of doping $CuGeO_3$ with nonmagnetic impurities.

One approach to solving these problems, which follow from the experimental data, is to use the quantum Monte Carlo method based on the trajectory algorithm.[23] The main idea of the algorithm is to transform a D-dimensional quantum problem into a $(D+1)$-dimensional classical problem by introducing ''temporal'' sections in the space of the imaginary time $0<\tau<1/T$ and to carry out the Monte Carlo procedure in the ''imaginary-time–coordinate'' space.

2. THE MODEL

Let us consider a two-dimensional lattice with spins $S=1/2$ localized at the lattice sites and the spin–spin bonds between the nearest spins distributed anisotropically. According to the experimental data of Nishi *et al.*,[18] the exchange interaction along one of the axes is roughly ten times stronger than along the other: $J_b/J_c=0.1$. The exchange interaction J_c is alternated along the c axis, which we denoted by z, i.e., $J^c_{1,1+1}=J_0+\delta$ and $J^c_{1+1,1+2}=J_0-\delta$. This nonuniformity in the exchange interaction is due to lattice distortion, $J^c_{1,1+1}-J^2_{1+1,1+2}=\lambda'(u_1-u_{1+1})$ (here u is the displacement of the atom from the equilibrium position), and to the anharmonicity of the vibrations. The Hamiltonian is

$$H=-\frac{1}{2}\sum_{i,j=1}^{L}\left\{J^{z(b)}_{i,j}S^z_iS^z_j+J^{x,y(b)}_{i,j}\frac{S^+_iS^-_j+S^-_iS^+_j}{2}\right\}$$

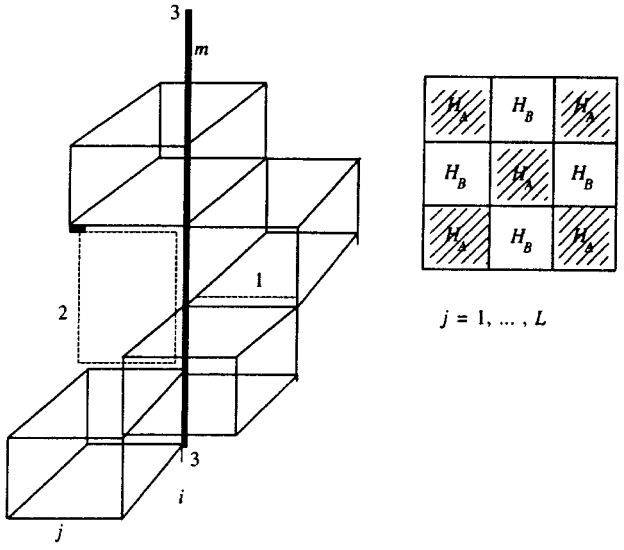


FIG. 1. Partitioning of a lattice into four-spin clusters described by the Hamiltonians H_A and H_B. The lattice is three-dimensional and consists of spatial coordinates (i,j) and the ''time'' (m). Local (1), loop (2) (dashed line), and global (the heavy straight line 3) rotations of spins were performed in the lattice.

$$-\frac{1}{2}\sum_{i,j=1}^{L}\left\{(J^{z(c)}_{i,j}+(-1)^j\delta^z)S^z_iS^z_j+(J^{x,y(b)}_{i,j}+(-1)^j\delta^{x,y})\frac{S^+_iS^-_j+S^-_iS^+_j}{2}\right\}-\sum_{i=1}^{N}h^zS^z_i, \qquad (1)$$

where $J^b<0$ is the anisotropic interaction along the b axis, $J^{z(b)}>J^{x,y(b)}$; $J^c<0$ is the anisotropic interaction along the c axis, $J^{z(c)}>J^{x,y(b)}$; $\delta^{z(x)}$ is the measure of alternation of the exchange interaction along the c axis; $H=h^z/J^c$ is the external magnetic field along the axis external magnetic field along the c axis; and L is the linear dimension of the lattice $(N=L\times L)$.

To transform a D-dimensional quantum system into a $(D+1)$-dimensional classical system, we partition the Hamiltonian into terms with different types of coupling and use the Trotter formula for decomposing an exponential operator:[24]

$$e^{A1+A2+A3+\cdots+Ap}=\lim_{m\to\infty}[e^{A1/m}e^{A2/m}e^{A3/m}\cdots e^{Ap/m}]^m, \qquad (2)$$

where m is a positive integer known as the Trotter number. Suzuki[25] proved an equivalence theorem on the correspondence between the partition function of a D-dimensional quantum system and the partition function of a $(D+1)$-dimensional classical system. The partitioning of the Hamiltonian into terms is arbitrary; often it is the even and odd terms that are grouped separately. For a fixed m the condition (2) is approximate, and increasing the size of the cluster makes the calculation errors smaller, as shown in Ref. 26. We partition the Hamiltonian (1) into four-spin clusters (Fig. 1) with overlapping values of the exchange interactions J, which enter into the clusters A and B with a weight equal to 1/2:

$$H=H^A+H^B=\sum_{i,j=1}^{L}(H^A_{2i-1,2j-1}+H^A_{2i,2j})$$
$$+\sum_{i,j=1}^{L}(H^B_{2i-1,2j}+H^B_{2i,2j-1}), \quad (3)$$

and the mth approximation of the partition function $Z(m)$ has the form

$$Z(m)=\mathrm{Tr}\left[\left\{\exp\left(-\frac{\beta H^A}{m}\right)\exp\left(-\frac{\beta H^B}{m}\right)\right\}^m\right], \quad (4)$$

where $\beta=1/k_BT$. When m tends to infinity, $Z(m)$ becomes the exact partition function in the thermodynamic limit. Since all four-spin clusters commute with each other in every part of the Hamiltonian, H^A or H^B, we have the following expression for $Z(m)$:

$$Z(m)=\mathrm{Tr}[(L_AL_B)^m]$$
$$=\mathrm{Tr}\left[\left\{\prod_{i,j=1}^{L}\exp\left(-\frac{\beta H^A_{2i-1,2j-1}}{m}\right)\right.\right.$$
$$\times\exp\left(-\frac{\beta H^A_{2i,2j}}{m}\right)\exp\left(-\frac{\beta H^B_{2i-1,2j}}{m}\right)$$
$$\left.\left.\times\exp\left(-\frac{\beta H^B_{2i,2j-1}}{m}\right)\right\}^m\right]. \quad (5)$$

Let us form a complete orthogonal set of state vectors $\sigma_r=\{S_{i,j,r}\}$, $i,j=1,2,\ldots,L$, $r=1,2,\ldots,4m$. Then

$$Z(m)=\sum_{\sigma_r}\langle\sigma_1|L_A|\sigma_2\rangle\langle\sigma_2|L_B|\sigma_3\rangle\cdots(\sigma_{2m}|L_A|\sigma_1\rangle. \quad (6)$$

The expression for the classical partition function of an $N\times 4m$ three-dimensional lattice can be written as

$$Z(m)=\sum_{\{\sigma_r\}}\prod_{\langle i,j,r\rangle}\exp(-\beta E(j,i,r)). \quad (7)$$

Here $E(i,j,r)$ is the energy of a block of eight spins and is determined by the matrix element

$$\exp(-\beta E(i,j,r))$$
$$=\langle S_{i,j,r}S_{i+1,j,r}S_{i,j+1,r}S_{i+1,j+1,r}|$$
$$\times\exp(-\beta H^{A,B}(i,j))|S_{i,j,r+1}S_{i+1,j,r+1}$$
$$\times S_{i,j+1,r+1}S_{i+1,j+1,r+1}\rangle. \quad (8)$$

This 16×16 matrix $H^{A,B}(i,j)$ consists of four independent matrices $T_1=(2\times 2)$, $2T_2=(4\times 4)$, and $T_3=(6\times 6)$. The eigenvalues and eigenvectors of the matrices T_1 and T_2 can be calculated analytically and those of the matrix T_3, numerically.

There are three types of spin flip that are used in the Monte Carlo procedure: loop, local, and global (Fig. 1). All spins on the specified lines rotate according to the transition probability, which is determined by the change in the local energy (8). The zero matrix elements (8) correspond to infinite energy and are excluded from the procedure. Spin rotations on horizontal lines have an infinitesimal probability of occurring and do not participate in the procedure either. Monte Carlo calculations use periodic boundary conditions in the Trotter direction and on the lattice. The linear dimension of the lattice L was set at 40, 48, and 64 and m, at 16, 24, and 32. The number of Monte Carlo steps per spin varied from 3000 to 10 000. A single Monte Carlo step is defined as the rotation of all spin on an $L\times L\times 4m$ lattice. The energy E and the specific heat C were calculated by the following formulas:

$$E=\left\langle\frac{1}{2}\sum_{i,j,r}F^r_{x,y}\right\rangle, \quad F^r_{i,j}=-\frac{\partial(\ln\rho^r_{i,j})}{\partial\beta}, \quad C=\frac{dE}{dT}. \quad (9)$$

Here the $\rho^r_{i,j}$ are the matrix elements of the local density matrix ($i,j=1,\ldots,L$ and $r=1,\ldots,m$). The sum is over $L\times L\times m$ eight spin clusters, and the angle brackets denote a thermodynamic average. The magnetization M and the longitudinal susceptibility χ are defined as follows:

$$M=\left\langle\sum_{i,j,r}M^r_{i,j}\right\rangle, \quad \chi=\frac{M}{H},$$
$$M^r_{i,j}=\frac{1}{4m}\sum_{h_x,h_y=0}^{1}(S^r_{i+h_x,j+h_y}+S^{r+1}_{i+h_x,j+h_y}), \quad S_i=\pm 1. \quad (10)$$

We calculated the longitudinal spin–spin and four-spin correlation functions, $R(r)=\langle S^z_0S^z_r\rangle$ and $\langle S^z_iS^z_{i+1}S^z_{i+r}S^z_{i+r+1}\rangle$, and their Fourier transforms along the sides and diagonals of the lattice normalized to $1/S^2$ and $1/S^4$ (S is the spin), respectively:

$$S^z(q)=\frac{1}{L}\sum_{r=1}^{L}\exp(-iqr)\langle S^z_0S^z_r\rangle,$$
$$\langle t_qt_{-q}\rangle=\frac{1}{L}\sum_{r=1}^{L}\exp(-iqr)\langle S^z_0S^z_1S^z_rS^z_{r+1}\rangle. \quad (11)$$

We define the thermodynamic average of the spin at a site as $\sigma=\lim_{r\to\infty}\sqrt{|\langle S^z_0S^z_r\rangle|}$. The correlation radius ξ of spin–spin interaction and the pre-exponential exponent η are defined as follows:

$$R(r)=A/r^\eta\exp(-r/\xi), \quad (12)$$

where $R(r)$ is the normalized correlation function, $R(r)=|\langle S^z(0)S^z(r)\rangle-\langle S^z\rangle^2|$.

Let us now calculate the correlation functions of the normal type for spin operators, $\langle S^+(0)S^-(r)\rangle$, by the Hirsch method.[27] The idea of this method is that the world lines are broken in the Trotter direction at a distance $r=m$ and the wave functions in the S^z-representation are compared at this distance, and then their asymptotic behavior is determined for $m\to 0$. Calculating these correlations involves a new Monte Carlo procedure with free boundary conditions in the Trotter direction and doubles the computation time.

The statistical error in the Monte Carlo calculations was estimated by standard methods. The average value was calculated, the instantaneous value was stored, and after completion of the Monte Carlo procedure the standard deviation was calculated. The error was found to vary from 0.1 to 2%. A systematic error emerges because m is finite; the error is proportional to roughly $A/(mT)^2$.

3. EFFECT OF THE ANISOTROPY OF THE EXCHANGE INTERACTION ON THE STABILITY OF THE ANTIFERROMAGNETIC STATE IN ALTERNATION OF THE EXCHANGE INTERACTION

When the exchange interaction in a magnetic material is anisotropic, there is a tendency of long-range antiferromagnetic order of the Néel type to set in in the material. Alternation of the exchange interaction forms a dimer state, where the pairs of the nearest spins form a singlet. An isolated pair has a gap between the singlet and the triplet, $E_c = J^{x,y}$. The correlation functions of the nearest neighbors in a singlet and in an anisotropic antiferromagnet without alternation ($\delta = 0$) are related as follows, respectively: $\langle S_0^+ S_1^- \rangle = 2\langle S_0^z S_1^z \rangle$, and $\langle S_0^z S_1^z \rangle \gg \langle S_0^+ S_1^- \rangle$. According to the model of resonant valence bonds, pairs of spins coupled by strong interactions ($J_0 + \delta$) are in the singlet state, and between the other pairs of spins there is no interaction along the c axis. Then the four-spin correlation function in the longitudinal components, represented by the product of these pairs $q = \langle S_0^z S_1^z S_r^z S_{r+1}^z \rangle - \langle S_0^z S_1^z S_{r+1}^z S_{r+2}^z \rangle$, is much larger along odd distances than along even. Between the antiferromagnetic state and the state of resonant valence bonds there is a mixed state, where between the singlets positioned at a distance equal to the correlation radius ξ there are correlations in the longitudinal components of spin. At a certain distance these correlations tend to zero, $\langle S_0^z S_r^z \rangle \to 0$, and the spins do not interact. In this way the magnetic materials is partitioned into unrelated microregions, in which ordinary spin excitations can propagate. According to Schulz and Ziman,[28] the limited size of these regions causes a gap $E_c \sim 1/\xi$ to occur in the excitation spectrum. A dimer state becomes energy-preferable to Néel order because of the large contribution of entropy into the free energy; for instance, the entropy satisfies $S \sim \ln W$, where $W = C_N^d = N!/d!(N-d)!$, with d the number of dimers.

To find the dependence of the region of stability of the antiferromagnetic and dimer states on the measure of alternation, we must calculate the above characteristics at low temperature, $T/J = 0.06$ and 0.125, for a number of values of the parameter of anisotropy of the exchange interaction: $\Delta = 1 - J^{x,y}/J^z = 0.1$, 0.2, 0.25, 0.3, 0.4, and 0.5. The critical value of alternation, δ_c, at which long-range antiferromagnetic order is disrupted, is determined by the disappearance of magnetization at a site and can be found from the inflection point of the dimerization parameter q. Dimers are also formed in magnetically ordered states. For instance, the spin correlation function in the transverse components increases along the c axis and decreases along the b axis. The energy of a dimer state is higher than the energy of the antiferromagnetic state and increases with the alternation of the exchange interaction as a power function, $E - E_{AF} = A\delta^\beta$, where the exponent β increases with the anisotropy of the exchange interaction and at $\Delta = 0$ is equal to 1.5(1). After the transition to a dimer state is completed, the correlations between the spins in the b-direction weaken drastically: $\langle S_0^z S_1^z \rangle$ decreases severalfold. This is due to the formation of singlet pairs along the c axis. To determine the magnetic structure, we printed out the instantaneous values of the magnetization at a site and the spin correlation functions $\langle S_0^z S_1^z \rangle$ at a site. Possibly, the solitons in neighboring chains are shifted in relation to each other.

Knowing the dependence of the spin–spin correlation functions on the distance in the three directions, c, b, and along the diagonal, we can find the correlation radius and the average dimensions of the ordered region of the soliton type. The correlation radius diverges in the vicinity of the critical value δ_c as a power function: $\xi = 5(4)/(\delta - \delta_c)^{0.50(5)}$. Within 10% this function is universal and is independent of the anisotropy of the exchange interaction. For the isotropic case the exponent is different: $\xi = 5/\delta^{0.70(4)}$. The nonuniformity of the exchange interaction in value enhances quantum fluctuations and leads to a second-order antiferromagnet–dimer-state phase transition in δ. The phase boundary of this transition can be approximated by the power law $\Delta = A(\delta - B)^\alpha$ with the fitting parameters $A = 1.0 \pm 0.8$, $B = 0.0 \pm 0.01$, and $\alpha = 0.58 \pm 0.06$.

4. THE THERMODYNAMIC CHARACTERISTICS OF A DIMER STATE AS FUNCTIONS OF TEMPERATURE

For $CuGeO_3$ the anisotropy Δ of the exchange interaction amounts to 0.25, so that from the relationship $\Delta \sim \delta^{0.58}$ we can estimate the critical value of alternation of the exchange interaction along the c axis at $\delta_c \approx 0.09$; above this value a dimer state is realized in $CuGeO_3$. For a given anisotropy of the exchange interaction we can find the dependence of the temperature of the transition from a dimer state to a paramagnetic state on δ and the dependence $E_g(\delta)$ of the energy gap between the singlet and triplet states from the calculated magnetization vs. field-strength curves $M(H)$. By comparing the experimental values T_c^{expt} and $M^{\text{expt}}(H)$ with the calculated values we can find the values of the exchange interaction J_{c0} and alternation δ. The critical temperature T_c, at which all correlations between the dimers disappears, is determined from the four-spin correlation by the maximum in dq/dT. The excitation of the singlet is the triplet with $S^z = 0$, which is the cause of a strong decrease in the correlation function in the transverse components along the c axis as the temperature increases. Along the b axis, where there is no alternation of the exchange interaction, the temperature behavior of correlations between nearest neighbors is similar to the behavior in a magnetically ordered phase.

In the majority of papers discussed in Ref. 29 it is assumed that the main excitation in a dimer state is a spinon, when one dimer is torn apart and a system of two oppositely directed spin is formed, with the spins located at certain distances from each other, which is equivalent to a soliton excitation in a one-dimensional chain of spins. This is true when the chain has a degenerate ground state of singlets at the sites (12)(34)(56)$\cdots$ and of another type (23)(45)(67)$\cdots$ (see Ref. 21). In the model with alternating bonds, an amount of energy ΔE equal to 3δ is lost when the singlet is transferred to a neighboring site. The probability W of triplet formation is proportional to $\exp(-J^{x,y}/T)$. If the average distance between triplets, $p \sim 1/W$, exceeds the characteristic size of a dimer correlation, $p > \xi$, there is no interaction between the triplets. The approximation of the correlation radius depicted in Fig. 2a by power and exponential curve leads to a more complicated excitation pattern. For an aniso-

tropic antiferromagnet ($\delta=0$) above T_N one can observe a typical power-law behavior: $\xi=0.50(5)/(T-0.134(8))$. In the dimer state in the vicinity of a dimer-state–paramagnet transition ($T>0.6T_c$), the approximation $\xi=A/T\ \exp(\delta/T)$ produces the smallest error. Above T_c the temperature dependence of the correlation radius is independent of the alternation of the exchange interaction. Possibly, at T_c there is crossover of the temperature dependence of ξ from exponential to power-like, where the crossover depends on a change in the nature of the excitations.

The gain in energy of the dimer state is achieved due to exchange in the transverse components of the spins, $E-E_{AF}\sim\delta^{1.5}$. Hence the smaller the value of the correlation function $\langle S_0^+ S_1^-\rangle$, the smaller the energy in absolute value. The emergence of a spin moment at a site (say, a triplet with $S^z=0$) fixes the short-range magnetic order but does not substantially change the correlation function in the longitudinal components. Hence the Fourier transforms of this function are practically independent of temperature in the dimer phase, including the Fourier transform in the reciprocal lattice vector, $S^z(Q)$ (Fig. 2b). In an anisotropic antiferromagnet, $S^z(Q)$ drastically varies above the Néel temperature. In $CuGeO_3$, the intensity of inelastic neutron scattering in the vector Q for $T<14$ K is also temperature-independent.[17] Figure 2c depicts the temperature dependence of the intensity of inelastic neutron scattering and the results of Monte Carlo calculations with alternation of the exchange interaction, $J_{0c}=127$ K and $\delta=0.15$. The agreement with experiment becomes better if we take a smaller value of the inhomogeneity of the exchange interaction, $\delta=0.11$. This, however, requires using a larger lattice and hence a larger number of Monte Carlo steps, which is impossible with the current computer. The dependence on δ of the temperature of the dimer-state–paramagnet transition is well approximated by the power function $T_c(\delta)=0.55(4)(\delta-0.082(6))^{0.50(3)}$.

The data on specific heat support the assumption that there is a gap between the ground and excited states (Fig. 3b). The approximation of the temperature dependence of $C(T)$ by a two-level system (of the Schottky type) does not describe the results even qualitatively. The specific heat in a dimer state is well approximated by an exponential function similar to the one for an anisotropic antiferromagnet, $C=A\ \exp(-E_g/2T)$, where E_g is the gap between the ground and excited states. From this approximation it follows that the ratio E_g/T_c is approximately 2.1(1) and agrees well with the experimental data, $E_g^{\rm expt}/T_c^{\rm expt}=2.14$ (see Ref. 4), and so does the temperature behavior of $C(T)$ for $T<T_c$, which is depicted in normalized units in Fig. 3c. Above the transition temperature the susceptibility increases (Fig. 3a), since dimers ($S=0$) still exist and as a result of heating are gradually transformed into triplets. The magnetic state may be considered a dimer gas in a paramagnetic matrix. There are two possible explanations of the experimental data on specific heat and susceptibility above T_c. The first is that the dimer-state–paramagnet transition is of a spin–Peierls type and that the structural transition below $T_c=14$ K alternates the exchange interaction. Then for $T>14$ K the specific heat, the susceptibility and the neutron scattering intensity in $CuGeO_3$ vary with temperature in the same way as they do in

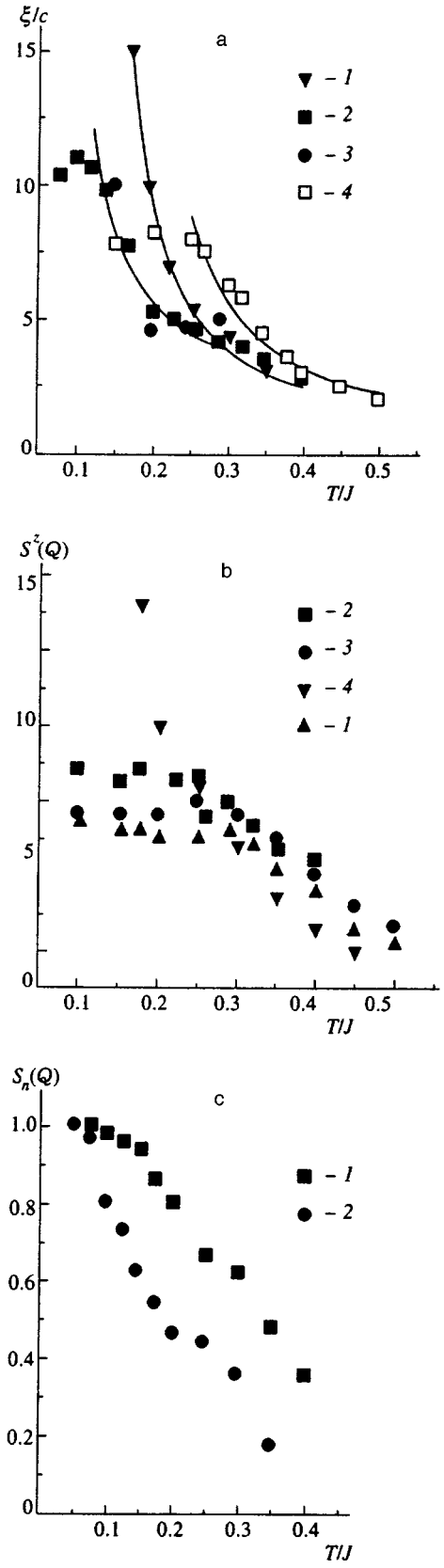


FIG. 2. (a) Temperature dependence of the correlation radius ξ along the c axis for $\delta=0$ (*1*), 0.3 (*2* and *3*), and 0.5 (*4*) at $L=64$ (*2*) and 48 (*3*). (b) Temperature dependence of the Fourier transform $S^z(Q)$ in the reciprocal lattice vector $Q=\pi/c$ for $\delta=0$ (*1*), 0.3 (*2* and *3*), and 0.5 (*4*) at $L=64$ (*2*) and 40 (*3*). (c) Temperature dependence of the normalized intensity of inelastic neutron scattering in $Q=\pi/c$, $S_n(Q)=S(Q,T)/S(Q,T=4\ \text{K})$ (*2*; data taken from Ref. 17), and the results of Monte Carlo calculations for $\delta=0.15$ (*1*) at $J_c=127$ K.

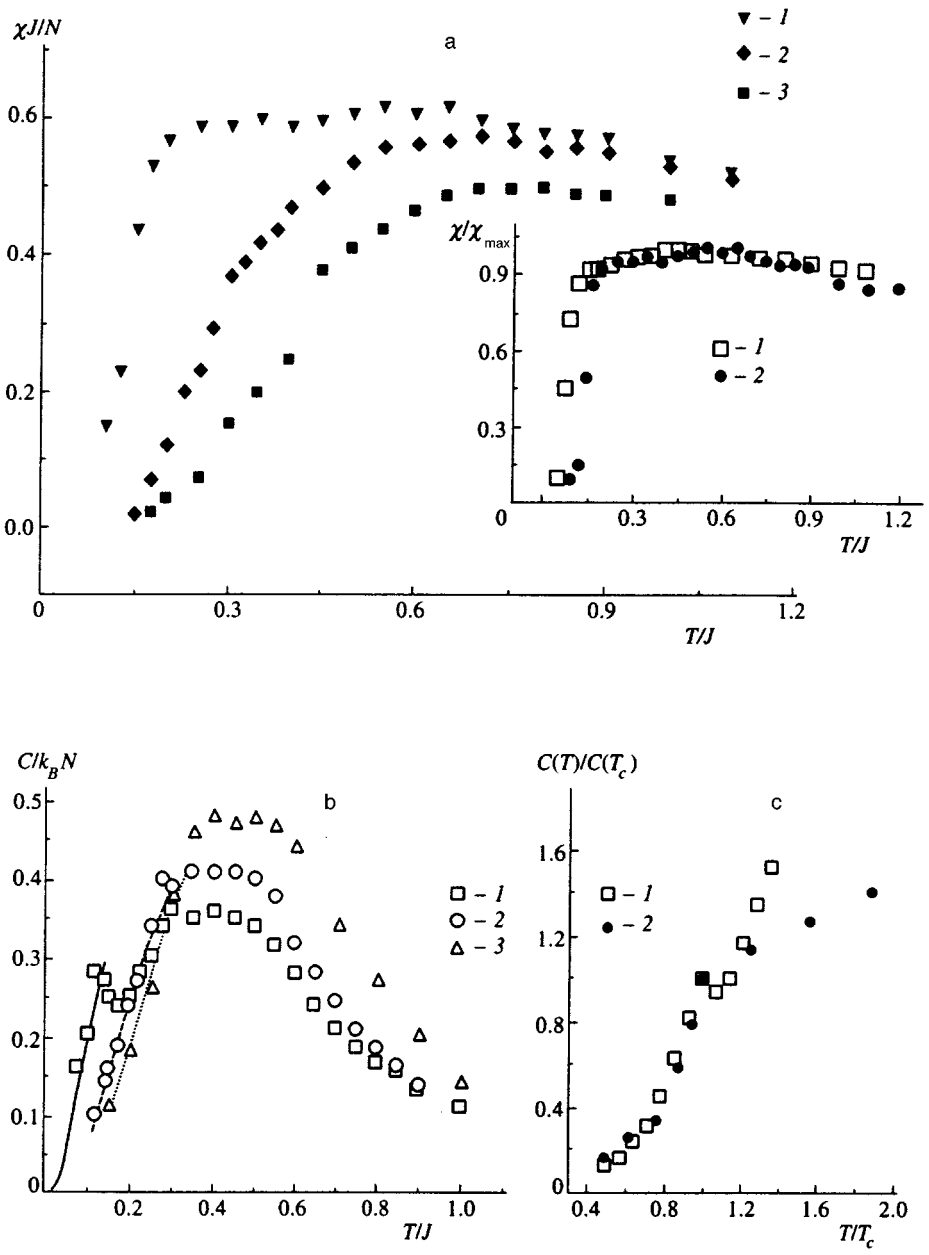


FIG. 3. Temperature dependence of the susceptibility in the field $H/J=0.01$ (a) and the specific heat (b) for $\delta=0$ (*1*), 0.3 (*2*), and 0.5 (*3*). The inset in Fig. 3a depicts the temperature dependence of the normalized susceptibility $\chi/\chi_{\max}$ of $CuGeO_3$ (*1*; data taken from Ref. 5), and the results of Monte Carlo calculations for $\delta=0.15$ (*2*). (c) Temperature dependence of the normalized specific heat $C(T)/C(T_c)$: (*1*) the results of the experiment of Kobayashi *et al.*,[4] and (*2*) the results of Monte Carlo calculations.

an anisotropic antiferromagnet. The second is that the nonlinear anharmonicity of lattice vibrations drives the value of alternation of the exchange interaction up with increasing temperature. In this case the specific heat will increase under heating above $T>14$ K, which qualitatively agrees with the experimental data (Fig. 3c). The temperature behavior of susceptibility calculated for $\delta=0.15$ on the assumption that there is a spin–Peierls transition and that calculated on the assumption that δ is temperature-dependent (corresponding to $T_c(\delta)$, i.e., susceptibility is taken at T_c) coincide when expressed in normalized units $\chi/\chi_{\max}$ and are in good agreement with the experimental data depicted in the inset to Fig. 3a.

The field dependence of the magnetization along the c axis, depicted in Fig. 4, can be used to determine the critical field H_c and the jump in magnetization, ΔM, for different values of δ. For $H>H_c$ the correlated dimer state disappears. The critical field H_c is the field strength at which the derivative of the order parameter of the dimer state, dq/dH, is at its maximum (Fig. 4b) and long-range ferromagnetic order sets in, $\langle S_0^z S_{r=L/2}^z\rangle\neq 0$, although the correlations between the nearest neighbors remain negative up to higher magnetic field strengths (Fig. 4a). Hence in such fields $M(H)$ depends on the alternation of the exchange interaction. The saturation field H_{c2} coincides with the classical value $H_{c2}=2zS(J_c+J_b)$. The magnitude of the jump in magnetization (the inset in Fig. 4c) increases according to a power law: $\Delta M=0.05(\delta-0.08)^{0.5}$. Such a dependence follows from the fact that the size of the region of correlating dimers amounts to $\xi_c\xi_b$ and the emergence of one triplet is this region yields a magnetic moment of order $\sim 2/\xi_c\xi_b$; for instance, at $\delta=0.11$ we have $\xi_c=28$, $\xi_b\approx 5$, and $\Delta M\sim 0.013\mu_B$. Plugging the experimental value $\Delta M=0.025\mu_B$ (see Ref. 8) into the theoretical dependence ΔM vs. δ, we can find the value of alternation of the ex-

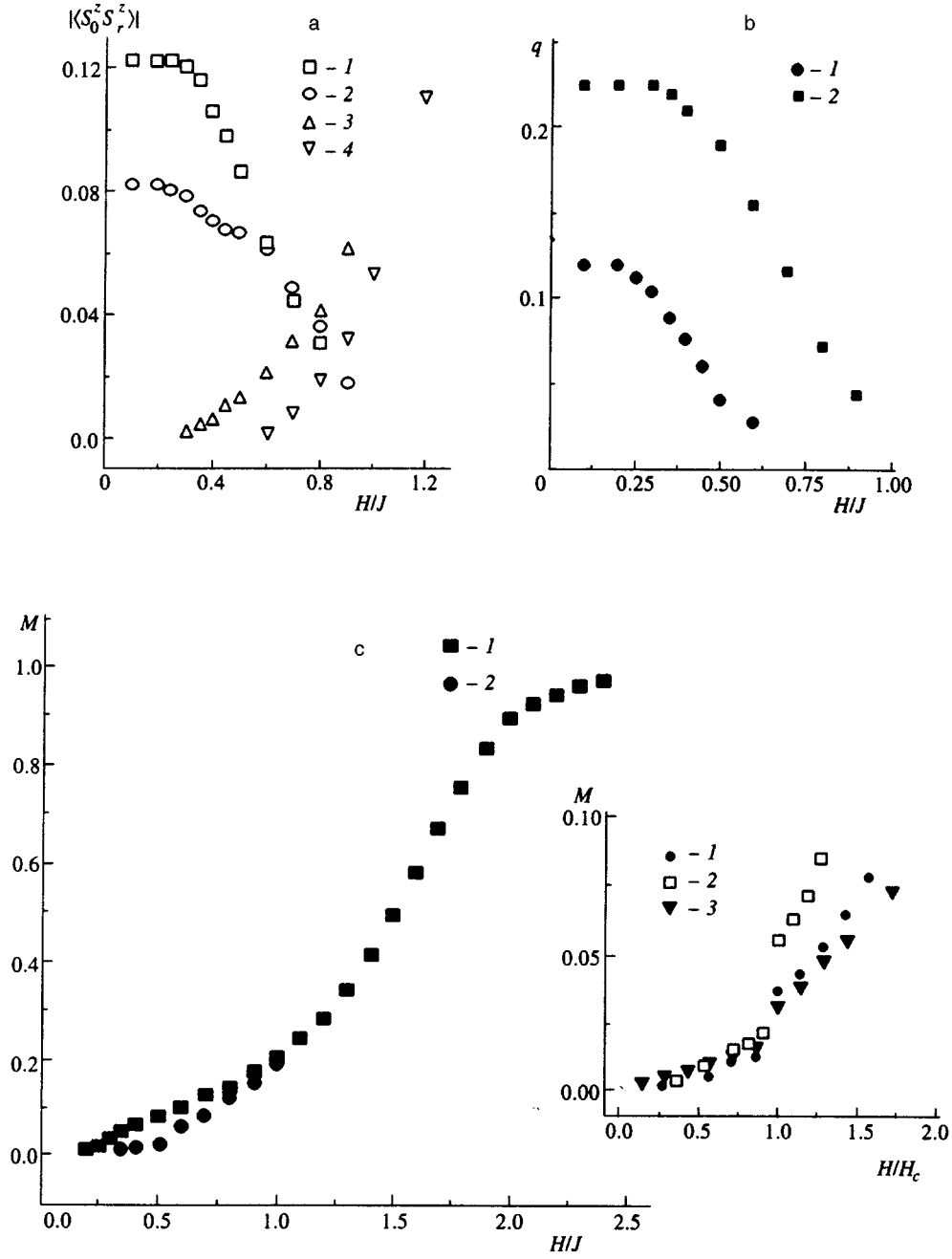


FIG. 4. (a) Field dependence of the spin correlation functions for the dimer state for $\delta=0.15$ (*1* and *3*) and 0.3 (*2* and *4*) at distances $r=1$ (*1* and *2*) and $r=32$ (*3* and *4*) along the *c* axis. (b) Field dependence of the dimerization parameter q for $\delta=0.15$ (*1*) and 0.3 (*2*). (c) Field dependence of magnetization M for $\delta=0.3$ (*1*) and 0.6 (*2*). The inset in Fig. 4c depicts the dependence of M on the reduced field strength H/H_c for $\delta=0.15$ (*1*) and 0.3 (*2*) and the experimental data (*3*) taken from Ref. 8.

change interaction, $\delta=0.11(1)$, in $CuGeO_3$ and the transition temperature $T_c/J=0.11$, which corresponds to the value of the exchange interaction $J_{0c}=127$ K.

The magnetic structure being investigated can help predict effects to which diamagnetic dilution over the sites can lead. For instance, substitution of zinc atoms for copper atoms at the critical concentration $x=0.015-0.02$ (Ref. 30) and substitution of Ga for Cu (Ref. 31) in $CuGeO_3$ lead to an increase in susceptibility in the low-temperature region and to formation of antiferromagnetic order, according to antiferromagnetic resonance data.[32] Substitution of Si atoms for Ge atoms at $x=0.7\%$ also leads to formation of antiferromagnetic order at low temperatures.[33] A silicon atom is smaller than a germanium atom, and according to theoretical estimates[17] the exchange interaction J_c is approximately zero. Since a Ge atom affects two Cu–O–Cu bonds, silicon doping breaks two bonds. When zinc is substituted for copper, bonds along the *c* axis are also broken. This disrupts the motion of dimers and diminishes the entropy contribution to the energy. According to Monte Carlo calculations, the energy of a dimer state at $\delta=0.11$ exceeds the energy of a ferromagnetic state by 3%, which corresponds to an energy of roughly ~ 2 K per spin. Dilution at such concentrations leads to a similar (in order of magnitude) decrease in the energy: $E_{DS}(x)/E_{DS}(0)\approx 1-2x$. In diluted $CuGeO_3$, the *c* axis is the easy-magnetization axis and the spin-flop field amounts to $H_c\approx 1$ to 2 T (Ref. 34). If the impurity does not substantially alter the anisotropy fields, the value of the moment at a site can be estimated from the fact that $H_c/J\sim z\sigma\sqrt{2\Delta}$, i.e., $\sigma\sim 0.01\mu_B$, and the Néel temperature

of a diluted anisotropic antiferromagnet, which is of order of the exchange interaction field, $T_c \sim 2$ K; this is in fairly good agreement with the experimental data.[30,31] There is also another way to estimate the effective moment. For instance, when one spin is removed from the correlating antiferromagnetic region of size $\xi_c \xi_b$, an odd number of spins are left, and the contribution of each such spin to the effective moment is roughly $1/\xi_c \xi_b \sim 0.0072 \mu_B$ at the critical concentration $x_c = 0.7\%$. Lack of oxygen in $CuGeO_3$ can lead to such effects, and because of the disruption of chemical bonds a diamagnetic copper atom may form. Petrakovskiĭ *et al.*[7] showed how thermal treatment can affect the magnetic properties of $CuGeO_3$. Lack of oxygen acts like a paramagnetic impurity: divergence of the susceptibility is observed and the M vs H dependence is nonlinear.

Thus, in a two-dimensional Heisenberg model with anisotropic bonds and alternated exchange interaction a correlating dimer state can form, i.e., the magnetic material is separated into finite regions with antiferromagnetic order, regions characterized by the presence of an energy gap between the ground and excited states. For $CuGeO_3$ we have found the dimensions of such a region along the c and b axes, $\xi_c \approx 28c$ and $\xi_b \approx 5b$, the value of the exchange interaction, $J_c = 127$ K, and the alternation of the exchange interaction, $\delta = 0.11 J_c$, for $T < 14$ K. The paramagnetic susceptibility and the specific heat for $T < T_c$ are well described by this model. The linear increase in the specific heat in the paraphase could be caused by nonlinear anharmonicity; for instance, the nonalternation of the exchange interaction increases because of the nonlinear spin–phonon interaction. The small value of the magnetic moment at the critical value of the magnetic field, H_c, is related to the formation of one triplet $(S = 1)$ in the correlating region whose size is $\xi_c \xi_b \approx 140bc$, and this region can be a soliton. The soliton gas has a critical temperature of the transition to the paraphase, and the Fourier spectrum of the pair spin correlation function in this state is weakly temperature-dependent and qualitatively coincides with the temperature dependence of the integral intensity of inelastic neutron scattering.

This work was supported by the Krasnoyarsk Regional Science Fund (Project No. 6F0004).

[1] M. Nishi, O. Fujita, and J. Akimitsu, Phys. Rev. B **50**, 6508 (1994).
[2] O. Fujita, J. Akimitsu, M. Nishi, and K. Kakurai, Phys. Rev. Lett. **74**, 1667 (1995).
[3] L. C. Bunel, T. M. Brill, I. Zaliznyak, J. P. Boucher, and J. P. Renard, Phys. Rev. Lett. **69**, 1699 (1992).
[4] T. C. Kobayashi, A. Koda, H. Honda, C. U. Hong, K. Amaya, T. Asano, Y. Ajiro, M. Mekato, and T. Yosida, Physica B **211**, 205 (1995).
[5] G. A. Petrakovskiĭ, K. A. Sablina, A. M. Vorotynov, A. I. Kruglik, A. G. Klimenko, A. D. Balayev, and S. S. Aplesnin, Zh. Éksp. Teor. Fiz. **98**, 1382 (1990) [Sov. Phys. JETP **71**, 772 (1990)].
[6] M. Hase, I. Terasaki, and K. Uchinokura, Phys. Rev. Lett. **70**, 3651 (1993).
[7] G. A. Petrakovskiĭ, A. I. Pankrats, K. A. Sablina, A. M. Vorotynov, D. A. Velikanov, and A. D. Vasil'ev, H. Szymczak, and S. Kolesnik, Fiz. Tverd. Tela (St. Petersburg) **38**, 1857 (1996) [Phys. Solid State **38**, 1025 (1996)].
[8] H. Hory, M. Furusawa, S. Sugai, M. Honda, T. Takeuchi, and K. Kindo, Physica B **211**, 180 (1995).
[9] J. P. Pouget, L. R. Regnault, M. Arn, B. Hennion, J. P. Renard, V. Veillet, G. Dhalenne, and A. Revcolevschi, Phys. Rev. Lett. **72**, 4037 (1994).
[10] Q. J. Harris, Q. Feng, R. J. Birgerneau, H. Hirota, K. Kakurai, J. E. Lorenzo, G. Shirane, M. Hase, K. Uchinokura, H. Kojima, I. Tanaka, and Y. Shibuya, Phys. Rev. B **50**, 12606 (1994).
[11] M. Nishi, O. Fujita, J. Akimitsu, K. Kakurai, and Y. Fujii, Phys. Rev. B **52**, R6959 (1995-II).
[12] T. Wei, A. J. Heeger, M. B. Salamon, and G. E. Delker, Solid State Commun. **21**, 595 (1977).
[13] J. Riera and A. Dobry, Phys. Rev. B **51**, 16098 (1995).
[14] G. Castilla, S. Chakravarty, and V. J. Emery, Phys. Rev. Lett. **75**, 1823 (1995).
[15] Y. Nonomura and M. Suzuki, J. Phys. A **27**, 1127 (1994).
[16] H. Nojiri, Y. Shimamoto, N. Miura, M. Hase, K. Uchinokura, H. Kojima, I. Tanaka, and Y. Shibuya, Physica B **211**, 184 (1995).
[17] B. Roessli, P. Fisher, J. Schefer, W. Buhrer, A. Furrer, T. Vogt, G. Petrakovskii, and K. Sablina, J. Phys.: Condens. Matter **6**, 8469 (1994).
[18] M. Nishi, O. Fujita, and J. Akimitsu, Technical Report of ISSR, Ser. A **2759**, 1 (1993).
[19] K. Hirota, D. E. Cox, J. E. Lorenzo, G. Shirane, J. M. Tranquada, M. Hase, K. Uchinokura, H. Kojima, Y. Shibuja, and I. Tanaka, Phys. Rev. Lett. **73**, 736 (1994).
[20] S. S. Aplesnin, Fiz. Tverd. Tela (St. Petersburg) **38**, 1868 (1996) [Phys. Solid State **38**, 1031 (1996)].
[21] D. Khomskii, W. Geertsma, and M. Mostovoy, ''Elementary excitations exchange interaction and spin-Peierls transition in $CuGeO_3$,'' Invited talk at the 21st Int. Conf. on Low Temperature Physics, Prague, August 8–14, 1996, submitted to Czech. J. Phys.
[22] M. Hase, L. Terasaki, and K. Uchinokura, Phys. Rev. Lett. **70**, 3651 (1993).
[23] H. Raedt and A. Lagendijk, Phys. Rep. **127**, 233 (1985).
[24] M. Suzuki, J. Stat. Phys. **43**, 883 (1986).
[25] M. Suzuki, Prog. Theor. Phys. **56**, 1454 (1976).
[26] T. Tsuzuki, Prog. Theor. Phys. **73**, 1352 (1985).
[27] J. E. Hirsch and R. L. Sugar, Phys. Rev. B **26**, 5039 (1982).
[28] H. A. Schulz and T. A. Ziman, Europhys. Lett. **18**, 355 (1992).
[29] Yu. A. Izyumov, M. I. Katsnel'son, and Yu. N. Skryabin, *The Magnetism of Collectivized Electrons* [in Russian], Nauka, Moscow (1994).
[30] K. Uchinokura, M. Hase, and Y. Sasago, Physica B **211**, 175 (1995).
[31] G. A. Petrovskiĭ, A. M. Vorotynov, K. A. Sablina, A. I. Pankrats, and D. A. Velikanov, Fiz. Tverd. Tela (St. Petersburg) **38**, 3430 (1996) [Phys. Solid State **38**, 1871 (1996)].
[32] A. I. Smirnov, V. N. Glazkov, A. N. Vasil'ev, L. I. Leonyuk, S. M. Coad, D. McK. Paul, G. Dhalenne, and A. Revcolevschi, JETP Lett. **64**, 305 (1996).
[33] J. P. Renard, Europhys. Lett. **30**, 475 (1996).
[34] M. Poirier, R. Beaudry, M. Castonguay, M. L. Plumer, G. Quirion, F. S. Razavi, A. Revcolevschi, and G. Dhalenne, Phys. Rev. B **52**, R6971 (1995-II).

Translated by Eugene Yankovsky

Anomalous Hall effect in granular alloys

A. V. Vedyaev, A. B. Granovskiĭ, and A. V. Kalitsov

M. V. Lomonosov Moscow State University, 119899 Moscow, Russia

F. Brouers

Institut de Physique, Sart Tilman, Liège 4000, Belgium
(Submitted 7 April 1997)
Zh. Éksp. Teor. Fiz. **112**, 2198–2209 (December 1997)

A theoretical study is performed of the anomalous Hall effect in granular alloys with giant magnetoresistance. The calculation is carried out within the Kubo formalism and the Green's function method. The mechanism of asymmetric scattering of the spin-polarized current carriers is considered with allowance for a size effect associated with scattering not only by one grain, but also with more complicated processes of transport among two and three grains. It is shown that scattering of conduction electrons by the interfaces of the grains and the matrix has a substantial effect on the magnitude of the anomalous Hall effect and determines its sign. In general, correlation between the quantities ρ_H and ρ^2 is absent, where ρ_H is the Hall resistivity and ρ is the total resistivity of the granular alloy. However, numerical calculation shows that for certain values of the model parameters $\rho_H \sim \rho^{3.8}$ and for these same parameter values the amplitude of the giant magnetoresistance reaches 40%, which is found to be in quantitative agreement with the experimental data for $Co_{20}Ag_{80}$ alloys [P. Xiong, G. Xiao, J. Q. Wang *et al.*, Phys. Rev. Lett. **69**, 3220 (1992)]. It is also shown that increasing the resistivity of the matrix leads to a significant growth in the anomalous Hall effect, and more substantial growth for alloys with small grain size, which is in good agreement with experiment [A. B. Pakhomov, X. Yan, and Y. Xu, J. Appl. Phys. **79**, 6140 (1996); [X. N. Jing, N. Wang, and A. B. Pakhomov, Phys. Rev. B **53**, 14032 (1996)]. © *1997 American Institute of Physics.* [S1063-7761(97)02012-X]

1. INTRODUCTION

Magnetic granular alloys are inhomogeneous magnets in which ferromagnetic metallic grains similar in size to a single-domain magnet are randomly distributed in the paramagnetic or diamagnetic matrix of a metal or insulator. Although the study of electron transport phenomena in such systems was begun more than 20 years ago,[1] the many anomalies in the dependence of the various kinetic effects (resistivity, thermal conductivity, thermoelectric power, Hall effect) on the magnetic field, concentration and size of the grains, temperature, etc. in both metal–metal alloys and metal–insulator composites were only recently discovered (see, for example, Refs. 2–6).

The presence of spin-dependent scattering of conduction electrons in the bulk and on the surface of the grains, spin-dependent potential barriers at the intergranular contacts, along with classical and quantum size effects, all of which are characteristic of magnetically inhomogeneous systems, together with the inhomogeneity in the size and shape of the grains are the reason, on the one hand, for the wide variety of observed properties and, on the other, for substantial difficulties in their theoretical interpretation.

By now, the main mechanisms of giant magnetoresistance in metal–metal granular alloys have been reliably established and explained. In a number of papers (see, for example, Refs. 7–9), it has been shown that giant magnetoresistance in granular alloys is of the same nature as in multilayer structures and spin-valve sandwiches, i.e., it is related to spin-dependent scattering of electrons in the bulk and on the surface of the grains. The simple approach to the description of transport phenomena in these alloys proposed by Zhang and Levy[8] is based on the concept of self-averaging of the probability of scattering of spin-polarized current carriers, and is valid provided $l \gg r_0$, where l is the mean free path of the electrons and r_0 is the radius of the grains or intergranular distance. For real granular alloys with giant magnetoresistance this condition is obviously not met. Recently, Vedyaev *et al.*[9] proposed a new quantum-statistical approach to the description of giant magnetoresistance in granular alloys, based on the Kubo formalism and the Green's function method, which does not make use of the concept of self-averaging and allows one to explain the observed dependence of giant magnetoresistance on the size of the grains.

A no less interesting effect, revealing unusual behavior in granular alloys, is the anomalous Hall effect. The Hall field in ferromagnets can be written in the form[10,11]

$$E_y = R_0 B_z j_x + 4\pi R_s M_z j_x, \tag{1}$$

where B_z is the magnetic induction, M_z is the magnetization, j_x is the current density, R_0 is the coefficient of the normal Hall effect, due to the action of the Lorentz force, and R_s is the coefficient of the anomalous Hall effect. The latter coefficient is associated with the effect of spin–orbit coupling on the scattering of spin-polarized current carriers. For homogeneous crystalline ferromagnets with small-to-moderate impurity content, the anomalous Hall coefficient R_s^b, due to the grain material, is described well by the relation

$$R_s^b = a\rho + b\rho^2, \tag{2}$$

where ρ is the resistivity and a and b are constants of arbitrary sign, where the first term is associated with skew scattering and the second with both skew scattering and the side-jump mechanism.[11] The coefficient R_s^b is usually called the anomalous Hall constant,[10,11] since it is independent of the magnetization and the magnetic field. It can be taken for granted that the main anomalous Hall carriers in alloys of transition metals are d-like electrons, since spin–orbit coupling is significantly greater for them than it is for s-like electrons, and the relatively small mean free path of the d states presents no obstacle to the appearance of the anomalous Hall effect.[11]

Experimental studies of the anomalous Hall effect in metal–metal and metal–insulator granular alloys have uncovered a number of facts that have not found even qualitative explanation within the framework of the well-developed theory of this effect in homogeneous alloys or in the effective medium theory for composites.[12] We note some of these results.

1. In the system $Co_{20}Ag_{80}$, after annealing, which alters the size of the grains and the amplitude of giant magnetoresistance, instead of a correlation of type (2) for the residual (4.2 K) anomalous Hall effect, the relation $R_s \sim \rho^{3.7}$ is observed.[4] For the granular alloys $(Co_{70}Fe_{30})_xAg_{1-x}$ it has also been established that a correlation of type (2) does not hold, either for variation of the grain concentration or for variation of the temperature.[13]

2. At 77 K, the Hall coefficient $R_s > 0$ for the granular alloys $(Co_{70}Fe_{30})_xAg_{1-x}$ whereas for the homogeneous alloys $Co_{20}Ag_{80}$, $R_s^b < 0$ (Ref. 13).

3. The coefficient R_s for the granular alloys $Co_{20}Ag_{80}$ depends markedly on the magnitude of the magnetic field, and in a nonmonotonic fashion.[3]

4. In granular metal–insulator alloys the coefficient R_s grows abruptly with decreasing concentration of ferromagnetic metallic grains, and reaches giant values (three to four orders of magnitude greater than R_s^b) near the percolation threshold.[5,6]

Efforts recently undertaken[14,15] to explain some of these results for alloys with giant magnetoresistance in terms of the Zhang–Levy model[8] were based on the concept of self-averaging and were of a purely qualitative character. In particular, within the framework of the Zhang–Levy model, it is not possible, for the same values of the parameters characterizing scattering from the grain surfaces, to explain correlation of the type $R_s \sim \rho^{3.7}$ and giant magnetoresistance amplitudes of up to 40% in $Co_{20}Ag_{80}$ alloys at low temperature.

In the present paper, within the framework of the Kubo formalism and the Green's function method, we construct a theory of the residual anomalous Hall effect of granular alloys. Of primary concern for us will be the effect of scattering at the grain–matrix interface on the possible types of correlation between R_s and ρ. As in previous studies on the theory of the anomalous Hall effect in inhomogeneous alloys,[11] we restrict the discussion to the mechanism of skew scattering and show that in this case it is possible to explain most of the observed effects.

2. STATEMENT OF THE PROBLEM AND CALCULATION OF THE HALL CONDUCTIVITY

By definition (1), the anomalous Hall coefficient of a granular alloy is

$$R_s = \frac{\sigma_{xy}}{4\pi M_z}\,\rho^2, \tag{3}$$

where σ_{xy} is the off-diagonal, linear (in the spin–orbit coupling) part of the conductivity tensor (expression (3) assumes that $\sigma_{xy} \ll \sigma_{xx}$ and $\sigma_{xx} = \sigma_{yy} = \rho^{-1}$). For systems with giant magnetoresistance, the quantity ρ obviously depends on the magnetic field, but the ratio σ_{xy}/M_z can also be a function of the magnetic field by virtue of the fact that a magnetic field alters the character of spin-dependent scattering in the system. The effect of a magnetic field on R_s is considered in Ref. 16, but here we calculate the coefficient R_s and the anomalous Hall resistivity $\rho_H = 4\pi R_s M_z$ for the state of magnetic saturation, for which, as a rule, the anomalous Hall effect is observed.

The granular alloy model proposed here consists in the following. We assume that all ferromagnetic grains are identical spheres with radius r_0; the type, concentration, and distribution of the impurities in the bulk of the grains and near the grain–matrix interfaces may vary; both the d-like and s-like conduction electrons are characterized by a quadratic dispersion law, and their scattering by impurities is weak, $(k_F l)^{-1} \ll 1$; the effect of potential barriers on passage of an electron through an interface is not considered.

For weak scattering and a quadratic dispersion law, the Kubo formula for the antisymmetric part of the conductivity tensor takes the form

$$\sigma_{xy} = \frac{e^2\hbar}{4\pi\Omega}\,\mathrm{Tr}\langle v_x \tilde{G}^{\mathrm{ret}}(\epsilon) v_y \tilde{G}^{\mathrm{adv}}(\epsilon)\rangle_{\epsilon=\epsilon_F}, \tag{4}$$

where $\tilde{G}$ is the total Green's function with allowance for spin–orbit coupling, and angular brackets denote quantum-statistical averaging. In the linear approximation in the spin–orbit coupling

$$\tilde{G} = G + GTG + GH^{so}G, \tag{5}$$

where T is the matrix operator of isotropic scattering by impurities, and H^{so} is the spin–orbit coupling operator, whose matrix elements have the form[11]

$$H^{so}_{\mathbf{k}\mathbf{k}'} = \frac{1}{N}\sum_l e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{R}_l} H_l^{so}(\mathbf{k}-\mathbf{k}'), \tag{6}$$

$$H_l^{so}(\mathbf{k}-\mathbf{k}') = is(\mathbf{k}\cdot\mathbf{k}')_z a_0^2 \frac{M_z}{M_s}\,\lambda_l^{so}(\mathbf{k}-\mathbf{k}'), \tag{7}$$

where a_0 is the lattice parameter and λ_l is the spin–orbit coupling parameter for the impurity located at the site $\mathbf{R}_l$; λ_l is proportional to the scattering potential of this impurity V_l (Ref. 11), and the index s is equal to $+1(-1)$ for states with spin aligned (anti-aligned) with the magnetization (up-spin and down-spin).

Then it follows from Eqs. (4)–(7) that to lowest order in the scattering potential V_l

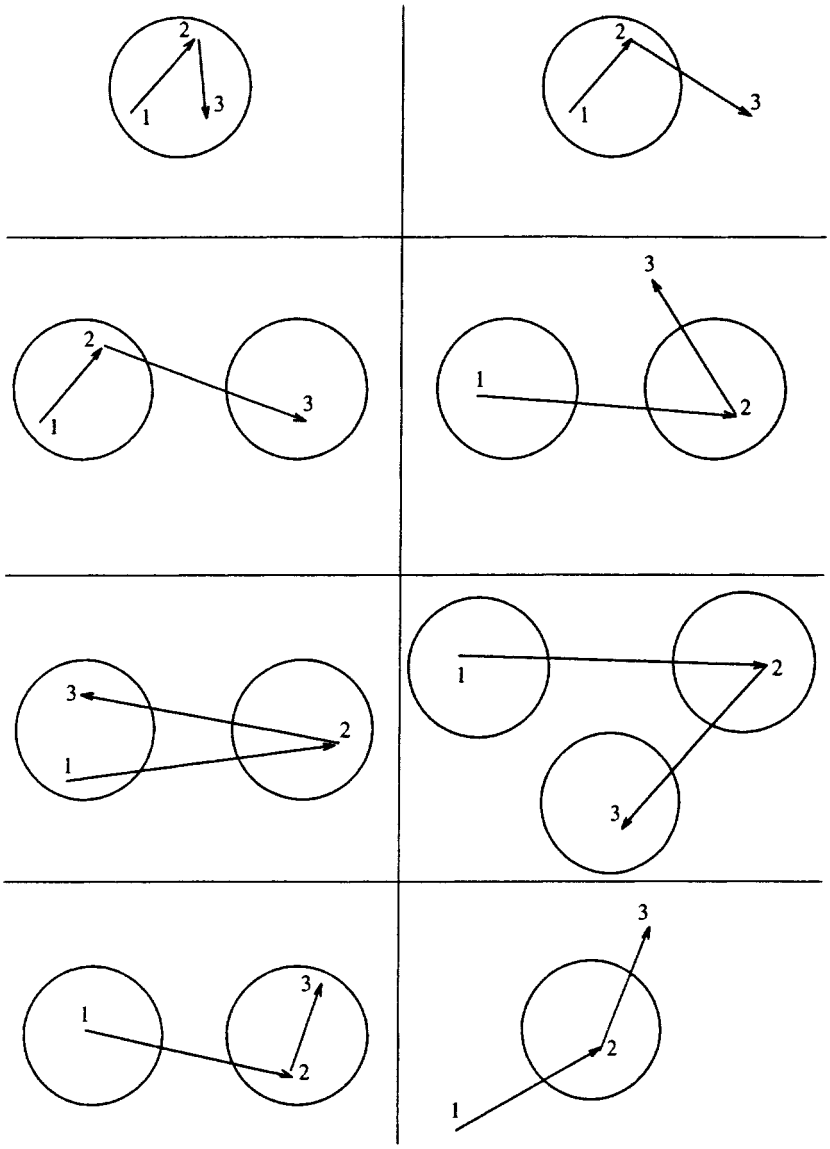


FIG. 1. Diagrams for calculation of the Hall conductivity σ_{xy} of a granular alloy.

$$\sigma_{xy}^{(s)} = -2sa_0^2 \frac{M_z}{M_s} \frac{e^2\hbar}{4\pi\Omega} \frac{\hbar^2}{m^2} N(\epsilon_F) \sum_{\mathbf{k},\mathbf{k}',\mathbf{k}''} k_x^2 k_y'^2 |G_k|^2 \times |G_{k'}|^2 \langle\langle \lambda^{so}_{(l)\mathbf{k},\mathbf{k}'} V^l_{\mathbf{k}'\mathbf{k}''} V^l_{\mathbf{k}''\mathbf{k}} \rangle\rangle_{k=k'=k''=k_F}, \quad (8)$$

where $N(\epsilon_F)$ is the density of states at the Fermi level and the double brackets denote configurational averaging. In what follows we will consider the contribution to σ_{xy} of only one subband, e.g., the subband of states with up-spin. Transforming to the r representation, we obtain an expression for the nonlocal Hall conductivity of a granular alloy:

$$\sigma_{xy}(\mathbf{r},\mathbf{r}') = -2a_0^2 \frac{M_z}{M_s} \frac{e^2\hbar\Omega}{4\pi} \frac{\hbar^2}{m^2} \left(\frac{\pi\hbar}{4e^2}\right)^2 \left(\frac{2m}{\hbar^2}\right)^4 N(\epsilon_F) \times \int d\mathbf{r}'' \sigma_{xx}(\mathbf{r},\mathbf{r}'') \langle\langle \lambda_l^{so}(\mathbf{r}'') V_l^2(\mathbf{r}'') \rangle\rangle \sigma_{xx}(\mathbf{r}'',\mathbf{r}'), \quad (9)$$

where $\sigma_{xx}(\mathbf{r},\mathbf{r}')$ is the nonlocal conductivity of the alloy. The method of calculating $\sigma_{xx}(\mathbf{r},\mathbf{r}')$ is explained in detail in Ref. 9, and is based on finding the Green's functions $G(\mathbf{r},\mathbf{r}')$, where the points $\mathbf{r}$ and $\mathbf{r}'$ are found either in the same grain, in the matrix, in neighboring grains, or finally, in a grain and in the matrix. The calculation of $\sigma_{xy}(\mathbf{r},\mathbf{r}')$ is more complicated, since the integrand in Eq. (9) is a function of the coordinates of three points $(\mathbf{r},\mathbf{r}',\mathbf{r}'')$ and is therefore defined by the sum of the contributions of the eight different scattering processes depicted in Fig. 1.

We stress that since we assume that the matrix is nonmagnetic, $H^{so}=0$ for an impurity in the matrix, by virtue of which the point $\mathbf{r}''$ can lie only in a ferromagnetic grain (in its bulk or on its surface). Therefore, taking advantage of the form of the Green's functions and the technique for calculating $\sigma_{xx}(\mathbf{r},\mathbf{r}')$ used in Ref. 9, it is possible to calculate all eight contributions (Fig. 1); then, integrating over $\mathbf{r}''$ and $\mathbf{r}'$ in expression (9) and averaging over $\mathbf{r}$, we find the contributions of the grains and the matrix to σ_{xy}. The contribution σ_{xy}^m of the grains is defined by the diagrams (Fig. 1) in which point 1 is found in a grain, and the contribution σ_{xy}^{nm} of the

matrix—by the diagrams in which point 1 lies in the matrix. Finally, these two contributions are different for impurities in the bulk of the grains ($\sigma_{xy}^{m(b)}, \sigma_{xy}^{nm(b)}$) and on their surface ($\sigma_{xy}^{m(s)}, \sigma_{xy}^{nm(s)}$). The final expressions for the corresponding contributions have the form

$$\sigma_{xy}^{m} = \sigma_{xy}^{m(b)} + \sigma_{xy}^{m(s)}, \tag{10}$$

$$\sigma_{xy}^{nm} = \sigma_{xy}^{nm(b)} + \sigma_{xy}^{nm(s)}, \tag{11}$$

where

$$\sigma_{xy}^{m(b)} = -2\frac{a_0^2 M_z}{M_s}\frac{e^2\hbar\Omega k_F^4}{4\pi}\frac{\hbar^2}{m^2}\frac{\lambda_m^{so}(1-2x_m)}{\sqrt{x_m(1-x_m)}}\left[\frac{\epsilon_F}{\pi k_F N(\epsilon_F) l_m}\right]^{3/2}\left(\frac{2m}{\hbar^2}\right)^4 N(\epsilon_F)\left\{ l_m^2 - \frac{3l_m^4 I_0}{4r_0}F^2\left(\frac{2r_0}{l_m}\right) + \frac{3l_m^3 l_{nm} I}{8r_0^2}\right.$$
$$\times\exp\left(-\frac{2r_0}{l_m}\right)F\left(\frac{2r_0}{l_m}\right) + \frac{3N_1 l_m^3 I}{16 r_0 R^2}\exp\left(2\left(\frac{1}{l_{nm}}-\frac{1}{l_m}\right)r_0\right)\exp\left(-\frac{2R}{l_{nm}}\right)F\left(\frac{2r_0}{l_m}\right)\left[l_m^2 F\left(\frac{2r_0}{l_m}\right)\exp\left(2\left(\frac{1}{l_{nm}}-\frac{1}{l_m}\right)r_0\right)\right.$$
$$\left. - l_{nm}^2 F\left(\frac{2r_0}{l_{nm}}\right)\right] + \frac{3l_m^5 l_{nm} N_1}{64 r_0^2 R^2}\exp\left(-\frac{2R}{l_{nm}}\right)\exp\left(2\left(\frac{2}{l_{nm}}-\frac{3}{l_m}\right)r_0\right)F\left(\frac{2r_0}{l_m}\right)F_1\left(\frac{4r_0}{l_m}\right) + \frac{N_1 r_0^2}{R^4}\exp\left(-\frac{4R}{l_{nm}}\right)$$
$$\times\exp\left(4\left(\frac{1}{l_{nm}}-\frac{1}{l_m}\right)r_0\right)\left[l_m^4 F^2\left(\frac{2r_0}{l_m}\right)\exp\left(4\left(\frac{1}{l_{nm}}-\frac{1}{l_m}\right)r_0\right) - l_{nm}^4 F^2\left(\frac{2r_0}{l_{nm}}\right)\right] + \frac{3N_2 l_m^2 I_1}{16\pi r_0 R^4}\exp\left(-\frac{4R}{l_{nm}}\right)$$
$$\times\exp\left(6\left(\frac{1}{l_{nm}}-\frac{1}{l_m}\right)r_0\right)F\left(\frac{2r_0}{l_m}\right)\left[l_m^2 F\left(\frac{2r_0}{l_m}\right)\exp\left(2\left(\frac{1}{l_{nm}}-\frac{1}{l_m}\right)r_0\right) - l_{nm}^2 F\left(\frac{2r_0}{l_{nm}}\right)\right] + \frac{3N_1 l_m^6}{64 r_0^2 R^2}\exp\left(-\frac{2R}{l_{nm}}\right)$$
$$\left.\times\exp\left(4\left(\frac{1}{l_{nm}}-\frac{1}{l_m}\right)r_0\right)\left(1-\exp\left(-\frac{2r_0}{l_m}\right)\right)F\left(\frac{2r_0}{l_m}\right)F_1\left(\frac{4r_0}{l_m}\right)\right\}, \tag{12}$$

$$\sigma_{xy}^{m(s)} = -2\frac{a_0^3 M_z}{M_s}\frac{e^2\hbar\Omega k_F^4}{4\pi}\frac{\hbar^2}{m^2}\frac{\lambda_s^{so}(1-2x_s)}{\sqrt{x_s(1-x_s)}}\left[\frac{\epsilon_F}{\pi k_F N(\epsilon_F) l_s}\right]^{3/2}\left(\frac{2m}{\hbar^2}\right)^4 N(\epsilon_F)\left\{\frac{3l_m^4}{4r_0^3}\exp\left(-\frac{4r_0}{l_m}\right)F^2\left(\frac{2r_0}{l_m}\right) + \frac{3l_m^3 l_{nm}}{8r_0^3}\right.$$
$$\times\exp\left(-\frac{2r_0}{l_m}\right)\left(1-\exp\left(-\frac{4r_0}{l_m}\right)\right)F\left(\frac{2r_0}{l_m}\right) + \frac{3N_1 l_m^3}{16 r_0^2 R^2}\exp\left(2\left(\frac{1}{l_{nm}}-\frac{1}{l_m}\right)r_0\right)\exp\left(-\frac{2R}{l_{nm}}\right)F\left(\frac{2r_0}{l_m}\right)\left(1-\exp\left(-\frac{4r_0}{l_m}\right)\right)$$
$$\times\left[l_m^2 F\left(\frac{2r_0}{l_m}\right)\exp\left(2\left(\frac{1}{l_{nm}}-\frac{1}{l_m}\right)r_0\right) - l_{nm}^2 F\left(\frac{2r_0}{l_{nm}}\right)\right] + \frac{3N_1 l_m^4 l_{nm}}{8 r_0^2 R^2}\exp\left(-\frac{2R}{l_{nm}}\right)\exp\left(2\left(\frac{2}{l_{nm}}-\frac{3}{l_m}\right)r_0\right)$$
$$\times\sinh^2\left(\frac{2r_0}{l_m}\right)F\left(\frac{2r_0}{l_m}\right) + \frac{3r_0 N_1}{4R^4}\exp\left(-\frac{4R}{l_{nm}}\right)\exp\left(4\left(\frac{1}{l_{nm}}-\frac{1}{l_m}\right)r_0\right)\left[l_m^4 F^2\left(\frac{2r_0}{l_m}\right)\exp\left(4\left(\frac{1}{l_{nm}}-\frac{1}{l_m}\right)r_0\right) - l_{nm}^4 F^2\left(\frac{2r_0}{l_{nm}}\right)\right]$$
$$+ \frac{3N_2 l_m^2 I_2}{16\pi r_0 R^4}\exp\left(6\left(\frac{1}{l_{nm}}-\frac{1}{l_m}\right)r_0\right)\exp\left(-\frac{4R}{l_{nm}}\right)F\left(\frac{2r_0}{l_m}\right)\left[l_m^2 F\left(\frac{2r_0}{l_m}\right)\exp\left(2\left(\frac{1}{l_{nm}}-\frac{1}{l_m}\right)r_0\right) - l_{nm}^2 F\left(\frac{2r_0}{l_{nm}}\right)\right]$$
$$\left.+ \frac{3N_1 l_m^5}{8 r_0^2 R^2}\exp\left(4\left(\frac{1}{l_{nm}}-\frac{1}{l_m}\right)r_0\right)\exp\left(-\frac{2R}{l_{nm}}\right)\left(1-\exp\left(-\frac{2r_0}{l_m}\right)\right)\sinh^2\left(\frac{2r_0}{l_m}\right)F\left(\frac{2r_0}{l_m}\right)\right\}. \tag{13}$$

$$\sigma_{xy}^{nm(b)} = -2\frac{a_0^2 M_z}{M_s}\frac{e^2\hbar\Omega k_F^4}{4\pi}\frac{\hbar^2}{m^2}\frac{\lambda_m^{so}(1-2x_m)}{\sqrt{x_m(1-x_m)}}\left[\frac{\epsilon_F}{\pi k_F N(\epsilon_F) l_m}\right]^{3/2}\left(\frac{2m}{\hbar^2}\right)^4 N(\epsilon_F)\frac{3l_m^3 l_{nm}^2}{32((R-r_0)^3 - r_0^3)}$$
$$\times\exp\left[4\left(\frac{1}{l_{nm}}-\frac{1}{l_m}\right)r_0\right]F_1\left(\frac{4r_0}{l_m}\right)\left[\exp\left(-\frac{4r_0}{l_{nm}}\right)-\exp\left(-\frac{2R}{l_{nm}}\right)\right], \tag{14}$$

$$\sigma_{xy}^{nm(s)} = -2\frac{a_0^3 M_z}{M_s}\frac{e^2\hbar\Omega k_F^4}{4\pi}\frac{\hbar^2}{m^2}\frac{\lambda_s^{so}(1-2x_s)}{\sqrt{x_s(1-x_s)}}\left[\frac{\epsilon_F}{\pi k_F N(\epsilon_F) l_s}\right]^{3/2}\left(\frac{2m}{\hbar^2}\right)^4 N(\epsilon_F)\frac{3l_m^2 l_{nm}^2}{4((R-r_0)^3 - r_0^3)}$$
$$\times\exp\left[4\left(\frac{1}{l_{nm}}-\frac{1}{l_m}\right)r_0\right]\sinh^2\left(\frac{2r_0}{l_m}\right)\left[\exp\left(-\frac{4r_0}{l_{nm}}\right)-\exp\left(-\frac{2R}{l_{nm}}\right)\right], \tag{15}$$

where

$$I=\int_0^{r_0}\left[1-\exp\left(-\frac{4\rho}{l_m}\right)\right]\rho^{-1}d\rho,$$

$$I_0=\int_{r_0}^{\infty}\exp\left(-\frac{4\rho}{l_m}\right)\rho^{-2}d\rho,$$

$$I_1=2\pi\int_{-1}^{1}d(\cos\ \gamma)\times\int_0^{r_0}\exp\left(-\frac{2\rho[\cos((\pi/3)-\gamma)-\cos\ \gamma]}{l_m}\right)\rho^2 d\rho,$$

$$I_2=2\pi r_0^2\int_{-1}^{1}d(\cos\ \gamma)\times\exp\left(-\frac{2r_0[\cos((\pi/3)-\gamma)-\cos\ \gamma]}{l_m}\right),$$

$$N_1=2c\left(3\left(\frac{4\pi}{3c}\right)^{2/3}+1\right),\quad N_2=\frac{9c}{2}\left(\frac{4\pi}{3c}\right)^{1/3},$$

$$F(x)\equiv\cosh\ x-\frac{\sinh\ x}{x},\quad F_1(x)\equiv\cosh\ x-x,$$

l_{nm}, l_m, and l_s are the mean free path of the electron in the matrix, in the grains, and on the surface of the grains, respectively, R is the mean intergranular distance, c is the volume concentration of the grains, x_m and x_s are the impurity concentrations, and λ_m^{so} and λ_s^{so} the spin–orbit coupling constants in the bulk and on the surface of the grains, respectively. To find the resulting Hall conductivity of the granular alloy, because of the different conductivity of the grains and matrix it is necessary to use the effective-field approximation.[12] The desired expression then takes the form

$$\sigma_{xy}=\left\langle\left\langle\frac{\sigma_{xy}^i}{(\sigma^i+2\sigma)^2}\right\rangle\right\rangle\langle\langle(\sigma^i+2\sigma)^{-2}\rangle\rangle^{-1},\qquad(16)$$

where $i=m,nm$. In general, the Hall conductivity σ_{xy} and the ordinary conductivity can be mediated by different groups of electrons, so we denote the corresponding parameters in the expression for σ_{xy}^i by the subscript (d), and in the ordinary conductivity σ^i by the subscript (s). It is convenient to normalize the results of calculation for the Hall conductivity of a granular alloy $\rho_H=4\pi R_s M_z$ by the corresponding value $\rho_H^b=4\pi R_s^b M_z^b$ for a homogeneous massive ferromagnetic material in which the type, concentration, and distribution of the impurities are the same as in the bulk of the grains; here we also assume that $M_z=cM_z^b$.

3. RESULTS OF CALCULATION; DISCUSSION

As follows from an analysis of the derived analytic expressions (10)–(16), scattering by the surfaces of the grains, the classical size effect (whose magnitude depends on the ratio of the grain radius r_0 to the characteristic mean free path l), and the conductivity of the matrix all have a substantial effect on the Hall resistivity of a granular alloy.

The type, concentration, and distribution of impurities near the grain–matrix interface determine the corresponding mean free path l_s, the magnitude of the parameter p_s of spin-dependent scattering from the surface, and the surface Hall resistivity $\rho_H^s=4\pi R_s^s M_z^s$, which by definition is equal to the Hall resistivity of a massive ferromagnet with the same type, concentration, and distribution of impurities as the surface layer of the grains. As follows from expression (13), the sign of p_H^s is in general not the same as that of p_H^b; therefore, for strong surface scattering the Hall resistivity of a granular alloy ρ_H can have the opposite sign of the Hall resistivity ρ_H^b of the parent ferromagnetic metal or alloy. Such a situation was recently observed[13] for granular $(Co_{70}Fe_{30})_xAg_{1-x}$, alloys which is a direct confirmation of the significant role of surface scattering in granular systems.

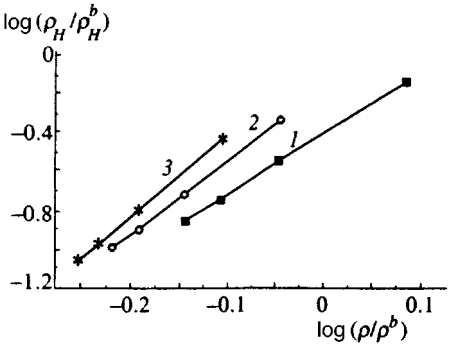


FIG. 2. Correlation between the Hall resistivity ρ_H/ρ_H^b and total resistivity ρ/ρ^b of a granular alloy, having the form $(\rho_H/\rho_H^b)\sim(\rho/\rho^b)^n$; $c=0.2$, $p_b=0.2$, $p_s=0.52$, $l_{m(s)}=120$ Å, $l_{m(d)}=20$ Å, $l_{nm}=200$ Å, $\rho_H^s/\rho_H^b=1$, $r_0=20$–80 Å. The power n depends on the nature of scattering by the surfaces of the grains: $l_s/a_0=2$, $n=3.1$ (curve *1*); $l_s/a_0=4$, $n=3.8$ (curve *2*); $l_s/a_0=6$, $n=4.3$ (curve *3*).

Interestingly enough, if spin-dependent surface scattering always amplifies the giant magnetoresistance, then it can both increase (for $\rho_H^s/\rho_H^b>0$) and decrease (for $-2<\rho_H^s/\rho_H^b<0$) the Hall resistivity. Therefore, there is in general no universal correlation for granular alloys between ρ_H and ρ or giant magnetoresistance. We have found by numerical calculation that for alloys with giant magnetoresistance, grain concentration $c=20\%$, and mean grain size varying between 20 and 60 Å, the relation $\rho_H\sim\rho^n$ does hold, where the exponent n depends substantially on the nature of the surface scattering (Fig. 2). For $l_s/a_0=4$ we have $n=3.8$, and the amplitude of the giant magnetoresistance, calculated in the same model and with the same parameters, reaches 40%. This is found to be in quantitative agreement with the results of Ref. 4 for the anomalous Hall effect and giant magnetoresistance of $Co_{20}Ag_{80}$ alloys. We recently obtained a power-law dependence $\rho_H\sim\rho^n$ with $n=3.7$ for the Zhang–Levy model, but for model parameters that did not describe giant magnetoresistance.[14]

Let us now consider how the size of the grains affects the anomalous Hall effect (see Fig. 3), and whether this influence can be described by some simple relation. On the basis of an analysis of the theory of the size effect for the case of thin films, proceeding from qualitative arguments the authors of Ref. 15 proposed the following relation for the Hall resistivity of a spherical grain:

$$\rho_H=\rho_H^b+\rho_H^s B(l,r_0)P^2\frac{l}{r_0}\left(1+P^s\frac{l}{r_0}\right),\qquad(17)$$

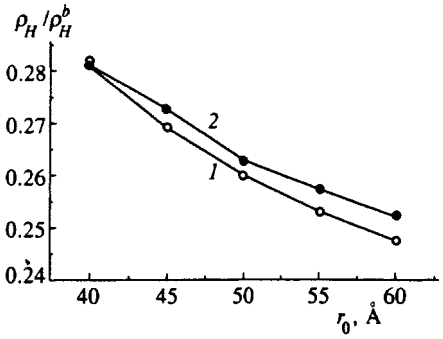


FIG. 3. Dependence of the Hall resistivity ρ_H/ρ_H^b of a granular alloy on the size of the grains r_0 ($c=0.2$, $p_b=0$, $p_s=0$, $l=l_{m(s)}=l_{m(d)}=l_{mn}=l_s/a_0=50$ Å, $\rho_H^s/\rho_H^b=1$); *1*) relation (17), *2*) relations (10)–(16).

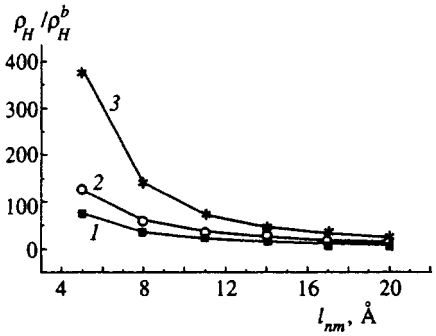


FIG. 5. Dependence of the Hall resistivity ρ_H/ρ_H^b of a granular alloy on the mean free path of an electron in a nonmagnetic matrix l_{nm} ($c=0.2$, $p_b=0.2$, $p_s=0.52$, $l_{m(s)}=150$ Å, $l_{m(d)}=20$ Å, $l_s/a_0=2$, $\rho_H^s/\rho_H^b=1$); *1*) $r_0=50$ Å, *2*) $r_0=30$ Å, *3*) $r_0=10$ Å.

where P^s is the Fuchs–Sondheimer surface reflection parameter and $B(l,r_0)\approx 0.2$ for $l/r_0\leqslant 1$. Figure 3 plots the results of a calculation of the Hall resistivity of a granular alloy as a function of grain size employing relation (17) and the more consistent theory (10)–(16). As could be expected, the two approaches yield the same trends, namely that ρ_H grows as the grain size r_0 is decreased, and all the more so for strong surface scattering. Note that the proposed theory allows not only for the size effect associated with scattering in one grain, but also more complicated transport processes between two and three grains which, of course, are not described by relation (17).

Above we assumed that electronic states with definite spin polarization (for concreteness, anti-aligned with the resulting magnetization) are responsible for the anomalous Hall effect, and we did not distinguish between the mean free path of the anomalous Hall carriers, $l_{(d)}$, and the mean free path of the *s*-like states, $l_{(s)}$, with the same spin polarization. We now show that the possible difference between $l_{(d)}$ and $l_{(s)}$ is not fundamental. Figure 4 plots the results of a calculation of the dependence of ρ_H/ρ_H^b on $l_{(d)}$ for fixed mean free path of the *s* states $l_{(s)}=100$ Å. It can be seen that this dependence is weak and monotonic, and that even for small values of r_0 and $l_{(d)}=l_s/10$ the Hall resistivity is decreased by not more than twofold. This has to do with the fact that the Hall conductivity σ_{xy} depends on l more weakly than the ordinary conductivity σ, since σ_{xy} arises only in the higher Born approximations in the scattering potential. However, the situation can change if a potential barrier (absent or substantially reduced for the *s* states) exists for the *d* states at the grain–matrix interface. This would lead to significant changes in ρ_H/ρ_H^b, which are not observed in experiment. Unfortunately, since the quantity ρ_H/ρ_H^b depends strongly on the surface scattering and not only on the type of carriers, it is difficult to judge the type of carrier of the anomalous Hall effect from the magnitude of ρ_H. For this reason it is also difficult to judge the existence of potential barriers for them in granular systems.

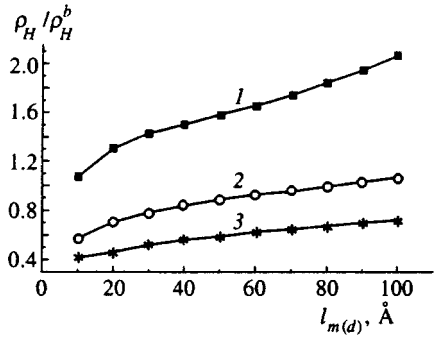


FIG. 4. The Hall resistivity ρ_H/ρ_H^b of a granular alloy for fixed mean free path of the *s* states plotted as a function of the mean free path of the *d* states ($c=0.2$, $p_b=0.2$, $p_s=0.52$, $l_{m(s)}=150$ Å, $l_{nm}=200$ Å, $l_s/a_0=2$, $\rho_H^s/\rho_H^b=10$); *1*) $r_0=50$ Å, *2*) $r_0=30$ Å, *3*) $r_0=10$ Å.

Let us now consider how the conductivity of the matrix affects ρ_H/ρ_H^b. With decreasing mean free path l_{nm} in the matrix (Fig. 5), the Hall resistivity can grow by more than two orders of magnitude, the effect being most pronounced for small grains ($r_0=10$ Å). These trends are found to be in agreement with experimental data on the giant anomalous Hall effect in the granular systems (NiFe)SiO_2 (Refs. 4 and 5) and Fe(SiO_2) (Ref. 16) near the percolation threshold. Such a large effect arises due to the fact that in the case of a high-resistivity matrix, due to shunting, an electron travels mainly via grains where, in multiple collisions with the surfaces of the grains under conditions of the size effect, the skewness of its own motion increases many-fold. Obviously, along with this, the structure of the percolation cluster and the nature of the contact between the grains will play a role. These two factors are not taken into account here, but can lead to additional enhancement of the anomalous Hall effect.

The proposed quantum-statistical model makes it possible to explain the many observed regularities of the anomalous Hall effect in granular alloys. All previous efforts[12,14,17] were based on the concept of self-averaging and were of a purely qualitative character. It is worth noting that the proposed model contains a number of parameters. Such parameters as the concentration and size of the grains and the impurity concentration are easily determined experimentally. The characteristic mean free paths of the electron can also be reliably determined from the experimental data on the resistivity and the anisotropic magnetoresistance,[8,9,11,14] and the spin–orbit coupling parameters according to calculations of the anomalous Hall effect for crystalline ferromagnets[11] are

very close to the spectroscopic data for the spin–orbit coupling of the isolated atoms.

In conclusion, we dwell briefly on a question in the theory of the anomalous Hall effect that has been already discussed for 40 years, specifically, what is the dominant mechanism of this effect? Despite the fact that it has been noted more than once in the literature that the side-jump mechanism does not permit an explanation of much of the data for disordered crystalline[11] and amorphous[18,19] alloys, and that under no conditions can its role exceed that of skew scattering,[11] a number of authors continue to take it to be the main mechanism (see, for example, Ref. 20). In the present paper we have proposed a theory of the anomalous Hall effect that explains the experimental data for granular alloys taking into account only the mechanism of skew scattering, which is an additional confirmation of its dominant role.

ACKNOWLEDGMENTS

This work was carried out with the financial support of the Russian Fund for Fundamental Research (Grant No. 96-02-681a), INTAS (Grant No. 93-0718), and NATO (Grant No. HTECH. LG 951527).

[1] Y. I. Gittleman, Y. Goldstein, and S. Bozovski, Phys. Rev. B **5**, 3609 (1972).
[2] A. Milner, A. Gerber, B. Groisman *et al.*, Phys. Rev. Lett. **76**, 475 (1996).
[3] H. Sato, H. Henmi, Y. Kobayashi *et al.*, J. Appl. Phys. **76**, 6919 (1994).
[4] P. Xiong, G. Xiao, J. Q. Wang *et al.*, Phys. Rev. Lett. **69**, 3220 (1992).
[5] A. B. Pakhomov, X. Yan, and Y. Xu, J. Appl. Phys. **79**, 6140 (1996).
[6] X. N. Jing, N. Wang, A. B. Pakhomov, Phys. Rev. B **53**, 14032 (1996).
[7] A. E. Berkowitz, J. R. Mitchell, M. J. Carey *et al.*, Phys. Rev. Lett. **68**, 3745 (1992).
[8] S. Zhang and P. M. Levy, J. Appl. Phys. **73**, 5315 (1993).
[9] A. Vedyayev, B. Mevel, N. Ryzhanova *et al.*, J. Magn. Magn. Mater. **136**, 91 (1996).
[10] S. V. Vonsovskiĭ, *Magnetism* [in Russian], Nauka, Moscow (1971).
[11] A. V. Vedyaev, A. B. Granovskiĭ, O. A. Kotel'nikova, *Kinetic Phenomena in Disordered Ferromagnetic Alloys* [in Russian], Moscow State Univ. Press, Moscow (1992).
[12] A. Granovsky, A. V. Vedyaev, and F. Brouers, J. Magn. Magn. Mater. **136**, 229 (1994).
[13] M. V. Prudnikova, A. B. Granovsky, and V. N. Prudnikov, in *Proceedings of the Russian–Japanese Joint Seminar ''Physics and Modeling of Intelligent Materials and their Applications'' (PMIMA)*, Moscow (1996), p. 85.
[14] A. Granovsky, F. Brouers, A. Kalitsov, and M. Chshiev, J. Magn. Magn. Mater. **166**, 193 (1997).
[15] F. Brouers, A. Granovsky, A. Sarychev, and A. Kalitsov, submitted to Physica A (1997).
[16] B. Zhao and X. Yan, in *Abstracts of the 41st Annual Conference on Magnetism and Magnetic Materials*, Atlanta, Georgia (1996).
[17] A. Granovskiĭ, A. Kalitsov, and F. Brouers, JETP Lett. **65**, 509 (1997).
[18] A. V. Vedyaev and A. B. Granovskiĭ, Fiz. Met. Metalloved. **58**, 1084 (1984).
[19] M. Prudnikova, A. Granovsky, and V. Prudnikov, J. Magn. Magn. Mater. **166**, 201 (1997).
[20] S. Zhang, Phys. Rev. B **51**, 3632 (1995).

Translated by Paul F. Schippnick

Temperature dependence of surface impedance of $Tl_2Ba_2CaCu_2O_{8-\delta}$ and $YBa_2Cu_3O_{6.95}$ single crystals measured in the microwave band

A. A. Zhukov, M. R. Trunin, A. T. Sokolov, and N. N. Kolesnikov

Institute of Solid State Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia
(Submitted 5 June 1997)
Zh. Éksp. Teor. Fiz. **112**, 2210–2222 (December 1997)

The real part R_s and the imaginary part X_s of the surface impedance $Z_s=R_s+iX_s$ of $Tl_2Ba_2CaCu_2O_{8-\delta}$ and $YBa_2Cu_3O_{6.95}$ single crystals have been measured with high precision at frequency $\omega/2\pi=9.4$ GHz in the temperature range $0<T<140$ K. In the $Tl_2Ba_2CaCu_2O_{8-\delta}$ crystal a linear temperature dependence $R_s(T)$ has been found for $T\leqslant 50$ K, and the magnetic field penetration depth $\lambda(4.2\text{ K})=X_s(4.2\text{ K})/\omega\mu_0\approx 3760$ Å has been measured. Along with well known features of the function $Z_s(T)$ in high-quality $YBa_2Cu_3O_{6.95}$ single crystals, such as the linearity of $\lambda(T)$ and $R_s(T)$ for $T<T_c/3$ and a maximum of $R_s(T)$ at $T\sim T_c/2$, the linearity range of $\lambda(T)$ extends to $T\simeq 50$ K, and this curve has a plateau in the range $60<T<85$ K. The curve of $R_s(T)$ in both the superconducting and normal states of $YBa_2Cu_3O_{6.95}$ is well described by a two-fluid model with the electron–phonon mechanism of quasiparticle relaxation. A formula describing the curve of $\lambda^2(0)/\lambda^2(T)$ throughout the studied temperature range is also given. © *1997 American Institute of Physics.* [S1063-7761(97)02112-4]

1. INTRODUCTION

Measurements of the surface impedance $Z_s=R_s+iX_s$ of high-T_c superconductors in the microwave band yield information about the scattering mechanism, density of states, and nature of the superconducting paring in these compounds. In particular, much attention is being focused on the symmetry of the order parameter in high-T_c superconductors.[1] This discussion was prompted by the observation of a linear temperature dependence of the magnetic field penetration depth, $\lambda(T)=X_s(T)\omega\mu_0$, in the ab-plane of $YBa_2Cu_3O_{6.95}$ (YBCO) single crystals in the low-temperature range.[2] Such a dependence, which is very unusual in the case of isotropic s-pairing, was readily interpreted in terms of the $d_{x^2-y^2}$ symmetry of the order parameter,[3] when the Fermi surface has lines of zero order parameter. Later research demonstrated that $\Delta\lambda(T)$ in YBCO can be linear with temperature in an s-type two-band model,[4] when the superconducting state in the band of Cu–O chains becomes gapless owing to the magnetic impurity scattering, and also in models with a mixed $(d+s)$ symmetry.[5] To date, linear curves of $\lambda(T)$ in the range $T<30$ K have been detected not only in YBCO high-quality single crystals fabricated by different techniques,[6–8] but also in $Bi_2Sr_2CaCu_2O_8$ (BSCCO)[9] and $Ba_{0.6}K_{0.4}BiO_3$ (BKBO) perfect single crystals.[10] Another common feature of high-T_c single crystals detected in numerous experiments[6–10] is the linearity of the surface resistance with temperature in the low-temperature range, $\Delta R_s(T)\propto T$. At frequencies of about 10 GHz the curve of ΔR_s was linear in the range up to $T\sim T_c/2$ (Refs. 9,10) (here T_c is the critical temperature), whereas in YBCO single crystals[6–8] this curve is linear in the range of up to $T\leqslant T_c/4$, and at higher temperatures the curve of $R_s(T)$ has a broad peak at about $T_c/2$. In high-T_c superconductors the shape of $R_s(T)$ is largely determined by scattering of quasiparticles. Calculations for the case of inelastic scattering due to antiferromagnetic spin fluctuations in the $d_{x^2-y^2}$ model[11] or an anisotropic s-type model taking into account electron–phonon scattering[12] showed a peak on the curve of $R_s(T)$ in the intermediate temperature range. However, until now no microscopic model predicts a linear temperature dependence of $R_s(T)$.

The paper describes an experimental technique that allows us to measure the surface resistance $R_s(T)$ and reactance $X_s(T)$ versus temperature in HTSC single crystals at a frequency of 9.4 GHz over a broad temperature range. This technique has been used to measure for the first time the surface impedance of $Tl_2Ba_2CaCu_2O_{8-\delta}$ (TBCCO) single crystals as a function of temperature. Curves of $R_s(T)$ and $X_s(T)$ measured in YBCO samples have been analyzed in terms of the two-fluid model suggested in our earlier publications.[8,10]

2. EXPERIMENTAL TECHNIQUE

The surface impedance of small samples was measured by the ''hot-finger'' technique suggested by Sridhar and Kennedy.[13] Its underlying idea is to place a sample fixed to a sapphire rod at the center of a cylindrical superconducting cavity resonating in the TE_{011} mode, i.e., at the maximum of a homogeneous microwave magnetic field. By varying the sample temperature and measuring first the Q-factor and resonant frequency shift Δf of the cavity with a sample inside and then those of the empty cavity (Q_0 and Δf_0), one can determine the surface resistance R_s and reactance X_s of the sample as functions of temperature.

This method requires that the experimental facility should satisfy two basic conditions. First, since HTSC single crystals are small and their losses in the superconducting state are low, the empty cavity should have high Q_0 so that the temperature dependence of the sample parameters can be detected against the background of the cavity parameters.

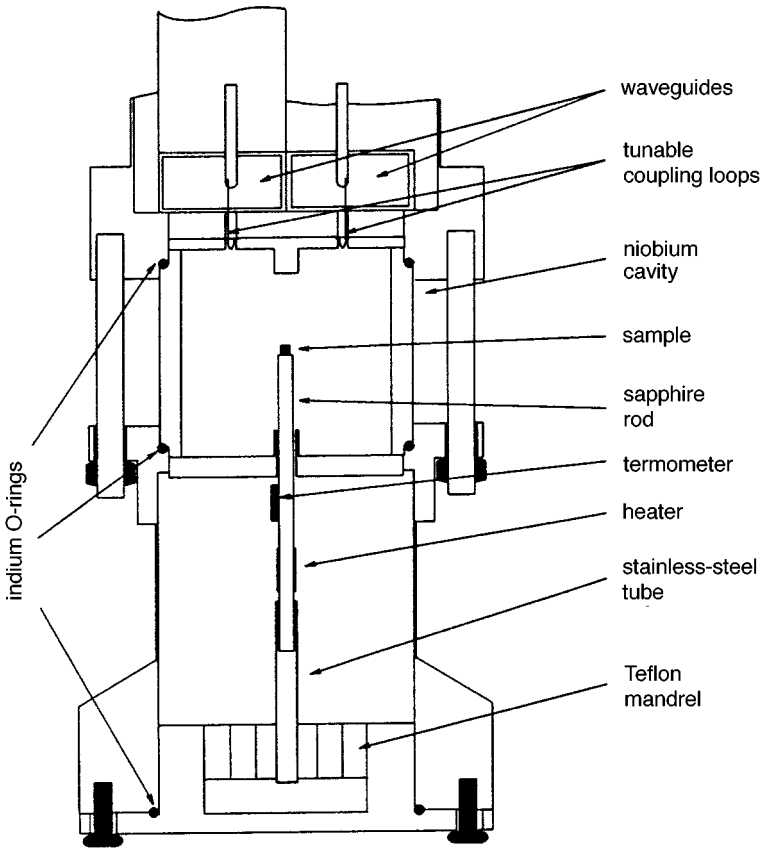


FIG. 1. Design of the microwave cavity unit.

Second, our measurements of Q and Δf by sweeping the microwave frequency demand that the microwave generator should be highly stable.

We have used a cylindrical cavity with a diameter and height of 42 mm fabricated from niobium (Fig. 1). The cavity was immersed in the liquid helium, so it was always in the superconducting state. The TE_{011} mode was driven at a frequency of 9.42 GHz. Since this mode is degenerate (oscillations at the TM_{111} mode have the same frequency), niobium posts with a diameter and height of 5 mm were made on both the upper and lower lids of the cavity, which resulted in a difference between the TM_{111} and TE_{011} modes greater than 10 MHz. The sapphire rod supporting the sample was thermally insulated from the cavity walls owing to the high-vacuum condition inside the cavity, which was sealed by indium O-rings. The lower end of the sapphire rod was inserted into a stainless-steel tube, which was, in turn, supported by an aligning Teflon mandrel. This thermal insulation allowed the cavity to maintain a Q of about 10^7 when the temperature of the sample and rod was raised to $T \sim 140$ K. The sample temperature was controlled by a computerized circuit whose block diagram is shown in Fig. 2. Measurements from the thermometer T were processed by an analogue-to-digital converter (ADC) and fed to a computer. The latter compared the measured and prescribed temperatures and transmitted the difference to a digital-to-analogue converter (DAC). The resulting driving signal was fed to amplifier Amp_1 and corrected the thermal power generated by a heater H. The temperature was maintained at the prescribed value to within 0.3% throughout the temperature range $4 < T < 140$ K.

The high stability of the microwave frequency ($\approx 10^9$) was ensured by a microwave synthesizer equipped with a phase-locked frequency control (PLFC) loop. The microwave signal was fed through an isolator I_1, then a fraction of the microwave power was conducted by a directional coupler (DC) to a frequency converter (FC). There the oscillations with frequency 9.4 GHz were translated to the lower-frequency band (around 50 MHz) and fed to a PLFC unit, where the lower frequency was compared with that generated by a reference frequency synthesizer in the megahertz band. Thus a signal proportional to the phase difference was produced to drive the microwave generator, and the feedback loop was closed. The frequency was swept by changing that of the reference frequency synthesizer, which was driven by the output of a 20-bit DAC.

An electromagnetic wave generated by the microwave synthesizer was conducted via a rectangular waveguide, through the attenuator At_1 and isolator (rectifier) I_2 to the cavity coupled to the waveguides through tunable coupling loops. Downstream of the cavity, the wave was conducted through the isolator I_3 and attenuator At_2 to diode D operated in the quadratic regime. The diode output amplified by amplifier Amp_2 was converted by an ADC and fed to the computer.

By varying the frequency f_{sw} from the microwave synthesizer and measuring the voltage across the diode, we plotted the microwave power transmitted through the cavity as a

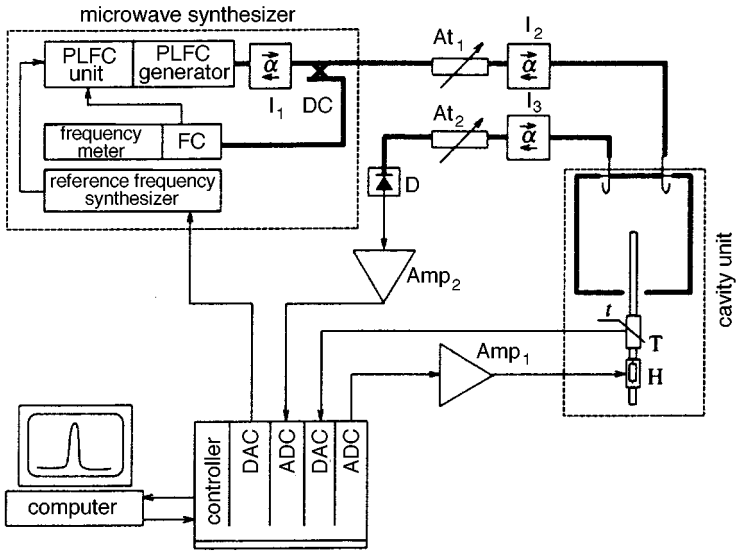


FIG. 2. Box diagram of the facility measuring Q-factor and shift of the resonant frequency as functions of temperature.

function of frequency, which was described by the conventional resonant formula:

$$P(f_{sw}) = \frac{P_0}{4(f_{sw}-f_i)^2/f_i^2 + 1/Q_i^2}, \tag{1}$$

where P_0 is a constant independent of the frequency f_{sw}, and $f_i = f$ and $Q_i = Q$ when the cavity contains a sample and $f_i = f_0$, $Q_i = Q_0$ when the cavity is empty. The resonant frequency f_i of the cavity was derived from the peak position on the curve of $P(f_{sw})$, and the value of Q was derived from the FWHM of the transmission curve using the formula $Q_i = f_i/\delta f_i$, where δf_i is a transmission bandwidth. The Q of about 10^7 was measured to within 1%, and the resonant frequency error was within 10 Hz.

3. EXPERIMENTAL RESULTS

In our experiments, we studied TBCCO and YBCO single crystals, grown by techniques described elsewhere.[8,14] They were shaped as plates with characteristic dimensions of $1\times1\times0.1$ mm (TBCCO) and $1.5\times1.5\times0.1$ mm (YBCO). A sample was fixed on the end of the sapphire rod so that its $\hat{c}$-axis was aligned with the microwave magnetic field and high-frequency currents circulated in the ab-plane.

Figure 3 shows Q (squares) and the shift of the resonant frequency (circles) of the cavity versus temperature with a TBCCO sample inside. The surface resistance R_s and reactance X_s were derived from measured curves of $Q(T)$ and $\Delta f(T)$ using the relations

$$R_s(T) = \Gamma_s[Q^{-1}(T) - Q_0^{-1}(T)], \quad X_s(T) = -\frac{2\Gamma_s}{f_0}[\Delta f(T) - \Delta f_0(T)] + X_0, \tag{2}$$

where Γ_s is the geometrical factor of the sample and X_0 is an additive constant. The parameter X_0 was calculated by equating the imaginary and real parts of the impedance in the normal state, $R_s(T_c) = X_s(T_c)$. The factor Γ_s for a sample with known dimensions was calculated assuming that the magnetic field strength on the sample surface was equal to the field amplitude at the center of the empty cavity. We calculated the geometrical factor of one sample using Eq.

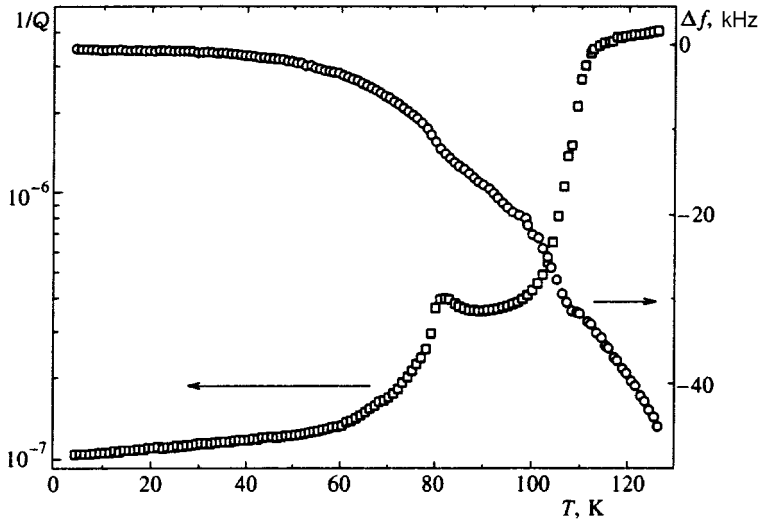


FIG. 3. Measurements of $1/Q$ and Δf versus temperature in a TBCCO single crystal.

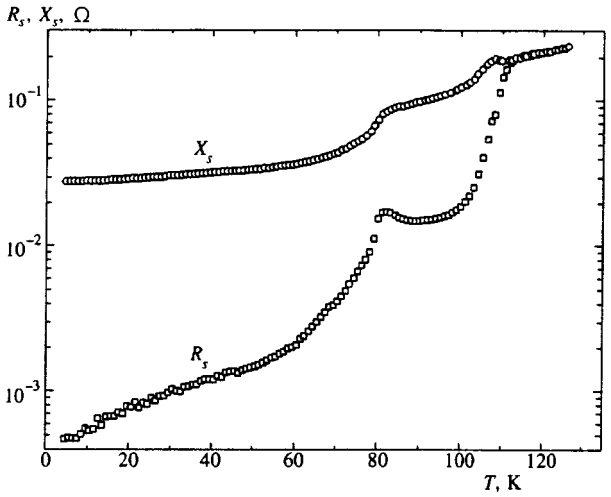


FIG. 4. Surface resistance R_s and reactance X_s of a TBCCO single crystal measured as functions of temperature.

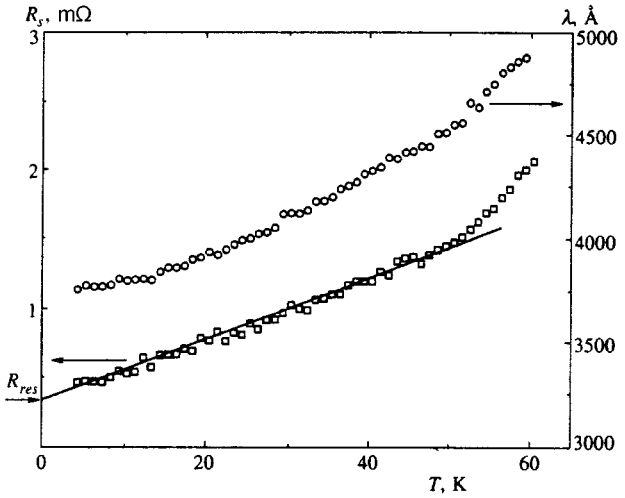


FIG. 5. Surface resistance R_s and field penetration depth $\lambda = X_s/\omega\mu_0$ in a TBCCO single crystal in the low-temperature range. The value $R_s(0\ \mathrm{K}) \equiv R_{\mathrm{res}}$ is marked by the arrow on the left.

(2), measured values of $Q(T_c)$ and the resistivity $\rho(T_c) = 2R^2(T_c)/\omega\mu_0$. This value of Γ_s derived from experimental data was 40% lower than the calculated one. In our subsequent analysis, we took into account this correction to calculations of the geometrical factor.

The surface resistance $R_s(T)$ and reactance $X_s(T)$ of the TBCCO single crystal as functions of temperature are plotted in Fig. 4. For $T \geqslant T_{c1} = 112$ K these parameters are equal, $R_s(T) = X_s(T)$, which corresponds to the normal skin-effect condition. The value $R(T_{c1}) \simeq 0.19\ \Omega$ corresponds to $\rho(T_{c1}) \simeq 100\ \mu\Omega \cdot$cm. On both the $R_s(T)$ and $X_s(T)$ curves, the features corresponding to two superconducting transitions with critical temperatures $T_{c1} \simeq 112$ K and $T_{c2} \simeq 81$ K are clearly seen. According to Simonov and Molchanov,[15] the first transition occurs in the $(Tl_{1.85}Cu_{0.15})Ba_2(Ca_{0.875}Tl_{0.125})Cu_2O_8$ (or 2212) phase, and the second in the $TlBa_2(Ca_{0.87}Tl_{0.13})Cu_2O_7$ (or 1212) phase, which, in contrast to the 2212 phase, has no TlO(3) planes. A single crystal TBCCO containing only one phase has not been fabricated as yet.

Figure 5 shows how the magnetic field penetration depth $\lambda(T) = X_s(T)/\omega\mu_0$ and surface resistance $R_s(T)$ of the TBCCO single crystal depend on the temperature in the low-temperature range. The value $\lambda(4.2\ \mathrm{K}) \approx 3760$ Å is approximately twice as large as the value of $\lambda(0)$ in the 2212 phase with the critical temperature $T_c \approx 100$ K.[16] The function $R_s(T)$ is linear over the interval $4 < T \leqslant 50$ K. By extrapolating this curve to $T = 0$, we obtain the residual resistance $R_{\mathrm{res}} \equiv R_s(0) \approx 320\ \mu\Omega$. The small value of R_{res} indicates a good quality of the crystal surface. In most other samples, the curve of $R_s(T)$ was at a constant level of $R_{\mathrm{res}} > 1\ \mathrm{m}\Omega$ in the range $T < 30$ K.

The surface impedance components of a YBCO single crystal with critical temperature $T_c \simeq 92$ K versus temperature are plotted in Fig. 6. The normal state has $R_s(T) = X_s(T)$, $R_s(T_c) \simeq 0.11\ \Omega$, and $\rho(T_c) \simeq 35\mu\Omega \cdot$cm. The features of $R_s(T)$ and $\lambda(T)$ behavior in the superconducting state is illustrated by Figs. 7 and 8. The linear section of $R_s(T)$ at $T \leqslant 20$ K is followed by a broad peak in the interval around $T \sim 40$ K (Fig. 7). Such a behavior of $R_s(T)$ is typical of high-quality YBCO single crystals.[6–8] A linear section of the curve of $\lambda(T)$ in the range of $T < 30$ K was also observed previously.[2,6–8] On the curve for the YBCO crystal in Fig. 8 the parameter $\Delta\lambda(T) \propto T$ is linear in temperature in the range of up to $T \approx 50$ K. By extrapolating the curve of $\lambda(T)$ to $T = 0$, we obtain the value $\lambda(0) \simeq 1400$ Å, which is characteristic of the ab-plane in YBCO. On the interval $60 < T < 85$ K the magnetic field penetration depth is almost constant with temperature. A plateau on the curves of $\lambda(T)$ was detected in all YBCO crystals grown by the same technique,[8] but the positions of these plateaus with respect to $T_c/2$ were different. For example, in the data of Ref. 8 the plateau was in $35 < T < 65$ K. Microwave measurements of epitaxial films[17] and YBCO single crystals[7] also demonstrated a feature in the curve of $\lambda(T)$ at $T \sim 60$ K.

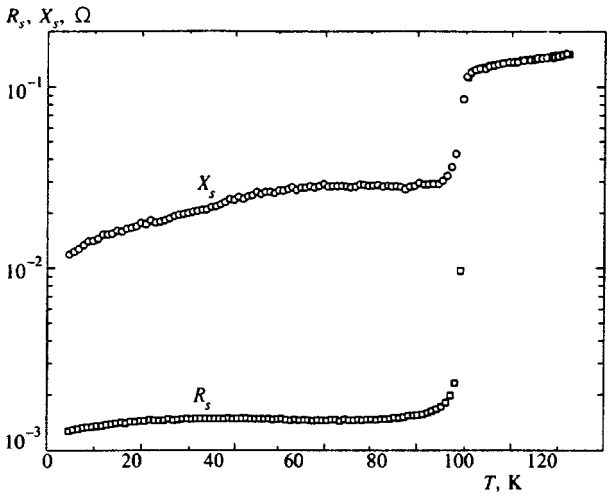


FIG. 6. Surface resistance R_s and reactance X_s of a YBCO single crystal versus temperature.

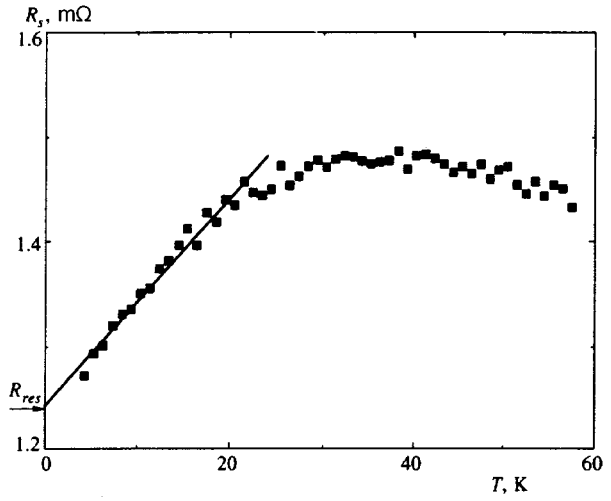


FIG. 7. Surface resistance $R_s(T)$ of the YBCO single crystal in the low-temperature range. The straight line is a guide for the eye. The value $R_s(0) \equiv R_{\rm res}$ is marked by an arrow.

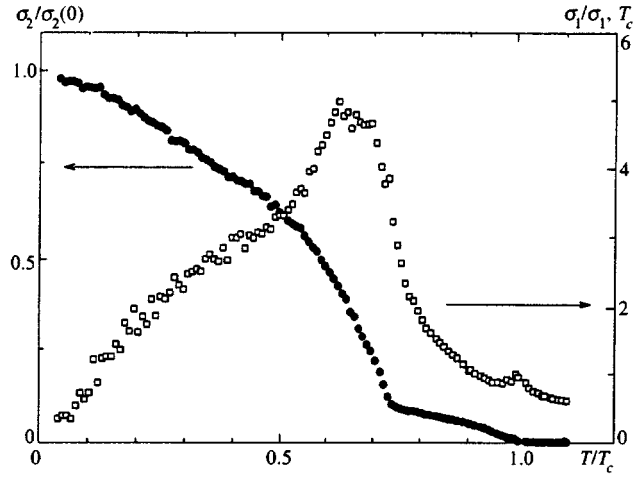


FIG. 9. Components of the TBCCO single crystal conductivity plotted against temperature. The temperature dependent part of $R_s(T)$ in Eq. (5) was determined by subtracting from the measured values of R_s given in Figs. 4 and 5 the residual resistance $R_{\rm res}$.

4. DISCUSSION

In discussing experimental curves of $Z_s(T)$, it is convenient to consider concurrently the complex conductivity $\sigma_s \equiv \sigma_1 - i\sigma_2$. The conductivity is related to the impedance through a simple local formula

$$Z_s = R_s + iX_s = \left(\frac{i\omega\mu_0}{\sigma_1 - i\sigma_2} \right)^{1/2}. \tag{3}$$

Given Eq. (3), we can obtain for normalized parameters general equations relating both the real and imaginary parts of the impedance to the complex conductivity and vice versa in both the superconducting and normal states:

$$\frac{R_s(T)}{R_s(T_c)} = \sqrt{\frac{\sigma(T_c)(\varphi^{1/2}-1)}{\sigma_2\varphi}},$$
$$\frac{X_s(T)}{X_s(T_c)} = \sqrt{\frac{\sigma(T_c)(\varphi^{1/2}+1)}{\sigma_2\varphi}}, \tag{4}$$

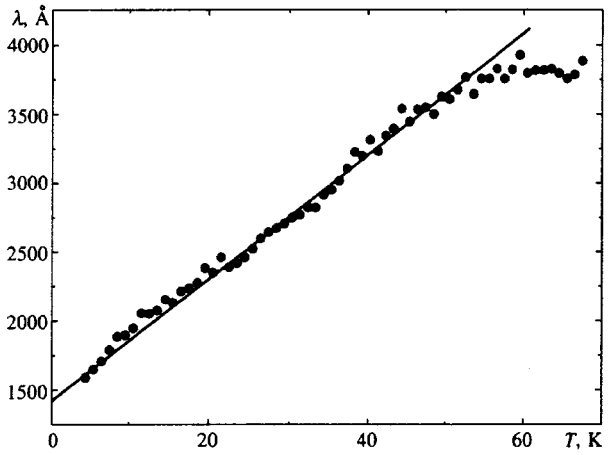


FIG. 8. Magnetic field penetration depth $\lambda(T) = X_s(T)/\omega\mu_0$ in the YBCO crystal in the low-temperature range.

$$\frac{\sigma_1(T)}{\sigma(T_c)} = \frac{4R_s^2(T_c)R_sX_s}{(R_s^2+X_s^2)^2},$$
$$\frac{\sigma_2(T)}{\sigma_2(0)} = \frac{\lambda^2(0)}{\lambda^2(T)} = \frac{X_s^2(0)(X_s^2-R_s^2)}{(R_s^2+X_s^2)^2}. \tag{5}$$

Here $R_s(T_c) = X_s(T_c)$ and $\sigma(T_c) = \sigma_1(T_c)$ are the impedance and conductivity at $T = T_c$, and $X_s(0)$ and $\sigma_2(0)$ are the same parameters at zero temperature; $\lambda = (1/\omega\mu_0\sigma_2)^{1/2}$, $\varphi = 1 + (\sigma_1/\sigma_2)^2$. Figure 9 shows the ratios defined by Eq. (5) for the TBCCO single crystal derived from the experimental data on $R_s(T)$ and $X_s(T)$ plotted in Figs. 4 and 5. The curves in Fig. 9 have rather complex shapes defined by the contributions of two superconducting thallium phases 2212 and 1212. Nonetheless, on the base of our experimental data and previous microwave measurements of the impedance,[6–10] we can describe some features in $Z_s(T)$ and $\sigma_s(T)$ common for all high-quality YBCO, BSCCO, BKBO, and TBCCO single crystals in the ranges of low, intermediate, and subcritical temperatures.

Low temperatures, $T \ll T_c$. The common features are linear sections of the curve of surface resistance $R_s(T)$ and the real part of the conductivity $\sigma_1(T)$. The reactance $X_s \propto \lambda(T)$ and imaginary part of the conductivity $\sigma_2(T)$ are also linear with temperature in all these HTSC materials, except TBCCO crystal, although the curves of $\lambda(T)$ in Fig. 5 and $\sigma_2(T)$ in Fig. 9 have clearly defined linear sections on the interval $20 < T < 50$ K. Our measurements of TBCCO are, in a sense, similar to the data on the first high-quality YBCO single crystals,[18] in which $R_s(T)$ and $\sigma_1(T)$ were linear, whereas the functions $\lambda(T)$ and $\sigma_2(T)$ were thought to be quadratic.

Intermediate temperatures, $T \sim T_c/2$. The common feature of $R_s(T)$ curves for high-quality YBCO crystals manufactured by different techniques is a broad peak in the region about $T \sim 40$ K. Such a peak has not been detected in other HTSC materials, which, unlike YBCO, have tetragonal structures. Measurements of these crystals at ~ 10 GHz are

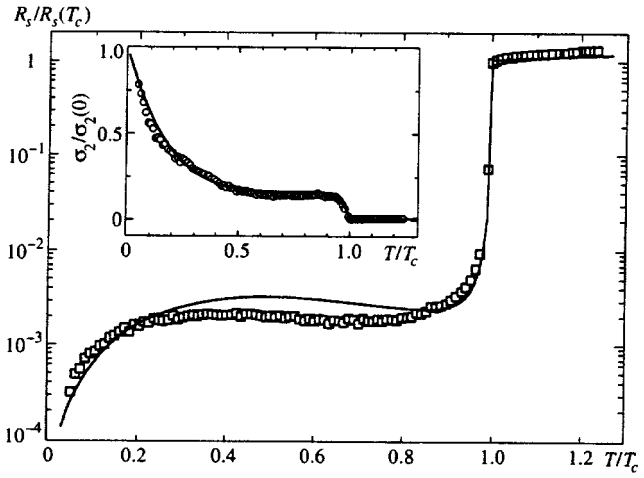


FIG. 10. Comparison between calculations (solid lines) and measurements (dots) of the surface resistance $R_s/R_s(T_c)$ and imaginary part of the conductivity $\sigma_2/\sigma_2(0)$ (insert) of the YBCO single crystal. The temperature dependent part of R_s was obtained by subtracting the residual resistance $R_{\rm res}$ from R_s given in Figs. 6 and 7.

characterized by a linear growth in $R_s(T)$ up to $T\sim T_c/2$. In this work and in recent experiments[7,8] with YBCO single crystals the features were observed in curves of $\lambda(T)$ and $\sigma_2(T)$. The insert to Fig. 10 shows $\sigma_2/\sigma_2(0)$ versus temperature (circles) derived from the experimental data of Fig. 6 using Eq. (5). This curve has a plateau in the intermediate range extending to $T\simeq 0.9T_c$. Note also that the peak of $R_s(T)$ in Figs. 6, 7, and 10 is less pronounced than usual.[6,8,18] Srikanth *et al.*[7] detected a notable increase in $R_s(T)$ in the temperature range following the peak.

Temperatures close to T_c, $T\to T_c$. After a transition to the superconducting state, the surface resistance $R_s(T)$ drops rapidly. At frequencies near 10 GHz, R_s of high-quality YBCO crystals drops by a factor of one hundred or more when the temperature decreases by 1 K below T_c. Near T_c, the curves of $\sigma_2(T)$ are very steep. As can be seen in Fig. 10, the derivative $[1/\sigma_2(0)]d\sigma_2(T)/dT$ equals -4 at $T=T_c$. An identical or similar value of this derivative was obtained in experiments with YBCO single crystals.[6–8]

At present no microscopic theory describing these features of the temperature dependence of the surface impedance or microwave conductivity of HTSC crystals is available. Even the linear section of $R_s(T)$ in the low-temperature range, which is typical of all high-quality HTSC crystals, could not be interpreted in terms of the models with $d_{x^2-y^2}$- or anisotropic *s*-pairing. At the same time, experimenters' attention is focused on the intermediate temperature range and the interval close to T_c. For example, new features of the conductivity detected recently in high-quality YBCO crystals[7,19,20] indicate, in the opinion of the authors, a multicomponent nature of the order parameter. In the absence of a generally accepted microscopic theory, it seems important to have a phenomenological description of the microwave response which would describe features common to all high-T_c superconducting materials. Previously[8,10] we suggested a two-fluid model describing curves of $Z_s(t)$ in the range of $0<T\leqslant 1.3T_c$. Here we apply this model to measurements of the YBCO sample.

In the two-fluid model, a crystal with carrier density n undergoes a transition from the normal to superconducting state at $T=T_c$. As a result of this transition, the superconducting carriers with density n_s and normal carriers with density n_n with equal charges e and masses m coexist, the total density being $n=n_s(t)+n_n(t)$ at all reduced temperatures $t\equiv T/T_c\leqslant 1$. In the microwave band ($\omega\tau\ll 1$) the conductivity components in the two-fluid model are expressed in the very simple form

$$\sigma_1=\frac{e^2\tau}{m}n_n,\quad \sigma_2=\frac{e^2}{m\omega}n_s, \tag{6}$$

where τ is the relaxation time of normal carriers. Given the measured values of $\sigma_2(t)$, we derive $n_s(t)$ and $n_n(t)=n-n_s(t)$ using Eq. (6). Therefore, the only missing function of temperature needed for determination of $\sigma_1(t)$ in Eq. (6) and impedance $Z_s(t)$ in Eqs. (3) and (4) is $\tau(t)$. A reciprocal statement can be formulated, namely, the function $\tau(T)$ can be derived from the experimental data for $R_s(T)$ and $\sigma_2(t)$ on the base of the two-fluid model.

Previously we proved[8] that all the features of $R_s(T)$ in YBCO single crystals discussed above can be described in terms of the electron–phonon scattering mechanism applied to quasiparticles, when $\tau(T)$ in Eq. (6) is expressed by the Bloch–Grüneisen formula:

$$\frac{1}{\tau(t)}=\frac{1}{\tau(T_c)}\frac{\beta+t^5\mathscr{J}_5(\kappa/t)/\mathscr{J}_5(\kappa)}{1+\beta},$$

$$\mathscr{J}_5\left(\frac{\kappa}{t}\right)=\int_0^{\kappa/t}\frac{z^5e^z dz}{(e^z-1)^2}, \tag{7}$$

where $\kappa=\Theta/T_c$ (Θ is the Debye temperature) and $\beta\approx\tau(T_c)/\tau(0)\ll 1$ are dimensionless parameters. For $\kappa\gg 1$ we derive from Eq. (7) the well-known Debye formula $1/\tau(t)\propto t^5$.

By putting $\beta=0.2$ and $\kappa=1.2$ in Eq. (7) and using measured values of $\sigma_2(t)$ (circles in the insert to Fig. 10) and $\omega\tau(T_c)=[\rho(T_c)\sigma_2(0)]^{-1}=0.004$, we derive from Eqs. (6) and (2) the ratio $R_s(t)/R_s(T_c)$ shown by the solid line in Fig. 10. As we noted in our previous publication,[10] the peak of $R_s(T)$ at $T\sim 40$ K typical of YBCO single crystals can be described only if $1/\tau(t)\propto t^5$. If a term quadratic in temperature and corresponding to electron–electron scattering is added to the numerator in Eq. (7), the peak is ''smoothed out,'' as is the case in high-T_c superconductors with a tetragonal structure.

The measured function $\sigma_2(t)/\sigma_2(0)=\lambda^2(0)/\lambda^2(t)=n_s(t)/n$ plotted in Fig. 10 can be also described in terms of the two-fluid model. At $t=1$ the derivative of this function with respect to t equals -4, which is in agreement with the number deriving from the Gorter–Casimir two-fluid model $[n_s(t)\propto(1-t^4)]$ and, as follows from theoretical calculations,[21] corresponds to an average electron–phonon coupling constant larger than unity. The linear sections of $\lambda(T)$ and $R_s(T)$ curves in YBCO, BSCCO, and BKBO

single crystals in the low-temperature range are adequately described by the function $n_s(t) \propto (1-t)^{\alpha}$, where α is a numerical parameter.[10] Therefore, at an arbitrary temperature $T \leqslant T_c$ the function $n_s(t)$ with the asymptotic limits in the low-temperature range and near the critical temperature discussed above can be described by the following formula:

$$n_s/n = (1-t)^{\alpha}(1-\delta) + \delta(1-t^{4/\delta}), \tag{8}$$

where $0 < \delta < 1$ is a weighing factor. The curve representing function (8) in Fig. 10 at $\alpha = 5.5$ and $\delta = 0.15$ is in fair agreement with the experimental values of $\sigma_2(t)/\sigma_2(0)$ in Fig. 10. The entire set of curves $\sigma_2(t)$ measured in YBCO crystals grown by the same technique is approximated by Eq. (8) with the parameter $\alpha \approx 5.5$ almost constant and δ varying between 0.1 and 0.5.

5. CONCLUSIONS

The paper reports on the first ever investigation of the temperature dependence of the real part $R_s(T)$ and imaginary part $X_s(T)$ of the surface impedance of a TBCCO single crystal. We have detected a linear section of the $R_s(T)$ curve in the range of $T \leqslant 50$ K and measured the magnetic field penetration depth $\lambda(4.2\ \mathrm{K})$. Common features of the microwave impedance and complex conductivity of all high-quality HTSC single crystals and specific features of these parameters as functions of temperature have been described. The curves of $R_s(T)$ and $\lambda(T)$ measured in YBCO single crystals are adequately described by the two-fluid model with the electron–phonon mechanism of quasiparticle scattering, the superconducting carrier density being described by a linear function of temperature at $T \ll T_c$ and by the Gorter–Casimir function near T_c.

This work was supported by Russian Fund for Fundamental Research (Project 97-02-16836) and the Superconductivity Program sponsored by the Government of Russia (Grant 96-060).

[1] V. M. Loktev, Fiz. Nizk. Temp. **22**, 3 (1996) [Low Temp. Phys. **22**, 1 (1996)].
[2] W. N. Hardy, D. A. Bonn, D. C. Morgan *et al.*, Phys. Rev. Lett. **70**, 3999 (1993).
[3] P. J. Hirschfeld and N. Goldenfeld, Phys. Rev. B **48**, 4219 (1993); H. Won and K. Maki, Phys. Rev. B **49**, 1397 (1994).
[4] S. D. Adrian, M. E. Reeves, S. A. Wolf *et al.*, Phys. Rev. B **51**, 6800 (1995); A. A. Golubov, M. R. Trunin, A. A. Zhukov *et al.*, JETP Lett. **62**, 496 (1995).
[5] R. Combescot and X. Leyronas, Phys. Rev. Lett. **75**, 3732 (1995); C. O'Donovan and J. P. Carbotte, Phys. Rev. B **52**, 4568 (1995); H. Kim and E. J. Nicol, Phys. Rev. B **52**, 13576 (1995).
[6] D. A. Bonn, S. Kamal, K. Zhang *et al.*, Phys. Rev. B **50**, 4051 (1994); J. Mao, D. H. Wu, J. L. Peng *et al.*, Phys. Rev. B **51**, 3316 (1995); T. Jacobs, S. Sridhar, C. T. Rieck *et al.*, J. Phys. Chem. Solids **56**, 1945 (1995).
[7] H. Srikanth, B. A. Willemsen, T. Jacobs *et al.*, Phys. Rev. B **55**, R14733 (1997).
[8] M. R. Trunin, A. A. Zhukov, G. A. Emel'chenko *et al.*, JETP Lett. **65**, 938 (1997).
[9] T. Jacobs, S. Sridhar, Q. Li *et al.*, Phys. Rev. Lett. **75**, 4516 (1995); S. F. Lee, D. C. Morgan, R. J. Ormeno *et al.*, Phys. Rev. Lett. **77**, 735 (1996).
[10] M. R. Trunin, A. A. Zhukov, G. É. Tsydynzhapov *et al.*, JETP Lett. **64**, 832 (1996).
[11] P. J. Hirschfeld, W. O. Putikka, and D. J. Scalapino, Phys. Rev. B **50**, 4051 (1994).
[12] A. A. Golubov, M. R. Trunin, A. A. Zhukov *et al.*, J. Phys. I (France) **6**, 2275 (1996).
[13] S. Sridhar and W. L. Kennedy, Rev. Sci. Instrum. **54**, 531 (1988).
[14] N. N. Kolesnikov, M. P. Kulakov, Yu. A. Osip'yan *et al.*, Sverkhprovodimost': Fizika, Khimiya, Tekhnika **4**, 957 (1991).
[15] V. S. Simonov and V. N. Molchanov, Kristallografiya **41**, 31 (1996) [Crystallogr. Rep. **41**, 26 (1996)].
[16] H. Ning, H. Duan, P. D. Kirven *et al.*, J. Supercond. **5**, 503 (1992).
[17] N. Klein, N. Tellmann, H. Schulz *et al.*, Phys. Rev. Lett. **71**, 3355 (1993).
[18] D. A. Bonn, R. Liang, T. M. Riseman *et al.*, Phys. Rev. B **47**, 11314 (1993).
[19] R. Cagnon, S. Pu, B. Ellman *et al.*, Phys. Rev. Lett. **78**, 1976 (1997).
[20] C. J. Stevens, D. Smith, C. Chen *et al.*, Phys. Rev. Lett. **78**, 2212 (1997).
[21] A. A. Mikhailovsky, S. V. Shulga, A. E. Karakozov *et al.*, Solid State Commun. **80**, 511 (1991); R. T. Collins, Z. Schlesinger, and F. Holtzberg, Phys. Rev. B **43**, 3701 (1991); O. V. Dolgov, E. G. Maksimov, A. E. Karakozov *et al.*, Solid State Commun. **89**, 827 (1994).

Translation provided by the Russian Editorial office.

Contributions to the theory of ferromagnetism in the degenerate Hubbard model

R. O. Zaĭtsev

Kurchatov Institute, 123182 Moscow, Russia
(Submitted 14 July 1997)
Zh. Éksp. Teor. Fiz. **112**, 2223–2236 (December 1997)

The possibility of ferromagnetic ordering in a generalized Hubbard model with allowance for degeneracy and for infinite Hubbard energy is studied. The region of existence of ferromagnetism for electron density greater than 1 is determined. It is shown that for electron density less than 1 ferromagnetism exists only in special cases when the Fermi surface passes near van Hove singularities. © *1997 American Institute of Physics.* [S1063-7761(97)02212-9]

1. INTRODUCTION

The ferromagnetic properties of transition metals are best described by the Hubbard model[1] with strong repulsion between *d*-electrons located in the same unit cell. The long-range part of the Coulomb interaction is assumed to be negligible, since it is compensated by the screening effect of the *s* electrons. The Hubbard energy is assumed to the largest energy parameter,[2] and from the outset it is assumed to be infinite.

If $s-d$ hybridization is neglected, then the magnetic properties of transition elements are determined mainly by the *d*-electron subband, whose width can be expressed in terms of hopping integrals. For simplicity and clarity, in the case of interest to us, that of a cubic crystal, a model with zero off-diagonal and identical diagonal hopping integrals will be used:

$$\hat{H}=-\sum_{\mathbf{r},\mathbf{r}',\sigma,\lambda;\mathbf{r}\neq\mathbf{r}'} t_\lambda(\mathbf{r}-\mathbf{r}')\hat{a}^+_{\mathbf{r},\sigma,\lambda}\hat{a}_{\mathbf{r},'\sigma,\lambda} - \sum_{\mathbf{r},\sigma,\lambda}(\mu+\sigma H)\hat{a}^+_{\mathbf{r},\sigma,\lambda}\hat{a}_{\mathbf{r},\sigma,\lambda}. \tag{1}$$

Here μ is the chemical potential, $\sigma=\pm$ is the spin index, H is the external magnetic field, the crystal index λ takes on three values ($\lambda=xy,yz,zx$) if the t_{2g} shell is filled, and two values ($\lambda=3z^2-r^2$, x^2-y^2) if the e_g shell is filled.

It will be shown that for electron density less than 1, ferromagnetism exists only in special cases for which the Fermi surface passes near van Hove singularities. This method makes it possible to explain the existence of ferromagnetism in Ni (see, for example, Refs. 3 and 4).

The present work is devoted to a study of the conditions under which ferromagnetism appears that are not associated with the existence of van Hove singularities. The calculations are performed in the one-loop approximation for electron or hole densities greater than 1.

2. DENSITY RANGE $N<1$

In the limiting case of infinite Hubbard energy, it is convenient to express the creation and annihilation operators in terms of the $\hat{X}$ operators describing transitions between vacant (0) and one-particle ($|0,\lambda\rangle$) states

$$\hat{a}_{\mathbf{r},\sigma,\lambda}=\hat{X}^{(0|\sigma,\lambda)}_{\mathbf{r}},\quad \hat{a}^+_{\mathbf{r},\sigma,\lambda}=\hat{X}^{(\sigma,\lambda|0)}_{\mathbf{r}}. \tag{2}$$

To find the average occupation numbers $n_{\sigma,\lambda}$ we employ the simplest one-loop approximation, in which the self-energy part $\Sigma_{(\sigma)}$ does not depend on the momentum $\mathbf{p}$:

$$n^{(\sigma)}_\lambda=f^{(\sigma)}_\lambda\sum_{\mathbf{p}} n_F(\xi^{(\sigma,\lambda)}_{\mathbf{p}}). \tag{3}$$

Here $n_F(\epsilon)$ is the Fermi distribution and $\xi^{(\sigma)}_{\mathbf{p}}=f^{(\sigma)}_\lambda t_{\mathbf{p}}+\Sigma(\sigma)-\mu-\sigma H$ is the one-particle excitation energy.

On account of the cubic symmetry the end multiplier can be expressed in terms of the average occupation numbers,[5] and does not depend on the number λ of the single-particle state:

$$f^{(\sigma)}_\lambda=1-\sum_s^k n^{(\hat{\sigma})}_s-\sum_{s\neq\lambda}^k n^{(\sigma)}_s =1-kn^{(\overline{\sigma})}-(k-1)n^{(\sigma)},\qquad \overline{\sigma}=-\sigma. \tag{4}$$

This is also true of the self-energy part

$$\Sigma(\sigma)=-\sum_{\mathbf{p}} t_{\mathbf{p}}[kn_F(\xi^{(\overline{\sigma})}(\mathbf{p}))+(k-1)n_F(\xi^{(\sigma)}(\mathbf{p}))], \tag{5}$$

where k is the number of crystal components.

The end multiplier can be expressed in terms of the average number of particles $n=\Sigma_{\lambda,\sigma}n^{(\sigma)}_\lambda$ and the average magnetic moment per cell $M=\Sigma_{\lambda,\sigma}[n^{(\sigma)}_\lambda-n^{(\overline{\sigma})}_\lambda]$.

In zero field, all quantities can be expressed in terms of the average number of particles n per cell:

$$n^{(\pm)}_\lambda=\frac{n}{4k},\quad f=1-n\left(1-\frac{1}{2k}\right),\quad n=2kf\sum_{\mathbf{p}} n_F(\xi_{\mathbf{p}}).$$

We obtain the spin susceptibility by differentiating Eq. (3). Using the obvious condition $\delta n^{(+)}=-\delta n^{(-)}$, we find the derivative

$$\chi=\frac{\delta M}{\delta H}=\frac{-2kfD_0}{(1-D_1)[(1-n)/f]-fD_1}, \tag{6}$$

where

$$D_m=\sum_{\mathbf{p}} t^m_{\mathbf{p}}n'_F(\xi_{\mathbf{p}}). \tag{7}$$

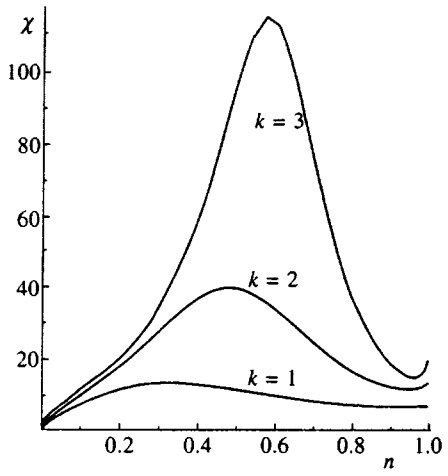


FIG. 1. Magnetic susceptibility ($T=0$) as a function of the electron density for various degrees of degeneracy.

In the limit $T=0$, all moments D_m can be expressed in terms of the initial density of states $\rho(\epsilon)=\Sigma\,\delta(t_{\mathbf{p}}-\epsilon)$.

The band calculations for the fcc nickel lattice show a very sharp peak in the density of states for energies corresponding to filling of the degenerate hole-type ($k=2$) $3z^2-r^2$ and x^2-y^2 states.

For this reason, it can be expected that the denominator of the susceptibility (6) will vanish for some finite interval $\mu<0$, which at $T=0$ corresponds to the hole density range $0<n_h<0.8$. The experimentally observed saturation magnetic moment in ferromagnetic nickel corresponds to $n_h=0.6$ (see Fig. 1).

This result holds for $n<1$. In this range the susceptibility differs from that of an ideal Fermi gas by a factor much greater than 1. In the three-dimensional model under study, the denominator vanishes at the limit of the range—$n_e=1$. However, the numerator also vanishes at that point, so that as a result the susceptibility near this density acquires a substantial temperature dependence,[6,7] which corresponds to the magnetism of localized one-particle states.

3. E_g ELECTRONS. DENSITY RANGE $1<N<2$

The situation in which the system resonates between one- and two-particle states is much more interesting. It is convenient here to introduce a new chemical potential $\mu-U\to\mu$ and assume that there are no vacant states at all.

The single-particle states $\hat{a}_\sigma^+|0\rangle$ and $\hat{b}_\sigma^+|0\rangle$ have spin 1/2. Lower-energy two-particle states $^3A_2$ have spin $S=1$:

$$\hat{a}_\sigma^+\hat{b}_\sigma^+|0\rangle\ \ (S_z=\sigma),\quad \frac{\hat{a}_\uparrow^+\hat{b}_\downarrow^+ + \hat{a}_\downarrow^+\hat{b}_\uparrow^+}{\sqrt{2}}|0\rangle\ \ (S_z=0). \tag{8}$$

For simplicity, higher-energy ^1E and $^1A_1$ states are neglected.

In a finite external magnetic field, the variations in the end multipliers depend on variations in both the one-particle (n_{I}) and two-particle (n_{II}) occupation numbers. Taking into consideration the symmetry of the system under interchange of states a and b, the variations in the end multipliers are given by

$$f_1^{(\sigma)}=n_{\mathrm{II}}^{(\sigma)}+n_{\mathrm{I}}^{(\sigma)},\quad \delta f_1^{(\sigma)}=\delta n_{\mathrm{II}}^{(\sigma)}+\delta n_{\mathrm{I}}^{(\sigma)},$$
$$f_2^{(\sigma)}=n_{\mathrm{II}}^{(0)}+n_{\mathrm{I}}^{(\overline{\sigma})},\quad \delta f_2^{(\sigma)}=\delta n_{\mathrm{I}}^{(\overline{\sigma})}=-\delta n_{\mathrm{I}}^{(\sigma)}. \tag{9}$$

Thus, in contrast to the preceding ''one-particle'' case, now it is necessary to have two independent equations to find the one- and two-particle variations.

To obtain these equations, we employ the part of the expansion of the annihilation operator that corresponds to transitions between one- and two-particle states:

$$\hat{a}_{\mathbf{r}\sigma}=\hat{Z}_{\mathbf{r}}=b_1\hat{X}_{\mathbf{r}}^{(0,\sigma|\sigma,\sigma)}+b_2\hat{X}_{\mathbf{r}}^{(0,\hat{\sigma}|A_0)}, \tag{10'}$$

where $b_1=1$ and $b_2=1/\sqrt{2}$.

We multiply this expression by an arbitrary linear combination of conjugate X operators

$$\hat{Y}_{\mathbf{r}}=\beta_1\hat{X}_{\mathbf{r}}^{(\sigma,\sigma|0,\sigma)}+\beta_2\hat{X}_{\mathbf{r}}^{(A_0|0,\overline{\sigma})}. \tag{10a}$$

Averaging the separate T-products over states with a fixed temperature and chemical potential, in the one-loop approximation we find a relation between the two-particle occupation numbers n_{II}, the Fourier components $\hat{G}_\omega(\mathbf{p})$ of the virtual one-particle Green's function, and the end multipliers f_k.

The inverse single-particle Green's function matrix in turn can be expressed in terms of the end multipliers f_k and in terms of the self-energy matrix $\hat{\Sigma}$, which in the one-loop approximation is independent of both frequency and momentum and in our model reduces immediately to the diagonal form

$$\hat{G}_\omega^{k,s}(\mathbf{p})=[\,\delta_{k,s}(i\omega-\Sigma_s+\mu+\sigma H)-f_k b_k t_{\mathbf{p}} b_s\,]^{-1}. \tag{11}$$

We calculate the T-products in the one-loop approximation:

$$-\langle\hat{T}(\hat{Z}_{\mathbf{r}}(\tau)\hat{Y}_{\mathbf{r}}(\tau+0))\rangle=b_1\beta_1\langle X^{(\sigma,\sigma|\sigma,\sigma)}\rangle+b_2\beta_2\langle X^{(A_0|A_0)}\rangle$$
$$=b_1\beta_1 n_{\mathrm{II}}^{(\sigma)}+b_2\beta_2 n_{\mathrm{II}}^{(0)}$$
$$=T\sum_{\omega,k,s,\mathbf{p}} b_k G_\omega^{(k,s)}(\mathbf{p})\beta_s f_s e^{i\omega\delta}. \tag{12}$$

Here δ is a small positive correction and f_s is a terminal factor.

Differentiating Eq. (12) with respect to the magnitude of the magnetic field gives two equations. Under the condition $b_1\beta_1+b_2\beta_2=0$ we find an equation that does not depend on the applied magnetic field:

$$\delta n_{\mathrm{II}}^{(\sigma)}(1-K_0)-2K_0\delta n_{\mathrm{I}}^{(\sigma)}$$
$$=\delta f_1^{(\sigma)}(1-K_0)+\delta f_2^{(\sigma)}(1+K_0)$$
$$=A(\mu)[\,\delta\Sigma_1(\sigma)-\delta\Sigma_2(\sigma)], \tag{13}$$

where the electron (hole) density lies in the range $1<n_e<2$,

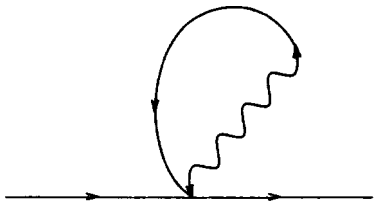


FIG. 2. One-loop self-energy part. The nonwavy line represents separate components of the virtual Green's function, and the wavy line represents the Fourier components of the hopping integrals.

$$K_0=\sum_{\mathbf{p}} n_F(\xi_{\mathbf{p}})=4\frac{n_e-1}{2+n_e},\quad \xi_{\mathbf{p}}=b^2 f t_{\mathbf{p}}-\mu,$$

$$f=\frac{2+n_e}{12},\quad A(\mu)=\sum_{\mathbf{p}}\frac{n_F(\xi_{\mathbf{p}})-n_F(-\mu)}{b^2 t_{\mathbf{p}}}. \tag{14}$$

For $\beta_k=b_k$ we obtain an equation for the susceptibility,

$$\begin{aligned}\delta n_{\mathrm{II}}^{(\sigma)}&=\delta f_2^{(\sigma)}+\delta f_1^{(\sigma)}\\ &=K_0\sum_k b_k^2\delta f_k^{(\sigma)}+fD_0\sum_k b_k^2\delta\Sigma_k(\sigma)\\ &\quad+fb^2D_1\sum_k b_k^2\delta f_k^{(\sigma)}-b^2fD_0\sigma\delta H,\end{aligned} \tag{15}$$

where

$$D_n=\sum_{\mathbf{p}} t_{\mathbf{p}}^n n_F'(\xi_{\mathbf{p}}),\quad b_1^2=1,\quad b_2^2=1/2,\quad b^2=3/2.$$

To study the one-loop self-energy diagrams $\Sigma_{1,2}$, it is sufficient to calculate the separate loops and then sum them taking account of the commutation rules, which determine the nonzero vertex parts of the kinematic interaction (see Refs. 8 and 9 as well as Fig. 2).

In the simplest model without hybridization, we have only diagonal self-energy parts

$$\begin{aligned}\Sigma_1^{(a,\sigma)}&=-A_2^{(\bar\sigma)}+B_1^{(\sigma)},\\ \Sigma_2^{(a,\sigma)}&=-A_1^{(\bar\sigma)}+B_2^{(\bar\sigma)}+A_2^{(\bar\sigma)}+B_2^{(\sigma)},\\ \Sigma_1^{(b,\sigma)}&=-B_2^{(\bar\sigma)}+A_1^{(\sigma)},\\ \Sigma_2^{(b,\sigma)}&=-B_1^{(\bar\sigma)}+A_2^{(\bar\sigma)}+B_2^{(\bar\sigma)}+A_2^{(\sigma)}.\end{aligned} \tag{16}$$

Here

$$A_s^{(\sigma)}=T\sum_{n,\omega,\mathbf{p}} t_{\mathbf{p}}^{s,n}G_\omega^{n,s}(\mathbf{p})$$

is the sum of products of the matrix elements of the hopping operator $\hat t(\mathbf{p})$ with the matrix elements of the virtual Green's function (11), referred to a given projection of the spin and a given a state.

In a cubic crystal, the variation of the self-energy part does not depend on the number of the atomic state (a or b), but it does change sign when the spin projection changes sign.

Direct calculations show that the variation in the self-energy part depends only on the transition number, $\delta\Sigma_1^{(a,\sigma)}=\delta\Sigma_2^{(b,\sigma)}=\delta\Sigma_1(\sigma)$ and $\delta\Sigma_2^{(a,\sigma)}=\delta\Sigma_2^{(b,\sigma)}=\delta\Sigma_2(\sigma)$.

We obtain two equations for the variations $\delta\Sigma_k$ directly from their definition in terms of the integrals of the Green's function (11)—the so-called one-loop approximation:

$$\begin{aligned}\delta\Sigma_k^{(\sigma)}&=-\delta\Sigma_k^{(-\sigma)}=-(QW_{k,n}-D_1U_{k,n})\delta\Sigma_n^{(\sigma)}\\ &\quad+b^2D_2U_{k,n}\delta f_n^{(\sigma)}-\sigma D_1R_k\delta H.\end{aligned} \tag{17}$$

The numerical matrices $\hat U$ and $\hat W$, as well as the vector $\mathbf{R}$, can be expressed in terms of the matrix $\hat S$ constructed in accordance with Eqs. (16)

$$\begin{aligned}&\hat S=\begin{pmatrix} b_1^2=1 & b_2^2=1/2\\ b_1^2=1 & -b_2^2=-1/2\end{pmatrix},\\ &R_k=\sum_n S_{k,n},\quad U_{k,n}=\frac{R_kb_n^2}{\mathbf{b}^2},\quad W_{k,n}=U_{k,n}-S_{k,n}.\end{aligned} \tag{18}$$

In this case $R_k=\Sigma_nS_{k,n}=(3/2,1/2)$. In the matrix $\hat W$ the elements in the first row equal 0, $W_{1,1}=W_{1,2}=0$, and the elements of the second row sum to zero, $W_{2,1}=-W_{2,2}=-2/3$,

$$U_{1,1}=1,\quad U_{1,2}=\frac{1}{2},\quad U_{2,1}=\frac{1}{3},\quad U_{2,2}=\frac{1}{6}. \tag{19}$$

The inhomogeneous term is found to be proportional to the vector $\mathbf{R}=(3/2,1/2)$.

Equations (13) and (17) imply a relation that is independent of both the external field and the end multipliers f_k:

$$b_2^2(\delta\Sigma_1+\delta\Sigma_2)=[b_1^2+Q(\mu)](\delta\Sigma_1-\delta\Sigma_2). \tag{20}$$

Here $\xi_{\mathbf{p}}=\Sigma_0+t_{\mathbf{p}}b^2f-\mu$, $Q(\mu)=[K_0-n_F(-\mu)]/fb^2$, $b^2=3/2$, $b_1^2=1$, and $b_2^2=1/2$.

Two of Eqs. (9), which relate the variations of the occupation numbers and the end multipliers, must be added to this equation. Thus, our problem is to solve the system of four equations (13), (15), (17), and (20).

The condition that this equation be solvable is that the determinant of the following matrix not vanish:

$$\begin{pmatrix} 1-b_1^2(K_0+fb^2D_1) & 1-b_2^2(K_0+fb^2D_1) & -fD_0b_1^2 & -fD_0b_2^2\\ 1-K_0 & 1+K_0 & -A(\mu) & A(\mu)\\ -D_2b^2b_1^2 & -D_2b^2b_2^2 & 1-D_1b_1^2 & -D_1b_2^2\\ 0 & 0 & b_2^2-b_1^2-Q & b_1^2+b_2^2+Q\end{pmatrix}.$$

Vanishing of the determinant of this matrix signifies ferromagnetic instability.

Direct calculations at $T=0$ lead to the equation

$$K_0(1-K_0)(Q+b^2)$$
$$=-A(\mu)D_2b_2^2+D_1[f(Q+b^2)(b_1^2-b_2^2+b^2K_0)$$
$$+K_0(1-K_0)(Qb^2+b_1^4+2b_1^2b_2^2-b_2^4)]. \quad (21)$$

All coefficients depend on the chemical potential. For a semielliptical band model, these results can be expressed in terms of the angular parameter α, which varies from 0 to 2π. In the limit $T=0$ we have

$$D_0=-\frac{2\sin(\alpha/2)}{f\pi b^2}, \quad D_1=\frac{\sin\alpha}{f\pi b^2}, \quad D_2D_0=D_1^2, \quad b^2=\frac{3}{2},$$

$$A(\mu)=\frac{2}{\pi b^2}\left\{\ln\left|\tan\left(\frac{\pi-\alpha}{4}\right)\right|+\sin\alpha/2\right\}, \quad K_0=\frac{\alpha-\sin\alpha}{2\pi},$$

$$Q(\mu)=\frac{K_0-\theta(\alpha-\pi)}{fb^2}, \quad n_e=2\frac{2+K_0}{4-K_0},$$

$$f=\frac{n_e+2}{12}, \quad K_0=\frac{n_e-1}{3f}. \quad (22)$$

The logarithmic divergence of the coefficient $A(\mu)$ in the limit $\mu\to 0$ in Eq. (21) is compensated by the fact that the determinant D_2 vanishes in the limit $\alpha\to\pi$. Solving Eq. (21) numerically gives the unique solution $\alpha_0=2.56$, $n_e=1.26$, corresponding to the onset of ferromagnetic instability in the range $1<n_e<1.26$.

4. t_{2g} ELECTRONS. DENSITY RANGE $1<N<2$

In this case the system resonates among six single-particle states $\hat a_\sigma^+|0\rangle$, $\hat b_\sigma^+|0\rangle$, and $\hat c_\sigma^+|0\rangle$ and nine lower-energy two-particle states $^3A_2$:

$$\hat a_\sigma^+\hat b_\sigma^+|0\rangle \ (S_z=\sigma), \quad \frac{\hat a_\uparrow^+\hat b_\downarrow^++\hat a_\downarrow^+\hat b_\uparrow^+}{\sqrt{2}}|0\rangle \ (S_z=0), \quad (23a)$$

$$\hat b_\sigma^+\hat c_\sigma^+|0\rangle \ (S_z=\sigma), \quad \frac{\hat b_\uparrow^+\hat c_\downarrow^++\hat b_\downarrow^+\hat c_\uparrow^+}{\sqrt{2}}|0\rangle \ (S_z=0), \quad (23b)$$

$$\hat c_\sigma^+\hat a_\sigma^+|0\rangle \ (S_z=\sigma), \quad \frac{\hat c_\uparrow^+\hat a_\downarrow^++\hat c_\downarrow^+\hat a_\uparrow^+}{\sqrt{2}}|0\rangle \ (S_z=0). \quad (23c)$$

These three groups of relations are interrelated by the cyclic substitution $A\to B\to C\to A$.

In the simplest model without hybridization we have only the diagonal self-energy parts:

$$\begin{aligned}
\Sigma_1^{(a,\sigma)}&=-A_2^{(\overline\sigma)}+B_1^{(\sigma)}-C_1^{(\sigma)}-C_2^{(\overline\sigma)},\\
\Sigma_2^{(a,\sigma)}&=-A_1^{(\overline\sigma)}+B_2^{(\overline\sigma)}+B_2^{(\sigma)}+A_2^{(\overline\sigma)}-C_1^{(\overline\sigma)}-C_2^{(\sigma)},\\
\Sigma_3^{(a,\sigma)}&=-A_4^{(\overline\sigma)}+C_1^{(\sigma)}-B_1^{(\sigma)}-B_2^{(\overline\sigma)},\\
\Sigma_4^{(a,\sigma)}&=-A_3^{(\overline\sigma)}+C_2^{(\overline\sigma)}+C_2^{(\sigma)}+A_4^{(\overline\sigma)}-B_1^{(\overline\sigma)}-B_2^{(\sigma)},
\end{aligned} \quad (24)$$

Here

$$A_s^{(\sigma)}=T\sum_{n,\omega,\mathbf{p}}t_{\mathbf{p}}^{s,n}G_\omega^{n,s}(\mathbf{p})$$

is the sum of the products of the matrix elements of the transition matrix $\hat t(\mathbf{p})$ with the elements of the virtual Green's function matrix $\hat G_\omega(\mathbf{p})$, referred to a given projection of the spin and a given a state. We obtain six other equations by means of the cyclic substitutions $A\to B\to C\to A$ and $a\to b\to c\to a$.

Note that for a cubic crystal, the variation in the self-energy part does not depend on the number of the atomic A, B, or C state, but it does change sign when the spin projection change sign: $\delta\Sigma_s^{(n,\sigma)}=\delta\Sigma_s(\sigma)$. It is also evident that each term on the right-hand side is proportional to the squared genealogical coefficient: $A_k^{(\sigma)}=B_k^{(\sigma)}=C_k^{(\sigma)}=g_k^2Q$, where $g_k^2=(1,1/2,1,1/2)$.

Substituting these numbers into the right-hand side of Eqs. (24) yields the same result for all three self-energy parts calculated in a vanishing external magnetic field.

To find the first correction to the magnitude of the magnetic field, we employ the general equations (17). The real S matrix in these equations can be constructed on the basis of Eqs. (24):

$$\hat S=\begin{pmatrix}
0 & 2g_2^2=1 & 0 & 0\\
2g_1^2=2 & -2g_2^2=-1 & 0 & 0\\
0 & g_2^2=1/2 & 0 & g_4^2=1/2\\
g_1^2=1 & -g_2^2=-1/2 & g_3^2=1 & -g_4^2=-1/2
\end{pmatrix},$$

$$R_k=(1,1,1,1). \quad (25)$$

The matrix $U_{k,i}=R_kg_i^2/g^2$ thus possesses elements that do not depend on the row number. Therefore equations (17) possess a particular solution that does not depend on the transition number:

$$\delta\Sigma_k^{(a,\sigma)}=\delta\Sigma_l^{(b,\sigma)}=\delta\Sigma_m^{(c,\sigma)}=\delta\Sigma(\sigma)=-\delta\Sigma(-\sigma).$$

Using these relations and the explicit expression for the one-particle Green's function, we obtain an equation for $\delta\Sigma$:

$$\delta\Sigma(\sigma)=D_1\delta\Sigma(\sigma)+D_2\sum_{k=1}^4 g_k^2\delta f_k^{(\sigma)}-\sigma\delta HD_1. \quad (26)$$

Here we use the same notation as in Eq. (15), but in this case

$$g^2=\sum_{k=1}^4 g_k^2=3.$$

The equation of state, differentiated with respect to the external field, has the same general form as Eq. (15):

$$2\delta n_{\rm II}^{(\sigma)}=(K_0+fg^2D_1)\sum_{k=1}^4 g_k^2\delta f_k^{(\sigma)}+fg^2\delta\Sigma(\sigma)D_0$$
$$-fg^2\sigma\delta HD_0. \quad (27)$$

However, in this equation

$$f=\frac{4-n_t}{18}, \quad K_0=\sum_{\mathbf{p}}n_F(\xi_{\mathbf{p}})=\frac{n_t-1}{9f}.$$

This becomes a complete set of equations if a relation is established between the variations of the one- and two-particle states. In our case this equation is independent of both the variation in the external field and the variation in the self-energy part

$$(1-K_0)\,\delta n_{\rm II}^{(\sigma)}=2K_0\,\delta n_{\rm I}(\sigma). \tag{28}$$

This equation must be supplemented with two relations from (9) that relate the variations in the occupation numbers and end multipliers:

$$\delta n_{\rm II}^{(\sigma)}=\delta f_1^{(\sigma)}+\delta f_2^{(\sigma)}, \quad \delta f_2^{(\sigma)}=-\delta n_{\rm I}^{(\sigma)}. \tag{29}$$

Thus, our problem is to solve the system of three equations (26), (27), and (28) with the additional conditions (29).

The condition for this system to have a solution is that the determinant of the following matrix not vanish:

$$\begin{pmatrix} 1-g_1^2(K_0+fg^2D_1) & 1-g_2^2(K_0+fg^2D_1) & -fD_0g^2/2 \\ 1-K_0 & 1+K_0 & 0 \\ -2D_2g_1^2 & -2D_2g_2^2 & 1-D_1 \end{pmatrix}. \tag{30}$$

Vanishing of the determinant of this matrix signifies ferromagnetic instability.

At $T=0$ we obtain

$$K_0(1-K_0)=D_1[f(1+3K_0)+K_0(1-K_0)]. \tag{31}$$

All coefficients depend on the chemical potential. For a semielliptical band model, the integrals can be expressed in terms of the angular parameter α. In the limit $T=0$, we again have Eqs. (22), but with a different equation of state:

$$K_0=\frac{\alpha-\sin\,\alpha}{2\pi}, \quad n_t=2\,\frac{2K_0+1}{2+K_0}, \quad f=\frac{4-n_t}{18}, \quad g^2=3. \tag{32}$$

Equation (31) has only one solution, at $\alpha=2.51$. Ferromagnetic instability is therefore present at all $0<\alpha<2.51$, which corresponds to the density range $1<n_e<1.4$.

5. t_{2g} ELECTRONS. DENSITY RANGE $2<N<3$

Neglecting hybridization among $3d$ and $4sp$ states, we examine the case of an almost half-filled t subband, in which electronic states resonate between two- and three-particle states. The lowest three-particle state has $S=3/2$ and is fourfold degenerate in the spin projection:

$$\hat a_\sigma^+\hat b_\sigma^+\hat c_\sigma^+|0\rangle, \quad S_z=3\sigma/2,$$

$$\frac{1}{\sqrt3}\,(\hat a_{\overline\sigma}^+\hat b_\sigma^+\hat c_\sigma^+|0\rangle+\hat a_\sigma^+\hat b_{\overline\sigma}^+\hat c_\sigma^+|0\rangle+\hat a_\sigma^+\hat b_\sigma^+\hat c_{\overline\sigma}^+|0\rangle),$$

$$S_z=\sigma/2. \tag{33}$$

The three lowest triplet states with spin 1 were constructed from the three different products of pairs of creation operators (see Eq. (23)).

The expansion in terms of the X transition operators between the two- and three-particle states of lowest energy is determined by the three genealogical coefficients:

$$\hat a_{\mathbf{r}\sigma}=\hat X_{\mathbf{r}}^{(0,\sigma,\sigma|3\sigma/2)}+\sqrt{\frac{2}{3}}\,\hat X_{\mathbf{r}}^{(A(yz,xz)|\sigma/2)}+\frac{1}{\sqrt3}\,\hat X_{\mathbf{r}}^{(0,\overline{\sigma,\sigma}|\sigma/2)}. \tag{34}$$

From Eq. (34) we obtain an expansion of the two other annihilation operators by means of cyclic substitution.

In the absence of a field, all average occupation numbers and end multipliers can be expressed in terms of n_t, the mean number of electrons per cell. Taking account of the degree of degeneracy, we have

$$3n_{\rm II}+4n_{\rm III}=1, \quad 18n_{\rm II}+12n_{\rm III}=n_t. \tag{35}$$

To find the equation of state, we express the occupation numbers of the three-particle states in terms of the one-particle Green's function at coinciding points

$$n_t=18n_{\rm II}+b^2f\sum_{\sigma,k,\mathbf{p}}n_F(b^2f\epsilon_{\mathbf{p}}^{(k)}-\mu). \tag{36}$$

Here $b^2=2$ is the sum of the squares of all genealogical coefficients in the expansion (34). We determine the excitation spectrum in terms of the poles of the one-particle Green's function, and we express $n_{\rm II}$ in terms of n_t by means of (36). Summing over the spin index, we obtain the equation of state:

$$n_t=2+4fK_0, \quad f=\frac{5n_t-6}{36},$$

$$K_0=\frac{1}{3}\sum_{\mathbf{p},k=1,2,3}n_F(\xi_{\mathbf{p}}^{(k)}), \quad \xi_{\mathbf{p}}^{(k)}=b^2f\epsilon_{\mathbf{p}}^{(k)}-\mu. \tag{37}$$

The equations for the variations of the three-particle occupation numbers $\delta n_{\rm III}^{(3\sigma/2)}$ and $\delta n_{\rm III}^{(\sigma/2)}=-\delta n_{\rm III}^{(-\sigma/2)}$ can be obtained from the general equation for the average value of T-products of the annihilation operator (34) by a linear combination of three conjugate operators with arbitrary coefficients β_s:

$$b_1\beta_1n_{\rm III}^{(3\sigma/2)}+b_2\beta_2n_{\rm III}^{(\sigma/2)}+b_3\beta_3n_{\rm III}^{(-\sigma/2)}$$

$$=T\sum_{1\leqslant k,n\leqslant3}\sum_{\omega,\mathbf{p}}b_kG_\omega^{k,n}(\mathbf{p})\beta_nf_n. \tag{38}$$

In the one-loop approximation, the matrix elements of the Green's function can be expressed in terms of the inverse matrix (11).

Variation of this relation leads to

$$b_1\beta_1\delta n_{\rm III}^{(3/2)}+b_2\beta_2\delta n_{\rm III}^{(1/2)}+b_3\beta_3\delta n_{\rm III}^{(-1/2)}$$

$$=K_0[b_1\beta_1\delta f_1+b_2\beta_2\delta f_2+b_3\beta_3\delta f_3]+(\mathbf{b}\cdot\boldsymbol{\beta})f\,\delta G. \tag{38$'$}$$

If the vector $\boldsymbol{\beta}$ is assumed to be orthogonal to the vector $\mathbf{b}$, then $\mathbf{b}\cdot\boldsymbol{\beta}=0$ and two relations that do not depend explicitly on the variation of the magnetic field can be obtained. The equation relating the variations of the occupation numbers also contains the variation of the self-energy parts $\delta\Sigma_k$. We find the first equation under the conditions $b_3\beta_3=b_1\beta_1$ and $b_2\beta_2=-2b_1\beta_1$:

$$(1-K_0)(\delta n_{\rm III}^{(3/2)}-3\,\delta n_{\rm III}^{(1/2)})=A(\mu)(\delta\Sigma_1-2\,\delta\Sigma_2+\delta\Sigma_3). \tag{39}$$

Setting $\beta_2=0$ and $b_3\beta_3=-b_1\beta_1$, we obtain the second equation

$$(1-K_0)(\delta n_{\rm III}^{(3/2)}+\delta n_{\rm III}^{(1/2)})-2K_0\delta n_{\rm II}=A(\mu)(\delta\Sigma_1-\delta\Sigma_3). \tag{40}$$

Here

$$A(\mu)=\sum_{\mathbf{p}}\ \frac{n_F(\xi_{\mathbf{p}})-n_F(-\mu)}{b^2 t_{\mathbf{p}}}.$$

The variations of the terminal factors can be expressed in terms of the variations of the occupation numbers:

$$\delta f_1=\delta n_{\rm III}^{(3/2)}+\delta n_{\rm II},\quad \delta f_2=\delta n_{\rm III}^{(1/2)},$$

$$\delta f_3=\delta n_{\rm III}^{(-1/2)}-\delta n_{\rm II}.$$

Using the additional condition $\delta n_{\rm III}^{(-1/2)}=-\delta n_{\rm III}^{(1/2)}$, we find the inverse relations

$$\delta n_{\rm III}^{(3/2)}=\delta f_1+\delta f_2+\delta f_3,\quad \delta n_{\rm III}^{(1/2)}=\delta f_2,$$

$$\delta n_{\rm II}=-\delta f_2-\delta f_3. \tag{41}$$

The variation δG of the virtual Green's function contains three types of terms:

$$b_1^2\delta n_{\rm III}^{(3\sigma/2)}+b_2^2\delta n_{\rm III}^{(\sigma/2)}+b_3^2\delta n_{\rm III}^{(-\sigma/2)}$$

$$=K_0\sum_{k=1,2,3} b_k^2\delta f_k^{(\sigma)}+f\sum_{k=1,2,3} b_k^2\delta\Sigma_k^{(\sigma)}D_0$$

$$+b^2 f\sum_{k=1,2,3} b_k^2\delta f_k^{(\sigma)}D_1-b^2 f\sigma\delta H D_0. \tag{42}$$

Here

$$D_k=\sum_{\mathbf{p}}\ t_{\mathbf{p}}^k n_F'(\xi_{\mathbf{p}}).$$

To study the one-loop self-energy diagrams, it is sufficient to calculate the separate loops and then sum them taking account of the commutation rules, which determine the nonzero vertex parts of kinematic origin (see Fig. 2).

In our very simple model without hybridization, we have only diagonal self-energy parts:

$$\Sigma_1^{(a,\sigma)}=-A_3^{(\overline{\sigma})}+B_1^{(\sigma)}+C_1^{(\sigma)},$$

$$\Sigma_2^{(a,\sigma)}=-A_2^{(\overline{\sigma})}+B_2^{(\sigma)}+C_2^{(\sigma)}+A_3^{(\overline{\sigma})}+B_3^{(\overline{\sigma})}+C_3^{(\overline{\sigma})}, \tag{43}$$

$$\Sigma_3^{(a,\sigma)}=-A_1^{(\overline{\sigma})}+A_2^{(\overline{\sigma})}+B_2^{(\overline{\sigma})}+C_2^{(\overline{\sigma})}+B_3^{(\sigma)}+C_3^{(\sigma)}.$$

Here

$$A_s^{(\sigma)}=T\sum_{n,\omega,\mathbf{p}} t_{\mathbf{p}}^{s,n}G_\omega^{n,s}(\mathbf{p})$$

is the sum of the products of the matrix elements of the transition matrix $\hat{t}(\mathbf{p})$ with the elements of the virtual Green's function matrix $\hat{G}_\omega(\mathbf{p})$, referred to a given spin projection and a given a state. We find the remaining six equations by cyclic substitution: $A\to B\to C\to A$.

In the absence of a field the integrals A, B, and C differ only in their end multipliers: $A_k=B_k=C_k=b_k^2\nu$, where ν is a correction to the chemical potential, so that at $H=0$ all self-energy parts (43) equal $5\nu/3$.

We obtain directly from Eq. (43) three equations for $\delta\Sigma$ in terms of integrals of the Green's function (11)—the so-called one-loop approximation:

$$\delta\Sigma_k^{(\sigma)}=-\delta\Sigma_k^{(-\sigma)}=-(F_{k,n}^{(0)}-D_{k,n}^{(1)})\delta\Sigma_n^{(\sigma)}+b^2D_{k,n}^{(2)}\delta f_n^{(\sigma)}$$
$$-\sigma\delta H R_k D_1. \tag{44}$$

In the presence of a field, the equations for the variations can be written as a set of three equations of the general form (12), with coefficients that can be expressed in terms of the elements of the numerical matrix $\hat{S}$:

$$\hat{S}=\begin{pmatrix} 2b_1^2=2 & 0 & b_3^2=1/3 \\ 0 & 3b_2^2=2 & -3b_3^2=-1 \\ b_1^2=1 & -3b_2^2=-2 & 2b_3^2=2/3 \end{pmatrix},$$

$$R=\left(\frac{7}{3},\,1,\,-\frac{1}{3}\right). \tag{45}$$

The matrices

$$\hat{D}^{(n)}=\sum_{\mathbf{p}}\ t_{\mathbf{p}}^n n_F'(\xi_{\mathbf{p}})\hat{U}$$

differ in their temperature factors, and are proportional to the same matrix $\hat{U}_{n,m}=R_n b_m^2/b^2$:

$$\begin{pmatrix} \frac{(2b_1^2+b_3^2)b_1^2}{b^2}=\frac{7}{6} & \frac{(2b_1^2+b_3^2)b_2^2}{b^2}=\frac{7}{9} & \frac{(2b_1^2+b_3^2)b_3^2}{b^2}=\frac{7}{18} \\ \frac{3(b_2^2-b_3^2)b_1^2}{b^2}=\frac{1}{2} & \frac{3(b_2^2-b_3^2)b_2^2}{b^2}=\frac{1}{3} & \frac{3(b_2^2-b_3^2)b_3^2}{b^2}=\frac{1}{6} \\ \frac{(b_1^2-3b_2^2+2b_3^2)b_1^2}{b^2}=-\frac{1}{6} & \frac{(b_1^2-3b_2^2+2b_3^2)b_2^2}{b^2}=-\frac{1}{9} & \frac{(b_1^2-3b_2^2+2b_3^2)b_3^2}{b^2}=-\frac{1}{18} \end{pmatrix}. \tag{46}$$

Here the following values of the squared genealogical coefficients have been used: $b_1^2=1$, $b_2^2=2/3$, $b_3^2=1/3$, and $b^2=\Sigma_k b_k^2=2$.

The operator $\hat{F}^{(0)}$ has the form $\hat{F}^{(0)}=Q(\mu)\hat{W}$, where

$$Q(\mu)=\frac{1}{fb^2}\sum_{\mathbf{p}}\left[n_F(\xi_{\mathbf{p}})-n_F(-\mu)\right]$$

$$=\frac{1}{fb^2}\left[K_0-n_F(-\mu)\right].$$

In the matrix $\hat{W}$, the elements of each row sum to zero:

$$\begin{pmatrix} U_{1,1}-2b_1^2=-\frac{5}{6} & U_{1,2}=\frac{7}{9} & U_{1,3}-b_3^2=\frac{1}{18} \\ U_{2,1}=\frac{1}{2} & U_{2,2}-3b_2^2=-\frac{5}{3} & U_{2,3}+3b_3^2=\frac{7}{6} \\ U_{3,1}-b_1^2=-\frac{7}{6} & U_{3,2}+3b_2^2=\frac{17}{9} & U_{3,3}-2b_3^2=-\frac{13}{18} \end{pmatrix}.$$

The inhomogeneous term is proportional to the three-component quantity $\mathbf{R}=(7/3,\ 1,\ -1/3)$. Here we have introduced the following notation for the squared genealogical coefficients: $b_1^2=1$, $b_2^2=2/3$, $b_3^2=1/3$, and $b^2\Sigma_k b_k^2=2$. (47)

Since the effects of an external field and the relationship to the variation of the end multipliers are governed by the vector $\mathbf{R}$, we can obtain general relations only for the variations $\delta\Sigma_k$. To this end we multiply both sides of Eqs. (44) by the components of the vectors $\mathbf{N}^{(\pm)}$, which are orthogonal to the vector $\mathbf{R}$ (i.e. $\mathbf{R}\cdot\mathbf{N}^{(\pm)}=0$):

$$\delta\Sigma_k^{(\sigma)}N_k^{(\pm)}=Q(\mu)N_k^{(\pm)}W_n^k\delta\Sigma_n^{(\sigma)}. \tag{48}$$

Here $\lambda=\pm$, and the third equation can be obtained by multiplying by the vector $\mathbf{M}$, whose components satisfy the condition that they are orthogonal to the result obtain by operating with the operator $\hat{W}$. In other words, $M_nW_n^kH_k=0$ for arbitrary vector $\mathbf{H}$. This possibility stems from the fact that the determinant of the matrix $\hat{W}$ equals 0, while none of the principal minors vanishes. As a result we obtain

$$\delta\Sigma_k^{(\sigma)}M_k-S_nD_1\delta\Sigma_n^{(\sigma)}$$

$$=b^2S_nD_2\delta f_n^{(\sigma)}-\sigma\delta H(M_kR_k)D_1. \tag{49}$$

Here $S_n=M_kU_n^k$.

We employ below the three auxiliary vectors

$$\mathbf{N}^{(+)}=(10,-21,7),\quad \mathbf{N}^{(-)}=(0,1,3),$$

$$\mathbf{M}=(1,-2/3,-1),\quad \mathbf{S}=(1,2/3,1/3)=b_k^2.$$

Ferromagnetic instability is a consequence of the resulting set of equations having no solution. Simple calculations give the following condition for the onset of ferromagnetic instability:

$$[3K_0(1-K_0)-2fD_1(2+3K_0)]$$

$$\times\{18+52Q+9Q^2-D_1[26+68Q-35Q^2]\}$$

$$=2D_2\{fD_0(2+3K_0)[14+40Q-11Q^2]$$

$$-2A(\mu)[14+40Q+9Q^2]\}. \tag{50}$$

At zero temperature, when $D_0D_2=D_1^2$, Eq. (50) becomes

$$[3K_0(1-K_0)-b^2f_tD_1(2+3K_0)]$$

$$\times[18+52Q+9Q^2+4D_1(3+7Q-6Q^2)]$$

$$=-4D_2A(\mu)(14+40Q+9Q^2). \tag{51}$$

The calculation of the coefficients in Eq. (51) for a semielliptical-band model at $T=0$ shows that ferromagnetic instability exists in the narrow density range $2<n_t<2.141$.

6. CONCLUSIONS

In summary, in the limiting case of infinite Hubbard energy, ferromagnetic ordering emerges over narrow ranges of electron or hole density, each of which exceeds 1. This is due to the fact that for densities less than 1, the system resonates between vacant and one-particle states, and it exhibits the maximum tendency toward ferromagnetic ordering for an extremely small number of electrons. The scattering amplitude calculated for opposing values of the spin projection is positive, which leads to high paramagnetic susceptibility without the onset of ferromagnetic ordering.[10]

For a number of particles (or holes) greater than 1, the system resonates between N- and $N\pm1$- particle (Hund) states, which have the maximum possible value of the spin. Ferromagnetic instability results from the existence of finite spin polarization under the additional condition that the scattering amplitude of excitations with opposing spins is positive.

Experimental data show that electron states in pure iron (Fe) fall between the configurations $3d^24(sp)^6$ and $3d^34(sp)^5$. The saturation magnetic moment equals $2.2\mu_B$, which agrees qualitatively with the results of Sec. 5.

[1] J. Hubbard, Proc. R. Soc. London, Ser. A **296**, 82, 100 (1967).
[2] J. Goodenough, Phys. Rev. **120**, 67 (1960).
[3] E. C. Stoner, Phil. Mag. **22**, 81 (1936).
[4] J. Hubbard, Proc. Phys. Soc. London **84**, 455 (1964).
[5] R. O. Zaĭtsev, Zh. Éksp. Teor. Fiz. **70**, 1100 (1976) [Sov. Phys. JETP **43**, 574 (1976)].
[6] Y. Nagaoka, Phys. Rev. **147**, 392 (1966).
[7] A. Suto, Phys. Rev. B **43**, 8779 (1991).
[8] F. Dyson, Phys. Rev. **102**, 1217,1230 (1956).
[9] R. O. Zaĭtsev, Phys. Lett. A **134**, 199 (1988).
[10] J. Hubbard and K. P. Jain, J. Phys. C **2**, 1650 (1968).

Translated by M. E. Alferieff

Generation of soliton packets in a two-level laser

A. A. Zabolotskiĭ

Institute of Automation and Electrometry, Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia

(Submitted 28 March 1997)
Zh. Éksp. Teor. Fiz. **112**, 2237–2251 (December 1997)

A variant of perturbation theory is constructed for a system of nearly integrable equations. Perturbations of a special type are considered, which makes it possible to represent the system in the form of compatibility condition for ''deformed'' linear systems. The corresponding deformation of the Whitham equations is found. The mathematical apparatus is used to theoretically examine the generation of a sequence of solitons in a two-level laser. The generation process is described by a system of Maxwell–Bloch equations with pumping of the upper level and with allowance for some relaxation effects. The dynamics of the transformation of the initial perturbation into a sequence of solitons under pumping is studied. Finally, the various generation regimes are analyzed and compared with the experimental data. © *1997 American Institute of Physics.* [S1063-7761(97)02312-3]

1. INTRODUCTION

Describing the generation of laser pulses in nonlinear media is one of the most important problems of nonlinear physics. Much work, both theoretical and experimental, has been done in this field (see the literature cited in Refs. 1–4). The theoretical work mostly deals with the behavior of isolated pulses. At the same time, it was discovered in experiments that for large relaxation times dense packets of pulses can be generated in a system with pumping. Such regimes, especially the nature of the generated pulse packets and their characteristics, have yet to be studied in detail. The study of nonlinear processes in amplifying media is related to solving evolutionary equations. In most cases for such models there is no analytical analysis of the generation of packets of ultrashort pulses. Partially this is due to the analytical difficulties that emerge when one tries to describe the behavior of dense packets of pulses with many degrees of freedom. Some of these difficulties can be overcome by employing models that are exactly solvable or are close to being integrable. The most detailed information about the evolution of the fields in nonlinear media can be obtained by employing the inverse scattering method.[5] This method was used to study the characteristics of long laser amplifiers[6] and to analyze the self-similar asymptotic behavior of the solutions of Raman scattering[7,8] and Maxwell–Bloch models in media with nondegenerate[9,10] and degenerate[11] transitions. Some researchers (see Refs. 12–14) also studied the mixed boundary-value problem for a model of stimulated Raman scattering. It was assumed that initially the medium is partially inverted. Both numerically and by the inverse scattering method it was found that the solution has two components, a soliton component and a nonsoliton component, and that at large times the solution exhibits a nonsoliton self-similar asymptotic behavior. Formally, the equations used in Refs. 8, 12–14 to describe stimulated Raman scattering are equivalent to the Maxwell–Bloch equations for a two-level laser, in which a weak pulse propagating in the medium ''absorbs'' energy previously stored by the inverted medium.

Most lasers use a different amplification scheme. For instance, the operation of a broad class of gas, ion, and solid-state lasers can be described by the Maxwell–Bloch equations for a two-level medium in which the higher level is being pumped.[1,4] It was noticed in experiments that the initial stage of pulse generation in amplifying media can in certain conditions be approximated by a modulated periodic solution. In this paper, a modulated periodic (single-phase) solution is employed to explain the generation of solitons in a laser. Such an approach is used to describe the modulational instability, which sets in when a step pulse propagates in a nonlinear medium. In the case of the Korteweg–de Vries equation, this approach makes it possible to describe fairly accurately the transformation of a plane wave into a soliton packet. The heuristic method is based on the solution of the system of Whitham equations for the slowly varying parameters of a periodic wave.[15,5] This approach has been justified for several models both numerically and analytically.[5]

The goal of the present investigation is to explain the generation of soliton packets in a two-level laser. To this end a perturbation theory is developed in order to solve integrable systems of equations of the Ablowitz–Kruskal–Newell–Segur (AKNS) type.[5] Here perturbations of a special type, i.e., perturbations that can be incorporated into the model by ''extending'' the partial derivatives, are taken into account. The limits of applicability of this perturbation theory are partially compensated by the relative simplicity of the final formulas and the importance of the problems being solved by the theory. In developing the perturbation theory we will use the results obtained for integrable models with a variable spectral parameter.[16] Such models emerge in different areas of physics. Belinskiĭ and Zakharov[17,18] developed the mathematical apparatus for constructing N-soliton solutions for integrable equations of the theory of gravity, and Mikhaĭlov and Yaremchuk[19,20] developed the necessary tools for constructing soliton solutions for the cylindrically symmetric Heisenberg model and a model of the ''main chiral field.'' What sets the approach used in this paper apart from the exact methods of finding the solutions used in the cited papers is that here we study the dynamics of a dense packet

of a large number of pulses. The interaction of the pulses with each other and the effect of pumping and relaxation lead to modulation of the pulse packet, a process described by a periodic wave with slowly varying parameters. A perturbation-theory mathematical apparatus is developed in order to describe the slow deformation of the periodic wave. The apparatus can be used for a certain class of models (see below). To this end an appropriate generalization of the Whitham equations is carried out. The solution of these equations is used to analyze the nature and dynamics of pulse packets generated in a two-level laser from a weak initial plane wave. Both the stationary and time-dependent generation regimes are examined. It is found that generation of soliton packets is possible. The results are compared with the experimental data.

The plan of the paper is as follows. In Sec. 2 a perturbation theory is developed for the general AKNS system. In Sec. 3 the deformed Whitham equations are derived. These equations take into account perturbations for the single-phase solution. Section 4 is devoted to applying the mathematical apparatus in describing the generation of pulse packets within the framework of the Maxwell–Bloch equations with allowance for pumping and relaxation. Two types of solution are studied here. The first describes the time-dependent process of the trailing edge of a step pulse splitting into a packet of solutions and the amplification that accompanies this process. The range of applicability of this solution is discussed. The second solution describes the passage of the system to a state of steady generation of solitons. The results are compared with the experimental data. Finally, it is established that the second regime agrees qualitatively with the experimental results.

2. PERTURBATION THEORY FOR DEFORMED INTEGRABLE SYSTEMS

The mathematical apparatus used in this section is based primarily on the results of Ref. 16. Let us examine the deformation of the integrable equations that emerges as a result of ''extending'' the partial derivatives. Such a procedure is known to lead to integrable equations with a variable spectral parameter, provided that certain conditions are met.[16] Most often these conditions cannot be realized in physical experiments. We will show that they can be made less stringent or can even be dropped if the deformations are weak and slow. Below we will see that for slow deformations this approach makes it possible to study a range of parameters that cannot be reached by reduction of the integrable model.

The evolutionary equations emerge as the compatibility condition for the following overdetermined systems of equations:

$$\Phi_\xi = U\Phi, \quad \Phi_\eta = V\Phi. \tag{1}$$

Here U, V, and Φ are matrix functions of ξ, η, and the spectral parameter. Generally, U and V are rational functions of the spectral parameter λ and have the form

$$U(\lambda,\xi,\eta) = u_0 + \sum_{n=1}^{N_1} \frac{u_n(\xi,\eta)}{\lambda-\lambda_n},$$

$$V(\lambda,\xi,\eta) = v_0 + \sum_{n=1}^{N_2} \frac{v_n(\xi,\eta)}{\lambda-\mu_n}, \tag{2}$$

where the simple poles λ_n and μ_m do not coincide. The compatibility condition is

$$U_\eta - V_\xi + [U,V] = 0. \tag{3}$$

Let us ''extend'' the partial derivatives:

$$\frac{\partial}{\partial\eta} \to D_\eta = \frac{\partial}{\partial\eta} + F(\lambda)\,\frac{\partial}{\partial\lambda},$$

$$\frac{\partial}{\partial\xi} \to D_\xi = \frac{\partial}{\partial\xi} + G(\lambda)\,\frac{\partial}{\partial\lambda}. \tag{4}$$

Instead of the compatibility condition (3) we have

$$D_\eta U - D_\xi V + [U,V] = 0. \tag{5}$$

Suppose that

$$F = -\varepsilon \sum_{m=1}^{N_1} \frac{c_m}{\lambda-\lambda_m}, \quad G = -\varepsilon \sum_{m=1}^{N_2} \frac{b_m}{\lambda-\lambda_m}. \tag{6}$$

Then the evolutionary equations satisfying (5) become

$$\partial_\eta u_0 - \partial_\xi v_0 + [u_0, v_0] = 0,$$

$$\frac{\partial u_n}{\partial\eta} + \left[u_n, \sum_{k=1}^{N_2} \frac{v_k}{\lambda_n-\mu_k}\right] = \varepsilon \sum_{m=1}^{N_2} \frac{b_m u_n + c_n v_m}{(\lambda_n-\mu_m)^2},$$

$$\frac{\partial v_n}{\partial\xi} + \left[v_n, \sum_{k=1}^{N_1} \frac{u_k}{\mu_n-\lambda_k}\right] = \varepsilon \sum_{m=1}^{N_1} \frac{c_m v_n + b_n u_m}{(\lambda_m-\mu_n)^2}. \tag{7}$$

The compatibility condition requires that

$$F_\eta + GF_\lambda = G_\xi + FG_\lambda; \tag{8}$$

these relationships were obtained in Ref. 16 for exactly integrable models. In real physical systems the conditions (8) impose restrictions, and in most cases the physical parameters cannot be realized. Let us show that in an approximate theory the conditions (8) can be dropped.

Suppose that F and G are slow functions of their arguments:

$$F = \varepsilon f(\varepsilon\eta, \varepsilon\xi, \lambda), \quad G = \varepsilon g(\varepsilon\eta, \varepsilon\xi, \lambda), \tag{9}$$

where ε is a small parameter. The condition (8) is met for all f and g to within terms of order ε^2. With this accuracy the system of equations (7) satisfies the compatibility condition (5) for arbitrary perturbations on the right-hand side. Such perturbations can be studied by the inverse scattering method with allowance for the ''extension'' (4). We need only replace the Lax representation (1) with

$$D_\xi\Phi = U\Phi, \tag{10}$$

$$D_\eta\Phi = V\Phi. \tag{11}$$

This representation makes it possible to use the inverse scattering method for the perturbed system. In particular, by employing this method we can build a generalization of the Whitham equation directly in diagonal form.

3. DEFORMED WHITHAM EQUATIONS

As noted earlier, it is often impossible to analyze analytically the behavior of a packet consisting of a large number of pulses. Sometimes the dynamics of such a packet can be described by a slow packet modulation, which in turn is described by variations in the parameters of the packet of pulses. Here the parameters obey the Whitham equations. In deriving the Whitham equations one assumes that the scale of modulation with respect to both variables is much larger than the corresponding characteristics of the pulses comprising the packet. The inverse scattering method reduces the initially complex problem of deriving the Whitham equations to a universal procedure, and the resulting equations are diagonal. An important advantage of the suggested perturbative approach is that it allows deformed Whitham equations to be constructed diagonal form, too. Several ways of constructing Whitham equations for N-phase solutions have been suggested (see, e.g., Refs. 21–23). Here we will use a direct generalization of the approach proposed by Flaschka *et al.*[23] to the case of deformed systems.

Let us suppose that the general deformed AKNS system can be written in the form of the compatibility condition for the following overdetermined systems of equations:

$$D_\xi \Phi = \begin{pmatrix} L_{11} & L_{12} \\ L_{21} & -L_{11} \end{pmatrix} \Phi, \tag{12}$$

$$D_\eta \Phi = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & -A_{11} \end{pmatrix} \Phi. \tag{13}$$

Following Marchenko,[21] we introduce the quadratic functions

$$f = \frac{i}{2}(\phi_1\psi_2 + \phi_2\psi_1), \quad g = \phi_1\psi_1, \quad h = \phi_2\psi_2, \tag{14}$$

where $\phi_{1,2}$ and $\psi_{1,2}$ are distinct solutions of the systems (12) and (13). These functions satisfy the following system:

$$\begin{aligned} &D_\eta f = i(A_{12}h - A_{21}g), \quad D_\xi f = i(L_{12}h - L_{21}g), \\ &D_\eta g = 2iA_{12} f + 2A_{11}g, \quad D_\xi g = 2iL_{12} f + 2L_{11}g, \\ &D_\eta h = -2iA_{21} f - 2A_{11}h, \quad D_\xi h = -2iL_{21}f - 2L_{11}h. \end{aligned} \tag{15}$$

Clearly, Eqs. (15) imply that $P(\lambda) = f^2 - gh$ is constant, i.e., $D_\xi P(\lambda) = 0$ and $D_\eta P(\lambda) = 0$. The shape of the periodic wave is determined by the dependence of P on the spectral data. An N-phase solution is determined by the following polynomial:

$$f^2 - gh = P(\lambda) = \prod_{k=1}^{2N+2} (\lambda - \lambda_k) = \sum_{j=0}^{2N+2} P_j \lambda^j. \tag{16}$$

The roots λ_k are fixed by the initial conditions. The following representation is valid for the general N-phase solution for the AKNS system:

$$f = \sum_{k=0}^{N+1} f_k \lambda^k, \quad g = l(\lambda) \sum_{k=1}^{N+1} g_k \lambda^k, \quad h = p(\lambda) \sum_{k=1}^{N+1} h_k \lambda^k. \tag{17}$$

Here $l(\lambda)$ and $p(\lambda)$ are functions of the spectral parameter, such that $l(\lambda)$ $(p(\lambda))$ are common factors of L_{12} (L_{21}) and A_{12} (A_{21}). Following Refs. 21 and 24, we assume that

$$g = l(\lambda) \prod_{k=1}^{N+1} [\lambda - \mu_k(\xi, \eta)], \tag{18}$$

where μ_k are additional spectral variables. Plugging (18) into (15) and setting $\lambda = \mu_k$, we arrive at a system of equations for μ_k:

$$D_\xi \mu_k = \frac{2iL_{12}(\mu_k)\sqrt{P(\mu_k)}}{\Pi_{j\neq k}(\mu_k - \mu_j) l(\mu_k)}, \tag{19}$$

$$D_\eta \mu_k = \frac{2iA_{12}(\mu_k)\sqrt{P(\mu_k)}}{\Pi_{j\neq k}(\mu_k - \mu_j) l(\mu_k)}. \tag{20}$$

The next stage is to construct the Whitham equations for the deformed system, i.e., a system that can be represented by the compatibility condition for the linear systems (12) and (13). To this end we generalize the results of Refs. 23 and 24 directly. From (12) and (13) it follows that

$$D_\eta \left(\frac{L_{12}}{g} \right) = D_\xi \left(\frac{A_{12}}{g} \right). \tag{21}$$

We introduce a new normalization: $f^2 - gh = 1$. Then (21) can be written as

$$D_\eta \frac{L_{12}(\lambda)\sqrt{P(\lambda)}}{\Pi_{k=1}^{N+1}(\lambda - \mu_k) l(\lambda)} = D_\xi \frac{A_{12}(\lambda)\sqrt{P(\lambda)}}{\Pi_{k=1}^{N+1}(\lambda - \mu_k) l(\lambda)}. \tag{22}$$

The average over the rapid oscillations can be done by replacing the integral over the phase by an integral over the additional variables μ_k. The corresponding equations are found via a Jacobi transformation.[22,23] The Whitham equations are obtained as the conditions for the vanishing of the coefficients of the terms that become singular in the limit $\lambda \to \lambda_k$ (the λ_k are the roots of the polynomial P). A direct check shows that these conditions for the vanishing of the coefficients of the derivatives with respect to the variables ξ and η are the same as for the extended derivatives, i.e., the generalized equations remain diagonal. We write the final equations in the form

$$\frac{\partial \lambda_k}{\partial \eta} + V_k \frac{\partial \lambda_k}{\partial \xi} + F(\lambda_k) + V_k G(\lambda_k) = 0, \tag{23}$$

where

$$V_k = \left\langle \frac{2iL_{12}(\mu_k)}{\Pi_{j\neq k}(\mu_k - \mu_j) l(\mu_k)} \right\rangle \left\langle \frac{2iA_{12}(\mu_k)}{\Pi_{j\neq k}(\mu_k - \mu_j) l(\mu_k)} \right\rangle^{-1}. \tag{24}$$

The formal expression for the Whitham velocity (24) contains averages over the phases $W_j = k_j\xi + \omega_k\eta + w_0$, where w_0, k_j, and ω_k are constants. Equations (19) and (20) contain perturbing terms of order ε. Hence, if the equations are used for averaging, these terms yield a correction of order ε. We will assume that the spectral parameters λ_k vary slowly: their derivatives with respect to the variables are of order $O(\varepsilon)$. Note that this assumption agrees with the adopted hydrodynamic approximation. For this reason, allowance for

the perturbing terms in Eqs. (19) and (20) leads to an $O(\varepsilon^2)$ correction, which is ignored in the first approximation.

4. SOLITON GENERATION IN A TWO-LEVEL LASER

Let us use the above mathematical apparatus to analyze the generation of pulse packets in a two-level laser system. Some nonlinear processes that occur in laser systems can be described in terms of a two-level model in which the upper level is being pumped. Pumping can be taken into account by introducing a coefficient C (see below) into the right-hand side of the Bloch equation for the population difference. Such a model is used to describe gas and ion lasers and, less frequently, dye lasers.[4] The pumping results in a constant replenishment of the upper level. Relaxation effects play a certain role as the system sets in in the stationary regime. Allowing for such effects is important when we examine long pulse packets, even if the length of each pulse is small compared to the relaxation time. Note that the literature contains no analysis of the nature and dynamics of generation of ultrashort-pulse packets and of the effect of pumping and relaxation on this process. One reason, as noted earlier, lies in the substantial analytical difficulties encountered in such analysis. The mathematical apparatus developed in Secs. 2 and 3 makes the study of the behavior of pulse packets in amplification much simpler.

In this section we discuss two qualitatively different regimes of generation of soliton packets in a laser. The first is time-dependent and is accompanied by the development of modulational instability and simultaneous amplification. In the second regime the system reaches the steady regime of generation very fast.

The Maxwell–Bloch equation s for a two-level medium are

$$\partial_\tau Q+\gamma_2 Q+2i\nu Q=-\frac{id_{12}}{\hbar}EN, \tag{25}$$

$$\partial_\tau N_3+\gamma_1(N_3-N_0)=\frac{id_{12}}{2\hbar}(Q^*E-QE^*)+C_0, \tag{26}$$

$$\partial_z E=iN_0\frac{2\pi\omega_0}{c_l}\langle Q\rangle_\Gamma. \tag{27}$$

Here d_{12} is the dipole moment of the transition, N_3 is the difference of level populations, Q is the off-diagonal term of the density matrix, N_0 is the number density of resonant atoms, $\gamma_{1,2}$ are the relaxation constants, τ is the time lag, ω_0 is the transition frequency, and c_l is the speed of light. The angle brackets stand for averaging over the velocity distribution of the ensemble of atoms: $\langle Q\rangle_\Gamma=\int Q(\nu)\Gamma(\nu)d\nu$. Here we examine the case where the generated pulses are much shorter than the relaxation times $\gamma_{1,2}^{-1}$. Pumping is modeled by the function $C_0(z)$ on the right-hand side of Eq. (26). Below we introduce the following notation:

$$t=\tau\sqrt{\Omega},\quad x=\frac{z\sqrt{\Omega}}{c_l},\quad \Omega=\frac{2\pi N_0\omega_0 d_{12}}{\hbar},\quad C=\frac{C_0}{\sqrt{\Omega}}.$$

Let us apply the procedure discussed in Sec. 3 to the system of Maxwell–Bloch equations (25)–(27). The Lax representation for this system has the form of (12) and (13), where

$$L_{11}=-i\lambda,\quad L_{12}=\frac{E}{2},\quad L_{21}=\frac{E^*}{2},$$

$$A_{11}=\frac{i}{4}\left\langle\frac{N_3}{\lambda-\nu}\right\rangle_\Gamma,\quad A_{12}=-\frac{i}{4}\left\langle\frac{Q}{\lambda-\nu}\right\rangle_\Gamma,$$

$$A_{21}=-\frac{i}{4}\left\langle\frac{Q^*}{\lambda-\nu}\right\rangle_\Gamma. \tag{28}$$

The system (25)–(27) admits a Lax representation either when $C(x)$ is arbitrary and $\gamma_{1,2}=0$ holds or for $\gamma_1=\gamma_2\neq 0$ and $C=0$ (see Ref. 16). Nonzero values of C and $\gamma_1=\gamma_2$ are possible in certain conditions,[16] but it is impossible to meet these conditions in the real physical world. Within the approximate theory developed above such conditions can be avoided. More than that, we can study a range of parameters in which reduction of the integrable model with a variable spectral parameter is not achieved. For instance, in the present approach we can study effects associated with slow arbitrary variations of pumping in time.

We select the following perturbation:

$$F(\lambda)=\varepsilon\frac{c(x,t)}{\lambda},\quad G(\lambda)=\varepsilon\gamma\lambda,$$

where ε is a small parameter. The first function corresponds to pumping and the second, to relaxation effects. The condition $\gamma_1=\gamma_2$ is met for a gas of metal atoms. We assume that the theory makes it possible to establish the rules valid for distinct relaxation constants γ_1 and γ_2, too. The present author does not know of any general N-phase solutions of the Maxwell–Bloch model. However, such solutions clearly must have the same structure as the solutions of the nonlinear Schrödinger equation. The difference can only be in the values of the phase velocities. Let us use the well-known single-phase solution of the unperturbed model of Marchenko[21] and Kamchatnov and Pavlov[25] for the field strength. For the two pairs of complex-conjugate roots of the polynomial P, i.e., for $\lambda_{1,3}=\alpha\pm i\beta$ and $\lambda_{2,4}=\alpha_0\pm i\beta_0$, we have

$$\begin{aligned}|E(x,t)|^2&\\ &=|E(0,0)|^2[(\beta+\beta_0)^2-4\beta\beta_0\mathrm{sn}^2\\ &\times\{\sqrt{(\beta+\beta_0)^2+(\alpha-\alpha_0)^2}W,\kappa\}],\end{aligned} \tag{29}$$

where $\kappa=4\beta\beta_0/[(\beta+\beta_0)^2+(\alpha-\alpha_0)^2]$. The phase W can be found from (19) and (20):

$$W=t+xV_0^{-1}+t_0,\quad V_0=-\int\frac{\Gamma(\nu)}{4\Pi_{i=1}^4(\lambda_i-\nu)}d\nu. \tag{30}$$

For a single-phase solution we must set $N=2$ and $l(\lambda)=1$ in (15) and (17). From Eq. (21) we find that

$$D_\eta\left(\frac{\sqrt{P(\lambda)}}{\lambda-\mu}\right)=D_\xi\left[\frac{\sqrt{P(\lambda)}}{V_0}\left(\frac{1}{\lambda-\mu}-\left\langle\frac{1}{\lambda-\nu}\right\rangle\right)\right], \tag{31}$$

where the angle brackets stand for averaging over the period T of the rapid oscillations:

$$T=\int dW=\int \frac{d\mu}{\varepsilon(c/\mu+\gamma\mu)+\sqrt{P(-\mu)}}. \tag{32}$$

Formula (32) follows from (19) and (20). As noted earlier, the perturbing terms provide a contribution to the Whitham equations, which we ignore. Hence in (32) we can put $c=\gamma=0$. Then

$$T=\frac{2K(k)}{\sqrt{(\lambda_1-\lambda_3)(\lambda_2-\lambda_4)}},$$

where $K(k)$ is the complete elliptic integral of the first kind with modulus k,

$$k^2=\frac{(\lambda_1-\lambda_2)(\lambda_3-\lambda_4)}{(\lambda_1-\lambda_3)(\lambda_2-\lambda_4)},$$

and the λ_k are the roots of the polynomial P, with $\lambda_1>\lambda_2>\lambda_3>\lambda_4$. We assume that the λ_k are slow functions of their variables: $\lambda_k=\lambda_k(\varepsilon z,\varepsilon t)$. Averaging according to the formula

$$\left\langle\frac{1}{\lambda-\mu}\right\rangle=\frac{1}{T}\int\frac{d\theta}{\lambda-\mu}=\frac{1}{T}\int\frac{d\mu}{(\lambda-\mu)\sqrt{P(-\mu)}}, \tag{33}$$

we arrive at the deformed Whitham equation:

$$\partial_x\lambda_n-\frac{1}{V_n}\partial_t\lambda_n=\frac{\gamma}{V_n}\lambda_n+\frac{c}{\lambda_n}, \tag{34}$$

where

$$\frac{1}{V_n}=\frac{1}{V_0}\left[1-\left(\lambda_n\left\langle\frac{1}{\lambda_n-\mu}\right\rangle\right)^{-1}\right],$$

$$\left\langle\frac{1}{\lambda_1-\mu}\right\rangle=\frac{(\lambda_2-\lambda_3)E(k)+(\lambda_1-\lambda_2)K(k)}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)K(k)},$$

$$\left\langle\frac{1}{\lambda_2-\mu}\right\rangle=\frac{(\lambda_1-\lambda_2)K(k)+(\lambda_1-\lambda_4)E(k)}{(\lambda_1-\lambda_2)(\lambda_3-\lambda_4)K(k)},$$

$$\left\langle\frac{1}{\lambda_3-\mu}\right\rangle=\frac{(\lambda_2-\lambda_4)E(k)+(\lambda_2-\lambda_3)K(k)}{(\lambda_2-\lambda_3)(\lambda_3-\lambda_4)K(k)},$$

$$\left\langle\frac{1}{\lambda_4-\mu}\right\rangle=\frac{(\lambda_1-\lambda_3)E(k)+(\lambda_1-\lambda_4)K(k)}{(\lambda_1-\lambda_4)(\lambda_3-\lambda_4)K(k)}, \tag{35}$$

with V_0 given by (30). For $\Gamma(\nu)=\delta(\nu)$ we have $V_0=1/4\sqrt{P_0}$, where $P_0=\lambda_1\lambda_2\lambda_3\lambda_4$. In (35) $E(k)$ is the complete elliptic integral of the second kind with the same modulus k. The above formulas generalize the Whitham equations obtained by Kamchatnov and Pavlov[25] to the case of perturbed systems.

5. SOLUTIONS OF THE WHITHAM EQUATIONS FOR AN AMPLIFYING MEDIUM

The evolutionary equations (34) have different types of solution. Let us describe the possible scenario of soliton generation when an infinitely long square pulse is injected into the system. We assume that in the process of evolution of the pulse in the medium, the leading edge of the pulse decays into a sequence of solitons. As the packet of solitons travels through the active medium, the solitons become amplified. Suppose for simplicity that γ is a constant and c depends only on x. Then, using the ansatz

$$\lambda_k=e^{-\gamma t}\sqrt{\zeta_k+2\int c(x)dx}, \tag{36}$$

we can reduce Eqs. (34) to ordinary homogeneous equations for the hidden parameters $\zeta_k(x,t)$. Thus, the problem reduces to describing modulational instability of a step pulse propagating in an attenuator and to a corresponding change of parameters.

Modulational instability of a step pulse has been examined by many researchers[5] (see also Ref. 26). The common way to tackle the problem is to use the heuristic Whitham approach based primarily on the single-phase solution of the model. This solution is modulated so the plane-wave and soliton solutions can be matched. It is often assumed in the process that the positive asymptotic solution consists of a sequence of solitons with an amplitude that increases as $x\to\infty$. For the Maxwell–Bloch model this assumption is not always valid. Let us take the unperturbed spectral problem (13). Suppose that an infinitely long rectangular pulse of field $E(0,t)$ with an amplitude $2A$ is injected into a semi-infinite unperturbed medium at the point $x_0=0$. The solution of the spectral problem (13) can easily be found. The continuous spectrum related to the infinitely long pulse consists of the real axis and a segment of a straight line perpendicular to the real axis and symmetric to this axis. The vertex of the segment lying in the imaginary part of the spectrum corresponds to a value of the spectral parameter ξ_t equal to iA. The imaginary part of the spectrum corresponds to the soliton asymptotic limits. If we consider a semi-infinite step as the limit of a rectangular pulse, we can easily show that a soliton related to a point $\xi_1<iA$ of the spectrum that lies in the $[0,+iA]$ interval has a smaller amplitude and a higher velocity.

Generally, in analyzing the characteristics of the generated solitons the following conditions must be met: (a) the Jost functions (the solutions of the spectral problem (12)) must be analytic in the upper half-plane, and (b) v_s must be smaller than c_l, where v_s is the phase velocity of a soliton. A simple analysis of the phase W in (30) shows that a soliton with a larger amplitude has a smaller phase velocity. As a result, the decay of such a step pulse is accompanied by forward-attenuated pulsations. Numerical calculations have shown that there is a limited range of initial conditions under which intense solitons can be generated at the leading edge of the step. This is possible only if there is partial inversion of the medium: $Q(x,0)$ is nonzero.

Another method of generating solitons at the leading edge consists in injecting into the system a periodic wave of a special type: its associated spectrum consists of segments of curves symmetric with respect to the real axis and positioned in such a way that the distance from the point (0,0) in the complex plane to a point in the spectrum on a segment of a curve increases as the point moves from the segment's vertex toward the real axis.

Let us examine the case of finite initial polarization. With initial inversion of the medium we must account for the contribution of the continuous spectrum to the generation process, since this spectrum participates in the formation of the intense pulses. For the Maxwell–Bloch model with pumping and without relaxation (see Eqs. (25)–(27)), the contribution of the continuous spectrum to the asymptotic solution was calculated by Burtsev *et al.*,[27] who found that the field has a self-similar asymptotic solution with an amplitude that grows in proportion to $\sqrt{x}$. This solution depends on the self-similar variable $\sigma=\sqrt{2x(c_l t-x)}$ (see Ref. 27).

Here is an estimate of the phase velocity of an amplified soliton. For a fixed point σ_0 at the leading edge we find that

$$x-c_l t=-\sigma_0 t^{-1}+o(t^{-1}),$$

where σ_0 is a constant. We estimate the soliton velocity at the leading edge of the packet. To this end we construct a solution corresponding to the point ζ_1 of the spectrum. There are several methods of building soliton solutions,[5] and using them is easy. A single-soliton solution of the model (25)–(27) for a zero field at $x\to\pm\infty$ has the form

$$E(x,t)=4\ \mathrm{Im}\ \zeta_1\frac{\exp(-2i\ \mathrm{Im}\ \phi+ia_0)}{\cosh(2\ \mathrm{Re}\ \phi+a_1)},$$

$$\phi=\frac{\sqrt{2cx+\zeta_1}(1-e^{-\gamma t})}{\gamma}-\frac{\sqrt{2cx+\zeta_1}-\sqrt{\zeta_1}}{4c}. \tag{37}$$

The asymptotic behavior of this solution is

$$x=c_l t=c_1-c_2 t^{-1/2}+o(t^{-1/2}),$$

where $a_{0,1}$ and $c_{1,2}$ are constants, with

$$c_1=\frac{c_l}{\gamma}\left(1+\frac{\gamma}{4c_l}\right),\quad c_2=\frac{\sqrt{c_l}}{\sqrt{8}(1+\gamma/4c_l)}.$$

The above reasoning suggests that for large times the soliton packet and the fraction of the generated pulses related to the continuous spectrum separate in time. The soliton packet lags behind. Analyzing the phase W, we find that solitons can split away from the trailing edge of the pulse and the distance between such solitons slowly increases, with the last solution having the largest amplitude (Fig. 1). In this case the conditions (a) and (b) for generating intense solitons at the trailing edge of the pulse are met. The Whitham equations can be used to describe the decay of the trailing edge of a step pulse. This process is described by the dynamics of the hidden spectral data ζ_k, the roots of the polynomial P. Suppose that the roots ζ_1 and ζ_3 of the polynomial P are fixed: $\zeta_1=\alpha_0+i\beta_0$ and $\zeta_3=\alpha_0-i\beta_0$, with α_0 and β_0 real constants. The dynamics of the other two (moving) roots, ζ_2 and ζ_4 is described by the self-similar solution of Eqs. (34). Suppose that $\zeta_2=\alpha+i\beta$ and $\zeta_4=\alpha-i\beta$. Then Eqs. (38) reduce to

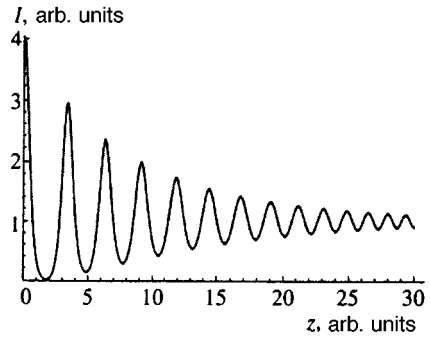


FIG. 1. Transformation of the trailing edge into a packet of pulses without allowance for amplification and relaxation. The diagram shows the dependence of the field's intensity I on the spatial coordinate z.

$$\frac{t}{x}=\frac{1}{\sqrt{P_0}}\left\{1-\frac{1}{\alpha+i\beta}\times\frac{2i\beta[\alpha_0-\alpha+i(\beta_0-\beta)]K(\kappa)}{[\alpha_0-\alpha+i(\beta_0-\beta)]K(\kappa)-[\alpha_0-\alpha+i(\beta_0+\beta)]E(\kappa)}\right\}. \tag{38}$$

Separating the imaginary and real parts in (38), we find that

$$R(\kappa)=\frac{E(\kappa)}{K(\kappa)}=\frac{\alpha(\alpha_0^2+\alpha^2+\beta_0^2+\beta^2)-2\alpha(\alpha_0\alpha+\beta_0\beta)}{\alpha(\alpha_0^2+\alpha^2+\beta_0^2+\beta^2)-2\alpha_0(\alpha^2+\beta^2)}, \tag{39}$$

$$\left(\frac{t}{x}\sqrt{P_0}-1\right)(\alpha^2+\beta^2)(\alpha_0-\alpha)^2[1-R(\kappa)]+[\beta_0-\beta+(\beta_0+\beta)R(\kappa)]^2=4\beta(\alpha_0-\alpha)(\alpha_0\beta-\alpha\beta_0)(1-R(\kappa))+[\beta_0-\beta+(\beta_0+\beta)R(\kappa)](\beta\beta_0-\beta^2+\alpha\alpha_0-\alpha^2). \tag{40}$$

Equations (38) and the expression for the modulus κ of the elliptic functions,[28]

$$\kappa^2=\frac{4\beta\beta_0}{(\alpha_0-\alpha)^2+(\beta_0+\beta)^2}, \tag{41}$$

implicitly specify the dependence of the spectral data on the self-similar variable x/t. This dependence has the shape usually found in related problems.[26] As x/t increases, the roots ζ_2 and ζ_4 monotonically change from ζ_1 and ζ_3, respectively, to zero values. The dynamics of pulse packets (without amplification) corresponding to such behavior of the roots is depicted in Fig. 1. Numerical calculations show that if we allow for amplification and relaxation, a weak periodic wave in the time-dependent regime is amplified and transforms into a dense packet of narrow pulses (solitons); see Fig. 2. Asymptotically the amplitude of these solitons diverges like $\sqrt{x}$.

Here is another mechanism of soliton generation in which relaxation plays a leading role (together with amplification). A distinctive feature of this mechanism is that the regime of generation of asymptotic solitons of the same am-

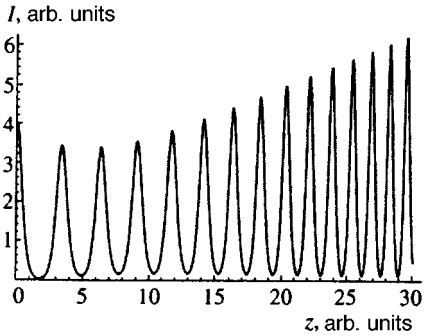


FIG. 2. The same as in Fig. 1 but with allowance for amplification and relaxation; $\varepsilon c = 0.05$ and $\varepsilon\gamma = \varepsilon\gamma_{1,2} = 0.05$.

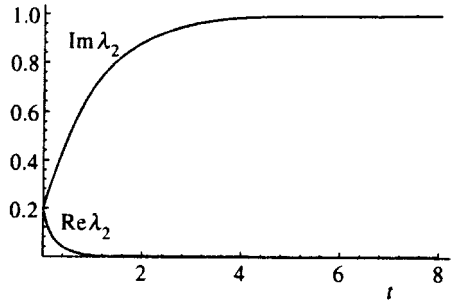


FIG. 3. Time dependence of the real and imaginary parts of the spectral parameter λ_2 in the second soliton generation mechanism. Time is measured in units of $(\epsilon\gamma)^{-1}$.

plitude becomes quasistationary very fast. Let λ_s (λ_s^*) be the value of the spectral parameter λ_1 (λ_3) at which the right-hand side of Eq. (34) vanishes. This point is a stable focal point in the phase plane. Figure 3 depicts the dependence of the real and imaginary parts of λ_2 on the variable x. For the initial field we take a low-amplitude harmonic wave. Suppose (as we did earlier) that two roots, λ_1 and λ_3, vary in an interval whose origin corresponds to the plane-wave solution and whose end point, to the soliton limit. Then, as λ_2 varies from zero to λ_s, the harmonic wave is transformed into a sequence of asymptotic solitons of the same amplitude (Fig. 4). Numerical analysis of the phase velocity of the leading soliton shows that the velocity monotonically tends to the speed of light as the pumping C increases.

In real experiments it is often impossible to observe the details of formation of a pulse packet in one pass of the cavity. At the same time, investigating the initial stage in amplification is important for studies of the characteristics of the generated pulses. Knowledge of the dependence of the group velocity on the coordinates, time, and pumping makes it possible to use the laser for further amplification and transformation of the generated pulses more effectively. Hence establishing the mechanism of soliton generation and its features is important for practical reasons.

Here is a qualitative comparison of our results with the experimental data taken from the literature. Apolonskiĭ[29] studied the dependence of the velocity of the generated pulses on the amplification length. He found, in particular, that the pulse velocity increases monotonically with the amplification length. This property is common for the two generation regimes considered above. In the first case the system does not reach the stationary regime. The amplitude of the generated pulses asymptotically increases in proportion to $\sqrt{x}$. The second mechanism of generation predicts that the system reaches the quasistationary soliton generation regime very fast. The soliton solution is an attractor. For strong soliton pulses to form it is enough to inject a packet of small height but large length into the medium.

Let us now estimate the values of the physical parameters for which soliton generation can be observed. For a carbon-dioxide laser the relaxation time T_2 is approximately 10 ps and the dipole moment d_{12} of the vibrational–rotational transition is roughly 3×10^{-20} esu. Suppose that the gas pressure is 10 atm. As a result of amplification, pulses several picoseconds long and with a radiation intensity of order 1 TW cm^{-2} are emitted by the laser. This means that tens or even hundreds of ultrashort pulses simultaneously exists over the length of the laser, which amounts to tens of centimeters. The amplification of these pulses can be described as modulation of the parameters of the packet as a whole. Here, if we assume that amplification is described by the second mechanism of generation, the duration of one pass through the active medium of the laser is sufficient for the stationary regime to be attained. For Nd:YAG and He–Ne lasers the ratio of the soliton size (in time) to the time it takes the soliton to travel through the cavity is estimated via experiments at 0.02–0.01 (see Ref. 4). In these experiments it was discovered that the soliton amplitude reaches a value close to the stationary one during a time less than the time of one pass through the cavity. If we assume that the transverse relaxation time T_2 is the minimum relaxation parameter, a crude estimate shows that in the second mechanism of generation the time it takes the pulse packet to pass through the cavity once is sufficient for these lasers to reach the stationary regime (see the literature cited in Ref. 4).

We believe that the last generation scenario agrees qualitatively with the experimental data on generation of pulses in gas and ion lasers.[4] The relationship between pulse length and relaxation time used in numerical calculations holds in experiments involving gas and ion lasers. With the second mechanism one can find the dependence of the amplitude of the generated solitons on C. In ion lasers, C is determined by the current used to excite the ions. Figure 5 depicts the dependence of the amplitude of the generated asymptotic soli-

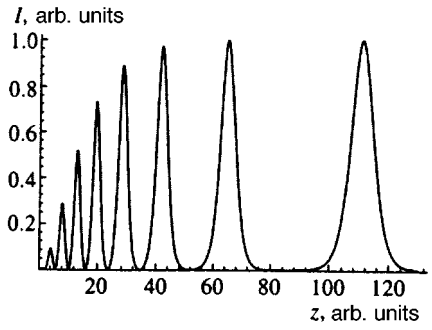


FIG. 4. Transformation of a weak initial harmonic wave to a soliton packet in the second soliton generation mechanism.

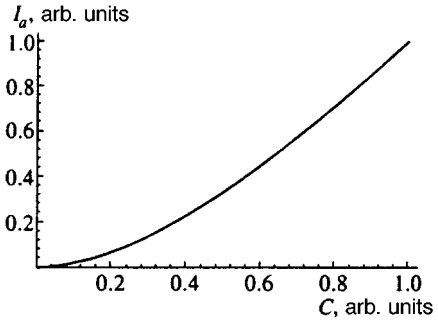


FIG. 5. The square of the amplitude of asymptotic (as $z \to \infty$) solitons, I_a, as a function of pumping C.

tons on C. This dependence can be used to verify the theory qualitatively and must hold for pulses whose intensity is far from saturation.

In this paper we have proposed an approximate approach to analyzing the dynamics of dense pulse packets propagating in nonlinear media. We believe that this method makes it possible to study analytically the initial stages of coherent nonlinear processes in nonlinear media (including amplifying media). In more general cases we must incorporate N-phase solutions and allow for saturation. This can be done, at least partially, by generalizing the results of the present work. The results can also be used to describe Raman scattering and four-wave mixing[30,31] with allowance for relaxation and pumping, and in the adopted approach it is possible to avoid imposing formal restrictions on physical parameters.[32]

I am grateful to Heinz Steudel for discussing the topics examined in this paper. I am also extremely grateful to A. A. Apolonskiĭ for numerous stimulating discussions of the laser experiments and the corresponding theoretical aspects. The work was sponsored by the Russian Fund for Fundamental Research (Grant No. 95-02-04392) and Deutsche Forschungsgemeinschaft (Grant 426 RUS 113/89/0(R,S)).

[1] Y. R. Shen, *The Principles of Nonlinear Optics*, Wiley, New York (1984).
[2] A. I. Maimistov, A. M. Basharov, S. O. Elyutin, and Yu. M. Sklyarov, Phys. Rep. **191**, 1 (1991).
[3] A. C. Newell and J. V. Moloney, *Nonlinear Optics*, Addison-Wesley, Redwood City, CA (1992).
[4] A. A. Apolonsky, A. A. Zabolotskii, V. P. Drachev, and E. I. Zinin, Proc. SPIE **2041**, 385 (1993).
[5] S. V. Manakov, S. P. Novikov, L. P. Pitaevskiĭ, and V. E. Zakharov, *Theory of Solitons: The Inverse Scattering Method*, Consultants Bureau, New York (1984).
[6] S. V. Manakov, Zh. Éksp. Teor. Fiz. **83**, 68 (1982) [Sov. Phys. JETP **56**, 37 (1982)].
[7] I. R. Gabitov, V. E. Zakharov, and A. V. Mikhaĭlov, JETP Lett. **37**, 279 (1983).
[8] I. R. Gabitov, V. E. Zakharov, and A. V. Mikhaĭlov, Zh. Éksp. Teor. Fiz. **86**, 1204 (1984) [Sov. Phys. JETP **59**, 703 (1984)].
[9] I. R. Gabitov, V. E. Zakharov, and A. V. Mikhaĭlov, Teor. Mat. Fiz. **63**, 11 (1985).
[10] S. V. Manakov and V. Yu. Novokshenov, Teor. Mat. Fiz. **69**, 40 (1986).
[11] I. R. Gabitov and S. V. Manakov, Phys. Rev. Lett. **50**, 495 (1983).
[12] H. Steudel, Quantum Opt. **2**, 387 (1990).
[13] D. J. Kaup and C. R. Menyuk, Phys. Rev. A **42**, 1712 (1990).
[14] C. R. Menyuk, Phys. Rev. A **47**, 2235 (1993).
[15] G. Whitham, *Linear and Nonlinear Waves*, Wiley, New York (1974).
[16] S. P. Burtsev, A. V. Mikhaĭlov, and V. E. Zakharov, Teor. Mat. Fiz. **70**, 323 (1987).
[17] V. A. Belinskiĭ and V. E. Zakharov, Zh. Éksp. Teor. Fiz. **75**, 1955 (1978) [Sov. Phys. JETP **48**, 985 (1978)].
[18] V. A. Belinskiĭ and V. E. Zakharov, Zh. Éksp. Teor. Fiz. **77**, 3 (1979) [Sov. Phys. JETP **50**, 1 (1979)].
[19] A. V. Mikhaĭlov and A. I. Yaremchuk, JETP Lett. **36**, 95 (1982).
[20] A. V. Mikhailov and A. I. Yaremchuk, Nucl. Phys. B **82**, 202 (1982).
[21] V. A. Marchenko, Mat. Sb. (N.S.) **95**, 331 (1974).
[22] B. A. Dubrovin, V. B. Matveev, and S. P. Novikov, Russian Math. Surveys **31**, 59 (1976); E. Date and S. Tanaka, Prog. Theor. Phys. Suppl. **59**, 107 (1976); E. Date, Prog. Theor. Phys. **59**, 265 (1978).
[23] H. Flaschka, M. G. Forest, and D. W. McLaughlin, Commun. Pure Appl. Math. **68**, 739 (1980).
[24] M. G. Forest and D. W. McLaughlin, J. Math. Phys. **27**, 1248 (1982).
[25] A. M. Kamchatnov and M. V. Pavlov, Zh. Éksp. Teor. Fiz. **107**, 44 (1995) [JETP **80**, 22 (1995)].
[26] A. M. Kamchatnov, submitted to Phys. Rep. (1997).
[27] S. P. Burtsev, I. R. Gabitov, and V. E. Zakharov, in *Plasma Theory and Nonlinear and Turbulent Processes in Physics*, World Scientific, Singapore (1988), p. 897.
[28] *Handbook of Mathematical Functions*, edited by M. Abramowitz and I. A. Stegun, National Bureau of Standards Applied Mathematics Series 55, Washington, D.C. (1964).
[29] A. A. Apolonskiĭ, Opt. Spektrosk. **61**, 564 (1986) [Opt. Spectrosc. **61**, 564 (1986)].
[30] A. A. Zabolotskii, Physica D **40**, 283 (1989).
[31] A. A. Zabolotskii, Phys. Rev. A **50**, 3384 (1994).
[32] S. P. Burtsev and I. R. Gabitov, Phys. Rev. A **49**, 2065 (1994).

Translated by Eugene Yankovsky

Dynamics of two-dimensional radiating vortices described by the nonlinear Schrödinger equation

I. A. Ivonin

Russian Research Center ''Kurchatov Institute,'' 123182 Moscow, Russia
(Submitted 16 May 1997)
Zh. Éksp. Teor. Fiz. **112**, 2252–2262 (December 1997)

The paper considers the dynamics of dark charged solitons (vortices) described by the two-dimensional (2D) nonlinear Schrödinger equation (NSE) with a repulsive potential. The dynamics of these point-like vortices in the NSE is quite different in comparison with the vortices in an incompressible liquid because of the possibility of wave-like emission of energy, momentum, and angular momentum in the first case. Another important feature is the characteristic scale of the problem, namely the screening parameter. Related problems of the collapse of a vortex dipole and the decay of a multicharged vortex in a region bounded by an absolutely reflecting shell are investigated both analytically and numerically. The conditions and scaling of a vortex dipole collapse and the limitations on the decay of a multicharge dipole in a bounded region are obtained. © *1997 American Institute of Physics.* [S1063-7761(97)02412-8]

1. INTRODUCTION

The nonlinear Schrödinger equation with a repulsive potential

$$i\Psi_t+\frac{1}{2}\Delta\Psi-U(|\Psi|^2)\Psi=0 \tag{1}$$

describes[1] propagation of modulated ion-acoustic waves in plasma ($U=|\Psi|^2$), nonlinear waves in optical fibers with a ''normal'' dependence of the refractive index on the light intensity ($U=|\Psi|^2+\alpha|\Psi|^4$), three-dimensional diffraction patterns of a laser beam transmitted through a diffraction grating and scattering media ($U=|\Psi|^2$), etc. Moreover, Eq. (1) has been applied recently to describe excitations in a Bose-condensate.[2–4]

In a one-dimensional case, this equation can be integrated[5] if $U=|\Psi|^2$; it has a set of particular solutions with one parameter in the form of dark solitons,[1,6] i.e., exponentially screened depressions in the field amplitude $|\Psi|^2$ with a nonzero asymptotic background value. The depth of the depressions (the parameter of the solution) uniquely determines the velocity of dark solitons. An important property of these particular solutions is that in the case of the nonintegrable NSE they are, like ordinary solitons,[7] nonlinear attractors, i.e., their interaction with linear waves of small amplitudes does not change their shapes, but only increases their amplitudes.[8]

In the two-dimensional case, one-dimensional dark solitons are unstable against transverse excitations.[9] Growth of this self-focusing instability leads to the formation of bounded two-dimensional solutions.[3] To interpret these local solutions, the ''hydrodynamic'' ansatz (Madelung) is quite convenient. In particular, in field $\Psi=\sqrt{\rho}\exp(i\phi)$, the squared amplitude ρ and phase ϕ have the sense of the density and velocity potential of some iso-frozen compressible medium,

$$\frac{\partial\rho}{\partial t}+\mathrm{div}(\rho\nabla\phi)=0,$$

$$\frac{\partial\phi}{\partial t}+\frac{1}{2}(\nabla\phi)^2=-h, \tag{2}$$

with nonlocal specific enthalpy h. In what follows, we select for definiteness the potential in Eq. (1) in the form $U=|\Psi|^2-1$ with the background (unperturbed) field amplitude and density $|\Psi|^2=\rho=1$. In this case, the enthalpy in the Bernoulli equation (2) has the form

$$h=\rho-1-\frac{\nabla^2\sqrt{\rho}}{2\sqrt{\rho}}.$$

The uniqueness condition on the field Ψ requires only that the velocity potential be defined to within $2\pi N$. The possible branch points of ϕ correspond to point vortices with a discrete amplitude N and velocity potential in the vicinity of a vortex $\phi=N\theta$, where θ is the azimuthal angle. A uniquely defined velocity potential corresponds to purely potential flows, such as acoustic waves.

Conservation (the freezing-in condition) of point vortex amplitudes gives rise to conservation of the topological charge N of the corresponding dark solitons described by Eq. (1). Unlike one-parameter one-dimenional dark solitons, ''charged'' ($N\neq 0$) two-dimensional dark solitons have no such parameter because the presence of a branch point implies that the maximum depth (a drop to $\Psi=0$) is attained at the vortex center. However, once generated, a vortex cannot disappear except through collapse with another vortex of opposite charge. Potential fields of arbitrary amplitudes cannot eliminate an isolated vortex, and this property is the source of the keen interest in studying the dynamics of charged solitons.

A vortex pair (a combination of two dark 2D solitons with opposite charges) has, like 1D solitons, one parameter, namely the distance between the vortices. This parameter uniquely determines the integrals of motion and velocity of a uniformly propagating dipole[2] and is convenient in describing evolution of the variable vortex solutions discussed below.

Linearization of the continuity Bernoulli equations (2) about a constant background $\rho=1$ yields the phase velocity of acoustic waves

$$V_{\mathbf{k}}=\sqrt{1+\frac{\mathbf{k}^2}{2}}>1,$$

propagating in the unperturbed medium. Therefore we can conclude that stable localized steady solutions of Eq. (1) moving at a supersonic velocity $V>C_s^0=1$ are forbidden, since otherwise the Cherenkov resonance condition would be easily satisfied, and emission of waves would lead to variability.

At the same time, an isolated point vortex generates in its environment rotations of the medium with linear velocity $V\propto 1/r$, varying with the distance between the vortex and the observation point. Therefore, there are no solutions describing dipoles moving at a constant velocity for which the distance between the vortices is smaller than a critical value. This statement was verified[2] by a numerical calculation of parameters of the uniformly moving solutions: it turned out that there are no singly charged vortex dipoles moving at constant velocities $V/C_s^0>1/\sqrt{2}$. In the following section, this statement will be confirmed by analytical calculations.

Vortices in a dipole moving at a large velocity copiously shed potential waves and approach one another. In the next section, such a motion of a vortex pair with due account of radiation will be analyzed numerically using the NSE with a repulsive quadratic potential. This allows us to determine the scaling of the collapse at the point z_0 at time t_0 of two oppositely charged dark solitons separated by a distance $L(t)$ and moving in the z direction: $L\sim\sqrt{z_0-z}$, $z_0-z=(t_0-t)/\sqrt{2}$. It is noteworthy that the complex Ginzburg–Landau equation, whose structure is very similar to that of NSE, also predicts collapse of two vortices in a dipole when the distance between them becomes sufficiently small.[10] The analogy between these cases is, obviously, not direct, because vortices described by NSE are very different from those described by the Ginzburg–Landau complex equation (where vortices repel one another at large distances, and the interaction intensity between them decays exponentially slowly[10]).

Another interesting problem is a decay of an isolated multicharged ($N\gg 1$) dark soliton. A dark solution with a charge $N>1$ is inherently unstable and decays to vortices of unit amplitude initially located at vertices of a regular polygon, as has been confirmed numerically many times.[11] Then, as in hydrodynamics, if the number of vortices satisfies $N>7$, a ''polygon instability'' develops,[12] and the vortices are distributed randomly over a certain area. If the number of vortices is large, the flow pattern is similar to the velocity field due to a vortex of finite dimensions. Given appropriate boundary conditions (a rigid shell of finite dimensions),[13] a Kelvin vortex (a vortex shaped as a circular cylinder with a uniform distribution of vorticity) is Lyapunov-stable in three dimensions even if emission of waves is taken into account. Thus, the decay of a multiply charged dark soliton in a bounded region can at least be retarded considerably, if not stopped completely. Details of this process will be discussed in the last section of the paper.

Thus, in this study we have attempted to investigate both analytically and numerically the nonlinear dynamics of both oppositely and similarly charged topologically stable 2D dark NSE solitons (vortices) with a repulsive potential.

2. COLLAPSE OF A VORTEX DIPOLE IN THE NSE MODEL

Unlike the dynamics of point vortices in an incompressible liquid, the vortex dynamics in the NSE model is not determined solely by their charges and instantaneous positions. The motion of an isolated vortex is controlled by the potential fields generated at its site. The problem of vortex dynamics is formulated in a self-consistent manner only if we know that these potential waves (their amplitudes, wave vectors, and phases) are due to the evolution of vortices. A similar case is evolution of radiation emitted by charged particles. The problem of their dynamics is well-posed only in the absence of an external electromagnetic field.

The similarity between the 2D dynamics of charged particles interacting through the Coulomb potential and dynamics of vortices is complete only in a incompressible liquid, when the Hamiltonians of a quasi-stationary system of charges and of a system of point-like vortices in the Lamb representation are identical.[14] If radiation is taken into account, at least, in the first approximation in the parameter V/C_s, this similarity persists. So, the Maxwell equations combined with the gauge condition are transformed to linearized hydrodynamic equations (2) if charges are replaced with vortex intensities, and the magnetic field (which has only one component in the 2D configuration) are replaced with the deviation of the density from its background value ($\rho/\rho_0=B_z$). The electric field vector directed at 90° with respect to the magnetic field determines the vortex velocity:

$$\mathbf{V}=\mathbf{e}_z\times\mathbf{E}.$$

Specifically, the divergence of this vector yields, on one hand, a continuity equation linearized about a constant background density ρ_0, and on the other hand, an equation for the electric field circulation. The equation for the magnetic field circulation corresponds to the Euler equation linearized about a constant background density. In addition, the equation for the electric field divergence in combination with the gauge condition yields a relationship between the vorticity (vortex intensity) and a function of the velocity flux (corresponding to the electric field potential) in the form of retarded potentials. Thus, the electrodynamic equation for an electric field generated by arbitrarily moving electric charges[15] is a full analogue of an integral equation in the form of retarded potentials for a hydrodynamic velocity field generated by mobile emitting vortices. This equation (it is not given in the paper because it is too complicated[15]) is closed, i.e., the field derived from this equation uniquely determines the velocities of all vortices. In the specific case of a uniformly propagating pair of similarly charged vortices ($N=1$) with a distance L between them, we obtain an equation for the vortex velocity

$$V=\frac{1}{L}\frac{1}{\sqrt{1-V^2}} \tag{3}$$

(recall that $C_s^0=1$ if the potential in Eq. (1) is selected in the form $U=|\Psi|^2-1$).

It is worth noting that Eq. (3) is approximate since the sound velocity is not constant, $C_s^2\equiv\rho+\Delta\sqrt{\rho}/4\sqrt{\rho}$, and applies only to the case $L\gg 1$. Nonetheless, it yields a qualitatively correct result even when $L\sim 1$, as will be proved below.

The only difference between Eq. (3) and the usual formula for a vortex pair velocity in an incompressible liquid is the presence of the relativistic denominator in the former. In qualitative terms, this denominator leads to the following effects. First, there are no stationary moving vortex pairs with L smaller than a critical value. This critical distance $L=2$ corresponds to the critical velocity $V/C_s^0=1/\sqrt{2}$, which is identical to the numerical result in Ref. 2, which is a pleasant surprise since it confirms that Eq. (3) yields qualitatively correct results even beyond its applicability range. Second, for $L>2$ there are two solutions, of which one asymptotically goes over to the corresponding solution for an incompressible liquid as $L\to\infty$, and the second (with unstable direction of vortex propagation) has a velocity tending to the speed of sound.

The absence of solutions with vortex pairs uniformly moving along the z-axis with the distance between them smaller than the critical value can be proved by linearizing the NSE [Eq. (1)] near $\Psi=0$ (near the vortex pair) and expanding smooth linear perturbations in a Fourier series. As a result, we have the relation

$$k_z^2V^2=(1-\mathbf{k}^2/2)^2,$$

which has real roots k_z for $V<1$, and $k_{xm}=m\pi/L$ only if $L>\pi/2$.

For $L\gg 1$ Eq. (3) can also be derived by linearizing the NSE [Eq. (1)]. In fact, at large distances from the vortices $\Psi\to 1$. By linearizing Eq. (1) about this value and neglecting derivatives of higher order,[2,16] we obtain a formula for the velocity potential ϕ:

$$(\nabla^2-V^2\,\partial^2/\partial z^2)\,\phi=0.$$

Therefore, after the relativistic transformation of coordinates, $\sigma=z/\sqrt{1-V^2}$, the potential $\phi(x,\sigma)$ becomes harmonic and, with due account of sources, specifically vortex pairs of unit intensity, it is expressed as

$$\phi=\arctan\left(\frac{\varsigma}{x+L/2}\right)-\arctan\left(\frac{\varsigma}{x-L/2}\right).$$

The gradient of this potential averaged at the site of one vortex yields an expression for the velocity V identical to Eq. (3).

Solutions of Eq. (1) with uniformly moving vortices are invariant with respect to the inversion of the propagation direction.[2] This means that they move in a specially configured inhomogeneous background field Ψ which provides a coherent return to the vortex pair. If the field Ψ in front of a vortex dipole is not specially configured (for example, the pair propagates in the initially constant background field), the vortices lose their energy to wave generation (they produce the specially configured background), and they no longer move uniformly. As the energy decreases, the vortices in a dipole approach one another. If the distance between them drops to the critical value (at which the vortex pair velocity reaches the local sound velocity), a local Cherenkov resonance becomes possible, and the state of uniform motion cannot be restored. The vortex pair energy is lost to radiation, and the distance between the vortices drops until they annihilate. As follows from computer simulations described below, this collapse takes a finite time comparable to the propagation time of waves emitted by the moving vortices through the screening length. Therefore, it does not matter as regards the possibility of collapse what fraction of emitted energy is transmitted to infinity and what fraction comes back, since the vortices will have annihilated by the time it returns.

In order to determine the scaling of the collapse of two oppositely charged dark solitons modeled by the NSE, we simulated the process on a computer using a second-order finite-difference ''predictor–corrector'' Euler scheme with respect to time and space. The studied area was covered with a 100×100 mesh, which corresponded to more than ten zones over the screening length in Eq. (1). The vortex strength in each mesh was derived from the circulation of the velocity $\nabla\phi$ on its boundary. The vortex position was defined as the center of the resulting vortex area. The transverse size of this area in the simulation was at most two or three meshes, which allowed us to determine the vortex position fairly accurately. Figure 1 shows contours of constant field density $|\Psi|^2$ at three sequential moments ($t_1=0.75$, $t_2=1.75$, and $t_3=2.5$). The initial distance $L_0=1$ between vortices of unit amplitude equalled one half of the critical value, which guaranteed the collapse of the vortices. One can see that the dipole becomes more asymmetric with respect to the reversal of the direction of motion owing to Cherenkov radiation of potential waves. It was found that the collapse of vortices with opposite charges moving along the z-axis takes a finite time, and in the final stage the distance L decreases with time as $\sqrt{t_0-t}$ near the moment of collapse t_0 (in Fig. 1 $t_0\approx t_3=2.5$). Figure 2 shows calculations of the function $L(t_0-t)$. Oscillations due to emission of waves with lengths comparable to the distance L between the vortices are clearly seen. The dashed curve plots the best fit of the function $L(t_0-t)$ to a power function with exponent $\alpha=0.5$. The uncertainty in this parameter due to the uncertainty in the time of collapse is $\delta\alpha=0.1$. Figure 3 plots the position of the vortex pair z_0-z measured with respect to the site of collapse z_0 as a function of the time to collapse (as can be seen in Fig. 1, $z_0\approx 1.7$). The dashed line in this graph corresponds to uniform motion with the critical velocity $V/C_s^0=1/\sqrt{2}$, which is the maximum possible value according to Eq. (3). One can see that the pair velocity is very close to the critical value right up until the moment of collapse.

3. DECAY OF A MULTICHARGED DARK NSE SOLITON

The boundary condition on the field Ψ in Eq. (1) with a repulsive potential implies that $|\Psi|^2$ tends to a constant background. The phase of the field can be arbitrary. This allows us to consider sets of vortices with unbalanced

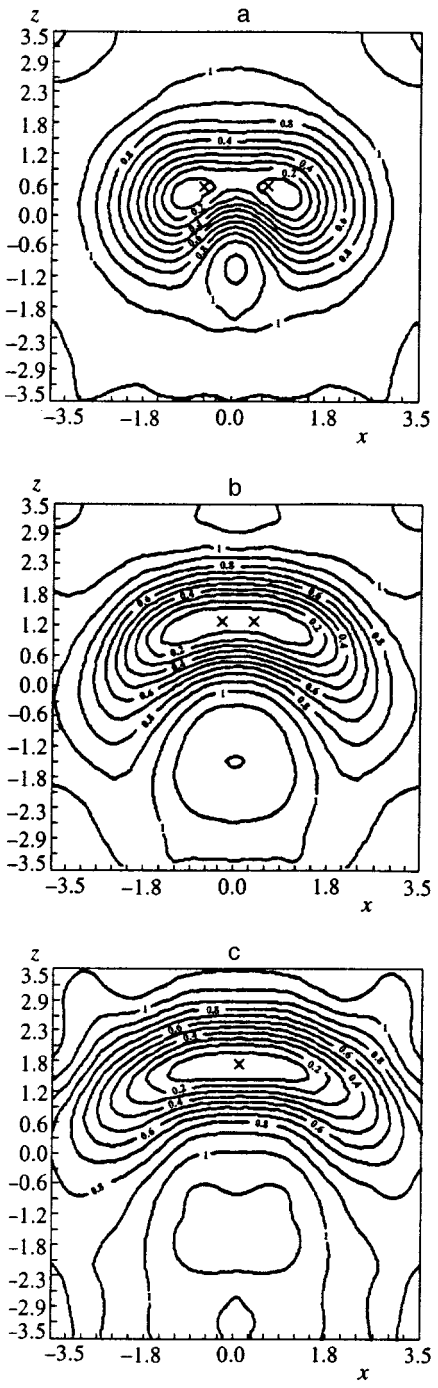


FIG. 1. Contours of constant $|\Psi|^2$ in the process of collapse of a pair of oppositely charged dark NSE solitons at successive times: (a) $t_1=0.75$; (b) $t_2=1.75$; (c) $t_3=2.5$. The crosses mark point-like vortex positions.

charge, i.e., charged dark 2D solitons. As in the case of any nonlinear differential equation, the important parameters in investigation of Eq. (1) include its integrals [Eq. (2)], namely the energy E, number of particles N, momentum $\mathbf{P}$, and angular momentum M:

$$E=\int \frac{1}{2}(\nabla\Psi)^2+\frac{1}{2}(|\Psi|^2-1)^2 d^2r,$$

$$N=\int (|\Psi|^2-1)d^2r, \tag{4}$$

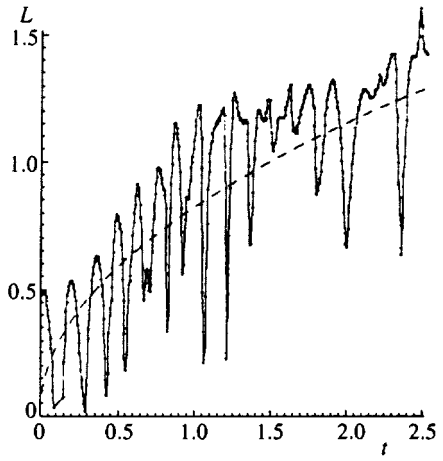


FIG. 2. Distance between dark NSE solitons with opposite charges as a function of the time to collapse at $t_0=2.5$. The dashed line shows the best fit based on a power law with exponent 0.5.

$$\mathbf{P}=\frac{i}{2}\int (\Psi^*\nabla\Psi-\Psi\nabla\Psi^*)d^2r,$$

$$M=\frac{i}{2}\int [\mathbf{r}(\Psi^*\nabla\Psi-\Psi\nabla\Psi^*)]_z d^2r,$$

which are named in analogy with hydrodynamic integrals. Given a nonzero background field Ψ on the boundary, the boundary condition for the phase should be determined consistently so as to provide conservation of the integrals in Eq. (4). The hydrodynamic impenetrability, or solid-wall, condition,

$$\mathbf{V}|_{\mathbf{n}}\equiv\nabla\phi|_{\mathbf{n}}\equiv\frac{i}{2|\Psi|^2}(\Psi^*\nabla\Psi-\Psi\nabla\Psi^*)|_{\mathbf{n}}=0$$

is sufficient for conservation of the integrals in Eq. (4).

Dark solitons with like charges gyrate around a fixed point, similarly to hydrodynamic point vortices. The energy of the vortex system drops as the distances between them increase, so dark solitons with like charges move apart if

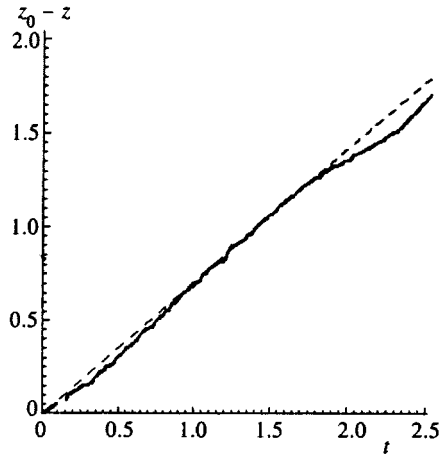


FIG. 3. Vortex pair position (with respect to the collapse site $z_0=1.7$) as a function of time to collapse. The dashed line corresponds to uniform motion with critical velocity $V=1/\sqrt{2}$.

emission of potential waves is taken into account. Even a single dark soliton with a charge $N>1$ is unstable and decays to vortices of unit amplitude located initially at the vertices of a regular N-gon, as has been verified repeatedly in numerical simulations.[11] The size of this N-gon increases with time owing to emission of energy (in the acoustic approximation as $\propto t^{1/2(N+1)}$, see Ref. 12). At the same time, by analogy with hydrodynamics, a polygon instability develops when the number of vortices satisfies $N>7$,[12] and the vortices are randomly distributed over a certain area. Far from this area, the velocity field due to these vortices has only an asymptotic rotational component, so the flow pattern as a whole is similar to the velocity field generated by a vortex of finite size.

In an unbounded area, the spreading of the vortices is unlimited. In a region bounded by solid walls, waves generated by the vortices are reflected from the boundary and return to the vortex zone. In the subsequent radiation, the wave momentum is imparted to the vortices, which can stabilize their spreading. Indirect evidence in favor of this stabilization can be found, for example in Lamb's monograph.[14] It was shown previously[13] using Arnold's method of iso-frozen variations[17] that, given appropriate boundary conditions (a solid wall of finite dimensions), a vortex shaped as a circular cylinder with a uniform distribution of vorticity is Lyapunov-stable in three dimensions, even when radiation is taken into account. The Lyapunov functional $\Phi=E-\Omega M$ is composed of such integrals as the energy E and angular momentum M. Specifically, the vortex energy drops as the distances between the vortices increases, whereas, the angular momentum increases. The decrease in the energy as a function of the angular momentum is fully balanced (the variation of functional Φ is zero) when the Lagrangian factor Ω equals the average angular velocity of the region occupied by the vortices. Thus, definiteness of the sign of the second variation of Φ (and consequently, the stability of the dimensions of the region with uniform distribution of charged dark NSE solitons) is ensured, as in Ref. 13, when the solid wall is taken to be of appropriate size. Specifically, the radius R of the solid wall should be such that the inequality

$$C_s^2>\left(\Omega r-\Omega\frac{a^2}{r}\right)^2, \quad a<r<R, \quad \Omega=\frac{N}{a^2} \tag{5}$$

holds everywhere outside the circular area of radius a occupied by N vortices. The physical meaning off this condition is that the flow should be ''subsonic'' in a reference frame rotating with angular velocity Ω. On the boundary of the vortex area ($r=a$), the right-hand side of inequality (5) is zero. The field density $\rho\equiv|\Psi|^2$ in the vortex area is almost zero, but the local sound velocity $C_s^2\equiv\rho+\Delta\sqrt{\rho}/4\sqrt{\rho}$ on this boundary is finite owing to the large perturbations in the field density. Therefore condition (5) is satisfied near the vortex area. Far from this area we have $C_s^0=1$, so condition (5) is fully satisfied if

$$R\leqslant\frac{a^2}{2N}+\sqrt{a^2+\left(\frac{a^2}{2N}\right)^2}.$$

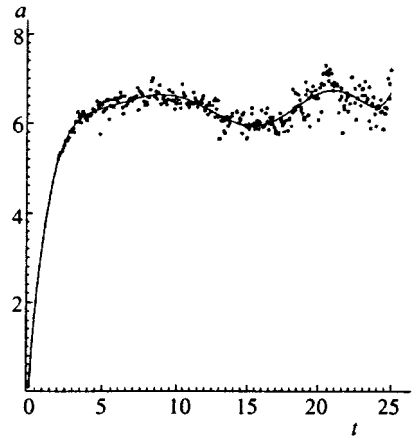


FIG. 4. Vortex area radius as a function of time.

At a low packing density ($N/RC_s^0\ll1$) the radius a of the stable area occupied by the vortices can be much smaller than the boundary radius R:

$$a_{\min}\simeq\frac{R}{\sqrt{\frac{R}{N}+1}}\sim\sqrt{NR}\ll R. \tag{6}$$

In other words, Eq. (6) indicates that, if the radius R of the shell containing a system of dark NSE solitons is much smaller than their total charge N, all dark solitons can be stably localized in the region with dimensions considerably smaller than those of the shell, i.e., $a_{\min}\sim\sqrt{NR}$.

This statement is illustrated in Fig. 4, which shows calculations of the decay of an isolated multicharged dark NSE soliton in the region bounded by a solid wall. The abscissa in Fig. 4 is time, and the ordinate is the rms radius a of the localization region containing unit vortices resulting from this decay. The input data in these calculations are the wall radius $R=25$ and the dark soliton charge $N=11$. It is clear that the size of the vortex zone does not grow beyond $a\simeq7$, which is even smaller than the estimate by Eq. (6). Figure 5 shows contours of constant $|\Psi|^2$ at $t=10$, when the growth of the vortex zone is saturated. The graph clearly shows an almost circular region, which persists at all subsequent times in this calculation and is occupied by vortices undergoing nonuniform motion.

4. CONCLUSIONS

In this paper, we have discussed the interaction of dark two-dimensional NSE solitons with either like or opposite charges. The hydrodynamic (Madelung) analogy between these dark solitons and point vortices of discrete amplitude in a compressible medium has proved to be very useful. A singly charged isolated dark soliton cannot be destroyed by any potential waves, which makes it a very attractive subject of research.

The electromagnetic analogy between dark solitons and electric charges emitting electromagnetic waves has allowed us to determine the translational velocity of a dipole pair of dark solitons [Eq. (3)]. There is a maximum critical velocity for uniformly moving dipoles, which is in both qualitative and quantitative agreement with the numerical calculations

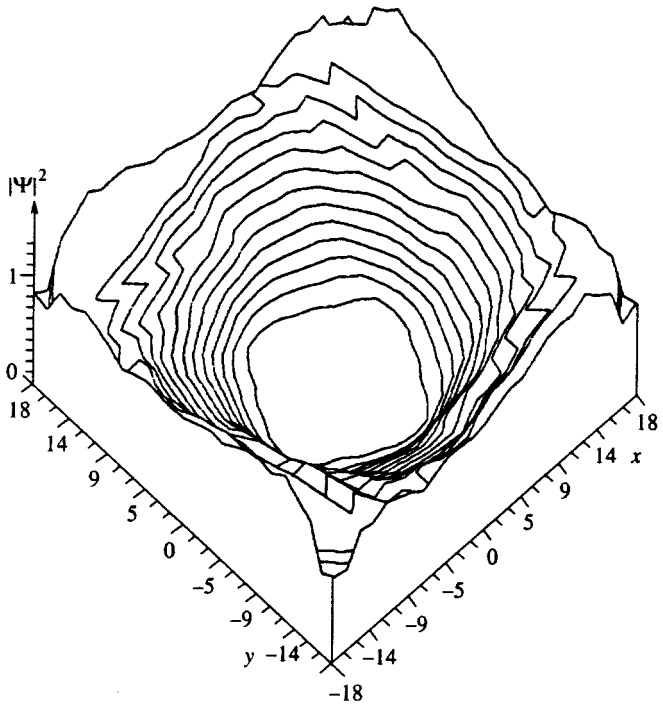


FIG. 5. Lines of equal field density $|\Psi|^2$ at moment $t=10$.

of Ref. 2. The impossibility of uniform propagation of dark charged solitons with characteristic dimensions smaller than the screening length determined by Eq. (1) has been proved. This also applies to the three-dimensional NSE (for example, to solutions with vortex rings). The computer simulation demonstrates what happens to a vortex pair of dark solitons with a distance between them smaller than the critical value: they begin to emit waves and annihilate after a finite time interval. During their collapse, the translational velocity of the vortex pair is finite and less than the critical velocity $V/C_s^0=1/\sqrt{2}$, i.e., the maximum value determined by Eq. (3). The distance between the vortices averaged over the oscillations is $L\propto(t_0-t)^{\alpha}$ where the exponent is $\alpha=0.5\pm0.1$, and in the final stage of collapse near the time t_0 it decreases at an increasing rate.

The problem of decay of a multicharged ($N\gg1$) dark NSE soliton in a region bounded by a solid wall (''impenetrable'' boundary condition) has been solved both analytically and numerically. The vortex decay does not yield a uniform spatial distribution of the resulting vortices; instead, they are uniformly spread over a region whose dimension is in some cases $\sim\sqrt{NR}$ [Eq. (6)], which is much smaller than the wall dimension R.

I would like to express gratitude for advice and helpful discussions to K. V. Chukbar and V. V. Yan'kov. The work was supported by the Russian Fund for Fundamental Research (Grant No. 96-02-17249a) and the Russian State Program on Fundamental Problems of Nonlinear Dynamics.

[1] *Solitons in Action*, ed. by K. Lonngren and A. Scott, Academic Press, New York (1981).
[2] C. F. Jones and P. H. Roberts, J. Phys. A: Math. Gen. **15**, 2599 (1982).
[3] C. Josserand and Y. Pomeau, Europhys. Lett. **30**, 43 (1995).
[4] M. H. Anderson, J. R. Ensher, M. R. Matthews, C. E. Wieman, and E. A. Cornell, Science **269**, 198 (1995).
[5] V. E. Zakharov and A. B. Shabat, Zh. Éksp. Teor. Fiz. **64**, 1627 (1973) [Sov. Phys. JETP **64**, 823 (1973)].
[6] I. V. Barashenkov and E. V. Panova, Physica D **69**, 114 (1993).
[7] A. I. D'yachenko, V. E. Zakharov, A. N. Pushkarev, V. F. Shvets, and V. V. Yan'kov, Zh. Éksp. Teor. Fiz. **96**, 2026 (1989) [Sov. Phys. JETP **69**, 1144 (1989)].
[8] I. A. Ivonin, Zh. Éksp. Teor. Fiz. **103**, 107 (1993) [JETP **76**, 57 (1993)].
[9] E. A. Kuznetsov and S. K. Turitsyn, Zh. Éksp. Teor. Fiz. **94**, 119 (1988) [Sov. Phys. JETP **67**, 1583 (1988)].
[10] H. Sakaguchi, Prog. Theor. Phys. **82**, 7 (1989).
[11] J. Neu, Physica D **43**, 385, 407 (1990).
[12] I. M. Gryanik, Izv. Akad. Nauk SSSR, FAO **19**, 203 (1983).
[13] I. A. Ivonin, Fiz. Plazmy **18**, 581 (1992) [Sov. J. Plasma Phys. **18**, 302 (1992)].
[14] H. Lamb, *Hydrodynamics*, 6th ed., Dover, New York (1945).
[15] L. D. Landau and E. M. Lifshits, *Field Theory*, Pergamon Press, Oxford (1988).
[16] E. A. Kuznetsov and J. J. Rasmussen, Phys. Rev. E **51**, 4479 (1995).
[17] V. I. Arnol'd, *Mathematical Methods of Classical Mechanics*, Springer, New York (1989).

Translation provided by the Russian Editorial office.

ERRATA

Erratum: Spin liquid in an almost ferromagnetic Kondo lattice [JETP 85, 399–414 (August 1997)]

K. A. Kikoin, M. N. Kiselev, and A. S. Mishchenko

Kurchatov Institute, 123182 Moscow, Russia

[S1063-7761(97)02512-2]

The correct title of this paper should read as follows: ''Spin liquid in an almost antiferromagnetic Kondo lattice.''

Erratum: Inelastic light scattering by electrons and plasmons in metals [JETP 85, 370–375 (August 1997)]

L. A. Falkovsky

L. D. Landau Institute for Theoretical Physics, Russian Academy of Sciences, 117334 Moscow, Russia

S. Klama

Institute of Molecular Physics, Polish Academy of Sciences, Smoluchowskiego 17, 60-179 Poznan, Poland

Zh. Éksp. Teor. Fiz. **112**, 2263 (December 1997)

[S1063-7761(97)02612-7]

At the end of section ''5. CONCLUSIONS'' the following sentence should be added: ''The other (S.K.) acknowledges the financial support of the Polish Research Committee of Sciences (Grant No. 2 P03B 101 12).''

 1063-7761/97/121239-01$10.00