Pure nuclear resonant scattering of synchrotron radiation by a multilayer structure: theoretical analysis for [169]Tm

V. G. Kon*)

Kurchatov Institute Russian Scientific Center, 123182 Moscow, Russia

A. I. Chumakov and R. Rüffer

European Synchrotron Radiation Facility, BP 220, F-38043 Grenoble Cedex, France
(Submitted 29 October 1997)
Zh. Éksp. Teor. Fiz. **114**, 3–22 (July 1998)

The possibility of observing pure nuclear resonant scattering of synchrotron radiation by a multilayer structure containing the [169]Tm isotope is analyzed theoretically. The main problem is the need to suppress the enormous background of radiation scattered by electrons. Two methods for the destructive interference of a synchrotron radiation beam in reflection at grazing incidence by a layered system containing Tm nuclei in one of the layers are considered, and their efficiency as applied to the conditions of third-generation synchrotron radiation sources, such as in the European Synchrotron Radiation Facility (ESRF), is calculated. An electron scattering suppression efficiency parameter is formulated as the ratio of the integrated nuclear scattering intensity (with a time delay) to the total prompt electron scattering intensity in assigned ranges of angles and energies. In the first method thin films of a special type on a substrate, viz., GIAR films, are used. In the second method a new effect, which is termed the Bragg antipeak effect and involves the destructive interference of a wave that is Bragg-diffracted in a multilayer superlattice and a wave reflected on the upper boundary of the sample, is employed. The physical properties of the Bragg antipeak effect are considered, and it is found that its efficiency is sufficient for practical use. © *1998 American Institute of Physics.* [S1063-7761(98)00107-3]

1. INTRODUCTION

Since the first proposals[1,2] to use synchrotron radiation to excite low-lying nuclear levels and the first clear-cut observation of nuclear resonant scattering of synchrotron radiation,[3] a great deal of attention has been focused on the problem of suppressing the nonresonant background of electron scattering. In principle, the pulsed temporal structure of synchrotron radiation permits the use of a "time window" to separate the delayed nuclear resonant scattering and the pulsed electron scattering. However, the extremely narrow energy width of nuclear levels (10^{-7}–10^{-10} eV) makes it possible to use a time window only after strong preliminary monochromatization of the synchrotron radiation with the aim of effectively suppressing the nonresonant electron scattering.

Several approaches to solving the problem of enormous nonresonant scattering have been successfully developed in recent years. The compact high-resolution monochromators with an energy width of the order of several millielectron volts[4–7] and avalanche photodiode detectors developed[8,9] have permitted the performance of investigations in a more universal forward-scattering experimental setup[10] in addition to the experimental setup in which pure nuclear Bragg diffraction is realized.[3] These advances provided the base for the rapid development of nuclear hyperfine spectroscopy using synchrotron radiation. In the case of Bragg diffraction, the alternative technique of the polarization analysis of scattered radiation has been used to suppress nonresonant radiation under the conditions of the hyperfine splitting of nuclear levels.[11,12]

The development of new sources of synchrotron radiation permitted a sharp increase in the spectral density of synchrotron radiation. The third-generation sources of synchrotron radiation, such as ESRF (France), APS (U.S.A.), and SPring-8 (Japan), provide beams with intensities of order 10^{13} photons/s/eV. All the advantages of nuclear hyperfine spectroscopy can be utilized only with ultrahigh monochromatization, which is still not achievable with electron-reflection monochromators. Therefore, nuclear resonant scattering must be used in the monochromatization of the radiation. Although the use of pure nuclear Bragg diffraction is very effective, it is applicable only to highly perfect, isotope-enriched single crystals. Therefore, a great deal of attention has been focused on the development of artificial, purely nuclear reflecting elements, mainly on the basis of multilayer antireflection films and multilayer superlattices. The operation of such elements is based on efficient nuclear resonant scattering under conditions where electron scattering is suppressed. In the case of GIAR (grazing-incidence antireflection) films electron scattering is suppressed through the destructive interference of waves reflected from the upper and lower faces of the film,[13–16] and in the case of multilayer superlattices it is achieved owing to the use of an electron-forbidden diffraction reflection under conditions where the atomic and nuclear scattering parameters differ.[17–19]

The study of nuclear resonant scattering using synchrotron radiation sources on both the theoretical and experimental levels has been confined hitherto to the ^{57}Fe isotope. The only exception is the work described in Refs. 16 and 20, where the ^{119}Sn isotope was considered. The present paper describes a theoretical analysis of several methods for suppressing electron scattering in the case of the nuclear transition with energy 8.410 keV in ^{169}Tm. The relatively large width of this nuclear transition (1.1×10^{-7} eV) creates the conditions for intense nuclear scattering of synchrotron radiation. On the other hand, the short natural lifetime (5.8 ns) makes high preliminary monochromatization especially necessary. Crystals containing Tm atoms rarely have a sufficiently perfect crystal lattice; therefore, the use of Bragg diffraction has been demonstrated in only one paper.[21] Suitable monochromatization methods can be based on the employment of conditions for near-grazing reflection by layered samples containing Tm atoms in one of the layers. The technology for fabricating multilayer structures has been developed to a fairly high level.[22] We examined the possibilities of using multilayer structures in which layers consisting of Tm atoms alternate with layers of Mo, Ti, or Fe atoms. Multilayer structures in which the layers just mentioned are combined can easily be fabricated.

Since the degree of disparity between the refractive indices in these combinations differs strongly from one to another, layered systems of different structure and different techniques for suppressing the nonresonant scattering must be used. In the Tm/Mo combination one layer of Tm on a Mo substrate or one layer of Tm with one layer of Mo on a glass substrate is sufficient. The thickness of the layers is found from the ''damping-stabilized solution'' in the theory of GIAR films.[13] A similar solution was found in the case of a Ti film on a Tm substrate. In the case of Tm layers and Fe layers we propose a new method of destructive interference, which combines two scattering channels: strong reflection by the surface of a sample with grazing incidence and Bragg diffraction in a multilayer superlattice. The interference of these two channels is totally different for nuclear resonant scattering and electron nonresonant scattering. We found conditions under which these two processes suppress one another in the case of electron scattering with the resultant suppression of reflection of the synchrotron radiation beam at the Bragg angle, i.e., the Bragg antipeak effect. On the other hand, in the case of nuclear scattering they can reinforce one another and produce an ordinary Bragg peak.

To characterize each method for suppressing electron reflection, we introduced a parameter which takes into account both the decrease in the electron reflection of the synchrotron radiation pulse and the level of nuclear resonant scattering (which is delayed during the scattered radiation) following a short pulse of synchrotron radiation. This parameter is defined with allowance for the possible angular width of the synchrotron radiation beam and the presently attainable level of premonochromatization.

The next section presents the basic principles of the theory of reflection by layered structures, which are used in a quantitative analysis. In particular, the physical nature of the suppression of reflection is considered, and simple condi-
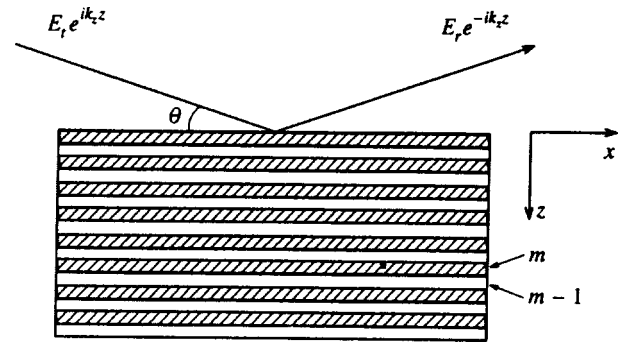

FIG. 1. Geometry of the grazing-incidence reflection of a synchrotron radiation beam by a multilayer structure.

tions for observing it are formulated. An analysis of the specific examples of (Tm, Mo) and (Ti, Tm) layered systems, as well as (Tm, Fe) multilayer superlattices, is presented in the third section. A discussion of the possibilities of the practical implementation of the conditions found is given in the last section.

2. THEORY OF THE SUPPRESSION OF ELECTRONIC SCATTERING BY MULTILAYER STRUCTURES

2.1. General formulas

The simplest method for calculating the reflectivity of a multilayer structure is based on the Parratt recurrence formula[23] for the reflection amplitude from a boundary between two neighboring layers. Let us consider the magnitude of the electric field vector in the mth layer, which we write in the form

$$E_m(x,z,t)=E(x,t)[E_{tm}\exp(ik_{zm}z)+E_{rm}$$
$$\times\exp(-ik_{zm}z)], \qquad (1)$$

where $E(x,t)=\exp(ik_x x-i\omega t)$, ω is the angular frequency of the x radiation, and $\mathbf{k}=(k_x,0,k_z)$ is the wave vector. Formula (1) is applicable to each layer taken individually, and the coordinate z is measured from the upper boundary of each layer within the layer. The surface of the layers coincides with the xy plane. To illustrate the idea behind the calculation, let us consider the simple case of a σ-polarized plane wave with an electric field directed along a normal to the scattering plane (i.e., along the y axis). The geometric parameters in the xz scattering plane are shown in Fig. 1. If θ is the angle between the direction of the incident plane wave and the surface of the multilayer, then

$$k_x=\lambdabar^{-1}\cos\theta, \quad k_{zm}=\lambdabar^{-1}\sqrt{\sin^2\theta+\chi_m}, \quad \lambdabar=\lambda/2\pi, \qquad (2)$$

where λ is the wavelength of the radiation and $\chi_m=\varepsilon_m-1$ is the complex susceptibility of the mth layer (ε_m is the dielectric constant). In the case of grazing incidence we can take $\sin\theta\approx\theta$.

We introduce the complex reflection amplitude $R_m=E_{rm}/E_{tm}$ as the ratio of the amplitude of the reflected wave E_{rm} to the amplitude of the refracted wave E_{tm}. The Parratt formula is

$$R_m = C_m \frac{F_m + R_{m-1}}{1 + F_m R_{m-1}}, \tag{3}$$

where

$$C_m = \exp(2ik_{zm}d_m), \quad F_m = \frac{k_{zm} - k_{zm-1}}{k_{zm} + k_{zm-1}}, \tag{4}$$

d_m is the thickness of the mth layer. The layers are numbered from the lower to the upper boundary of the multilayer (see Fig. 1). The complex susceptibility of the medium is accurately specified by the known expression

$$\chi = \chi' + i\chi'' = -\frac{\lambda^2 r_0}{\pi} N[Z + \Delta f' - i\Delta f'']$$

$$= -\lambda^2 C \frac{\rho}{A}[Z + \Delta f' - i\Delta f''], \tag{5}$$

where $r_0 = e^2/mc^2$, N is the number of atoms per unit volume of the material, $\Delta f' - i\Delta f''$ is the complex dispersion correction to the atomic x-ray scattering factor, which is simply equal to the number of electrons Z in the case of grazing incidence, $C = 0.5402 \times 10^{-5}$, ρ is the density in g/cm^3, A is the atomic weight in atomic units, and λ is the wavelength in angstroms. We recall that the susceptibility is twice the complex refractive index.

Formula (3) was written for a perfect multilayer structure with a smooth surface on each boundary between the layers. A real surface is generally rough. In this study we assume that the roughness takes the form of uncorrelated upward or downward displacements of finite smooth portions of the surface (along the z axis), while the mean z coordinate of the boundary in space remains unchanged. Such a type of roughness can be taken into account by introducing a variable layer thickness $d'_m = d_m + \Delta$, where Δ is a random value. Assuming that the distribution of the values of Δ with the width D is Gaussian, we should average the phase factor C_m. As a result of this averaging, C_m is replaced by $C_m W_m$, where

$$W_m = \frac{2}{D\sqrt{\pi}} \int_{-\infty}^{\infty} d\Delta \, \exp\left(2ik_{zm}\Delta - \frac{4\Delta^2}{D^2}\right)$$

$$= \exp\left(-\frac{1}{4}k_{zm}^2 D^2\right) = \exp(-2k_{zm}^2 \sigma^2). \tag{6}$$

Here

$$\sigma^2 = \langle \Delta^2 \rangle = \frac{2}{D\sqrt{\pi}} \int_{-\infty}^{\infty} d\Delta \Delta^2 \exp\left(-\frac{4\Delta^2}{D^2}\right) = \frac{D^2}{8}. \tag{7}$$

The factor W_m can be called the Debye–Waller factor,[24] because its dependence on the change in the value of the layer thickness is similar to the dependence of the Debye–Waller factor on the displacement of atoms from their equilibrium positions in a crystal lattice. This approach differs from the one in Ref. 25, where a different model of roughness was considered. Below we shall take into account the Debye–Waller factor directly in C_m and redefine C_m in the following manner:

$$C_m = \exp(2ik_{zm}d_m - 2(k_{zm}\sigma)^2). \tag{8}$$

2.2. Damping-stabilized solutions

We begin by considering the simple case of a multilayer with one layer of Tm of thickness d containing resonant nuclei on a substrate of some other material. We take Mo as an example; therefore, the system under consideration can be described by the formula Tm(d)Mo. In this case we have only two boundaries, and we must use the Parratt formula only twice. We introduce the amplitudes of the field E: E_{ta} and E_{ra} in the air directly above the surface and E_{tl} and E_{rl} in the Tm layer above the interface with the substrate. We are interested in the total reflection amplitude $R_{\mathrm{tot}} = E_{ra}/E_{ta}$, while the reflection amplitude from the internal boundary $R_1 = E_{rl}/E_{tl}$. In this case the Parratt formula leads to the relation

$$R_{\mathrm{tot}}(\theta) = C_0(\theta) \frac{R_0(\theta) + R_1(\theta)C_1(\theta, d_1)}{1 + R_0(\theta)R_1(\theta)C_1(\theta, d_1)}, \tag{9}$$

where $R_0(\theta)$ is the reflection amplitude from the sample surface, and $R_0(\theta)$ and $R_1(\theta)$ equal

$$R_0(\theta) = \frac{\theta - \theta_1}{\theta + \theta_1}, \quad R_1(\theta) = \frac{\theta_1 - \theta_2}{\theta_1 + \theta_2},$$

$$\theta_1 = \sqrt{\theta^2 + \chi_1}, \quad \theta_2 = \sqrt{\theta^2 + \chi_2}. \tag{10}$$

Here χ_1 is the susceptibility in the layer, and χ_2 is the susceptibility of the substrate. The values of C_0 and C_1 were determined in (8) for $d_0 = 0$.

As we know, the critical angle for total external reflection is specified by the formula $\theta_{cm} = (2\delta_m)^{1/2}$, where δ_m is the damping of the refractive index in layer m, which is related to the susceptibility by the expression $\chi/2 = -\delta + i\beta$. The substrate has a larger electron density than the layer; therefore, it has a larger value for the critical angle for total external reflection $\theta_{c2} > \theta_{c1}$. We are interested in the range of angles near the second critical angle θ_{c2}, where the magnitude of the reflection amplitude from the internal boundary, $R_1(\theta)$, is greater than the reflection amplitude from the surface, $R_0(\theta)$. Our goal is to find the conditions under which the total reflection amplitude from the sample equals zero. The solution of this problem was found for the first time in Ref. 13 and was termed the "damping-stabilized solution."

Such solutions correspond to a zero value for the numerator in (9). We write the complex reflection amplitudes, explicitly distinguishing between the modulus and phase:

$$R_0 = a_0 \exp(i\varphi_0), \quad R_1 = a_1 \exp(i\varphi_1). \tag{11}$$

This enables us to formulate the necessary conditions in the form of a system of two real equations, which specify the conditions for impedance matching and for destructive interference:

$$a_0 = a_1 \exp(-2(\sigma/\check{\lambda})^2(\theta^2 + \chi_1'))\exp(-2d\theta_1''/\check{\lambda}),$$

$$\varphi_0 = \varphi_1 - 2(\sigma/\check{\lambda})^2 \chi_1'' + 2d\theta_1'/\check{\lambda} - m\pi, \quad m = \pm 1, \pm 3, \dots . \tag{12}$$

Each of these equations can be treated as an independent equation for the layer thickness d. In addition, the solutions can be represented by analytical expressions, which give the angular dependence of thickness. We shall call the respective thicknesses the impedance-matching thickness d_{im} and the destructive-interference thickness d_{di}:

$$d_{im}(\theta) = -\frac{\ln(a_0/a_1) + 2(\sigma/\varkappa)^2(\theta^2 + \chi_1')}{2\,\theta_1''/\varkappa},$$

$$d_{di}(\theta) = \frac{\varphi_0 - \varphi_1 + m\pi + 2(\sigma/\varkappa)^2\chi_1''}{2\,\theta_1'/\varkappa}. \tag{13}$$

The complete solution of the problem is obtained at the points where these curves cross the plot of the angular dependence. The coordinates of the crossing points specify both the required angular orientation of the sample and the layer thickness. Since only positive values of the layer thickness have physical meaning, the condition $a_0 < a_1$ must be satisfied for success.

It is also interesting to consider more complicated systems with a large number of free parameters and to find the conditions for suppressing electron reflection in systems with some additional properties, for example, systems which are stable toward variation of the layer thickness, which can be preferable from the standpoint of their fabrication technology. Let us consider a system in which a Mo layer of thickness d_2 is located under a Tm layer of thickness d_1 on a glassy SiO_2 substrate, i.e., the $Tm(d_1)Mo(d_2)SiO_2$ system. Using the Parratt formula three times, we can obtain the total reflection amplitude in this case in the following form:

$$R_{tot}(\theta) = C_0(\theta)\frac{R_0(\theta) + R_1^+(\theta)C_1(\theta, d_1)}{1 + R_0(\theta)R_1^+(\theta)C_1(\theta, d_1)},$$

$$R_1^+(\theta) = \frac{R_1(\theta) + R_2(\theta)C_2(\theta, d_2)}{1 + R_1(\theta)R_2(\theta)C_2(\theta, d_2)}. \tag{14}$$

Here the functions $R_0(\theta)$ and $R_1(\theta)$ are defined by formula (10) except that the subscript 1 now refers to the Mo layer and, in addition,

$$R_2(\theta) = \frac{\theta_2 - \theta_3}{\theta_2 + \theta_3}, \quad \theta_3 = \sqrt{\theta^2 + \chi_3}, \tag{15}$$

where the subscript 3 refers to the substrate and $C_2(\theta, d_2)$ is the transfer function of the wave for the Mo layer. The conditions for the suppression of electron reflection can be written in the form (12). However, in this case the real functions a and φ have somewhat different definitions:

$$R_0(1 + R_1R_2C_2) = a_0\exp(i\varphi_0),$$

$$R_1 + R_2C_2 = a_1\exp(i\varphi_1). \tag{16}$$

Thus, all the parameters a_0, a_1, φ_0, and φ_1 now depend on the thickness d_2 of the Mo layer. Therefore, solutions differing in the value of d_2 can be found.

2.3. Bragg antipeak effect

In the case of a (Tm, Fe) multilayer the susceptibility difference is small. This leads to a small difference in the

critical angles for total external reflection. As a result, the damping-stabilized solutions considered above are impossible. On the other hand, the small susceptibility difference opens up a new way to achieve destructive interference between a wave reflected by the surface of a sample and a wave that is Bragg-diffracted in a multilayer superlattice having a periodic element in the form of a pair of Tm and Fe layers of different thickness, which is repeated n times. We again introduce the amplitudes of the field E: E_{ta} and E_{ra} in the air above the surface and E_{tt} and E_{rt} in the top layer directly beneath the surface. We are interested in the total reflection amplitude $R_{tot} = E_{ra}/E_{ta}$, while the reflection amplitude from the multilayer superlattice can be defined as $R_{ml} = E_{rt}/E_{tt}$.

In this case the Parratt formula gives the relation

$$R_{tot}(\theta) = C_0(\theta)\frac{R_0(\theta) + R_{ml}(\theta)}{1 + R_0(\theta)R_{ml}(\theta)}, \tag{17}$$

where $R_0(\theta)$ is the reflection amplitude from the surface, which is defined in (10), and χ_1 is the susceptibility of the top (Tm) layer of the multilayer. Near the critical angle for total external reflection, this reflection channel is also significant and leads to appreciable beam reflection for a material with a large electron density.

The only possibility for making the amplitude $R_{tot}(\theta)$ equal to zero for a multilayer superlattice arises in the case of the destructive interference of two channels, where $R_{ml} = -R_0$. This condition can be satisfied only if $|R_0| < 1$. Therefore, the multilayer superlattice sought need not be strongly reflective, and we are interested in the case of a small reflection amplitude from one period of the multilayer superlattice with a negative sign, which is opposite to R_0. Just such a condition is realized in the case of the (Tm, Fe) multilayer superlattice under consideration. An exact analytical formula describing the reflection amplitude from a multilayer superlattice was obtained in Ref. 26. In the general case of a multilayer superlattice having n periods (unit cells), it can be written in the following manner:

$$R_{ml}^{(n)}(\theta) = \frac{r}{f_+}\frac{1 - A\exp(in\psi)}{1 - B\exp(in\psi)}, \quad A = \frac{r - R_s f_+}{r - R_s f_-}, \quad B = A\frac{f_-}{f_+},$$

$$f_\pm = 1 - \sqrt{t\bar{t}}\exp(\pm i\psi/2), \quad \exp(\pm i\psi/2) = v \mp \sqrt{v^2 - 1},$$

$$v = \frac{1 + t\bar{t} - r\bar{r}}{2\sqrt{t\bar{t}}}. \tag{18}$$

Here R_s is the reflection amplitude directly above the first periodic element of the multilayer. Below the multilayer superlattice under consideration there can be another superlattice or system of layers with different parameters. The parameters r, $\bar{r}$, t, and $\bar{t}$ are the reflection and transmission amplitudes for one period in the forward and reverse directions. The explicit expressions for these quantities depend on the structure of a period in the multilayer superlattice. We are interested in the simple case with two layers in a period. We shall use the labels 1 and 2 to distinguish between the parameters of the upper and lower layers and thus have notation similar to the case considered above. The corresponding formulas can be written in the form

$$\bar{r}=r_{\mathrm{ml}}\left[1-\frac{t_{\mathrm{ml}}C_2}{1-r_{\mathrm{ml}}^2C_2}\right], \quad r=\bar{r}C_1, \quad t=\bar{t}=\frac{t_{\mathrm{ml}}(C_1C_2)^{1/2}}{1-r_{\mathrm{ml}}^2C_2},$$

$$r_{\mathrm{ml}}=\frac{\theta_1-\theta_2}{\theta_1+\theta_2}, \quad t_{\mathrm{ml}}=\frac{4\theta_1\theta_2}{(\theta_1+\theta_2)^2}=1-r_{\mathrm{ml}}^2, \tag{19}$$

where θ_j and C_j are defined in (8) and (10). We note that the results of a calculation using the analytical formulas are in complete agreement with the approach based on the repeated used of the Parratt recurrence formula for each boundary, but the computer time is significantly less, especially for a large number of periods. In addition, in the case of a thick multilayer superlattice, we obtain an analytical solution in the simpler form $R_{\mathrm{ml}}=r/f_+$ directly from (18), provided the function ψ is defined with a positive imaginary part.

The condition for Bragg diffraction in this approach requires that the complex quantity $t\bar{t}$ have a phase $2\pi n$. At grazing incidence the z component of the wave vector is small; therefore, the roughness of the boundary between the layers does not play a significant role. When the reflection amplitude from one boundary r_{ml} is small due to the small difference between the susceptibilities of the layers, the transmission amplitude t_{ml} is close to unity [see (19)]. Therefore, in an approximation we obtain the Bragg condition for first-order diffraction in the form $\mathrm{Re}(d_1\theta_1+d_2\theta_2)=\lambda/2$, where d_1 and d_2 are the thicknesses of the upper and lower layers of a single period in the multilayer superlattice. As we know, in the case of the electron scattering of hard x rays the imaginary part of the susceptibility is much smaller than the real part. Therefore, for angles of incidence greater than the critical angle, the reflection amplitude from the surface $R_0 = -\chi_1/4\theta^2$ has an approximately real value. The reflection amplitude from the boundary between layers in a multilayer superlattice has the approximate value $r_{\mathrm{ml}}\approx(\chi_1-\chi_2)/4\theta^2$. The small difference between the real parts of the complex susceptibilities can be accompanied by a normal difference between the imaginary parts [this is precisely what occurs in the case of a (Tm, Fe) multilayer superlattice]. As a result r_{ml} is a complex quantity with an appreciable phase. Taking this into account, we arrive at the conclusion that the optimum structure for the repeating element in a multilayer superlattice is a combination of layers of different thickness.

To observe the Bragg antipeak effect, i.e., suppression of the total electron reflectivity as a consequence of destructive interference between a surface reflection and a reflection from the multilayer superlattice near the Bragg angle, we must first ensure the condition $|R_{\mathrm{ml}}|\approx|r_{\mathrm{ml}}|F>|R_0|$, where F is a factor which characterizes the degree of increase in the reflection amplitude from the multilayer superlattice compared with reflection on one boundary. For a small number of periods N and equal thicknesses $d_1=d_2$ of layers with smooth surfaces we obtain simply $F=2N$. In the opposite case of a large number of layers, the value of F is restricted by the absorption and roughness of the surface and is approximately

$$F\approx\alpha/(1-|t\bar{t}|)\approx\alpha/M,$$

$$M=(\mu_1d_1+\mu_2d_2)/\theta+8(\theta\sigma/\lambda)^2,$$

μ_i is the absorption coefficient in layer i, and α is the absolute value of the structure factor of the multilayer superlattice, which can have a value between 0 and 2. The estimate obtained shows that appreciable reflection by a multilayer superlattice is confined to a finite range of angles, where absorption, on the one hand, and roughness, on the other hand, are not significant. In the case of a high level of roughness, this region can have zero dimensions, i.e., destructive interference can occur.

The second condition is that the phase difference between the complex quantities R_{ml} and R_0 must be equal to π. As was shown above, the phase of R_0 is approximately equal to zero. The phase of r_{ml} is approximately equal to π if absorption is neglected and in the case of equal layer thicknesses provided the upper layer has a smaller electron density. However, the absorption is significant, and therefore the phase of r_{ml} differs from π in the general case. There are two possibilities to compensate for this difference. The first arises at an angle shifted slightly from the Bragg angle. Since the reflection phase varies very rapidly near the Bragg angle, the necessary shift is small. For example, in the kinetic approximation and with the neglect of absorption, the reflection amplitude can be approximated by the expression

$$R_{\mathrm{ml}}=r\frac{\sin(N\varphi/2)}{\sin(\varphi/2)}\exp(i[N-1]\varphi/2),$$

$$\varphi=2\,\mathrm{Re}(\theta_1d_1+\theta_2d_2)/\lambda. \tag{20}$$

It is readily noted that near the Bragg maximum the phase of the reflection amplitude from a multilayer superlattice varies approximately N times faster than does the phase for a single period.

The second possibility is to use a multilayer superlattice with unequal layer thicknesses. In this case the phases of the multipliers C_1 and C_2 are not equal to π even at the Bragg angle, and the additional phase can cancel the deviation of the phase of r_{ml} from zero value. This case is preferable from the standpoint of the conception of the Bragg antipeak effect because destructive interference takes place here precisely at the Bragg angle. This leads to a large value for the nuclear resonant reflectivity. On the other hand, in this case the width of the dip on the electron reflectivity curve is narrower than in the former case owing to the property of Bragg diffraction just mentioned, i.e., the rapid variation of the phase of the reflection amplitude.

2.4. Nuclear resonant scattering

The susceptibility of a material is described by formula (5) only in the case of an electron channel for scattering a synchrotron radiation beam. The absence of electron reflection creates conditions for pure nuclear scattering by Tm nuclei in a layered system. When nuclear resonant scattering occurs, an additional contribution, which depends on the radiation frequency and describes the resonant interaction of Tm nuclei with synchrotron radiation, must be included in the susceptibility. In the simple case of a single unsplit line, the additional term is described by the formula

$$\Delta\chi_{\text{Tm}} = -\frac{\lambdabar N \sigma_0 \eta f_{\text{LM}}}{2w+i}, \quad w = \frac{\hbar\omega}{\Gamma_0},$$

$$\sigma_0 = 2\pi\lambdabar^2 \frac{2I_e+1}{2I_g+1} \frac{1}{\alpha+1}. \tag{21}$$

Here N is the number of Tm atoms per unit volume, σ_0 is the nuclear scattering cross section at resonance, η is the fraction of the resonant isotope, f_{LM} is the Lamb–Mössbauer factor, I_e and I_g are the values of the nuclear spin in the excited and ground states, respectively, α is the gamma-electron conversion coefficient, the frequency ω is measured relative to the resonance value, and Γ_0 is the natural resonance bandwidth.

The new term in the susceptibility of the Tm layer, which depends on frequency, causes the total reflection amplitude to also depend on frequency. In the case of strictly monochromatic radiation with a varying frequency, it becomes possible to measure the dependence of the reflection amplitude on frequency near resonance. On the other hand, when a short pulse of synchrotron radiation is scattered, a time-delayed component appears in the scattered radiation because of the resonant frequency dependence of the reflection amplitude. To calculate the delayed part of the scattered radiation, it is convenient to introduce a response function of the system to the instantaneous excitation of nuclei by a very short pulse. As we know,[27] this function is defined by the Fourier transform of the reflection amplitude:

$$G(t,\theta) t_0 = \int \frac{dw}{2\pi} R(w,\theta) \exp\left(-\frac{iwt}{t_0}\right), \quad t_0 = \frac{\hbar}{\Gamma_0}. \tag{22}$$

Here t_0 is the lifetime of the excited state of the nucleus. The simplest way to isolate the resonant scattering is to measure only the delayed radiation after a time window of width t_w with the detector covered within the time window $0 < t < t_w$. In this case the nuclear and electron resonant reflectivities can be defined by

$$P_n(\theta, t_w) \frac{\Gamma_0}{\hbar} = \int_{t_w}^{\infty} dt |G(t,\theta)|^2, \quad P_e(\theta) = |R(\infty,\theta)|^2. \tag{23}$$

Here and below we have omitted the subscript ''tot'' on the total reflection amplitude.

Below we shall confine ourselves to consideration of the special case of a very short time window, $t_w \to 0$. The integrated reflectivity of the delayed synchrotron radiation can then be calculated without Fourier transformation by utilizing Parseval's theorem:

$$P_n(\theta)\frac{\Gamma_0}{\hbar} = \int_{+0}^{\infty} dt |G(t,\theta)|^2$$

$$= \frac{\Gamma_0}{\hbar}\int \frac{dw}{2\pi}|R(w,\theta) - R(\infty,\theta)|^2. \tag{24}$$

On the other hand, the response function for short delay times can be determined from the asymptotic expansion of the frequency dependence of the reflection amplitude, i.e., if

$$\lim_{|w|\to\infty} R(w,\theta) = R(\infty,\theta) + R_1(\theta)\frac{1}{w} + R_2(\theta)\frac{1}{w^2} + \ldots, \tag{25}$$

then

$$G(t,\theta)\underset{t\to 0}{=} R(\infty,\theta)\delta(t) - iR_1(\theta) - R_2(\theta)\frac{t}{t_0} + \ldots$$

$$\approx R(\infty,\theta)\delta(t) - iR_1(\theta)$$

$$\times\exp\left(-\frac{t}{t_0}\left(\frac{\gamma}{2} + iw_0\right)\right) + \ldots, \tag{26}$$

where we have used the notation $R_2/R_1 = w_0 - i\gamma/2$. The quantity $R(\infty)$ is simply the electron part of the reflection amplitude, and R_1 specifies the nuclear part of the scattered radiation immediately after the pulse, while R_2 assigns the decay rate of the excited state for short delay times. The parameter γ characterizes the decay acceleration effect[27] (it is equal to unity if acceleration is not observed). Formula (26) permits direct evaluation of the integrated intensity during a short time window immediately after a pulse of synchrotron radiation. Thereafter a short time window can be taken into account in the integrated nuclear reflectivity (24) by simple subtraction.

The quantity of charged (nuclear) radiation depends on how strongly the near-resonant reflection amplitude differs from the off-resonance value. In a simple system with one surface the frequency dependence of the susceptibility causes the critical angle to also become dependent on frequency. However, at small grazing angles ($\theta \ll \theta_c$) and at large grazing angles ($\theta \gg \theta_c$) the reflection amplitude depends weakly on the critical angle and thus depends weakly on frequency. Therefore, it should be expected that the intensity of the delayed radiation integrated over time will have a maximum precisely in the range of angles near the critical angle $|\theta - \theta_c| \ll \theta_c$. As was shown in Ref. 28, in first-order perturbation theory with respect to the nuclear addition to the susceptibility (the kinematic approximation for nuclear scattering) the time dependence of the delayed intensity for a nuclear system mimics the time dependence for an isolated nucleus. The angular dependence of the intensity is identical for all times following a pulse (or for all values of the frequency mismatch from resonance). The corresponding coefficient can be related to the effective number of nuclei participating in coherent reflection of the delayed radiation. Therefore, the suppression of electron scattering precisely in the range of angles near the critical angle, where nuclear reflection is strongest, is of greatest interest from the standpoint of obtaining pure nuclear scattering.

2.5. Efficiency of pure nuclear scattering

We assume that radiation which has passed through the premonochromator system has a nonzero intensity within the angular range $\Delta\theta$ and the frequency range $\Delta\omega$. We are interested in the integrated intensity of the scattered radiation in these ranges,

TABLE I. Some parameters of the elements considered.

	Z	A, au	ρ, g/cm^3	$\Delta f'$	$\Delta f''$	$-\chi'$, 10^{-6}	χ'', 10^{-6}
Ti	22	47.88	4.53	0.269	1.68	24.728	1.866
Fe	26	55.85	7.86	-0.786	2.94	41.648	4.856
Mo	42	95.94	10.2	-0.087	2.54	52.300	3.169
Tm	69	168.93	9.29	-9.94	3.90	38.120	2.517

$$\bar{P}(\theta_0) = \int_{\Delta\theta} d\theta' \int_{\Delta\omega} d\omega \, P(\omega, \theta_0 + \theta'), \qquad (27)$$

where θ' is the angular shift away from the value θ_0, which corresponds to the complete suppression of electron reflection, and the frequency ω is measured relative the resonance value for ^{169}Tm. In the case of the nonresonant (prompt) part of the radiation, there is no frequency dependence; therefore,

$$\bar{P}_e = \Delta\omega \int_{\Delta\theta} d\theta' P_e(\theta_0 + \theta'). \qquad (28)$$

In the case of the nuclear resonant (delayed) part of the radiation, we can neglect the angular dependence in an approximation and write

$$\bar{P}_n = \Delta\theta P_n(\theta_0)\Gamma_0/\hbar, \qquad (29)$$

where $P_n(\theta)$ is defined by formula (24). The ability of different reflecting systems to produce pure nuclear scattering can be estimated by the ratio between these two quantities. In the small range of angles near the electron reflectivity minimum the angular dependence of the reflection amplitude has an approximately parabolic shape, and the integral is calculated analytically. As a result, we can introduce the efficiency parameter of pure nuclear scattering in the form

$$E_s = \frac{\bar{P}_n}{\bar{P}_e} \approx \frac{3\Gamma_0}{\hbar\Delta\omega} \frac{P_n(\theta_0)}{P_e(\theta_0 + \Delta\theta/2) + 2P_e(\theta_0)}. \qquad (30)$$

In the ensuing analysis of specific cases we shall use the following particular values of the parameters appearing in this formula for the conditions in ESRF: $\eta\Delta\omega \approx 3 \times 10^{-2}$ eV, $\Gamma_0 = 10^{-7}$ eV, $\Delta\theta = 12$ μrad. The efficiency parameter depends on the thickness of the layers of the system; therefore, each individual method for isolating pure nuclear scattering can also be characterized in terms of the sensitivity toward variation of the layer thickness.

For numerical estimates of the efficiency of various layered systems we used susceptibility values calculated from Eqs. (5) and (21), which are presented in Table I. The contribution of the nuclear resonant scattering of Tm to the susceptibility was calculated with the parameters

$$\eta = f_{\rm LM} = 1, \quad \lambdabar N\sigma_0 = 2 \times 10^{-5}. \qquad (31)$$

3. SPECIFIC EXAMPLES

3.1. (Tm, Mo) layered system

In the case of the Tm(d)Mo system the susceptibility difference is fairly large; therefore, damping-stabilized solutions are possible. Figure 2 shows the angular dependence of

$d_{\rm im}$ and $d_{\rm di}$ for three values of the roughness parameter: $\sigma = 0$, 5, and 10 Å. It is readily noted in the range of grazing angles of incidence that roughness up to 5 Å has a weak influence on the curves and that the $d_{\rm di}$ curves corresponding to a phase difference of π scarcely vary for the roughness values considered. We used the value $\sigma = 5$ Å in the ensuing calculations. On the other hand, regardless of the roughness level, there are two solutions. There are no solutions with other values of the phase difference (3π or more).

Figure 3 shows how the efficiency depends on the layer thickness for the damping-stabilized solutions found, as well as the angular dependence of the electron and nuclear reflectivities for the layer thickness corresponding to the maximum efficiency value. As follows from the calculations, the solution at small thicknesses has a doubled peak efficiency, but is more sensitive to variation of the thickness (Figs. 3a and 3c). The angular dependence for the thicknesses $d = 97$ Å and 158 Å is shown in Fig. 3b and 3d. The solution with the larger thickness has a higher level of nuclear reflectivity and therefore corresponds to an angle of incidence in the range between the critical angles for Tm ($\theta_c = 6.17$ mrad) and Mo ($\theta_c = 7.23$ mrad). At the same time, the dip on the electron reflectivity curve is narrower, and when the angular divergence of the beam (even 12 mrad) is taken into account, this also leads to a higher level of electron reflectivity. The practical choice of the appropriate variant is determined by the possibility of creating a layer with a
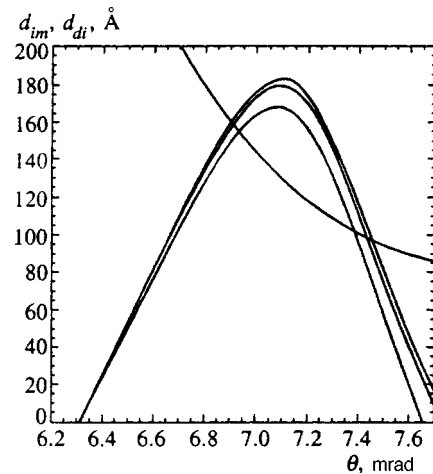

FIG. 2. Graphical solutions of the equations of impedance matching and destructive interference for determining the layer thickness and the beam angle of incidence in the case of a Tm(d)Mo layered structure for three values of the roughness parameter: $\sigma = 0$, 5, and 10 Å. The $d_{\rm im}(\theta)$ curves have maxima which decrease with increasing σ and the $d_{\rm di}(\theta)$ curves practically coincide.
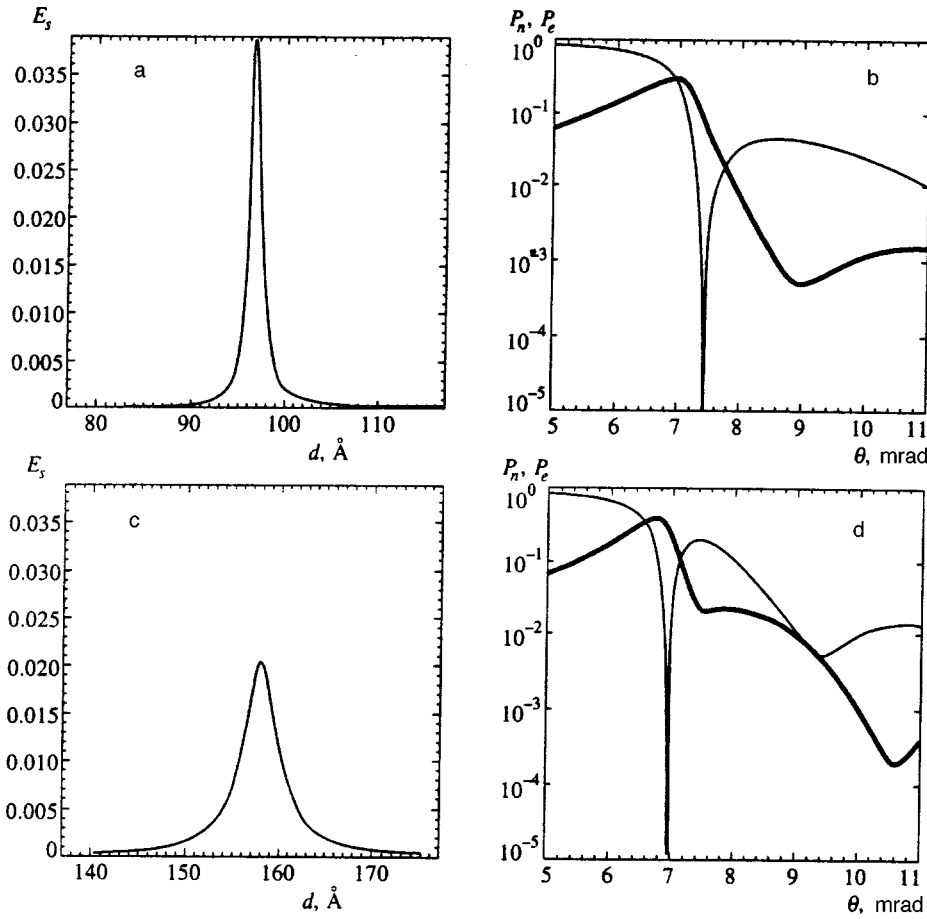
FIG. 3. Dependence of the efficiency on thickness (a and c) and angular dependence (b, d, decimal logarithms) of the electron and nuclear (thick lines) reflectivities in the two possible solutions for a $Tm(d)Mo$ layered structure with $d = 97$ Å and 158 Å.

sufficient degree of control over the thickness or the power of the source.

A more complicated system, which contains two layers and can be described by the formula $Tm(d_1)Mo(d_2)SiO_2$, has more free parameters. If the phase difference equals π, two solutions for the thickness d_1 of the Tm layer exist for each value of the thickness d_2 of the Mo layer, which we denote by $d_1^{(1)}$ and $d_1^{(2)}$ in order of increasing values. Calculations show that the solution $d_1^{(1)}$ for fairly large values of d_2 scarcely depends on the thickness of the Mo layer, while the value of $d_1^{(2)}$ decreases with a simultaneous increase in the efficiency as d_2 is diminished. Accordingly, the angle of incidence of the synchrotron radiation beam is also increased. When the thickness of the Mo layer is diminished further, the values of $d_1^{(1)}$ and $d_1^{(2)}$ approach one another. There is a critical value of d_2, at which the two solutions coincide, and then they vanish entirely as d_2 is subsequently diminished. Such behavior is specified by the strong influence of the additional reflectivity on the Mo/SiO_2 boundary. Figure 4 shows plots of the dependence of the efficiency on d_1 for various values of d_2 near the critical point. Attention should be focused on the very high efficiency level: up to 0.4 for the left-hand peaks and up to 0.2 for the right-hand peaks. However, destructive interference is very sensitivity to variation of the thickness. Of course, there is similar sensitivity toward variation of the susceptibility or roughness parameters.

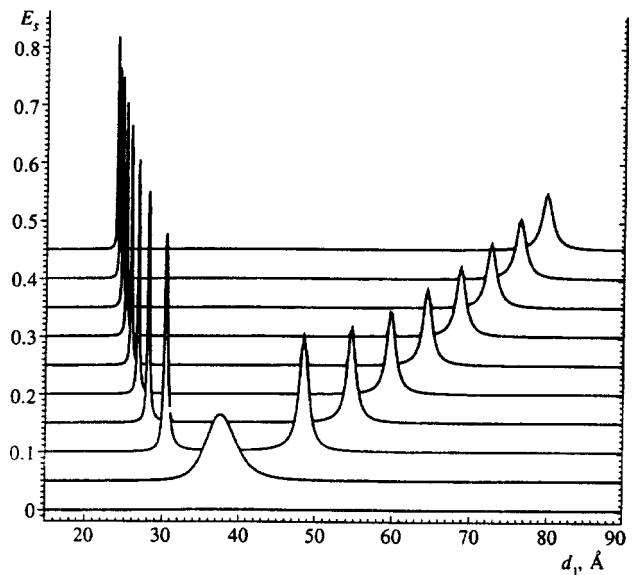Since the thicknesses of the layers are small, the angle of

FIG. 4. Dependence of the efficiency on thickness for a $Tm(d_1)Mo(d_2)SiO_2$ multilayer in the region of small values of d_2 near the critical point. For clearer viewing each curve has been shifted by 0.5 along the vertical axis. The lowest curve (which remains at zero) corresponds to $d_2 = 21$ Å. The value of d_2 for each successive curve is 1 Å greater than for the preceding curve. The upper curve corresponds to $d_2 = 30$ Å.
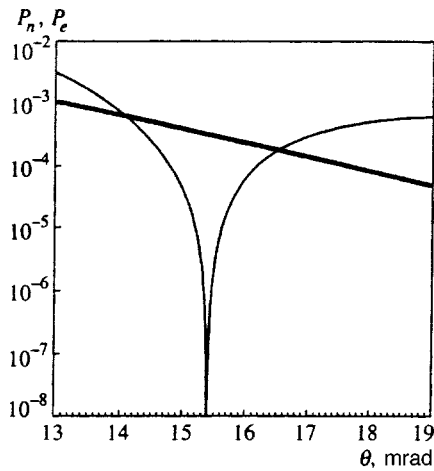
FIG. 5. Angular dependence of the electron and nuclear (thick line) reflectivities for a $Tm(d_1)Mo(d_2)SiO_2$ multilayer with $d_1 = 28.2$ Å and $d_2 = 24$ Å.

incidence of the beam appreciably surpasses the critical angle for total external reflection. As an example, Fig. 5 shows plots of the angular dependence of the electron and nuclear reflectivities for a layered system with $d_1 = 28.2$ Å and $d_2 = 24$ Å. It is easy to see that a high efficiency level is achieved owing to the very broad dip in electron reflectivity, while the level of nuclear reflectivity is quite small. A fairly powerful source of synchrotron radiation and a technology for fabricating multilayer structures with high-precision layers are needed to utilize this case.

The system under consideration also has solutions with a phase difference of 3π in the range of values of d_2 from 0 to 140 Å. From the standpoint of the stability of the efficiency parameter as the thickness is varied, there is interest in the range of values of d_2 near the maximum, where the two solutions have similar values of d_1. Figure 6 shows the (d_1,d_2) dependence of the efficiency in this range. As follows from the figure, an efficiency greater than 0.01 can be

obtained within a region having the dimensions $\Delta d_1 = 20$ Å and $\Delta d_2 = 7$ Å near the point $d_1 = 220$ Å, $d_2 = 136$ Å.

3.2. (Ti, Tm) layered system

The systems considered above contain a Tm layer as the upper layer. This specifies a high level of nuclear reflectivity. On the other hand, the susceptibility difference cannot exceed a certain limit. In particular, it is impossible to find a substance which would have a susceptibility twice as large as that of Tm. Such a case is roughly realized if Tm nuclei are found in a substrate coated from above by a Ti layer (see Table I), i.e., in a system described by the formula $Ti(d)Tm$. When the phase difference equals π, there are again two solutions on the plots of the angular dependence of d_{im} and d_{di} for this system. Figure 7 shows plots of the thickness dependence of the efficiency for these solutions. The efficiency maximum is at $d = 78.5$ and 211 Å, respectively. The solution with the larger thickness corresponds to the nuclear reflectivity maximum on the angular dependence, which is slightly weakened by absorption in the upper Ti layer. This is because the angle of incidence of the synchrotron radiation beam in this case is smaller than the critical angle for Tm and lies between the values of the critical angles for Tm and Ti (the critical angle for Ti $\theta_c = 4.97$ mrad). However, the dip in electron reflectivity is extremely narrow, leading to a low level of efficiency (0.002) and weak sensitivity toward variation of the thickness.

It is noteworthy that there is a correlation between the orientation of the d_{im} and d_{di} curves at the crossing point on the angle-thickness plane [an example is given in Fig. 2 for the $Tm(d)Mo$ system] and the width of the efficiency peak as a function of thickness, on the one hand, as well as the width of the dip in electron reflectivity as a function of the angle, on the other hand. A large angle between the d_{im} and d_{di} curves in the direction of the angle axis specifies a narrow dip in electron reflectivity, while a small angle in the direc-
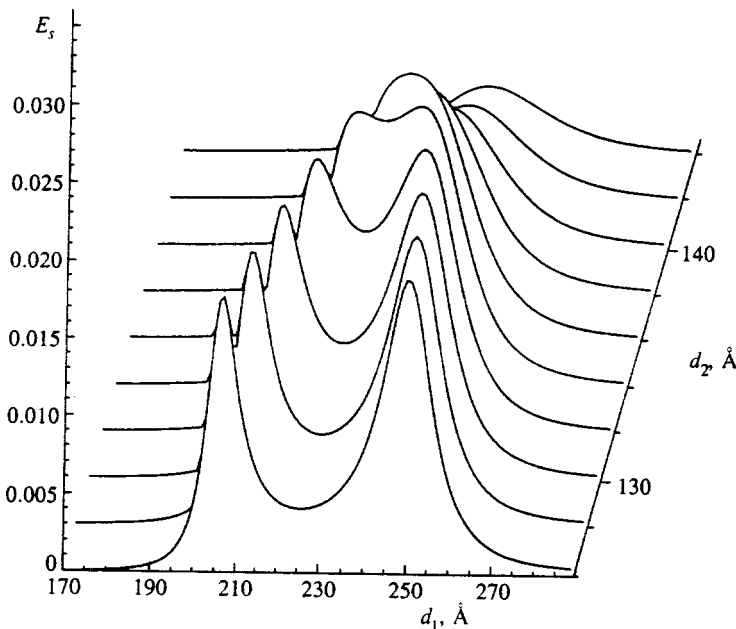

FIG. 6. Dependence of the efficiency on thickness for a $Tm(d_1)Mo(d_2)SiO_2$ multilayer in the region of large thicknesses near the critical point.
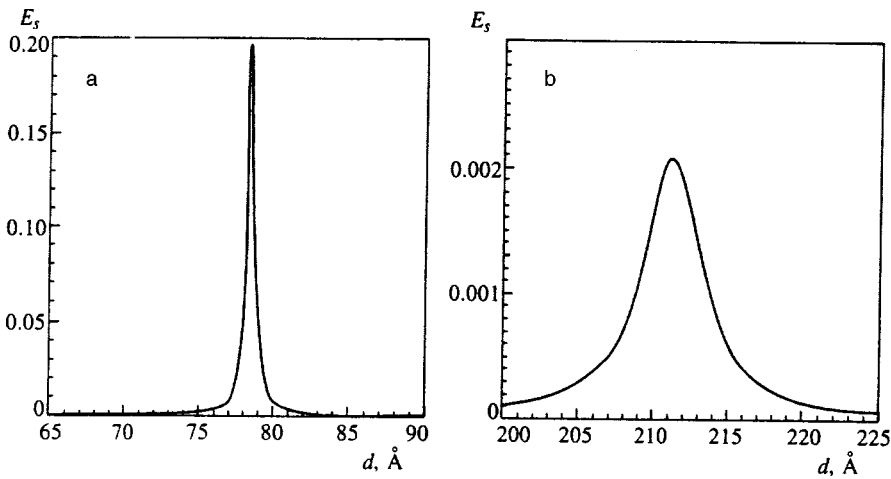
FIG. 7. Dependence of the efficiency on thickness for the two possible solutions (a and b) in the case of a Ti(d)Tm layered structure.

tion of the thickness axis leads to a broad peak on the plot of the dependence on thickness. The large slope of the d_{di} curve at small angles is the cause of the narrow dip in electron reflectivity in the case of large thicknesses. Thus, the angle-thickness diagrams clearly demonstrate why the case of small layer thicknesses leads both to an extremely high efficiency (0.2) of pure nuclear reflection and to a high sensitivity toward variation of the thickness (Fig. 7a). The level of nuclear reflectivity here is roughly the same as for large thicknesses, but the angular width of the dip in electron reflectivity is very large. Thus, this case can be preferable if there is a way to ensure a thin layer during fabrication of the sample.

3.3. (Tm, Fe) multilayer superlattice

A totally different situation arises in the case of layered systems containing Tm and Fe. Because of the small difference between their susceptibilities, the reflectivity of the Tm-Fe interface cannot make up for the strong reflectivity of the Tm-air interface. As a result, damping-stabilized solutions are impossible. Instead, a new Bragg antipeak effect is possible. As was noted above, the possibility of destructive interference between the reflection amplitude from the surface of a sample and the reflection amplitude from a multilayer superlattice depends strongly on the roughness of the interfaces between the layers independently of the roughness of the surface. This is because roughness strongly decreases the reflection amplitude from a multilayer superlattice regardless of the number of periods. For a multilayer superlattice with ideal interfaces between the layers there is a possibility for destructive interference over a broad range of angles, including fairly large angles, at which high-order Bragg peaks appear. Nevertheless, only the smallest of the possible values of the angle of incidence of the synchrotron radiation beam are of interest from the standpoint of obtaining a high level of nuclear reflectivity at the Bragg angle.

For multilayer superlattices with a $Tm(d_1)Fe(d_2)$ period and rough interfaces between the layers, only the first-order Bragg peak can provide a reflection amplitude comparable to reflection from the surface. Figure 8 shows the dependence of the efficiency on thickness for a superlattice containing 10 periods with a roughness parameter $\sigma = 3$ Å. The efficiency

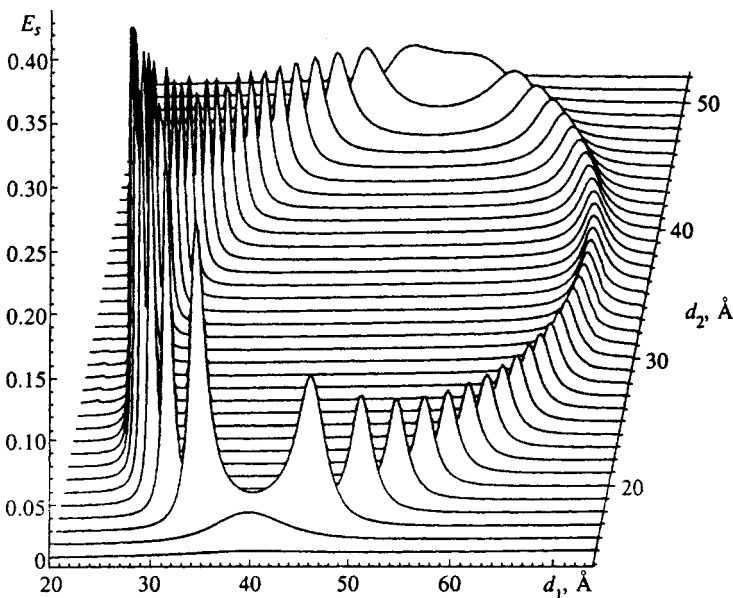FIG. 8. Dependence of the efficiency on the layer thicknesses d_1 and d_2 for a $[Tm(d_1)Fe(d_2)]*10/SiO_2$ superlattice.

rises sharply on the closed line, along which the conditions for the complete suppression of electron reflection are satisfied. The side with large values of d_1 corresponds to the suppression of electron reflection precisely at the center of the Bragg peak, while on the left-hand side (small values of d_1) with higher and narrower peaks, suppression occurs on the large-angle slope of the Bragg maximum. The angular position of the Bragg peak can be calculated approximately from the values of d_1 and d_2 using the simple formula

$$(\theta^2 - \theta_c^2)^{1/2} = \lambda/2(d_1 + d_2),$$

where θ_c is the mean value of the critical angles of Tm and Fe ($\theta_c = 6.45$ mrad).

For all values of the layer thicknesses within the closed line, the electron reflectivity curve has a dip on the angular dependence, but it is too small to provide a high level of efficiency, since the minimum value is above zero. As the calculations show, the size of the region within the closed line depends strongly on the number of periods in the multilayer superlattice up to a certain limiting number, which is governed by absorption, as well as on the level of roughness. For example, the conditions for destructive interference vanish entirely for the value of the roughness parameter $\sigma = 5$ Å. However, when the Bragg antipeak effect occurs, the efficiency of pure nuclear reflection is fairly high and stable toward variation of the thickness. Thus, this case can also be utilized to obtain pure nuclear scattering of synchrotron radiation.

4. CONCLUSION

The purpose of the present work was to find and analyze conditions under which electron reflection by a multilayer structure can be suppressed to a considerable extent while there is a high level of nuclear resonant reflection for ^{169}Tm. This is needed to set off the very narrow energy range (10^{-7} eV) corresponding to the nuclear resonant scattering of a synchrotron radiation beam by Tm nuclei. To evaluate the effect, we have introduced an efficiency parameter as the ratio of the intensity of the radiation scattered by nuclei (with a delay) to the electron scattering intensity (the prompt scattering) in energy and angular ranges that are attainable in third-generation synchrotron radiation sources. In principle, an efficiency greater than 0.0001 is sufficient for measurements without a time window.

The method used in the theoretical analysis is based on the Fresnel formulas for reflection on an interface between two homogeneous media. This method was developed by Parratt for smooth boundaries and extended here to the case of rough boundaries. A set of possible layered structures capable of providing the required efficiency level has been found. Cases where a Tm layer is accompanied by a layer with a considerably larger electron density (Mo) or with a considerably smaller electron density (Ti), as well as with a similar electron density (Fe), can be utilized. In the last case a new method for suppressing electron scattering, viz., the Bragg antipeak effect, has been found. As a rule, a high efficiency level is accompanied by strong sensitivity toward variation of the parameters. Only the sensitivity toward

variation of the layer thicknesses has been analyzed in detail, but the sensitivity toward variation of the susceptibility or surface roughness parameters can be of the same order.

From the standpoint of the stability of pure nuclear scattering with respect to variation of the parameters, cases with intermediate efficiency values (0.01–0.1) can be preferable. Such conditions have been found in a $Tm(d_1)Mo(d_2)SiO_2$ multilayer with relatively thick layers and in a $[Tm(d_1)Fe(d_2)]^*n/SiO_2$ multilayer superlattice with a low level of interfacial roughness.

*)E-mail: kohn@kurm.polyn.kiae.su

[1] L. Ruby, J. Phys. (Paris) **35**, C6-209 (1974).

[2] R. L. Mössbauer, in *Proceedings of the International Union of Crystallography*, *Madrid*, *April, 1974*, p. 463.

[3] E. Gerdau, R. Rüffer, H. Winkler, W. Tolksdorf, C. P. Klages, and J. P. Hannon, Phys. Rev. Lett. **54**, 835 (1985).

[4] G. Faigel, D. P. Siddons, J. B. Hastings, P. E. Haustein, J. R. Grover, J. P. Remeika, and A. S. Cooper, Phys. Rev. Lett. **58**, 2699 (1987).

[5] T. Ishikawa, Y. Yoda, K. Izumi, C. K. Suzuki, X. W. Zhang, M. Ando, and S. Kikuta, Rev. Sci. Instrum. **63**, 1015 (1992).

[6] T. S. Toellner, T. Mooney, S. Shastri, and E. E. Alp, Proc. SPIE **1740**, 218 (1992).

[7] A. I. Chumakov, J. Metge, A. Q. R. Baron, H. Grünsteudel, H. F. Grünsteudel, R. Rüffer, and T. Ishikawa, Nucl. Instrum. Methods Phys. Res. A **383**, 642 (1996).

[8] S. Kishimoto, Nucl. Instrum. Methods Phys. Res. A **309**, 603 (1991).

[9] A. Q. R. Baron and S. L. Ruby, Nucl. Instrum. Methods Phys. Res. A **343**, 517 (1994).

[10] J. B. Hastings, D. P. Siddons, U. van Bürck, R. Hollatz, and U. Bergmann, Phys. Rev. Lett. **66**, 770 (1991).

[11] D. P. Siddons, U. Bergmann, and J. B. Hastings, Phys. Rev. Lett. **70**, 359 (1993).

[12] T. S. Toellner, E. E. Alp, W. Sturhahn, T. M. Mooney, X. Zhang, M. Ando, Y. Yoda, and S. Kikuta, Appl. Phys. Lett. **67**, 1993 (1995).

[13] J. P. Hannon, N. V. Hung, G. T. Trammel, E. Gerdau, M. Mueller, R. Rüffer, and H. Winkler, Phys. Rev. B **32**, 5068, 5081 (1985).

[14] J. P. Hannon, G. T. Trammel, M. Mueller, E. Gerdau, R. Rüffer, and H. Winkler, Phys. Rev. B **32**, 6363, 6374 (1985).

[15] R. Röhlsberger, E. Gerdau, M. Harsdorff, O. Leupold, E. Lüken, J. Metge, R. Rüffer, H. D. Rüter, W. Sturhahn, and E. Witthoff, Europhys. Lett. **18**, 561 (1992).

[16] E. E. Alp, T. M. Mooney, T. Toellner, W. Sturhahn, E. Witthoff, R. Röhlsberger, E. Gerdau, H. Homma, and M. Kentjana, Phys. Rev. Lett. **70**, 3351 (1993).

[17] A. I. Chumakov, G. V. Smirnov, A. Q. R. Baron, J. Arthur, D. E. Brown, S. L. Ruby, G. S. Brown, and N. N. Salashchenko, Phys. Rev. Lett. **71**, 2489 (1993).

[18] R. Röhlsberger, E. Witthoff, E. Lüken, and E. Gerdau, in *Physics of X-Ray Multilayer Structures, Optical Society of American Technical Digest*, Jackson Hole, WY (1992), Vol. 7, p. 178.

[19] T. S. Toellner, W. Sturhahn, R. Röhlsberger, E. E. Alp, C. H. Sowers, and E. E. Fullerton, Phys. Rev. Lett. **74**, 3475 (1995).

[20] H. Homma, M. Kentjana, E. E. Alp, T. M. Mooney, E. Witthoff, and T. Toellner, J. Appl. Phys. **72**, 5668 (1992).

[21] W. Sturhahn, E. Gerdau, R. Hollatz, R. Rüffer, H. D. Rüter, and W. Tolksdorf, Europhys. Lett. **14**, 821 (1991).

[22] N. N. Salashchenko, Yu. Ya. Platonov, and S. Yu. Zuev, Nucl. Instrum. Methods Phys. Res. A **359**, 114 (1995).

[23] L. G. Parratt, Phys. Rev. **45**, 359 (1954).

[24] A. G. Michette, *Optical System for Soft X-Rays*, Plenum Press, New York (1986).

[25] L. Nevot and P. Croce, Rev. Phys. Appl. **15**, 761 (1980).

[26] V. G. Kohn, Phys. Status Solidi B **187**, 61 (1995).

[27] Yu. M. Kagan, A. M. Afanasev, and V. G. Kohn, J. Phys. C: Solid State Phys. **12**, 615 (1979).

[28] A. Q. R. Baron, J. Arthur, S. L. Ruby, A. I. Chumakov, G. V. Smirnov, and G. S. Brown, Phys. Rev. B **50**, 10 354 (1994).

Translated by P. Shelnitz

Fluctuational spatial dispersion in achiral liquid crystals

M. V. Gorkunov and M. I. Ryazanov*[)]

Moscow State Engineering-Physics Institute (Technical University), 115409 Moscow, Russia
(Submitted 23 January 1998)
Zh. Éksp. Teor. Fiz. **114**, 182–190 (July 1998)

An investigation is made of the spatial dispersion of the permittivity of achiral liquid crystals
with large-scale fluctuations. It is observed that for large correlation lengths the terms
with spatial dispersion are anomalously large. The specific form of these terms is obtained for
orientational fluctuations in a nematic and fluctuations of the deformations of the smectic
layers in a smectic-A in an orienting magnetic field. It is shown that these effects may be observed
optically by means of accurate measurements of the angular dependence of the refractive
indices of electromagnetic waves. © *1998 American Institute of Physics.*
[S1063-7761(98)01307-9]

1. INTRODUCTION

Local coupling between the electrical induction and the field may be impaired in the presence of spatial fluctuations of the permittivity. The spatial dispersion obtained by allowing for these fluctuations is naturally called the fluctuation dispersion. This fluctuation spatial dispersion has not been analyzed previously, so that a material with these properties is an interesting topic for discussion.

It has been established that allowance for spatial dispersion in nongyrotropic media for frequencies some distance from the absorption line yields corrections of order the square of the ratio of the characteristic microscopic length (such as the average electron displacement under the action of the field, the lattice constant, and so on) to the wavelength of the field.[1,2] If a medium with large-scale fluctuations is considered, its characteristic length is the correlation length, so that for a large correlation length the corrections arising from the spatial dispersion should be quite significant. This effect is analyzed subsequently in liquid crystals. The similar situation encountered near the critical point in liquids will be studied elsewhere.

By their very nature liquid crystals are materials with pronounced fluctuations. Calculations of the dielectric properties of liquid crystals, for example, using local field theory,[3] usually only take account of microscopic fluctuations of the position and orientation of the nearest-neighbor molecules. These fluctuations actually play a major role and strongly influence the calculated principal values of the permittivity tensor. It is clear, however, that the corresponding spatial dispersion should be relatively low.

Some types of fluctuations in liquid crystals have large macroscopic dimensions: their correlation lengths are many times greater than the intermolecular distances.[4] In particular, near the phase transition point the correlation length of the pretransition fluctuations depends strongly on temperature, increasing as the transition point is approached. The orientational fluctuations in a nematic or the fluctuations of the smectic layer deformation in a smectic-A have macroscopic correlation lengths over the entire temperature range

where the phase exists. In this case, it is typically found that in the absence of external orienting factors, the correlation lengths are formally infinitely large. In practice, however, the surfaces, and the external electric and magnetic fields always have a substantial orienting influence in liquid crystals, which leads to finite correlation lengths, the latter becoming controllable parameters of the liquid crystals.[5]

Since long-wavelength fluctuations in nematics and smectic-A liquid crystals in fact induce local changes in the direction of the principal optic axis, it is possible to study the fluctuational spatial dispersion in these by assuming that the material is an inhomogeneous locally uniaxial medium. To illustrate the results, we use the well-known correlation functions of liquid crystals in a magnetic field.

Calculations of the electromagnetic wave spectra show that the fluctuations influence the angular dependences of the refractive indices of the liquid crystals. A similar effect was considered previously for ordinary waves in a nematic[6] where numerical results were obtained for very large correlation lengths. An analysis in terms of the spatial dispersion is clearly valid for correlation lengths shorter than the wavelength of light, i.e., in well-oriented liquid crystals. In this approach the corrections to the spectra of the ordinary and extraordinary waves have a simple analytic form.

The estimates made show these corrections caused by fluctuational spatial dispersion can be observed by optical methods.

2. SPATIAL DISPERSION IN A LIQUID CRYSTAL WITH FLUCTUATIONS OF THE DIRECTOR ORIENTATION

We shall analyze a locally uniaxial medium with the fluctuational direction **n** of the principal optic axis:

$$\varepsilon_{ij}(\mathbf{r},\omega) = \varepsilon_a(\omega)\delta_{ij} + (\varepsilon_e(\omega) - \varepsilon_o(\omega))n_i(\mathbf{r})n_j(\mathbf{r}). \quad (1)$$

Assuming that the deviations of the vector **n** from its average **n**$_0$ are small, we have

$$\mathbf{n} = \mathbf{n}_0 + \delta\mathbf{n}, \quad \mathbf{n}_0 \cdot \delta\mathbf{n} \approx 0. \quad (2)$$

Using linear terms in $\delta\mathbf{n}$, we rewrite formula (1) in the form

$$\varepsilon_{ij}(\mathbf{r},\omega) = \varepsilon_{ij}(\omega) + \delta\varepsilon_{ij}(\mathbf{r},\omega), \tag{3}$$

where

$$\varepsilon_{ij}(\omega) = \varepsilon_o(\omega)\delta_{ij} + \varepsilon_a(\omega)e_i e_j,$$

$$\delta\varepsilon_{ij}(\mathbf{r},\omega) = \varepsilon_a(\omega)(e_i\delta\mathbf{n}_j + \delta\mathbf{n}_i e_j),$$

and $\varepsilon_a = n_0^2(\varepsilon_e - \varepsilon_o)$ is the optical anisotropy of the liquid crystal averaged over the fluctuations. The unit vector $\mathbf{e}$ is directed along $\mathbf{n}_0$ and thus defines the principal optic axis of the liquid crystal as a whole.

The wave equation for the electric field components

$$\nabla \operatorname{div} \mathbf{E} - \Delta \mathbf{E} = \frac{\omega^2}{c^2}\mathbf{D} \tag{4}$$

after Fourier transforming and using Eq. (3) yields

$$\left[q^2\delta_{ij} - q_i q_j - \varepsilon_{ij}(\omega)\frac{\omega^2}{c^2} \right] E_j(\mathbf{q},\omega)$$

$$= \frac{\omega^2}{c^2}\int d^3 p\, \delta\varepsilon_{ij}(\mathbf{p},\omega) E_j(\mathbf{q}-\mathbf{p},\omega). \tag{5}$$

Integrating Eq. (5), we can easily obtain

$$\left[q^2\delta_{ij} - q_i q_j - \varepsilon_{ij}(\omega)\frac{\omega^2}{c^2} \right] E_j(\mathbf{q},\omega)$$

$$= \frac{\omega^4}{c^4}\int d^3 p \int d^3 p'\, \delta\varepsilon_{ij}(\mathbf{p},\omega)\delta\varepsilon_{kl}(\mathbf{p}',\omega)$$

$$\times\Bigg[(\mathbf{q}-\mathbf{p})^2\delta_{jk} - (\mathbf{q}-\mathbf{p})_j(\mathbf{q}-\mathbf{p})_k$$

$$- \varepsilon_{jk}(\omega)\frac{\omega^2}{c^2} \Bigg]^{-1} E_l(\mathbf{q}-\mathbf{p}-\mathbf{p}',\omega). \tag{6}$$

Since the optical anisotropy of a liquid crystal is usually weak, in calculations of the fluctuation corrections it can be neglected wherever it is not of fundamental importance. In particular, on the right-hand side of formula (6) we can set

$$\left[(\mathbf{q}-\mathbf{p})^2\delta_{jk} - (\mathbf{q}-\mathbf{p})_j(\mathbf{q}-\mathbf{p})_k - \varepsilon_{jk}(\omega)\frac{\omega^2}{c^2} \right]^{-1}$$

$$\simeq \frac{\varepsilon_o(\omega/c)^2\delta_{jk} - (\mathbf{q}-\mathbf{p})_j(\mathbf{q}-\mathbf{p})_k}{\varepsilon_o(\omega/c)^2[(\mathbf{q}-\mathbf{p})^2 - \varepsilon_o(\omega/c)^2]}. \tag{7}$$

Averaging over the fluctuations assuming a homogeneity of the averaged liquid crystal yields

$$\langle \delta\varepsilon_{ij}(\mathbf{p},\omega)\delta\varepsilon_{kl}(\mathbf{p}'\omega) \rangle$$

$$= \langle \delta\varepsilon_{ij}(\mathbf{p},\omega)\delta\varepsilon_{kl}(-\mathbf{p},\omega) \rangle \delta^{(3)}(\mathbf{p}+\mathbf{p}'). \tag{8}$$

Substituting expressions (7) and (8) into formula (6) gives the wave equation which contains the following tensor as the permittivity

$$\varepsilon_{ij}(\mathbf{q},\omega) = \varepsilon_{ij}(\omega) + \frac{1}{\varepsilon_o}\int d^3 p\, \langle \delta\varepsilon_{ij}(\mathbf{p},\omega)\delta\varepsilon_{kl}(-\mathbf{p},\omega) \rangle$$

$$\times\frac{\varepsilon_o(\omega/c)^2\delta_{jk} - (\mathbf{q}-\mathbf{p})_j(\mathbf{q}-\mathbf{p})_k}{(\mathbf{q}-\mathbf{p})^2 - \varepsilon_o(\omega/c)^2}. \tag{9}$$

Using Eq. (3), we can easily obtain

$$\langle \delta\varepsilon_{ij}(\mathbf{p},\omega)\delta\varepsilon_{kl}(-\mathbf{p},\omega) \rangle = \varepsilon_a^2[e_i e_k g_{jl}(\mathbf{p}) + e_i e_l g_{jk}(\mathbf{p})$$

$$+ e_j e_k g_{il}(\mathbf{p}) + e_j e_l g_{ik}(\mathbf{p})], \tag{10}$$

where $g_{ms}(\mathbf{p})$ is the Fourier transform of the correlation function

$$g_{ms}(\mathbf{r}) = \langle \delta n_m(\mathbf{r}')\delta n_s(\mathbf{r}'-\mathbf{r}) \rangle. \tag{11}$$

In fact, formula (9) only takes into account the contribution of two-photon scattering processes at fluctuations. Multiphoton scattering can justifiably be neglected because the perturbation of the permittivity is proportional to the dielectric anisotropy, so that the discarded terms contain the superfluous factor ε_a^2 which is almost always small. In addition, the orientational elasticity of the liquid crystal also has the result that the deviations of the director from the equilibrium direction are small, making the corrections to the permittivity even smaller and the role of the multiphoton processes even more negligible.

The wave vectors $|\mathbf{p}| \lesssim \xi^{-1}$ make some contribution to the integral in Eq. (9), where ξ is the correlation length over which $g_{ms}(\mathbf{r})$ decreases. Since the macroscopic description is valid for wavelengths

$$\lambda \gg \xi, \tag{12}$$

in Eq. (9) we can obviously assume $q \ll p$, $(\omega/c) \ll p$, i.e., an expansion in terms of small q and (ω/c) is valid.

The terms in this expansion which do not depend on $\mathbf{q}$ contribute to the permittivity without spatial dispersion. This contribution is obviously small compared with the contributions of the short-wavelength correlations[3] and thus we shall denote the entire wave-vector-independent component of the permittivity as $\varepsilon_{ij}(\omega)$, assuming that this takes into account all the fluctuation corrections and is actually observed experimentally. Integration of terms linear in $\mathbf{q}$ gives zero since the medium, and with it the correlation function, is assumed to have a center of inversion. The quadratic terms determine the spatial dispersion of the permittivity so that

$$\varepsilon_{ij}(\mathbf{q},\omega) = \varepsilon_{ij}(\omega) + \frac{1}{\varepsilon_o}\int d^3 p\, \langle \delta\varepsilon_{ij}(\mathbf{p},\omega)\delta\varepsilon_{kl}(-\mathbf{p},\omega) \rangle \frac{1}{p^2}$$

$$\times\Bigg[-q_j q_k + (q_j p_k + p_j q_k)\frac{2(\mathbf{q}\cdot\mathbf{p})}{p^2}$$

$$+ p_j p_k\frac{q^2 p^2 - 4(\mathbf{q}\cdot\mathbf{p})^2}{p^4} \Bigg]. \tag{13}$$

Since we have $\delta\mathbf{n} \perp \mathbf{e}$, only those components of the tensor $g_{ms}(\mathbf{p})$ corresponding to directions perpendicular to $\mathbf{e}$ are nonzero. Additionally, we can assume that its symmetry is the same as that of an averaged liquid crystal, as is the case, e.g., in an orienting external field. We can then write

$$g_{ms}(\mathbf{p}) = (\delta_{ms} - e_m e_s)u(p_\parallel, p_\perp) + \frac{p_{\perp m}p_{\perp s}}{p_\perp^2}v(p_\parallel, p_\perp), \tag{14}$$

where $\mathbf{p}_\parallel$ and $\mathbf{p}_\perp$ are the longitudinal and perpendicular components of the vector $\mathbf{p}$ relative to the $\mathbf{e}$ optic axis.

Substituting Eq. (14) into Eqs. (10) and (9) yields explicit forms of the terms with spatial dispersion

$$\varepsilon_{ij}(\mathbf{q},\omega) = \varepsilon_{ij}(\omega) + \frac{\varepsilon_a^2(\omega)}{\varepsilon_o(\omega)}\gamma_{ijkl}q_k q_l,$$

$$\gamma_{ijkl}q_k q_l = q^2[a\,\delta_{ij} + b e_i e_j] + q_\parallel^2[d\,\delta_{ij} + f e_i e_j]$$
$$+ q_\parallel(q_i e_j + e_i q_j)g + q_i q_j h, \qquad (15)$$

and the frequency-independent constants

$$a = -u_1 + 2u_2 + \frac{1}{2}v_2,$$

$$b = 2u_1 - 4u_2 + \frac{1}{2}v_0 + v_1 - \frac{3}{2}v_2,$$

$$d = 6u_1 - 6u_2 - u_0 - \frac{1}{2}v_0 + \frac{5}{2}v_1 - \frac{5}{2}v_2,$$

$$f = 2u_0 - 6u_1 + 4u_2 - \frac{3}{2}v_1 + \frac{3}{2}v_2,$$

$$g = -3u_1 + 4u_2 + \frac{1}{2}v_0 - 3v_1 + 3v_2,$$

$$h = v_2 - v_1 \qquad (16)$$

are expressed in terms of the integrals

$$u_\alpha = \int d^3p\,\frac{u(p_\parallel,p_\perp)}{p^2}\left(\frac{p_\parallel}{p}\right)^{2\alpha},$$

$$v_\alpha = \int d^3p\,\frac{v(p_\parallel,p_\perp)}{p^2}\left(\frac{p_\parallel}{p}\right)^{2\alpha}, \qquad \alpha = 0,1,2. \qquad (17)$$

These have dimensions of the square of length. Their smallness compared with λ^2 is responsible for the smallness of the terms with spatial dispersion.

3. NEMATIC LIQUID CRYSTAL IN AN ORIENTING MAGNETIC FIELD

An infinite nematic in an orienting magnetic field $\mathscr{H}$ has finite correlation lengths of orientational fluctuations, and these are proportional to $\mathscr{H}^{-1}$. If the anisotropy of the magnetic susceptibility is $\chi_a > 0$, the optic axis is oriented in the direction of the field and the correlation function is given by[5]

$$u(p_\parallel,p_\perp) = \frac{k_B T}{K_1 p_\perp^2 + K_3 p_\parallel^2 + \chi_a\mathscr{H}^2},$$

$$u(p_\parallel,p_\perp) + v(p_\parallel,p_\perp) = \frac{k_B T}{K_2 p_\perp^2 + K_3 p_\parallel^2 + \chi_a\mathscr{H}^2}, \qquad (18)$$

where K_α are the corresponding elastic moduli of the liquid crystal.

Since these moduli usually have comparatively similar values, we can assume for simplicity in the so-called single-constant approximation that $K_1 = K_2 = K_3 = K$. Then, we have $v = 0$ and

$$u(p_\parallel,p_\perp) = u(p) = \frac{k_B T}{K(p^2 + \xi^{-2})}, \qquad (19)$$

where the correlation length is $\xi = \sqrt{K/\chi_a}\,\mathscr{H}^{-1}$. Thus, $v_\alpha = 0$ and

$$u_0 = 2\pi^2\frac{k_B T}{K}\xi, \quad u_1 = \frac{1}{3}u_0, \quad u_2 = \frac{1}{5}u_0. \qquad (20)$$

The permittivity with allowance for the spatial dispersion has the form

$$\varepsilon_{ij}(\mathbf{q},\omega) = \varepsilon_{ij}(\omega) + \frac{\varepsilon_a^2(\omega)u_0}{15\varepsilon_o(\omega)}\{q^2(\delta_{ij} - 2e_i e_j)$$
$$- q_\parallel^2(3\delta_{ij} - 12e_i e_j) - 3q_\parallel(q_i e_j + e_i q_j)\}. \qquad (21)$$

Estimating $T \sim 400$ K, $K \sim 10^{-6}$ dyn, $q \sim 10^4$ cm^{-1}, $\varepsilon_o \sim 1$, and $\varepsilon_a^2 \sim 0.1$, we have in order of magnitude

$$\frac{\varepsilon_a^2}{\varepsilon_o}u_0 \sim 10^{-3}(\xi q), \qquad (22)$$

i.e., since $(\xi q) < 1$, the terms with spatial dispersion may have values of 10^{-3}. Assuming that $\chi_a \sim 10^{-7}$, we also find that inequality (12) which restricts the validity of the macroscopic analysis of the fluctuations to fairly well-oriented liquid crystals, requires a strong orienting magnetic field $\mathscr{H} > 10^4$ Oe.

4. ACHIRAL SMECTIC LIQUID CRYSTAL IN A MAGNETIC FIELD

We shall analyze a single-domain smectic-*A* oriented by a magnetic field. Since the distortions of the position order in the layer structure are only important for us because they change the local direction of the optic axis, the question of violations of the position order caused by instability of a one-dimensionally ordered structure can be left to one side. An important factor is that the orientational nematic order in smectics is fairly high.

The principal optic axis in a smectic-*A* is directed along the normal to the surface of the smectic layers. Thus, if the layers are deformed by long-wavelength fluctuations, we have

$$\mathbf{n}(\mathbf{r}) = \mathbf{n}_0 - \frac{\partial}{\partial\mathbf{r}_\perp}w(\mathbf{r}), \qquad (23)$$

where $w(\mathbf{r})$ is the deviation of the layer from the equilibrium smectic plane. The correlation function is then

$$g_{ms}(\mathbf{p}) = -p_{\perp m}p_{\perp s}\langle|w(\mathbf{p})|^2\rangle, \qquad (24)$$

i.e., in a smectic-*A*

$$u(p_\parallel,p_\perp) = 0, \quad v(p_\parallel,p_\perp) = -p_\perp^2\langle|w(p_\parallel,p_\perp)|^2\rangle. \qquad (25)$$

In an orienting magnetic field we have

$$\langle|w(p_\parallel,p_\perp)|^2\rangle = \frac{k_B T}{Bp_\parallel^2 + K(p_\perp^4 + \chi_a\mathscr{H}^2 p_\perp^2)}, \qquad (26)$$

where B is the bulk modulus of the smectic layers.[5]

Introducing the correlation lengths perpendicular and parallel to the optic axis,

$$\xi_\perp = \sqrt{\frac{K}{\chi_a}}\frac{1}{\mathscr{H}}, \quad \xi_\parallel = \xi_\perp^2\sqrt{\frac{B}{K}}, \qquad (27)$$

we can write

$$\nu(p_\parallel, p_\perp) = -\frac{k_B T p_\perp^2 \xi_\perp^4}{K(p_\parallel^2 \xi_\parallel^2 + p_\perp^2 \xi_\perp^2 + p_\perp^4 \xi_\perp^4)}. \tag{28}$$

Since the smectic layers are usually weakly compressible, the value of B is such that the parameter $\sqrt{K/B}$ is small and has the order of the molecular lengths, i.e.,

$$\xi_\parallel \gg \xi_\perp. \tag{29}$$

The integrals ν_α are then

$$\nu_0 = -\frac{2\pi^2 k_B T}{\sqrt{KB}} \ln\left(\frac{2\xi_\parallel}{\xi_\perp}\right), \quad \nu_1 = -\frac{\pi^2 k_B T}{\sqrt{KB}}, \quad \nu_2 = \frac{1}{2}\nu_1. \tag{30}$$

A comparison with Eq. (19) shows that in a smectic-A the spatial dispersion is $\xi_\perp \sqrt{B/K}$ times lower than that in a nematic. An estimate for $B \sim 10^8$ erg/cm^3 and $\xi_\perp \sim 10^{-5}$ cm gives the characteristic order $\sim 10^{-5}$.

The physical reasons for these differences between smectics and nematics become clear if we take into account that the spatial dispersion is determined by the integrals over correlated regions of the liquid crystals. In a smectic-A these regions are highly elongated in the direction of the principal optic axis, and since condition (12) should be satisfied for all their characteristic dimensions, the integrated contribution from the almost spherical regions in a nematic is obviously substantially greater. In other words, the macroscopic description of the fluctuations in a smectic in terms of the spatial dispersion is valid when these effects are very weak.

However, it is known that near the second-order smectic-A–nematic phase transition, the smectic layers become smeared out so that the elastic modulus B is reduced substantially. In this case, the shape of the correlated regions approaches spherical. It is easy to see that at these temperatures the spatial dispersion is comparatively high. For example for $\xi_\parallel = \xi_\perp = \xi$, we have

$$\nu_0 = -2\pi^2 \frac{k_B T}{K}, \quad \nu_1 = \frac{1}{3}\nu_0, \quad \nu_2 = \frac{1}{5}\nu_0. \tag{31}$$

In this case, the permittivity (15) is given by

$$\varepsilon_{ij}(\mathbf{q}, \omega) = \varepsilon_{ij}(\omega) + \frac{\varepsilon_a^2(\omega)\nu_0}{15\varepsilon_o(\omega)}$$

$$\times \left\{ q^2\left(\frac{3}{2}\delta_{ij} + 8e_i e_j\right) - q_\parallel^2\left(\frac{5}{2}\delta_{ij} + 3e_i e_j\right) \right.$$

$$\left. + \frac{3}{2}q_\parallel(q_i e_j + e_i q_j) - 2q_i q_j \right\}, \tag{32}$$

and the spatial dispersion has the same order of magnitude as in a nematic.

5. SPECTRUM OF ELECTROMAGNETIC WAVES IN A UNIAXIAL MEDIUM WITH SPATIAL DISPERSION

The lightwave spectrum is given by the nulls of the determinant

$$|q^2 \delta_{ij} - q_i q_j - (\omega/c)^2 \varepsilon_{ij}(\mathbf{q}, \omega)| = 0. \tag{33}$$

As a result of substituting Eq. (15), this equation separates into two: the dispersion equation for the ordinary waves

$$q^2 = \left(\frac{\omega}{c}\right)^2 \left\{ \varepsilon_o(\omega) + \frac{\varepsilon_a^2(\omega)}{\varepsilon_o(\omega)}(aq^2 + d(\mathbf{q}\cdot\mathbf{e})^2) \right\} \tag{34}$$

and that for the extraordinary waves

$$q^4 G(\omega, \theta) + (\omega/c)^2 q^2 F(\omega, \theta) = (\omega/c)^4 \varepsilon_o(\omega)\varepsilon_e(\omega), \tag{35}$$

where θ is the angle between the vectors $\mathbf{q}$ and $\mathbf{e}$ and

$$G(\omega, \theta) = \left(\frac{\omega}{c}\right)^2 \frac{\varepsilon_a^2(\omega)}{\varepsilon_o(\omega)} \{ a + h + (b + d + 2g)$$

$$\times \cos^2 \theta + f \cos^4 \theta \}, \tag{36}$$

$$F(\omega, \theta) = \varepsilon_o(\omega)\sin^2 \theta + \varepsilon_e(\omega)\cos^2 \theta$$

$$- \left(\frac{\omega}{c}\right)^2 \frac{\varepsilon_a^2(\omega)}{\varepsilon_o(\omega)} \{ [a+b]\varepsilon_o(\omega) + [a+h]\varepsilon_e(\omega)$$

$$+ [f+d+h+2g]\varepsilon_o(\omega)\cos^2 \theta$$

$$+ (d-h)\varepsilon_e(\omega)\cos^2 \theta \}. \tag{37}$$

The terms attributable to the spatial dispersion are small. It has been shown that these have the order 10^{-3} so that the approximate solutions of these equations are valid. For the ordinary waves, Eq. (34) gives

$$q_0^2 = \frac{\omega^2}{c^2}\left\{ \varepsilon_o + \varepsilon_a^2 \frac{\omega^2}{c^2}(a + d \cos^2 \theta) \right\}, \tag{38}$$

and for the extraordinary waves Eq. (35) yields

$$q_e^2 = \frac{\varepsilon_o \varepsilon_e (\omega/c)^2}{\varepsilon_o \sin^2 \theta + \varepsilon_e \cos^2 \theta}\left[1 + \left(\frac{\omega}{c}\right)^2 \frac{\varepsilon_a^2}{\varepsilon_o}(a + b \sin^2 \theta \right.$$

$$\left. + d \cos^2 \theta + f \cos^2 \theta \sin^2 \theta) \right]. \tag{39}$$

In the corrections caused by the spatial dispersion, we neglected the optical anisotropy of the liquid crystals as before.

Thus, when the spatial dispersion is taken into account, the ordinary wave spectrum becomes weakly anisotropic. The angular dependence of the extraordinary wave spectrum also changes.

6. DISCUSSION OF RESULTS

Incoherent scattering at long-wavelength fluctuations is considered to be the main reason for the poor transparency of liquid crystals. The corresponding imaginary correction to the permittivity is given by the pole of the integral in Eq. (9). Since the pole is positioned at $p^2 \sim \varepsilon(\omega/c)^{-2} \ll \xi^{-2}$, the corresponding part of the integral is

$$\frac{\lambda \int d^3 p\, p^{-2} g(\mathbf{p})}{g(p \to 0)} \tag{40}$$

times smaller than that corresponding to the spatial dispersion of the permittivity. Converting to coordinate-dependent Fourier transforms, we have the ratio

$$\frac{\varepsilon_{SD}}{\varepsilon''} \sim \frac{\lambda \int d^3 r \; r^{-1} g(\mathbf{r})}{\int d^3 r \; g(\mathbf{r})}, \tag{41}$$

i.e., the imaginary part is λ/ξ times smaller than the terms with spatial dispersion.

The factor allows us to make a preliminary estimate of the spatial dispersion in terms of the transparency of the liquid crystals. For example, assuming that the characteristic distances over which light is scattered significantly are of the order of a few millimeters, we obtain $\varepsilon'' \sim 10^{-3} - 10^{-4}$, i.e., terms with spatial dispersion should have an order no less than 10^{-3}. For liquid crystals of lower transparency these terms will be even greater, provided of course that the correlation lengths are shorter than the wavelengths of light.

In nongyrotropic solid crystals the spatial dispersion is usually of the order 10^{-6}. Thus, this dispersion in liquid crystals with long-wavelength fluctuations can be described as anomalously high.

The anisotropy of the refractive index for the ordinary waves, produced as a result of the spatial dispersion, and the corrections to the angular dependence of the refractive index of the extraordinary waves can obviously be measured experimentally. That is to say, these effects can be used to study fluctuation processes in liquid crystals by optical measurements of the angular dependences of the refractive indices. A combination of this method with the conventional measurements of the cross sections for incoherent scattering at fluctuations can be used to determine the correlation lengths and amplitudes of the fluctuations separately, opening up new prospects for detailed studies of fluctuation processes in liquid crystals.

*)E-mail: ryazanov@theor.mephi.msk.su

[1] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media*, transl. of 1st Russ. ed. (Pergamon Press, Oxford, 1960) [Russ. original, later ed., Nauka, Moscow, 1992)].

[2] V. M. Agranovich and V. L. Ginzburg, *Spatial Dispersion in Crystal Optics and the Theory of Excitons* (Wiley, New York, 1967) [Russ. original, Nauka, Moscow, 1965).

[3] E. M. Aver'yanov and M. A. Osipov, Usp. Fiz. Nauk **160**(5), 89 (1990) [Sov. Phys. Usp. **33**, 365 (1990)].

[4] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, Part. 1, 3rd. ed. (Pergamon Press, Oxford, 1980) [Russ. original, Part 1, 3rd. ed., Nauka, Moscow, 1976].

[5] P. G. de Gennes and J. Prost, *The Physics of Liquid Crystals* (Clarendon Press, Oxford, 1993).

[6] N. B. Baranova, B. Ya. Zel'dovich and V. S. Liberman, Zh. Éksp. Teor. Fiz. **99**, 1504 (1991) [Sov. Phys. JETP **72**, 841 (1991)].

Translated by R. M. Durham

Reduction in the rate of phonon scattering by a spatially correlated system of iron ions and low-temperature ''anomaly'' of the thermoelectric phenomena in HgSe:Fe crystals

I. G. Kuleev,*) A. T. Lonchakov, I. Yu. Arapova, and G. I. Kuleev

Institute of Metal Physics, Ural Branch of the Russian Academy of Sciences, 620219 Ekaterinburg, Russia
(Submitted 30 July 1997)

A new effect of the reduction in the rate of phonon scattering by the spatially correlated system of iron ions in HgSe:Fe crystals is detected experimentally and calculated theoretically. The thermoelectric power is measured using HgSe:Fe samples with different iron content in the temperature range 7.5–60 K. It is found that the dependence of the thermoelectric power on iron content exhibits remarkable features at $T < 10$ K: the quantity $|\alpha(N_{Fe})|$ increases as the iron concentration increases to $N_{Fe} = 5 \times 10^{18}$ cm^{-3}, reaches a maximum at $N_{Fe} \approx (1-2) \times 10^{19}$ cm^{-3}, but then monotonically decreases with further increases in N_{Fe}. It is shown that the observed increase in the thermoelectric power is due to a reduction in the rate of phonon scattering by the spatially correlated system of Fe^{3+} ions. This new effect is analyzed theoretically, and the theoretical results are compared with the experimental data. © *1998 American Institute of Physics.* [S1063-7761(98)01407-3]

1. INTRODUCTION

There is interest in electron-transfer phenomena in HgSe systems doped with iron because these compounds exhibit some remarkable properties.[1-9] An extraordinary ''anomaly'' is the substantial increase in electron mobility at liquid-helium temperatures as the iron concentration increases from $N_{Fe} = N^* = 4.5 \times 10^{18}$ cm^{-3} to $N_{Fe} = 2 \times 10^{19}$ cm^{-3} (Ref. 1). It has been shown[5-9] that the anomalous behavior of the dependence of the electron mobility and thermomagnetic effects on iron content and temperature is due to the formation of a mixed-valence state of Fe^{2+} and Fe^{3+} ions at the Fermi level and the spatial ordering of positive charges on iron ions because of Coulomb repulsion. Since at $N_{Fe} > N^*$ the concentration of conduction electrons and trivalent iron ions stabilizes, $n_e = N_{Fe}^{3+} = N^*$, the concentration $N_0 = N_{Fe} - N_{Fe}^{3+}$ of lattice-neutral Fe^{2+} ions, which serve as vacant positions for the redistribution of d holes, increases with increasing iron content. At the same time, the degree of spatial ordering of Fe^{3+} ions increases, the rate of conduction-electron scattering decreases, and the mobility increases. A variant of the short-range correlation model was proposed in Ref. 5. This model made it possible to quantitatively describe the dependence of the thermoelectric and thermomagnetic effects on iron content and temperature.[5-9]

However, spatial ordering of Fe^{3+} ions leads not only to a reduction in the rate of electron scattering, but also to changes in the nature of the dependence of the electron relaxation time τ on the energy ε. The analysis of the $\tau(\varepsilon)$ in Ref. 6, which was based on the short-range correlation model developed in Ref. 5, showed that upon passage from the weak-correlation region ($N_0/N_{1+} \ll 1$, where $N_{1+} = N_{Fe}^{3+}$) to the strong-correlation region ($N_0/N_{1+} > 1$) the derivative $d\tau/d\varepsilon$ changes its sign from ''plus'' to ''minus,'' causing a change in the signs of the Nernst–Ettingshausen effects as

the iron content increases. Experimental studies[6,7] have corroborated this important theoretical finding.

Up to now, the focus has been on the properties of HgSe:Fe crystals which are determined primarily by the relaxation of electron momentum on the correlated system of Fe^{3+} ions, the alloy potential, etc.[1-9] Far less attention has been paid to effects associated with the influence of the phonon system on electron transport in HgSe crystals with an admixture of mixed-valence iron. One such effect is phonon drag, which plays an important role in the behavior of thermoelectric and thermomagnetic effects at low temperatures.[6,7,10] In the present paper we show that the increase in the degree of ordering in the correlated system of Fe^{3+} ions with increasing iron content in HgSe:Fe crystals leads not only to a reduction in the rate of electron scattering and a change in the nature of the τ vs ε dependence,[6] but also to a reduction in the rate of phonon scattering by the spatially ordered system of Fe^{3+} ions. We will see that at low temperatures this effect leads to a substantial increase in the thermoelectric power as the iron concentration N_{Fe} increases in the range 5×10^{18} cm$^{-3} < N_{Fe} < 2 \times 10^{19}$ cm^{-3}. Note that the electron mobility exhibits an anomalous increase within the same iron-content range.[1-3] Below we show that this should be expected, since the physical reason for both ''anomalies'' is the spatial ordering of the trivalent Fe^{3+} ions in the mixed-valence system of iron ions. As far as we know, reduction in the rate of phonon scattering by a correlated system of charged centers has not been considered either in gapless HgSe semiconductors doped with transition elements or in other mixed-valence systems.

To observe this effect in experiments, we must ensure that (a) the contribution of phonon drag to the thermoelectric power, $|\alpha_{ph}|$, is larger than the diffusion contribution $|\alpha_e|$ and that (b) Rayleigh scattering is the main mechanism of

TABLE I.

| Sample No.	N_{Fe}, 10^{19} cm^{-3}	n_e, 10^{18} cm^{-3}	μ, 10^4 cm^2/V·s (4.2 K)	N_i, 10^{18} cm^{-3}
1	0.0	2.4	2.25	2.4
2	0.1	3.0	2.8	3
3	0.3	4.0	2.9	4.0
4	0.5	4.8	5.1	5.0
5	1.0	4.7	8.3	10.0
6	2.0	4.9	6.4	20.0
7	5.0	4.81	5.95	50.0
8	40.0	6.2	2.5	400.0

FIG. 1. Calculated (lines) and experimental (points) temperature dependence of the absolute value of the thermoelectric power for HgSe:Fe crystals with different iron content in samples 1 (○), 2 (●), 3 (□), 4 (■), 7 (△), and 8 (▲). The values of N_{Fe} and the phonon scattering parameters are listed in Tables I and II, respectively.

phonon momentum relaxation. Studies have shown[7] that HgSe:Fe crystals meet these requirements in the temperature range 5–15 K. The thermoelectric power was measured in samples of HgSe:Fe with different iron content in the temperature range 10–45 K in Ref. 7. It was found that the temperature dependence of $|\alpha|$ has a characteristic minimum at $T_{\min} = (20–30)$ K and that the phonon-drag contribution $|\alpha_{\mathrm{ph}}|$, whose value is determined by the phonon-electron interaction and the mechanism of phonon momentum relaxation, dominates at $T < T_{\min}$. It was also established that the Rayleigh mechanism of phonon scattering plays an important role in phonon momentum relaxation. However, only four samples with iron concentrations $N_{\mathrm{Fe}} = 0$, $N_{\mathrm{Fe}} = 0.1 \times 10^{19}$ cm^{-3}, $N_{\mathrm{Fe}} = 1 \times 10^{19}$ cm^{-3}, and $N_{\mathrm{Fe}} = 40 \times 10^{19}$ cm^{-3} were studied in Ref. 7, precluding a detailed investigation and detection of the nonmonotonic behavior of $|\alpha(N_{\mathrm{Fe}})|$. Below we report the results of measurements on a larger (than in Ref. 7) set of samples HgSe:Fe and a quantitative analysis of the dependence of the thermoelectric power on temperature and iron content.

2. EXPERIMENTAL RESULTS

We measured the thermoelectric power $\alpha(T)$ in eight HgSe:Fe crystals with iron concentrations ranging from zero to 4×10^{20} cm^{-3} in the temperature range 7.5–60 K. The main sample characteristics (iron content N_{Fe}, electron concentration n_e, and mobility μ) are listed in Table I. The average sample dimensions were $8 \times 2.0 \times 0.8$ mm. The measured temperature difference did not exceed 10% of the average sample temperature. Table I shows that in samples 2 and 3 with $N_{\mathrm{Fe}} < N^*$, where the Fermi level is below the donor level of iron, the electron concentration exceeds the iron concentration. This is related to the presence of charged intrinsic defects, whose concentration N_d in HgSe:Fe crystals is usually $(1–2) \times 10^{18}$ cm^{-3}. In this case the electron concentration n_e is equal to the total concentration of charged centers, $n_e = N_d + N_{1+} = N_i$. When $N_{\mathrm{Fe}} > N^*$, the Fermi level is pinned at the donor level of iron and the electron concentration is independent of N_{Fe}.

Figures 1 and 2 depict the temperature dependence of the absolute value of the thermoelectric power, $|\alpha|$. Clearly, the $|\alpha(T)|$ curves have minima. At $T > T_{\min}$ the thermoelectric power is determined primarily by the electron contribution $|\alpha_e(T)|$, which monotonically decreases with decreasing temperature. At $T < T_{\min}$ the main contribution to the
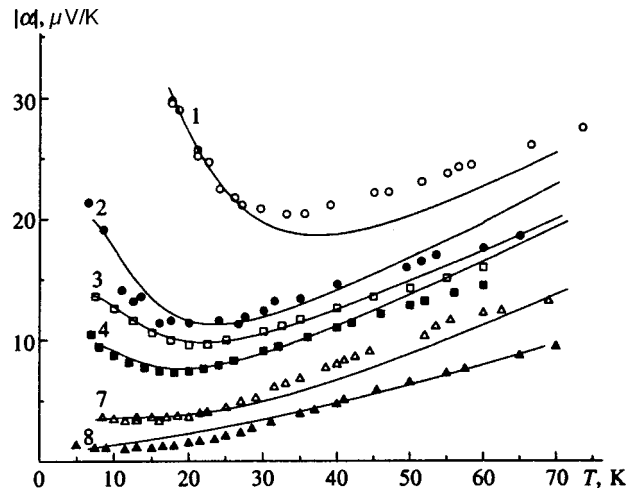
thermoelectric power is provided by phonon drag, and the phonon-drag thermoelectric power $|\alpha_{\mathrm{ph}}(T)|$ is determined by the phonon relaxation mechanism. As the iron content increases, the minima of $|\alpha(T)|$ are shifted toward lower temperatures, and $|\alpha(T)|$ decreases in the entire temperature range for all samples except samples 5 and 6 (Fig. 2), in which $|\alpha(T)|$ increases with decreasing temperature much faster than in sample 4. Hence at $T < 11$ K the values of the thermoelectric power for HgSe:Fe crystals with iron concentrations equal to 1×10^{19} cm^{-3} and 2×10^{19} cm^{-3} become larger than for crystals with $N_{\mathrm{Fe}} = 5 \times 10^{18}$ cm^{-3}, although it would seem that the increase in iron content should lead to an increase in the rate of the Rayleigh scattering of phonons and that thermoelectric power at a fixed temperature should decrease with increasing concentration of the scattering centers, especially since such a dependence is actually observed
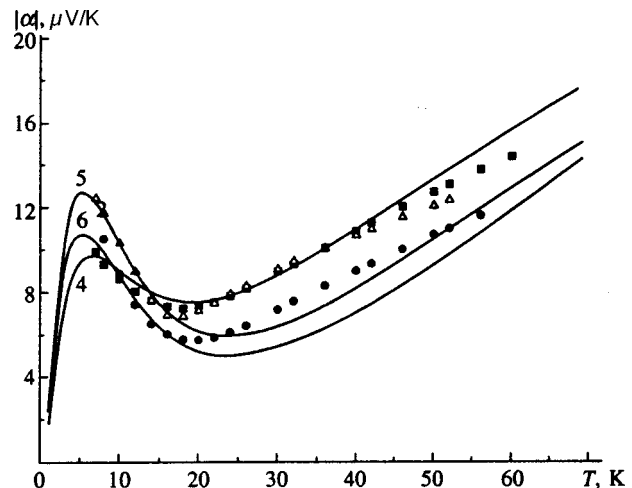

FIG. 2. Calculated (lines) and experimental (points) temperature dependence of the absolute value of the thermoelectric power for samples 4 (■), 5 (△), and 6 (●). The values of N_{Fe} and the phonon scattering parameters are listed in Tables I and II, respectively.
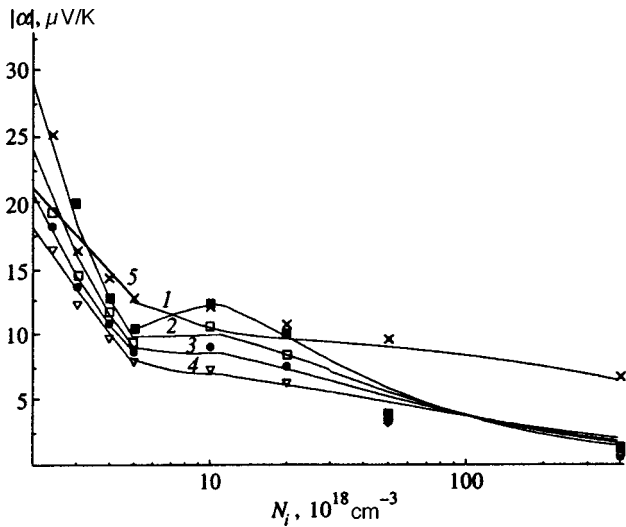
FIG. 3. Calculated (lines) and experimental (points) dependence of the absolute value of the thermoelectric power on iron content at $T = 7.5$ K ($\blacksquare$, curve *1*), 10 K ($\square$, curve *2*), 12 K ($\bullet$, curve *3*), 15 K ($\triangledown$, curve *4*), and 50 K ($\times$, curve *5*). The following average values of the phonon scattering parameters were used: $E_1 = 0.75$, $c_L = 0.1$, $c_H = 2$, $c_{R+} = 2$, and $c_{R0} = 0.15$.

for the other samples: Fig. 1 shows that at $T = 7.5$ K the thermoelectric power decreases by a factor greater than 20 as the iron concentration increases from 1×10^{18} cm^{-3} to 4×10^{20} cm^{-3}, and at $T = 20$ K this factor is greater than 25 for the transition from sample 1 to sample 8. This implies that at $T < 20$ K the Rayleigh mechanism of phonon scattering plays an important role in the relaxation of the momentum of the phonon system and determines the phonon-drag contribution to the thermoelectric power to a large extent.

The unusual dependence of the thermoelectric power $|\alpha|$ on iron content at different temperatures can be seen in Fig. 3. The concentration N_i for samples 1–3 is equal to the concentration of charged centers, while for samples 4–8 it is equal to N_{Fe}. Figure 3 shows that at $T < 12$ K the $|\alpha|$ vs N_i dependence is nonmonotonic: as N_i increases, $|\alpha|$ first decreases to a value $|\alpha_{min}| \sim 10$ μV/K at $N_i = 5 \times 10^{18}$ cm^{-3} and then increases, reaching its maximum $|\alpha_{max}| \approx 12.5$ μV/K at $N_{Fe} \approx (1-2) \times 10^{19}$ cm^{-3}. Thus, the increase in thermoelectric power in this concentration range amounts to about 2 μV/K, which is larger than the experimental error by a factor of four. A further increase in iron content leads to a monotonic decrease in $|\alpha(N_i)|$, which, obviously, is due to the increase in the rate of Rayleigh scattering by lattice-neutral Fe^{2+} ions. The thing is that when $N_{Fe} > N^*$, the Fermi level is pinned at the iron donor level, and as the iron content increases, the concentration of Fe^{3+} ions remains constant with only the concentration of Fe^{2+} ions increasing. As the temperature rises, the maximum on the $|\alpha(N_i)|$ curve disappears, although characteristic discontinuities remain on the curve up to 20 K. At $T = 50$ K the value of the thermoelectric power is determined by the electron contribution, and $|\alpha(N_i)|$ is a monotonically decreasing function.

Note that qualitatively the dependence of the thermoelectric power on iron content at $T < 10$ K is similar to the dependence of the electron mobility μ on N_{Fe} (see Ref. 3). Like $\mu(N_{Fe})$, the thermoelectric power $|\alpha(N_{Fe})|$ first de-

creases (with increasing N_{Fe}) up to the point where $N_{Fe} = 5 \times 10^{18}$ cm^{-3} and then increases, reaching its maximum at $N_{Fe} = (1-2) \times 10^{19}$ cm^{-3}. As the iron content increases still further, the mobility and thermoelectric power monotonically decrease. When $N_{Fe} < 5 \times 10^{18}$ cm^{-3}, the concentrations of trivalent iron ions and conduction electrons increase with N_{Fe}. As a result, the mobility and thermoelectric power decrease for two reasons: because of the increase in the probability of the scattering of electrons and phonons by Fe^{3+} ions, and because of a rise in the Fermi level and the nonparabolicity of the band structure of HgSe:Fe. When $N_{Fe} > 2 \times 10^{19}$ cm^{-3}, the concentration of lattice-neutral Fe^{2+} ions increases with iron content, leading to decreases in both $\mu(N_{Fe})$ and $|\alpha(N_{Fe})|$ due to an increase in the rate of scattering by these ions. The increase in electron mobility in the range of iron concentrations from 5×10^{18} to 2×10^{19} cm^{-3} is due to the decrease in the probability of electron scattering by charged centers as the degree of spatial ordering in the correlated system of Fe^{3+} ions increases.[5,6] Hence we assume that the increase in thermoelectric power in this concentration range is due to a decrease in the probability of phonon scattering by the correlated system of Fe^{3+} ions.

Below we shall analyze quantitatively the dependence of the thermoelectric power on temperature and iron content with allowance for phonon drag. The calculation takes into account the scattering of electrons by the correlated system of Fe^{3+} ions, the alloy potential, and acoustic phonons, as well as the main phonon scattering mechanisms. The phonon scattering by a spatially ordered system of Fe^{3+} ions is treated separately.

3. DIFFUSION COMPONENT OF THE THERMOELECTRIC POWER

As is well known,[10] the thermoelectric power observed in experiments at low temperatures is determined by the sum of the diffusion component $\alpha_e(T)$ and the phonon-drag component $\alpha_{ph}(T)$:

$$\alpha(T) = \alpha_e(T) + \alpha_{ph}(T). \tag{1}$$

In calculating the thermoelectric power one must take into account the nonparabolicity of the conduction band of HgSe crystals within the two-band Kane model with an energy $\varepsilon_g = 0.022$ eV and an effective electron mass at the band bottom $m_n = 0.02 m_0$, where m_0 is the free-electron mass. For a degenerate electron gas (the condition $\varepsilon_F \gg k_B T$ is sure to be met in the experiment), the electronic component of the thermoelectric power can be written as[1)]

$$\alpha_e(T) = -\frac{\pi^2 k_B^2 T}{3e\varepsilon_F}\left(\frac{3}{2}f_{gF} + D\right), \tag{2}$$

where

$$D = \varepsilon_F\left\{\frac{\partial}{\partial\varepsilon}\ln\frac{\tau(\varepsilon)}{m(\varepsilon)}\right\}_{\varepsilon=\varepsilon_F}, \quad m(\varepsilon) = m_n\left(1 + \frac{2\varepsilon}{\varepsilon_g}\right),$$

$$f_{gF} = \frac{\varepsilon_g + 2\varepsilon_F}{\varepsilon_g + \varepsilon_F}.$$

The main mechanisms of electron momentum relaxation in HgSe:Fe crystals at low temperatures are scattering by the correlated system of Fe^{3+} ions, by the alloy potential, by acoustic phonons,[5] and by intrinsic defects, whose concentration is about 1×10^{18} cm^{-3}. If all these mechanisms are taken into account, the expression for D can be written as follows:[7,8]

$$D = \frac{\varepsilon_g + 2\varepsilon_F}{2(\varepsilon_g + \varepsilon_F)} \left[\frac{K_+ + K_a - K_{ph}}{K_b + K_{ph}} \right] - \frac{4\varepsilon_F}{\varepsilon_g + \varepsilon_F},$$

$$K_a = \Lambda \left[2 \frac{N_{1+}}{N_+} \left(\frac{N_0}{N_+} \right)^{1/2} \left(\Phi_{+0} - \frac{\partial \Phi_{+0}}{\partial y} \bigg|_{y=1} \right) - \frac{N_0}{2N_+} \Lambda \right],$$

$$N_+ = N_d + N_{1+}, \tag{3}$$

$$K_b = \frac{N_r}{N_+} \Phi_{BH} + \frac{N_c}{N_+} \Phi_c + \Lambda \left[2 \frac{N_{1+}}{N_+} \left(\frac{N_0}{N_+} \right)^{1/2} \Phi_{+0} + \frac{\Lambda}{2} \frac{N_0}{N_+} \right],$$

$$K_{ph} = \frac{3E_1^2 m_F k_F k_B T}{4\hbar^2 \varepsilon_B(\varepsilon_F) \rho s^2},$$

$$K_+ = \frac{N_c}{N_+} \left(3\Phi_c - \frac{\partial \Phi_c}{\partial y} \bigg|_{y=1} \right) + \frac{N_r}{N_+} \left(3\Phi_{BH} - \frac{\partial \Phi_{BH}}{\partial y} \bigg|_{y=1} \right),$$

$$\Phi_{BH} = \ln(1 + b_s) - (1 + b_s^{-1})^{-1}.$$

Here $y = k/k_F$; E_1 is the deformation potential constant; ρ is the crystal density; s is the speed of sound; N_{1+} and N_0 are the concentrations of charged Fe^{3+} ions and lattice-neutral Fe^{2+} ions; $\Phi_{+0} = 1 - b_s^{-1} \ln(1 + b_s)$, where $b_s = (2k_F r_s y)^2$ (r_s is the Thomas–Fermi screening radius); $\hbar k_F$ is the Fermi momentum; $\varepsilon_B(\varepsilon_F) = m(\varepsilon_F) e^4 / 2\chi \hbar^2$ is the Bohr energy, where χ the dielectric constant; and Λ is the ratio between the electron–neutral-center and electron–charged-center coupling constants. According to the estimates in Ref. 5, $\Lambda = 0.1$, and

$$\Phi_c(k_F) = 2 \int_0^1 \frac{x^3 S(2k_F x)}{(x^2 + b_s^{-1})^2} \, dx, \tag{4}$$

where $S(q)$ is the structure factor, which is the measure of order in the donor system and is determined using the short-range correlation model proposed in Ref. 5. This variant of the short-range correlation model is valid for arbitrary (in magnitude) Fe^{3+}–Fe^{3+} correlations and is based on the following physical assumptions. When $N_{Fe} > N^*$, a mixed-valence state forms at the Fermi level, and it becomes possible for the positive charges on the iron ions (d holes) to redistribute among the crystal-lattice sites occupied by Fe^{2+} ions. As was shown in Ref. 11, the maximum free-energy gain from the ordering of Fe^{3+} ions is attained when the d holes closest to each other move apart. Hence around each Fe^{3+} ion there forms a correlation sphere of radius r_c, within which there are no other Fe^{3+} ions. As N_{Fe} increases, so does the number of vacant positions for redistributing the d holes. Consequently, the correlation-sphere radius and the degree of spatial ordering in the correlated system of Fe^{3+} ions increase, too. This makes it possible to approximate the system of Fe^{3+} ions by a system of hard spheres of diameter $d = r_c$ (Ref. 11). The degree of ordering in such a system is char-

acterized by the packing factor $\eta = \pi d^3 N_{1+}/6 = V_c N_{1+}/8$, which is equal to the ratio of the volume occupied by the hard spheres to the total volume of the system. The integral equation for the pair correlation function of the system of hard spheres can be solved exactly,[12] and the structure factor in (4) can be found without resorting to perturbation-theory techniques. The presence in the crystal of randomly positioned ions, i.e., intrinsic defects with a concentration N_d, reduces the correlation-sphere radius.[13] To determine the dependence of η on N_{Fe}, Tsidil'kovskiĭ *et al.*[13] derived the following equation:

$$\eta = \eta_L \exp\left(-p \frac{\eta}{\eta_L} \frac{N_d}{N_{1+}} \right) \left[1 - \exp\left(-\frac{\eta}{\eta_L} \frac{N_{Fe}}{N_{1+}} \right) \right]. \tag{5}$$

The value $p \approx 0.2$ was found by analyzing the experimental $\mu(N_{Fe} N_d)$ curves, for HgSe:Fe and HgSe:Ga crystals in Ref. 13. For concentrations $N_{Fe} > 2 \times 10^{19}$ cm^{-3} and $N_d \to 0$ the $\eta(N_{Fe})$ curve becomes saturated with η reaching a value $\eta_L = 0.45$, which corresponds to the state of a strongly correlated Coulomb liquid.[5]

In calculating the electron relaxation time in HgSe:Fe crystals with intrinsic defects, one must take into account the mutual effect of the mechanisms of electron scattering by charged centers of two types[14] and the Coulomb correlations between d holes and intrinsic defects.[13] Since intrinsic defects are immobile and d holes can migrate between crystal-lattice sites occupied by Fe^{2+} ions, correlation spheres also form around the fraction δN_d of the intrinsic defects located at distances larger than r_c. This fraction of intrinsic defects scatters electrons in the same way as the correlated system of Fe^{3+} ions does. The remaining intrinsic defects, $N_r = (1 - \delta)N_d$, which are located at distances smaller than r_c, scatter electrons just like a disordered collection of charged centers would. Tsidil'kovskiĭ *et al.*[13] showed that the parameter δ can be expressed in terms of the packing factor as follows:

$$\delta = \exp\left(-p \frac{\eta}{\eta_L} \frac{N_d}{N_{1+}} \right). \tag{6}$$

Such an approach made it possible to explain the dependence of the electron mobility on the content of iron and gallium in HgSe:Fe and HgSe:Ga crystals.[13] Here it is used to calculate the diffusion components of thermoelectric power in HgSe:Fe crystals with intrinsic defects.

As the temperature rises, the migration of d holes between the Fe^{2+} and Fe^{3+} ions increases, and the system of Fe^{3+} ions becomes increasingly chaotic, while the degree of spatial ordering and the correlation-sphere radius decrease. The effect of temperature on the degree of spatial ordering of the correlated system of Fe^{3+} ions can be taken into account in the soft-sphere approximation.[9] According to Ref. 9, the variation of the soft-sphere diameter $d(T)$ is given by the following expression:

$$d(T) = d_0 \left[1 - \frac{\Delta r(T)}{d_0} \right], \quad \frac{\Delta r(T)}{d_0} = \beta_1 [\sqrt{1 + \beta_2 T} - 1]. \tag{7}$$

Here d_0 is the hard-sphere diameter at $T = 0$, and β_1 and β_2 are parameters which depend in the general case on the con-

centrations N_+ and N_0, the packing factor, and the screening radius. These parameter were obtained in Ref. 9 for different values of N_{Fe} by comparing theoretical and experimental $\mu(T)$ curves, where μ is the electron mobility.

Thus, Eqs. (1)–(7) make it possible to analyze the dependence of the diffusive thermoelectric power on temperature and on iron content.

4. PHONON COMPONENT OF THE THERMOELECTRIC POWER

The contribution of phonon drag to the thermoelectric power can be expressed as follows:[7,15]

$$\alpha_{ph}(T) = -\left(\frac{k_B}{T}\right) A_{ph}(T). \tag{8}$$

In HgSe crystals, electrons interact not only with longitudinal phonons, but also with transverse phonons. As in Ref. 16, to simplify matters we introduce the mean speed of sound for acoustic phonons, $s = (1/3)(1/s_l + 2/s_t)^{-1}$, where s_l and s_t are the speeds of the longitudinal and transverse phonons, respectively. Then the phonon drag A_{ph} can be written as follows:[7]

$$A_{ph} = \frac{6m(\varepsilon)s^2}{k_B} \sum_{\mathbf{k'},\mathbf{q}} \frac{W(q)}{\nu_{ph}} \frac{dN_q^0}{dT}$$
$$\times \left(1 - \frac{\mathbf{k}\cdot\mathbf{k'}}{k^2}\right) \delta(\varepsilon_{\mathbf{k}} - \varepsilon_{\mathbf{k'}}) \delta_{\mathbf{k'}-\mathbf{k},\mathbf{q}}, \tag{9}$$

where $W(q) = \pi E_1^2 q/\rho s$, E_1 is the deformation potential constant, ρ is the crystal density, ν_{ph} is the long-wavelength phonon relaxation frequency, and $N_q^0 = [\exp(\hbar\omega_q/k_BT) - 1]^{-1}$ is the Planck distribution function.

The form of the function $A_{ph}(\varepsilon_F, T, N_{Fe}, N_+)$ depends on the electron-phonon coupling mechanism and the phonon relaxation frequency ν_{ph}. The calculations in Refs. 7 and 16 show that the main contribution to the long-wavelength phonon momentum relaxation in HgSe:Fe crystals at low temperatures is provided by the Herring and Rayleigh mechanisms and the scattering of phonons by electrons and sample boundaries:

$$\nu_{ph} = \nu_{phH} + \nu_{phR} + \nu_{phe} + \nu_{phL}. \tag{10}$$

Here $\nu_{phH} = \Lambda_H x^2$, where $x = q/q_T$ and $\Lambda_H = B_2 T^3 s^2 q_T^2$ with $q_T = k_B T/\hbar s$ the phonon thermal momentum and $B_2 = (3 \pm 0.8) \times 10^{-22}$ s/K^3, and

$$\nu_{phe} = \nu_{phe}^0 x, \quad \nu_{phe}^0 = \frac{1}{2\pi} \frac{E_1^2 m(\varepsilon)}{\rho\hbar^3} q_T,$$

$$\nu_{phL} = \frac{c_L s}{(L_1 L_2)^{1/2}} = \nu_{phL}^0 c_L, \quad c_L = \frac{f}{2-f},$$

where f is the fraction of phonons scattered diffusively by sample boundaries, and $L_1 L_2$ is the cross-sectional area of the sample. For phonon scattering by charged Fe^{3+} ions and lattice-neutral Fe^{2+} ions (the Rayleigh mechanism) we have

$$\nu_{phR} = \nu_{phR+} + \nu_{phR0}. \tag{11}$$

An analysis of the experimental data (see Sec. 2) showed that this mechanism is important when one wishes to determine the phonon component of the thermoelectric power in HgSe crystals. As in Ref. 16, in studying phonon scattering by randomly distributed neutral and charged centers we use the point-defect approximation. Then

$$\nu_{phR0} = \Lambda_{R0} N_0 x^4, \quad \Lambda_{R0} = A_0 s^4 q_T^4, \tag{12}$$

$$\nu_{phR+} = \Lambda_{R+} N_+ x^4, \quad \Lambda_{R+} = A_+ s^4 q_T^4. \tag{13}$$

By point defects we usually mean substitutional impurities, isotopes, vacancies, and interstitial atoms, whose perturbing effect is usually limited to a unit cell. In relation to long-wavelength acoustic phonons, a point defect behaves like a small region in the crystal (small compared to the phonon wavelength) with a different density and different elastic properties. Obviously, replacing the substitutional Fe^{2+} impurity ions by point defects is a good approximation. Since we do not know how the force constants change in response to such substitution, A_0 is a fitting parameter of the theory. By comparing the calculated dependence of the thermal conductivity of HgSe crystals with different donor impurity concentrations and the experimental data we arrive at the following estimate for the constant A_+: $A_+ = 12.1 \times 10^{-40}$ cm^4/s. According to Ref. 16, the values of A_+ for the different types of impurities may differ by a factor of 10 due to the difference in the phonon scattering cross sections.

After performing some simple manipulations on (8) and (9), we get

$$\alpha_{ph} = -\frac{k_B}{2\pi^2 T} \frac{q_T^3}{n_e} \nu_{phe}^0 \int_0^{x_{2k_F}} \frac{x^5 \exp(x)dx}{[\exp(x)-1]^2 \nu_{ph}(x)}, \tag{14}$$

where $x_{phF} = 2k_F/q_T$, and

$$\nu_{ph}(x) = c_L \nu_{phL}^0 + \nu_{phe}^0 x + c_H \Lambda_H x^2$$
$$+ c_{R+}(N_+ + c_{R0}N_0)\Lambda_{R+} x^4. \tag{14a}$$

In (14a) we have explicitly specified the fitting parameters c_H, c_{R+}, and c_{R0}, which, as we shall shortly see, characterize the difference between the properties of a HgSe:Fe crystal and those of a HgSe crystal with intrinsic defects.[16] For the values of the phonon scattering parameters obtained in Ref. 16 we have $c_H = c_{R+} = 1$, and the constant c_{R0} shows the extent to which the probability of phonon scattering by neutral Fe^{2+} ions is lower than that of scattering by charged Fe^{3+} ions.

Thus, Eqs. (8)–(14) make it possible to analyze the dependence of the phonon component of the thermoelectric power for randomly distributed scattering centers.

5. PHONON SCATTERING BY THE CORRELATED SYSTEM OF Fe^{3+} IONS

As in Ref. 16, we use the point-defect model to examine the phonon scattering by the correlated system of Fe^{3+} ions in HgSe:Fe crystals. We assume that the entire disturbance created by an Fe^{3+} ion replacing a Hg^{2+} ion at a lattice site is localized within a unit cell and amounts, as in the case of Fe^{2+} ions, to changes in elastic constants and mass. Although the values of $\Delta M/M$ for Fe^{3+} and Fe^{2+} ions are the

same, the constant A_+ characterizing the interaction between a phonon and an Fe^{3+} ion must differ from the constant A_0 characterizing the interaction between a phonon and a neutral center. The thing is that the potential of Fe^{3+} ions in HgSe:Fe is screened at distances $r_s \approx 5 \times 10^{-7}$ cm, while the average distance between the charged centers $R_+ \sim (N^*)^{-1/3}$ is approximately 6×10^{-7} cm. Hence the disturbance introduced by an Fe^{3+} ion (the changes in elastic constants and local density) encompasses many unit cells because $r_s \sim 10a_0$, where a_0 is the lattice constant. Since the Rayleigh scattering cross section is proportional to the square of the volume of the disturbed region,[17] the probability of a phonon being scattered by a Fe^{3+} ion is higher than the probability for scattering by a lattice-neutral Fe^{2+} ion.

In calculating the phonon relaxation frequency on the correlated system of Fe^{3+} ions, we allow for spatial ordering of the Fe^{3+} ions by employing the structure factor, as was done in Ref. 6 in calculations of the electron relaxation time. Then for $\nu_{ph\,Rc}$ we obtain

$$\nu_{ph\,Rc} = 2\pi N_+ s \int_0^\pi (1 - \cos\theta)$$

$$\times \sigma(q, \theta) S(q(1 - \cos\theta)) \sin\theta\, d\theta. \tag{15}$$

Here $S(q)$ is the structure factor, and $\sigma(q, \theta)$ is the cross section for the scattering of a phonon with a wave vector q by a Fe^{3+} ion, which depends on the angle θ between the directions of the incident and scattered phonons in the general case. After we average over the polarization vectors of the phonons, the dependence on θ disappears in the scattering cross section $\sigma(q, \theta)$ calculated in the Born approximation.[18,19] Here $\sigma(q, \theta) = \sigma(q)$, so that actually only S scattering has been taken into account. In this case, according to Refs. 18–20, for $\sigma(q)$ we have

$$\sigma(q) = \frac{V_0^2 \omega_q^4}{4\pi s^4} S^2, \quad S^2 = S_1^2 + (S_2 + S_3)^2, \tag{16}$$

where V_0 is the volume of the region in the crystal disturbed by the defect, and S_1, S_2, and S_3 characterize the contributions of the changes in the unit-cell mass and the force constants and of the deformation of the lattice to the phonon scattering cross section (see Refs. 18–20 for more details). Since for HgSe:Fe crystals only the constant S_j is known,

$$S_1 = \frac{\Delta M}{M} = \frac{M_{Fe} - M_{Hg}}{M_{Hg} + M_{Se}},$$

we write the expression for $\sigma(q)$ as

$$\sigma(q) = c_{R+} A_+ \left(\frac{\omega_q}{s}\right)^4, \tag{17}$$

where A_+ is the value of the parameter for HgSe crystals with intrinsic defects found in Ref. 16, and c_{R+} is a parameter of the theory, whose value must be determined from the experimental data. As Eqs. (15)–(17) imply, the latter parameter indicates the extent to which the probability of phonon scattering by Fe^{3+} ions differs from that of scattering by intrinsic defects in HgSe crystals.[16]

At low temperatures $(T \sim 10$ K$)$, long-wavelength phonons $(q \sim q_T \approx 6.71 \times 10^6$ cm$^{-1})$ provide the main contribution to α_{ph}. Hence, to simplify the ensuing calculations, we expand the structure factor S_q in a power series in q up to fourth-order terms inclusively. Then $\nu_{ph\,Rc}(q)$ can be calculated analytically:

$$\nu_{ph\,Rc} \approx \Lambda_{R+} N_+ x^4 \{S(0) + 2S_{(1)}(u_T x)^2 + \tfrac{16}{3} S_{(2)}(u_T x)^4\}$$

$$\equiv \Lambda_{R+} N_+ x^4 \tilde{S}(x), \tag{18}$$

where

$$u_T = dq_T, \quad S(0) = \frac{(1 - \eta)^4}{(1 + 2\eta)^2},$$

$$S_{(1)} = S^2(0)\, 4\eta \left(\frac{\beta}{5} + \frac{\gamma}{6} + \frac{\delta}{8}\right),$$

$$S_{(2)} = S^2(0) \left\{ S(0) \left[4\eta \left(\frac{\beta}{5} + \frac{\gamma}{6} + \frac{\delta}{8}\right) \right]^2 \right.$$

$$\left. - \frac{\eta}{5} \left(\frac{\beta}{7} + \frac{\gamma}{8} + \frac{\delta}{10}\right) \right\},$$

$$\beta = \frac{(1 + 2\eta)^2}{(1 - \eta)^4}, \quad \gamma = -\frac{6\eta(1 + 0.5\eta)^2}{(1 - \eta)^4}, \quad \delta = 0.5\eta\beta.$$

Figure 4a depicts the dependence of $\nu_{ph\,Rc}/\nu^*_{ph\,R+}$, where $\nu^*_{ph\,R+} = \Lambda_{R+} N^* x^4$, on the reduced wave vector x for various iron concentrations. We see that this ratio is weakly dependent on the wave vector when $x \lesssim x_{2k_F}$ and that the values of $\nu_{ph\,Rc}$ are smaller for $N_{Fe} = (1-2) \times 10^{19}$ cm^{-3} than for $N_{Fe} = 5 \times 10^{18}$ cm^{-3}. Note that the region of long wave vectors $(x > 1)$ is cut off exponentially due to the phonon distribution function [see Eq. (14)]. Hence, when the Rayleigh scattering of phonons provides the main contribution to phonon momentum relaxation, α_{ph} increases with the degree of spatial ordering of the correlated system of Fe^{3+} ions as long as the scattering by neutral centers is relatively weak. Figure 4b depicts the dependence of $\tilde{\nu}_{ph\,R} = \langle \nu_{ph\,R} \rangle / \langle \nu^*_{ph\,R+} \rangle$ on iron content at $T = 10$ K, and the angle brackets denote averaging:

$$\langle \nu_{ph\,R} \rangle = \frac{1}{x_{2k_F}} \int_0^{x_{2k_F}} \nu_{ph\,R}(x)\, dx. \tag{19}$$

Figure 4b shows that we should expect an increase in thermoelectric power as N_{Fe} increases from N^* to $N_{Fe} \simeq 1 \times 10^{19}$ cm^{-3} and a decrease in thermoelectric power for $N_{Fe} > 1 \times 10^{19}$ cm^{-3}, although the value of $|a_{ph}|$ is somewhat larger at $N_{Fe} \approx 2 \times 10^{19}$ cm^{-3} than at $N_{Fe} \approx 5 \times 10^{18}$ cm^{-3}.

Thus, the expression for the relaxation frequency upon phonon scattering in HgSe:Fe crystals with allowance for spatial ordering of the correlated system of Fe^{3+} ions can be written as follows:

$$\nu_{ph}(x) = c_L \nu^0_{ph\,L} + \nu^0_{ph\,e} x + c_H \Lambda_H x^2$$

$$+ c_{R+}(N_+ \tilde{S}(x) + c_{R0} N_0) \Lambda_{R+} x^4. \tag{20}$$

Equations (14), (19), and (20) make it possible to calculate the dependence of the phonon component of the thermoelec-

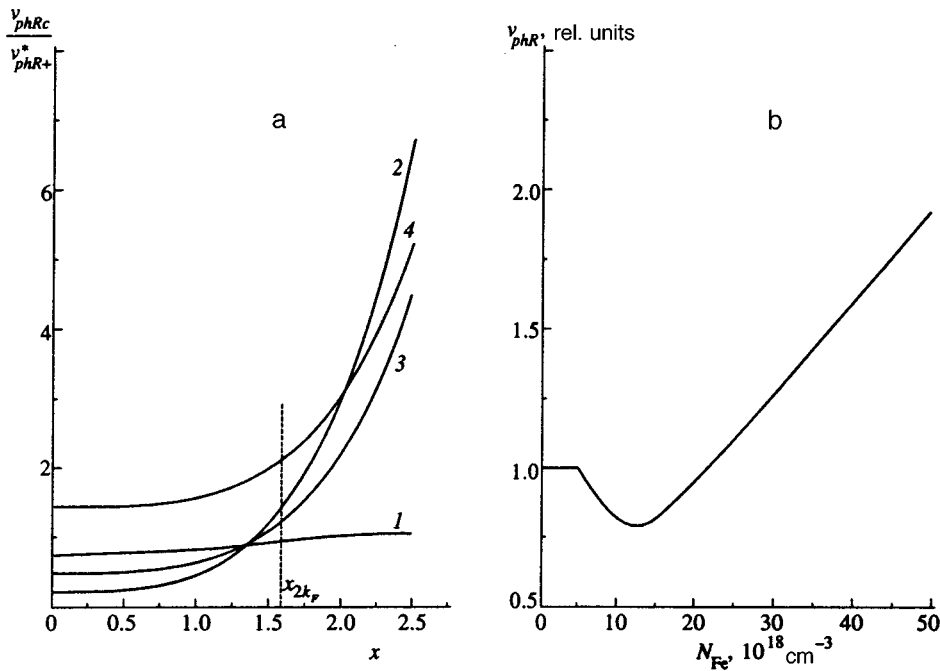FIG. 4. (a) Dependence of the ratio $\nu_{\text{ph}\,Rc}/\nu^{*}_{\text{ph}\,R+}$ on the reduced wave vector x at $T = 10$ K for different values of the iron concentration N_{Fe}, 10^{18} cm^{-3}: 5 (curve *1*), 10 (curve *2*), 20 (curve *3*), and 50 (curve *4*). (b) Dependence of the ratio $\tilde{\nu}_{\text{ph}\,R}$ on iron content at $T = 10$ K. The values of the parameters are $c_{R+} = 2$ and $c_{R0} = 0.15$.

tric power on temperature and iron content for HgSe:Fe crystals containing a mixed-valence Fe^{2+}–Fe^{3+} impurity.

6. DISCUSSION

In calculating the thermoelectric power we used the following values for the parameters: $m(\varepsilon_F) = 0.07m_0$ (m_0 is the free-electron mass), $s_l = 3 \times 10^5$ cm/s, $s_t = 1.65 \times 10^5$ cm/s, $s = 1.95 \times 10^5$ cm/s, and $\chi = 25$. The results of calculating $|\alpha(T)|$ are depicted by solid curves in Figs. 1 (samples 1–4, 7, and 8) and 2 (samples 4–6). Both figures show that the calculated plots of $|\alpha(T)|$ are in quantitative agreement with the experimental data. At $T < T_{\min}$, where the phonon component of the thermoelectric power is predominant, i.e., $|\alpha(T)| \approx |\alpha_{\text{ph}}(T)|$, our results fit the experimental data much better. This indicates that we have correctly allowed for the effect of spatial ordering in the correlated system of Fe^{3+} ions and the main phonon scattering mechanisms (see Fig. 2). Figure 2 depicts calculated plots of $|\alpha(T)|$ up to 1 K. The greatest reduction in the rate of phonon scattering by the correlated system of Fe^{3+} ions should be observed near the peak in the phonon thermoelectric power at $T_{\max} \approx 5$–6 K. At lower temperatures, phonon scattering by the sample boundaries is predominant and the effect is much weaker. As the temperature rises, for $T > T_{\max}$, the contribution of the phonon-phonon relaxation mechanism increases greatly, leading to a further reduction in the rate of phonon scattering by the correlated system of Fe^{3+} ions. Note that when $N_{\text{Fe}} > 1 \times 10^{19}$ cm^{-3}, the increase in the probability of Rayleigh phonon scattering with iron concentration leads to a drop in $|\alpha(T_{\max})|$, and for $N_{\text{Fe}} \geqslant 5 \times 10^{19}$ cm^{-3} no phonon peaks in the thermoelectric power are present.

At $T > T_{\min}$, where the main contribution to the thermoelectric power is provided by the diffusion component, the discrepancy between the calculated $|\alpha(T)|$ curves and the experimental data is much more evident. We believe this is due to the effect of intrinsic defects, whose concentration

varies from sample to sample, as well as to the approximate allowance for the nonparabolicity of the band structure of HgSe:Fe crystals within the two-band Kane model. Actually, the nonparabolicity of the band structure must be taken into account not only through the dependence of the effective electron mass $m(\varepsilon)$ on the energy ε, but also through the Bloch wave-function amplitudes characteristic of the Γ_8 band.[21,22] Such calculations, however, are beyond the scope of the present paper and require special studies.

Table II lists the values of the parameters characterizing phonon momentum relaxation. We see that for sample 1 with $N_{\text{Fe}} = 0$ the value of the parameter A_+ for Rayleigh phonon scattering coincides with the one obtained in Ref. 16 ($c_{R+} = 1$). For HgSe:Fe crystals we have $c_{R+} = 2$, i.e., the cross section for phonon scattering by Fe^{3+} ions is larger by a factor of $\sqrt{2}$ than the cross section for phonon scattering by intrinsic defects in HgSe crystals.[16] A comparison of c_{R+} and c_{R0} shows that the phonon scattering by lattice-neutral Fe^{2+} ions is weaker by a factor of nearly ten than the scattering by Fe^{3+} ions. This is because the probability of Rayleigh phonon scattering is proportional to the square of the volume of the disturbed region in the crystal lattice, which is much larger for a charged impurity than for a neutral impu-

TABLE II.

Sample No.	N_{Fe}, 10^{19} cm^{-3}	n_e, 10^{18} cm^{-3}	N_d, 10^{18} cm^{-3}	E_1, eV	c_L	c_H	c_R	c_{R0}
1	0.0	2.4	2.4	0.8	0.5	0.5	1	–
2	0.1	3.0	2.0	0.73	0.2	2	2	–
3	0.3	4.0	1.0	0.73	0.2	2	2	–
4	0.5	4.8	0.7	0.68	0.15	2.5	2	0.15
5	1.0	4.7	0.7	0.75	0.1	2	2	0.15
6	2.0	4.9	0.7	0.76	0.1	2	2	0.15
7	5.0	4.81	1	0.62	0.9	2	2	0.15
8	40	6.2	1	0.62	0.9	2	2	0.15

rity. The disturbing effect of a neutral impurity is usually localized within a unit cell, while the volume disturbed by a charged impurity in semiconductors may encompass a large number of unit cells. Oskot-skiĭ and Smirnov[23] also found that neutral impurities in semiconductors with a large dielectric constant scatter impurities with a much lower rate than do electrically charged impurities. Note that the values of the parameters c_{R+} and c_{R0} obtained by fitting the $|\alpha(T)|$ curves remain constant for all HgSe:Fe crystals, regardless of iron content. The values of the deformation potential constant E_1 that we obtained for samples with different iron content differ by $\pm 10\%$ from the value $E_1 \approx 0.7$ eV used in Ref. 16. Note that E_1 tends to decrease with increasing iron content. This is to be expected, since iron impurities, which have smaller radii than do Hg^{2+} ions, stabilize the HgSe crystal lattice,[3] which probably leads to changes in the elastic properties of HgSe:Fe crystals and, accordingly, to changes in the phonon spectrum. We believe this is related to a change in the value of the constant c_H as we go from HgSe crystals to HgSe:Fe crystals with low ($N_{Fe} \leqslant 5 \times 10^{18}$ cm^{-3}) and high ($N_{Fe} \geqslant 1 \times 10^{19}$ cm^{-3}) iron concentrations. The spread of values of the parameter c_L may be due to different degrees of diffusivity of phonon scattering by the boundaries of the samples investigated. Determining this parameter more precisely requires measuring the thermoelectric power at temperatures below T_{max} (see Fig. 2) and calculating the thermoelectric power with separation of the contributions of the transverse and longitudinal phonons.

The solid curves in Fig. 3 depicts the calculated dependence of $|\alpha|$ on N_{Fe} at fixed temperatures. We see that the results of the calculations of $|\alpha(N_{Fe})|$ for $T \leqslant 10$ K are in good agreement with the experimental data. However, as the temperature rises, the role of the diffusion component of the thermoelectric power increases, and the deviation of the theoretical curves from the experimental data increases. Note that the sharp drop in thermoelectric power for samples with $N_{Fe} < 5 \times 10^{18}$ cm^{-3} is due to an increase in the rate of Rayleigh scattering by randomly distributed charged centers, $N_i = N_{1+} + N_d$. The increase in thermoelectric power in the range of iron concentrations from 5×10^{18} cm^{-3} to 1×10^{19} cm^{-3} at $T = 7.5$ K amounts to about 20% of $|\alpha|$. According to our theoretical estimates, the maximum increase in thermoelectric power can be observed at temperatures corresponding to the phonon-peak temperature $T_{max} \approx 5$–6 K and may reach 40% of $|\alpha|$ at $N_{Fe} = 5 \times 10^{18}$ cm^{-3}. The slow decrease in the value of the thermoelectric power with increasing iron concentration near $N_{Fe} = 2 \times 10^{19}$ cm^{-3} is due to the increase in the probability of the Rayleigh scattering of phonons by lattice-neutral Fe^{2+} ions.

Figure 5 depicts calculated curves representing the dependence of the diffusion and phonon components of the thermoelectric power on N_{Fe} at different temperatures. At $N_{Fe} > N^*$, as the iron content increases and hence the degree of spatial ordering in the correlated system of Fe^{3+} ions increases, the value of $|\alpha_e(N_i)|$ decreases, while $|\alpha_{ph}(N_i)|$ varies nonmonotonically, and at $T = 12$ K the reduction in the rate of phonon scattering by the correlated system of Fe^{3+} ions clearly manifests itself on the $|\alpha_{ph}(N_i)|$ curve. However, $|\alpha_{ph}(N_i)|$ increases slower than the diffusion compo-
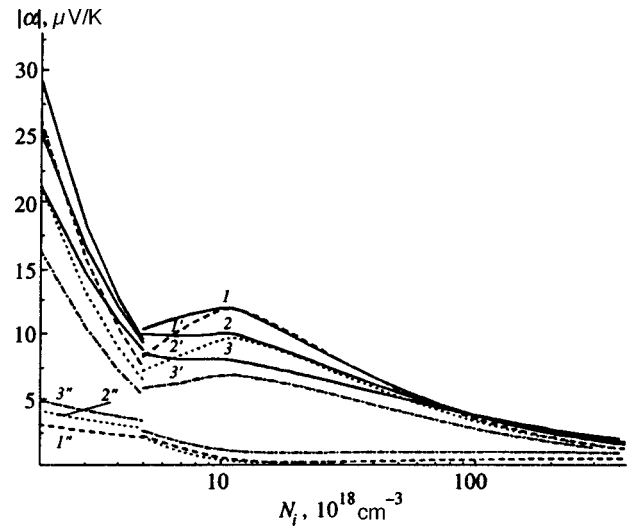

FIG. 5. Absolute values of the phonon (curves *1′*–*3′*), diffusive (curves *1″*–*3″*), and total (curves *1*–*3*) thermoelectric power as functions of iron content at $T = 7.5$ K (curves *1*, *1′*, and *1″*), 10 K (curves *2*, *2′*, and *2″*), and 12 K (curves *3*, *3′*, and *3″*). The values of the phonon scattering parameters are the same as in Fig. 3.

nent $|\alpha_e(N_i)|$ decreases, so that the total thermoelectric power at $T \geqslant 12$ K is a decreasing function of iron content.

We conclude this section by discussing of the approximations adopted in our calculations of the phonon-drag contribution. First, we used the approximation of the mean speed of sound for acoustic phonons. Second, in discussing the Rayleigh scattering of phonons by Fe^{3+} ions we used the point-impurity approximation, although the stringent inequality $r_s/\lambda \ll 1$, where λ_T is the wavelength of a thermal phonon, is not satisfied. It would have been more appropriate to study the scattering of phonons by lattice deformations caused by the random distribution of the charged centers and by the variation of the degree of lattice deformation as the Fe^{3+} ions become spatially ordered. As shown in Ref. 20, when $|(r_i - r_0)/r_0| \geqslant 0.1$ (r_i and r_0 are the ionic radii of the impurity and matrix atoms, respectively), the rate of phonon scattering by lattice deformations is higher than the rate of phonon scattering by local variations in mass and force constants. Note that the ratio for Fe^{2+} and Hg^{2+} ions is about 0.3. Unfortunately, for semiconductors this aspect has yet to be developed theoretically (see Refs. 19 and 20). Hence for semiconductors this phonon scattering mechanism requires further study.

Note that the reduction in the rate of phonon scattering by the correlated system of Fe^{3+} ions due to an increase in the degree of spatial ordering in the system of trivalent iron ions may lead not only to an ''anomalous'' increase in the thermoelectric power $|\alpha(N_{Fe})|$, but also to a substantial increase in the lattice thermal conductivity of HgSe:Fe crystals at low temperatures in the range of iron concentrations from 5×10^{18} cm^{-3} to 1×10^{19} cm^{-3}. Research into the thermal conductivity of HgSe:Fe crystals could yield additional information about the phonon scattering mechanism and about the effect of the spatial ordering of charged centers on the Rayleigh phonon scattering mechanism in mixed-valence systems.

Kuleev *et al.*

7. CONCLUSION

We have interpreted the experimentally established, unusual dependence of the thermoelectric power on iron content and temperature in HgSe:Fe crystals at low temperatures. We have shown that the nonmonotonic nature of the dependence of thermoelectric power on iron content, i.e., the increase in thermoelectric power in the range of iron concentrations from 5×10^{18} cm^{-3} to $(1-2) \times 10^{19}$ cm^{-3}, is due to a new effect, a reduction in the rate of phonon scattering by the spatially correlated system of Fe^{3+} ions. The good agreement between the curves representing the dependence of the thermoelectric power on iron content and temperature and the experimental data suggests that we have correctly taken into account the effect of the spatial ordering of the trivalent iron ions on the scattering of phonons and electrons, as well as the main mechanisms of momentum relaxation in the electron-phonon system.

*)E-mail: kuleev@imp.uran.ru

[1])The factor f_{gF}, whose value is about 1.5 for $\varepsilon_F = 210$ meV, was erroneously omitted in Eq. (4) for $\alpha_e(T)$ in Ref. 7.

[1] F. S. Pool, J. Kossut, U. Debska, and R. Reinfenberger, Phys. Rev. B **35**, 3900 (1987).

[2] Z. Wilamowski, K. Swiatek, T. Dietl, and J. Kossut, Solid State Commun. **74**, 833 (1990); Z. Wilamowski, Acta Phys. Pol. A **77**, 133 (1990).

[3] I. M. Tsidil'kovskiĭ, Usp. Phys. Nauk **162** (2), 63 (1992) [Sov. Phys. Usp. **35 (2)**, 85 (1992)].

[4] I. M. Tsidilkovskii and I. G. Kuleyev, Semicond. Sci. Technol. **11**, 625 (1996).

[5] I. G. Kuleev, I. I. Lyapilin, and I. M. Tsidil'kovskiĭ, Zh. Éksp. Teor. Fiz. **102**, 1652 (1992) [Sov. Phys. JETP **75**, 893 (1992)].

[6] I. G. Kuleev, I. I. Lyapilin, A. T. Lonchakov, and I. M. Tsidil'kovskiĭ, Zh. Éksp. Teor. Fiz. **103**, 1447 (1993) [JETP **76**, 707 (1993)].

[7] I. G. Kuleev, I. I. Lyapilin, A. T. Lonchakov, and, I. M. Tsidil'kovskiĭ, Fiz. Tekh. Poluprovodn. **28**, 937 (1994) [Semiconductors **28**, 544 (1994)].

[8] I. G. Kuleev, I. I. Lyapilin, A. T. Lonchakov, and I. M. Tsidil'kovskiĭ, Zh. Éksp. Teor. Fiz. **106**, 1205 (1994) [JETP **79**, 653 (1994)].

[9] I. G. Kuleev, I. I. Lyapilin, and I. M. Tsidil'kovskiĭ, Fiz. Tverd. Tela (St. Petersburg) **37**, 2360 (1995) [Phys. Solid State **37**, 1291 (1995)].

[10] I. M. Tsidil'kovskiĭ, *Thermomagnetic Effects in Semiconductors*, Academic Press, New York (1963).

[11] I. G. Kuleev, Fiz. Tverd. Tela (St. Petersburg) **39**, 250 (1997) [Phys. Solid State **39**, 219 (1997)].

[12] J. M. Ziman, *Models of Disorder*, Cambridge Univ. Press, Cambridge (1979).

[13] I. M. Tsidil'kovskiĭ, I. G. Kuleev, N. K. Lerinman, A. T. Lonchakov, L. D. Sabirzyanova, and S. Yu. Paranchich, Fiz. Tekh. Poluprovodn. **30**, 2113 (1996) [Semiconductors **30**, 1101 (1996)].

[14] I. M. Tsidil'kovskiĭ, I. G. Kuleev, and G. L. Shtrapenin, Dokl. Ross. Akad. Nauk **347**, 4472 (1996) [Phys. Dokl. **41** (4), 145 (1996)].

[15] B. M. Askerov, *Electron Transport Phenomena in Semiconductors* [in Russian], Nauka, Moscow (1985).

[16] C. R. Whitset, D. A. Nelson, J. G. Broerman, and R. Paxhia, Phys. Rev. B **7**, 4625 (1973).

[17] J. W. S. Rayleigh, *The Theory of Sound*, Dover, New York (1945).

[18] P. G. Klemens, Proc. Phys. Soc., London **68**, 1113 (1995).

[19] A. Griffin and J. Carruthers, Phys. Rev. **131**, 1976 (1963).

[20] B. M. Mogilevskiĭ and A. F. Chudnovskiĭ, *Thermal Conductivity of Semiconductors* [in Russian], Nauka, Moscow (1972).

[21] T. Dietl and W. Szymanska, J. Phys. Chem. Solids **39**, 1041 (1978).

[22] W. Szymanska, P. Boguslawki, and W. Zavadzki, Phys. Status Solidi B **65**, 641 (1974).

[23] V. S. Oskotskiĭ and I. A. Smirnov, *Defects in Crystals and Thermal Conductivity* [in Russian], Nauka, Leningrad (1972).

Translated by Eugene Yankovsky
Edited by P. Shelnitz

Phase transitions in self-assembled monolayers of alkanethiols containing the polar group

Y. V. Sukhinin*)

L. V. Kirensky Institute of Physics Russian Academy of Sciences, 660036 Krasnoyarsk, Russia
(Submitted 8 August 1997)
Zh. Éksp. Teor. Fiz. **114**, 208–221 (July 1998)

Equilibrium states of the system of self-assembled monolayers (SAMs) of n-alkanethiol molecules $HS(CH_2)_{n-1}(X)$ with polar group X chemiabsorbed on the Au(111) crystal surface are considered. The couplings between the atoms (C, H) of the n-alkanethiols are approximated by the Lennard–Jones potential. The couplings between the n-alkanethiols and the crystal surface are approximated by the 12-3 potential. The interactions of polar groups and the self-images with the metal substrate are taken into consideration. The phase-transition temperatures, the structural order and equilibrium tilt, and the twist and azimuthal angles of the macromolecules, which depend on the dipole moments, are found. © *1998 American Institute of Physics.* [S1063-7761(98)01507-8]

1. INTRODUCTION

The self-assembled monolayers (SAMs) are a comparatively new type of organic monolayers,[1–3] which are formed by spontaneous chemisorption of long-chain molecules from a solution to many different solid substrates (e.g., Au, Ag, Cu, Al, GaAs, and Si). The self-assembled monolayers are presently the focus of considerable attention for technological and fundamental reasons. They have potential applications in such areas as corrosion prevention, wear protection, sensing devices, and the formation of well-defined microstructures.[1,4,5] They also present an excellent opportunity for the study of two-dimensional, condensed, organic solids at the microscopic level. Chemisorption of the thiol headgroup to the surface results in long-range translational and orientational lattice structures. The stable monolayers have been studied extensively by transmission electron spectroscopy,[6–9] optical ellipsometry,[10–12] infrared spectroscopy,[13,14] electrochemistry,[15,16] and helium diffraction.[17,18] These monolayers form at a fixed surface density, which remains nearly constant with changing temperature. This fact simplifies the study of rotational and conformational states of SAMs.

The most thoroughly studied and robust SAM system is $CH_3(CH_2)_{n-1}SH$ adsorbed on the Au(111) crystal surface. Theoretical investigations of its properties have provided important insight into the nature of the long-range orderings of SAMs. To the extent possible, phenomenological methods,[19,20] molecular dynamic (MD) simulations,[21–26] and models of the SAMs such as the two-dimensional Ising model[27,28] have explained the ground state structure and thermal-equilibrium orientational states of the SAM. One of the objectives is to study SAMs with more complex molecular chains, i.e., self-assembled monolayers of alkanethols that contain a polar group. Molecular dynamic simulations of Langmuir–Blodgett monolayer with dipole group have already been considered.[29,30] However, these simulations, which are based on the so-called united atom model that treats CH_2 groups as single interaction sites, and allowance

for the packing patterns of alkyl chains with one or two molecules per unit cell, give rise to an incorrect monolayer structure, the tilt angle, and tilt direction for zero dipole moment.

In this paper we are considering the effect of incorporating the polar group into self-assembled alkanethiol monolayers on the phase transitions and the molecular structure of the phases. We use an all-atom model for hydrocarbon chain interactions. In order to avoid gauche or kink defect in alkyl chains we placed the polar group at the end of the chains.

2. MOLECULAR MODEL

The model describes the n-alkanethiol rigid chains that terminate the polar group. The SH headgroups of the alkanethiols form a $(\sqrt{3}\times\sqrt{3})R30°$ triangular lattice to adjust with (coordinate) Au(111) substrate. An all-atom model includes hydrogen which is connected by rigid-bond couplings. This model is based mainly on the molecular model which was accepted by Hautman and Klein[22,23] and Mar and Klein,[24] except for the interactions between the chains. The chain has a zig-zag form and consists of hydrocarbon groups CH_2, which begin with sulfur and end with a polar group which a dipole moment **d** directs along the chain axis. We assume that arbitrary dipole moment belongs to the molecular group CH_3. Hydrocarbon groups CH_2 and CH_3 are represented by single interaction sites that include hydrogens. We assume that the chain may freely rotate about the chain axis as a whole with the twisting angle as the dihedral angle between the plane consisting of the normal to the gold surface, and the chain axis and plane defined by trans segments of the zig-zag molecular chain. We also assume that the chain may rotate relative to the crystal surface in such a way that the sulfur does not take part in this rotation. This rotation is determined by two angles: the tilting angle θ and the tilt direction φ (the precession angle of the long molecular axis about the surface normal to the gold crystal surface) (Fig. 1).
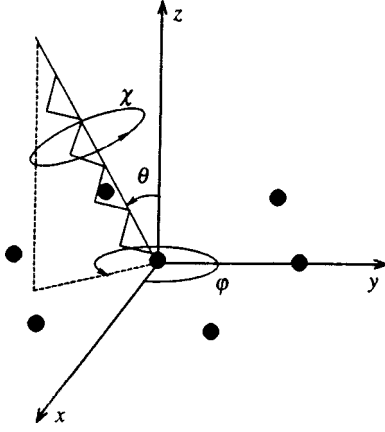
FIG. 1. θ is the tilt angle of the molecule, χ is the chain twist (rotation) angle, and φ characterizes the tilt direction along the surface plane.

Following Ref. 27, let us consider the Lennard–Jones interactions between the atoms H, C, and S

$$U(R) = 4\epsilon \left[\left(\frac{\sigma}{R} \right)^{12} - \left(\frac{\sigma}{R} \right)^6 \right]. \tag{1}$$

The Lennard–Jones (LJ) parameters ϵ and σ were chosen by fitting the potentials (1) to the van der Waals (vdW) potential $\exp(-r^{-6})$ in such a way that the potentials would have the same position, depth, and second derivatives at the point of minimum of the potentials. These parameters are listed in Tables I and II. We used the LJ potential since the vdW potential has a negative divergence for small distances, which means that all chains have a tendency to collapse on each other.

Following Hautman and Klein,[22] the interaction of the hydrocarbon groups CH_2 and CH_3 with the Au substrate was modeled by the 12-3 potential

$$V(z) = \frac{C_{12}}{(z-z_0)^{12}} - \frac{C_3}{(z-z_0)^3}, \tag{2}$$

where $C_{12} = 2.8 \times 10^7$ K Å^2, $C_3 = 17100$ K Å^3, and $z_0 = 0.86$ Å.

The dipole–dipole coupling of the dipoles $\mathbf{d}_i$ and $\mathbf{d}_j$ is

$$W(\mathbf{R}_{ij}) = \sum_{\alpha,\beta} W_{ij}^{\alpha,\beta} d_i^\alpha d_j^\beta, \tag{3}$$

where

$$W_{ij}^{\alpha,\beta} = \frac{\delta_{\alpha,\beta}}{R_{ij}^3} - \frac{3 R_{ij}^\alpha R_{ij}^\beta}{R_{ij}^5}, \tag{4}$$

and $\mathbf{R}_{ij}$ is the vector between the dipole moments; the magnitude of a dipole moment is measured in units of $1D = 4.8 \times 10^{-18}$ CGSE·cm.

The coordinates of the kth carbons in the local coordinate system with $\theta = 0$, $\varphi = 0$, and $\chi = 0$ for the zig-zag chain in Fig. 2 are

$$\mathbf{R}_{Ck} = \begin{cases} (r, \ 0, \ l_0 + (k-1)l/2), & k = 1,3,5,... \ . \\ (0, \ 0, \ l_0 + (k-1)l/2), & k = 2,4,6,..., \end{cases} \tag{5}$$

where the distances in the chain are shown in Fig. 2, and χ is the twisting angle. The coordinates of the hydrogen are

$$\mathbf{R}_{Hks} = \begin{cases} (r+h\cos\alpha, \ hs\sin\alpha, \\ (-h\cos\alpha, \ hs\sin\alpha, \\ l_0 + (k-1)/l/2), & k = 1,3,5,..., \\ l_0 + (k-1)l/2), & k = 2,4,6,... \ . \end{cases} \tag{6}$$

The coordinates of the dipole are the same as those of C_n.

In order to find the coordinates defining the carbon and hydrogen atoms of the n-thiol chain in the coordinate system of the substrate, it is necessary to use the transformations of rotations determined by the Euler matrix

$$\mathbf{T}(\varphi,\theta,\chi) = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}$$
$$\times \begin{pmatrix} \cos\chi & -\sin\chi & 0 \\ \sin\chi & \cos\chi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

This gives a transformation

$$\mathbf{R}(\varphi,\theta,\chi) = \mathbf{T}(\varphi,\theta,\chi)\mathbf{R}. \tag{7}$$

Experimental data[8,9] show that the chain is tilted to the next nearest neighbor (NNN) in the direction from the NNN direction ($\varphi \sim 10°$). Below we consider the tilted phase ($\theta > 0$) only. The number and equilibrium angular positions of the chains of the paraelectric phase depend on the symmetry of the system. We see that the one-body potential of the chain-S (1) and the chain-Au (2) is fourfold degenerate with respect to the angular positions of a chain $\mathbf{R}(\varphi,\theta,\chi)$, $\mathbf{R}(-\varphi,\theta,-\chi)$, $\mathbf{R}(\pi+\varphi,\theta,\chi)$, $\mathbf{R}(\pi-\varphi,\theta,-\chi)$. If the contribution of the potential of the chain from the straight chain of carbon atoms [in (5) $r=0$] is taken into account, then the chain-chain interactions (1) remove the mirror symmetry in the yz plane. Hence, the total one-body potential of a chain is twofold degenerate with respect to the angular positions

TABLE 1. Parameters of the van der Waals potential $A\exp(-Br) - C/r^6$.

A, 10^7 K			B, Å^{-1}			C, K Å^6			Refs.
A_{HH}	A_{HC}	A_{CC}	B_{HH}	B_{HC}	B_{CC}	C_{HH}	C_{HC}	C_{CC}	
0.33	1.4	6.0	4.08	4.08	3.08	2.5	6.3	16	31
1.50	1.5	1.5	5.00	4.13	3.42	2.2	7.0	21	32
1.10	1.1	1.1	4.64	3.94	3.42	2.7	7.2	17	33
0.92	1.1	1.3	4.64	3.94	3.42	2.2	6.9	19	34
0.46	6.8	2.1	4.54	4.56	3.58	2.1	6.4	20	35

TABLE II. The Lennard–Jones coupling parameters.

ϵ_{HH}, K	ϵ_{HC}, K	ϵ_{CC}, K	$\sigma_{0,HH}$, Å	$\sigma_{0,HC}$, Å	$\sigma_{0,CC}$, Å	Refs.
18	29	52	2.653	2.903	3.118	31
36	38	36	2.338	2.814	3.418	32
30	30	30	2.498	2.939	3.387	33
24	29	34	2.500	2.939	3.387	34
33	30	920	2.316	2.941	1.842	35

$\mathbf{R}(\varphi,\theta,\chi)$ and $\mathbf{R}(-\varphi,\theta,-\chi)$. This degeneracy was shown quantitatively in Ref. 27. The equilibrium state of the chains $\mathbf{R}(\varphi_0,\theta_0,\chi_0)$ and that of the mirror plane xz $\mathbf{R}(-\varphi_0,\theta_0,-\chi_0)$ is found from the minimum thermodynamic potential for $T>T_c$ and defined below.

In order to consider the phase transition with spontaneous breakdown of the symmetry we follow Vaks[36] and write the following expressions for the rotated coordinates of the atoms:

$$\mathbf{R}_s = \mathbf{R}(s\varphi,\theta,s\chi)$$

$$= \frac{1}{2}\left[\mathbf{T}(\varphi,\theta,\chi)\mathbf{R}+\mathbf{T}(-\varphi,\theta,-\chi)\mathbf{R}\right]$$

$$+ s\frac{1}{2}\left[\mathbf{T}(\varphi,\theta,\chi)\mathbf{R}-\mathbf{T}(-\varphi,\theta,-\chi)\mathbf{R}\right]$$

$$= \mathbf{R}_{\parallel}+s\mathbf{R}_{\perp}, \quad s=\pm 1. \tag{8}$$

Obviously, since $\mathbf{R}_{\parallel}=(\mathbf{R}_{+1}+\mathbf{R}_{-1})/2$, $\mathbf{R}_{\perp}=(\mathbf{R}_{+1}-\mathbf{R}_{-1})/2$ and $|\mathbf{R}_{+1}|=|\mathbf{R}_{-1}|$,

$$\mathbf{R}_{\parallel}\mathbf{R}_{\perp}=0,$$

and $\mathbf{R}_{\perp}$ is directed along the y axis, as shown in Fig. 3. As the basis vectors of the two-dimensional triangular lattice of sulfur atoms we chose the vectors

$$\mathbf{a}_1 = a(\sqrt{3}/2,1/2,0), \quad \mathbf{a}_2 = (0,1,0), \tag{9}$$

where the lattice constant $a=4.97$ Å. Now let us specify the coordinates of the atoms of the jth chain on the surface:

$$\mathbf{R}_{jgk} = \mathbf{R}_j+\mathbf{R}_{\parallel,gk}+s_j\mathbf{R}_{\perp,gk}, \tag{10}$$

where j runs over all sites of the triangular lattice, k runs over the atomic groups CH_2 and CH_3 along the chain, and g runs over the specific atomic groups ($g=C$, H_1, and H_2). Accordingly, the dipole moment of the jth chain is

$$\mathbf{d}_j = \mathbf{d}_{\parallel}+s_j\mathbf{d}_{\perp}, \tag{11}$$

where $\mathbf{d}_{\perp}\|y$, and $\mathbf{d}_{\parallel}\mathbf{d}_{\perp}=0$.

Substituting expressions (8) and (10) into the potential interactions (1) and (2), we obtain the following expression for the LJ coupling chain-chain energy:

$$\frac{1}{2}\sum_{i,j}\sum_{g,g'}\sum_{k,k'} U_{gg'}(\mathbf{R}_{ij}+\mathbf{R}_{0,gk}-\mathbf{R}_{0,g'k'}$$

$$+ s_i\mathbf{R}_{1,gk}-s_j\mathbf{R}_{1,g'k'}), \tag{12}$$

where $\mathbf{R}_{ij}=\mathbf{R}_i-\mathbf{R}_j$, and for the total chain-surface energy, which includes a coupling of the n-thiol atoms (C, H) with the sulfur atoms,

$$\sum_i\sum_g\sum_k V(R^z_{0,k}+s_iR^z_{1,k})$$

$$+\sum_{i,j}\sum_g\sum_{k,k'} U_{gS}(\mathbf{R}_{ij}+\mathbf{R}_{0,gk}+s_i\mathbf{R}_{1,gk}). \tag{13}$$

In accordance with Eq. (1), we introduce the notation

FIG. 2. Coordinates defining the model of the zig-zag n-thiol chain chemisorbed to a gold substrate; $l_0=1.58$ Å, $l=2.506$ Å, $r=0.878$ Å, $h=1.04$ Å, and $\alpha=55°$.

FIG. 3. Symmetrical positions of an atom of the n-thiol chains in the tilted phase.

$$U_{g,g'} = 4\epsilon_{g,g'}\left[\left(\frac{\sigma_{g,g'}}{R}\right)^{12} - \left(\frac{\sigma_{g,g'}}{R}\right)^6\right], \tag{14}$$

and the notation U_{gS} means the interaction of the chain's atoms with the sulfur.

A single dipole spaced apart from a metal feels an interaction with its self-image, so a dipole-dipole part of the chain-chain energy consists of dipole-dipole, dipole-image and image-image interactions. Substituting expressions (9) and (10) into (3), we obtain the energy of the dipole-dipole coupling

$$\sum_{i,j}\left[W(\mathbf{R}_{ij} + (s_i - s_j)\mathbf{R}_{\perp,C_n}) \right.$$
$$\left. + \frac{1}{2}\,W(\mathbf{R}_{ij} + \mathbf{Z}_{12} + (s_i - s_j)\mathbf{R}_{\perp,C_n})\right], \tag{15}$$

where $\mathbf{Z}_{12}$ is the vector between the real dipole and its self-image.

In order to simplify expressions (12), (13), and (15) we introduce the projection operators $s^{\pm} = (1 \pm s)/2$. Then for any function f of the two operators s_1 and s_2 there is an identity[36]

$$f(xs_1 + ys_2) = (s_1^+ + s_1^-)(s_2^+ + s_2^-)f(xs_1 + ys_2)$$
$$= s_1^+ s_2^+ f(x+y) + s_1^+ s_2^- f(x-y)$$
$$+ s_1^- s_2^+ f(-x+y) + s_1^- s_2^- f(-x-y). \tag{16}$$

Using this identity, we can write the following expression for the total energy of the SAM:

$$E = E_0 - \frac{1}{2}\sum_{ij} J_{ij}(\theta,\varphi,\chi)s_i s_j. \tag{17}$$

In accordance with Eq. (16), we obtain the expressions

$$E_0 = \frac{1}{4}\sum_{ij}\sum_{s,s'}(U_{s,s'}^{ij} + W_{s,s'}^{ij}) + \frac{1}{2}\sum_i\sum_s V_s^i, \tag{18}$$

$$J_{ij} = \frac{1}{4}\sum_{s,s'} ss' U_{s,s'}^{ij} + \frac{1}{4}\sum_{s,s'} ss' W_{s,s'}^{ij}, \tag{19}$$

where

$$U_{s,s'}^{ij} = \frac{1}{2}\sum_{gg'}\sum_{kk'} U_{gg'}(\mathbf{R}_{ij} + \mathbf{R}_{\|,gk}$$
$$- \mathbf{R}_{\|,g'k'} + s\mathbf{R}_{\perp,gk} - s'\mathbf{R}_{\perp,g'k'}),$$

$$W_{s,s'}^{ij} = W(\mathbf{R}_{ij} + (s-s')\mathbf{R}_{\perp,C_n})$$
$$+ \frac{1}{2}\,W(\mathbf{R}_{ij} + \mathbf{Z}_{12} + (s-s')\mathbf{R}_{\perp,C_n}),$$

$$V_s^i = \sum_g\sum_k V(R_{\|,k}^z + sR_{\perp,k}^z)$$

$$+ \sum_j\sum_g\sum_k U_{gS}(\mathbf{R}_{ij} + \mathbf{R}_{\|,gk} + s\mathbf{R}_{\perp,gk}). \tag{20}$$

The linear term $\Sigma_i B_i s_i$ is absent in Eq. (17) due to the symmetry.

3. MEAN-FIELD APPROXIMATION

According to (17), the thermodynamic potential of the SAM is

$$F = E_0(\theta,\varphi,\chi) - T \ln \mathrm{Tr}_{\{s_i\}}$$
$$\times \exp\left[\frac{1}{2}\sum_{ij} J_{ij}(\theta,\varphi,\chi)\frac{s_i s_j}{T}\right]. \tag{21}$$

A mean-field approximation of Eq. (21) is given by the expression[36]

$$F = E_0(\theta,\varphi,\chi) + \frac{1}{2}\sum_{ij} J_{ij}(\theta,\varphi,\chi)\langle s_i\rangle\langle s_j\rangle$$
$$- T\sum_i \ln\left[2\cosh\left(\sum_j J_{ij}(\theta,\varphi,\chi)\frac{\langle s_j\rangle}{T}\right)\right]. \tag{22}$$

We see from Eq. (22) that in the paraelectric phase ($\langle s_i\rangle = 0$) the minimum $E_0(\varphi,\theta,\chi)$ with respect to the angles gives equilibrium $\varphi_0,\theta_0,\chi_0$ SAM's chains. The order parameter $\langle s_i\rangle \neq 0$ is defined as a solution of the equations of state[37]

$$\langle s_i\rangle = \tanh\left[\sum_j J_{ij}(\theta,\varphi,\chi)\frac{\langle s_j\rangle}{T}\right]. \tag{23}$$

A substitution of solutions of this equation into (22) and self-consistent minimization of the thermodynamic potential over three angles give the complete equilibrium state of the SAM. Next, according to Eq. (23), the structure ordered phase is determined by the wave vector $\mathbf{q}_0$, for which

$$J_{\mathbf{q}} = \sum_j J_{ij} \exp(i\mathbf{q}\mathbf{R}_{ij}) \tag{24}$$

takes on the maximum value, and $T_c = J_{\mathbf{q}_0}$.[37]

The first item transform in (19) under nearest-neighbor chain-chain coupling is given by the expression[27]

$$U(\mathbf{q}) = 2J_1 \cos(2\pi\xi_1) + 2J_1 \cos[2\pi(\xi_1 - \xi_2)]$$
$$+ 2J_2 \cos(2\pi\xi_2), \tag{25}$$

where the wave vector $\mathbf{q} = \xi_1\mathbf{b}_1 + \xi_2\mathbf{b}_2$ is written about the basis reciprocal to (9)

$$\mathbf{b}_1 = 4\pi a^{-1}(1/\sqrt{3},0,0), \quad \mathbf{b}_2 = 2\pi a^{-1}(-1/\sqrt{3},1,0), \tag{26}$$

J_1 is the coupling constant along the vectors $\pm\mathbf{a}_1$, $\pm(\mathbf{a}_2 - \mathbf{a}_1)$, and J_2 is the coupling constant along the vectors $\pm\mathbf{a}_2$.

One points to fact that at $\mathbf{d} = 0$ expression (24) reduces to (25), maximums which give the following structure ordered phase and the transition temperature:[27]

1) $\xi_1 = 0$, $\xi_2 = 0$, $T_c = 4J_1 + 2J_2$ (ferro);
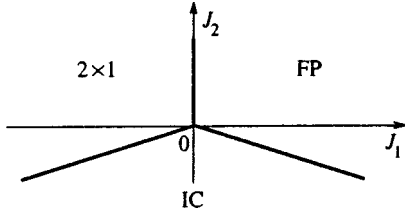
2) $\xi_1 = 0.5$ $\xi_2 = 0$, $T_c = -4J_1 + 2J_2$ (2×1);

FIG. 4. Phase diagram of the Ising model on the triangular lattice.

3) $\xi_1 = 0.5\xi_2$, $T_c = -\dfrac{J_1^2}{J_2} - 2J_2$, $\left|\dfrac{J_1}{2J_2}\right| \leqslant 1$ (IC).

$$\tag{27}$$

The phase diagram, which corresponds to Eq. (27), is shown in Fig. 4.

The dipole-dipole interaction transform of (19) is computed by Ewald's method[38] by fitting to the $2D$ lattice. Formally, the Fourier transform (4) can be found as follows:

$$W_{\mathbf{q},\mathbf{x}}^{\alpha,\beta} = -\frac{\partial^2}{\partial x_\alpha \partial x_\beta} \sum_{\mathbf{l}} \frac{e^{i\mathbf{q}\mathbf{l}}}{|\mathbf{l}-\mathbf{x}|}, \tag{28}$$
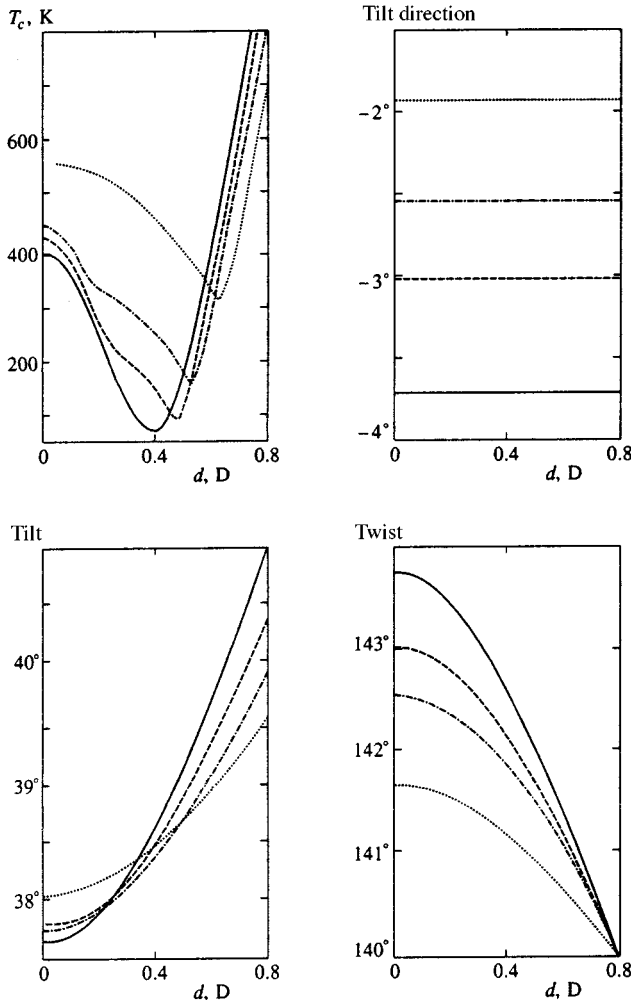
FIG. 5. Curves of the temperature transition T_c and the equilibrium angles of the n-thiols, plotted as functions of the dipole moment, are described by the solid ($n=8$), dashed ($n=10$), dot-dashed ($n=12$), and dotted ($n=17$) lines for the coupling parameters from Ref. 32.
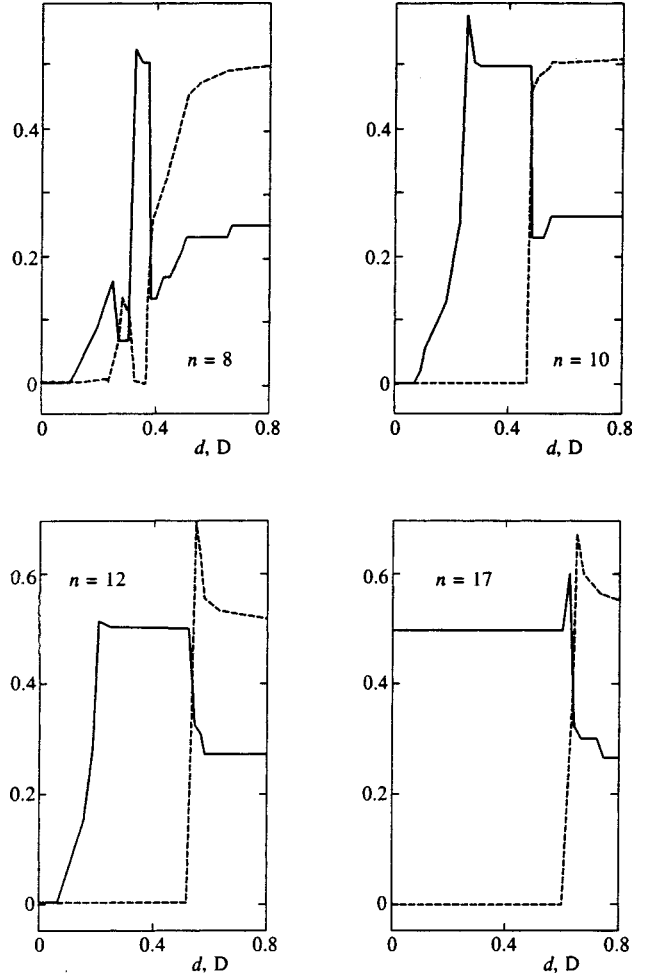
FIG. 6. The components of the wave vector of the ordered structure ξ_1 (solid) and ξ_2 (dashed) for the coupling parameters from Ref. 32.

where $\mathbf{l}=l_1\mathbf{a}_1+l_2\mathbf{a}_2, l_1, l_2$ is an integer, and $\mathbf{a}_1$ and $\mathbf{a}_2$ are defined by expressions (9). According to Ewald, the series of (28) for a plane lattice is presented by the sum of two series:

$$\sum_{\mathbf{l}} \frac{e^{i\mathbf{q}\mathbf{l}}}{|\mathbf{l}-\mathbf{x}|} = \sum_{\mathbf{l}} \frac{\mathrm{erfc}(R|\mathbf{l}-\mathbf{x}|)}{|\mathbf{l}-\mathbf{x}|} + \frac{\pi}{S_0} \sum_{\mathbf{g}} \frac{e^{i\mathbf{x}(\mathbf{q}+\mathbf{g})}}{|\mathbf{g}+\mathbf{q}|}$$

$$\times \left[e^{z|\mathbf{q}+\mathbf{g}|} \mathrm{erfc}\left(\frac{|\mathbf{q}+\mathbf{g}|}{2\mathbf{R}} + zR \right) \right.$$

$$\left. + e^{-z|\mathbf{q}+\mathbf{g}|} \mathrm{erfc}\left(\frac{\mathbf{q}+\mathbf{g}}{2\mathbf{R}} - zR \right) \right], \tag{29}$$

where

$$\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-y^2} dy.$$

Here $\mathbf{g}=g_1\mathbf{b}_1+g_2\mathbf{b}_2, g_1, g_2$ is an integer, $\mathbf{b}_1$ and $\mathbf{b}_2$ are defined by expressions (26), S_0 is the unit cell area, z is the component along the z axis of the vector $\mathbf{x}$, and R is the adjustable parameter of the velocity convergence of the series. Note that Eq. (29) was obtained in Ref. 29 for $\mathbf{q}=0$.

The results of self-consistent numerical minimization procedure of (22) are given in Figs. 5–8 and Table III. The first feature of the temperature transition is high sensitivity to
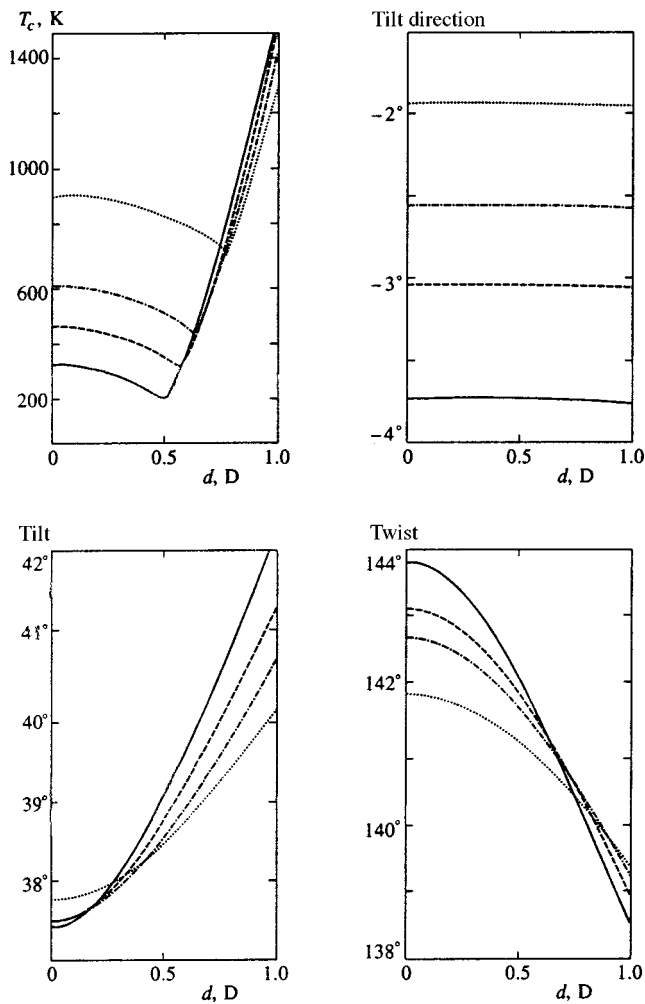
FIG. 7. The same as in Fig. 5 for Ref. 35.

FIG. 8. The same as in Fig. 6 for Ref. 35.

the choice of the coupling constants listed in Table II. For the choice of the coupling constants defined in Refs. 32 and 35 there is ordered phase sequence from the ferroelectric or 2×1 to the IC (Figs. 6 and 8). However, for the choice of coupling constants defined in Refs. 31, 33, and 34 there is IC phase (Table III), where the transition temperature T_c increases with increasing dipole moment.

4. DISCUSSION AND CONCLUSIONS

An advantage of using the Ising variable is that a rich variety of the couplings between atoms of the n-thiols and the couplings with the crystal surface is reduced to a few competing exchange parameters. For $d=0$ they allow one to establish a simple phase diagram of the system shown in Fig. 4. The SAMs are described by the Ising model on the triangular lattice with exact solution.[39] Ferroelectric, 2×1, and incommensurate phases are the only possible ordered states of the system of the n-thiols which are self-assembled on the crystal surface Au(111). For $d\neq0$ competition a LJ interaction and a dipole-dipole coupling can give rise to various combinations of the structures.

The most interesting behavior of the critical temperature and a sequence of ordered structures upon change in the

dipole moment has been found for the LJ's parameters from Refs. 32 and 35. The transition temperature dependence is that T_c with $d\neq0$ can be reduced when $d=0$. In particular, the lowest temperature of the phase transition is realized for the parameters taken from Refs. 32 and 34. T_c is very sensitive to the coupling parameters which change many times. Moreover, it is necessary to take into consideration that all coupling parameters listed in Table I are given for the case of three-dimensional crystals in which the distances between the atomic groups CH_2 may differ in comparison with the case of the SAM. Therefore, the coupling constants listed in Tables I and II should be considered carefully. As far as the structure of the ordered phase for any set coupling parameters is concerned, with increasing dipole moment the incommensurate phase is described by the modulation vector either near $\xi=(0.25,0.5)$ or near $\xi=(0.3,0.6)$.

The phase transition leads to a freezing of the jumps of the chains between twofold degenerate states with equilibrium azimutal and twist angles φ_0, and χ_0 is given in Figs. 5 and 7 and Table III for various references. The values of these angles agree with the experimental observations and theoretical considerations except for the azimutal angle which was determined experimentally to be $\varphi\sim10°$ for $d=0$ and $10<n<20$ (Ref. 9).

The final feature of the phase transition to the twisted

TABLE III. The critical temperatures, the wave vector, and the equilibrium angles.

n	T_c, K	ξ_1	ξ_2	θ_0	χ_0	φ_0	Refs.
				$d = 0.5$ D			
8	1249.6	0.259	0.517	−3.88	35.55	143.75	31
12	1701.6	0.259	0.525	−2.68	35.11	142.95	”
16	2156.9	0.259	0.529	−2.05	34.90	142.50	”
8	688.5	0.259	0.520	−3.82	36.21	143.75	33
12	855.6	0.273	0.538	−2.64	35.80	143.00	”
16	1028.7	0.273	0.549	−2.02	35.61	142.58	”
8	619.7	0.259	0.525	−3.81	36.68	143.37	34
12	745.3	0.273	0.546	−2.63	36.23	142.69	”
16	878.0	0.273	0.561	−2.01	36.01	142.30	”
				$d = 1.0$ D			
8	2857.1	0.253	0.510	−3.79	38.34	141.53	31
12	3224.2	0.259	0.519	−2.63	37.04	141.61	”
16	3642.2	0.259	0.524	−2.02	36.39	141.55	”
8	2227.0	0.253	0.512	−3.77	39.05	141.19	33
12	2311.7	0.267	0.524	−2.60	37.72	141.44	”
16	2442.9	0.267	0.532	−1.99	37.08	141.46	”
8	2154.5	0.267	0.516	−3.77	39.71	140.52	34
12	2200.6	0.267	0.528	−2.60	38.29	140.91	”
16	2290.4	0.267	0.537	−1.99	37.59	141.01	”

ordered phase is that it is a first-order transition. The reason is analogous to the effect of elastic media on the order of the phase transition,[36,40] where the spontaneous ordering gives rise to a distortion of the crystal, which in turn leads to a slight increase in the exchange integrals. Hence, the Curie temperature from the paraelectric phase turns out to be lower than the Curie temperature from the ordered phase. Similarly, in the system of the n-thiol's chains a spontaneous twist ordering of the chains give rise to a change in the tilt. This in turn leads to a change of the exchange integrals defined by Eq. (19).

*)E-mail: zeos@zeos.krascience.rssi.ru

[1] J. D. Swalen, D. L. Allara, J. D. Andrade *et al.*, Langmuir **3**, 932 (1987).
[2] L. H. Dubois and R. G. Nuzzo, Annu. Rev. Phys. Chem. **43**, 437 (1992).
[3] R. G. Nuzzo, F. A. Fusco, and D. L. Allara, J. Am. Chem. Soc. **109**, 2358 (1987).
[4] P. E. Laibinis and G. M. Whitesides, J. Am. Chem. Soc. **114**, 9022 (1992).
[5] A. Kumar, H. A. Biebuyck, and G. M. Whitesides, Langmuir **10**, 1498 (1994).
[6] L. S. Strong and G. M. Whitesides, Langmuir **4**, 546 (1988).
[7] C. E. D. Chidsey and D. N. Loiacono, Langmuir **6**, 682 (1990).
[8] P. Fenter, P. Eisenberger, and K. S. Liang, Phys. Rev. Lett. **70**, 2447 (1993).
[9] P. Fenter, P. Eisenberger, K. S. Liang, and A. Eberhard, J. Chem. Phys. **106**, 1600 (1997).

[10] C. D. Bain, E. B. Troughton, Y.-T. Tao *et al.*, J. Am. Chem. Soc. **111**, 321 (1989).
[11] G. M. Whitesides and P. E. Laibinis, Langmuir **6**, 87 (1990).
[12] S. D. Evans, E. Urankar, A. Ulman, and N. Ferris, J. Am. Chem. Soc. **113**, 4121 (1991).
[13] R. G. Nuzzo, L. H. Dubois, and D. L. Allara, J. Am. Chem. Soc. **112**, 558 (1990).
[14] R. G. Nuzzo, E. M. Korenic, and L. H. Dubois, J. Chem. Phys. **93**, 767 (1990).
[15] T. T. T. Li and M. J. Weaver, J. Am. Chem. Soc. **106**, 6107 (1984).
[16] M. D. Porter, T. B. Bright, D. L. Allara *et al.*, J. Am. Chem. Soc. **109**, 3559 (1987).
[17] N. Camillone III, C. E. D. Chidsey, G. Liu *et al.*, J. Chem. Phys. **94**, 8493 (1991).
[18] N. Camillone III, C. E. D. Chidsey, G. Liu *et al.*, J. Chem. Phys. **98**, 3503 (1993).
[19] R. Nagarajan and E. Ruckenstein, Langmuir **7**, 2934 (1991).
[20] V. M. Kaganer, M. A. Osipov, and I. R. Peterson, J. Chem. Phys. **98**, 3512 (1993).
[21] C. M. Knobler and R. C. Desai, Annu. Rev. Phys. Chem. **43**, 207 (1992).
[22] J. Hautman and M. L. Klein, J. Chem. Phys. **91**, 4994 (1989).
[23] J. Hautman and M. L. Klein, J. Chem. Phys. **93**, 7483 (1990).
[24] W. Mar and M. L. Klein, Langmuir **10**, 188 (1994).
[25] S. Shin, N. Collazo, and S. A. Rice, J. Chem. Phys. **98**, 3469 (1993).
[26] M. A. Moller, D. J. Tildesley, K. S. Kim *et al.*, J. Chem. Phys. **94**, 8390 (1991).
[27] A. F. Sadreev and Y. V. Sukhinin, Phys. Rev. B **54**, 17966 (1996).
[28] A. F. Sadreev and Y. V. Sukhinin, J. Chem. Phys. **107**, 2643 (1997).
[29] V. V. Kislov, Y. A. Kriksin, and I. V. Taranov, Radiotekh. Elektron. **38**, 539 (1993).
[30] V. V. Kislov, Y. A. Kriksin, and I. V. Taranov, Radiotekh. Elektron. **41**, 241 (1996).
[31] L. S. Bartell, J. Chem. Phys. **32**, 827 (1960).
[32] A. I. Kitaigorodsky, Tetrahedron **14**, 230 (1961).
[33] A. I. Kitaigorodsky, Molecular Crystals and Molecules, Academic Press, New York and London (1973), Chap. 2.
[34] A. I. Kitaigorodsky, Adv. Struct. Res. 3, 204 (1970).

[35] T. Wasiutyński and T. Luty, Acta Phys. Pol. A **45**, 551 (1974).

[36] V. G. Vaks, *Introduction to Microscopic Theory of Ferroelectrics,* Nauka, Moscow (1973) (in Russian).

[37] R. Blinc and B. Žekš, *Soft Modes in Ferroelectrics and Antiferroelectrics,* North-Holland Publ., Amsterdam (1974).

[38] J. M. Ziman, *Principles of the Theory of Solids, 2nd ed.,* Cambridge (1972).

[39] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics,* Academic Press, New York (1982), Chap. 11.

[40] O. K. Rice, J. Chem. Phys. **22**, 1535 (1954).

Dynamics of cold atoms in a quadrupole magnetic trap with an orbiting potential

V. G. Minogin, M. V. Subbotin,*) and P. N. Melent'ev

Institute of Spectroscopy, Russian Academy of Sciences, 141092 Troitsk, Moscow Region, Russia
(Submitted 10 November 1997)
Zh. Éksp. Teor. Fiz. **114**, 23–36 (July 1998)

We study the dynamics of atoms confined to a quadrupole magnetic trap with an orbiting potential. For typical values of the experimental parameters of the trap, the rotating magnetic field is shown to produce high-frequency modulation of atomic momenta with an amplitude comparable to the widths of the momentum distributions for the lowest oscillation states of atoms in the time-averaged potential. We find the quantum-statistical momentum and position distributions of atoms and show that at temperatures much higher than the effective vibrational temperature of the atoms in the trap the quantum-statistical momentum and position distributions are Gaussian. We also establish that at temperatures comparable to the effective vibrational temperature of the atoms the quantum-statistical momentum distribution has an annular structure in the trap's symmetry plane, which is due to the deep modulation of the atomic momenta caused by the rotating magnetic field. © *1998 American Institute of Physics.* [S1063-7761(98)00207-8]

1. INTRODUCTION

Lately there has been an upsurge of interest in developing experimental means for the trapping and deep cooling of atoms in magnetic traps with the aim of observing quantum-statistical effects, including Bose-Einstein condensation.[1-15] Among the various types of magnetic traps used in preparing ultracold ensembles of atoms, one of the first and most important from the practical standpoint was the quadrupole magnetic trap with an orbiting potential (TOP). In this type of trap the magnetic field confining the atoms consists of two parts. A stationary quadrupole magnetic field generates a potential. In the central part of the trap this potential localizes atoms with a negative projection of the magnetic moment onto the local direction of the magnetic field vector. However, the potential well created by the quadrupole magnetic field is not perfect. Near its bottom the strength of the quadrupole magnetic field is zero, and the atoms moving in this region can change the sign of the projection of the magnetic moment onto the magnetic field vector as a result of nonadiabatic Majorana transitions and leave the trap.[16] To eliminate this channel for the loss of atoms, another, uniform, magnetic field is applied. This field rotates with a high frequency in the plane perpendicular to the symmetry axis of the quadrupole field. The presence of this rotating magnetic field causes the potential well to rotate about the trap's symmetry axis and to create an orbiting potential. The time-averaged value of the orbiting potential at the trap's center is nonzero and thus suppresses the loss of atoms due to Majorana transitions.[9]

Up to now the motion of individual atoms in a TOP trap has been discussed in the literature using a time-averaged stationary harmonic potential.[9] This approach, however, ignores the fast (compared to the natural oscillation frequencies of the trap) oscillations of atomic momenta associated with the rotating magnetic field.

In this paper we examine the dynamics of noninteracting atoms in a TOP trap using a time-dependent Schrödinger equation that incorporates a magnetic potential oscillating in time.

An analysis of the dynamics of noninteracting atoms in a TOP trap shows that the rotating magnetic field generates oscillations of the atomic momenta with the frequency of the rotating field and an amplitude which, for typical values of trap parameters, is comparable to the widths of momentum distributions of atoms in a time-averaged harmonic potential. While the quantum-statistical distribution of the atomic coordinates remains Gaussian, the momentum distribution at low temperatures acquires an annular structure in the trap's symmetry plane at all temperatures. The annular structure of the momentum distribution becomes appreciable at atomic ensemble temperatures of order of the effective vibrational temperature of the atoms in the time-averaged harmonic potential.

2. MAGNETIC DIPOLE POTENTIAL

In the central region of a TOP trap the total magnetic field $\mathbf{B}_t$ consists of two parts:

$$\mathbf{B}_t = \mathbf{b} + \mathbf{B}. \tag{1}$$

Here $\mathbf{b}$ is the stationary quadrupole field proper,[16]

$$\mathbf{b} = -\tfrac{1}{2} a (x\mathbf{e}_x + y\mathbf{e}_y) + az\mathbf{e}_z, \tag{2}$$

which is expressed in (2) in terms of the gradient a of the quadrupole field at point $x=y=z=0$, the $\mathbf{e}_i$ are unit vectors along axes $i=x, y, z$, and $\mathbf{B}$ is a uniform magnetic field rotating with a frequency Ω in the xy plane:[16]

$$\mathbf{B} = B_0(\mathbf{e}_x \cos \Omega t + \mathbf{e}_y \sin \Omega t), \tag{3}$$

where B_0 is the amplitude of the rotating field.

In the field (1) an atom having a negative projection μ_t of the magnetic moment $\boldsymbol{\mu}$ onto the direction of the field, $\mu_t = -\mu$ ($\mu > 0$), possesses an additional potential energy $U = U(\mathbf{r}, t)$:

$$U = -\boldsymbol{\mu} \cdot \mathbf{B}_t = \mu B_t$$

$$= \mu \sqrt{(B_0 \cos \Omega t - \tfrac{1}{2} ax)^2 + (B_0 \sin \Omega t - \tfrac{1}{2} ay)^2 + a^2 z^2}. \tag{4}$$

We see that the time-dependent magnetic-dipole interaction potential rotates about the z axis in such a way that its minimum moves in the xy plane along a circular orbit of radius $r_0 = 2B_0/a$.

To analyze the motion of atoms in the potential (4), it is convenient to introduce the cylindrical coordinates $x = \rho \cos \varphi$, $y = \rho \sin \varphi$, and z and write the potential in a form corresponding to the standard expression of the integrand for elliptic integrals of the second kind:

$$U = U_m \sqrt{1 - e^2 \sin^2 \xi}, \tag{5}$$

where U_m is the maximum value of the potential oscillating in time, e determines the modulation depth of the oscillating potential and is known in the theory of elliptic integrals as the modulus, and ξ is the effective phase:

$$U_m = U_L \sqrt{\left(1 + \frac{\rho}{r_0}\right)^2 + \frac{4z^2}{r_0^2}}, \tag{6}$$

$$e^2 = \frac{4\rho/r_0}{(1 + \rho/r_0)^2 + 4z^2/r_0^2}, \tag{7}$$

$$\xi = \frac{1}{2}(\Omega t - \varphi - \pi). \tag{8}$$

In (6), $U_L = \mu B_0$ is the energy of magnetic interaction at the central point of the trap ($\mathbf{r} = 0$) and determines the Larmor precession frequency

$$\omega_L = \frac{\mu B_0}{\hbar}. \tag{9}$$

Note that the square of the modulus does not exceed unity. The maximum value of the square of the modulus,

$$(e^2)_{\max} = \frac{2}{1 + \sqrt{1 + 4z^2/r_0^2}} \leqslant 1 \tag{10}$$

is attained for a given z on a circle of radius $\rho = \rho_m$:

$$\rho_m = r_0 \sqrt{1 + \frac{4z^2}{r_0^2}}. \tag{11}$$

It is convenient to write the potential (5) in the form of a Fourier series:

$$U = U_m \left[\frac{1}{2} a_0 + \sum_{n=1}^{\infty} (-1)^n a_n \cos n(\Omega t - \varphi) \right], \tag{12}$$

where the Fourier coefficients

$$a_s = \frac{4}{\pi} \int_0^{\pi/2} \sqrt{1 - e^2 \sin^2 \xi} \, \cos(2s\xi) d\xi \tag{13}$$

($s = 0, 1, 2, \dots$) can be expressed in terms of elliptic integrals of the first and second kind, $K(e)$ and $E(e)$. The first three Fourier coefficients are

$$a_0 = \frac{4}{\pi} E(e),$$

$$a_1 = \frac{4}{3\pi e^2} [(2 - e^2)E(e) - 2(1 - e^2)K(e)], \tag{14}$$

$$a_2 = -\frac{4}{15\pi e^4} [(16 - 16e^2 + e^4)E(e)$$

$$- 8(2 - 3e^2 + e^4)K(e)].$$

We now note that since the modulus e is bounded, according to (10), for the Fourier coefficients we need only expressions for moderate values of the modulus, $e \leqslant 1$. Using the well-known expansions of the complete elliptic integrals $E(e)$ and $K(e)$ in powers of e^2 and keeping only term to the fourth order in e inclusively, we can write the following approximate expression for the potential (5):

$$U = U_0 - U_1 \cos(\Omega t - \varphi) - U_2 \cos 2(\Omega t - \varphi), \tag{15}$$

where

$$U_0 = \frac{1}{2} U_m a_0 = U_m \left(1 - \frac{1}{4} e^2 - \frac{3}{64} e^4 \right),$$

$$U_1 = U_m a_1 = \frac{1}{4} U_m e^2 \left(1 + \frac{1}{4} e^2 \right), \tag{16}$$

$$U_2 = -U_m a_2 = \frac{1}{64} U_m e^4.$$

The dependence of U_0, U_1, and U_2 on the coordinates ρ and z is depicted in Fig. 1.

Below we use the magnetic-dipole interaction potential in the form (15) to establish the equations of motion of an atom in a TOP trap.

3. GENERAL EQUATIONS OF MOTION

The dynamics of an atom in the potential (15) is described by a Schrödinger equation, which in the system of coordinates we have adopted assumes the form

$$i\hbar \frac{\partial \Phi}{\partial t} = -\frac{\hbar^2}{2M} \Delta \Phi + U_0 \Phi - [U_1 \cos(\Omega t - \varphi) + U_2$$

$$\times \cos 2(\Omega t - \varphi)] \Phi,$$

$$\Delta = \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2}{\partial \varphi^2} + \frac{\partial^2}{\partial z^2}. \tag{17}$$

We recall that the amplitudes U_0, U_1, and U_2 of the harmonics of the magnetic potential are functions of the coordinates ρ and z.

In Eq. (17) the two parts of the magnetic potential play different roles. The time-independent potential U_0 generates
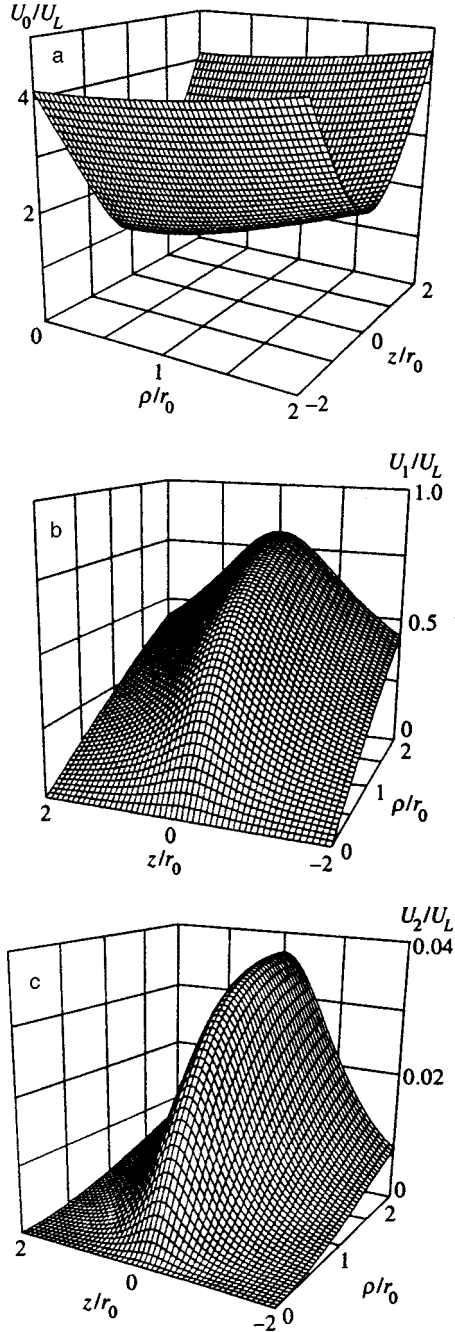
FIG. 1. The time-averaged value U_0 (a) and the amplitudes of the first two harmonics, U_1 (b) and U_2 (c), of the oscillating potential U as functions of the cylindrical coordinates ρ and z.

slow movements of the atom with characteristic times $\tau_{\mathrm{slow}} \simeq 1/\omega$, where ω specifies the characteristic natural oscillation frequencies of the atom in the potential U_0. The presence of the oscillating parts U_1 and U_2 in the potential leads to fast oscillations of the momentum and position of the atom with a characteristic time scale $\tau_{\mathrm{fast}} \simeq 1/\Omega$.

To explicitly separate the fast movements of the atom from the slow, it is convenient to pass to a representation similar to the standard interaction representation for time-independent potentials. For Eq. (17) the corresponding transition is accomplished by the following replacement of the wave function:

$$\Phi = \Phi \, \exp\left[\frac{i}{\hbar\Omega}\left(U_1 \sin(\Omega t - \varphi) + \frac{1}{2} U_2 \sin 2(\Omega t - \varphi)\right)\right]. \tag{18}$$

After (18) is plugged into (17), the equation of motion acquires terms that correspond explicitly to the fast oscillations of the momentum and position of the atom. There also emerges an addition U_{add} to the time-independent potential U_0, which is related to the mean energy of the fast oscillations of the atom:

$$
i\hbar \frac{\partial \Phi}{\partial t} = -\frac{\hbar^2}{2M} \Delta\Phi + (U_0 + U_{\mathrm{add}})\Phi - \frac{i\hbar}{2M\Omega}
$$

$$
\times\left[\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial U_1}{\partial\rho}\right) - \frac{U_1}{\rho^2} + \frac{\partial^2 U_1}{\partial z^2}\right]\sin(\Omega t - \varphi)\Phi
$$

$$
+\frac{i\hbar}{4M\Omega^2}\left[\frac{\partial U_1}{\partial\rho}\frac{\partial U_2}{\partial\rho} + \frac{2U_1 U_2}{\rho^2} + \frac{U_1}{\partial z}\frac{\partial U_2}{\partial z}\right]
$$

$$
\times\cos(\Omega t - \varphi)\Phi - \frac{i\hbar}{4M\Omega}\left[\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial U_2}{\partial\rho}\right)\right.
$$

$$
\left. - \frac{4U_2}{\rho^2} + \frac{\partial^2 U_2}{\partial z^2}\right]\sin 2(\Omega t - \varphi)\Phi - \frac{1}{4M\Omega^2}
$$

$$
\times\left[\left(\frac{\partial U_1}{\partial\rho}\right)^2 - \frac{U_1}{\rho^2} + \left(\frac{\partial U_1}{\partial z}\right)^2\right]\cos 2(\Omega t - \varphi)
$$

$$
\times\Phi - \frac{i\hbar}{M\Omega}\left[\frac{\partial U_1}{\partial\rho}\frac{\partial\Phi}{\partial\rho} + \frac{\partial U_1}{\partial z}\frac{\partial\Phi}{\partial z}\right]\sin(\Omega t - \varphi)
$$

$$
+\frac{i\hbar U_1}{M\Omega\rho^2}\frac{\partial\Phi}{\partial\varphi}\cos(\Omega t - \varphi) - \frac{i\hbar}{2M\Omega}
$$

$$
\times\left[\frac{\partial U_2}{\partial\rho}\frac{\partial\Phi}{\partial\rho} + \frac{\partial U_2}{\partial z}\frac{\partial\Phi}{\partial z}\right]\sin 2(\Omega t - \varphi)
$$

$$
+\frac{i\hbar U_2}{M\Omega\rho^2}\frac{\partial\Phi}{\partial\varphi}\cos 2(\Omega t - \varphi), \tag{19}
$$

where the additional time-independent potential is

$$
U_{\mathrm{add}} = \frac{1}{4M\Omega^2}\left[\left(\frac{\partial U_1}{\partial\rho}\right)^2 + \frac{1}{4}\left(\frac{\partial U_2}{\partial\rho}\right)^2 + \frac{U_1^2 + U_2^2}{\rho^2}\right.
$$

$$
\left. + \left(\frac{\partial U_1}{\partial z}\right)^2 + \frac{1}{4}\left(\frac{\partial U_2}{\partial z}\right)^2\right]. \tag{20}
$$

4. EQUATION OF MOTION OF COLD ATOMS

Let us use the equations of motion derived above to study the dynamics of cold atoms, i.e., atoms near the bottom of the potential well $U_0 + U_{\mathrm{add}}$. For cold atoms the coordinates ρ and z meet the condition $\rho, z \ll r_0$.

When the displacements of an atom from the trap's center are small, we can expand U_0, U_1, and U_2 in power series in the small coordinates ρ and z. Keeping all terms up to second order in ρ and z, we get

$$U_0 = U_L + \tfrac{1}{2}k(\rho^2 + 8z^2), \quad U_1 = 2kr_0\rho, \quad U_2 = \tfrac{1}{2}k\rho^2. \quad (21)$$

The elastic constant for the harmonic part of the time-independent potential is

$$k = \frac{U_L}{2r_0^2}. \quad (22)$$

According to (21), the potential U_0 generates harmonic oscillations of the atom along the x, y, and z axes with the frequencies

$$\omega_x = \omega_y = \omega_\rho = \sqrt{\frac{k}{M}}, \quad \omega_z = 2\sqrt{2}\,\omega_\rho. \quad (23)$$

At this point we note that according to the basic idea of a magnetic trap with an orbiting potential, all of the three fundamental frequencies, ω_ρ, Ω, and ω_L, are assumed to be related by the following inequalities

$$\omega_\rho \ll \Omega \ll \omega_L. \quad (24)$$

According to Eq. (19), the left-hand inequality in (24) is needed so that the action of the time-averaged potential $U_0 + U_{\mathrm{add}}$ would be stronger than the action of the harmonics with the amplitudes U_1 and U_2. The right-hand inequality in (24) excludes the possibility of nonadiabatic Majorana transitions due to rotation of the magnetic field **B**. In a typical experiment $\omega_\rho/2\pi \sim 10$–100 Hz, $\Omega/2\pi \sim 10^4$ Hz, and $\omega_L/2\pi \sim 10^7$ Hz, so that $\omega_\rho/\Omega \sim 10^{-2}$–$10^{-3}$ and $\Omega/\omega_L \sim 10^{-3}$ (Ref. 15).

Now we plug the expressions for U_1 and U_2 from (21) into Eqs. (19) and (20). In the case of cold atoms, the additional potential proves to be much smaller than U_0:

$$U_{\mathrm{add}} = \frac{2k^2 r_0^2}{M\Omega^2}\left(1 + \frac{\rho^2}{r_0^2}\right) = \left(U_L + \frac{1}{8}k\rho^2\right)\left(\frac{\omega_\rho}{\Omega}\right)^2 \ll U_0. \quad (25)$$

Neglecting this small correction to the time-averaged potential U_0, from Eq. (19) we arrive at the following equation of motion of cold atoms:

$$i\hbar\frac{\partial\Phi}{\partial t} = -\frac{\hbar^2}{2M}\Delta\Phi + U_0\Phi + \frac{k^2 r_0}{M\Omega^2}\rho\Phi$$

$$\times \cos(\Omega t - \varphi) - i\frac{2\hbar k r_0}{M\Omega}$$

$$\times\left[\frac{\partial\Phi}{\partial\rho}\sin(\Omega t - \varphi) - \frac{1}{\rho}\frac{\partial\Phi}{\partial\varphi}\cos(\Omega t - \varphi)\right]$$

$$- i\frac{\hbar k\rho}{2M\Omega}\left[\frac{\partial\Phi}{\partial\rho}\sin 2(\Omega t - \varphi)\right.$$

$$\left. - \frac{1}{\rho}\frac{\partial\Phi}{\partial\varphi}\cos 2(\Omega t - \varphi)\right]. \quad (26)$$

Two facts are worth noting. The third term on the right-hand side of Eq. (26), which specifies the oscillating potential of the atom in the interaction representation, has the same structure as the third term on the right-hand side of Eq.

(17). Hence, we could exclude this term from Eq. (26) by a repeated passage to the interaction representation, i.e., by a replacement of type (18):

$$\Phi = \widetilde{\Phi}\,\exp\left[-i\frac{k^2 r_0}{\hbar M\Omega^3}\rho\,\sin(\Omega t - \varphi)\right]. \quad (27)$$

Actually, there is no need for (27), since exclusion of the third term in (26) leads to corrections that are small compared to the other terms as $(\omega_\rho/\Omega)^2 \ll 1$. This follows directly from the fact that the ratio of the exponent in (27) to the first term in the exponent in (18) is

$$\frac{k^2 r_0 \rho}{\hbar M\Omega^3}\frac{U_1}{\hbar\Omega} = \frac{1}{2}\left(\frac{\omega_\rho}{\Omega}\right)^2 \ll 1. \quad (28)$$

Thus, the third term on the right-hand side of Eq. (26) can simply be dropped, since it is very small compared to the other terms. If we now bear in mind that the amplitude of the harmonic oscillations of an atom in the potential U_0 from (21) is of the order of the amplitude ρ_0 of the zero-point oscillations,

$$\rho_0 = \sqrt{\frac{\hbar}{M\omega_\rho}}, \quad (29)$$

we see that the terms in Eq. (26) that oscillate with the doubled frequency 2Ω are small compared to terms oscillating with the frequency Ω in the ratio

$$\frac{\rho}{2r_0} \simeq \frac{\rho_0}{2r_0} = \sqrt{\frac{\omega_\rho}{2\omega_L}} \ll 1. \quad (30)$$

At frequencies used in experiments the left-hand side of Eq. (30) is of order 10^{-3}. If we ignore the terms oscillating with the frequency 2Ω in Eq. (26) and go back to the Cartesian system of coordinates x, y, and z, we arrive at a very simple equation:

$$i\hbar\left[\frac{\partial\Phi}{\partial t} + v_r\left(\frac{\partial\Phi}{\partial x}\sin\Omega t - \frac{\partial\Phi}{\partial y}\cos\Omega t\right)\right]$$

$$= -\frac{\hbar^2}{2M}\Delta\Phi + U_0\Phi, \quad \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}, \quad (31)$$

where, as we can easily see, the quantity

$$v_r = \frac{p_r}{M} = \frac{2\omega_\rho^2 r_0}{\Omega} = \sqrt{2\frac{\omega_L\omega_\rho}{\Omega^2}}\,v_0, \quad (32)$$

expressed in terms of the zero-point oscillation velocity

$$v_0 = \frac{p_0}{M} = \sqrt{\frac{\hbar\omega_\rho}{M}} \quad (33)$$

($p_0 = \sqrt{M\hbar\omega_\rho}$ is the momentum of the zero-point oscillations) specifies the velocity of the fast oscillations of the atom due to fast oscillations of the potential. Here we stress that since in Eq. (31) we dropped the term that oscillates with the doubled orbiting frequency of the potential, together with Eq. (31) we must use the complete wave function (18) in which the phase oscillates only with a frequency Ω, i.e., the function
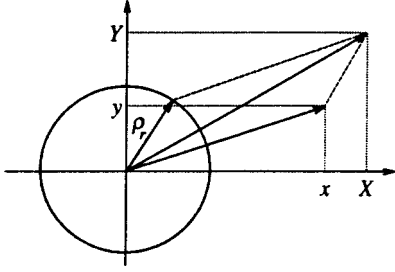
FIG. 2. The true coordinates of an atom, x and y, and the effective coordinates X and Y in the rotation plane of the potential.

$$\Psi = \Phi \, \exp\left[i \frac{p_r}{\hbar} \rho \, \sin(\Omega t - \varphi) \right]. \tag{34}$$

We now note that formally the left-hand side of Eq. (31) appears to be the total (convective) time derivative, which allows for the variation of the wave function Φ in time and due to oscillations of the atom with the velocities $v_r \sin \Omega t$ and $-v_r \cos \Omega t$ along the x and y axes, respectively. Bearing all this in mind, we can introduce the new coordinates X and Y and the new time variable T via the following relationships:

$$X = x + \rho_r \cos \Omega t, \quad Y = y + \rho_r \sin \Omega t, \quad T = t, \tag{35}$$

where ρ_r has the meaning of the amplitude of the fast variations of the coordinate of the atom (Fig. 2),

$$\rho_r = \frac{v_r}{\Omega} = 2\left(\frac{\omega_\rho}{\Omega}\right)^2 r_0 = \rho_0 \frac{\omega_\rho}{\Omega} \sqrt{2 \frac{\omega_L \omega_\rho}{\Omega^2}}. \tag{36}$$

In terms of the new variables, the equation of motion (31) reduces, after neglect of the corrections to the potential of order $\rho_r/\rho_0, (\rho_r/\rho)^2 \ll 1$ on the right-hand side, to the equation of a three-dimensional asymmetric oscillator

$$i\hbar \frac{\partial \Phi}{\partial T} = -\frac{\hbar^2}{2M} \Delta \Phi + U_0 \Phi, \tag{37}$$

where the Laplacian is defined in terms of the coordinates X, Y, and z:

$$\Delta = \frac{\partial^2}{\partial X^2} + \frac{\partial^2}{\partial Y^2} + \frac{\partial^2}{\partial z^2}$$

Note that for the typical frequencies mentioned earlier, the dimensionless parameter

$$S = \sqrt{\frac{\omega_L \omega_\rho}{\Omega^2}} \tag{38}$$

is of order unity. In view of this, the fast spatial modulation of the wave function due to oscillations of the effective coordinates X and Y has, according to (35), a small relative depth of order $\rho_r/\rho_0 \sim \omega_\rho/\Omega \ll 1$. If we ignore this very shallow modulation, we must assume that the effective coordinates X and Y coincide with the true coordinates of the atom x and y and, therefore, that the function Φ is specified by the equation (37) of a three-dimensional oscillator with a Laplacian defined, as usual, with respect to the coordinates x, y, and z.

Equation (37) is the final equation of motion of cold atoms. It contains no information about the shallow high-frequency modulation of the wave function on the spatial scale ρ_r. At the same time, the expression (34) for the complete wave function has retained the information about the fast modulation of the velocity (momentum) of the atom, whose relative depth is of order unity. This fact, as we shall see in Sec. 5, leads to an annular statistical velocity (momenta) distribution of cold atoms.

5. QUASISTATIONARY STATES AND STATISTICAL DISTRIBUTIONS

In the Cartesian coordinates x, y, and z, Eq. (37) has the well-known steady-state solutions:

$$\Phi = \Phi_{klq}(\mathbf{r}, t) = \chi_k(x)\chi_l(y)\chi_q(z)\exp\left[-i\left(\omega_L + \frac{E_{klq}}{\hbar} \right)t \right]. \tag{39}$$

Here $\chi_\alpha(\xi) = \chi_k(x)$, $\chi_l(y)$ are the eigenfunctions of linear oscillators along the axes $\xi = x, y$, and $\alpha = k, l = 0, 1, 2, \ldots$, i.e.,

$$\chi_\alpha(\xi) = \frac{1}{\sqrt{2^\alpha \alpha! \sqrt{\pi}\rho_0}} \exp\left(-\frac{\xi^2}{2\rho_0^2} \right) H_\alpha\left(\frac{\xi}{\rho_0} \right), \tag{40}$$

where $H_\alpha(\eta)$ is a Hermite polynomial, and the eigenfunctions $\chi_q(z)$ of the linear oscillator along the z axis are given by (40) after the replacements $\xi \rightarrow z$ and $\rho_0 \rightarrow \rho_{0z} = \rho_0/\sqrt{2\sqrt{2}}$. The energy eigenvalues of a three-dimensional asymmetric oscillator are

$$E_{klq} = \hbar\omega_\rho(k + l + \tfrac{1}{2}) + 2\sqrt{2}\hbar_\rho(q + \tfrac{1}{2}). \tag{41}$$

According to (34), the complete wave functions in the coordinate representation have the meaning of quasistationary wave functions:

$$\Psi_{klq}(\mathbf{r}, t) = \chi_k(x)\chi_l(y)\chi_q(z)$$
$$\times \exp\left[\frac{ip_r(x \sin \Omega t - y \cos \Omega t)}{\hbar} \right.$$
$$\left. -i\omega_L t - \frac{iE_{klq}t}{\hbar} \right]. \tag{42}$$

In the momentum representation

$$\Phi(\mathbf{p}, t) = (2\pi\hbar)^{-3/2} \int \Phi(\mathbf{r}, t)\exp\left(-\frac{i\mathbf{p}\cdot\mathbf{r}}{\hbar} \right) d^3r, \tag{43}$$

the complete quasistationary wave functions are

$$\Phi_{klq}(\mathbf{p}, t) = \chi_k(p_x - p_r \sin \Omega t)\chi_l$$
$$\times (p_y + p_r \cos \Omega t)\chi_q(p_z)$$
$$\times \exp\left[-i\omega_L t - \frac{iE_{klq}t}{\hbar} \right], \tag{44}$$

where $\chi_\alpha(p)$ are the eigenfunctions of the linear oscillators along the x and y axes in the momentum representation. These functions directly reflect the high-frequency modulation of the momentum of the atom with a relative depth $p_r/p \simeq 1$.

Now we use (42) and (44) to study the quantum-statistical distributions of the position and momentum of an atom. At a given temperature T the statistical distribution of the position of an atom,

$$w(\mathbf{r}) = w(x)w(y)w(z), \tag{45}$$

is defined by the well-known Bloch formula.[17] The distribution of the position of an atom along the axes $\zeta = x,\ y$ is[18]

$$w(\zeta) = \frac{\sum_{\alpha=0}^{\infty} \exp\left[-\dfrac{\hbar\omega_\rho(\alpha+1/2)}{k_B T}\right]\chi_\alpha^2(\zeta)}{\displaystyle\sum_{\alpha=0}^{\infty} \exp\left[-\dfrac{\hbar\omega_\rho(\alpha+1/2)}{k_B T}\right]}$$

$$= \frac{1}{\sqrt{\pi}a_0}\exp\left(-\frac{\zeta^2}{a_0^2}\right), \tag{46}$$

where

$$a_0 = \rho_0 \sqrt{\coth\frac{\hbar\omega_\rho}{2k_B T}} \tag{47}$$

is the half-width of the distribution. The distribution of the coordinate of an atom along the z axis is defined by the same formulas (46) and (47) after the replacements $\xi \to z$, $\rho_0 \to \rho_{0z} = \rho_0/\sqrt{2\sqrt{2}}$, and $\omega_\rho \to 2\sqrt{2}\omega_\rho$.

The quantum-statistical distribution of atomic momentum is given by an expression containing additional averaging over the high-frequency oscillations of the momentum:

$$w(\mathbf{p}) = \langle w(p_x,t)w(p_y,t)\rangle_t w(p_z). \tag{48}$$

Here, according to the Bloch formula, the partial distributions along the x, y, and z axes are

$$w(p_x,t) = \frac{1}{\sqrt{\pi}q_0}\exp\left[-\frac{(p_x - p_r\sin\Omega t)^2}{q^2}\right],$$

$$w(p_y,t) = \frac{1}{\sqrt{\pi}q_0}\exp\left[-\frac{(p_y + p_r\cos\Omega t)^2}{q^2}\right], \tag{49}$$

$$w(p_z,t) = \frac{1}{\sqrt{\pi}q_{0z}}\exp\left[-\frac{p_z^2}{q_{0z}^2}\right],$$

where

$$q_0 = p_0\sqrt{\coth\frac{\hbar\omega_\rho}{2k_B T}}, \quad q_{0z} = 2\sqrt{2}p_0\sqrt{\coth\frac{\sqrt{2}\hbar\omega_\rho}{2k_B T}}. \tag{50}$$

For the two-dimensional distribution of the momentum in the xy plane with allowance for time averaging we have the following expression:

$$w(p_x,p_y) = w(p) = \langle w(p_x,t)w(p_y,t)\rangle_t$$

$$= \frac{1}{\pi q_0^2}\exp\left[-\frac{p^2+p_r^2}{q_0^2}\right]\frac{1}{\pi}$$

$$\times \int_0^\pi \cosh\left(2\frac{pp_r}{q_0^2}\sin\tau\right)d\tau, \tag{51}$$
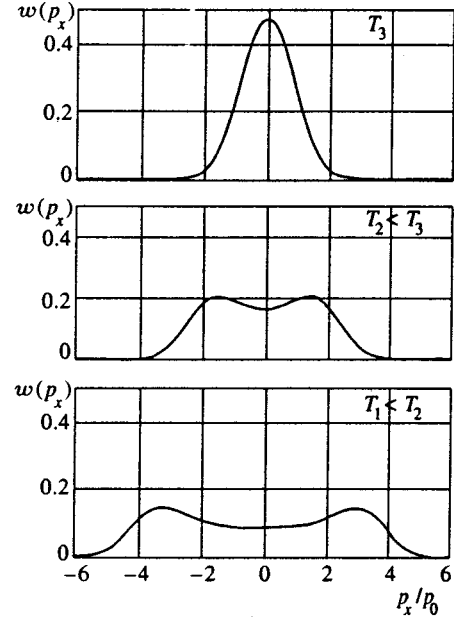
FIG. 3. Quantum-statistical distributions of a momentum component for a single atom. The distributions correspond to the temperatures $T_1 = \hbar\omega_\rho/2k_B$, $T_2 = 2\hbar\omega_\rho/k_B$, and $T_3 = 25\hbar\omega_\rho/k_B$ and the parameter $S = 3$.

which is normalized to one atom (just as the other distributions are):

$$\int w(p_x,p_y)dp_x dp_y = 2\pi\int w(p)p\,dp = 1. \tag{52}$$

According to (45) and (46), the coordinate distribution has an ordinary Gaussian shape in this case.

The distributions of the momentum of an atom in the xy plane for three temperatures are depicted in Fig. 3. Clearly, at low temperatures, where the modulation depth is close to unity, the distribution acquires an annular structure in the xy plane.

According to (52), the annular structure of the momentum distribution emerges when, as the temperature decreases, the dimensionless parameter

$$d = \frac{p_r}{q_0} = S\sqrt{2\tanh\frac{\hbar\omega_\rho}{2k_B T}} \tag{53}$$

(S is another dimensionless parameter) increases to values of order unity. This becomes especially evident if we write the statistical distribution as a power series in the momenta:

$$w(p_x,p_y) = w(p) = \frac{1}{\pi q_0^2}\exp\left(-\frac{p^2+p_r^2}{q_0^2}\right)$$

$$\times\left(1 + \frac{p_r^2}{q_0^4}p^2 + \frac{p_r^2}{4q_0^8}p^4 + \cdots\right). \tag{54}$$

6. ANHARMONIC OSCILLATION SPECTRUM

Up to this point the motion of atoms in a trap was considered in the harmonic approximation (21). Let us now examine the spectrum of natural frequencies of an atom with allowance for anharmonicity.

Retaining the terms up to the eighth order in e in the expression for U_0 in (16),

$$U_0 = U_m\left[1 - \frac{1}{4}e^2 - \frac{3}{64}e^4 - \frac{5}{256}e^6 - 7\left(\frac{5}{128}\right)^2 e^8\right], \quad (55)$$

and expanding (55) in a power series up to the fourth order in the small displacements ρ/r_0 and z/r_0, we get

$$U_0 = U_L + \frac{1}{2}k(\rho^2 + 8z^2) + \frac{k}{r_0^2}\left(\frac{1}{32}\rho^4 + \rho^2 z^2 - 4z^4\right). \quad (56)$$

For the oscillating parts of the potential, U_1 and U_2, we can keep the expressions (21) for them.

As can be seen from the previous section, allowing for the anharmonicity of the oscillations of an atom only requires solving Eq. (37) with potential (56). As usual, we use standard perturbation theory for this purpose. Since the anharmonic term in (56) depends only on the cylindrical coordinates ρ and z, in the case at hand it is convenient to take the solution of Eq. (37) in the cylindrical coordinates ρ, z, and φ as the zeroth-approximation solution. Thus, instead of the solutions (39) we write

$$\Phi_{nmq}(\rho, \varphi, z) = u_{nm}(\rho)e^{im\varphi}\chi_q(z)$$

$$\times \exp\left[-i\left(\omega_L + \frac{E_{nmq}}{\hbar}\right)t\right]. \quad (57)$$

Here the radial wave functions ($n = 0, 1, 2, \ldots$; and $m = 0, \pm 1, \pm 2, \ldots$) are

$$u_{nm}(\rho) = \sqrt{\frac{n!}{\pi(n + |m|)!}}\frac{1}{\rho_0}\left(\frac{\rho}{\rho_0}\right)^{|m|}$$

$$\times \exp\left(-\frac{\rho^2}{2\rho_0^2}\right)L_n^{|m|}\left(\frac{\rho^2}{\rho_0^2}\right), \quad (58)$$

where the $L_n^{|m|}(\eta)$ are associated Laguerre polynomials and ρ_0 defined in (29). The energy eigenvalues for the solutions (57) are

$$E_{nmq} = \hbar\omega_\rho(2n + |m| + 1) + 2\sqrt{2}\hbar\omega_\rho(q + \tfrac{1}{2}). \quad (59)$$

Using the known expressions for the diagonal matrix elements of the second and fourth powers of the coordinates ρ and z, we arrive at the following expression for the first-order correction to the energy eigenvalues (59) due to anharmonicity:

$$\Delta E_{nmq} = \frac{k}{r_0^2}\langle nmq|\frac{1}{32}\rho^4 + \rho^2 z^2 - 4z^4|nmq\rangle$$

$$= 2\hbar\omega_\rho\left(\frac{\omega_\rho}{\omega_L}\right)\left\{\frac{1}{32}[(n+1)(n+2) + (n+|m|)\right.$$

$$\times(5n + |m| + 3)] + \frac{1}{2\sqrt{2}}(2n + |m| + 1)$$

$$\times\left(q + \frac{1}{2}\right) - \frac{3}{4}\left(q^2 + q + \frac{1}{2}\right)\right\}. \quad (60)$$

The foregoing relations imply that the anharmonicity of U_0 becomes essential when the atomic displacements are much larger than the amplitude of the zero-point oscillations, $\rho, z \sim \rho_0\sqrt{\omega_L/\omega_\rho}$, and that the energy spectrum ceases to be equidistant at large quantum numbers, $n, |m|, q \sim \sqrt{\omega_L/\omega_\rho} \simeq 10^3$.

7. CONCLUSION

The theory developed in this paper shows that the motion of atoms in a magnetic trap with an orbiting potential consists of two parts: slow oscillatory motion in a time-averaged potential and fast oscillations with a frequency determined by the rotation frequency of the magnetic field. For typical experimental values of the three fundamental frequencies, ω_ρ, Ω, and ω_L, the amplitude of the fast oscillations of the coordinates of an atom is negligible compared to the amplitude of the slow oscillations in the averaged potential. On the other hand, the amplitude of the high-frequency oscillations of the momentum of an atom is, in general, comparable to the width of the momentum distribution for the fundamental oscillation state of the atom in the trap potential. This deep modulation of atomic momenta becomes especially significant at low temperatures of the atoms, where the average thermal energy becomes comparable to the oscillation energy of atoms in the averaged potential. Accordingly, at low temperatures the momentum distribution of atoms generally acquires an annular structure in the trap's symmetry plane.

*)E-mail: subbotin@lls.isan.troitsk.ru

[1] V. S. Bagnato, G. P. Lafyatis, A. G. Martin, E. L. Raab, R. N. Ahmad-Bitar, and D. E. Pritchard, Phys. Rev. Lett. **58**, 2194 (1987).

[2] H. F. Hess, G. P. Kochanski, J. M. Doyle, N. Masuhara, D. Kleppner, and T. J. Greytak, Phys. Rev. Lett. **59**, 672 (1987).

[3] E. L. Raab, M. Prentiss, A. Cable, S. Chu, and D. E. Pritchard, Phys. Rev. Lett. **59**, 2631 (1987).

[4] P. L. Gould, P. D. Lett, P. S. Julienne, W. D. Phillips, H. R. Thorsheim and J. Weiner, Phys. Rev. Lett. **60**, 788 (1988).

[5] A. G. Martin, K. Helmerson, V. S. Bagnato, G. P. Lafyatis, and D. E. Pritchard, Phys. Rev. Lett. **61**, 2431 (1988).

[6] H. F. Hess, Phys. Rev. B **34**, 3476 (1986).

[7] N. Masuhara, J. M. Doyle, J. C. Sandberg, D. Kleppner, T. J. Greytak, H. F. Hess, and G. P. Kochanski, Phys. Rev. Lett. **61**, 935 (1988).

[8] K. B. Davis, M.-O. Mewes, M. A. Joffe, M. R. Andrews, and W. Ketterle, Phys. Rev. Lett. **74**, 5202 (1995).

[9] W. Petrich, M. H. Anderson, J. R. Ensher, and E. A. Cornell, Phys. Rev. Lett. **74**, 3352 (1995).

[10] M. H. Anderson, J. R. Ensher, M. R. Matthews *et al.*, Science **269**, 198 (1995).

[11] C. C. Bradley, C. A. Sackett, J. J. Tollett, and R. G. Hulet, Phys. Rev. Lett. **75**, 1687 (1995).

[12] K. B. Davis, M.-O. Mewes, M. R. Andrews, N. J. van Druten, D. S. Durfee, D. M. Kurn, and W. Ketterle, Phys. Rev. Lett. **75**, 3969 (1995).

[13] M.-O. Mewes, M. R. Andrews, N. J. van Druten, D. M. Kurn, D. S. Durfee, and W. Ketterle, Phys. Rev. Lett. **77**, 416 (1996).

[14] D. S. Jin, J. R. Ensher, M. R. Matthews, C. E. Wieman, and E. A. Cornell, Phys. Rev. Lett. **77**, 420 (1996).

[15] M.-O. Mewes, M. R. Andrews, N. J. van Druten, D. M. Kurn, D. S. Durfee, C. G. Townsend, and W. Ketterle, Phys. Rev. Lett. **77**, 988 (1996).

[16] T. H. Bergeman, P. McNicholl, J. Kycia, H. Metcalf, and N. Balazs, J. Opt. Soc. Am. B **6**, 2249 (1989).

[17] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, Vol. 1, 3rd ed., Pergamon Press, Oxford (1980) [Russ. orig., Part 1, 3rd ed., Nauka, Moscow (1976), p. 107].

[18] S. Flügge, *Practical Quantum Mechanics*, Vol. 1, Springer, Berlin (1971), p. 107.

Multiwell potentials of endohedral complexes of icosahedral symmetry

A. B. Roĭtsin,*) A. A. Klimov, and L. V. Artamonov

Institute of Semiconductor Physics, Ukrainian National Academy of Sciences, 252650 Kiev 28, Ukraine
(Submitted 29 August 1997)
Zh. Éksp. Teor. Fiz. **114**, 222–238 (July 1998)

We examine multiwell potentials that correspond to the displacements of off-center ions in complexes with icosahedral symmetry (dodecahedrons, icosahedrons, fullerenes, etc.) along symmetry directions (toward vertices, midpoints of edges, and centers of faces). An expression is derived for an effective Hamiltonian, which describes the behavior of endohedral complexes with off-center ions placed in external electric fields of arbitrary strength and orientation. We find the eigenvalues of this Hamiltonian and calculate the intensities of the lines of all possible transitions between tunneling levels. We also predict and analyze the spectra of paraelectric resonances in the absence of an external static electric field (zero-field resonances) and in the presence of such a field. Finally, we provide recommendations for detecting these effects and discuss the specific features of the effects and the possibility of studying them. © *1998 American Institute of Physics.* [S1063-7761(98)01607-2]

1. INTRODUCTION

The discovery of fullerenes and the synthesis of fullerene-based crystals, or fullerites, has led to new avenues of research being developed in different areas of science.[1] There has also been an upsurge of interest in symmetries whose elements include rotations through angles that are integral multiples of $2\pi/5$, in particular in the icosahedral group Y_h (Ref. 2). Not only the unique structure and symmetry of these new carbon formations, but also the unusual diverse properties of such substances have drawn much attention. For instance, it was found that the doping of fullerites with atoms of other elements can give rise to semiconducting and metallic (including superconducting) properties.[3]

It has proved possible to implant atoms (molecules) of elements of different groups of the periodic table (up to lanthanides and even uranium[4-6]) directly into fullerenes. Such structures became known as endohedral fullerenes and are denoted by $M@C_N$, where M is the implanted atom (ion) or group of atoms, and N is the number of carbon atoms in the fullerene. Using endohedral fullerenes to produce condensed materials opens new possibilities to solid-state physics. In particular, Cioslowski and Nanayakkara[7] found that implanting polar molecules in fullerenes may serve as a base for the production of a new class of ferroelectric crystals. Wang *et al.*[8] pointed to the possibility of the emergence of a new class of high-T_c superconductors with specific electron-phonon coupling based on endohedral fullerenes (in contrast to the impurity fullerites already tested, where the impurity is intercalated between fullerenes[3]).

The large diameter of the fullerene cage (in the case of C_{60} the diameter is roughly 7.1 Å) points to the possibility of the appearance of the off-center effect which has been observed in crystals with local symmetry that is lower than icosahedral symmetry.[9-11] This effect consists of a displacement of the potential-energy minimum for the interaction of the guest particle and the fullerene cage from the center of

symmetry of the latter and is usually present when the atomic radius of the particle is small compared to the radius of the cavity in which the particle is implanted. These ideas have been corroborated by calculations. For instance, Wang *et al.*[8] calculated the equilibrium positions (r_{min}) of a large number of ions of the periodic table implanted in C_{60} and found that for many of these ions $r_{min} \neq 0$, i.e., the atom is not at the center of symmetry of C_{60}. Joslin *et al.*[12] pointed out, among other things, that in the case of Li^+ ions $r_{min} \approx 1.4$ Å, while the calculations done by Cioslowski and Fleischmann[13] and Ballester and Dunlap[14] for the Na^+ ion yielded $r_{min} \approx 0.7$ Å. In heavier alkali-metal ions (e.g., K^+; Ref. 15) and inert gases,[16-18] $r_{min} = 0$.

The fullerene C_{60} and a number of higher fullerenes C_N ($N = 80$, 180, 240, 540, and 960) have icosahedral symmetry.[19,20] Moreover, the calculations of Tang *et al.*[20-22] showed that stable giant fullerenes ($N \approx 10^4$) with icosahedral symmetry are possible. Bearing all this in mind, we can assume that the off-center effect manifests itself most vividly and often in higher fullerenes. Clusters consisting of other atoms and having different symmetries, including the icosahedral symmetry, have also been studied.[23-25] Jinlong *et al.*[26] studied the icosahedral cluster $M@Co_{12}$, which consists of 12 cobalt atoms with an atom of various elements (M) of the iron group, from Ti to Ni, inside it. They showed that all clusters except those with Ti, Mn, and Co atoms, which have a closed electronic shell inside the cluster, allow a displacement from the center of the icosahedron. It has also been found that the clusters S_{12} (Ref. 27) and B_{12} (Ref. 28) have an icosahedral structure. Finally, we note that long before fullerenes were discovered, Bersuker *et al.*[29] studied the configurational instability of clusters with icosahedral symmetry due to the Jahn–Teller effect.

Off-center effects are accompanied by the presence of several (N) equivalent equilibrium positions. In the case of impurity ions, these positions are displaced from the center of symmetry, while in the case of dipolar molecules there

can be both displacements and dipole orientation along equivalent directions. Since both the potential wells and the wave functions of a particle moving in them overlap, inter-well tunneling of the particle and splitting of the N-fold degenerate state occur. The characteristic system of tunneling energy levels emerging in the process produces various phenomena, including paraelectric resonance (PER),[9,10] which can be used to study off-center effects directly.

Thus, the results of research point to the possibility of off-center effects manifesting themselves in many endohedral complexes with icosahedral symmetry. These effects lead to multiwell potentials and a system of tunneling levels between which quantum transitions can occur. In this paper we study the energy structure of off-center ions in the intracrystalline electric field generated by different configurations of atoms, such as an icosahedron, a dodecahedron, a truncated dodecahedron, which has been termed a roklar, and a fullerene. We also investigate the effect of an external static electric field and the transitions between tunneling levels induced by an external variable electric field.

2. ENERGY LEVELS: A GENERAL TREATMENT

Earlier studies[9,10] of highly symmetric systems (the O_h group) have shown that off-center atoms are displaced along symmetry directions of the [100], [110], or [111] type, i.e., toward faces, edges, and vertices. We shall thus study the symmetry directions of displacement with allowance for the results of the calculations of $r_{\min}$ in Refs. 8, 12–18. We shall examine in detail one of the most complicated multiwell potentials, viz., a 30-well potential, which simultaneously corresponds to displacements of an off-center ion toward the vertices of a roklar and the midpoints of the edges of an icosahedron, a dodecahedron, and a fullerene (in the last case the edges are those that connect two hexahedrons). In all these cases the equilibrium positions i form a roklar (Fig. 1). The corresponding reducible representation Π_{30} maps the centers of the potential wells into one another. On the basis of the character table of the irreducible representations of the Y_h group (Ref. 30), we can write the

$$\Pi_{30} = A_g + F_{1u} + F_{2u} + G_g + G_u + 2H_g + H_u, \qquad (1)$$
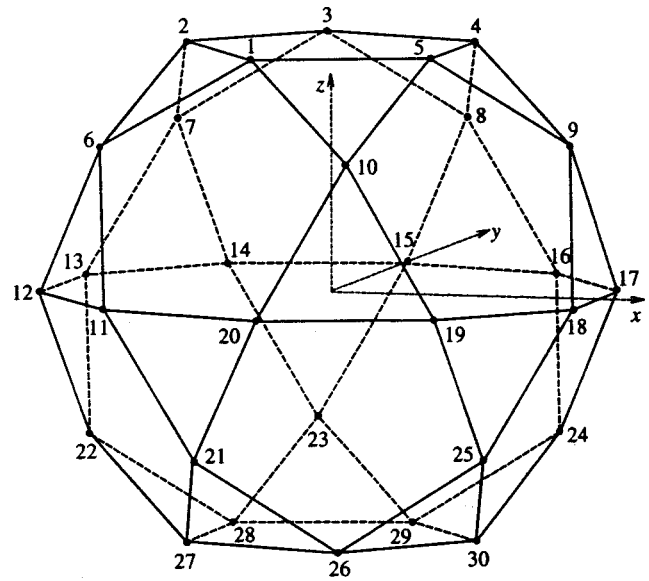
FIG. 1. Roklar (truncated dodecahedron): 32 faces, (12 pentagons and 20 triangles), 30 vertices, and 60 edges. The equilibrium positions i of off-center ions (directions of dipoles) in the case of a 30-well potential in endohedral complexes with icosahedral symmetry. The points i are at vertices of the roklar.

which points to the nature and number of tunneling levels.

Let us examine the resonant transitions in the system of levels (1) in the absence of an external static electric field $\mathbf{E}_0$, i.e., zero-field paraelectric resonance. In such a process the transitions are initiated by a variable electric field, whose intensity vector $\mathbf{e}$ is transformed according to the F_{1u} irreducible representation. An analysis of the direct products of the irreducible representations of the Y_h group (Ref. 30) leads to the allowed transitions depicted in Fig. 2a; here the transitions are initiated by a vector $\mathbf{e}$ arbitrarily oriented in space.

When the field $\mathbf{E}_0$ is switched on, the tunneling levels split, the number of the new levels depending on the orientation of $\mathbf{E}_0$. To make the analysis of the spectra more convenient, the field is usually oriented along the symmetry directions of the geometric figures, so that some symmetry elements are conserved. For symmetry directions we select
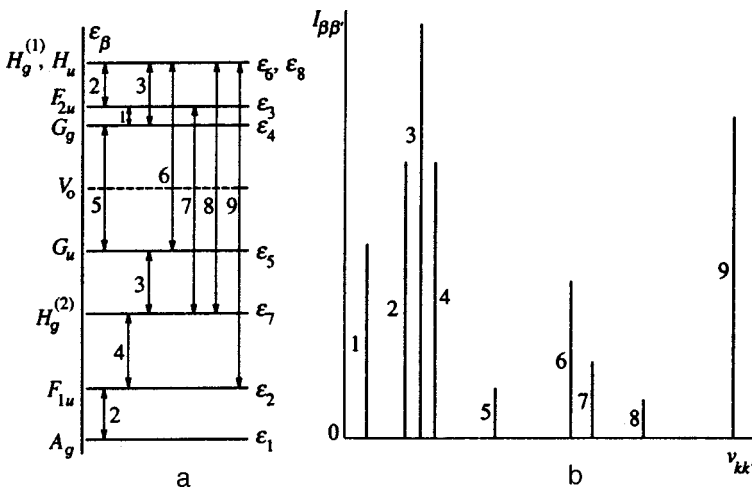

FIG. 2. Tunneling levels in the case of a 30-well potential corresponding to displacements of an off-center ion toward the vertices of a roklar, the midpoints of the edges of an icosahedron and a dodecahedron, and the midpoints of the edges of a fullerene that connect two hexahedrons. (a) Possible resonant transitions in zero-field paraelectric resonance. (b) Relative line intensities $I_{\beta\beta'}$.

TABLE I. Point groups of geometric figures in the presence of an electric field $\mathbf{E}_0$.

Geometric figure	Direction from center to		
	face	vertex	edge
Icosahedron	C_{3v}	C_{5v}	C_{2v}
Dodecahedron	C_{5v}	C_{3v}	C_{2v}
Roklar	C_{5v}/C_{3v}	C_{2v}	C_s
Fullerene	C_{5v}/C_{3v}	C_s	C_s/C_{2v}

the directions from the inversion center to the vertices, centers of the faces, and midpoints of the edges of the various geometric figures. Table I lists the symmetry subgroups that emerge for such directions of the field. The presence of two point groups in one cell in the case of fullerene is due to the two types of faces and edges. In each case the first group refers to a pentagon and an edge connecting a pentagon and a hexagon, while the second group refers to a hexagon and an edge connecting two hexagons. Similarly, for a roklar the first and second groups refer to a pentagon and a triangle, respectively. Table II lists the data on the nature of level splitting induced by the field $\mathbf{E}_0$. Each number in Table II shows how many times the irreducible representation listed in the heading of the respective column under a subgroup notation is contained in the irreducible representation of the Y_h group listed in the heading of the respective row. We see that to establish the selection rules in an arbitrary case, each irreducible representation of any point subgroup must be examined only once or twice. The results of the calculations are depicted in Fig. 3, where the arrows denote allowed electric-dipole transitions, and the labels $\perp$ (x,y) and $\parallel$ (z) next to the arrows indicate the direction of the components of $\mathbf{e}$. The absence of labels means that a transition is possible for any direction of $\mathbf{e}$.

3. PERTURBATION MATRICES AND EFFECTIVE HAMILTONIAN

The method of deriving the generalized effective Hamiltonian for centers with a multiwell potential and arbitrary local symmetry has been described in Ref. 31. In our case the Hamiltonian is a 30×30 matrix $\mathbf{M}_{30}$, whose elements are defined in the basis set of the symmetrized functions ψ_j^β corresponding to the irreducible representations β of the Y_h group. We write the initial perturbation operator as $\hat{\mathbf{W}} = \hat{\mathbf{W}}_K + \hat{\mathbf{W}}_E$, where $\hat{\mathbf{W}}_K$ and $\hat{\mathbf{W}}_E$ are the operators of the interaction of an off-center ion with, respectively, the intracrystalline electric field and an external electric field. The operator $\hat{\mathbf{W}}_K$ is invariant under all operations of the Y_h group, and $\hat{\mathbf{W}}_E$ has the form $\hat{\mathbf{W}}_E = -\hat{\mathbf{d}}\cdot\mathbf{E}$, where $\mathbf{E}$ is the sum of the external electric fields $\mathbf{E}_0 + \mathbf{e}$ acting on the off-center ion in the general case. Using the perturbation-matrix method[31,32] and the matrices of the irreducible representations of the Y_h group (Ref. 33), we find the nonzero perturbation matrices $\mathbf{M}(\beta\times\beta')$ of the operator $\hat{\mathbf{W}}$ for all pairs of the irreducible representations β. The results of the calculations are systematized in Appendix A. The off-diagonal matrices $(\beta\neq\beta')$ in Appendix A do not contain the matrix elements of the operator $\hat{\mathbf{W}}_K$, since it is invariant under all transformations of the group. Its nonzero matrix elements are present only in square matrices of the type $\mathbf{M}(\beta\times\beta)$, where they occupy positions along the principal diagonal and where all the matrix elements for each irreducible representation are equal. The matrices $\mathbf{M}(\beta\times\beta)$ do not contain the matrix elements of the operator $\hat{\mathbf{W}}_E$, since $\hat{\mathbf{d}}$ is an odd operator with respect to inversion.

The matrix Hamiltonian sought can be written in the form of a combination of perturbation matrices:

TABLE II. Splitting of tunneling levels by a field $\mathbf{E}_0$ applied along symmetry directions.

Y_h	C_{5v}				C_{3v}			C_{2v}				C_s	
	A_1	A_2	E_1	E_2	A_1	A_2	E	A_1	A_2	A_3	A_4	A_1	A_2
	z		x,y		z		x,y	z	y		x	x,y	z
A_g	1	0	0	0	1	0	0	1	0	0	0	1	0
F_{1u}	1	0	1	0	1	0	1	1	1	0	1	2	1
F_{2u}	1	0	0	1	1	0	1	1	1	0	1	2	1
G_g	0	0	1	1	1	1	1	1	1	1	1	2	2
G_u	0	0	1	1	1	1	1	1	1	1	1	2	2
H_g	1	0	1	1	1	0	2	2	1	1	1	3	2
H_u	0	1	1	1	0	1	2	1	1	2	1	2	3
Orders of secular equations	5	1	6	6	7	3	10	10	7	6	7	17	13

Note: A—one-dimensional irreducible representations, *E*—two-dimensional irreducible representations. The letters x, y, and z below the irreducible representations indicate that the respective components of the polar vector belong to these representations.

$$\mathbf{M}_{30}=\begin{pmatrix} \mathbf{M}(1\times1) & \mathbf{M}(1\times2) & 0 & 0 & 0 & 0 & 0 & 0 \\ & \mathbf{M}(2\times2) & 0 & 0 & 0 & \mathbf{M}(2\times6) & \mathbf{M}(2\times7) & 0 \\ & & \mathbf{M}(3\times3) & \mathbf{M}(3\times4) & 0 & \mathbf{M}(3\times6) & \mathbf{M}(3\times7) & 0 \\ & & & \mathbf{M}(4\times4) & \mathbf{M}(4\times5) & 0 & 0 & \mathbf{M}(4\times8) \\ & & & & \mathbf{M}(5\times5) & \mathbf{M}(5\times6) & \mathbf{M}(5\times7) & 0 \\ & & & & & \mathbf{M}(6\times6) & \mathbf{M}(6\times7) & \mathbf{M}(6\times8) \\ & & \text{c.c.} & & & & \mathbf{M}(7\times7) & \mathbf{M}(7\times8) \\ & & & & & & & \mathbf{M}(8\times8) \end{pmatrix}, \tag{2}$$

where c.c. stands for the complex conjugates, and a zero indicates that all the matrix elements of the block are zeros. The rows and columns in (2) have been numbered according to the irreducible representations in the following order: A_g, F_{1u}, F_{2u}, G_g, G_u, $H_g^{(1)}$, $H_g^{(2)}$, and H_u. In Eq. (2) there are 12 parameters $\alpha_{\beta\beta'}$ describing the effect of the external field $\mathbf{E}$ and 9 parameters $\lambda_{\beta\beta}$ characterizing the intracrystalline electric field.

It is impossible to find exact analytical expressions for the eigenvalues ε of the operator (2) in a general form for an arbitrary value and orientation of the field $\mathbf{E}_0$. Hence it is advisable to study approximate solutions by perturbative techniques in the presence of degeneracy[34] for two possible cases, viz., $\hat{\mathbf{W}}_K \gg \hat{\mathbf{W}}_E$ and $\hat{\mathbf{W}}_E \gg \hat{\mathbf{W}}_K$. For symmetry directions of the field $\mathbf{E}_0$ the secular equation corresponding to (2) splits into equations of lower order. The numbers in the last row of Table II indicate the order of such equations, which correspond to specific irreducible representations of the subgroup. In the case of two-dimensional representations there are two identical equations. For instance, for $\mathbf{E}_0$ oriented along the z axis, the Y_h symmetry reduces to C_{5v}, and the secular equation splits into four sixth-order equations (two for each of the irreducible representations E_1 and E_2), one fifth-order equation for the A_1 irreducible representation, and one first-order equation for the A_2 irreducible representation:

$$A_1 : \varepsilon\{\varepsilon^4 - \varepsilon^2(E_{12}^2 + E_{26}^2 + E_{27}^2 + E_{27}^2 + E_{36}^2 + E_{37}^2)$$
$$+ E_{12}^2(E_{36}^2 + E_{37}^2) + E_{37}^2(E_{26}^2 + E_{36}^2)$$
$$- 2E_{26}E_{27}E_{36}E_{37}\} = 0,$$

$$A_2 : \varepsilon = 0, \text{ where } E_{\beta\beta'} = \alpha_{\beta\beta'}E_{0z}. \tag{3}$$

The third-order equations with respect to ε^2 for the E_1 and E_2 irreducible representations are more complicated, and for the sake of brevity we shall not present them here. In deriving Eqs. (3), to simplify matters we set $\lambda_{\beta\beta}=0$, which corresponds to the zeroth approximation for the case where $\hat{\mathbf{W}}_E \gg \hat{\mathbf{W}}_K$.

When $\hat{\mathbf{W}}_K \gg \hat{\mathbf{W}}_E$, the zeroth approximation is determined by the values of $\lambda_{\beta\beta}$ with $\beta \neq 6, 7$ and the expression $(\lambda_{66}+\lambda_{77})/2 \pm \sqrt{(\lambda_{66}-\lambda_{77})^2/4+\lambda_{67}^2}$. For this case, using the data of Appendix A, we can calculate exactly the squares of the matrix elements of the transitions within the system of levels depicted in Fig. 2a. We have e_{12}^2, $(5/2)e_{26}^2$, $(5/2)e_{27}^2$, $(5/3)e_{36}^2$, $(5/3)e_{37}^2$, $(5/2)e_{56}^2$, $(5/2)e_{57}^2$, $2e_{34}^2$, $4e_{45}^2$, $10e_{68}^2$,

$10e_{78}^2$, and $(5/2)e_{48}^2$, respectively, for the $1\leftrightarrow2$, $2\leftrightarrow6$, $2\leftrightarrow7$, $3\leftrightarrow6$, $3\leftrightarrow7$, $5\leftrightarrow6$, $5\leftrightarrow7$, $3\leftrightarrow4$, $4\leftrightarrow5$, $6\leftrightarrow8$, $7\leftrightarrow8$, and $4\leftrightarrow8$ transitions, where $e_{\beta\beta'} = \alpha_{\beta\beta'}e$ with $e^2 = e_x^2 + e_y^2 + e_z^2$.

4. SYMMETRIZED FUNCTIONS AND PARAMETERS OF THE HAMILTONIAN

We use what is known as the well approximation, which is the oscillatory analog of the MO-LCAO method in quantum chemistry.[35] Let φ_i, $i=1,2,\ldots,30$, be the well function of the particle in the ith quasiequilibrium state (Fig. 1). The functions φ_i obey the reducible representation Π_{30}, which specifies the transformation of these functions into one another under the operations of the Y_h group. Now we expand Π_{30} in the irreducible representations[32,36] using the matrices of the irreducible representations.[33] As a result we arrive at the linearly independent, symmetrized functions ψ_j^β which are normalized to unity:

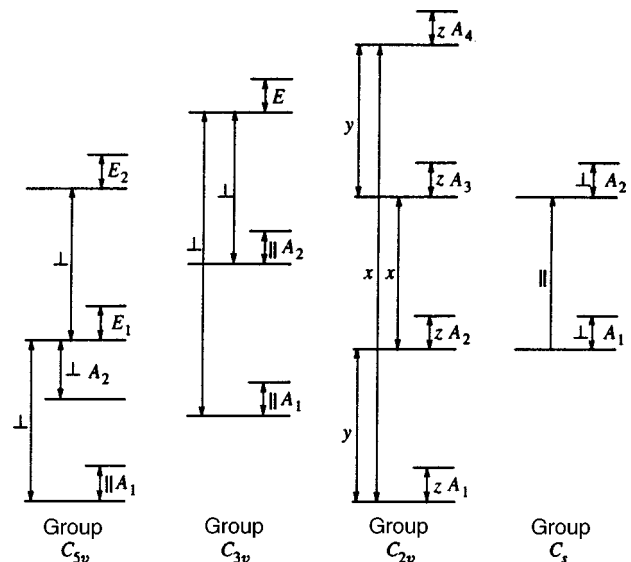$$\psi_j^\beta = D^\beta \sum_{i=1}^{30} k_{ij}^\beta \varphi_i. \tag{4}$$

FIG. 3. Allowed transitions between tunneling levels in the presence of an external constant electric field applied along symmetry directions.

The coefficients k_{ij}^{β} and D^{β} for each of the 30 functions are given in Appendix B. Of the functions (4) only ψ_j^6 and ψ_j^7 are not mutually orthogonal:

$$\int (\varphi_j^6)^* \varphi_j^7 d\tau = \frac{\Sigma_{m=0}^8 n_m' \sigma_m}{\sqrt{S_6 S_7}},$$

where $n_{0,8}' = -2$, $n_{1,7}' = 3p-1$, $n_{2,5}' = 2$, $n_{3,6}' = -(3p+1)$, and $n_4' = 4$. For $\lambda_{67} \neq 0$, a linear combination of these functions determines the regular (mutually orthogonal) zeroth-approximation wave functions of the terms $H_g^{(1)}$ and $H_g^{(2)}$.

Plugging (4) into the expression for $\lambda_{\beta\beta}$, we obtain

$$\lambda_{\beta\beta} = \frac{\Sigma_{m=0}^8 n_m^\beta V_m}{S_\beta},$$

where $V_{0-8} = \int \varphi_1^* \hat{\mathbf{W}}_K \varphi_i d\tau$, with $i = 1, 2, 3, 7, 8, 14, 15, 23$, and 29 for $m = 0, 1, \ldots, 8$, respectively;

$$\lambda_{67} = \frac{\Sigma_{m=0}^8 n_m' V_m}{\sqrt{S_6 S_7}}.$$

Similarly, for the integrals $\alpha_{\beta\beta'} = \int \varphi_j^{\beta*} \hat{\mathbf{d}}_z \varphi_j^{\beta'} d\tau$ we have

$$\alpha_{\beta\beta'} = -2D^\beta D^{\beta'} \sum_{m=0}^{12} R_m^{\beta\beta'} \mu_m. \tag{5}$$

Here $R_m^{\beta\beta'}$ is a linear combination of pair products of the coefficients k_{ij}^β, and μ_m are expressions of the type $\mu_{i,j} - \mu_{i,-j}$, where $\mu_{ij} = \int \varphi_i^* \hat{\mathbf{d}}_z \varphi_j d\tau$, and $-j$ denotes a well that is inversion-symmetric to the jth well by inversion. The matrix elements of $\hat{\mathbf{d}}_z$ that are related to the terms $H_g^{(1)}$ and $H_g^{(2)}$ are specified by linear combinations of two expressions of type (5).

Let us estimate the overlap integrals σ_m and the average dipole moments μ_m. For the well function we take an oscillatory function normalized to unity of the type $\varphi_i = (l/\pi)^{3/4} \times \exp(-lr_i^2/2)$, where $r_i^2 = (x-x_i)^2 + (y-y_i)^2 + (z-z_i)^2$, x_i, y_i, and z_i are the coordinates of the center of the ith well, and $l = m\omega/\hbar$ with m denoting the mass of an off-center ion and ω denoting the frequency of oscillations in the potential well. Assuming that $r_{\min} = 1.4$ Å, $m = 1.2 \times 10^{-23}$ g, $\omega = 3 \times 10^{12}$ rad/s (Ref. 37), and $\hat{\mathbf{d}} = z$, we obtain, e.g., for the maximum values of σ_m and μ_m, the following estimates: $\sigma_1 = 0.37$ and $\mu_0 = \mu_{1,1} = 0.85$. For comparison we note that $\sigma_0 = 1$ and $\mu_{12} = \mu_{1,12} - \mu_{1,17} = 0.033$.

Next we consider the tunneling approximation,[9,10] where all the off-diagonal matrix elements of the operator $\hat{\mathbf{W}}_K$ and the overlap integrals are assumed small and where the only integrals that are retained are those corresponding to the nearest wells. We also assume that all the off-diagonal matrix elements of $\hat{\mathbf{W}}_E$ are equal to zero. In this approximation, $\lambda_{\beta\beta} = V_0 + \alpha_\beta U$ and $\lambda_{67} = -[V_0 + (3p-1)(V_0\sigma_1/4 - V_1)/2]/2$, where $U = V_0\sigma_1 - V_1$ and $\alpha_\beta = -4, -(p+1)$, $p-1, 1, -1, (3p+1)/4, -1/2$, and 2 for $\beta = 1,2,\ldots,8$, respectively. The energy levels ε_β corresponding to the irreducible representations that are contained in (1) only once are given by the following equality: $\varepsilon_\beta = \lambda_{\beta\beta}$. For $\beta = 6$ and 7, solving the appropriate second-order equation, we find that $\varepsilon_{6,7} = V_0 \pm 2U$. This scale was used in building the

tunneling-level diagram in Fig. 2a. For the transition frequencies ν_{34}, ν_{21}, ν_{63}, ν_{57}, ν_{84}, ν_{72}, ν_{45}, ν_{65}, ν_{37}, ν_{87}, and ν_{62}, where $\nu_{kk'} = \varepsilon_k - \varepsilon_{k'}$, we have, respectively (in units of U), $p-2$, $3-p$, $3-p$, 1, 1, $p-1$, 2, 3, $p+1$, 4, and $3+p$. These values have been plotted in Fig. 2b. Note that in this approximation the levels of H_u and $H_g^{(1)}$ coincide (accidental degeneracy) and there are two pairs of coincident frequencies. Keeping only the local electric dipole moment μ_0 in (5) and using the relations for the squares of the transition matrix elements in zero-field paraelectric resonance, we can also obtain expressions for the line intensities $I_{\beta\beta'}$ in the tunneling approximation. These data (in units of $e^2\mu_0^2$) are depicted in Fig. 2b together with the corresponding transition-frequency data. Finally, we note that these parameters also appear in the equations that determine the energy structure of an off-center ion in the presence of an external static electric field [e.g., Eq. (3)]. Hence they can also be used to estimate the dependence of the position of the levels on the magnitude and direction of $\mathbf{E}_0$.

5. MICROWAVE SPECTROSCOPY OF OFF-CENTER IONS

The number of possible displacements of an off-center ion in various clusters with icosahedral symmetry is limited (countable). Here, irrespective of the nature of the clusters (icosahedron, dodecahedron, roklar, or fullerene), by virtue of their symmetry alone, there is a finite number of sets of potential wells that are the same for all types of clusters. Hence solving the problem of the number and nature of the possible tunneling levels, the allowed quantum transitions between these levels, and the paraelectric resonance spectrum for any type of cluster amounts to examining only a few cases, each of which relates to a certain set (number) of potential wells. Having a relatively small set of solutions in the form of paraelectric resonance spectra and comparing each solution belonging to this set with experimental data, we can acquire knowledge about the nature of the off-center effect and hence about the values of the parameters of the energy structure and directions of displacement of the ions.[9,10]

This approach is similar to the one used in electron paramagnetic resonance (EPR) studies, where, in view of the fact that the value of spin S is bounded (usually $S \leq 7/2$), the number of magnetic sublevels is finite, and the problem amounts to finding S and the spin-Hamiltonian parameters from a comparison of the experimental data with the theoretical EPR spectra obtained as a result of diagonalizing a finite set of spin Hamiltonians. The similarity to EPR becomes even more striking if we introduce an effective spin S_{eff} using the well-known formula $2S_{\mathrm{eff}} + 1 = n$, where n is the total multiplicity of the degeneracy of all the tunneling levels, i.e., the number of potential wells. In this case we can say that the problem has been reduced to determining S_{eff} and the parameters of the spin Hamiltonian (2) from experimental data.

We see that to study off-center effects we must have prior knowledge of the theoretical paraelectric resonance spectrum corresponding to each set of potential wells from the available finite collection. This was accomplished in the present work. In it we have studied the extremely compli-
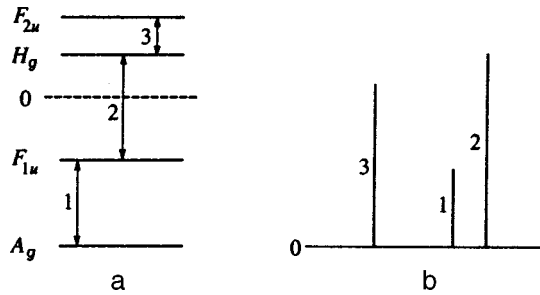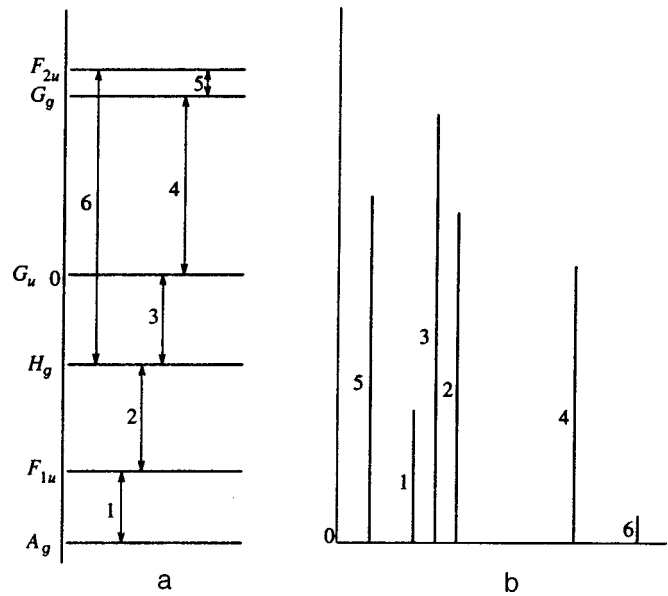
FIG. 4. Tunneling levels in the case of a 12-well potential corresponding to displacements of an off-center ion toward the vertices of an icosahedron, the centers of the faces of a dodecahedron, and the centers of the pentagonal faces of a roklar and a fullerene. The meaning of (a) and (b) is the same as in Fig. 2.

cated case of a 30-well potential. The same method was used to study two other cases that are relatively simple compared to the 30-well case: 12- and 20-well potentials. The former corresponds to displacements of an off-center ion from the center toward the vertices of an icosahedron, the centers of the faces of a dodecahedron, and the pentagonal faces of a roklar and a fullerene. The latter corresponds to displacements of an off-center ion toward the vertices of a dodecahedron, the trigonal faces of an icosahedron and a roklar, and the hexagonal faces of a fullerene. However, irrespective of the specific configuration of the cage atoms, the multiwell potential is an icosahedron in the former case and a pentagonal dodecahedron in the latter. The respective expansions of the reducible representations are

$$\Pi_{12} = A_g + F_{1u} + F_{2u} + H_g,$$
$$\Pi_{20} = A_g + F_{1u} + F_{2u} + G_g + G_u + H_g. \qquad (6)$$

A comparison of (1) and (6) shows that these two cases can be considered special cases of the scheme studied in detail in this paper. For instance, the energy structure corresponding to (6) is contained in the diagram of Fig. 2a. The results of similar calculations for 12- and 20-well potentials are depicted in Figs. 4 and 5, respectively. A comparison of Figs. 2, 4, and 5 shows that the energy structures and the paraelectric resonance spectra corresponding to different multiwell potentials differ substantially. On the basis of (6) and Table II we can easily obtain the analogs of Fig. 3 for 12- and 20-well potentials. All these figures are also different. The differences make it possible to identify the displacement directions for the off-center ions in different clusters. This collection of three sets of potential wells essentially exhausts the possible cases of the displacement of an off-center ion in any cluster and is sufficient for analyzing paraelectric resonance spectra. In other words, the theoretical results presented above can be related to a specific direction of displacement of an off-center ion in any cluster studied in experiments using the data on its paraelectric resonance spectra.

6. DISCUSSION OF RESULTS. CONCLUSIONS

1. Since we intend to apply our results to the analysis of the experimental data, we note the following. At least two

FIG. 5. Tunneling levels in the case of a 20-well potential corresponding to displacements of an off-center ion toward the vertices of a dodecahedron, the trigonal faces of an icosahedron and a roklar, and the hexagonal faces of a fullerene. The meaning of (a) and (b) is the same as in Fig. 2.

cases must be distinguished here. The first deals with orientationally disordered condensed substances, where the paraelectric centers have random orientations. Here the axes of the centers are distributed in space at random, and to describe the paraelectric resonance spectrum we must first average the transition frequencies (the resonant electric fields) over the angles characterizing the orientation of the axes of the complexes relative to the coordinate system associated with the external electric fields. Such averaging leads to additional, inhomogeneous broadening of the paraelectric resonance lines, as in EPR.[38] The second corresponds to similar orientation of centers that are at equivalent or slightly inequivalent positions. This is possible in some supercooled liquids or crystals of the fullerite type. In this case all the expressions considered above for the transition frequencies manifest themselves directly, and the broadening of lines is caused by other factors, among which the most important for paraelectric resonance is the defect structure of the condensed phase. Note, however, that at $\mathbf{E}_0 = 0$ the zero-field paraelectric resonance spectra in both cases coincide, since the spectra are independent of the orientation of the field **e** inducing the transitions. This fact broadens the class of substances that can be studied, adding the gaseous phase, solutions, and powders.

2. If we compare our results with those obtained earlier for different cluster configurations and their symmetries,[9–11] we will see that in the case of configurations with icosahedral symmetry, new, previously undetected, features emerge in the energy structure, the resonant transitions between tunneling levels, and paraelectric resonance spectra. An important feature of icosahedral symmetry even in comparison to the O_h group is the abundance of equivalent potential minima along with the resultant diversity and large number of tunneling levels and possible transitions between such levels for both $\mathbf{E}_0 = 0$ and $\mathbf{E}_0 \neq 0$. We also note that the multi-

well nature of the potential results, as Eq. (1) clearly shows, in a situation in which the energy structure incorporates several identical irreducible representations, which from a purely theoretical viewpoint leads to additional features and complications. Such situations have never been encountered before.

3. We stress that all possible realizations of paraelectric resonance can be described on the basis of the generalized Hamiltonian (2), which contains the maximum possible number of theoretical parameters allowed by space symmetry and time-reversal symmetry. Here the relationships for the frequencies, intensities, and other characteristics of the spectral lines become much simpler if one uses the tunneling approximation, which is advisable when the first experiments are described. The Hamiltonian (2) can serve as a basis for studying other properties of matter that are determined by the presence of off-center impurities in endohedral complexes. Among these, polarization effects should be mentioned first.

4. If the off-center ions have magnetic properties, i.e., their nuclei or electrons have nonzero spins, additional possibilities of studying them emerge. In this case, in addition to paraelectric resonance, one can use NMR or EPR spectroscopy.[9–11] The off-center nature of the ions manifests itself in such resonances in a decrease in local symmetry as the temperature drops and, as a result, in an increase in the number of spin-Hamiltonian constants, in changes in the hyperfine structure of the spectrum, and other effects. There have been reports about magnetic resonances in endohedral fullerenes. For instance, Shinohara *et al.*[39] and Kato *et al.*[40] used EPR spectroscopy to study La and Sc impurities in the fullerene C_{82}. They believe that the impurity atoms are triply positively charged and have a total angular momentum $J = 1/2$. A hyperfine structure consisting of eight lines emerges because of the interaction with ^{139}La and ^{45}Sc nuclei (the spin of each is 7/2) and is clearly visible. Independently, Moro *et al.*[41] found that Gd and Eu atoms in C_{60} are in the Gd^{3+} and Eu^{2+} states, which are characterized by a high angular momentum ($J = 7/2$). These data point to the possibility of detecting EPR on high-spin ions in fullerenes, in particular, in C_{60}. When the off-center ions become displaced, the symmetry of their environments lowers to C_{5v}, C_{3v}, C_{2v}, or C_s, depending on the direction of displacement. The spin Hamiltonians for these subgroups differ significantly from each other and from the spin Hamiltonians of the Y_h group,[33] making it possible, in addition, to establish the occurrence of a displacement and to determine the displacement direction. Estimates show[8] that a Gd^{3+} ion implanted in C_{60} is displaced from its position of equilibrium by 1.4 Å, while the other high-spin ion, Eu^{2+} is located at the center of C_{60}. High-spin ions are most suitable for detecting EPR in highly symmetric systems.[33,42] Using Gd^{3+} and Eu^{2+} for this purpose permits the study of both the position of ions displaced from the center of C_{60} and their central position.

5. As for the choice of a specific system for studying the effects described in this paper, we note, first, that by 1994 almost one-third of the elements of the periodic table had been used in the formation of endohedral complexes,[1] and the number constantly grows. At the same time, theoreticians considered the possibility of the existence of stable complexes when atoms and molecules are implanted in the respective cages (see, e.g., Refs. 8 and 18). In particular, Son and Sung[43] showed that along with endohedral complexes there can be stable exohedral complexes, in which an atom (ion) attaches itself to the cage on the outside and occupies one of a group of equivalent symmetry positions. Tunneling transitions of such external atoms between these positions with the resultant formation of a system of tunneling levels are possible. From symmetry considerations, the number of sets of equivalent positions for external atoms coincides with the number of such sets for atoms tunneling inside the cluster. Hence our results can be applied to exohedral complexes as well. One of the features of impurity atoms is their ability to stabilize the complex[13,27,44] even when the empty cage in unstable. This fact increases the number of endohedral complexes with the icosahedral symmetry considered in this paper.

The fullerene C_{60} occupies a central place in studies of the possibility of off-center effects occurring in endohedral complexes with icosahedral symmetry. Here such effects can be expected for the following implanted donor atoms (ions): Li^+, Na^+, Mn^{2+}, Ca^{2+}, Sr^{2+}, Yb^{2+}, Y^{3+}, La^{3+}, Ce^{3+}, Pr^{3+}, Nd^{3+}, Pm^{3+}, Gd^{3+}, Ac^{3+}, Th^{3+}, and U^{3+}; among the acceptor atoms (ions), F^- merits attention; and among the inert gases, He and Ne should be mentioned. Among the higher fullerenes we should mention C_{80} with a La_2 impurity molecule, which stabilizes the entire complex, imparting icosahedral symmetry to its cage. In this case one should expect orientational tunneling, which was observed earlier in the case of OH^- and CN^- impurity molecules in alkali-metal halide crystals.[9,10] Other complexes with icosahedral symmetry include B_{12}, Si_{12}, Co_{12}, and $C_{20}H_{20}$. Some of these have a very large cage diameter (e.g., in Si_{12} it amounts to approximately 5 Å), which is favorable for the appearance of off-center effects. For instance, estimates show that when V, Cr, Fe, or Ni atoms are implanted in the Co_{12} cluster, they are displaced from the center of the icosahedron due to the Jahn–Teller effect.

Note, however, that these data, which were taken from different sources, can be used only as estimates and sometimes contradict each other. For instance, according to Son and Sung,[43] the ions of He and Ne may move away from their centrosymmetric positions, while according to Pang and Brisse[17] and Breton *et al.*,[18] the equilibrium positions of these ions are at the fullerene's center. According to Wang *et al.*,[8] on the one hand, and Joslin *et al.*[12] and Ballester and Dunlap,[14] on the other, the displacements of Li^+ differ by a factor of seven. There are also discrepancies in the data concerning the possibility of the displacement of the Na^+ ion from the center of C_{60} (see Refs. 8 and 14). And although such discrepancies do not occur very often, they are a direct indication of the complexity of microcalculations of the structure and properties of endohedral complexes and of the need for direct experiments to detect and study off-center effects. Paraelectric-resonance experiments are prime examples of such experiments.

APPENDIX A

The matrices of the perturbation operator $\hat{\mathbf{W}}$ are

$$\mathbf{M}(1\times 2)=\frac{\alpha_{12}}{q}(P \quad qE_z \quad -P^*),$$

$$\mathbf{M}(2\times 6)=\frac{\alpha_{26}t}{2}\begin{pmatrix} P & E_z & -P^*/s & 0 & 0 \\ 0 & P/q & 2E_z/t & -P^*/q & 0 \\ 0 & 0 & P/s & E_z & -P^* \end{pmatrix},$$

$$\mathbf{M}(3\times 6)=\frac{\alpha_{36}}{2t}\begin{pmatrix} E_z & P^* & 0 & 0 & P \\ 0 & -P/q & tE_z & P^*/q & 0 \\ -P^* & 0 & 0 & P & -E_z \end{pmatrix},$$

$$\mathbf{M}(5\times 6)=\frac{\alpha_{56}}{4}\begin{pmatrix} -3P^* & 0 & 0 & 2P & -2E_z \\ -P & 4E_z & sP^* & 0 & 0 \\ 0 & 0 & -sP & 4E_z & P^* \\ 2E_z & 2P^* & 0 & 0 & -3P \end{pmatrix},$$

$$\mathbf{M}(3\times 4)=\frac{\alpha_{34}}{4}\begin{pmatrix} P & P^* & 0 & -2E_z \\ 0 & qP & qP^* & 0 \\ 2E_z & 0 & -P & -P^* \end{pmatrix},$$

$$\mathbf{M}(4\times 5)=\alpha_{45}\begin{pmatrix} -E_z & 0 & -P & 0 \\ 0 & -E_z & 0 & P \\ -P^* & 0 & E_z & 0 \\ 0 & P^* & 0 & E_z \end{pmatrix},$$

$$\mathbf{M}(6\times 8)=\alpha_{68}\begin{pmatrix} -2E_z & P^* & 0 & 0 & 0 \\ P & -E_z & sP^*/2 & 0 & 0 \\ 0 & sP/2 & 0 & sP^*/2 & 0 \\ 0 & 0 & sP/2 & E_z & P^* \\ 0 & 0 & 0 & P & 2E_z \end{pmatrix},$$

and $\mathbf{M}(\beta\times\beta)=\lambda_{\beta\beta}\mathbf{I}$, where $\mathbf{I}$ is the identity matrix, $P=E_x+iE_y$, $q=\sqrt{2}$, $t=\sqrt{3}$, and $s=\sqrt{6}$. The presence of a common factor in front of a matrix means that all matrix elements must be multiplied by it. Also,

$$\alpha_{12}=-\int (\psi_1^{(1)})^*\hat{\mathbf{d}}_z\psi_2^{(2)}d\tau,$$

$$\alpha_{26}=-\int (\psi_2^{(2)})^*\hat{\mathbf{d}}_z\psi_3^{(6)}d\tau,$$

$$\alpha_{36}=-\int (\psi_2^{(3)})^*\hat{\mathbf{d}}_z\psi_3^{(6)}d\tau,$$

$$\alpha_{56}=-\int (\psi_2^{(5)})^*\hat{\mathbf{d}}_z\psi_2^{(6)}d\tau,$$

$$\alpha_{34}=-\int (\psi_3^{(3)})^*\hat{\mathbf{d}}_z\psi_1^{(4)}d\tau,$$

$$\alpha_{45}=-\int (\psi_4^{(4)})^*\hat{\mathbf{d}}_z\psi_4^{(5)}d\tau,$$

$$\alpha_{68}=-\frac{1}{2}\int (\psi_5^{(6)})^*\hat{\mathbf{d}}_z\psi_5^{(8)}d\tau,$$

$$\lambda_{\beta\beta}=\int (\psi_i^{(\beta)})^*\hat{\mathbf{W}}_K\psi_i^{(\beta)}d\tau.$$

TABLE III. Coefficients for symmetrized functions of various irreducible representations (IR).

									k_{ij}^β							
	i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
		29	30	26	27	28	24	25	21	22	23	16	17	18	19	20
IR	j															
A_g	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
F_{1u}	1	$-c_7^*$	c_9	a_6	c_9^*	$-c_7$	$-c_{11}^*$	c_{12}	c_{12}^*	$-c_{11}$	$-a_1$	$-c_{21}^*$	b_3	c_{21}	c_{19}	c_{19}^*
	2	1	1	1	1	1	a_3	a_3	a_3	a_3	a_3	0	0	0	0	0
F_{2u}	1	c_8^*	$-c_{10}^*$	a_7	$-c_{10}$	c_8	$-c_{12}$	c_{11}	c_{11}^*	$-c_{12}^*$	a_1	c_{22}^*	b_2	$-c_{22}$	c_{20}^*	c_{20}
	2	1	1	1	1	1	$-a_4$	$-a_4$	$-a_4$	$-a_4$	$-a_4$	0	0	0	0	0
G_g	1	c_4	$-c_6$	1	$-c_6^*$	c_4^*	$-c_6^*$	c_4^*	c_4	$-c_6$	1	c_4^*	1	c_4	$-c_6^*$	$-c_6$
	2	c_6^*	$-c_4$	-1	$-c_4^*$	c_6	$-c_4^*$	c_6	c_6^*	$-c_4$	-1	c_6	-1	c_6^*	$-c_4^*$	$-c_4$
G_u	1	c_3^*	$-c_1^*$	$-b_6$	c_1	$-c_3$	c_5	$-c_2$	c_2^*	$-c_5^*$	$-b_1$	$-c_4^*$	1	$-c_4$	$-c_6^*$	c_6
	2	$-c_5$	$-c_2^*$	b_1	c_2	c_5^*	$-c_3$	$-c_1^*$	c_1	c_3^*	$-b_6$	c_6	1	c_6^*	c_4^*	$-c_4$
H_{g1}	1	c_{17}^*	$-c_{18}^*$	a_8	$-c_{18}$	c_{17}	c_{16}	$-c_{15}$	$-c_{15}^*$	c_{16}^*	$-a_5$	$-c_{15}$	$-a_5$	$-c_{15}^*$	c_{16}	c_{16}^*
	2	c_{14}^*	$-c_{13}$	$-a_2$	$-c_{13}^*$	c_{14}	$-c_{30}^*$	c_{25}	c_{25}^*	$-c_{30}$	$-a_9$	$-c_{29}$	a_{10}	$-c_{29}^*$	c_{26}^*	c_{26}
	3	1	1	1	1	1	$-a_4$	$-a_4$	$-a_4$	$-a_4$	$-a_4$	a_3	a_3	a_3	a_3	a_3
H_{g2}	1	$-c_{15}^*$	c_{16}^*	$-a_5$	c_{16}	$-c_{15}$	$-c_{18}$	c_{17}	c_{17}^*	$-c_{18}^*$	a_8	$-c_{15}$	$-a_5$	$-c_{15}^*$	c_{16}	c_{16}^*
	2	$-c_{29}^*$	c_{26}	a_{10}	c_{26}^*	$-c_{29}$	$-c_{13}^*$	c_{14}	c_{14}^*	$-c_{13}$	$-a_2$	c_{25}	$-a_9$	c_{25}^*	$-c_{30}^*$	$-c_{30}$
	3	a_3	a_3	a_3	a_3	a_3	1	1	1	1	1	$-a_4$	$-a_4$	$-a_4$	$-a_4$	$-a_4$
H_u	1	$-c_{28}$	c_{24}	$-b_4$	$-c_{24}^*$	c_{28}^*	c_{27}^*	$-c_{23}$	c_{23}	$-c_{27}$	b_5	c_{13}	$-a_2$	c_{13}^*	c_{14}	$-c_{14}^*$
	2	$-c_{27}$	$-c_{23}^*$	b_5	c_{23}	c_{27}^*	c_{28}	c_{24}^*	$-c_{24}$	$-c_{28}^*$	b_4	$-c_{14}$	$-a_2$	$-c_{14}^*$	$-c_{13}^*$	c_{13}
	3	0	0	0	0	0	0	0	0	0	0	-1	1	-1	1	-1

Note: The coefficients for $i=1-15$ are listed. The coefficients for the numbers below these values of i coincide in the case of even irreducible representations and differ in sign in the case of odd irreducible representations.

TABLE IV. Values of B^β and n_m.

β	B^β	$n_{0,8}$	$n_{1,7}$	$n_{2,6}$	$n_{3,5}$	n_4
1	30	1	4	4	4	4
2	$5p$	$\pm(p-1)$	±4	$\pm2(p-1)$	$\pm2(3-p)$	0
3	$5p$	$\pm(p+1)$	∓4	$\pm2(p+1)$	$\mp2(3+p)$	0
4	30	1	-1	-1	-1	4
5	-30	±1	±1	∓3	∓1	0
6	10	4	$-(3p+1)$	$3p-1$	2	-8
7	10	4	2	$-(3p+1)$	$3p-1$	-8
8	-10	±1	∓2	0	±2	0

The lower index of a function denotes its ordinal number in the basis set of the corresponding irreducible representation, which is denoted by the upper index. Here and below, to make the notation convenient, we replace the standard notation of irreducible representations in the formulas by numbers: $A_g \to 1$, $F_{1u} \to 2$, $F_{2u} \to 3$, $G_g \to 4$, $G_u \to 5$, $H_g^{(1)} \to 6$, $H_g^{(2)} \to 7$, and $H_u \to 8$. Also, $d\tau$ is the volume element. The matrices $\mathbf{M}(2\times7)$, $\mathbf{M}(3\times7)$, $\mathbf{M}(5\times7)$, $\mathbf{M}(4\times8)$, and $\mathbf{M}(7\times8)$ coincide, respectively, with the matrices $\mathbf{M}(2\times6)$, $\mathbf{M}(3\times6)$, $\mathbf{M}(5\times6)$, $\mathbf{M}(5\times6)$, and $\mathbf{M}(6\times8)$ if we replace the multipliers in the latter by, respectively, α_{27}, α_{37}, α_{57}, α_{48}, and α_{78}, whose explicit expressions are similar to the respective $\alpha_{\beta\beta'}$. The expression for $\lambda_{\beta\beta}$ also incorporates two similar irreducible representations belonging, however, to different levels, viz., $\beta = 6 = H_g^{(1)}$ and $\beta = 7 = H_g^{(2)}$; here the index i in $\lambda_{\beta\beta}$ is any one of the indices within the irreducible representation β. The parameters $\lambda_{\beta\beta}$, α_{12}, α_{26}, α_{27}, α_{36}, α_{37}, and α_{34} are real, and the parameters α_{56}, α_{57}, α_{45}, α_{48}, α_{68}, and α_{78}, imaginary. The choice of the coordinate system is shown in Fig. 1.

APPENDIX B

In addition,

$$k_{i3}^{2,3} = -(k_{i1}^{2,3})^*, \quad k_{i3,4}^{4,5} = (k_{i2,1}^{4,5})^*,$$

$$k_{i5}^{6,7,8} = (k_{i1}^{6,7,8})^*, \quad k_{i4}^{6,7,8} = -(k_{i2}^{6,7,8})^*,$$

$D^\beta = (B^\beta S_\beta)^{-1/2}$, $S_\beta = \Sigma_{m=0}^8 n_m^\beta \sigma_m$, where $\sigma_m \neq 0$ are the different possible overlap integrals (they are presented below in decreasing order): $\sigma_{1-8} = \int \varphi_1^* \varphi_i d\tau$, where $i = 2$, 3, 7, 8, 14, 15, 23, and 29 (here and in other places in the text, for i we have given several ''representative'' wells for the respective σ_m) for $m = 1, 2, \ldots, 8$, respectively; $\sigma_0 = 1$.

$$a_1 = \frac{1}{q}, \quad a_2 = \frac{1}{s}, \quad a_{3,4} = \frac{(p\mp1)}{2}, \quad a_5 = 2a_2,$$

$$a_{6,7} = \frac{a_{3,4}}{q}, \quad a_8 = 4a_2, \quad a_{9,10} = \frac{(3p\mp1)}{(2s)},$$

$$b_{1,6} = \frac{i}{N^\mp}, \quad b_{2,3} = \frac{iR^\pm}{2}, \quad b_{4,5} = \frac{b_{2,3}}{t};$$

$$c_{1,2} = \frac{1}{3\pm p} + \frac{i}{qR^\pm}, \quad c_{3,5} = \frac{1}{(p\pm1)} + \frac{i}{qU^\pm},$$

$$c_{4,6} = \frac{1}{p\pm1} + \frac{iR^\pm}{2q}, \quad c_{7,8} = 1 + \frac{iN^\mp}{2q},$$

$$c_{9,10} = \frac{1}{q(3\pm p)} + \frac{iR^\mp}{4}, \quad c_{11,12} = \frac{c_{4,6}}{q},$$

$$c_{13,14} = \frac{c_{4,6}}{s}, \quad c_{15,16} = \frac{2c_{4,6}}{s}, \quad c_{17,18} = \frac{4c_{4,6}}{s},$$

$$c_{19,20} = \frac{(p+iN^\mp)}{2q}, \quad c_{21,22} = \frac{p}{q(p\pm1)} + \frac{iR^\pm}{4},$$

$$c_{23,24} = \frac{c_{19,20}}{t}, \quad c_{25,26} = \frac{(7\pm p)/q + i\sqrt{65\mp19p}}{4t},$$

$$c_{27,28} = \frac{c_{21,22}}{t}, \quad c_{29,30} = \frac{4\pm p + i\sqrt{25\mp2p}}{2s};$$

$$p = \sqrt{5}, \quad N^\pm = \sqrt{5\pm2p},$$

$$R^\pm = \sqrt{5\pm p}, \quad U^\pm = \sqrt{25\pm11p}.$$

*)E-mail: roitsin@roklar.semicond.kiev.ua

[1] A. V. Eletskiĭ, Usp. Fiz. Nauk **164**, 1007 (1994) [Phys. Usp. **37**, 927 (1994)]; Yu. A. Osip'yan, Vestn. Ross. Akad. Nauk **66**, 597 (1996).

[2] A. B. Roĭtsin, Priroda **8**, 10 (1993).

[3] V. M. Loktev, Fiz. Nizk. Temp. **18**, 217 (1992) [Sov. J. Low Temp. Phys. **18**, 149 (1992)].

[4] F. D. Weiss, J. L. Elkind, S. C. O'Brien *et al.*, J. Am. Chem. Soc. **110**, 4464 (1988).

[5] D. S. Bethune, R. D. Johnson, J. R. Salem, *et al.*, Mon. Nat. **1**, 67 (1993).

[6] A. V. Eletskiĭ and B. M. Smirnov, Usp. Fiz. Nauk **165**, 977 (1995) [Phys. Usp. **38**, 935 (1995)].

[7] J. Cioslowski and A. Nanayakkara, Phys. Rev. Lett. **69**, 2871 (1992).

[8] Y. Wang, D. Tomanek, and R. S. Ruff, Chem. Phys. Lett. **208**, 79 (1993).

[9] M. D. Glinchuk, V. G. Grachev, M. F. Deĭgen, *et al.*, *Electrical Effects in Microwave Spectroscopy* [in Russian], Nauka, Moscow (1981).

[10] U. Kh. Kopvillem and R. S. Saburova, *Paraelectric Resonance* [in Russian], Nauka, Moscow (1982).

[11] V. S. Vikhnin, in *Microwave Spectroscopy of Solids* [in Russian], A. B. Roĭtsin (ed.), Naukova Dumka, Kiev (1992), p. 258.

[12] C. G. Joslin, J. Yang, C. G. Gray *et al.*, Chem. Phys. Lett. **208**, 86 (1993).

[13] J. Cioslowski and E. D. Fleischmann, J. Chem. Phys. **94**, 3730 (1991).

[14] J. L. Ballester and B. I. Dunlap, Phys. Rev. A **45**, 7985 (1992).

[15] C. G. Joslin, J. Yang, C. G. Gray *et al.*, Chem. Phys. Lett. **211**, 587 (1993).

[16] G. Cardini, P. Procacci, P. R. Salvi *et al.*, Chem. Phys. Lett. **200**, 39 (1992).

[17] L. Pang and F. Brisse, J. Phys. Chem. **97**, 8562 (1993).

[18] J. Breton, J. Gonzales-Platas, and C. Girardet, J. Chem. Phys. **99**, 4036 (1993).

[19] B. Shanker and J. Applequist, J. Phys. Chem. **98**, 6486 (1994).

[20] A. C. Tang, F. Q. Huang, Q. S. Liu *et al.*, Chem. Phys. Lett. **227**, 579 (1994).

[21] A. C. Tang and F. Q. Huang, Phys. Rev. B **51**, 13 830 (1995); **52**, 17 435 (1995).

[22] A. C. Tang, F. Q. Huang, and R. Z. Lui, Phys. Rev. B **53**, 7442 (1996).

[23] I. D. Morokhov, V. I. Petinov, L. I. Trusov, and V. F. Petrunin, Usp. Fiz. Nauk. **133**, 653 (1981) [Sov. Phys. Usp. **24**, 295 (1981)].

[24] W. A. de Heer, Rev. Mod. Phys. **65**, 611 (1993).

[25] M. Brack, Rev. Mod. Phys. **65**, 677 (1993).

[26] Y. Jinlong, X. Chuanyun, X. Shangda, and W. Kelin, Phys. Rev. B **48**, 12 155 (1993).

[27] A. Bahel and M. V. Ramakrishna, Phys. Rev. B **51**, 13 849 (1995).

[28] D. Li and W. Y. Ching, Phys. Rev. B **52**, 17 073 (1995).

[29] I. B. Bersuker, V. Z. Polinger, and V. P. Khlopin, Theor. Chim. Acta **48**, 87 (1978).

[30] A. B. Roĭtsin, Fiz. Tverd. Tela (St. Petersburg) **35**, 2547 (1993) [Phys. Solid State **35**, 1262 (1993)].

[31] A. B. Roĭtsin, Fiz. Tverd. Tela (St. Petersburg) **37**, 1594 (1995) [Phys. Solid State **37**, 865 (1995)].

[32] A. B. Roĭtsin, *Some Applications of Symmetry Theory in Problems of Microwave Spectroscopy* [in Russian], Naukova Dumka, Kiev (1973).

[33] A. B. Roĭtsin, L. V. Artamonov, and A. A. Klimov, Zh. Éksp. Teor. Fiz. **110**, 1392 (1996) [JETP **83**, 767 (1996)].

[34] A. B. Roĭtsin, Ukr. Fiz. Zh. **19**, 1216 (1974).

[35] I. B. Bersuker, *Electronic Structure and Properties of Transition Metal Compounds: An Introduction to the Theory*, Wiley, New York (1996).

[36] G. Ya. Lyubarskii, *The Application of Group Theory in Physics*, Pergamon Press, Oxford (1960).

[37] B. I. Dunlap, J. L. Ballester, and P. P. Schmidt, J. Phys. Chem. **96**, 9781 (1992).

[38] A. B. Roĭtsin, in *Microwave Spectroscopy of Solids* [in Russian], A. B. Roĭtsin (ed.), Naukova Dumka, Kiev (1992), p. 89.

[39] H. Shinohara, H. Sato, Y. Saito *et al.*, J. Phys. Chem. **96**, 3571 (1992).

[40] T. Kato, Sh. Suzuki, K. Kikuchi *et al.*, J. Phys. Chem. **97**, 13 425 (1993).

[41] L. Moro, R. S. Ruoff, C. H. Becker *et al.*, J. Phys. Chem. **97**, 6801 (1993).

[42] A. B. Roĭtsin, A. A. Klimov, and L. V. Artamonov, Fiz. Tverd. Tela (St. Petersburg) **38**, 3190 (1996) [Phys. Solid State **38**, 1743 (1996)].

[43] M. S. Son and Y. K. Sung, Chem. Phys. Lett. **245**, 113 (1995).

[44] A. Kobayashi, Sh. Nagase, and T. Akasaka, Chem. Phys. Lett. **245**, 231 (1995).

Translated by Eugene Yankovsky
Edited by P. Shelnitz

Roughening of a Si(100) surface induced by the adsorption of oxygen near the solid-oxide nucleation threshold

V. D. Borman,[*)] Yu. Yu. Lebedinskiĭ, and V. I. Troyan

Moscow Institute of Engineering Physics, 115409 Moscow, Russia
(Submitted 12 November 1997)
Zh. Éksp. Teor. Fiz. **114**, 239–262 (July 1998)

An investigation of the processes on a Si(100) surface interacting with oxygen near the solid-oxide nucleation threshold using x-ray photoelectron spectroscopy and atomic-force microscopy is described. The nucleation threshold is the boundary between the temperature and oxygen pressure regions where a phase transition with the formation of a submonolayer oxide and a roughening transition caused by oxygen adsorption occur. Near the nucleation threshold, either a random rough relief or a quasiperiodic structure is formed on a surface coated with chemisorbed oxygen. The formation of the rough relief due to oxygen adsorption has been interpreted within the theory of phase transitions as a result of vacancy clustering. A model that allows one to describe the dynamics of processes on the surface near the nucleation threshold in qualitative and in some cases in quantitative terms has been suggested. © *1998 American Institute of Physics.* [S1063-7761(98)01707-7]

1. INTRODUCTION

Effects resulting from the interaction between oxygen and a silicon surface have attracted a lot of attention recently since these processes are interesting from the viewpoint of both fundamental research[1–6] and various technological applications.[7–10] It has been known[11–16] that the formation of a submonolayer oxide on a Si surface proceeds like a first-order phase transition. As the temperature increases near the oxide nucleation threshold, formation of the solid oxide is replaced by generation of the volatile oxide SiO.[6,15] The escape of oxygen from the surface with the volatile oxide leads to a lower chemisorbed-oxygen coverage θ_{ch} of the surface, and when θ_{ch} is lower than the threshold value θ_{th}, the oxidation phase transition is terminated.

Recent investigations indicated that the general scheme of processes on a silicon surface interacting with oxygen near the nucleation threshold can be rather complicated. For example, observations using scanning tunneling microscopy indicated that, during a short exposure to oxygen ($\theta_{ch} \approx 10^{-2}$) at room temperature, defects resembling vacancies, vacancy clusters, and adsorbed atoms from the crystal appear on the Si surface.[17,18] At high temperatures ($T = 800$–950 K) and low oxygen pressures ($p = 10^{-7}$–10^{-8} Torr), the surface of the terraces becomes rough on the atomic scale.[1,3,19] It is known[6] that only the desorbed volatile oxide SiO is generated on a silicon surface under these conditions. Therefore, it seems that a roughening transition occurs on the surface in the region of high temperatures and low pressures. This raises the question of whether the nucleation threshold is only the boundary of the (p,T) region where the solid oxide is produced or whether it is the boundary between the regions where the oxidation phase transition and the surface roughening transition take place.

It is generally accepted that vacancies are generated on a

Si surface as lattice atoms leave the surface within desorbed SiO molecules.[1] At $T = 300$ K, however, the rate of SiO desorption is negligible (the activation energy for SiO desorption is[6] $E = 3.5$ eV); therefore, the emergence of vacancies[17,18] at $T = 300$ K cannot be attributed to this process. We previously[20] suggested an alternative mechanism of vacancy generation. According to Ref. 20, the attraction between oxygen adatoms and vacancies leads to a decrease in the energy of vacancy formation (E_v) by ΔE ($\Delta E < 0$), which is proportional to the oxygen concentration ($|\Delta E| \propto \theta_{ch}$). Therefore, E_v decreases and may become closer to the temperature when the surface is exposed to oxygen and θ_{ch} increases. The number of vacancies in this case should increase spontaneously without overcoming an activation barrier (vacancy instability). The attractive interaction between the vacancies in a crystal with a sufficient vacancy concentration[21] can lead to the formation of vacancy clusters, which can be observed at $T = 300$ K.[17,18] The concentration of adsorbed oxygen, however, could not be measured in previous investigations of surface morphology using scanning tunneling microscopy and high-vacuum reflection electron microscopy.[2,3,17–19,22] It was suggested[23,24] that surface roughness can result from vacancy clustering, and that the relief amplitude is determined by the number of clustered vacancies in the near-surface layer at a given point on the surface.

It follows from the above statements that there is still no clear picture of the processes on a silicon surface interacting with oxygen. The nature of the solid-oxide nucleation threshold, the mechanism of oxide formation, the generation of vacancies, and surface roughening have also remained unclear.

This paper presents the results of an experimental investigation of the dynamics of the formation of a submonolayer oxide and a rough relief on a Si(100) surface near the nucle-

ation threshold. The rough relief was studied by atomic-force microscopy. The surface morphology was also monitored by scanning electron microscopy. The oxygen concentration on the surface was measured using a technique based on x-ray photoelectron spectroscopy. This technique allowed us to monitor the accumulation of chemisorbed oxygen and oxide production in real time during exposure of the surface to oxygen.

We found that sequential changes in the way in which oxygen interacts with silicon take place in a narrow temperature range of width $\Delta T \sim 30$ K as the temperature is increased in the (p,T) region near the nucleation threshold studied. The regime in which the solid oxide is produced is replaced by oxidation of the rough surface, and at higher temperatures the rough relief is formed in the absence of the oxide (Sec. 3). The position of the nucleation threshold in the (p,T) region studied was derived from measurements of the characteristic time (τ_{ox}) for formation of the solid oxide as a function of temperature. Near the nucleation threshold the onset of formation of the solid oxide is delayed until the oxygen exposure time t_{th}. When the exposure time $t < t_{th}$, chemisorbed oxygen is accumulated on the surface, and the time t_{th} increases with temperature. We also found that a rough relief is formed in the presence of adsorbed oxygen on the surface. Our results are in disagreement with earlier predictions[23,24] of a narrow roughness size distribution, since the relief observed is characterized by the presence of sparse quasiperiodic deep pits against a background of small-scale fluctuations of the surface height under certain conditions.

The processes on a silicon surface interacting with oxygen near the oxide nucleation threshold are analyzed in Secs. 4 and 5. The experimental data can be described in terms of a competition between the two phase transitions. These are a first-order phase transition with growth of a submonolayer oxide and a phase transition with surface roughening induced by oxygen adsorption. In analyzing the growth dynamics of the submonolayer oxide near the threshold and in determining the threshold temperature (pressure) (Sec. 4), we use the Volmer–Weber–Zel'dovich theory, taking into account the diffusive interaction between oxide islands and the blocking of the surface by them.[12,16] The dynamics of surface roughening is analyzed in terms of the theory developed previously[23,24] as a result of vacancy clustering (Sec. 5). The model of processes on the silicon surface near the oxide nucleation threshold allows us to explain the switchover between oxygen-silicon interaction regimes and the features of the surface relief discovered.

2. EXPERIMENTAL TECHNIQUES

The experiments were performed on an XSAM-800 x-ray photoelectron spectrometer. X-ray photoelectron spectra were excited by the K_α line from a Mg source (1283.6 eV). The spectrometer resolution at the Au $4f_{1/2}$ line was 0.9 eV. The energy scale was calibrated using the C $1s$ line (E_b = 284.6 eV). The sample was n-type Si(100) and had dimensions of $4 \times 8 \times 0.3$ mm^3. The sample surface was chemically cleaned by etching in a 5% HF solution immediately before setting it in the spectrometer chamber. Before

the measurements the sample was annealed in ultrahigh vacuum in the spectrometer chamber (10^{-9} Torr, 1300 K). After this processing, we had a Si(100) surface atomically clean within the sensitivity of x-ray photoelectron spectroscopy and Auger spectroscopy.

During the measurements in the temperature range 800–1200 K, the sample was heated by an electric current which was passed through a wafer of the silicon being tested with a thickness of ≈ 300 μm. The current stabilization system allowed us to maintain the sample temperature constant to within ± 1 K during the measurements. This accuracy of temperature stabilization was confirmed by the reproducibility of the measurements of the total oxygen line intensity versus exposure time performed at close temperatures differing by $\Delta T = 1$ K.

The technique for studying the kinetics of the initial oxidation of the Si(100) surface was based on measurements of the amplitude of the O$1s$ oxygen peak as a function of exposure time in oxygen in a real-time mode and has been described in detail elsewhere.[12] The relationship between the measured amplitudes of the O$1s$ (I_O) and Si $2p$ (I_{Si}) lines, on the one hand, and the oxygen concentration on the Si(100) surface, on the other, i.e., the calibration curve, was determined using the technique described in a previous publication.[12] Additional evidence in favor of the adequacy of our technique is provided by the agreement between our results and the data by Lutz *et al.*,[25] some of which were obtained under the same conditions as in our experiments.

The surface roughness due to exposure to oxygen was tested using a Solver P4-SPM atomic-force microscope. The spatial resolution attained in a normal direction was ± 0.3 nm, and the resolution in a plane was ± 1 nm. The angle between the sample surface and the (100) crystal surface was within 0.1°. This accuracy is confirmed by the image of the original surface using atomic-force microscopy (see Fig. 4a). Note that the actual spatial resolution of 0.3 nm attained in the normal direction did not allow us to detect atomic steps on the surface and 2D oxide islands. In order to study the relief formed, we transported samples processed in the ultrahigh-vacuum chamber of the electron spectrometer to the P4-SPM microscope. In the process, the sample surface was coated by an oxide film with thickness < 2 nm. The value 2 nm determined the possibility of studying short-wavelength fluctuations of the surface height on the relief formed. The surface morphology was also monitored using a DSM-960 scanning electron microscope.

In order to identify oxygen and silicon states in the initial oxidation stage, we used high-resolution spectra of the O$1s$ and Si$2p$ lines. As in earlier experiments,[11,13,14,25–28] we found that silicon atoms near the surface exist in four oxidation states: Si^{n+}, where $n = 1,2,3,4$ (Fig. 1, curves *a* and *b*). A comparison between the Si $2p$ spectra of clean and exposed surfaces (Fig. 1) shows that the latter contains a shoulder on the high-binding-energy side relative to the Si $2p$ line of clean Si (99 eV). According to Ref. 14 and Fig. 1, this shoulder is due to Si atoms in different oxidation states: Si^{1+}, Si^{2+}, Si^{3+}, and Si^{4+}. The dotted lines in Fig. 1 show the respective components of the Si $2p$ spectrum.[14] The oxygen spectrum around the O$1s$ line peaking at 531.7 eV also
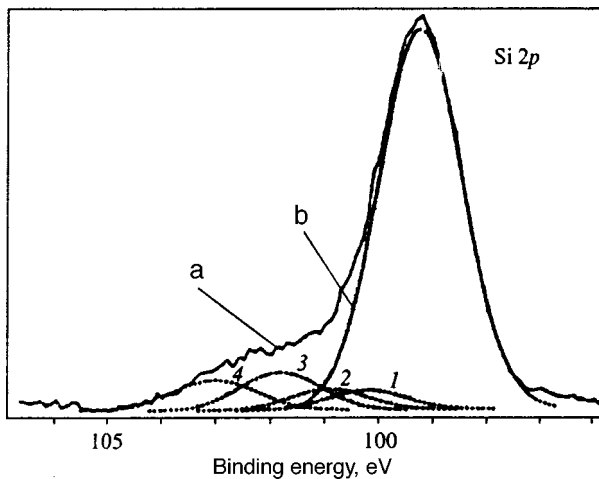
FIG. 1. X-ray photoelectron spectra of Si $2p$ on a Si(100) surface in different oxidation states Si^{n+} ($n = 1-4$, dotted curves *1*, *2*, *3*, and *4*) after exposure to oxygen ($t = 100$ s, $p = 10^{-6}$ Torr, $T = 921$ K) (curve a) and clean surface (curve b).

has a shoulder on the high-binding-energy side at $T = 300$ K. Decomposition of this spectrum into two Gaussian lines reveals that the shoulder is due to a peak corresponding to a binding energy of 533.5 eV with an amplitude four times smaller than that of the main peak at 531.7 eV. According to the interpretation suggested by Hollinger *et al.*,[29] these two peaks correspond to the oxide-like bridging oxygen and atomic nonbridging oxygen.

The Si $2p$ spectrum corresponding to the same exposure to oxygen ($\varepsilon = 100$ L, where 1 L $= 10^{-6}$ Torr·s) at room temperature is similar to the spectrum taken from a clean Si surface with a small contribution of Si^{1+} and Si^{2+} states. In other words, Si^{4+} states which correspond to pure SiO_2, do not appear at room temperature after relatively small exposures. At higher temperatures, the O$1s$ spectrum is an almost symmetrical peak at $E_b = 531.8$ eV with a small wing (hump) on the high-binding-energy side. At the same time, Si^{3+} and Si^{4+} states corresponding to a silicon oxide emerge in the Si$2p$ spectrum (Fig. 1). This is in agreement with the data obtained by Hollinger *et al.*[29] The numbers of oxygen atoms in the oxide and chemisorbed states were determined using a technique described elsewhere.[12] The technique is based on measurements of the intensity of the Si$2p$ line, which falls off owing to the reaction between oxygen and the silicon surface. At submonolayer values of the oxygen coverage θ, where the oxide film thickness is smaller than the mean free path of Si $2p$ photoelectrons, the intensity I_{Si} decreases linearly as θ increases. Assuming that the intensities of the Si_{2p}^{3+} and Si_{2p}^{4+} lines, which correspond to silicon oxide states (Fig. 1), drop because they are blocked by oxygen in the oxide state (θ_{ox}) and that the Si_{2p}^{1+} and Si_{2p}^{2+} lines are diminished by oxygen in the chemisorbed state (θ_{ch}), we can easily determine the ratio between the oxygen atoms in both the chemisorbed and oxide states. Thus, simultaneous measurements of the Si$2p$ and O$1s$ lines allow us to determine the fractions of oxygen in the oxide and chemisorbed states, as well as the number of silicon atoms in a silicon oxide.

3. EXPERIMENTAL RESULTS

Our experiments were performed in the region of p and T near the oxide nucleation threshold. The line of this threshold was determined empirically. In earlier experiments[30,31] the solid-oxide nucleation threshold was detected at relatively high temperatures $T > 1100$ K and pressures $p > 10^{-5}$ Torr. Lander and Morrison[30] found that the LEED pattern typical of a clean Si surface disappeared after the passage across this threshold. Later experiments demonstrated that volatile SiO is desorbed from the Si surface at temperatures near the threshold and that its flux increases with the temperature.[6] Therefore, the existence of this boundary is attributed to a decrease in the coverage of the surface with chemisorbed oxygen (θ_{ch}) to values below the threshold θ_{th} since oxygen atoms leave the surface within SiO molecules. This suggestion is supported by the temperature dependence $p_{th}(T) = p_0 \exp(-E_0/T)$, where[31] $E_0 = 3.80 \pm 0.2$ eV is close to the activation energy for the desorption of SiO, which equals[6,32] $E = 3.5 \pm 0.1$ eV. Hereafter the Boltzmann constant $k = 1$ in our formulas. Since the difference $\theta_{ch} - \theta_{th}$ drops near the nucleation threshold, one can expect that, in accordance with the theory of first-order phase transitions, the growth rate of the oxide phase should decrease, and the characteristic time τ_{ox} for the formation of a monolayer of the solid oxide on the surface should increase considerably. The position of the nucleation threshold was not determined in the present work in the region of pressures and temperatures investigated [$T = 850-950$ K, $p = (4-10) \times 10^{-7}$ Torr].

Figure 2 shows measurements of the total oxygen coverage $\theta = \theta_{ch} + \theta_{ox}$ in both the chemisorbed and oxide states on the Si surface at an oxygen pressure of 6×10^{-7} Torr as a function of exposure time in a narrow temperature range of width $\Delta T \approx 35$ K. It is clearly seen at the lowest temperature $T = 890$ K of this range that θ achieves a limiting value after a long exposure time. This temperature corresponds to the formation of an oxide monolayer on the surface ($\theta = \theta_{ox} \approx 1$, $\theta_{ch} \approx 0$), which is supported by observations of Si^{4+} and Si^{3+} states of oxidized silicon. Note that the sticking coefficient of oxygen molecules on a silicon oxide surface is three orders of magnitude lower than the value on a clean silicon surface; therefore, the total coverage observed after an exposure time $t = 800$ s at $T = 890$ K is determined only by the oxide, i.e., $\theta \approx \theta_{ox}$ and $\theta_{ch} \approx 0$. At the maximum temperature $T = 925$ K of the range studied, no oxide is produced on the surface during an exposure time of $t = 810$ s, and the final oxygen coverage on the surface is within 0.15 ($\theta_{ox} \lesssim 0.15$). Therefore, the temperature range 919–925 K, in which the transition from the regime of solid oxide monolayer formation to the chemisorption regime with a small value of θ_{ch} occurs, is the region of the solid-oxide nucleation threshold at $p = 6 \times 10^{-7}$ Torr. At higher pressures the threshold region shifts toward higher temperatures. This shift corresponds to the temperature dependence of the nucleation threshold pressure p_{th}.

The absence of the solid oxide on the surface at the maximum temperature T_{max} in this experiment can be ascribed either to the fact that $T_{max} > T_{th}$ at this oxygen pressure
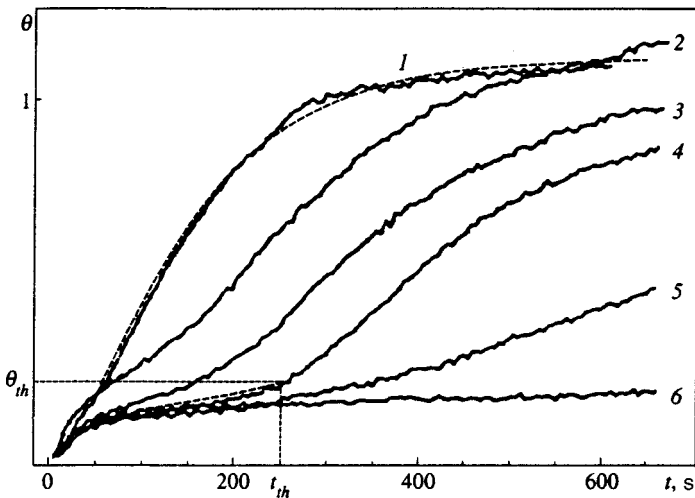
FIG. 2. Oxygen coverage $\theta = \theta_{ch} + \theta_{ox}$ of the Si(100) surface as a function of time near the oxide nucleation threshold ($p = 6 \times 10^{-7}$ Torr): (*1*) 890 K; (*2*) 906 K; (*3*) 910 K; (*4*) 912 K; (*5*) 915 K; (*6*) 925 K.

or to the insufficient measurement time in comparison with the characteristic time τ_{ox} for the formation of an oxide monolayer. Figure 3 shows the time τ_{ox}, in which the total oxygen coverage achieves $\theta = 0.7$, plotted against temperature. The graph clearly indicates that τ_{ox} increases rapidly as the temperature approaches $T_{th} = 905$ K. Therefore, the temperature $T_{th} = 905$ K can be defined as the threshold temperature for oxide nucleation at a pressure $p = 6 \times 10^{-7}$ Torr.

As follows from measurements of Si$2p$ and O$1s$ spectra at temperatures $902 < T < 920$ K, the oxygen on the surface exists in both the chemisorbed and oxidized states, and the surface is rough (Fig. 4). In this region, oxidation in an oxygen atmosphere leads to irreversible changes in the surface morphology. In repeated experiments with surface oxidation in the same regime after removing the submonolayer oxide by heating the sample to temperatures $T > 1200$ K in vacuum at $p = 10^{-9}$ Torr, the previously recorded $\theta(t)$ curves shown in Fig. 2 could not be reproduced.

Note that the temperature $T = 902$ K coincides within the experimental error with the temperature at which the time of oxide monolayer formation dramatically increases ($T = 905$ K in Fig. 3). Thus, there is a temperature range near the threshold in which the solid oxide is produced concurrently with the surface roughening. At lower temperatures only a solid oxide layer is formed on the surface, and at higher temperatures only the rough relief is formed under conditions for oxygen chemisorption.

Now let us discuss the changes in the oxygen accumulation dynamics on the surface as the temperature varies from its maximum to the minimum in the range studied near the nucleation threshold. Figure 2 clearly shows that at the minimum and maximum temperatures the oxygen coverage $\theta(t)$ is described by saturating curves. At the minimum temperature the $\theta(t)$ curves are characterized by the time τ_{ox} for oxide monolayer formation, and at the maximum temperature they are characterized by the time τ_{ch} for the achievement of a steady density of chemisorbed oxygen. At intermediate temperatures, the $\theta(t)$ curves have an additional characteristic time t_{th}, after which the rate of oxygen accumulation alters considerably. After exposure times longer than t_{th}, there are states on the silicon surface characteristic

of the solid oxide (Si^{3+} and Si^{4+}). For this reason, the oxygen coverage θ_{ch} at which the accumulation rate jumps can be considered to be the threshold coverage for the onset of the silicon-oxidation phase transition near the nucleation threshold. Note that the oxygen accumulation rate at the onset of oxide layer growth, $t \geqslant t_{th}$, depends weakly on temperature. As is shown by Fig. 2, t_{th} increases with temperature.

Atomic-force micrographs of the Si(100) surface after exposure to oxygen under the conditions of experiments near the nucleation threshold are shown in Fig. 4. For an initially smooth surface, the measured change in the surface height over a scanning length of 3×10^3 nm is 0.2–0.3 nm (Fig. 4a), which is within the spatial resolution of the instrument in the normal direction. The characteristic feature of the surface relief after an 800-s exposure to oxygen at $p = 6 \times 10^{-7}$ Torr and $T = 925$ K is a periodic pattern of sparse deep pits against a background of small-scale fluctuations of the surface height. The average pit depth is $h \approx 30$ nm, and the distance between them $l \approx 800$ nm. The rough surface relief turns up after exposure to oxygen throughout the re-
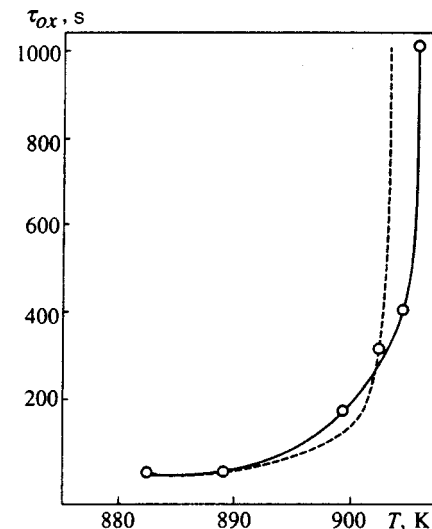
FIG. 3. Oxide growth time τ_{ox} as a function of temperature at $p = 6 \times 10^{-7}$ Torr. The dashed curve was constructed using Eq. (3).
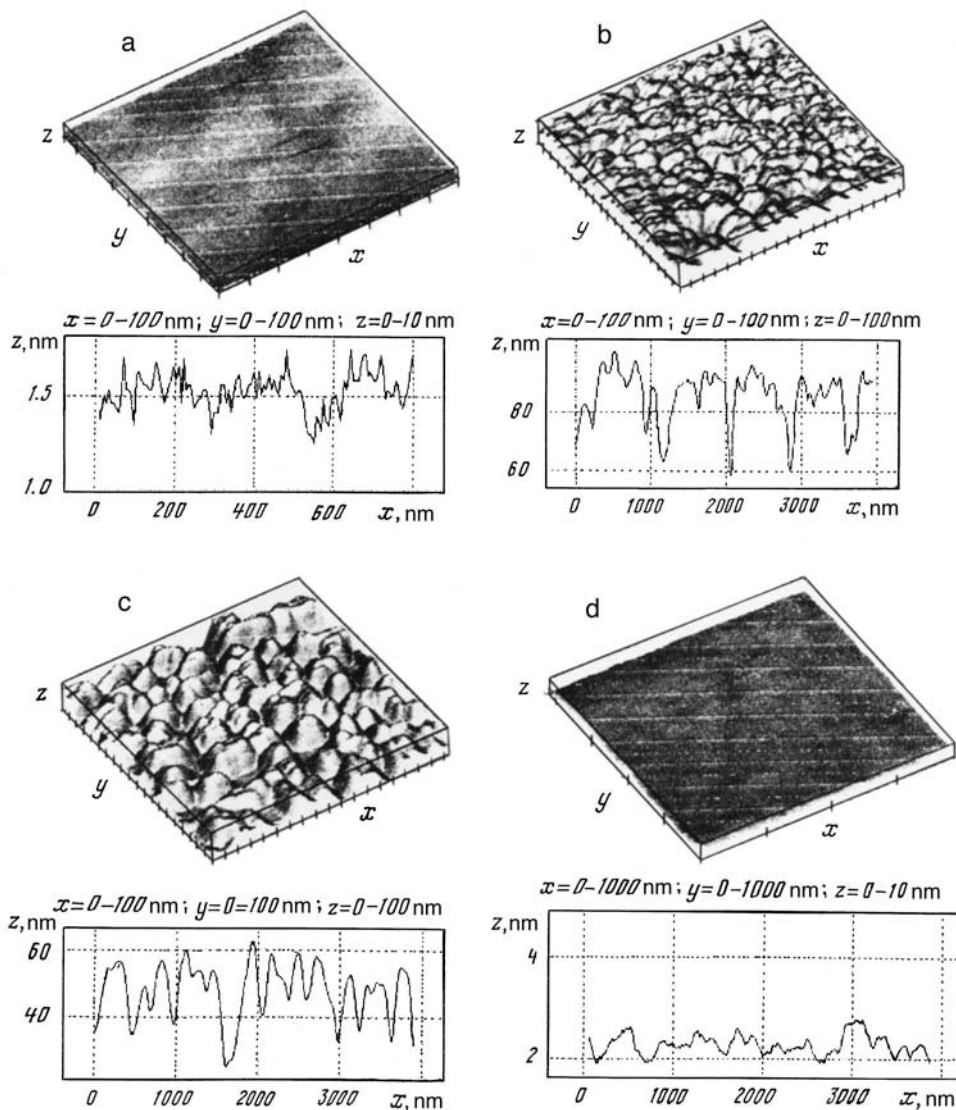
FIG. 4. Image obtained using an atomic-force microscope and profiles of rough Si(100) surfaces at an oxygen pressure $p = 6 \times 10^{-7}$ Torr: (a) initial surface; (b–d) after heating in oxygen at 915 K (b), 925 K (c), and 945 K (d).

gion of parameters tested near the nucleation threshold (see Fig. 3). The characteristic relief amplitude over a plane with dimensions of 4×4 μm is <30 nm, and the characteristic linear dimension along the plane is <250 nm. At higher temperatures, $T > 940$ K, the surface remains smooth within the experimental uncertainty after exposure to oxygen (Fig. 4).

Hence it follows that the regimes of interaction between oxygen atoms and a Si(100) surface change sequentially near the oxide nucleation threshold as the temperature increases. The regime in which the solid oxide is produced on the surface is replaced by the regime of oxide formation on the rough surface, and at higher temperatures only surface roughening in the presence of chemisorbed oxygen takes place. Note also that, near the threshold, the oxide is not generated immediately when the surface is exposed to oxygen, but only after a certain delay t_{th}. Until this moment, oxygen is accumulated on the surface only in the chemisorbed state.

4. GROWTH OF A SUBMONOLAYER OXIDE NEAR THE OXIDE NUCLEATION THRESHOLD

The growth of a submonolayer oxide has been studied previously[12,16] within the phenomenological Volmer–Weber–Zel'dovich theory of first-order phase transitions. According to the concepts based on this theory, a submonolayer oxide grows in the form of 2D islands, and this process has been observed in experiments.[14,18,25,26,32,33] The growth of islands in the vertical direction can be neglected since the oxygen sticking coefficient on a SiO_2 surface is three orders of magnitude lower than on a clean Si(100) surface.[34] Adsorbed oxygen atoms are brought to island perimeters owing to surface diffusion. Oxygen atoms are trapped by oxide islands, and silicon lattice atoms, which are abundant in the material, are incorporated into the islands concurrently. Therefore, the island growth rate is determined by the degree of supersaturation of the adsorbed oxygen.

The phase transition with formation of a submonolayer oxide proceeds when oxygen is constantly supplied to the

surface from the gaseous phase. Owing to the diffusive interaction between islands mediated by the concentration field of adsorbed oxygen and blocking of the surface by growing islands, the critical island dimension increases with the exposure time. The growth rate is limited by the capture of oxygen atoms on the island boundaries. The oxygen coverage θ_{ox} due to the islands on the surface as a function of the exposure time t near the nucleation threshold at $\theta_{ch} \gtrsim \theta_{th}$ is given by the expression[12,16]

$$\theta_{ox} = \left(1 - \frac{\theta_{th}}{\theta_0}\right)\tanh^2\frac{t}{\tau_{ox}}, \tag{1}$$

$$\tau_{ox} = \frac{A}{\sqrt{\pi N \theta_{0\nu}}}(T_{th} - T)^{-1/2}, \quad \theta_0 = Qs\tau, \tag{2}$$

where $\tau = (\alpha + Qs\Omega)^{-1}$, α is the oxygen escape rate, Ω is the oxygen adatom cross section, N is the surface density of oxide islands, ν is the rate of adsorbed oxygen capture on the island perimeters, θ_{th} and T_{th} are the threshold coverage of the surface with oxygen adatoms and the threshold temperature for oxide nucleation, τ_{ox} is the characteristic oxide growth time, $Q = p/\sqrt{2\pi mT}$ is the flux of oxygen from the gaseous phase on the surface, s is the sticking coefficient, and A is a constant independent of $T_{th} - T$. The nucleation threshold temperature is given by the expression

$$p = p_0 \exp(-E_\alpha/T_{th}), \tag{3}$$

where E_α is the activation energy for oxygen escape from the surface. With $E_\alpha = 3.8 \pm 0.2$ eV Eq. (3) adequately describes the experimental data obtained by Smith and Ghidini[31] at higher temperatures and pressures than those in our experiments.

The parameters of our experiments ($p \approx 10^{-6}$ Torr and $T \approx 900$ K) satisfy the conditions under which the interaction between oxide islands is significant, $L \approx l$ ($L = \sqrt{D\tau}$ is the characteristic diffusion length of oxygen adatoms over time τ, and $l \approx N^{-1/2}$ is the distance between islands) and island growth is limited by the time for the incorporation of oxygen into the oxide, $\nu L/D \ll 1$, and these conditions were used in deriving Eq. (2). In fact, when $L \approx l \approx 10^{-6}$ cm (the island density is estimated following Refs. 1 and 3 to be $N \approx 10^{12}$ cm^{-2}), the diffusion coefficient of adatoms $D \approx L^2/\tau \approx 10^{-11}$ cm^2/s [$\alpha \approx 10^{-1}$ s^{-1} (Ref. 6) and $\Omega sQ \approx 10^{-2}$ s^{-1}], which corresponds to a reasonable value of the activation energy $E \approx 1.7$ eV. The inequality $\nu \ll D/L \approx 10^{-5}$ cm/s is also satisfied with an activation energy $E = 1.8$ eV for oxygen capture by an island, which is associated with the passage of a silicon atom from the lattice to the oxide, whose activation energy $E \approx 1.83$ eV.

It follows from Eqs. (2) and (3) that the coverage of the SiO$_2$ layer on the surface tends to zero and the oxidation time increases rapidly ($\tau_{ox} \to \infty$) as the temperature approaches the nucleation threshold ($T \to T_{th}$). This is a general property of first-order phase transitions and should not depend on the assumptions made.

Figure 2 clearly shows that the dotted curve defined by Eq. (1) is in satisfactory agreement with experimental data on the boundary of the oxide nucleation region. This equa-

tion also describes experimental data obtained at lower temperatures, far from the threshold.[12] The measurements of $\tau_{ox}(T)$ (Fig. 3) and the theoretical curve given by Eq. (2) are in qualitative agreement. The measurement of T_{th}, however, is higher than the theoretical values estimated from Eq. (3). The $\theta_{ox}(t)$ curves for $T > 915$ K are in disagreement with the calculations. In this temperature range the onset of the oxidation process is delayed. The oxide growth takes place only at $t \gtrsim t_{th}$ and $\theta_{ch} \gtrsim \theta_{th}$. Finally, one can see in Fig. 2 that the oxidation rate $k = d\theta_{ox}/dt$ is almost constant with the temperature ($T \to T_{th}$ and $\theta_{ch} \to \theta_{th}$) near the nucleation threshold. This observation contradicts Eq. (2), from which follows $k = (\theta_{ch} - \theta_{th})^{-3/2}$, and hence $k \to 0$ as $\theta_{ch} \to \theta_{th}$.

It seems that these features of the oxidation process should be attributed to silicon surface roughening at temperatures $T > 915$ K. If we describe surface roughening in terms of vacancy clustering,[23,24] we must analyze the dynamics of the accumulation of vacancies in the process of oxygen adsorption in order to account for these anomalies. When the exposure time $t < t_{th}$, the adsorbed oxygen is accumulated on the surface. It is known that vacancies are generated when lattice atoms leave the surface within SiO molecules.[1] These molecules form as a result of the short-range interaction between oxygen and Si adatoms from the crystal.[6] Silicon adatoms from the crystal turn up when Si lattice atoms leave the first atomic layer owing to thermal activation and occupy positions of adsorbed atoms on terraces, while vacancies are generated in the first atomic layer. Hence, as a result of SiO desorption, excess vacancies should be generated in the first atomic layer. Inclusion of the long-range interaction between vacancies and oxygen adatoms in the mean-field approximation leads to the following expression for the equilibrium density $n_v^{(0)}$ of vacancies due to oxygen adsorption:[20]

$$n_v^{(0)} = n_0 \exp(-E_{eff}/T),$$
$$E_{eff} = E_0 - \theta_{ch}T \exp(U_{a-v}/T), \tag{4}$$

where U_{a-v} is the depth of the rectangular potential well modeling the interaction between oxygen adatoms and vacancies. The value of U_{a-v} is unknown. Below we shall use $U_{a-v} \approx 0.1 - 0.5$ eV in our calculations.

In accordance with Eq. (4), the equilibrium vacancy density increases with θ_{ch}. When the oxygen coverage on the surface is

$$\theta_{ch}^v = \frac{E_0}{T} \exp\left(-\frac{U_{a-v}}{T}\right), \tag{5}$$

the effective energy of vacancy formation is close to the temperature, and the number of vacancies and adsorbed lattice atoms can increase spontaneously without overcoming an activation barrier (vacancy instability).

Thus, there are three species on the surface near the nucleation threshold and at small exposure times $t < t_{th}$ before the onset of oxidation, namely vacancies, oxygen adatoms, and silicon adatoms from the crystal, whose densities, n_v, n_a, n_s, are different from their equilibrium values. The kinetic equations describing their accumulation on the surface, which consists of plane terraces and steps, for $t < t_{th}$ can be written as follows:
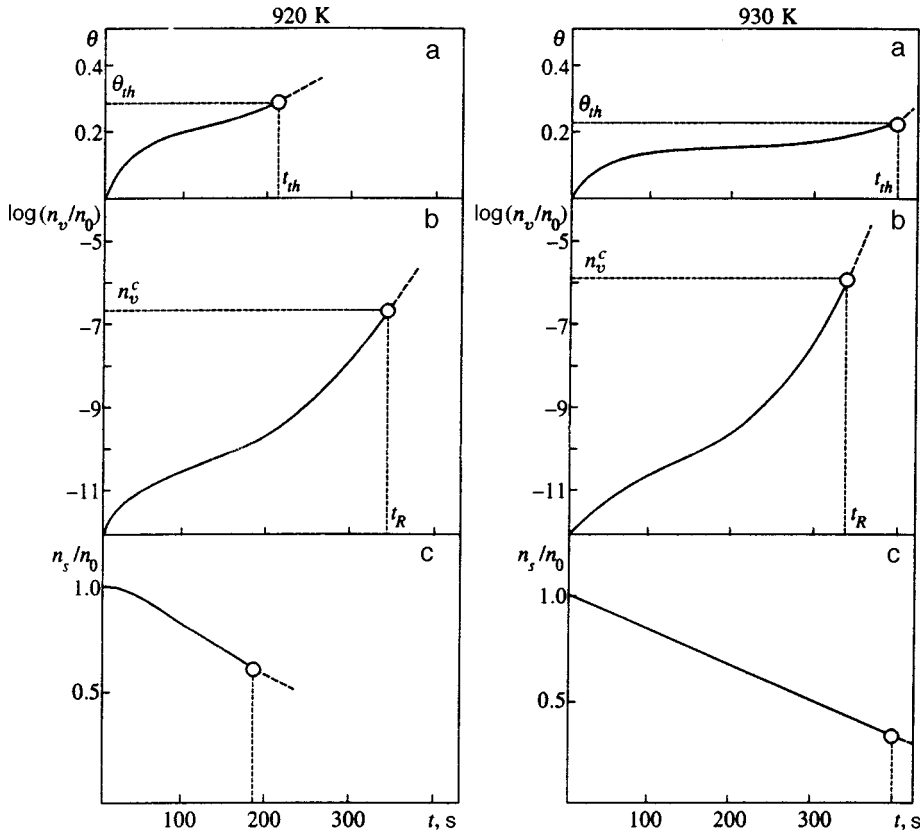
FIG. 5. (a) Chemisorbed-oxygen coverage θ, (b) surface density n_v of vacancies, and (c) surface density n_s of adsorbed atoms as functions of exposure time at $p = 10^{-6}$ Torr and temperatures of 920 K and 930 K; t_{th} is the exposure time at which condition (5) is satisfied.

$$\dot{n}_a = Qs(1 - \theta_{\text{ch}}) - \varepsilon n_a n_s,$$

$$\dot{n}_v = \gamma_1 n_v^{(0)} n_s^{(0)} - \gamma n_v n_s - \omega n_v, \tag{6}$$

$$\dot{n}_s = -\varepsilon n_a n_s + \gamma_1 n_v^{(0)} n_s^{(0)} - \gamma n_v n_s + G.$$

Here ε, γ, and γ_1 are the SiO desorption rate and the rates of vacancy/Si-adatom pair generation and annihilation, respectively, $n_v^{(0)}$ and $n_s^{(0)}$ are the equilibrium densities of vacancies and Si adatoms. The generation of SiO molecules incorporated into the first line of Eq. (6) is due to the short-range interaction[6] between oxygen and silicon adatoms ($\varepsilon n_a n_s$). The last term in the equation for $\dot{n}_v$ takes into account the vacancies moving into the crystal bulk, $\omega = D_v / a^2$ [$D \propto \exp(-E_b/T)$]. The flux G of silicon atoms from steps to terraces in the approximation of a small frequency of atom migration from a position on a step to a position on a break in comparison with the frequency of atom migration to a terrace is given by the expression[35]

$$G = g(n_s^{(0)} - n_s), \quad g = \omega_0 \frac{1}{a_k \lambda n_0} \exp\left(-\frac{E_g}{T}\right), \tag{7}$$

where ω_0 is the Debye frequency, λ^{-1} is the number of steps per unit length on the surface, a_k^{-1} is the number of breaks per unit length of a step, E_g is the potential barrier on the way from a step to a terrace, and n_0 is the density of surface atoms of the crystal.

When no oxygen is supplied to the surface ($Q=0$, $n_a = 0$), the system of equations (6) yields the equilibrium solution

$$n_s = n_s^{(0)}, \quad n_v = n_v^{(0)}. \tag{8}$$

By solving Eq. (6) numerically, one can analyze the kinetics of the densities of different species on the silicon surface until the onset of formation of the solid oxide ($t \leqslant t_{\text{th}}$). The desorption of SiO molecules, generation (annihilation) of silicon-adatom/vacancy pairs, and penetration of vacancies into the bulk are activation processes. For the numerical solution we took the following activation energies of these processes: $E_s = 3.6$ eV, $E_\gamma = E_{\gamma_1} = 1.2$ eV, and $E_b = 1.5$ eV. These values are within the spread of values quoted in different publications.[6,36,37] The energy E_g has been estimated at 2.8 eV.[36]

The coverage of the surface with oxygen adatoms, $\theta_{\text{ch}} = n_a / n_0$, as a function of exposure time is given in Fig. 2 (dashed line). It is clear that measurements of the total oxygen coverage of the surface, $\theta + \theta_{\text{ox}}$, are in satisfactory agreement with calculations of $\theta(t)$ at times before the onset of oxidation, $t \leqslant t_{\text{th}}$. Hence, there is a good reason to suppose that at $t \leqslant t_{\text{th}}$ all oxygen is in the chemisorbed state. This assumption is supported by the absence of Si^{3+} and Si^{4+} states in the x-ray photoelectron spectra at $t < t_{\text{th}}$.

Calculations of the oxygen coverage and the concentrations of vacancies and silicon adatoms as functions of exposure time for temperatures of 920 and 930 K at $p = 10^{-6}$ Torr are given in Fig. 5. It is seen that the numbers of vacancies and oxygen adatoms increase with the exposure time at both temperatures near the oxide nucleation threshold, whereas the number of silicon adatoms drops. This is caused by the formation of volatile SiO molecules. The
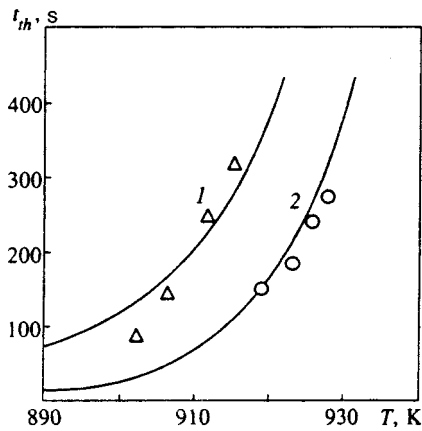
FIG. 6. Temperature dependence of the time t_{th} in which the chemisorbed-oxygen coverage achieves the threshold value at pressures of 6×10^{-7} Torr (*1*) and 10^{-6} Torr (*2*). Points—experimental data for these pressures.

curves also demonstrate that at lower temperatures (920 K) near the threshold the time t_{th} at which the threshold oxygen coverage is achieved $[\theta_{ch}(t_{th}) = \theta_{th}]$ is smaller than t_R, when θ_{ch} attains the value θ_{ch}^v specified by Eq. (5). Under these conditions, the oxide grows on the surface at a low vacancy density, and the surface may remain atomically smooth. At higher temperatures (930 K) the inverse condition is satisfied, $t_{th} > t_R$. In this case, spontaneous growth of the vacancy density starts earlier than oxidation, and n_v can achieve a value sufficient for vacancy clustering, and a rough relief is formed on the surface. This process will be discussed in Sec. 5.

Calculated plots of the time at which the nucleation-threshold oxygen coverage θ_{th} is achieved as a function of temperature at two pressures ($p = 6 \times 10^{-7}$ Torr and 10^{-6} Torr) are shown in Fig. 6. It is seen that a higher temperature corresponds to the same t_{th} at the higher pressure, in agreement with experimental data plotted in the same figure. Note that the flux of SiO molecules as a function of exposure time calculated at different temperatures are in qualitative agreement with experimental data.[6,32] This result confirms the correctness of taking into account the main surface processes included in Eq. (6).

Let us discuss features of the silicon oxidation process near the nucleation threshold. The slow variation of θ_{ch} at $\tau \lesssim t_{th}$ leads us to a conclusion that the phase transition with the formation of SiO_2 has a quasistationary character near the threshold. Therefore, in describing the conditions for the onset of oxidation, one can perform averaging in the first line in Eq. (6), which describes the variation of n_a, by representing $n_s(t)$ as a slowly descending linear function of time for $t \sim t_{th}$. As a result of this averaging, the concentration $n_s(t)$ can be replaced by $\langle n_s(t_{th}) \rangle = [n_s^{(0)} + n_s^{th}(t_{th})]/2$. The quasistationary solution of Eq. (6) for $\dot{n}_a$ allows us to determine the condition for the onset of oxidation: $\theta_{ch}(t_{th}) = \theta_{th}$. Equation (2) for the oxide-phase coverage θ_{ox} can be easily generalized to the case of n_s and n_v weakly deviating from their equilibrium values in the region near the threshold by replacing α in the first line of Eq. (6) for $\dot{n}_a$ and in the subsequent

formulas with $\alpha = \varepsilon \langle n_s(t_{th}) \rangle$. Then the value of θ_{ox} as a function of pressure, temperature, oxygen coverage, and exposure time in the threshold region is expressed by Eqs. (1) and (2), but oxidation starts at the time $t = t_{th}(T,p)$, as was observed in the experiments (Fig. 2). The experimental fact that the oxidation rate k does not decrease near the threshold, which contradicts the theoretical predictions, can be attributed to the closeness of t_{th} and t_R. For this reason, the two phase transitions, namely oxidation and surface roughening, are observed simultaneously, in agreement with the experimentally established oxidation of a rough surface, which should have excess vacancies. Since the solid oxide SiO_2 is formed by the short-range interaction between a vacancy and an oxygen atom,[20] the oxidation-limiting rate of incorporation (ν) of an oxygen atom into a vacancy on an island periphery should increase under these conditions.

5. Si(100) SURFACE ROUGHENING NEAR THE NUCLEATION THRESHOLD DUE TO ADSORBED OXYGEN

Let us analyze the conditions under which a rough relief is generated and its geometrical parameters after exposure of the surface to oxygen. The rough relief is generated under a constant flow of oxygen molecules incident on the surface, where the condition (5) for the activationless (spontaneous) generation of vacancies is satisfied. For this reason, a silicon surface adsorbing oxygen is a source rather than a drain of vacancies, whereas a flow of vacancies into the bulk is always impeded by an activation barrier. The bulk density $n_v(t,z)$ of isolated vacancies in the near-surface layer as a function of the distance z from the surface ($z=0$) and the time t of exposure to oxygen before the onset of surface roughening is determined by the standard solution[38] of the diffusion equation

$$\frac{\partial n_v}{\partial t} = D_v \frac{\partial^2 n_v}{\partial z^2}$$

with the boundary condition $n_v(t,0) = (n_v)_{z=0}/a$ [where $(n_v)_{z=0}$ is the time-dependent surface vacancy density, which satisfies Eq. (6)] and the initial condition $n_v(0,z) = n_v^{(0)}$, where $n_v^{(0)}$ is the equilibrium bulk vacancy density.

The surface vacancy density as a function of exposure time has been derived in Sec. 4 without taking into account the interaction between vacancies. The conditions for their clustering should be determined with allowance for this interaction. The interaction between vacancies ($v-v$) in the bulk was discussed in an earlier publication.[21] In a nonideal crystal, which contains defects in the form of substitutional impurities, spatial correlations of vacancies give rise to an effective long-range ($r \gg a$) attraction between vacancies. At distances $r \sim a$ vacancies repel one another since one cannot place two vacancies at one site. It is well known that near-surface relaxation and reconstruction take place in silicon, and when oxygen is adsorbed at $T \geq 900$ K, dissolved oxygen atoms are detected in a near-surface layer with a thickness $\lesssim 100$ nm.[37] We assume, therefore, that the interaction between vacancies in the near-surface layer is similar to that in the bulk of a nonideal crystal. Note that observations of

vacancy clusters on the silicon surface using scanning tunneling microscopy[2,17,18] give evidence in favor of attraction between vacancies.

Hereafter, as in earlier publications,[23,24] we describe the interaction between two vacancies using the model potential

$$V(\mathbf{r}) = \begin{cases} \infty, & r<a, \\ -\varepsilon, & a \leqslant r \leqslant b, \\ 0, & r>b. \end{cases} \tag{9}$$

In our estimates, we assume[21] that the parameter a in Eq. (9) is of the order of the interatomic distance, the attraction region is limited by the condition $b \leqslant 10a$, and $\varepsilon \sim 0.1$ eV.

Surface roughening due to the adsorption of oxygen atoms was described as a phase transition with vacancy clustering in our previous publications.[23,24] A more detailed account of the theory deserves a separate publication. In this section we only interpret our experimental data in terms of conditions under which the silicon surface is roughened and discuss a plausible mechanism of this process, features of the rough relief detected in experiments, and causes of the switchover between oxygen-silicon interaction regimes near the solid-oxide nucleation threshold.

The phase transition with vacancy clustering was described earlier in terms of the Landau theory.[24] The order parameter ξ is the deviation δn_v of the vacancy density $\tilde{n}(t,\mathbf{r})$ due to the formation of vacancy clusters in the near-surface layer from the average density n_v of isolated vacancies:

$$\xi(t,\mathbf{r}) \equiv \delta n_v(t,\mathbf{r}) = \tilde{n}_v(t,\mathbf{r}) - n_v . \tag{10}$$

The vacancies generated accumulate on the surface during the exposure time $t \sim t_R$ and diffuse into the near-surface layer of the crystal. The concentration n_v is a function of both the distance to the surface z and the exposure time t, and it is determined by the ratio between the rates of vacancy accumulation (see Sec. 4) and diffusion. Let us first discuss the phase transition with vacancy clustering in a spatially homogeneous system, where the mean concentration of isolated vacancies is uniform but time-dependent:

$$n_v = \overline{n_v(t,z)}^z = n_v(t).$$

This approximation allows us to account for the switchover between the oxygen-silicon interaction regimes (oxidation to roughening), the upper temperature at which a rough relief is formed, and the features of the rough surface. As will be shown below, the characteristic time of the phase transition with vacancy clustering (τ^{RT}) under the experimental conditions is much shorter than the accumulation time t_R of isolated vacancies. Therefore, we also assume that n_v is constant (or quasistationary) over the small characteristic time $t \sim \tau^{RT}$ in which the fluctuations $\delta n_v(t,\mathbf{r})$ develop, and the argument t ($t \sim t_R \gg \tau^{RT}$) in the function $n_v(t)$ determines only the value n_v at the moment t_R of the onset of vacancy clustering. Then

$$\xi(t,\mathbf{r}) = \delta n_v(t,\mathbf{r}) = \tilde{n}_v(t,\mathbf{r}) - n_v(t_R). \tag{11}$$

Under this definition of the order parameter, the quantity ξ/n_0 (where n_0 is the density of lattice atoms) is the fraction of atoms replaced by clustered vacancies in a near-surface layer of thickness H, where the phase transition with vacancy clustering takes place. Since open pores are generated (see below), the depth h of the rough relief at a given point on the surface is proportional to the density of clustered vacancies at this point:

$$h = \xi H/n_0 . \tag{12}$$

Therefore, the order parameter is the amplitude of surface roughness times an arbitrary multiplier.

The Landau theory[39] considers a phenomenological expansion of the free energy in terms of the order parameter near the transition point. The density-functional formalism[23,24] developed by Devyatko and Tronin[40] for a system of interacting classical Brownian particles in a medium not far from equilibrium allows one to calculate the coefficients of this expansion. The equation of motion for the Fourier transform $\xi_k(t)$ of the order parameter reads as follows:[23,24]

$$\frac{\partial \xi_k}{\partial t} = -D_v k^2 \frac{\partial F_k}{\partial \xi_k}, \tag{13}$$

$$F_k = \frac{1}{2}\lambda_k \xi_k^2 + \frac{1}{3}B\xi_k^3 + \frac{1}{4}\Gamma \xi_k^4 . \tag{14}$$

The terms proportional to ξ_k^3 and ξ_k^4 in Eq. (14) for F_k are ''stabilizing'' factors for the new state of the near-surface layer of thickness H [Eq. (12)] near the rough surface. The coefficients λ, B, and Γ in Eq. (14) for F_k are functions of ε and b, which are parameters of the model potential (9) of the interaction between vacancies:

$$\lambda_k = 1 - \frac{n_v}{n_v^c}f(kb), \quad B = \frac{1}{n_v^c}, \quad \Gamma = \frac{2}{(n_v^c)^2}, \tag{15}$$

where

$$n_v^c = \frac{3}{4\pi b^3} \frac{1}{e^{\varepsilon/T}-1}, \quad f(kb) = 3\left[\frac{\sin(kb)}{(kb)^3} - \frac{\cos(kb)}{(kb)^2}\right]. \tag{16}$$

In accordance with Eqs. (14)–(16), the free energy $\mathscr{F} = TF$ is a function of the concentration n_v of isolated vacancies and therefore depends on the oxygen pressure and temperature, as well as the exposure time. Using Eqs. (13)–(16), one can determine the conditions for the phase transition with vacancy clustering. To this end, one should analyze solutions of the equations

$$F_k(\xi_k) = 0, \tag{17}$$

$$\partial F_k/\partial \xi_k = 0. \tag{18}$$

In experiments at constant pressure and temperature, the concentration of isolated vacancies on an initially smooth surface increases with exposure time, whereas λ_k decreases [see Eq. (15)]. When n_v is sufficiently small and $\lambda_k > B^2/4\Gamma$, Eq. (18) has only one root $\xi_k = 0$, which corresponds, in accordance with Eq. (12), to the only stable state of the vacancy subsystem, namely the smooth surface. At $\lambda_k = B^2/4\Gamma$ a second minimum of the function $F_k(\xi_k)$ appears, and it corresponds to a metastable state of the surface. At even smaller

$\lambda_k = 2B^2/9\Gamma$ the curve of $F(\xi_k)$ touches the ξ_k axis at two points, namely $\xi_{k1} = 0$ and $\xi_{k2} = (2/3)B/\Gamma$. This means that the smooth ($\xi_k = 0$) and rough ($\xi_k = \xi_{k2} = n_v^c/3$) states of the surface have the same energy. These two states are separated by a potential barrier Q to oxide nucleation equal to the peak value of F at

$$\xi_{k3} = \frac{B}{6\Gamma} = \frac{n_v^c}{12}.$$

At $\lambda_k = 2B^2/9\Gamma$ the barrier height is

$$Q = TF(\xi_{k2}) = \frac{(n_v^c)^2 Ta^6}{18 \cdot 64} \simeq 10^{-2} \frac{Ta^6}{b^6} \frac{1}{e^{\varepsilon/T} - 1}. \tag{19}$$

Equation (19) indicates that, in the case of the long-range attraction between vacancies described by potential (9) for $b \sim 10a$ and $\varepsilon \sim 0.1$ eV at the temperatures $T \lesssim 10^3$ K, the quantity Q/T is much smaller than unity. Therefore, the transition from the smooth state of the surface with $\xi_k = 0$ to the rough state with $\xi_k = 2B/3\Gamma = n_v^c/3$ can proceed without overcoming a nucleation barrier. In this sense, surface roughening is similar to the relaxation of unstable states upon spinodal decomposition.[41]

It follows from the foregoing that the phase transition with vacancy clustering should start at an exposure time t_R, a temperature T, and a pressure p at which the concentration of isolated vacancies $n_v(t,T,p)$ satisfies, with allowance for Eqs. (17) and (18), the relation

$$\lambda_k = \frac{2B^2}{9\Gamma} = \frac{1}{9} = 1 - \frac{n_v(t_R, T, p)}{n_v^c} f(kb). \tag{20}$$

The function $f(kb)$ takes a maximum value equal to unity at $k = 0$ [Eq. (16)]. Therefore, the condition (20) for surface roughening with a wave vector k close to zero [when $f(kb) = 1$] is satisfied at the minimum vacancy concentration

$$n_v(t_R, T, p) = \frac{8}{9} n_v^c(T). \tag{21}$$

In the initial stage of the phase transition, where fluctuations in the vacancy concentration are small, we can retain only the term linear in ξ_k on the right-hand side of Eq. (13). Then the characteristic time τ_k for the kth mode of the order parameter in the initial stage of the phase transition is given by the expression

$$\tau_k = (\lambda_k D_v k^2)^{-1}. \tag{22}$$

Taking Eq. (16) into account, we can transform this expression for $kb \ll 1$ into

$$\tau_k^{-1} \approx \eta \left[\left(\frac{1}{\eta} - 1 \right) + \frac{(kb)^2}{10} \right] k^2 D_v, \quad \eta = \frac{n_v}{n_v^c}. \tag{23}$$

It follows from Eq. (23) that for $\eta \leqslant 1$, where $\lambda_k \geqslant 0$ for all small k, the quantity τ_k^{-1} is positive and that $\tau_k^{-1} \to 0$ as $kb \to 0$. For $0 < \lambda_k \leqslant 1/9$ this means that when the "energy" condition (20) for the phase transition is satisfied, the vacancy system is in the fluctuation region and that critical slowing takes place, i.e., the relaxation frequency tends to zero as the size of the fluctuations increases.
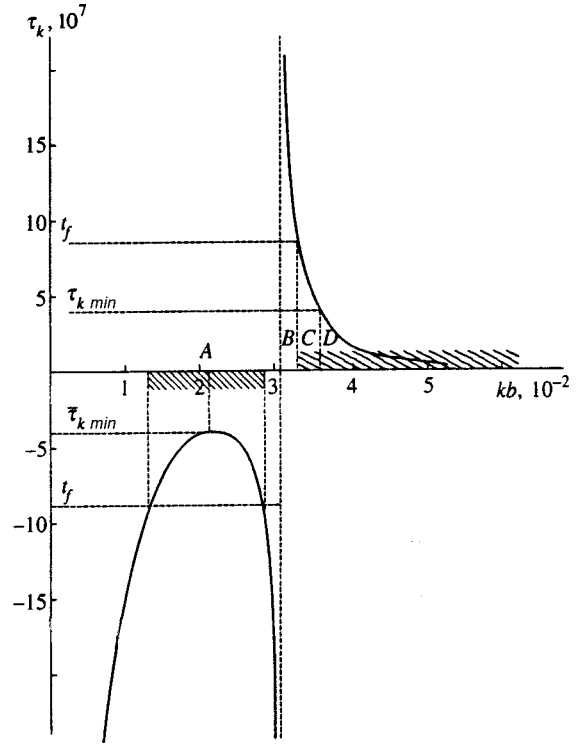
FIG. 7. Characteristic development time of the kth mode of the order parameter $\bar{\tau}_k = \tau_k (b^2/D_v)^{-1}$ versus kb calculated from Eq. (23) for $\eta - 1 = 10^{-4}$. The following points are marked on the abscissa axis: A—$(kb)_{\min}$; B—$(kb)_0$; C—$(kb)_f$; D—$(kb)_{\min}\sqrt{3}$.

For $\eta \geqslant 1$, when $\lambda_k < 0$, the quantity τ_k is negative and has the sense of the characteristic time for fluctuation growth (not decay) (Fig. 7). In accordance with Eq. (23), this increase takes place in modes with the wave vectors k which satisfy the inequalities

$$0 < (kb)^2 < (kb)_0^2 = 10(1 - 1/\eta). \tag{24}$$

These modes of the order parameter are unstable. The most rapidly developing mode has the wave vector

$$(kb)_{\min}^2 = 5(1 - 1/\eta), \tag{25}$$

and its characteristic time $|\tau_k|$ has the minimum $|(\tau_k)_{\min}|$:

$$|(\tau_k)_{\min}| = \frac{20}{3} \frac{b^2}{D_v} \frac{1}{(kb)_{\min}^4} \frac{1}{\eta}. \tag{26}$$

In accordance with Eq. (23), when $\eta \geqslant 1$ and $\lambda_k < 0$, the development of unstable modes with $k \approx k_{\min}$ should be accompanied by the development of small-scale fluctuations of the vacancy density (Fig. 7) with characteristic relaxation times $\tau_k \lesssim |\tau_{k \min}|$, whose wave vector k satisfies the inequality

$$(kb)^2 \geqslant 3(kb)_{\min}^2. \tag{27}$$

Thus, the resulting shape of the rough surface for $n_v \geqslant n_v^c$ ($\eta \geqslant 1$) should be determined by the development of unstable modes whose wave vectors belong to a narrow region near $k_{\min}$ and small-scale fluctuations with wave vectors satisfying Eq. (27).
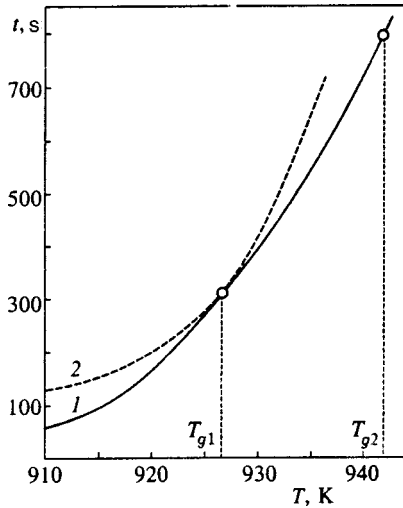
FIG. 8. Temperature dependence of the time t_R in which the critical concentration n_v^c is obtained (curve 2) and of the exposure time t_{th} for the onset of oxidation (curve 1).

The activation energy for vacancy diffusion in silicon is[36] $E_b \approx 1.5$ eV, and the diffusion coefficient at $T = 900$ K is $D_v \approx 10^{-7}$ cm^2/s. Therefore, it follows from our estimates that, if a rough relief with a maximum characteristic amplitude $L = 5 \times 10^2$ nm has been formed as a result of the phase transition with vacancy clustering (Fig. 4), we have $(kb)_{min} \approx 4 \times 10^{-2}$, $|\tau_{k\,min}| \approx 6$ s, and $\eta - 1 \approx 10^{-4}$ for $b = 3$ nm.

Note that the vacancy accumulation time $t_R \sim 10^2$ s (Fig. 5). Therefore, $t_R \gg |\tau_{k\,min}|$ under the experimental conditions, and the density n_v of isolated vacancies can be treated as quasistationary in determining the clustering conditions, as was done in deriving Eq. (13).

Let us discuss the switchover from the oxidation regime of the interaction between oxygen and silicon to surface roughening. The relation $n_v(t_R) = n_v^c$ ($\eta = 1$), which determines the condition for the formation of unstable developing modes due to vacancy clustering, yields, with consideration of Eq. (6) for $n_v(t)$ and Eq. (16) for $n_v^c(T)$, the time t_R as a function of oxygen pressure and temperature. The curve of $t_R(T)$ for $p = 1 \times 10^{-6}$ Torr corresponding to the condition $n_v(t_R) = n_v^c$ is plotted in Fig. 8. This figure also shows a plot of the exposure time $t_{th}(T)$ for the onset of oxidation (the time when the oxide nucleation threshold, $\theta_{ch} = \theta_{th}$, is achieved). The curves indicate that at low temperatures ($T < T_{g1}$) $t_R > t_{th}$, i.e., the oxidation process starts before the vacancy concentration needed for the development of unstable modes according to Eq. (23) is achieved. At low temperatures the rate G of the formation of adatoms from the crystal and the rate of SiO desorption are small.[6] The oxide islands formed after the exposure time t_{th} are drains for adsorbed oxygen atoms; therefore, their concentration decreases when the islands are formed, and the condition (5) for the spontaneous generation of vacancies can no longer be satisfied.

At $T > T_{g1}$ we have $t_R < t_{th}$, and the surface is roughened before the solid-oxide nucleation threshold is attained. In this case, the number of steps, which are sources of adatoms from the crystal, increases, and the average distance between them

drops; therefore, in accordance with Eqs. (6) and (7), the flux of SiO molecules desorbed from the surface should grow. As a result, the flux of SiO molecules desorbed from the rough surface is considerably larger (more than one order of magnitude) than the flux from the smooth surface.[6] In the long run, the concentration of adsorbed oxygen decreases as the surface is roughened, and the oxide nucleation threshold cannot be achieved.

The free parameters of the model are the parameters of the a-v and v-v interaction potentials: $U_{a-v} = 0.3$ eV, $U_{v-v} = 0.1$ eV, and $b = 10a$. Variation of these parameters leads to a shift of the temperature T_{g1} with respect to the threshold temperature T_{th}, but T_{g1} is still smaller than T_{th} since $\tau_{ox} \propto (T_{th} - T)^{-1/2}$. Therefore, the switchover from the regime of solid oxide formation to that of surface roughening should occur near the oxide nucleation threshold at a temperature lower than the oxidation threshold temperature at a given pressure. This is in agreement with our observations of the switchover between oxidation and surface roughening in the region of temperatures and pressures above the $p_{th}(1/T)$ curve. At $U_{a-v} = 0$, vacancies are generated in our model as a result of SiO desorption due to violation of the balance between formation of vacancies and adatoms from the crystal. In this case, the vacancy density n_v does not reach the critical value n_v^c needed for development of unstable modes at the kinetic parameters selected in our calculations.

The plot of T_{g1} versus pressure defines the boundary between the regimes of planar surface oxidation and surface roughening induced by oxygen adsorption. In the region of temperatures and pressures under consideration, this boundary replaces the threshold line for oxide formation specified by Eq. (3). Our model, however, does not yield the exact position of the new boundary $T_{g1}(p)$, since it does not apply to $t_R \sim t_{th}$. Around T_{g1}, oxide islands and surface roughness should develop simultaneously. This corresponds to the observed oxidation of the rough surface. In this region of temperatures and pressures, interaction between two order parameters describing the phase transitions with oxidation and roughening should be taken into account.

Our model also yields an upper temperature at which the surface roughness can develop. It is the temperature T_{g2} in Fig. 8. As the temperature rises, the rates of both SiO formation and desorption increase, and, consequently, the resulting oxygen concentration on the surface drops. At the same time, the oxygen coverage θ_{ch}^v, at which the effective energy of vacancy generation $E_{eff} \sim T$, also rises with T [see Eq. (6)]. Therefore, the condition (6) for the spontaneous generation and clustering of vacancies cannot be satisfied above the temperature T_{g2}. Then the crystal surface should remain smooth, in agreement with our experimental data (Fig. 4d). In this connection, some attention should be focused on the experimental results reported by Wurm *et al.*,[1] who established the temperature $T = 950$ K, below which terrace surfaces were roughened and above which only steps were affected, whereas the terrace surfaces remained smooth, at $p = 5 \times 10^{-8}$ Torr. Our model does not describe the roughening of steps, since they were taken into account only in the averaged source G of adatoms from the crystal in Eq. (7).

Let us discuss the geometrical parameters of a rough

surface (Fig. 4). Within our model the growth in the density of isolated vacancies should continue at times $t \sim \tau_k$ and when the condition for emergence of unstable modes, $n_v = n_v^c$, is satisfied owing to the high rate of the spontaneous generation of vacancies. Therefore, the formation of a rough surface with characteristic geometrical parameters continues as vacancies are accumulated on the surface and in the near-surface layer. As was noted above, formation of the rough relief leads to a notable increase in the flux of evaporated SiO by more than one order[6] and hence, to a decrease in the oxygen coverage θ_{ch} on the surface. For this reason, the oxygen coverage θ_{ch} can drop below θ_{ch}^v at a certain moment t_f after the onset of vacancy clustering. Then the condition for spontaneous vacancy generation is violated, and the vacancy concentration becomes smaller than the clustering threshold n_v^c. The flow of vacancies from the surface into the near-surface layer, where they form clusters, also varies during surface roughening. In order to determine the time t_f, during which the relief is formed, and the distribution of roughness amplitudes as a function of time, one should calculate again the densities n_a, n_s, and n_v over the time interval t_f with allowance for their dependence on the order parameter ξ, which is equivalent to taking into account the surface curvature. The experimental data needed for testing solutions of this problem are not available; therefore this problem and a more detailed analysis of the dynamics of surface roughening merit a separate study. Here we discuss only the simplest results deriving from our model, which allow us to understand the mechanisms forming the relief shown in Fig. 4b. The latter is characterized by sparse and almost periodic deep pits against a background of small-scale surface fluctuations.

It follows from the above statements that the rough surface is shaped under nonstationary conditions, in which the time-dependent source of vacancies turns off at the moment t_f. The presence of a developed rough relief indicates that $t_f > |\tau_{k\,\min}|$ and the vacancy concentration on the surface $(n_v)_{z=0}$ reaches a value satisfying the condition $n_v \geqslant n_v^c$ ($\eta \geqslant 1$) for unstable mode development. If we assume, as was done previously, that the vacancy concentration in a certain layer is independent of the distance z to the surface, then the rough surface is formed at times $t_f \geqslant |\tau_{k\,\min}|$ as a result of development of unstable modes with wave vectors k in the interval Δk around the point $(kb)_{\min}$ [Eq. (25)] and the relaxation of modes with wave vectors $(kb)_f \leqslant \sqrt{3}(kb)_{\min}$. As the ratio $t_f / |\tau_{k\,\min}|$ increases, the interval Δk widens and $(kb)_f$ drops. These conditions correspond to the relief shown in Fig. 4c, but not Fig. 4b, since the ratio between the characteristic distance L between the deep pits and the characteristic size $L_f = 2\pi/k_f$ of short-wavelength fluctuations is $L/L_f \geqslant 4$.

As a result of vacancy diffusion from the surface into the bulk, the density at a depth h from the surface is $(n_v)_h < (n_v)_{z=0}$. If pits with depth h are formed, this means that at this depth $(n_v)_h \geqslant n_v^c$ and $t_f \approx |\tau_{k\,\min}|$. Therefore, in accordance with Eq. (22), modes with k close to $k_{\min}$ become unstable at a distance h from the surface, and $(kb)_{\min}$ is smaller than on the surface since $\eta - 1$ is lower. Hence, open vacancy pores should be seen, and the characteristic ratio

L/L_f should be higher than $\sqrt{3}$, i.e., sparse deep pits should emerge against a background of small-scale fluctuations of the surface height. This corresponds to the surface shown by Fig. 4c.

The depth of the sparse pits can be estimated as the maximum depth H at which the vacancy concentration n_v^c corresponding to the onset of unstable mode development is achieved:

$$H \sim (t_f D_b)^{1/2}, \tag{28}$$

where D_b is the vacancy diffusion coefficient in the bulk. Since the rough surface with the maximum L corresponding to the average distance between deep pits is formed during the time t_f, then

$$t_f \sim L^2/D_s, \tag{29}$$

where D_s is the vacancy diffusion coefficient on the surface. It follows from Eqs. (28) and (29) that the ratio between the depth of the sparse pits to the distance between them is

$$H/L \sim D_b/D_s^{1/2}.$$

For silicon, the activation energies for diffusion are $E_b \approx 1.5$ eV (Ref. 37) and $E_s = 0.7$ eV (Ref. 36). Therefore, $L/H \approx 30$ at $T = 900$ K. This is in agreement with our observations of the relief formed (Fig. 4c).

Thus, the proposed model of surface roughening due to oxygen adsorption, as a result of a phase transition vacancy clustering allows us to account for the switchover between the oxygen-silicon interaction regimes, the existence of an upper temperature limit at which surface roughening takes place, and the reasons for the appearance of a relief with sparse deep pits against a background of short-wavelength surface fluctuations. Note that, in order to account for these features of the process, we needed neither calculations of the time t_f in which the relief is formed nor consideration of the density of isolated vacancies (or other particles) as a function of the order parameter ξ, i.e., the dependence of the intensity of sources/drains on the surface curvature. A tentative explanation of this fact is that the formation of the relief observed takes the time $t_f \geqslant |\tau_{k\,\max}|$, and the range Δk of unstable modes is little affected by changes in the sources/drains of vacancies due to surface roughening. This is the traditional approach to interpreting such features of the decay of unstable states which can be fully described by determining the conditions for the onset of decomposition and revealing unstable, rapidly growing modes of the order parameter.

*)E-mail: borman@bvd.mephi.ru

[1] K. Wurm, R. Kliese, Y. Hong *et al.*, Phys. Rev. B **50**, 1567 (1994).

[2] J. V. Seiple and J. P. Pelz, Phys. Rev. Lett. **73**, 999 (1994).

[3] A. Feltz, U. Memmert, and R. J. Behm, Surf. Sci. **314**, 34 (1994).

[4] Y. Ono, M. Tabe, and H. Kageshina, Phys. Rev. B **48**, 14 291 (1993).

[5] N. C. Lu, E. P. Gusev, E. Garfunkel *et al.*, Surf. Sci. Technol. **10**, 215 (1995).

[6] T. Engel, Surf. Sci. Rep. **18**, 94 (1993).

[7] Y. Enta, Y. Takegawa, T. Suemitsi *et al.*, Appl. Surf. Sci. **100/101**, 449 (1996).

[8] C. J. Sofield and A. M. Stoneham, Semicond. Sci. Technol. **10**, 215 (1995).

[9] M. Niwa, T. Kouzaki, M. Udagawa *et al.*, Jpn. J. Appl. Phys. **33**, 388 (1994).

[10] *Fundamental Aspects of Ultrathin Dielectrics on Si-Based Devices: Towards Atomic-Scale Understanding (Technical Program and Abstracts, Advance Research Workshop Supported by NATO Science Affairs Division)*, St. Petersburg, Russia (1997).

[11] F. Lutz, L. Kubler, and J. L. Bischoff, Phys. Rev. B **40**, 11 747 (1989).

[12] V. D. Borman, E. P. Gusev, Yu. Yu. Lebedinski, and V. I. Troyan, Phys. Rev. B **49**, 5415 (1994).

[13] G. Hollinger and F. J. Himpsel, Phys. Rev. B **28**, 3651 (1983).

[14] F. J. Himpsel, F. R. McFeely, A. Taleb-Ibrahimi *et al.*, Phys. Rev. B **38**, 6084 (1988).

[15] J. R. Engstrom, D. J. Bonser, and T. Engel, Surf. Sci. **268**, 238 (1992).

[16] V. D. Borman, E. P. Gusev, Yu. N. Devyatko *et al.*, Poverkhnost' **8**, 22 (1990).

[17] Ph. Avouris and D. Cahill, Ultramicroscopy **42–44**, 838 (1992).

[18] Ph. Avouris, In-Whan Lyo, and F. Bozso, J. Vac. Sci. Technol. B **9**, 424 (1991).

[19] K. Ohkubo, Y. Igari, S. Tomoda *et al.*, Surf. Sci. **260**, 44 (1992).

[20] V. D. Borman, E. P. Gusev, Yu. N. Devyatko, V. N. Tronin, and V. I. Troyan, Surf. Sci. Lett. **301**, L239 (1994).

[21] Yu. N. Devyatko and V. N. Tronin, Phys. Scr. **41**, 355 (1990).

[22] F. M. Ross, J. M. Gibson, and R. D. Twesten, Surf. Sci. **310**, 243 (1994).

[23] V. D. Borman, O. V. Tapinskaya, V. N. Tronin, and V. I. Troyan, Phys. Low-Dimens. Struct. **6**, 7 (1994).

[24] V. D. Borman, O. V. Tapinskaya, V. N. Tronin, and V. I. Troyan, JETP Lett. **60**, 718 (1994).

[25] F. Lutz, J. L. Bischoff, L. Kubler, and D. Bolmont, Phys. Rev. B **40**, 10 356 (1989).

[26] M. Tabe, T. T. Chiang, I. Lindau, and W. E. Spicer, Phys. Rev. B **34**, 2706 (1986).

[27] H. Ibach, H. D. Bruchmann, and H. Wagner, Appl. Phys. A **29**, 113 (1982).

[28] P. J. Grunthaner, M. H. Hecht, F. J. Grunthaner, and N. M. Johnson, J. Appl. Phys. **61**, 629 (1987).

[29] G. Hollinger, J. F. Morar, F. J. Himpsel *et al.*, Surf. Sci. **168**, 609 (1986).

[30] J. J. Lander and J. Morrison, J. Appl. Phys. **33**, 2089 (1962).

[31] F. W. Smith and G. Ghidini, J. Electrochem. Soc. **129**, 1300 (1982).

[32] M. P. D'Evelyn, M. M. Nelson, and T. Engel, Surf. Sci. **186**, 75 (1987).

[33] F. M. Leibsle, A. Samsavar, and T. C. Chiang, Phys. Rev. B **38**, 5780 (1988).

[34] P. Gupta, C. H. Mak, P. A. Coon, and S. M. George, Phys. Rev. B **40**, 7739 (1989).

[35] Ya. E. Geguzin and Yu. S. Kaganovskiĭ, *Diffusion Processes on Crystal Surfaces* [in Russian], Énergoatomizdat, Moscow (1984).

[36] A. Pimpinelli and J. Villain, Physica A **204**, 521 (1994).

[37] V. S. Vavilov, V. F. Kiselev, and V. N. Mukashev, *Defects in Silicon and on its Surface* [in Russian], Nauka, Moscow (1990).

[38] A. N. Tikhonov and A. A. Samarskii, *Equations of Mathematical Physics* [in Russian], Nauka, Moscow (1977).

[39] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, Pergamon Press, Oxford–New York (1980).

[40] Yu. N. Devyatko and V. N. Tronin, Zh. Éksp. Teor. Fiz. **98**, 1570 (1990) [Sov. Phys. JETP **71**, 880 (1990)].

[41] V. P. Skripov and A. V. Skripov, Usp. Fiz. Nauk **128**, 193 (1979) [Sov. Phys. Usp. **22**, 389 (1979)].

Magnetoelectric effect and toroidal ordering in $Ga_{2-x}Fe_xO_3$

Yu. F. Popov,*) A. M. Kadomtseva, G. P. Vorob'ev, V. A. Timofeeva,†) and D. M. Ustinin

M. V. Lomonosov Moscow State University, 119899 Moscow, Russia

A. K. Zvezdin and M. M. Tegeranchi

Institute of General Physics, Russian Academy of Sciences, 117942 Moscow, Russia
(Submitted 5 June 1997)
Zh. Éksp. Teor. Fiz. **114**, 263–272 (July 1998)

The field dependence of the magnetoelectric effect and longitudinal magnetostriction of $Ga_{2-x}Fe_xO_3$ single crystals is studied in magnetic fields up to 200 kOe in the temperature range from 4.2 to 300 K. It is shown that the magnetoelectric effect in these materials is determined mainly by the toroidal moment **T** and is not related to magnetostriction, as was previously theorized. A new method for determining the toroidal moment by measuring the electric polarization in a strong magnetic field is proposed. The value of the toroidal moment of the unit cell in $Ga_{1.15}Fe_{0.85}O_3$ is calculated: $\mathbf{T}=(T_a,0,0)$, where $T_a=24.155\mu_B$ Å per unit cell. Experimental data are analyzed using a theory of toroidal spin ordering, which gives good agreement with experiment. © *1998 American Institute of Physics.* [S1063-7761(98)01807-1]

1. INTRODUCTION

Since the early studies of the Fe–Ga oxides $Ga_{2-x}Fe_xO_3$ (Refs. 1–6), the interest in investigating them has been largely due to their combination of ferromagnetic and piezoelectric properties. The crystal symmetry of these materials is described by the C_{2v}^9 space group.[5] It should be noted that $Ga_{2-x}Fe_xO_3$ was the first ferromagnet in which a linear magnetoelectric effect was discovered[7,8] and that its magnetoelectric susceptibility exceeded the values previously observed in antiferromagnets. An off-diagonal magnetoelectric effect $P_b(H_c)$ was observed in Ref. 7. It was theorized that the magnetoelectric effect in $Ga_{2-x}Fe_xO_3$ may be caused by the combined action of the piezoelectricity and magnetostriction inherent to these materials. Testing this hypothesis, however, would require data on the magnetoelastic properties of $Ga_{2-x}Fe_xO_3$, which are presently not available in the literature. There are likewise no data from measurements of the magnetic properties of these crystals in strong magnetic fields. The purpose of our work was to study the magnetic, magnetoelastic, and magnetoelectric properties of $Ga_{2-x}Fe_xO_3$ in strong magnetic fields (in 200 kOe) and to establish the existence of toroidal spin ordering in $Ga_{2-x}Fe_xO_3$ which is allowed by the symmetry of these crystals.

2. EXPERIMENTAL RESULTS

A combined investigation of the magnetic, magnetoelastic, and magnetoelectric properties of $Ga_{2-x}Fe_xO_3$ ($x=0.7$, 0.85, and 1.0) single crystals was performed in strong pulsed magnetic fields up to 200 kOe in the temperature range from 4.2 to 300 K. The single crystals were grown in a melt of the $(Fe_2O_3–Ga_2O_3)–(Bi_2O_3–PbF_2)$ system with slow cooling from 1100 to 850 °C.[9] The magnetostriction measurements were performed using a quartz piezoelectric transducer glued onto the single-crystal sample. The magnetoelectric effect

was measured using the method described in detail in Ref. 10. The magnetic measurements in strong pulsed magnetic fields were performed by an induction method, and the measurements in static magnetic fields up to 12 kOe were performed using a torsional anisometer with autocompensation. As is shown in Fig. 1, the values of the magnetic moment and the Néel temperature increase with increasing Fe^{3+} ion concentration, in agreement with Ref. 3. Measurements of the magnetization and torque curves show that the easy magnetization direction in the compounds investigated is the c axis of the crystal with the strongest magnetic anisotropy in the bc plane. The relatively large value of the magnetization is apparently due to the uncompensated ferromagnetism, in agreement with the experimental results of neutron diffraction investigations.

A linear nondiagonal magnetoelectric effect ($P_b = \alpha_{bc}H_c$) was previously[7,8,11,12] predicted on the basis of a symmetry analysis and discovered experimentally in $Ga_{2-x}Fe_xO_3$ crystals. We measured not only the $P_b(H_c)$ curve, but also the $P_c(H_b)$ curve in pulsed magnetic fields up to 200 kOe for the purpose of discovering the difference between them, since a difference between the off-diagonal components of the magnetoelectric susceptibility tensor is critical for the existence of a toroidal moment. The investigations were carried out for the composition $Ga_{1.15}Fe_{0.85}O_3$, since there are neutron-diffraction data regarding the magnetic structure for just this composition.[13,14] The experimental results that we obtained are presented in Fig. 2. As can be seen from Fig. 2, the values of $P_b(H_c)$ (Fig. 2a) greatly exceed the values of $P_c(H_b)$ (Fig. 2b) obtained at the same magnetic fields. We note, however, that a comparison of the off-diagonal susceptibilities is insufficiently well posed in the present experiment, since the magnetic states of the crystal for **H∥c** and **H∥b** differ. To compare the different components of the off-diagonal susceptibilities in the same mag-
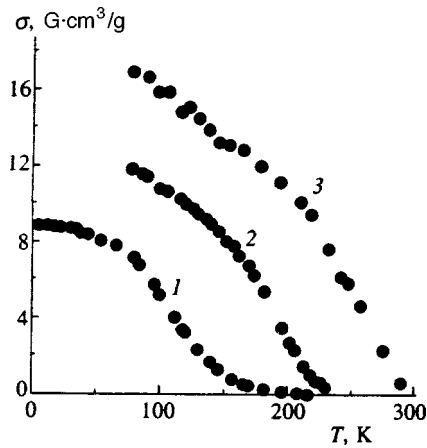
FIG. 1. Temperature dependence of the spontaneous specific magnetization of the $Ga_{2-x}Fe_xO_3$ system: *1*—$x=0.7$; *2*—$x=0.85$; *3*—$x=1$.

netic state of the crystal, the electric polarizations P_b and P_c were measured with the magnetic field oriented along the diagonal in the bc plane. The dependence of the components of the electric polarization tensor on magnetic field was treated by the least-squares method under the assumption of a power-law dependence of the polarization on field:

$$P_i = \sum_{n=0}^{\infty} d_n^i H^n .$$

The coefficients d_n^i for $n \geqslant 3$ were found to be nearly equal to zero. Figure 3 shows plots of $P_b(H_{bc})$ and $P_c(H_{bc})$ for a $Ga_{2-x}Fe_xO_3$ crystal with $x=0.85$ at $T=10$ K. The figure also shows the contributions to the magnetoelectric effect which are linear and quadratic with respect to the field. The linear contribution to the magnetoelectric effect was also observed when the $P_a(H_{ab})$ and $P_b(H_{ab})$ curves were measured in fields sufficient for reorienting the magnetization of the crystal from the c axis to the a axis (Fig. 4). From Figs. 3 and 4 it can be seen that $\alpha_{bc} \gg \alpha_{cb}$ and $\alpha_{ba} \gg \alpha_{ab}$. The temperature dependence of the off-diagonal components of the linear magnetoelectric coupling tensor obtained with allowance for the fact that the field is oriented along the diagonal in the bc plane is shown in Fig. 5. As will be shown below, the strong difference between the values of α_{bc} and α_{cb} observed is due to the existence of a toroidal moment in

this compound. Measurements of the diagonal magnetoelectric effect in $Ga_{1.15}Fe_{0.85}O_3$ along the b and c axis revealed a quadratic dependence of the electric polarization on magnetic field, in accordance with the theoretical predictions.

To establish a possible link between the linear off-diagonal magnetoelectric effect and the magnetoelastic interactions, magnetostriction measurements were also performed. Figure 6 shows the field dependence of the magnetostriction along the a, b, and c axes. According to Fig. 6, the components U_{aa} and U_{cc} of the longitudinal magnetostriction depend linearly on magnetic field, while U_{bb} exhibits a nonlinear dependence. A linear field dependence was also observed for the transverse components U_{bc} and U_{cb} of the magnetostriction. A theoretical analysis of the experimental results obtained is given below.

3. DISCUSSION OF RESULTS

1. According to Refs. 13 and 14, a $Ga_{2-x}Fe_xO_3$ crystal has a G_x ferrimagnetic structure with four inequivalent coordinates for the Fe1, Fe2, Ga1, and Ga2 ions. As follows from the neutron-diffraction data in Ref. 14, the probabilities of the accommodation of iron atoms in all four sites are nonzero and equal to

$$\zeta(1)=0.73, \quad \zeta(2)=0.87, \quad \zeta(3)=0.62, \quad \zeta(4)=0.08, \tag{1}$$

and the magnetic moment of iron equals

$$\mu_{Fe}=4.8\mu_B . \tag{2}$$

Such a structure creates difficulties when an attempt is made to examine the properties of a system with the aid of the antiferromagnetic vector. In this paper we describe the magnetic structure of crystals using a multipole expansion of the spin density. However, the contributions of the toroidal moments must now be taken into account in the multipole expansion.[15–20] It is known that an infinite series of multipole characteristics, viz., moments and the power functions of the radii for them, appears in the multipole expansion of the current density or the spin density after averaging over a physically small volume. As a rule, knowledge of only the lowest members of this series is important for a macroscopic description of a crystal; therefore, we confine ourselves to a description of a system on the basis of the multipole mo-
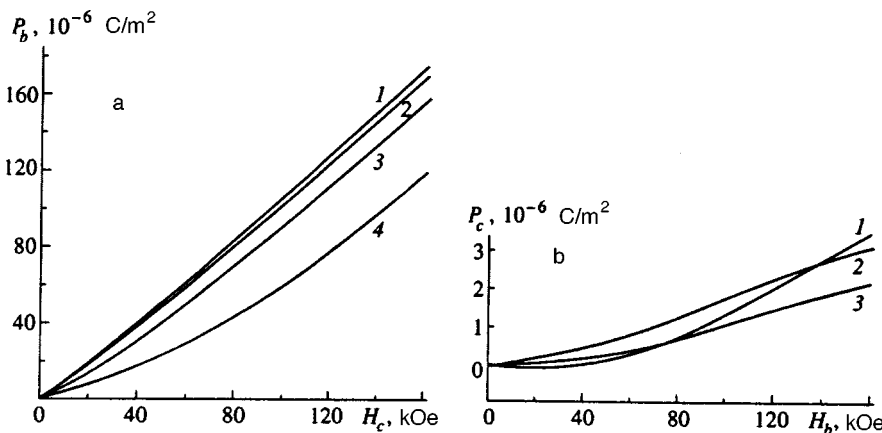
FIG. 2. Field dependence of the electric polarizations $P_b(H_c)$ (a) and $P_c(H_b)$ (b) of a $Ga_{1.15}Fe_{0.85}O_3$ crystal at various temperatures: a) *1*—35–62 K; *2*—10 K; *3*—102 K; *4*—150 K; b) *1*—35 K; *2*—10 K; *3*—150 K.
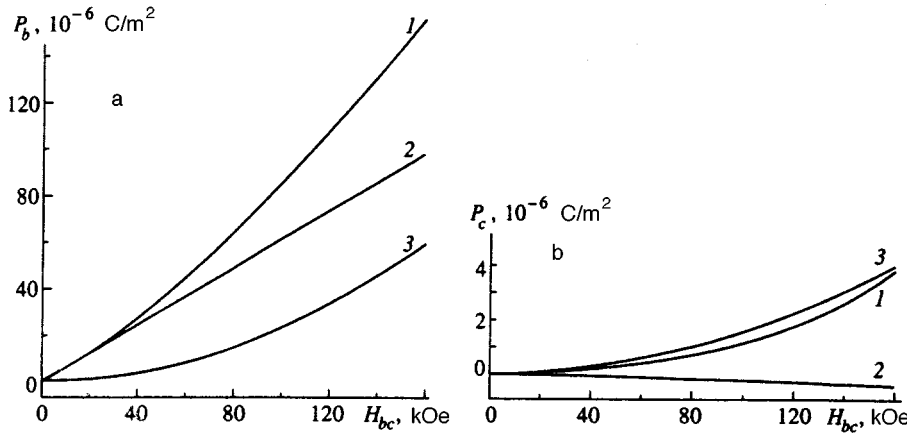
FIG. 3. Dependence of the electric polarizations $P_b(H_{bc})$ (a, curve *1*) and $P_c(H_{bc})$ (b, curve *1*), as well as the linear (curve *2*) and quadratic (curves *3*) contributions to the polarization derived from them, on the magnetic field H_{bc} for a $Ga_{1.15}Fe_{0.85}O_3$ crystal at 10 K.

ments up to the second power using magnetic and toroidal moments and the corresponding second-rank tensors (magnetic-quadrupole moments). The magnetic and toroidal multipole moments are defined by the familiar formulas (see Refs. 19 and 20)

$$\mathbf{M} = \frac{1}{2c} \int \mathbf{r} \times \mathbf{j}(\mathbf{r}) dv, \tag{3}$$

$$\mathbf{T} = \frac{1}{10c} \int (\mathbf{r}(\mathbf{r} \cdot \mathbf{j}) - 2r^2 \mathbf{j}) dv, \tag{4}$$

where $\mathbf{j}(\mathbf{r})$ is the current density,

$$\mathbf{j}(\mathbf{r}) = c \nabla \times \mathbf{S}(\mathbf{r}), \tag{5}$$

$\mathbf{S}(\mathbf{r})$ is the spin density, c is the speed of light, and the integration is carried out over a unit cell.

Plugging (5) into formulas (3) and (4), we obtain expressions for the magnetic and toroidal moments of a unit cell of a crystal:[19]

$$\mathbf{M} = 2\mu_B \sum_{\alpha} \mathbf{S}_{\alpha}, \tag{6}$$

$$\mathbf{T} = \frac{\mu_B}{2} \sum_{\alpha} \mathbf{r}_{\alpha} \times \mathbf{S}_{\alpha}, \tag{7}$$

where $\mathbf{S}_{\alpha}$ and $\mathbf{r}_{\alpha}$ are the spin moment and the radius vector of ion α in the unit cell of the crystal measured relative to the center of the unit cell, and the summation is carried out over all the ions in the unit cell.

2. It is known that the noncentrosymmetric $m2m$ crystal structure of a $Ga_{2-x}Fe_xO_3$ crystal can be transformed into a centrosymmetric mmm structure by means of fairly small displacements of the ions in the unit cell. In this case we can write

$$\mathbf{r}_{\alpha} = \mathbf{r}_{\alpha}^0 + \mathbf{r}_{\alpha}^1, \tag{8}$$

where $\mathbf{r}_{\alpha}^0$ is the radius vector of ion α in the centrosymmetric mmm structure and $\mathbf{r}_{\alpha}^1$ is the displacement of the respective iron ion from its symmetry center in the ideal mmm structure to its site in the real $m2m$ structure. The most important b components of $\mathbf{r}_{\alpha}^1$ for the ensuing analysis equal

$$r_{Ga}^1 = -0.0658 \pm 0.029 \text{ Å}, \quad r_{Fe}^1 = +0.23 \pm 0.058 \text{ Å},$$

and Eq. (7) can be represented in the form

$$\mathbf{T} = \frac{\mu_B}{2} \left(\sum_{\alpha} (\mathbf{r}_{\alpha}^0 \times \mathbf{S}_{\alpha}) + \mathbf{r}_{Ga}^1 \times \sum_{\alpha(Ga)} \mathbf{S}_{\alpha(Ga)} + \mathbf{r}_{Fe}^1 \right.$$

$$\left. \times \sum_{\alpha(Fe)} \mathbf{S}_{\alpha(Fe)} \right). \tag{9}$$

Substituting (6) and (7) in (9), we obtain

$$\mathbf{T} = \mathbf{T}^0 + \frac{1}{4} (\mathbf{r}^1 \times \mathbf{M}_{Ga} + \mathbf{r}^1 \times \mathbf{M}_{Fe}), \tag{10}$$

where
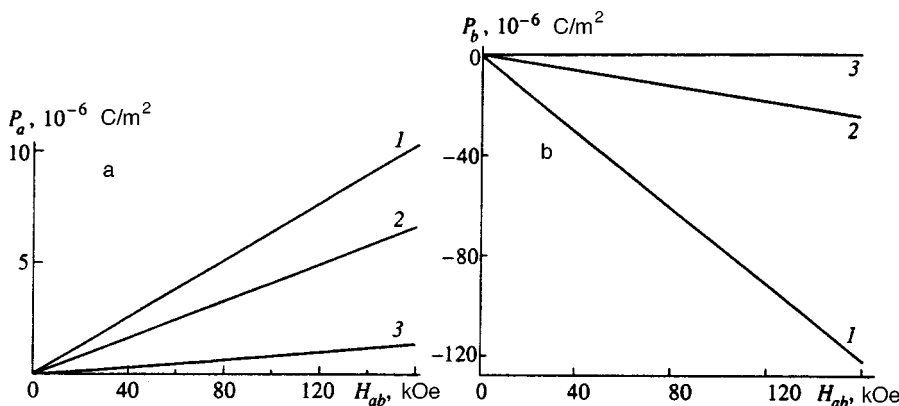
FIG. 4. Field dependence of the linear contributions to the electric polarizations $P_a(H_{ab})$ (a) and $P_b(H_{ab})$ (b) of a $Ga_{1.15}Fe_{0.85}O_3$ crystal at various temperatures: *1*—10 K; *2*—150 K; *3*—194 K.
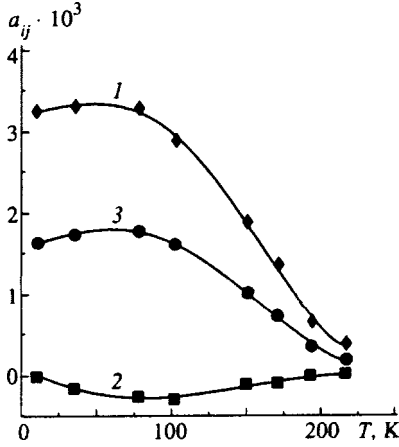
FIG. 5. Temperature dependence of the magnetoelectric susceptibilities of a Ga$_{1.15}$Fe$_{0.85}$O$_3$ crystal: 1—α_{bc}; 2—α_{cb}; 3—$\alpha_1^a T_a = (\alpha_{bc} - \alpha_{cb})/2$.
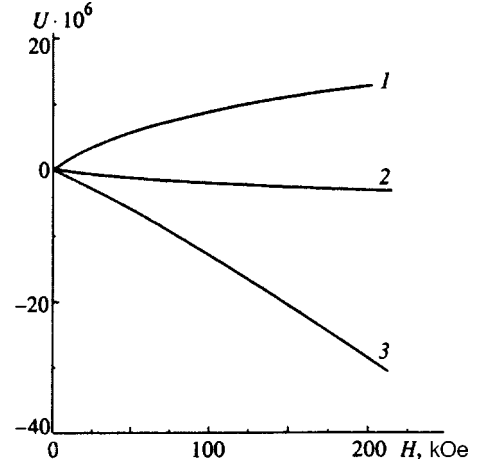

FIG. 6. Field dependence of the longitudinal magnetostriction along the b (1), c (2), and a (3) axes for a Ga$_{1.15}$Fe$_{0.85}$O$_3$ crystal at 78 K.

$$\mathbf{T}^0 = \frac{\mu_B}{2} \sum_\alpha \mathbf{r}_\alpha^0 \times \mathbf{S}_\alpha. \tag{11}$$

An estimate of the ratio of the second term to the first term in Eq. (10) with consideration of the experimental data in Refs. 13 and 14 yields the small quantity

$$\epsilon = |r^1| M / T^0, \tag{12}$$

and this provides some assurance that it is the appropriate expansion parameter in perturbation theory. Using the experimental results in Refs. 13 and 14, which specify S_α and r_α, and taking into account that $\mathbf{M} = (0,0,M_c)$, where $M_c = 3.84\mu_B$ per unit cell, we obtain[1] $\mathbf{T}^0 = (T_a, 0, 0)$, where $T_a = 24.155\mu_B$ Å per unit cell, whence $\epsilon \simeq 0.03$.

3. The electric polarization vector can be represented in the following form:

$$P_i = P_i^s + \alpha_{ij} H_j + \frac{1}{2}\beta_{ijk} H_j H_k, \tag{13}$$

where P_i^s is the spontaneous polarization vector, and α_{ij} and β_{ijk} are the magnetoelectric tensors. The second term in (13) describes the linear magnetoelectric effect. The tensor α_{ij} can be separated into symmetric and antisymmetric parts: $\alpha_{ij} = \alpha_{ij}^a + \alpha_{ij}^s$. The antisymmetric part α_{ij}^a can be represented in a linear approximation with respect to $\mathbf{M}$ and $\mathbf{T}$ as

$$\alpha_{ij}^a = \alpha_{ijk}^{a1} T_k + \alpha_{ijk}^{a2} M_k, \tag{14}$$

where α_{ijk}^{a1} and α_{ijk}^{a2} are, respectively, axial and polar third-rank tensors. The antisymmetric tensor α_{ijk}^{a2} vanishes for the

centrosymmetric *mmm* structure, while α_{ijk}^{a1} is conserved.[21] Therefore, to lowest order in the small parameter ϵ introduced above the tensor α_{ij}^a can be represented as

$$\alpha_{ij}^a = \alpha_{ijk}^{a1} T_k^0, \tag{15}$$

where $\mathbf{T}^0$ is defined in (11).

The symmetric tensor α_{ij}^s can be represented to first order in M and T as

$$\alpha_{ij}^s = \alpha_{ijk}^{s1} M_k + \alpha_{ijk}^{s2} T_k, \tag{16}$$

where α_{ijk}^{s1} and α_{ijk}^{s2} are polar and axial third-rank tensors, which are symmetric with respect to the indices ij. The formulas for the magnetoelectric effect can be written with allowance for (13)–(16) in the form (see also Table I)[2]

$$P_b = \{\alpha_{bcc} M_c + (\alpha_1^s + \alpha_1^a) T_a\} H_c + \{\alpha_{baa} M_a + (\alpha_2^s$$
$$- \alpha_2^a) T_c\} H_a + \eta_{bbb} P_b^s H_b + \beta_{bcc} H_c^2 + \beta_{baa} H_a^2$$
$$+ \beta_{bbb} H_b^2, \tag{17}$$

$$P_c = \{\alpha_{cbc} M_c + (\alpha_1^s - \alpha_1^a) T_a\} H_b + \beta_{ccb} H_c H_b, \tag{18}$$

$$P_a = \{\alpha_{aba} M_a + (\alpha_2^s + \alpha_2^a) T_c\} H_b + \beta_{aab} H_a H_b, \tag{19}$$

where, by definition,

$$\alpha_1^s = \frac{a_{bca} + \alpha_{cba}}{2}, \quad \alpha_1^a = \frac{a_{bca} - \alpha_{cba}}{2},$$

TABLE I. Irreducible representations of the C_{2v}^9 group.

	$2_z\sigma_x\sigma_y$	**M, H**	**T**	**P**	U_{mn}	**T**×**H**	$M_i H_j$	$H_i H_j$
Γ_1	$+1+1+1$		T_b	P_b	U_{aa}, U_{bb}, U_{cc}	$T_c H_a - T_a H_c$	$M_a H_a, M_b H_b, M_c H_c$	H_a^2, H_b^2, H_c^2
Γ_2	$+1-1-1$				U_{ac}	0	$M_c H_a, 0, M_a H_c$	$H_a H_c$
Γ_3	$-1-1+1$		T_c	P_c	U_{ab}	$T_a H_b - T_b H_a$	$M_b H_a, M_a H_b, 0$	$H_a H_b$
Γ_4	$-1+1-1$		T_a	P_a	U_{bc}	$T_b H_c - T_c H_b$	$0, M_c H_b, M_b H_c$	$H_b H_c$
Γ_1'	$+1+1+1$							
Γ_2'	$+1-1-1$	M_b, H_b						
Γ_3'	$-1-1+1$	M_a, H_a						
Γ_4'	$-1+1-1$	M_c, H_c						

$$\alpha_2^s = \frac{a_{abc} + \alpha_{bac}}{2}, \quad \alpha_2^a = \frac{a_{abc} - \alpha_{bac}}{2},$$

and β_{ijk} is a polar tensor, which is symmetric with respect to the indices ij.

4. In the following we assume, in accordance with the experimental conditions described above (Fig. 3), that the magnetic field vector is oriented in the bc plane of the crystal:

$$\mathbf{H} = (0, H_b, H_c). \tag{20}$$

This choice of the orientation of the magnetic field enables us to measure both polarizations, P_b and P_c, in the same magnetic structure. In addition, we assume that in a weak field

$$\mathbf{M} = (0, 0, M_c) \quad \text{and} \quad \mathbf{T} = (T_a, 0, 0), \tag{21}$$

in agreement with the experimental data in Ref. 11 and the calculations presented above in Sec. 2. Only the components α_{bca} and α_{cba} ($\alpha_{bcc} = \alpha_{cbc}$ by definition) are needed to determine P_b and P_c [see (17) and (18)].

Let us consider the contribution to the electric polarization $P(H)$ that is linear with respect to magnetic field. It follows[3)] from (17) and (21) that the contribution to the electric polarization that is linear with respect to magnetic field has the form

$$P_b^l - P_c^l = 2\alpha_1^a T_a H, \quad \text{or} \quad \alpha_1^a T_a = \frac{\alpha_{bc} - \alpha_{cb}}{2}. \tag{22}$$

Here it is assumed that $H_b = H_c = H$. It follows from (22) that the differences observed in Figs. 3 and 5 between $P_b(H)$ and $P_c(H)$, as well as the differences between α_{bc} and α_{cb}, are directly related to the existence of the toroidal moment T_a. Figure 5 also shows the temperature dependence of $\alpha_1^a T_a$, which is proportional to the toroidal moment.

5. When the magnetic field oriented in the ba plane of the crystal

$$\mathbf{H} = (H_a, H_b, 0) \tag{23}$$

is sufficiently large, reorientation of the magnetization of the crystal from the c axis to the a axis takes place. In this case the contribution to the electric polarization that is linear with respect to magnetic field can be represented, according to (17)–(19), in the form

$$P_b^l - P_a^l = 2\alpha_2^a T_c H, \quad \text{or} \quad \alpha_2^a T_c = \frac{\alpha_{ba} - \alpha_{ab}}{2}. \tag{24}$$

Here it is assumed that $H_a = H_b = H$. A conclusion regarding the field-induced reorientation of the toroidal moment from the a axis to the c axis when the magnetization of the crystal is reoriented from the c axis to the a axis can be drawn from Fig. 4 with consideration of (24).

6. Formulas for the longitudinal magnetostriction can be written in precisely the same manner as for the electric polarization (see Table I):

$$U_{cc} = (\lambda_{ccac}^a T_a + \lambda_{cccc}^p M_c) H_c + q_{cccc}^p H_c^2, \tag{25}$$

$$U_{aa} = (\lambda_{aaac}^a T_c + \lambda_{aaaa}^p M_a) H_a + q_{aaaa}^p H_a^2, \tag{26}$$

$$U_{bb} = \nu_{bbbb}^p P_b^s H_b + \lambda_{bbbb}^p M_b H_b + q_{bbbb}^p H_b^2. \tag{27}$$

The linear field dependence of the magnetostriction $U_{aa}(H_a)$ (see Fig. 6) means that the components M_a and (or) T_c of the magnetic and toroidal moments cannot be equal to zero when $\mathbf{H}$ is parallel to the a axis. It is interesting that if $\mathbf{r}^1 \to 0$ and if α_{bcc} and α_{cbc} in (17) and (18) and λ_{ccca}^a and λ_{aaac}^a in (25) and (26) tend to zero, then the linear magnetoelectric effect is determined only by the toroidal moment, and the longitudinal magnetostriction is determined only by the magnetic moment. Hence it follows that the main contribution to the linear magnetoelectric effect is made by the toroidal moment $\mathbf{T}$ and that the main contribution to the longitudinal magnetostriction is made by the magnetic effect. This means that there is no strict relationship between the linear magnetoelectric effect and the longitudinal magnetostriction in $Ga_{2-x}Fe_xO_3$.

4. CONCLUSION

The field dependence of the magnetoelectric effect and longitudinal magnetostriction of $Ga_{2-x}Fe_xO_3$ single crystals has been investigated in magnetic fields up to 200 kOe in the temperature range from 4.2 to 300 K. It has been shown that the magnetic order in $Ga_{2-x}Fe_xO_3$ crystals allows interpreting it as toroidal spin ordering. On the basis of an experimental finding, viz., significant asymmetry of the off-diagonal components of the linear magnetoelectric susceptibility tensor, it has been concluded that toroidal ordering exists. It has been shown that the linear magnetoelectric effect in $Ga_{2-x}Fe_xO_3$ is caused by a toroidal moment, rather than by magnetostriction, as was previously postulated.

*)E-mail: popov@plms.phys.msu.su
†)Deceased.

[1)]Just as in Refs. 13 and 14, the mean values of the magnetic moments of Fe^{3+} ions in the four inequivalent sites calculated in accordance with the probability distribution (1) were used in the calculation.

[2)]For generality, we assume here that the vector $\mathbf{T}$ can have a nonzero component T_c in a magnetic field of arbitrary orientation.

[3)]As is observed experimentally, it is assumed that the magnitude of the projection of the magnetic field onto the b axis is insufficient to cause significant deviation of the spins from the a axis of the crystal.

[1]J. P. Remeika, J. Appl. Phys. Suppl. **31**, 26 (1960).
[2]D. L. White, Bull. Am. Phys. Soc. **5**, 189 (1980).
[3]C. H. Nowlin and R. V. Jones, J. Appl. Phys. **34**, 1962 (1963).
[4]A. Pinto, J. Appl. Phys. **37**, 4372 (1966).
[5]E. A. Wood, Acta Crystallogr. **13**, 682 (1960).
[6]S. C. Abraham, J. M. Reddy, and J. L. Bernstein, J. Chem. Phys. **42**, 3957 (1965).
[7]G. T. Rado, in *Proceedings of the International Conference on Magnetism, Nottingham, 1964*, Institute of Physics and Physical Society, London (1965), p. 361.
[8]G. T. Rado, J. Appl. Phys. **37**, 1403 (1966).
[9]T. M. Perekalina, E. M. Smirnovskaya, I. S. Zheludev, V. A. Timofeeva, A. B. Bykov, Kristallografiya **32**, 795 (1987) [Sov. Phys. Crystallogr. **32**, 469 (1987)].

[10] Yu. F. Popov, A. K. Zvezdin, G. P. Vorob'ev, and A. M. Kadomtseva, JETP Lett. **57**, 69 (1993).

[11] G. T. Rado, Int. J. Magn. **6**, 121 (1974).

[12] G. T. Rado, Phys. Rev. Lett. **13**, 335 (1964).

[13] M. F. Bertaut, G. Buisson, J. Chappert, and G. Bassi, C. R. Acad. Sci. **260**, 3355 (1965).

[14] A. Delapalme, J. Phys. Chem. Solids **28**, 1451 (1967).

[15] Ya. B. Zel'dovich, Zh. Éksp. Teor. Fiz. **33**, 1531 (1957) [Sov. Phys. JETP **6**, 1184 (1958)].

[16] V. M. Dubovik and A. A. Cheshkov, Fiz. Elem. Chastits At. Yadra **5**, 791 (1974) [Sov. J. Part. Nucl. **5**, 318 (1975)].

[17] H. Schmid, Int. J. Magn. **4**, 337 (1973).

[18] A. A. Gorbatsevich and Yu. F. Tugushev, Zh. Éksp. Teor. Fiz. **85**, 1107 (1983) [Sov. Phys. JETP **58**, 643 (1983)].

[19] A. A. Gorbatsevich and Yu. V. Kopaev, Ferroelectrics **161**, 321 (1994).

[20] V. M. Dubovik and L. A. Tusunyan, Fiz. Elem. Chastits At. Yadra **14**, 1193 (1983) [Sov. J. Part. Nucl. **14**, 504 (1983)].

[21] R. R. Birss, *Symmetry and Magnetism*, North-Holland, Amsterdam (1964), p. 245.

Critical state of a thin superconducting plate in an inclined magnetic field

A. A. Zhukov,*) N. A. Joukov, and M. G. Mikheev†)

Physics Department, M. V. Lomonosov State University, 119899 Moscow, Russia

G. K. Perkins and A. D. Caplin

Centre for High Temperature Superconductivity, Blackett Laboratory, Imperial College, London SW7 2BZ, United Kingdom

Yu. V. Bugoslavsky

Institute of General Physics, 117942 Moscow, Russia
(Submitted 8 July 1997)
Zh. Éksp. Teor. Fiz. **114**, 273–290 (July 1998)

A solution is obtained for the Bean critical-state model in an inclined magnetic field which leads to localization of the irreversible magnetic moment induced by the shielding currents. An experimental investigation of $YBa_2Cu_3O_y$ single crystals and epitaxial films, and also plates of classical superconductors using vector magnetometry yields results fully consistent with the model calculations. The established geometric effects have an important influence on measurements of the angular dependence of the critical current and its anisotropy in a high-temperature superconductor. © *1998 American Institute of Physics.* [S1063-7761(98)01907-6]

1. INTRODUCTION

A characteristic feature of single crystals of high-temperature superconductors is that they have the form of thin plates. Various previous studies (Refs. 1–3, for example) have reported the observation of strong angular anisotropy of the irreversible magnetic moment, which was attributed to the anisotropy or the two-dimensional nature of the superconductivity in these materials. The importance of the sample geometry for the angular dependence of the magnetic moment was noted earlier in Refs. 4 and 5. However, no quantitative model which reflects this influence has been developed. In Sec. 2 we obtain a solution for the Bean critical-state model for an infinite, thin, superconducting parallelepiped in an inclined magnetic field. We also consider some generalizations to the case of more realistic model descriptions. In Sec. 3 the results of the calculations will then be compared with experimental investigations of the vector of the magnetic moment in high-temperature superconducting single crystals and in plates of conventional superconductors. Finally, in Sec. 4 we examine the role of geometric factors in the behavior of the magnetic characteristics of high-temperature superconducting single crystals, especially in measurements of the critical current anisotropy and various pinning mechanisms.

2. THEORY

Key factors which complicate the analysis of the current distribution for the critical state in a superconducting plate are the three-dimensional nature of the problem, the loss of symmetry when the magnetic field deviates from the principal axes, and the shape of the current–voltage characteristic.

Thus, we shall initially analyze the simplest case of the Bean model for an infinite parallelepiped and then its generalization.

2.1 Bean model for an infinite parallelepiped in an inclined magnetic field

The basic assumption of the Bean model is that the density **j** of the shielding currents inside a superconductor has a constant value j_c and their direction is determined by the magnetic prehistory. The steady-state case of fixed vortices was studied in the original work.[6] The real situation differs from steady-state because of vortex motion accompanying changes in the magnetic field or due to the thermally activated magnetic flux creep in the case of a fixed magnetic field. The changes in the magnetic flux accompanying the vortex motion induce an electric field **E** which is responsible for the generation of shielding currents. These currents can be described using local electrodynamics equations.[7–9] In addition, the influence of dynamic processes is frequently weak and in this case, the bean model[6] is a good approximation to describe the magnetic properties of superconductors. Formally, this corresponds to the solution of local electrodynamics equations with a step-like current–voltage characteristic $\mathbf{j} = j_c \mathbf{E}/E$.

As in the original model,[6] we shall assume that j_c does not depend on the magnitude and direction of the local magnetic induction **B** in the sample. We shall also confine our analysis to high values of the applied magnetic field **H** so that the influence of the self-field generated by the shielding currents can be neglected. In this case, the approximation $\mathbf{B} = \mu_0 \mathbf{H}$ is valid inside the superconductor, which corresponds to a uniform distribution of vortices parallel to the applied field **H**.

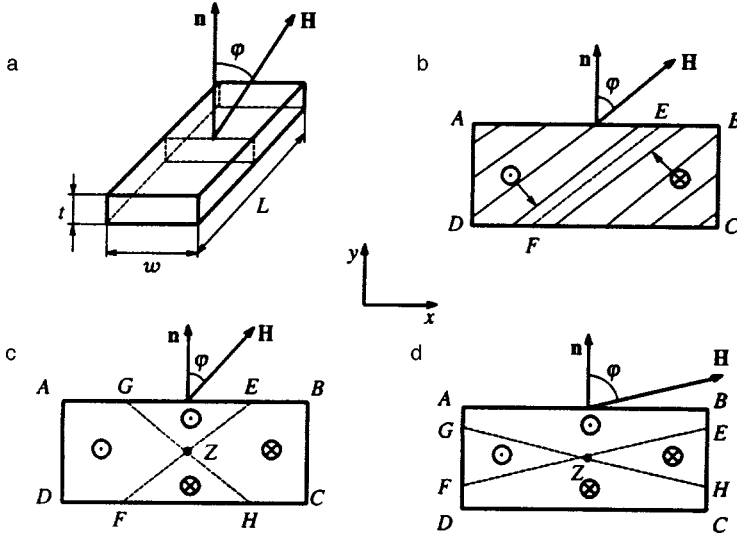Even with these approximations, a solution cannot be

FIG. 1. a) Diagram of sample; b) directions of shielding currents ($\odot$ and $\otimes$) and Lorentz force (arrows) acting on the vortices (bold lines); c) distribution of shielding currents for $\varphi < \varphi_c$; d) distribution of shielding currents for $\varphi > \varphi_c$.

obtained for a finite parallelepiped because of the loss of translational invariance. Thus, we shall consider an infinite parallelepiped (as the limiting case of a long plate whose length L is considerably greater than its width w and thickness) which reduces the three-dimensional problem to two-dimensional. The ratio of the width w to the height t of the cross section ($\kappa = w/t$), which we shall call the aspect ratio, is a key parameter in the following analysis.

We shall assume that the direction of the magnetic field rotates in the xy transverse plane and forms the angle φ with the normal **n** to the principal plane of the parallelepiped (Fig. 1). For an infinite plate, all the cross sections are equivalent as a result of translational invariance. From symmetry considerations, the vector **B** is always located in the transverse plane and as **H** increases, vortices enter from the corners A and C (Fig. 1b), move toward the center of the sample, and meet on the EF plane.[4] In this case, symmetry indicates that the shielding currents should be directed parallel to the z axis. The Lorentz forces, like the vortex velocity, are parallel to the xy plane but have opposite signs on different sides of the EF plane. In the Bean model the density of the shielding currents always has the same value j_c. Thus, the distribution of the shielding currents is antisymmetric relative to the EF plane. This conclusion, based on symmetry, will be confirmed below by solving the appropriate Maxwell equations.

The angular function of the magnetic moment differs depending on whether the angle of inclination of the magnetic field is smaller or larger than the angle of the diagonal of the parallelepiped cross section $\varphi_c = \arctan \kappa$. Thus, we shall consider these cases separately.

a) $\varphi \leq \varphi_c$. For our analysis we shall use a Cartesian coordinate system with the xy plane parallel to the plane of rotation of the magnetic field, and the y axis parallel to the principal normal **n** (Fig. 1c). The magnetic moment **m** of the shielding currents can be calculated using the relation

$$\mathbf{m} = \frac{1}{2} \int \mathbf{r} \times \mathbf{j} dV.$$

For **H** directed along the axes of symmetry of the parallelepiped, the closure currents give exactly the same contribu-

tion as the currents parallel to the long side.[7] In Sec. 2.3 we shall show that this statement is valid for an arbitrary direction of **H**. Thus, we have

$$m_y = L \int xj \, dS, \quad m_x = L \int yj \, dS,$$

where the integration is performed over the cross section of the sample $z = $ const.

From symmetry concepts, the region $GEZFH$ where the plane shown by the line GH corresponds to the $-\varphi$ direction (Fig. 1c) makes a zero contribution to the magnetic moment m_y. In consequence, the value of m_y can be calculated as twice the contribution of the trapezoidal region $AGZFD$:

$$m_y = 2L \int_{AGZFD} xj_c dx \, dy = \frac{Lt^3}{12} j_c (3\kappa^2 - \tan^2 \varphi). \quad (1)$$

To calculate the moment m_x, integration is performed over the region $GEZFH$:

$$m_x = 2L \int_{GEZ} yj_c dx \, dy = \frac{Lt^3}{6} j_c \tan \varphi. \quad (2)$$

The angle α between the direction of the magnetic moment **m** and the principal normal **n** to the plane is given by

$$\tan \alpha = \frac{2 \tan \varphi}{3\kappa^2 - \tan^2 \varphi}. \quad (3)$$

For $\varphi = 0$ we obtain the well-known result: $m_y = Lj_c tw^2/4$ and obviously $\alpha = 0$. For angles satisfying $\tan \varphi \ll \kappa$, we have $\tan \alpha \approx (2/3\kappa^2)\tan \varphi$. When the angle between the direction **H** and the normal reaches φ_c, the magnetic moment is deflected by the angle $\alpha_c = \arctan \kappa$, where $\alpha_c = \pi/2 - \varphi_c$. Thus, for a typical high-temperature superconducting single crystal with $\kappa \sim 10$ or more the irreversible magnetic moment is almost locked to the principal normal of the parallelepiped for the entire range of magnetic field directions $\varphi \leq \varphi_c$. At the boundary of this range, $m_y(\varphi_c) = 2m_y(0)/3$ and $m_x(\varphi_c) = m_y(\varphi_c)/\kappa$ and thus $|\mathbf{m}|$ will also not vary very substantially.

b) $\varphi \geqslant \varphi_c$. In the angular range $\varphi_c \leqslant \varphi \leqslant \pi/2$ (Fig. 1d), it is more convenient to measure the deviations of **H** and **m** from the principal plane, using the angles $\theta = \pi/2 - \varphi$ and $\beta = \pi/2 - \alpha$, respectively. Then, this case is equivalent to the previous one if we substitute $x \leftrightarrow y$, $\varphi \leftrightarrow \theta$, $t \leftrightarrow w$, and $\kappa \leftrightarrow 1/\kappa$, and as a result we obtain

$$m_x = 2L \int_{AGZEB} y j_c dx\, dy = \frac{Lw^3}{12} j_c \left(\frac{3}{\kappa^2} - \tan^2 \theta \right), \quad (4)$$

$$m_y = 2L \int_{EZH} x j_c dx\, dy = \frac{Lw^3}{6} j_c \tan \theta, \quad (5)$$

$$\tan \beta = \frac{2\kappa^2 \tan \theta}{3 - \kappa^2 \tan^2 \theta}. \quad (6)$$

If **H** (and also **m**) is parallel to the principal plane, the absolute value of the magnetic moment is $|\mathbf{m}| = L j_c w t^2/4$, which is κ times smaller than that for the magnetic field parallel to the principal normal. It should be noted that even a very small angular deviation of the magnetic field **H** induces a large deviation of the magnetic moment **m** from the principal plane: $\beta \approx 2\kappa^2 \theta/3$. A convenient parameter to describe the rotation of the magnetic moment is

$$\tau = \frac{d\beta}{d\theta} \bigg|_{\theta=0} = \frac{2\kappa^2}{3}. \quad (7)$$

For the critical angle $\theta_c = \pi/2 - \varphi_c$, the direction of the magnetic moment is determined by $\tan \beta_c = \kappa$, which is equivalent to the result obtained in case (a).

From Eqs. (4) and (5) in the lowest order of the expansion in terms of θ, we can obtain the following expression for the absolute value of the magnetic moment

$$m(\theta) \approx m(0) \left(1 + \frac{2}{9} \kappa^4 \theta^2 \right). \quad (8)$$

Thus $|\mathbf{m}|$ increases rapidly with increasing angle θ.

2.2. Magnetic flux creep and anisotropy

The step-like current–voltage characteristic used in the Bean model is not a good approximation for high-temperature superconducting materials, since these typically exhibit strong magnetic flux creep which leads to smoothing of the current–voltage characteristic. The power dependence $E/E_0 = (j/j_0)^n$ better describes the behavior of the experimental current–voltage characteristics with typical values of n between 5 and 30 for high-temperature superconducting single crystals and above 20 for conventional superconductors. In this case, the Bean model corresponds to the limit $n \to \infty$.

During magnetic measurements, changes in the magnetic field (which for simplicity we assume to take place at the constant rate $\dot{H} = dH/dt$) induce an electric field **E**. If the magnetic field is fixed, the electric field is formed as a result of magnetic flux creep. The equations of local electrodynamics and the current–voltage characteristics can be used to calculate the shielding current distribution. These currents change direction as the direction **E** changes and determine the behavior of the irreversible part of the magnetic

moment.[7–9] In addition, there are undamped superconducting currents which do not depend on the direction **E** and determine the reversible magnetization. However, this contribution to the magnetization is usually small and these currents will be neglected in the present study.

As we have already noted, we shall analyze the case when H appreciably exceeds the self-field of the shielding currents and thus their influence on E can be neglected. In addition, as a result of the homogeneity of **B** in the bulk of the sample for this case, we can easily take into account the dependence of the current–voltage characteristic on B.

For an infinite parallelepiped, the power current–voltage characteristic can easily be included in the previous analysis. As will be shown in Sec. 2.3, the electric field **E** parallel to the z axis has opposite signs on different sides of the EF plane in Fig. 1 and is given by

$$E_z = \mu_0 \dot{H}(x \cos \varphi - \sin \varphi). \quad (9)$$

By means of lengthy but elementary calculations for $\varphi > \varphi_c$ (Fig. 1d), we can obtain

$$
\begin{aligned}
m_y = 2L j_0 \left(\frac{w}{2} \right)^3 & \left(\frac{\mu_0 \dot{H} w \sin \theta}{2E_0} \right)^{1/n} \tan \theta \frac{n^2}{(2n+1)(n+1)} \\
& \times \Big\{ (a+1)^{2+1/n} + (a-1)^{2+1/n} - \frac{n}{3n+1}[(a+1)^{3+1/n} \\
& - (a-1)^{3+1/n}] \Big\}
\end{aligned} \quad (10)
$$

and

$$
\begin{aligned}
m_x = 2L j_0 \left(\frac{w}{2} \right)^3 & \left(\frac{\mu_0 \dot{H} w \sin \theta}{2E_0} \right)^{1/n} \frac{\tan \theta}{\kappa} \frac{n^2}{(2n+1)(n+1)} \\
& \times \Big\{ (a+1)^{2+1/n} + (a-1)^{2+1/n} - \frac{n}{3n+1}[(1+a^{-1}) \\
& \times (a+1)^{2+1/n} - (1-a^{-1})(a-1)^{2+1/n}] \Big\}.
\end{aligned} \quad (11)
$$

The angle β between **m** and the principal plane is given by

$$\tan \beta = \frac{2(2n+1)(n+1)a\kappa}{6n(n+1)a^2 - (2n+1)(n-1)}, \quad (12)$$

where $a = (\kappa \tan \theta)^{-1}$.

These complex expressions have two simple limiting cases. For $n \to \infty$, i.e., for the Bean model they are identical to the expressions derived in Sec. 2.1.b. For an Ohmic current–voltage characteristic ($n = 1$) they are simplified considerably:

$$m_y = \frac{Lw^3}{6} j_z \left(0, \frac{t}{2} \right) \tan \theta, \quad (13)$$

$$m_x = \frac{Lw^3}{6\kappa^2} j_z \left(0, \frac{t}{2} \right), \quad (14)$$

$$\tan \beta = \kappa^2 \tan \theta. \quad (15)$$

These results are exactly identical to Eqs. (4), (5), and (6) if the critical current j_c in the Bean model is replaced by the current density $j_z(0,t/2)$ flowing at the center of the principal plane of the parallelepiped.

A solution for the angles $\varphi < \varphi_c$ can easily be obtained by using the new notation for the parameters introduced above. For $n \to \infty$ this naturally corresponds to the expressions of the Bean model. For an Ohmic current–voltage characteristic these solutions are almost identical to Eqs. (1), (2), and (3) if we use the current density $j_z(w/2,0)$ flowing at the center of the lateral plane of the parallelepiped.

Assuming a small difference between the Bean and Ohmic limits, we can confirm that the angular dependence of the magnetic moment is almost unaffected by the type of current–voltage characteristic. Thus, we can predict that the results obtained earlier for the Bean model should accurately describe the experimental data.

We shall analyze the situation when the self-fields have little influence. In this case, the magnetic induction **B** may be considered to be constant inside the sample. The dependence of j_c (or j_z) on **B** can then easily be taken into account by using the value of j_c corresponding to the applied field $\mu_0 \mathbf{H}$ in the expressions obtained. This reflects the field dependence $j_c(B)$ and the pinning anisotropy determined by the direction of **B**. It is easily established that the direction of the vector **m** (see Eqs. (3), (6), (7), and (12)) remains almost constant for any dependence of j_c on **B**. Thus, the pinning anisotropy has no influence on the geometric localization effect. However, $|\mathbf{m}|$ (Eqs. (1), (2), (4), (5), (10), and (11)) does depend on the pinning anisotropy and may be used to study the angular dependence of the critical current.

2.3. Finite length and closure currents

We shall first show that the closure currents account for exactly half the contribution to the magnetic moment. By definition, this is given by

$$\mathbf{m} = \frac{1}{2} \int \mathbf{r} \times \mathbf{j}\, dV = \frac{1}{2}[\mathbf{e}_x(j_z y - j_y z) + \mathbf{e}_y(j_x z - j_z x)$$
$$+ \mathbf{e}_z(j_y x - j_x y)]dV, \tag{16}$$

where $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$ are the unit vectors of the Cartesian coordinate system.

It can be seen from this equation that each component of the magnetic moment **m** is determined by the sum of two contributions. Previously, we only took into account the contribution of the current component j_z and we asserted that the closure currents corresponding to the components j_x and j_y make exactly the same contribution. We shall prove this statement.

We shall consider a sample of arbitrary shape in which the shielding currents **j** have a steady-state distribution. In this case, it follows from the charge conservation law that div $\mathbf{j} = 0$ and in consequence, the field of the current **j** is solenoidal. Thus, the distribution of the shielding currents can be represented as closed, nonintersecting tubes which are not generally planar.

The component of the magnetic moment δm_x parallel to the x axis and generated by an elementary current tube may be expressed on the basis of Eq. (16) in the form

$$\delta m_x = \oint s\, dl(j_z y - j_y z), \tag{17}$$

where s is the cross-section area of an elementary tube, dl is the differential of the tube length, and integration is performed over the contour defined by the current tube. If we introduce the unit vector $\mathbf{e}_t$ in the direction of the elementary current tube, we can write the following expressions: $j_z = j\mathbf{e}_z \cdot \mathbf{e}_t$ and $dz = dl\mathbf{e}_z \cdot \mathbf{e}_t$ which, assuming $i = \mathbf{s} \cdot \mathbf{j}$, gives

$$\delta m_x = \frac{1}{2}i \oint dl(\mathbf{e}_z \cdot \mathbf{e}_t y - \mathbf{e}_y \cdot \mathbf{e}_t z) = \frac{1}{2}i \oint dl(dz\, y - dy\, z). \tag{18}$$

Taking into account that the area of the projection of the current tube S_{yz} is given by the equivalent expressions $\oint y\,dz \equiv -\oint z\,dy$, both terms in Eq. (18) make the same contribution to δm_z, which corresponds to the definition of the elementary magnetic moment $\delta m_x = iS_{yz}$, this proves the assertion formulated above.

We shall now analyze the distribution of shielding currents for a superconductor with an arbitrary current–voltage characteristic. If the characteristic does not depend on the magnetic field (as in the Bean model), the changes in the magnetic induction in the critical state are determined only by changes in the applied field **H**. If the rate of change of **H** is constant, the distribution of the shielding currents is steady-state. This assumption is a good description of the experimental situation if the changes in the magnetic flux caused by the self-field are negligible (in this case, the self-field need not necessarily be small).

We shall initially postulate that no space charges occur in the superconductor and their density is everywhere $\rho = 0$ (later, we shall show that these space charges may occur in certain cases). The electric field distribution can then be determined from the Maxwell equations using local electrodynamics:[8,9]

$$\text{div } \mathbf{E} = 0 \tag{19}$$

and

$$\text{curl } \mathbf{E} = -\mu_0 \frac{d\mathbf{H}}{dt} \tag{20}$$

with boundary conditions whereby the normal component of the electric field is zero at the surface of the sample, $E_n = 0$.

It is known that these equations have a unique continuous solution.[10] For an infinite parallelepiped where the field **H** is parallel to the y axis (Fig. 1), the electric field distribution is given by

$$E_z = \mu_0 \dot{H}_y x, \quad E_x = E_y = 0. \tag{21}$$

As a result of the linearity of Eqs. (19) and (20), it is easy to find a solution for an arbitrary direction of the magnetic field in the xy plane:

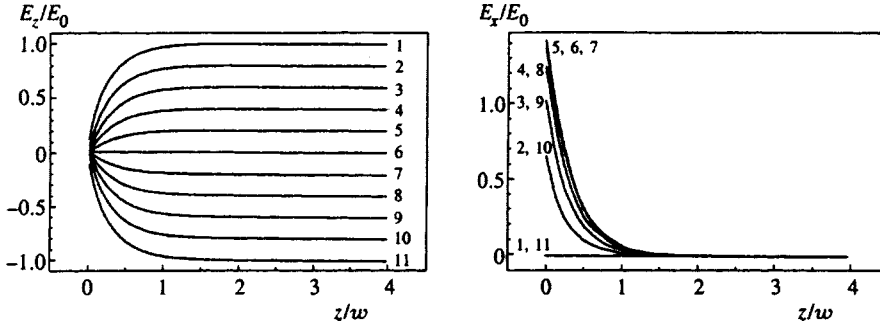$$E_z = -\mu_0 \dot{H}(y \sin \varphi - x \cos \varphi), \quad E_x = E_y = 0. \tag{22}$$

FIG. 2. Electric field in a semi-infinite super-conducting strip. The curves corresponds to $x/z=0$ (*1*), 0.1 (*2*), 0.2 (*3*), 0.3 (*4*), 0.4 (*5*), 0.5 (*6*), 0.6 (*7*), 0.7 (*8*), 0.8 (*9*), 0.9 (*10*), 1.0 (*11*); $E_0=\mu_0\dot{H}_y w/2$. For E_x curve *6* is almost the same as curves *5* and *7*. The pairs of curves *1*, *11*; *2*, *10*; *3*, *9*; *4*, *8*, and *5*, *7* are the same as a result of symmetry.

For a sample of finite length L, we can obtain an accurate solution in series form. When the field **H** is parallel to the y axis, this solution is given by

$$E_z(x,z)=\frac{\mu_0\dot{H}_y}{4\pi}\sum_{p,q=-\infty}^{\infty}(-1)^{p+q}[f(p+1,q)-f(p,q)$$

$$+g(p,q+1)-g(p,q)],\qquad(23)$$

where

$$f(p,q)=(pL-z)\ln\frac{[(q+1)w-x]^2+(pL-z)^2}{(qw-x)^2+(pL-z)^2},$$

$$g(p,q)=2(qw-x)\left[\arctan\frac{(p+1)L-z}{qw-x}-\arctan\frac{pL-z}{qw-x}\right].$$
$$(24)$$

For E_x a solution can be obtained by making the substitutions $q\leftrightarrow p$, $l\leftrightarrow w$, and $x\leftrightarrow z$ in Eqs. (23) and (24). The results of the calculations for a semi-infinite strip ($0<z<L\rightarrow\infty$, $p=0,-1$) are plotted in Fig. 2. These demonstrate that formula (21) accurately describes the behavior of **E** at the distance $\sim w$.

For a qualitative analysis we shall consider an approximate solution which shows good agreement with experiment[11] and satisfies Eqs. (19) and (20). In the central region the solution is given by Eq. (21) (**H** parallel to y). Near the ends of the parallelepiped $L/2>|z|>L/2-w/2+|x|$, the solution of the system (19) and (20) will be given by

$$E_x=-\mu_0\dot{H}_y\left(|z|-\frac{L}{2}+\frac{w}{2}\right)\text{sign}|z|,\quad E_z=E_y=0.\quad(25)$$

As a result of the linearity of Eqs. (19) and (20), expressions for **E** for an arbitrary direction of the magnetic field **H** in the xy plane can be obtained from a linear combination of Eqs. (21) and (25). Using the current–voltage characteristics from the Bean model, $\mathbf{j}=j_c\mathbf{E}/E$, and Eqs. (22), we can obtain the shielding current distribution shown in Fig. 1, which is consistent with the result obtained from symmetry concepts. However, if this current–voltage characteristic is applied to the distribution (25), which describes **E** in the closing regions of a finite parallelepiped, for directions of the magnetic field differing from the axes of symmetry, the condition

$$\text{div }\mathbf{j}=0,\qquad(26)$$

corresponding to the law of charge conservation in the steady-state case, is violated. We assume that in this case, the onset of magnetic field sweeping will be accompanied by the formation of space charges which should satisfy condition (26). This changes the distribution of **E** which should obey the condition div $\mathbf{E}=\rho(\mathbf{r})$ rather than Eq. (19). An analytic solution of this nonlinear equation cannot be obtained. It should be noted that for a Bean current–voltage characteristic with the magnetic field directed along the axes of symmetry, accurate solutions can be obtained and correspond to Eqs. (21) and (25). However, as a result of the nonlinearity of the current–voltage characteristic, a solution cannot be derived by linear superposition for an arbitrary angle of inclination.

An Ohmic current–voltage characteristic $\mathbf{j}=\sigma\mathbf{E}$ with isotropic conductivity σ appreciably simplifies the situation. In this case, a solution for **j** can be obtained trivially from Eqs. (21) and (25), and Eq. (19) guarantees that div $\mathbf{j}=0$. For **H** parallel to x, we can obtain

$$m_x=\sigma\mu_0\dot{H}_x\frac{Lw^3t}{12\kappa^2}\left(1-\frac{t}{4L}\right),\qquad(27)$$

and for **H** parallel to y, we have

$$m_y=\sigma\mu_0\dot{H}_y\frac{w^3Lt}{12}\left(1-\frac{w}{4L}\right).\qquad(28)$$

In the limit of a long sample, these expressions reduce to those obtained earlier for an infinite parallelepiped.

As a result of the linearity of the initial equations, the magnetic moment for the Ohmic current–voltage characteristic is determined by the linear superposition of Eqs. (27) and (28), and its direction is given by

$$\tan\beta=\kappa^2\tan\theta\frac{1-w/4L}{1-t/4L}.\qquad(29)$$

The influence of the ends of the sample is clear from a direct comparison with Eq. (15). As was to be expected, the ensuing correction is of the order of w/L or t/L. For the more realistic scenario of a current–voltage characteristic with $n>1$, similar calculations cannot be made. However, as we showed earlier for an infinite parallelepiped, the difference between the magnetic moment **m** for the Ohmic and Bean ($n\rightarrow\infty$) current–voltage characteristics is very small. Consequently, we suppose that the results obtained above on the

TABLE I. Parameters of samples studied.

Description	Material	Form	$L \times w \times t$, mm	$\kappa = w/t$	Comments
OZ	YBa$_2$CU$_3$O$_y$	Single crystal	$1.95 \times 1.15 \times 0.065$	17.7	With TB, plane of rotation $\perp$ TB
MK	YBa$_2$Cu$_3$O$_y$	Single crystal	$1.04 \times 0.97 \times 0.10$	9.7	Detwinned
WZ	YBa$_2$Cu$_3$O$_y$	Single crystal	$0.96 \times 0.90 \times 0.04$	22.5	Without TB, high purity
AH	YBa$_2$Cu$_3$O$_y$	Single crystal	$1.26 \times 0.57 \times 0.025$	22.8	Columnar defects
LF	YBa$_2$Cu$_3$O$_y$	Epitaxial film	$2.2 \times 0.9 \times 0.0003$	3×10^3	
SH	PbMoS$_6$	Polycrystal	$2.01 \times 1.72 \times 0.40$	4.3	
Sl	PbMoS$_6$	Polycrystal	$2.01 \times 0.67 \times 0.40$	1.68	Cut from SH
VS	V$_3$Si	Polycrystal	$6.00 \times 0.90 \times 0.35$	2.57	
VI	V$_3$Si	Polycrystal	$6.00 \times 0.90 \times 0.27$	3.33	Obtained by polishing from VS

Note: TB—twin boundary.

influence of the finite length for a linear current–voltage characteristic (Eqs. (27)–(29)) can also be applied to the Bean case.

High-temperature superconducting single crystals are usually close to square with $L \sim w$. In this case, the correction to the parameter τ, characterizing the ratio of the rotation rates of the magnetic moment **m** and the magnetic field **H**, when the direction of **H** is close to the plane, is small and does not exceed 25%.

3. EXPERIMENT

3.1. Vector magnetometry

In order to check the main results of Sec. 2, we need to obtain information on the magnitude and direction of the magnetic moment **m**. For this purpose we used a vibrating-sample magnetometer (Oxford Instruments 5^H) fitted with two independent detector coils such that the components of the magnetic moment parallel (m_{st}) and perpendicular (m_{ort}) to the applied field **H** could be measured simultaneously. Standard magnetometers only measure the component m_{st}. During the measurements the sample could be rotated about the third axis with an angular resolution of 0.01° and a reproducibility better than 0.03°.

All the results of the measurements given below were obtained from the hysteresis loops of the magnetic moment with the sample orientation fixed, when the magnetic field was swept at a constant rate of ~ 10 mT s^{-1} from a negative field considerably exceeding the remagnetization field, to 5 T and back to 0 T. Here we are only interested in the irreversible component of the magnetic moment **m** which reflects the behavior of the shielding currents.

3.2. Samples

We studied several YBa$_2$Cu$_3$O$_y$ single crystals with different dominant pinning mechanisms. These were samples with twin boundaries, detwinned samples, and samples with columnar radiation tracks (see Table I). Their preparation and detwinning was described earlier in Refs. 12 and 13). The different pinning systems act differently on the anisotropy of the critical current j_c so that we could study the influence of changes in anisotropy on the angle of inclination α. An epitaxial YBa$_2$Cu$_3$O$_y$ film was used as a sample with a very large aspect ratio $\kappa = w/t$ (Ref. 14).

Conventional polycrystalline superconductors are certainly preferable for a direct check of the calculated results. A sample of the Chevrel superconductor phase PbMoS$_6$ (we shall subsequently call this SV) was cut from an ingot with a diamond saw. After the first measurements, an additional cut was made to reduce its width and thus the aspect ratio. Further measurements were then made. A sample of V$_3$Si (VS) was cut from an ingot with a diamond saw. After the first cycle of measurements, its thickness was reduced by polishing with emery paper and a second cycle of measurements carried out.

3.3. Experimental results and discussion

3.3.1. High-temperature superconducting single crystals. Figure 3 shows typical behavior of the angle of inclination α of the magnetic moment relative to the normal to the *ab* plane of a YBa$_2$Cu$_3$O$_y$ single crystal having the form of a thin plate. Over a large range of angles φ of the magnetic field up to $\sim 87°$, the direction of the magnetic moment **m** is locked to the normal **n**, so that $\alpha \approx 0°$ holds (as a result of the inevitable precession when the sample is rotated through a large range of angles, some error, of the order of $\pm 1°$, arises in the determination of α). In a narrow range of angles where the direction of **H** passes near the plane of the plate, the direction of **m** is rapidly reversed almost by 180° as predicted theoretically. The results plotted in Fig. 3 do not depend on the magnetic field (if H is appreciably greater
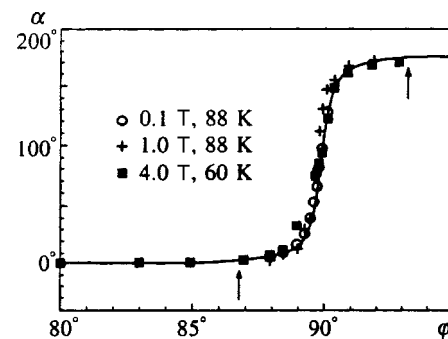

FIG. 3. Angle α determining the deviation of the vector of the irreversible magnetic moment from the c axis in a YBa$_2$Cu$_3$O$_7$ single crystal (OZ sample) when the direction φ of the applied magnetic field is rotated near the *ab* plane ($\varphi = 90°$) for various temperatures and magnetic fields, The solid curves corresponds to Eqs. (3) and (6) using the aspect ratio $\kappa = 17.7$; the arrows indicate the position of the critical angle φ_c.
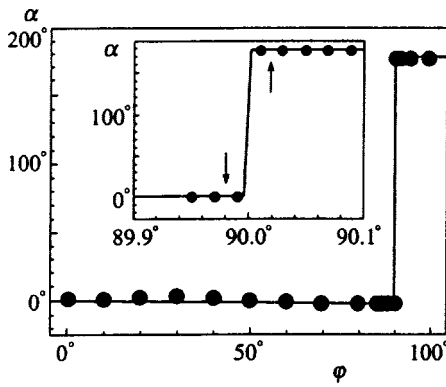
FIG. 4. Direction of the irreversible moment (using the same coordinates as in Fig. 3) in an epitaxial YBa$_2$Cu$_3$O$_7$ film (LF, $T = 50$ K, $\mu_0H = 0.5$ T). The solid curve gives the results of calculations using Eqs. (3) and (6) for the aspect ratio $\kappa = 3 \times 10^3$; the arrows indicate the critical angle φ_c.
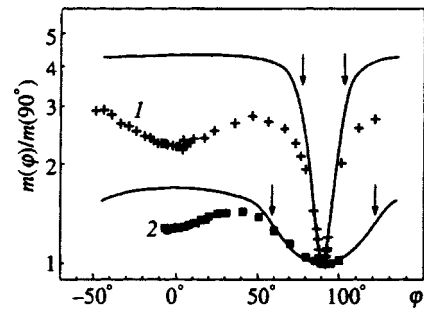

FIG. 6. Angular dependence of the magnitude of the magnetic moment in an PbMoS$_6$ sample ($T = 10$ K, $\mu_0H = 1.0$ T) in the initial state (*1*, SH) and after reducing the transverse dimension (*2*, SI). The solid curves give the calculations of $|\mathbf{m}| = \sqrt{m_x^2 + m_y^2}$ using Eqs. (1), (2), (4), and (5) for the aspect ratios $\kappa = 4.3$ and 1.68; the arrows indicate the critical angle φ_c.

than the penetration field H_p) or temperature even though the critical current j_c and its anisotropy change very substantially.[15,16] Moreover, all the YBa$_2$Cu$_3$O$_y$ samples studied, for which the magnitude and anisotropy of the critical current j_c differ substantially as a result of the predominance of different pinning mechanisms, exhibited similar behavior of $\alpha(\varphi)$ which may be described using the calculations from Sec. 2 with the measured aspect ratio and without using any fitting parameter. This confirms the dominant role of geometric effects which determine the angular behavior of the vector $\mathbf{m}$, observed as localization of its direction near the normal to the sample.

The results of the investigations of an epitaxial YBa$_2$Cu$_3$O$_y$ film (Fig. 4) reveal behavior exhibiting very strong geometric localization, when the direction of the magnetic moment $\mathbf{m}$ is reversed within a 0.02° step. These results demonstrate that even the high angular resolution of the magnetometer is insufficient to study the region of fast rotation.

3.3.2. Conventional superconductors. Measurements using conventional superconductors can be used to check these results under conditions where, unlike high-temperature superconductors, the anisotropy of the critical current is weak.

This allows us to analyze directly the dependence of $\mathbf{m}$ on the angle φ.

In this case, the experimental data for $\alpha(\varphi)$ also show good qualitative agreement with the theory without using any fitting parameter (Fig. 5). The variation of the absolute value of the magnetic moment $\mathbf{m}$ as a function of the angle φ (Fig. 6) demonstrates an important characteristic: a very narrow, sharp minimum of $|\mathbf{m}|$ when the magnetic field is directed near the plane of the plate. This dip is purely of geometric origin and is unrelated to the anisotropy of the superconducting parameters.

The results of the calculations presented in Sec. 2 suggest that for angles φ not very close to $\pi/2$, the value of $|\mathbf{m}|$ should be almost independent of the angle. The observed slight increase in the absolute value when $\mathbf{H}$ deviates from the direction $\varphi = 0$ may be attributed to uncontrollable anisotropy of the sample, formed during the preparation process, or to the influence of its finite length.

3.3.3. Angular rotation rate ratio τ. A very sensitive numerical parameter of the geometric model is the angular rotation ratio rate τ which characterizes the rotation of the magnetic moment $\mathbf{m}$ when the direction of the magnetic field $\mathbf{H}$ rotates near the plane of the plate. According to Eq. (7), the value of τ should be a simple quadratic function of κ. The results of the measurements for high-temperature superconducting samples and conventional superconductors (Fig.

FIG. 5. Rotation of the irreversible magnetic moment (using the same coordinates as in Fig. 3) in a PbMoS$_6$ sample ($T = 10$ K, $\mu_0H = 1.0$ T) in the initial state (SH, curve *1*) and after reducing the aspect ratio (SI, curve *2*). The solid curves give the results of the calculations using Eqs. (3) and (6) for the aspect ratios $\kappa = 4.3$ and 1.68, respectively. The arrows indicate the critical angle φ_c.
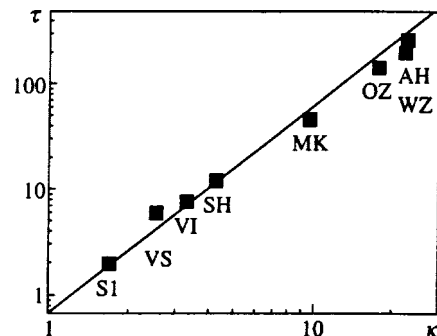

FIG. 7. Angular rotation rate ratio τ of angular rotation of the magnetic moment when $\mathbf{H}$ is directed near the plane of the plate for samples with different aspect ratios κ. The results of the calculations using Eq. (7) are given by the solid curve.
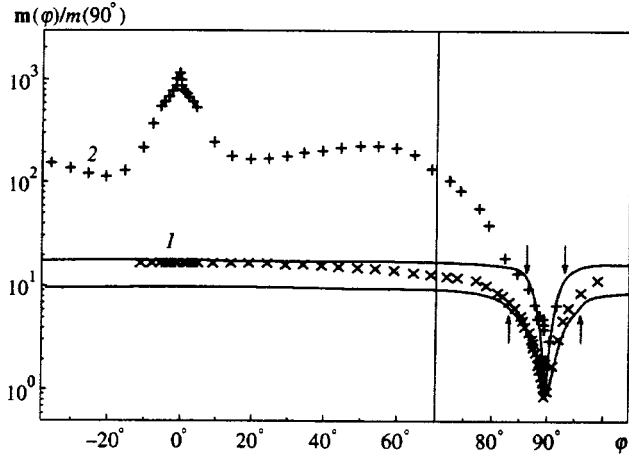
FIG. 8. Angular dependence of the magnitude of the magnetic moment $|\mathbf{m}|$ for YBa$_2$Cu$_3$O$_7$ single crystals (*1*—MK and *2*—OZ) for a fixed value of the reduced field $\varepsilon_\varphi \mu_0 H = 0.5$ T (calculated for $\Gamma^2 = 30$) at $T = 88$ K. The solid curves give the calculations of $|\mathbf{m}|$ using Eqs. (1), (2), (4), and (5) for the aspect ratios $\kappa = 9.7$ and 17.7; the arrows indicate the critical angle φ_c. A larger scale is used near $\varphi = 90°$.

7) show fairly good agreement with this equation (no points corresponding to the epitaxial film can be identified on this graph since the aspect ratio of the film $\kappa \sim 10^3$ corresponds to $\tau \sim 10^6$, which is far in excess of the experimental resolution). In addition, the slope of the dependence $\tau(\kappa^2)$ is close to 2/3, as is predicted by the proposed model.

4. GEOMETRIC EFFECTS AND MAGNETIC PROPERTIES OF HIGH-TEMPERATURE SUPERCONDUCTING SINGLE CRYSTALS

4.1. Angular scaling of the magnetization curves

As we have noted, in high-temperature superconducting single crystals over a wide range of angles of deflection of the magnetic field $\mathbf{H}$ from the c axis, the magnetic moment $\mathbf{m}$ remains almost parallel to the c axis, which indicates that shielding currents J_{ab} flowing in the ab plane predominate. The dependence of J_{ab} on the direction of the applied field is determined by the nature of the pinning mechanism. In general, this dependence is nontrivial although it is simplified appreciably for pinning by weak isotropic disorder. In this case, in an anisotropic superconductor the current J_{ab} will be constant if the parameter[17]

$$\varepsilon_\varphi B = \sqrt{\cos^2 \varphi + \Gamma^2 \sin^2 \varphi} \; B,$$

is constant, where Γ corresponds to the effective mass anisotropy (ε_φ determines the angular variations of the intrinsic vortex energy, the second critical field, the melting point of the vortex lattice, and the irreversibility field B_{irr} (Ref. 17). For an isotropic superconductor this condition corresponds to constant magnetic induction.

Thus, if this scaling law is accurate, the value of $|\mathbf{m}(\varphi)|$ measured for fixed $\varepsilon_\varphi H$ should vary very little with increasing φ and should reach 2/3 of the initial value at the critical angle φ_c (Sec. 2). Measurements for YBa$_2$Cu$_3$O$_7$ single crystals (Fig. 8, using the value $\Gamma^2 = 30$ obtained from the dependence of B_{irr}^{-2} on $\sin^2 \varphi$) showed that this behavior is indeed

observed for the detwinned single-crystal sample (MK). The solid curves in Fig. 8 give the $|\mathbf{m}(\varphi)|$ calculations using the proposed model. The experimental results show good agreement when the magnetic field is directed near the ab plane. Over a large range of angles, the magnetic moment is almost constant, as is to be expected for pinning at isotropic weak disorder (point defects or their fluctuations). The quantitative difference from the model calculations may be attributed to two factors. First, near CuO planes, it may be produced by intrinsic pinning which can reduce the value of the magnetic moment m ($\varphi = 90°$) to which the experimental curves are normalized. Second, this difference may be attributed to the finite length of the sample $L \approx w$, and also the influence of currents flowing along the c axis.

Conversely, for the twinned (OZ) crystal the dependence $m(\varphi)$ and thus J_{ab} decay rapidly when the direction of $\mathbf{H}$ deviates by $\sim 20°$ from the c axis. This behavior reflects the strong vortex pinning by twin boundaries.

For both samples, and for all the other samples studied by us, J_{ab} begins to depart from constant at $\varphi \approx 60°$, i.e., long before the angle φ_c is reached. We attribute this behavior to a kinked vortex structure when $\tan \varphi = \Gamma$ is reached.

Another manifestation of scaling at isotropic disorder is the universal angular behavior of the magnetization peak established empirically by Klein *et al.*[18] [1)]

$$m_{st}(\varphi, H/\cos \varphi) = m_{st}(0, H)\cos \varphi,$$

which is valid for angles of deviation of $\mathbf{H}$ from the c axis up to $\sim 60°$. For the crystal studied in these experiments the aspect ratio was $\kappa \sim 8$ so that in the range of angles up to $60°$, the direction of $\mathbf{m}$ deviates negligibly from the c axis. A conventional magnetometer with a single system of detector coils used in Ref. 18, measures the component m_{st} of the magnetic moment $\mathbf{m}$ parallel to the applied field $\mathbf{H}$:

$$m_{st} = |\mathbf{m}|\cos \varphi.$$

Since in YBa$_2$Cu$_3$O$_7$ the anisotropy is $\Gamma > 5$, in the range of angles studied we find

$$\varepsilon_\varphi B \sim B \cos \varphi.$$

The experimentally observed scaling law can easily be obtained from these two angular relations. Figure 9 gives the measured absolute value of the irreversible magnetic moment as a function of the reduced magnetic field $\varepsilon_\varphi B$ for various angles φ. Up to $60°$ this dependence shows almost ideal agreement. Then the magnetic moment begins to decrease rapidly. Thus, the angular behavior of the current J_{ab} observed by us indicates that the fishtail effect is associated with pinning at point defects or their clusters,[20] which induce isotropic weak disorder. The curves $J_{ab}(B)$ also have other peaks for which the scaling law examined above does not apply. These peaks are probably related to other pinning centers, such as twin boundaries.[19]

4.2. Analysis of critical current anisotropy

As a result of their layered structure, high-temperature superconducting materials are strongly anisotropic. The coupling between the CuO planes is fairly weak in YBa$_2$Cu$_3$O$_7$
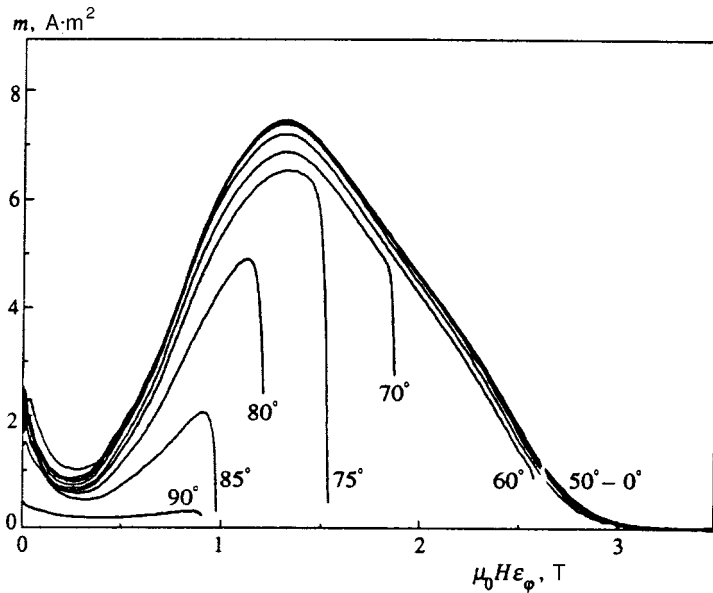
FIG. 9. Magnitude of the magnetic moment versus the reduced magnetic field for various angles φ for an MK sample at $T = 82$ K.

and very weak in BiSrCaCuO phases. The thermodynamic anisotropy Γ increases from ~ 5 in $YBa_2Cu_3O_7$ to more than 20 in BiSrCaCuO phases.

Here we shall analyze the influence of these effects on measurements of the anisotropy Ω of the (nonequilibrium) shielding currents, which is determined by the ratio of the critical currents J_{ab} and J_c flowing parallel and perpendicular to the CuO planes, respectively. This anisotropy naturally depends on the thermodynamic anisotropy. In addition, it is also influenced by the anisotropy of the pinning centers and the sample geometry.

Since high-temperature superconducting single crystals are usually thin plates with the c axis normal to the surface, geometric effects cause the shielding currents to flow predominantly parallel to the ab plane. An appreciable fraction of the currents flowing parallel to the normal to the plane (c crystallographic axis) only appears when **H** is within the angle $\sim 1/\kappa$ from the plane. To measure these currents it is necessary but not sufficient to operate within this angular range.

A solution of the anisotropic Bean model with the magnetic field parallel to the plane of the plate was obtained in Refs. 21 and 22. In this case, the distribution of the magnetic induction is determined by the linear profiles. The gradient of the magnetic induction **B** is determined by the transverse critical current. We shall analyze the case when **H** is applied in the x direction parallel to the plane of the plate (see Fig. 1). Two typical scenarios for flux penetration exist. In short samples ($L/t < \Omega$) the flux penetrates more rapidly in the z direction than in the y direction and the magnetic moment m_x is predominantly determined by the short-circuit currents parallel to the c axis. Only these samples are effective for measurement of the current anisotropy Ω. In long samples ($L/t > \Omega$) the flux penetrates more rapidly in the y direction. In this case, m_x is determined by the component J_{ab} parallel to the plane. Thus, in long samples the influence of the J_c component is not important.

Since high-temperature superconducting single crystals

are nearly square with $w \sim L$, long and short samples are defined according to whether κ is greater or smaller than Ω, respectively. The values of Ω in $YBa_2Cu_3O_7$ samples with a high oxygen content lie between 5 and 10 (see Refs. 15 and 16) and are lower than the κ values for all the single crystals investigated here (see Table I). Thus, all these crystals are effectively long and shielding currents flowing in the ab plane predominate. Thus, there is no need to include current anisotropy in our analysis. The agreement between the data and the geometric model, when the **H** direction is close to the ab plane (as can be seen from Fig. 7), confirm this conclusion.

However, samples with κ values less than ~ 10 are required to obtain information on the currents parallel to the c axis in $YBa_2Cu_3O_7$ single crystals. Such samples are fairly rare. In addition, this sample should be oriented with a precision better than one degree (Eqs. (4) and (5)) which is not usually achieved in SQUID measurements. Among the more anisotropic high-temperature superconductors, we can easily find samples with the required aspect ratio. However, the constraints on the orientation precision increase substantially and rapidly exceed the characteristic growth misorientation ($\sim 0.1°$) of high-temperature superconducting single crystals. For these reasons magnetic measurements can only yield the critical current anisotropy in these crystals for a fairly narrow range of Ω values.

5. CONCLUSIONS

A solution of the Bean model has been obtained for an infinite thin plate in an inclined magnetic field. An analysis was made of the influence of anisotropy, the finite slope of the current–voltage characteristics, and the closure currents in a finite samples and it was established that these factor weakly influence the results. Over almost the entire range of **H** directions, the magnetic moment **m** induced by the shielding currents is bound to the normal of the plate and the magnitude of **m** varies only negligibly. This localization of

the magnetic moment is of purely geometric origin and is in no way related to the anisotropy of the superconducting parameters of the sample. However, when the direction of **H** approaches the plane of the plate, **m** begins to vary rapidly in direction and magnitude.

The validity of the geometric model was investigated using vector magnetometry for various high-temperature superconducting samples and conventional superconductors. The results reveal very good agreement with the model calculations of the direction of **m**. However, |**m**| shows appreciable deviations from the calculations evidently because of the critical current anisotropy in the samples and the assumptions made in the model.

---

*)E-mail: zhukov@lt.phys.msu.su

†)E-mail: mikh@lt.phys.msu.su

1)This peak, which is frequently called a fishtail peak because of its shape, is characterized by a monotonic increase in the field corresponding to the peak position with a decrease in temperature and by universal scaling behavior of the field dependences of the magnetization and the relaxation rate.[19]

---

[1] W. Andrä, H. Danauand, and R. Hergt, Phys. Status Solidi **111**, 583 (1989).

[2] I. Felner, U. Yaron, Y. Yeshurun *et al.*, Phys. Rev. B **40**, 5239 (1989).

[3] H. Teshima, A. Oishi, H. Izumi *et al.*, Appl. Phys. Lett. **58**, 2833 (1991).

[4] F. Hellman, E. M. Gyorgy, and R. C. Dynes, Phys. Rev. Lett. **68**, 867 (1992).

[5] Yu. V. Bugoslavsky, A. A. Minakov, and S. I. Vasyurin, J. Appl. Phys. **79**, 1996 (1996).

[6] C. P. Bean, Phys. Rev. Lett. **8**, 250 (1962).

[7] E. H. Brandt, Phys. Rev. B **49**, 9024 (1994).

[8] E. H. Brandt, Phys. Rev. B **52**, 15 442 (1995); **54**, 4246 (1996).

[9] A. A. Zhukov, Solid State Commun. **82**, 983 (1992); A. A. Zhukov, H. Küpfer, V. A. Rybachuk *et al.*, Physica C **219**, 99 (1994); A. A. Zhukov, A. V. Volkozub, and P. A. J. de Groot, Phys. Rev. B **52**, 13 013 (1995).

[10] G. A. Korn and T. M. Korn *Mathematical Handbook for Scientists and Engineers* (McGraw-Hill, New York, 1961) [Russ. transl., Nauka, Moscow, 1971].

[11] Th. Schuster, H. Kuhn, E. H. Brandt *et al.*, Phys. Rev. B **52**, 10 375 (1995); Th. Schuster, H. Kuhn, and E. H. Brandt Phys. Rev. B **54**, 3514 (1996).

[12] Th. Wolf, W. Goldacker, B. Obst *et al.*, J. Cryst. Growth **96**, 1010 (1989).

[13] A. A. Zhukov, H. Küpfer, H. Claus *et al.*, Phys. Rev. B **52**, R9871 (1995).

[14] A. I. Golovashkin, E. V. Ekimov, S. I. Krasnosvobodtsev *et al.*, Physica C **162−164**, 715 (1989).

[15] V. V. Moshchalkov, A. A. Zhukov, V. A. Rybachuk *et al.*, Physica C **185−189**, 2191 (1991); H. Küpfer, A. A. Zhukov, A. Will *et al.*, Phys. Rev. B **54**, 644 (1996).

[16] J. F. Thomas, G. K. Perkins, D. E. Lacey *et al.*, Czech. J. Phys. **46**, 1775 (1996).

[17] G. Blatter, M. V. Feigel'man, V. B. Geshbenkein *et al.*, Rev. Mod. Phys. **66**, 1125 (1994).

[18] L. Klein, E. R. Yacoby, Y. Yeshurun *et al.*, Phys. Rev. B **49**, 4403 (1994).

[19] A. A. Zhukov, H. Küpfer, G. K. Perkins *et al.*, Phys. Rev. B **51**, 12 704 (1995); A. A. Zhukov, H. Küpfer, H. Claus *et al.*, Phys. Rev. B **52**, R9871 (1995).

[20] A. Erb, J.-Y. Genoud, F. Marti *et al.*, J. Low Temp. Phys. **105**, 1023 (1996).

[21] V. V. Moshchalkov, A. A. Zhukov, V. D. Kuznetsov *et al.*, Sverkhprovodimost' (KIAE) **2**(12), 84 (1989) [Superconductivity **2**(12), 98 (1989)].

[22] E. M. Gyorgy, R. B. van Dover, K. A. Jackson *et al.*, Appl. Phys. Lett. **55**, 283 (1989).

Translated by R. M. Durham

Crystal-field effects in the intermediate-valence compound YbCu$_2$Si$_2$

A. Yu. Muzychka*)

Joint Institute for Nuclear Research, 141980 Dubna, Moscow Region, Russia
(Submitted 24 July 1997)
Zh. Éksp. Teor. Fiz. **114**, 291–314 (July 1998)

Inelastic thermal-neutron scattering is used to study the intermediate-valence system YbCu$_2$Si$_2$. The magnetic scattering in two nonoverlapping ranges of transfer energies, $2 < \varepsilon < 5$ meV and $5 < \varepsilon < 100$ meV, is analyzed under the assumption that the regions influence each other only weakly. As a result, two sets of phenomenological crystal-field parameters are established, and their difference constitutes the experimental error in determining these parameters. A comparison of the fourth-order crystal field with other compounds belonging to the RCu$_2$Si$_2$ series (R stands for a rare-earth element) suggests that in YbCu$_2$Si$_2$ hybridization occurs between f electrons and copper electrons, in contrast to the heavy-fermion system CeCu$_2$Si$_2$, for which it was established earlier that hybridization occurs between f electrons and Si p electrons. © *1998 American Institute of Physics.* [S1063-7761(98)02007-1]

1. INTRODUCTION AND BRIEF REVIEW OF THE LITERATURE

YbCu$_2$Si$_2$ is an intermediate-valence compound. This is indicated by data on the x-ray absorption edge and on the electronic specific heat. The Sommerfeld constant γ amounts to 135 mJ/(K^2·mol) (Ref. 1), which is much larger than the values characteristic of normal metals (~ 1) and smaller than the values of γ in heavy-fermion systems. For instance, in CeCu$_2$Si$_2$, which is a heavy-fermion system, $\gamma = 1000$ mJ/(K^2·mol). The valence of the Yb ion in YbCu$_2$Si$_2$, determined from the x-ray absorption edge, is 2.9 (Refs. 2 and 3).

The problem of whether the crystal-field model can be applied to intermediate-valence systems has yet to be fully resolved, which comes as no surprise since the many-particle nature of the states in intermediate-valence systems is highly pronounced, and all attempts to describe such systems by a single-particle model should be fraught with many difficulties. And yet there is proof that such a description is possible. Goremychkin and Osborn[4] were the first to find experimental proof that the crystal-field model can be used to describe heavy-fermion systems in the case of CeCu$_2$Si$_2$. If the crystal fields for intermediate-valence systems could also be found, we would have proof of the existence of a specific form in which the many-particle nature of all rare-earth compounds with a strong hybridization interaction manifests itself, and we could perform a concrete analysis within the scope of this specific form.

For us it is especially important to determine the crystal field in the title compound because its determination would continue the study of crystal fields in the RCu$_2$Si$_2$ series (R stands for a rear-earth element)[5–8] performed for the purpose of establishing the crystal-field anomalies associated with hybridization. The main conclusion drawn in all these papers is that regular variation of the crystal field can be observed in the series of the isostructural compounds RCu$_2$Si$_2$: from an anomalous crystal field in the Ce system through an intermediate one in the Pr and Nd compounds (at high temperatures)

to a normal crystal field in the low-temperature Nd compound and in the Ho and Er compounds. We believe that we are on firm ground when we say that the normal, i.e., free from the effect of hybridization, fourth-order multipole of the crystal field is formed primarily by copper ligands, while the Si sphere normally contributes almost nothing to this order of the crystal field. A conception of the shape of the normal crystal field in this series provides a reference point in our search for the crystal field of the intermediate-valence system YbCu$_2$Si$_2$.

The Yb ion in the title compound is in an environment with tetragonal point symmetry, i.e., the symmetry of a site is fairly low. Five independent parameters are needed to determine the crystal field, and they can be reliably found only by directly measuring the crystal field in an inelastic thermal-neutron scattering experiment. Here the problem of determining the crystal field in an intermediate-valence system is sure to be complex because of the extremely broad inelastic peaks and the difficulties associated with separating the magnetic scattering, which carries information about the crystal field, from the phonon scattering. So far nobody has been able to measure a purely magnetic response in neutron experiments involving YbCu$_2$Si$_2$ in the entire range of energy transfers, and in interpreting a purely magnetic response separated in some way it is questionable whether separating inelastic scattering from quasielastic or inelastic Kondo scattering is possible. Hence in our attempts to determine the crystal field we must use all the information available, including the results of macroscopic studies of this compound.

According to the literature data, the inelastic scattering of thermal neutrons in YbCu$_2$Si$_2$ has been measured three times. Twice the measurements were performed by the same group of researchers.[9,10] One of the main conclusions drawn by these researchers was that the magnetic response cannot be described by a single quasielastic Lorentzian. A noncontradictory description of their data was possible only by assuming that there are several inelastic lines in the magnetic

response. This prompted them to suggest that there are inelastic transitions associated with the crystal field in the inelastic scattering of thermal neutrons. The energetic position of one of these transitions (≈ 30 meV) on the energy-transfer scale was found, and the result was highly reliable. Unfortunately, the sets of the crystal-field parameters found in Ref. 10 do not correlate very strongly with the results of measuring the static magnetic susceptibility in a $YbCu_2Si_2$ single crystal.[11] The calculated value of the anisotropy of χ obtained on the basis of the first variant of the crystal field[10] is much smaller than the experimental value. Another set yields an anisotropy of χ inverted in relation to the measured anisotropy.

In the last of these three experiments, Currat and Murani[12] measured the inelastic scattering of thermal neutrons in a $YbCu_2Si_2$ single crystal. The range of energy transfers from 8 to 50 meV was investigated. They found that in all scattering directions the magnetic response at $T = 4.3$ K had at least two broad inelastic components. The maximum of the first was approximately at 12–16 meV and the maximum of the second, at 30 meV. These researchers did not seek the crystal-field parameters.

On the basis of the data from the measurements of $\chi(T)$ performed by Shimizu et al.,[11] Zevin et al.[13] assumed that the leading contribution to the crystal field of $YbCu_2Si_2$ is provided by the quadrupole term. Indeed, for $B_2^0 = -0.56$ meV and a purely axial crystal-field symmetry, the calculated susceptibility is close to the measured value in the entire temperature range. The agreement can be improved by introducing a negative paramagnetic Curie temperature Θ, a characteristic feature of intermediate-valence systems. The agreement is best with $\Theta \simeq -30$ K. In addition, Zevin et al.[13] used the above crystal field to calculate the temperature-dependent component $Q(T)$ of the quadrupole moment by a method based on the noncrossing-diagram approximation (see, e.g., Refs. 14–17). The calculated dependence was fitted to the experimental data gathered from measurements of the quadrupole moment of the Yb ion in $YbCu_2Si_2$ by Mössbauer spectroscopy.[18] The fitting process involved varying the Kondo temperature T_K and the effective width Γ of the resonant layer of conduction electrons on the Fermi surface. This yielded $T_K = 200$ K and $\Gamma \simeq 50$ meV. The parameter T_K was found to be the most sensitive in the fitting process.

Although being outwardly convincing, the results of the calculations of $Q(T)$ and $\chi(T)$ done by Zevin et al.[13] pose several problems. First, why is it that with such a high value of T_K the temperature dependence of the static magnetic susceptibility is determined by the crystal field down to 50–30 K, while with $T \sim T_K$ or less, χ must become the temperature-independent Pauli susceptibility? Second, the negative paramagnetic Curie temperature Θ should coincide in order of magnitude with T_K, while it actually differs by a factor of approximately seven. Calculations done by the present author with $T_K = 30$ K by the method proposed in Ref. 13 show that there is not a single value of Γ at which the calculated $Q(T)$ curves could be made to resemble the experimental curves.

Moreover, the crystal field determined in Ref. 13 contra-

dicts the results of neutron experiments with a single crystal.[12] In the absence of spin and charge fluctuations, the crystal field at low temperatures yields only one transition between the crystal-field levels with a nonzero matrix element. The transition energy amounted to $\simeq 10$ meV. Zwicknagl et al.[19] calculated the magnetic response with allowance for the crystal field determined in Ref. 13 and strong spin and charge fluctuations. Allowance for the anisotropy of the fluctuations was found to lead to different energetic positions for the inelastic peak in the dynamic magnetic susceptibility in the directions along the c axis and perpendicular to that direction. Thus, for all other directions the magnetic response must have two peaks, while along the c axis and the direction perpendicular to that axis the magnetic response has only one peak. However, according to Currat and Murani,[12] at least two broad components are present in the magnetic response in all directions.

The conclusion that sums up this brief review of the literature is that the problem of finding the crystal field in the intermediate-valence system $YbCu_2Si_2$ is yet to be solved.

2. EXPERIMENT

A polycrystalline sample of $YbCu_2Si_2$ was prepared by arc melting in a water-cooled copper hearth with no weight loss during the melting process. After vacuum annealing at 700 °C, x-ray crystallographic analysis and neutron diffraction experiments showed that there are no other phases in the sample.

The inelastic thermal-neutron scattering experiment was conducted on a KDSOG-M time-of-flight spectrometer with inverse geometry attached to the IBR-2 pulsed reactor (Joint Institute for Nuclear Research, Dubna, Russia). The inelastic scattering of thermal neutrons was measured at three temperatures: 10 K, 80 K, and 300 K. The range of energy transfers investigated extended from 2 to 200 meV.

3. ANALYSIS OF RESULTS

3.1. Qualitative analysis of spectra

We found that the greatest difficulty in analyzing the results of inelastic thermal-neutron scattering measurements was separating the phonon and magnetic scattering components. Nevertheless, even if we resort only to qualitative analysis, we are able to draw certain conclusions about the nature of magnetic scattering: it is concentrated in two distinct ranges of the energy transfer ε. It is easy to assume that the low-energy feature ($\varepsilon \simeq 3 - 5$ meV), which is clearly visible in the spectra and unmistakably lies in the Debye range of phonon scattering, is of a purely magnetic nature. This assumption has been corroborated by the measurements of E. A. Goremychkin performed on the HET time-of-flight spectrometer with direct geometry at the ISIS spallation neutron source (RAL, UK). In contrast to the KDSOG spectrometer, the HET instrument makes it possible to substantially vary the momentum transfer at a fixed energy transfer and, due to this, to lower the phonon-to-magnetic component ratio substantially (the so-called angular suppression of phonon scattering). According to the HET data, this feature can be ob-
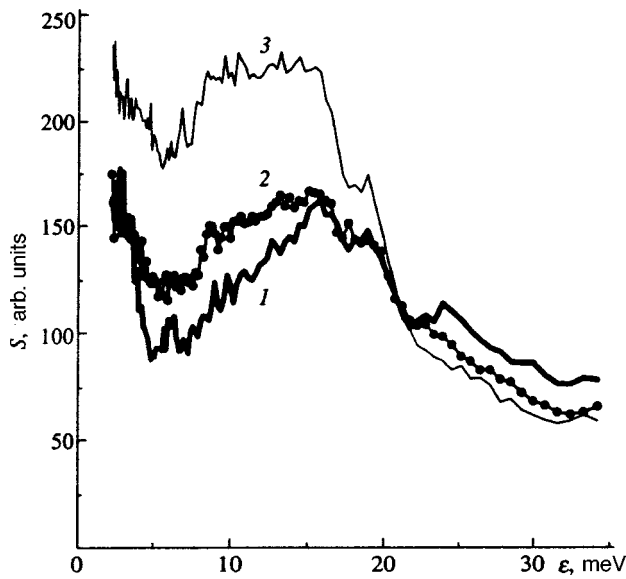
FIG. 1. Experimental scattering law obtained with the KDSOG-M spectrometer for $YbCu_2Si_2$ at $T = 10$ K (curve *1*), $T = 80$ K (curve *2*), and $T = 300$ K (curve *3*).

served only at small scattering angles.[20] The energy width of this feature almost coincides with its position and suggests that the possible high-energy ''tail'' is not a significant factor in shaping the spectrum in the range of energy transfers above 10 meV. Here the temperature dependence of the scattering in the range $20 < \varepsilon < 100$ meV is pronounced in the KDSOG-M spectra (Fig. 1), is of the magnetic nature (an increase in scattering intensity with decreasing temperature), and suggests that in this range of ε the magnetic component is predominant. The range $7 < \varepsilon < 20$ meV is the most difficult for analysis, since here phonon scattering is extremely strong (its intensity increases with temperature) and is mixed with strong magnetic scattering, whose tail extends up to $\varepsilon = 100$ meV.

3.2. "Primary" separation of the phonon component

Since the KDSOG-M spectrometer has an inverse geometry, it scans a broad range of energy transfers. However, with it one cannot substantially vary the momentum transfer for a fixed energy transfer, so that the only way to separate the phonon and magnetic components is to measure the magnetic compound and a nonmagnetic analog, in which the magnetically active ion is replaced, say, by La. After making an appropriate correction, the spectrum of the latter can be considered the phonon component of the former.

Figure 2 depicts the time-of-flight spectrum of $LaCu_2Si_2$ (Fig. 2a, filled circles) together with the spectra of the antiferromagnetic members of the series measured with the
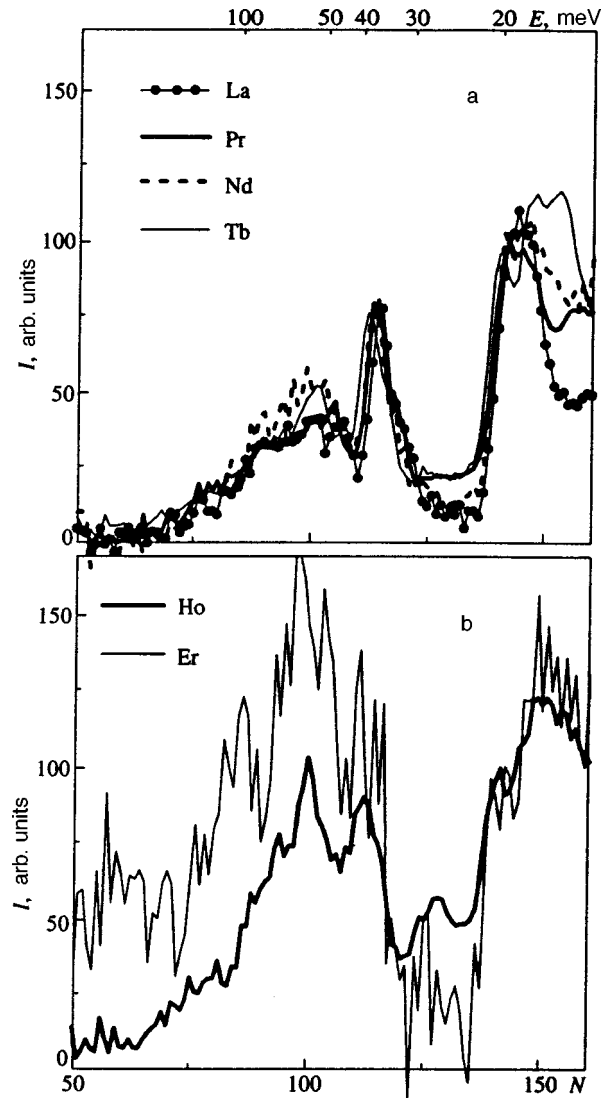
FIG. 2. Time-of-flight spectra of inelastically scattered thermal neutrons in antiferromagnetic compounds of the RCu_2Si_2 series recorded with the KDSOG-M spectrometer. The range of energy transfers is the one in which there is no magnetic scattering.

KDSOG-M spectrometer at $T = 80$ K. The value of the total splitting of the ground-state multiplet in the crystal field for all antiferromagnets coincided to within 0.5 meV and amounted to ≈ 11 meV (Ref. 7). The widths of the inelastic lines of the magnetic response do not exceed 1 meV. The range of energy transfers in Fig. 2 is chosen so as to show the sections of the spectra that are free of magnetic scattering. A correction related to the self-shielding of the sample was made for all spectra, i.e., they were all divided by a function $F(\varepsilon)$, which specifies the weakening of the scattered beam due to the finite thickness of the sample. In the spectrometer's geometry, this function has the form

$$F(\varepsilon) = \frac{\exp\{-\mu(E_0)d/\sin(45° + \theta)\} - \exp\{-\mu(E_0 + \varepsilon)d/\sin 45°\}}{\sin 45°[\mu(E_0 + \varepsilon)d/\sin 45° - \mu(E_0)d/\sin(45° + \theta)]}, \tag{1}$$

where E_0 is the final neutron energy, ε is the energy transfer, $\mu(E)$ is the energy-dependent linear beam-attenuation factor, which allows for all neutron scattering and absorption mechanisms, θ is the scattering angle, and d is the sample thickness.

In addition, all the spectra were normalized so that the intensities of the phonon feature at $\varepsilon = 20$ meV would coincide.

A comparison of Figs. 2a and 2b shows that for $\varepsilon > 25$ meV the phonon spectrum of $HoCu_2Si_2$ (the heavy curve in Fig. 2b) differs strongly from all the spectra depicted in Fig. 2a. The reason for this is unknown to us, since an analysis of the tabulated data shows that there is not a single scattering parameter for Ho that behaves abnormally. It is more important at this point to know to what extent the spectrum of $LaCu_2Si_2$ can be used as the phonon component of the spectrum of $YbCu_2Si_2$. In other words, we would like to know to what extent the features in the phonon spectrum of Ho that are not present in the spectra depicted in Fig. 2a are retained in the phonon spectrum of Yb.

The light curve in Fig. 2b depicts the spectrum of $ErCu_2Si_2$. The extremely poor statistics of the spectrum are due to the large absorption cross section for Er. The fact that the high-energy part of the spectrum is located above the low-energy part can be attributed to the insufficient self-shielding correction introduced by the function (1). Nevertheless, it is obvious that the Er spectrum is much closer to the Ho spectrum than it is to the spectra of lighter rarer-earth ions. This was a sufficient argument in favor of not taking the La spectrum as a base for imitating the phonon spectrum of $YbCu_2Si_2$, but using, for lack of anything better, the spectrum of a much closer neighbor in the rare-earth series, which was recorded with acceptable statistics, i.e., the spectrum of $HoCu_2Si_2$ (more precisely, the part free of magnetic scattering). The part of the phonon spectrum below the magnetic scattering was imitated by the Debye law with a Debye temperature equal to 20 meV. This part was matched to the remainder of the spectrum near the 160th time channel ($\varepsilon = 12$ meV).

The model phonon spectrum constructed in the way we have just described for $T = 80$ K was recalculated for other temperatures, $T = 10$ K and $T = 300$ K. The normalization constant of the model phonon spectrum was chosen so that the intensities of the features near the 100th time channel ($\varepsilon = 55$ meV) in the model spectrum and in the spectrum of $YbCu_2Si_2$ would coincide at $T = 300$ K (see Fig. 3a: the heavy curve is the model spectrum and the light curve is the spectrum of $YbCu_2Si_2$, $T = 300$ K). This choice of normalization can easily be explained: the higher is the temperature, the smaller is the fraction of magnetic scattering at high values of ε and the closer is the phonon spectrum (from below) to the total spectrum in this region. In this case we obtain the smallest possible value of the normalization constant for the phonon component.

Figure 3b depicts the time-of-flight spectra of the magnetic response of $YbCu_2Si_2$ at three temperatures obtained as a result of subtracting the model phonon spectra from the
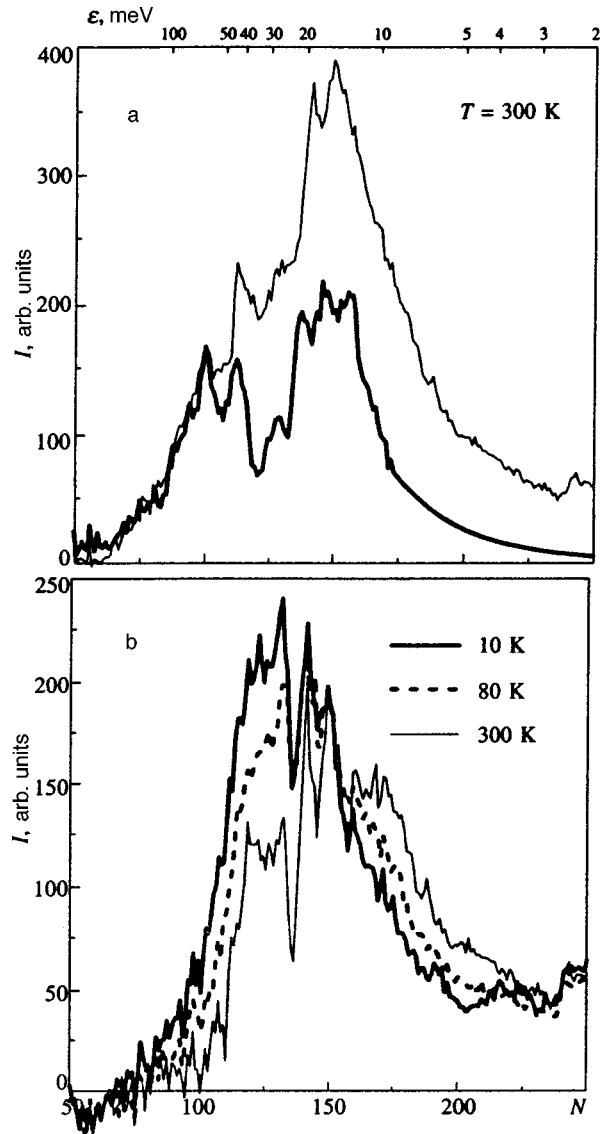

FIG. 3. (a) Time-of-flight thermal-neutron inelastic scattering spectrum of $YbCu_2Si_2$ obtained with the KDSOG-M spectrometer (light curve) and a model spectrum imitating the phonon component of the scattering and normalized as indicated in the text (heavy curve). (b) Results of subtracting model phonon spectra from inelastic thermal-neutron scattering spectra recorded at different temperatures.

experimental spectra of $YbCu_2Si_2$. The marked dips on these spectra appear because of the mismatch of the phonon features in the Ho and Yb spectra. If we mentally "smooth out" these features and pay attention only to the general outline of the curve, the temperature dynamics of the magnetic response in the time channels below the 200th ($\varepsilon > 5$ meV) appear to be as follows: at $T = 10$ K we have a wide unstructured line covering the range of energy transfers from 5 to 100 meV, at $T = 300$ K we have a structure consisting of three narrower components, and at $T = 80$ K we have an intermediate picture.

In what follows we call the magnetic-response spectra obtained in this way the "preliminary" magnetic-response spectra.

3.3. Grid analysis of "preliminary" magnetic-response spectra

In decomposing the "preliminary" magnetic response into spectral components we assume that the following assumptions hold:

1) the crystal-field levels are temperature-independent;

2) at a fixed temperature all inelastic lines in the spectrum have the same width;

3) the low-energy feature of the spectrum at $T = 10$ K, which is of a magnetic nature, in the compound under consideration is the analog of the quasielastic component in the case of a compound with a definite value of the total magnetic moment (an antiferromagnet or a heavy-fermion system);

4) the width of this feature is approximately equal to its position, i.e., ≈ 3 meV; then the "quasielastic" component provides an insignificant contribution to the magnetic scattering at $5 < \varepsilon < 100$ meV, which is mostly due to inelastic transitions in the crystal field;

5) the spectral function describing the shape of the inelastic line is Lorentzian.

The most "intriguing" part of the magnetic response is the quasielastic component: its width is much smaller than the width of inelastic lines, and its shape cannot be described by a Lorentzian (we shall elaborate on this below). Hence in analyzing the crystal field we focused on the range of energy transfers in which the quasielastic component is assumed to play a small role, $5 < \varepsilon < 100$ meV, i.e., in the 200th to the 75th time channels.

What complicates the problem of decomposing the "preliminary" magnetic-response spectra into spectral components is, of course, the fact that the spectra are highly undifferentiated, and the resolving power of the spectrometer has nothing to do with it. The spectra are resolved so poorly because of the large natural width of the spectral lines. On the other hand, matters are simplified considerably by the fact that at $T = 10$ K the inelastic scattering spectrum should contain no more than three lines, which correspond to transitions from the ground-state doublet to three excited doublets (it is natural to take into account the splitting of only the ground-state multiplet $^2F_{7/2}$, since the next multiplet, $^2F_{5/2}$, is separated from the ground-state multiplet by a huge energy gap of 1251.5 meV), while at $T = 300$ K the maximum number of lines is six.

Each spectral component corresponds to the transition Δ_{ij} from the ith crystal-field level to the jth level, i.e., to one of the terms in the dipole approximation of the law of scattering of unpolarized neutrons by a paramagnetic ion in a crystal field:[10]

$$S(\varepsilon, T) \propto \frac{\varepsilon/kT}{1 - \exp(-\varepsilon/kT)} \sum_{ij} \rho_i |\langle j|\hat{J}_\perp|i\rangle|^2$$

$$\times \frac{1 - \exp(-\Delta_{ij}/kT)}{\Delta_{ij}/kT} P(\varepsilon - \Delta_{ij}). \qquad (2)$$

The energetic position of the component Δ_{ij} corresponds to a peak in the spectral function P, and the transition intensity is determined by the thermal population ρ_i of the ith level and

the square of the matrix element $\langle j|\hat{J}_\perp|i\rangle$ of the operator of the component of the total angular momentum of the paramagnetic ion perpendicular to the scattering vector. In the case of a polycrystalline scatterer, this matrix element can be expressed in terms of the matrix elements of the standard angular-momentum operators:

$$|\langle j|\hat{J}_\perp|i\rangle|^2 = \tfrac{1}{3}|\langle j|\hat{J}_-|i\rangle|^2 + \tfrac{1}{3}|\langle j|\hat{J}_+|i\rangle|^2 + \tfrac{2}{3}|\langle j|\hat{J}_z|i\rangle|^2.$$

Both the positions and the intensities of the transitions are uniquely determined by the crystal-field operator, which for rare-earth ions in the case of tetragonal point symmetry has five terms:

$$\hat{H}_{CF} = B_2^0 \hat{O}_2^0 + B_4^0 \hat{O}_4^0 + B_6^0 \hat{O}_6^0 + B_4^4 \hat{O}_4^4 + B_6^4 \hat{O}_6^4, \qquad (3)$$

where $\hat{O}_l^m$ are the equivalent Stevens operators,[21] and $\{B_l^m\}$ is the set of crystal-field parameters sought.

The search for the crystal-field parameters is carried out by the standard method: the results of decomposition of the spectra yield a set of discrete data (the positions and intensities of the peaks), which can serve as a selection criterion for examining different crystal-field variants. The uncertainties in the positions and intensities of the peaks act as a gateway, through which the set of crystal-field parameters being tested either passes or does not. It is thus clear that in the present case the gateway is fairly "wide."

In sorting the various crystal-field variants it is convenient to use Walter's parametrization.[22] The convenience of this parametrization lies in the fact that the range of variation of each Walter parameter x_i is bounded: $|x_i| \leq 1$. Moreover, only four of the five Walter parameters are varied. The fifth parameter is the scale factor W, which is defined by relating the energetic position ε of a specific spectral component to a definite transition between crystal-field levels through the linear equation $\varepsilon = W\Delta_{ij}$, where Δ_{ij} is the transition from the ith level to the jth level of the crystal-field Hamiltonian taken for $W = 1$. In our case the energetic position of one of spectral components was $\varepsilon = 32$ meV. The fact that other neutron studies (see the Introduction) also detected a magnetic feature with a close energetic position suggests that this component is reliable and can be used to determine the scale factor W.

In sorting the crystal-field variants the entire parameter space $\{x_i\}$ was scanned with a mesh width for x_i equal to 0.02. The crystal-field transitions were calculated at each mesh point, and the results were compared with the parameters of the spectral components of the "preliminary" magnetic-response spectra. The uncertainties of the positions and intensities of these components were sure to be much larger than the errors in the calculated quantities due to the finite size of the grid cells. As a result of scanning, roughly two dozen unrelated regions of the space $\{x_i\}$ were selected, each corresponding to a definite region of the parameter space $\{B_l^m\}$.

3.4. Final isolation of the inelastic magnetic response

It goes without saying that the "preliminary" magnetic spectra can serve only as a very rough approximation of the YbCu$_2$Si$_2$ magnetic response. Moreover, the large number of
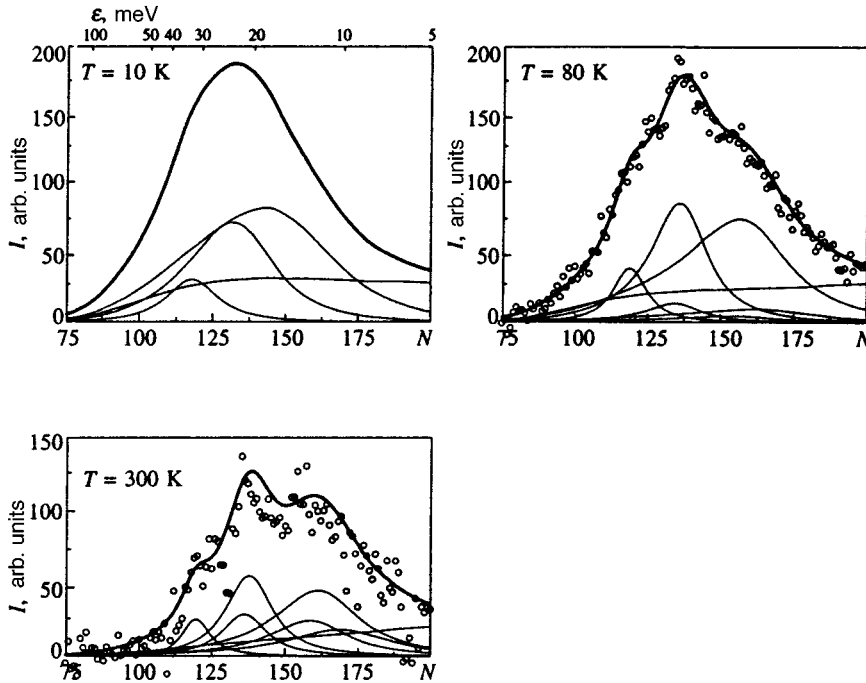
FIG. 4. Time-of-flight inelastic magnetic thermal-neutron scattering spectra recorded at different temperatures and corresponding to the first variant of the crystal field: points—magnetic spectrum isolated from the total experimental inelastic thermal-neutron scattering spectrum by the method described in the text; heavy curves—result of calculations based on the first set of crystal-field parameters; light curves—individual spectral components of the calculated spectrum.

regions in the crystal-field parameter space selected as result of grid analysis requires an effective selection criterion for selecting among these ''points.'' The criterion that we used consisted of the results of the least-squares fitting of the calculated inelastic ($\varepsilon > 5$ meV) magnetic-response spectrum containing both the magnetic and phonon components to the experimental spectra at $T = 80$ K and 300 K. In this process the spectra for both temperatures were fitted simultaneously. The difference between the total experimental spectrum and the calculated magnetic spectrum at $T = 10$ K, recalculated to the respective temperature, was taken as the phonon component of the calculated spectrum. The parameters $\{B_l^m\}$, the factor by which the magnetic response as a whole was multiplied, and the linewidths common to all inelastic transitions in the crystal field at a given temperature were varied in the fitting procedure. The high-energy tail of the low-energy feature was described as the tail of a normal quasielastic components, but the width of this component was varied independently of the width of the inelastic transitions.

The energy-transfer function

$$f(\varepsilon) = \frac{1 - \exp(-\varepsilon/kT_0)}{1 - \exp(-\varepsilon/kT_x)},$$

by which we must multiply the phonon spectrum to recalculate from $T_0 = 10$ K to T_x, increases with T_x. This causes enhancement of the ''noise'' present in the difference spectrum at 10 K and an increase in the spread of points in the calculated spectra as the temperature is increased, which may lead to difficulties in minimizing χ^2. Nevertheless, we hoped that if the starting point in the fitting procedure is close to the true set of crystal-field parameters, the procedure should converge to that set.

The least-squares fits obtained by starting the procedure in different selected regions were rated by two criteria: first, by the value of χ^2, and, second, by the stability of the pro-

cedure with respect to the region, i.e., if the region is not random but is near a true point, the procedure must converge in it.

Only one region was found to meet the stability criterion, and it also exhibited the lowest level of χ^2 as a result of fitting. A characteristic feature of this region, which sets it apart from all the other regions, is the large absolute value of the crystal-field parameter B_4^0. Figure 4 depicts time-of-flight spectra obtained as a result of least-squares fitting in this region. The heavy curves represent the total calculated magnetic response, the light curves represent the individual spectral components, and the points represent the result of subtracting the phonon component obtained by the method described earlier from the total experimental spectrum. The numerical results of the fitting procedure are listed in Table I. The plus-or-minus sign in front of the crystal-field parameters B_4^4 and B_6^4 reflects the fact that a neutron-scattering

TABLE I. Parameters of the calculated inelastic magnetic response used in describing the low-energy features of the spectrum at $T = 10$ K by a quasielastic Lorentzian (the first variant of the set of crystal-field parameters).

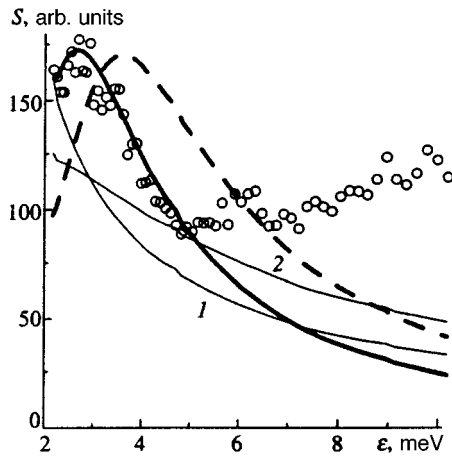| Variable parameters	Final fit
B_2^0	-0.21 meV
B_4^0	-0.21×10^{-1} meV
B_6^0	-0.10×10^{-4} meV
B_4^4	$\pm 0.46 \times 10^{-1}$ meV
B_6^4	$\pm 0.17 \times 10^{-2}$ meV
Width of quasielastic line at $T = 10$ K	0.5 meV
Width of inelastic line at $T = 10$ K	9.54 meV
Width of quasielastic line at $T = 80$ K	2.8 meV
Width of inelastic line at $T = 80$ K	5.65 meV
Width of quasielastic line at $T = 300$ K	3.3 meV
Width of inelastic line at $T = 300$ K	4.50 meV

FIG. 5. Low-energy part of the scattering law for YbCu$_2$Si$_2$ at $T = 10$ K: points—experimental spectrum obtained with the KDSOG-M spectrometer; *1*—quasielastic 0.5 meV wide Lorentzian corresponding to the quasielastic spectral component as shown in Fig. 4; *2*—quasielastic 2.25 meV wide Lorentzian with an intensity equal to one-third of the intensity corresponding to curve *1*; heavy solid line—result of calculations using the Kondo formula in Ref. 27: $T_K = 26.5$ K, the population of the f level is 0.85, the degeneracy of the ground state of the rare-earth ion is eightfold; heavy dashed line—same at $T_K = 35$ K.

experiment is insensitive to simultaneous changes in the signs of both parameters.

Below we assume that the set $\{B_l^m\}$ we have obtained is close to the true set.

3.5. Low-energy feature in the magnetic-response spectrum at $T = 10$ K

We recall that in decomposing the magnetic response into spectral components we assumed that the low-energy feature in the spectrum is an analog of the quasielastic component in compounds with a definite value of the magnetic moment and that the bulk of the intensity of ''quasielastic'' scattering in YbCu$_2$Si$_2$ should lie in the range $0 < \varepsilon < 10$ meV. The last assumption was based on a visual estimate of the width of the feature. Table I shows that in contrast to the spectra at 300 K and 80 K, the width of the quasielastic component of the spectrum recorded at $T = 10$ K is approximately 20 times smaller than the width of the inelastic lines. This fact merits an explanation.

The calculated intensity of the elastic component at $T = 10$ K in the response described by the set of parameters established earlier is related to the intensities of the elastic components as 1:0.75:0.53:0.19. The only possibility of preserving the very high intensity of the elastic line in the spectrum at $T = 10$ K (Fig. 4) in the range below 10 meV was to select a narrow width for the Lorentzian describing this line. As the width increases, the intensity ''spreads'' because of the effect of the detailed-balance factor $(\varepsilon/kT)/[1 - \exp(-\varepsilon/kT)]$ in the scattering law (2). Thus, the unjustifiably small width of the quasielastic component is the result of the attempt to describe this component by a Lorentzian and to assign a high intensity to it.

The points in Fig. 5 represent the low-energy part of the total scattering law obtained with the KDSOG-M spectrom-

eter for YbCu$_2$Si$_2$ at $T = 10$ K. The light curve *1* represents the Lorentzian of the quasielastic component of spectrum depicted in Fig. 4. Even if we were to assume that there exists a phonon background that shifts curve *1* upward, there is no way in which we can describe the characteristic peak at $\varepsilon \approx 3$ meV. As the Lorentzian width increases, the extent to which the Lorentzian does not coincide with the spectral feature grows due to both shape and intensity, as is demonstrated by curve *2*. This curve is a 2.25 meV wide elastic Lorentzian of the same intensity as curve *1*, but has been diminished threefold.

The fact that there is no way in which a Lorentzian can describe the low-energy part of the spectrum of a mixed-valence system should not come as a surprise, since it is known[14–17] that at low temperatures and small energy transfers the scattering by a rare-earth ion with an unstable $4f$ shell should contain an inelastic Kondo component, which is related to the transition from the ground state of a many-particle system consisting of $4f$ electrons and conduction electrons (a Kondo singlet) to excited states with a nonzero magnetic moment.[23–26]

Kuramoto and Muller–Hartman[27] proposed an analytical formula that describes the low-energy part of the dynamic magnetic susceptibility of a mixed-valence impurity at $T = 0$:

$$\frac{\operatorname{Im}\chi(\varepsilon)}{C} = \frac{\varepsilon}{(kT_K)^2} n \frac{\sin\alpha}{u^2(u^2 + 4\sin^2\alpha)}$$
$$\times \left\{ \sin\alpha \ln[(1-u^2)^2 + 4u^2\sin^2\alpha] \right.$$
$$\left. + u\left[\frac{\pi}{2} - \arctan\frac{1-u^2}{2u\sin\alpha}\right] \right\}, \qquad (4)$$

where $C = (g_J\mu_B)^2 J(J+1)/3$ is the Curie constant, T_K is the Kondo temperature, ε is the energy transfer, $u = \varepsilon/kT_K$, n_f is the population of the f shell, n is the degree of degeneracy of the rare-earth ion, and $\alpha = \pi(n_f/n)$.

Equation (4) was derived using the noncrossing-diagram approximation under the assumption that the environment of rare-earth ions is spherically symmetric, which plainly contradicts the real physical situation. Nevertheless, we expect that (4) gives a correct qualitative picture of the dependence of scattering on energy transfer for low ε and T in the case where the crystal-field splitting of the ground-state multiplet $\Delta_{CF} \geqslant kT_K$. When $n_f/n \geqslant 0.25$, Eq. (4) can be approximated by a sum of quasielastic and inelastic Lorentzians. When Eq. (4) was used, we assumed that $n_f = 0.85$ and $n = 8$ (degeneracy of the ground-state multiplet without a crystal field), knowingly avoiding this possibility. Here Eq. (4) was used to estimate the high-energy ''tail'' of the scattering related to the low-energy feature from below, since for small n_f/n Eq. (4) implies that the scattering intensity rapidly diminishes as ε increases.

Note that in addition to Eq. (4) being able to provide only a qualitative description of the spectrum, the situation with magnetic scattering at $\varepsilon < 5$ meV is indeterminate also due to the special way in which the magnetic component of the scattering is isolated: here we have a low-energy ''tail''

TABLE II. Parameters of the calculated inelastic magnetic response used in describing the low-energy feature of the spectrum at $T = 10$ K by Kondo scattering (the second variant of the set of crystal-field parameters).

Variable parameters	Final fit
B_2^0	-0.22 meV
B_4^0	-0.20×10^{-1} meV
B_6^0	-0.21×10^{-5} meV
B_4^4	$\pm 0.45 \times 10^{-1}$ meV
B_6^4	$\pm 0.48 \times 10^{-3}$ meV
Kondo temperature	3.9 meV (45.3 K)
Width of inelastic line at $T = 10$ K	9.71 meV
Width of quasielastic line at $T = 80$ K	1.5 meV
Width of inelastic line at $T = 80$ K	6.0 meV
Width of quasielastic line at $T = 300$ K	3.3 meV
Width of inelastic line at $T = 300$ K	4.50 meV

FIG. 6. Comparison of the Kondo component and inelastic-scattering intensities: light line—Kondo component, $T_K = 26.5$ K (Fig. 5); heavy dashed line—inelastic response calculated on the basis of the second variant of the crystal field at $T = 10$ K; heavy solid line—sum of the two; points—experimental scattering law obtained with the KDSOG-M spectrometer for YbCu$_2$Si$_2$ at $T = 10$ K.

of inelastic transitions in the crystal field and a phonon background. Both components of the spectrum, as is evident from the description of the processing procedure, are determined by magnetic scattering fitted in another range of energy transfers ($\varepsilon > 5$ meV). The only assumption that makes any sense in this situation is that at low temperatures the Kondo component (4) is the leading scattering component in the range of energy transfers $\varepsilon < 5$ meV.

Figure 5 depicts the result of the least-squares fitting of the calculated scattering law (the heavy solid curve) to the experimental scattering law obtained with the KDSOG-M spectrometer at $T = 10$ K (points). The fitting was carried out only in the range $2 < \varepsilon < 5$ meV. For the calculated scattering law we took only the Kondo component (4) multiplied by the factor $[1 - \exp(-\varepsilon/kT)]^{-1}$. In the fitting process we varied the Kondo temperature and the factor multiplying the calculated spectrum as a whole. As a result of the fitting, the value of kT_K was found to be equal to 2.25 meV ($T_K = 26.5$ K). We would like to believe that ignoring the other scattering components at these values of ε does not lead to serious errors in determining the Kondo temperature T_K, whose value is determined primarily by the position of the peak of the low-energy feature. The dashed curve in Fig. 5 represents the results of a calculation using Eq. (4) for $T_K = 35$ K.

The result of the procedure for fitting the inelastic part of the spectrum described in the previous section depends on the shape of the tail of the quasielastic component. The set $\{B_l^m\}$ listed in Table I was obtained under the assumption that this component is described by a Lorentzian. If we suppose that the quasielastic component is described by Eq. (4), the fitting results (see Table II) change, but they remain inside the region of the crystal-field parameter space determined by grid analysis. We believe that the difference between $\{B_l^m\}_1$ (Table I) and $\{B_l^m\}_2$ (Table II) characterizes the accuracy of determining the crystal-field parameters in the present paper. It is significant that the parameters B_4^0 and B_4^4 are almost the same in both sets.

Note that the crystal-field parameters and the Kondo scattering parameters (4) were obtained from two different least-squares fits in two nonoverlapping ranges of energy transfers. We intentionally avoided the question of matching
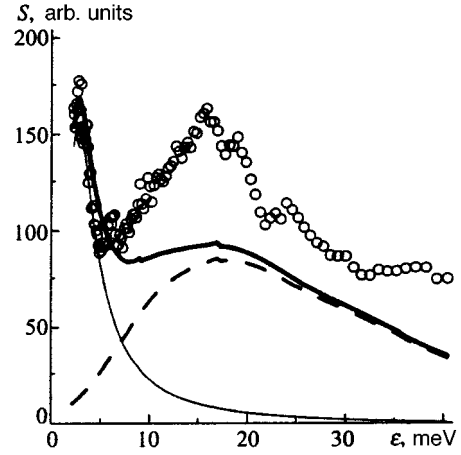
by assuming that one range of energy transfers has little effect on other ranges. As for fitting the magnetic spectrum over the entire range of energy transfers, we lack, first and most importantly, sufficient theoretical knowledge about the structure of the quasielastic component and, second, sufficient exact knowledge about the phonon contribution in the region of small energy transfers.

Nevertheless, to qualitatively compare the intensities of the quasielastic and inelastic components of the magnetic scattering, we fitted the intensity of the Kondo component shown in Fig. 5 (heavy solid line) in the presence of the inelastic magnetic response corresponding to the set $\{B_l^m\}_2$ to the experimental scattering law at $T = 10$ K in the range $2 < \varepsilon < 5$ meV. Only the common factor of the Kondo component was varied during the fitting. The results are depicted in Fig. 6: points—the total experimental scattering law obtained with the KDSOG-M spectrometer at $T = 10$ K; light line—the Kondo component of the scattering calculated by Eq. (4) at $T_K = 26.5$ K; heavy dashed line—the sum of three inelastic lines similar to those represented in Fig. 4, $T = 10$ K; and heavy solid line—the sum of the calculated Kondo component and the calculated inelastic components.

The intensity of each was found by the formula[10]

$$\chi_i(T) \propto \frac{\Delta_i}{1 - \exp(-\beta \Delta_i)} \sigma_{\text{mag}}^i(T)$$

$$\times \left[\int_{-\infty}^{\infty} \frac{\varepsilon}{1 - \exp(-\beta \varepsilon)} P(\varepsilon - \Delta_i, T) d\varepsilon \right]^{-1}, \quad (5)$$

which is valid for broad spectral lines, from the area σ_{mag} under the respective curve in Fig. 6 in the range of ε from 2 meV to 100 meV, i.e., the largest energy transfer for which traces of magnetic scattering are visible. In Eq. (5), χ is the intensity sought, and $P(\varepsilon - \Delta, T)$ is the spectral function corresponding to a feature whose energetic position is Δ.

The ratio of the intensity of the Kondo component to the overall intensity of the inelastic component was found to be

about 0.5, while the ratio of the intensity of the elastic line to the total intensity of the inelastic lines for $\{B_l^m\}_2$, which was used to calculate the inelastic component in Fig. 6, amounts to 0.91. If we assume that the compensation of the magnetic moment is reflected only in the intensity of the quasielastic scattering, which is fully described by the Kondo component just obtained, we can easily find that the reduction of the square of the total magnetic moment amounts to 21%.

The quasielastic scattering in spectra at higher temperatures was described only by a Lorentzian, since Eq. (4) can be used only at low temperatures.

4. DISCUSSION

The compound studied here contains variable-valence rare-earth ions as elements of the crystal matrix and is a Kondo lattice. However, the foregoing suggests that we were able to describe the main features of magnetic scattering by basing our reasoning exclusively on an incoherent or impurity approach. Below we continue our analysis along the same lines.

4.1. Comparison with the results of measuring macroscopic characteristics

Measurements of the static magnetic susceptibility of intermediate-valence systems suggest that at high temperatures its temperature dependence obeys the Curie-Weiss law:

$$\chi(T) = \frac{1}{3} N \mu_{\text{eff}}^2 \frac{1}{T + \Theta}, \qquad (6)$$

where N is the number of magnetic ions, μ_{eff} is the effective magnetic moment of a single ion, and $\Theta < 0$ is the negative paramagnetic Curie temperature. This experimental fact was described theoretically by Kojima et al.,[15] who used the noncrossing diagram approximation to calculate the static magnetic susceptibility of a variable-valence impurity.

Shimizu et al.[11] used the results of measuring χ in a single crystal (the experimental points on the plots of $\chi^{-1}(T)$ are depicted in Fig. 7 by filled and unfilled circles) to determine Θ for the directions along the crystallographic c axis and perpendicular to that axis: $\Theta_{\parallel} = -75$ K and $\Theta_{\perp} = -299$ K. Actually, in determining Θ from the high-temperature part of a $\chi^{-1}(T)$ curve there is a lot of arbitrariness in drawing the straight line for $\chi^{-1}(T) \propto (T + \Theta)$ due to the experimental error "corridor" and, more importantly, to the problem of correctly choosing the segment of the $\chi^{-1}(T)$ curve where the linear law holds.

Following the assumption of Sales and Wohlleben[28] that the absolute value of the negative paramagnetic temperature is close to the Kondo temperature, we equate $-\Theta_{\parallel}$ with the value of T_K obtained from an analysis of the low-energy feature of the magnetic scattering spectrum (see Sec. 3), i.e., $\Theta_{\parallel} = -26.5$ K. Note that the assumption in Ref. 28 is based on a phenomenological model of interconfigurational fluctuations and still requires rigorous mathematical substantiation.

The light solid line in Fig. 7 represents the Curie-Weiss law for the direction along the c axis obtained by connecting the point $(-26.5, 0)$ with the extrapolation of the experi-
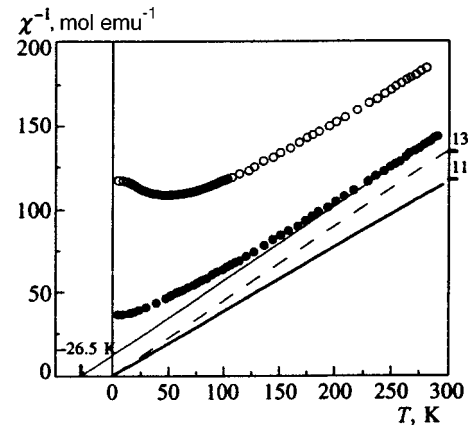

FIG. 7. Determining the high-temperature valence of the Yb ion in YbCu$_2$Si$_2$. Points—experimental data on the temperature dependence of the reciprocal static magnetic susceptibility of YbCu$_2$Si$_2$ (Ref. 11) (filled circles—along the crystallographic c axis, unfilled circles—in the perpendicular direction); light solid line—Curie-Weiss law with $\Theta = -26.5$ K corresponding to the effective magnetic moment of Yb in YbCu$_2$Si$_2$; dashed line—Curie law corresponding to the same effective magnetic moment of Yb in YbCu$_2$Si$_2$; heavy solid line—Curie law corresponding to the magnetic moment of Yb^{3+}.

mental $\chi_{\parallel}^{-1}(T)$ curve to $T = 300$ K. The dashed line represents the Curie law $\chi_{\parallel}^{-1}(T) \propto (\mu_{\text{eff}})^{-2} T$ with the same effective magnetic moment as in the Curie-Weiss law. The heavy solid line represents the Curie law for the Yb^{3+} ion ($\mu = 4.54 \mu_B$). The ratio of the square of the effective magnetic moment $(\mu_{\text{eff}})^2 = [\nu(T) - 2]\mu_{\text{Yb}^{3+}}^2$ to the square of the magnetic moment of the Yb^{3+} ion, $\mu_{\text{Yb}^{3+}}^2$, is equal to the reciprocal ratio of the slopes of the dashed and solid heavy straight lines, which thus equals $\nu(T) - 2$, where $\nu(T)$ is the valence of Yb in the compound under consideration at $T = 300$ K. Figure 7 shows that the high-temperature valence of Yb determined in this way is $2 + (117.2/135.3) = 2.87$, which is very close to the experimental value of 2.9 determined from the x-ray absorption edge.[2,3] Thus, the value $\Theta_{\parallel} = -26.5$ K and the assumption of Sales and Wohlleben[28] that $-\Theta_{\parallel}$ is close to T_K, on the basis of which this value was obtained, do not contradict the experimental facts. Note that the estimate of the magnetic-moment reduction at low temperature made at the end of Sec. 3 yields a low-temperature valence of Yb equal to 2.79, which appears to be a highly plausible result.

Figure 8 shows plots of $\chi^{-1}(T)$: experimental curves taken from Ref. 11 ($\chi_{\parallel}^{-1}(T)$—filled circles, and $\chi_{\perp}^{-1}(T)$—unfilled circles) and calculated curves (heavy solid line—based on the first set of crystal-field parameters, and dashed line—based on the second set of crystal-field parameters). The calculated curves have been corrected for the valence value obtained and have been shifted upward along the vertical axis so that the calculated and experimental $\chi^{-1}(T)$ curves would coincide at $T = 300$ K.

We were unable to find the set of parameters $\{B_l^m\}$ for which the anisotropy of χ would be determined only by the crystal field, although attempts to simultaneously fit the data on the static susceptibility and neutron scattering with variation of the crystal field were made. For this reason we as-
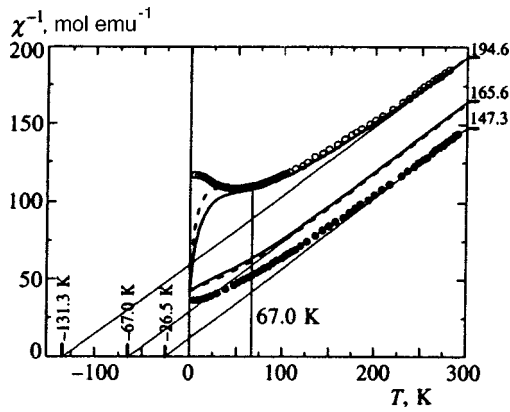
FIG. 8. Anisotropy of Θ: Points—experimental data on the temperature dependence of the reciprocal static magnetic susceptibility of YbCu$_2$Si$_2$ (Ref. 11) (filled circles—along the crystallographic c axis, unfilled circles—in the perpendicular direction); heavy lines—calculated temperature dependence of the reciprocal static magnetic susceptibility obtained on the basis of the first variant of the crystal field (solid line) and the second variant of the crystal field (dashed line) and shifted upward along the vertical axis of ordinates so that the calculated curve for the direction perpendicular to the c axis would coincide with the experimental curve.

sume that the anisotropy of the negative paramagnetic Curie temperature provided an additional contribution to χ. Figure 8 shows that $\Theta_\perp = -67$ K.

4.2. Comparison of the crystal fields in YbCu$_2$Si$_2$ and other compounds in the series

As noted in the Introduction, finding the crystal field in the title compound was a part of a systematic study of crystal fields in the RCu$_2$Si$_2$ series of compounds.[4–8] In their theoretical paper, Levy and Zhang[29] took the anisotropic hybridization interaction between the f electrons of rare-earth ions and conduction electrons into account by introducing an effective potential that contributed to the total crystal field. The main goal of the systematic study of the crystal fields in the members of the series is to experimentally determine this contribution to the crystal field of anomalous compounds by comparing them with normal compounds which exhibit antiferromagnetic properties.

The comparative analysis was based on the superposition model proposed by Newman and Ng.[30] In this model, the fourth and sixth orders of the crystal field (i.e., B_l^m with $l = 4$, 6) are determined only by the local environment of the rare-earth ions (ligands). Here the crystal-field potential is written as a superposition of the contributions of the different coordination spheres, and the contribution of each sphere is represented as a product of the geometric coordination factor $\bar{K}_{lm}$, which depends on the position of the ligand ions in the sphere, and what is known as the internal parameter of the respective sphere, $\tilde{A}_l(R_j)$, which is a measure of the participation of the respective coordination sphere in forming the corresponding crystal-field multipole:

$$B_l^m = \Theta_l \sum_{ij} \tilde{A}_l(R_j) K_{lm}(\theta_i, \varphi_i) = \Theta_l \sum_j \tilde{A}_l(R_j) \bar{K}_{lm}(R_j),$$

$$(7)$$

$$\bar{K}_{lm}(R_j) = \sum_i K_{lm}(\theta_i, \varphi_i),$$

where Θ_l is the Stevens factor, and R_j, θ_i, and φ_i are the coordinates of the ith ion in the jth coordination sphere.

A characteristic feature of the structure of compounds belonging to the series under consideration is that the local environment of the rare-earth ions comprises two coordination spheres, whose radii differ by less than 0.1 Å: the spheres of the Si and Cu ligands. The next coordination sphere is 0.5 Å away from them. As a result, for the fourth-order crystal field we can set up the system of equations

$$\frac{B_4^0}{\Theta_4} = \bar{K}_{40}(\mathrm{Si})\tilde{A}_4(\mathrm{Si}) + \bar{K}_{40}(\mathrm{Cu})\tilde{A}_4(\mathrm{Cu}),$$

$$(8)$$

$$\frac{B_4^4}{\Theta_4} = \bar{K}_{44}(\mathrm{Si})\tilde{A}_4(\mathrm{Si}) + \bar{K}_{44}(\mathrm{Cu})\tilde{A}_4(\mathrm{Cu}).$$

In the particular case of Yb, this system becomes

$$\frac{B_4^0}{-17.3 \times 10^{-4}} = -0.64\tilde{A}_4(\mathrm{Si}) - 2.17\tilde{A}_4(\mathrm{Cu}),$$

$$(9)$$

$$\frac{B_4^4}{-17.3 \times 10^{-4}} = -25.5\tilde{A}_4(\mathrm{Si}) + 5.10\tilde{A}_4(\mathrm{Cu}).$$

The values of $\bar{K}_{lm}$ were calculated from the structural data of Ref. 31 using the formulas on page 707 of Ref. 30.

By plugging the values of B_l^m found from experiments into (9) and solving the system we can find the values of the internal parameters $\tilde{A}_4(\mathrm{Si})$ and $\tilde{A}_4(\mathrm{Cu})$. As noted in Sec. 3, a neutron experiment does not makes it possible to determine the signs of B_4^4 and B_6^4, since it is insensitive to their simultaneous reversal.

If we choose the signs to be such that $B_4^4 > 0$ and $B_6^4 > 0$, (++) the system yields the following values of the fourth-order internal parameters:

$\tilde{A}_4(\mathrm{Si}) = -0.061$ meV, $\tilde{A}_4(\mathrm{Cu}) = -5.49$ meV for the first set $\{B_l^m\}$ and

$\tilde{A}_4(\mathrm{Si}) = -0.067$ meV, $\tilde{A}_4(\mathrm{Cu}) = -5.41$ meV for the second.

For the opposite choice of signs $(--)$ we have

$\tilde{A}_4(\mathrm{Si}) = -2.019$ meV, $\tilde{A}_4(\mathrm{Cu}) = -4.91$ meV for the first set $\{B_l^m\}$ and

$\tilde{A}_4(\mathrm{Si}) = -1.980$ meV, $\tilde{A}_4(\mathrm{Cu}) = -4.83$ meV for the second.

An entirely independent search for the crystal-field parameters in YbCu$_2$Si$_2$ in the (++) case yielded a very small value for $\tilde{A}_4(\mathrm{Si})$, which was characteristic of the normal crystal field in the series of isostructural compounds RCu$_2$Si$_2$ (see Refs. 5, 7, and 8). This simply cannot be a coincidence. Therefore, the (++) choice of signs is preferable.

The most characteristic feature of the crystal field of anomalous compounds belonging to the series, which was determined in previous studies, is the large (in absolute value) and negative (in sign) contribution of the Si sphere to the crystal field. Figure 9 shows a comparative diagram of the crystal fields (the internal parameters of the ligand
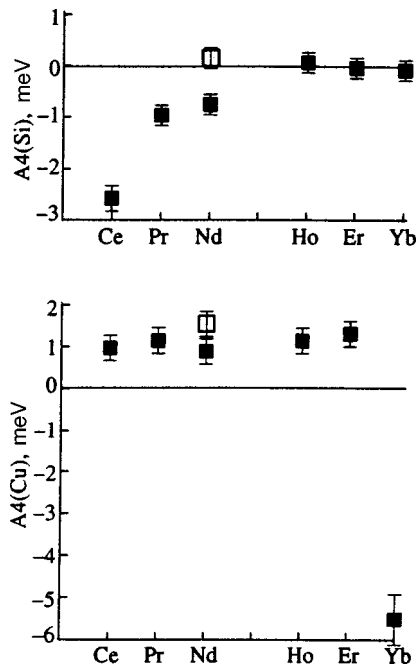
FIG. 9. Values of the parameters in the superposition model of ligand co-ordination spheres for various members of the RCu_2Si_2 series (in the case of R=Nd, an unfilled square corresponds to a low-energy value of a parameter and a filled square, to a high-temperature value).

FIG. 10. Structure of the wave functions of Yb^{3+} in the first variant of the crystal field.

spheres) in the series under consideration and the place occupied by Yb in it. We see that in the latter case the anomaly shifts with conservation of the sign from the Si sphere to the Cu sphere. Thus, as a result of our comparative analysis we arrive at the following conclusions:

1) just as $NdCu_2Si_2$, $PrCu_2Si_2$, and especially the heavy-fermion system $CeCu_2Si_2$, the intermediate-valence system $YbCu_2Si_2$ demonstrates a large anomaly in the fourth-order crystal field, which can be related to the hybridization contribution to the crystal-field potential;

2) the situation characteristic of the rare earths at the beginning of the series, i.e., in the case of rare-earth ions with a single f electron in the $4f$ shell (Ce), differs from the situation at the end of the series (Yb has one hole in the f shell instead of one electron): while in the former case there is strong hybridization of f electrons with Si p electrons, in the latter case there is hybridization between f electrons and copper electrons.

A system similar to (8) and (9) can be set up for the sixth-order crystal field and, knowing B_6^0 and B_6^4, we can find the values of the sixth-order internal parameters. One must bear in mind, however, that the relative error in determining B_6^0 and B_6^4 (i.e., the difference between the values in the first and second sets) is large. Table III lists all the data on the fourth- and sixth-order crystal fields for the first and second variants of the crystal field: the crystal-field parameters with the preferable signs of B_4^4 and B_6^4 and the corresponding values of the internal parameters of the superposition model.

Figure 10 depicts the structure of the wave functions for the first variant of the crystal field. In the case of the second variant, $\{B_l^m\}_2$, the differences are insignificant.

4.3. The intermediate-valence system $YbCu_2Si_2$ and anisotropic hybridization

Clearly, the assumption put forward in Sec. 4.1 that the negative paramagnetic temperature Θ is anisotropic does not allow the data on the static magnetic susceptibility of a single crystal to serve as an absolute criterion for determining whether the crystal field we found is the true one. Doubts regarding the validity of the assumption that Θ is anisotropic automatically lead to doubts regarding the crystal field found, and vice versa.

On the other hand, there is no unique interpretation of the anisotropy of Θ. This phenomenon may be due to the coherent mechanism of the interstitial interaction and is thus an inherent feature of the Kondo lattice. It may also be due to the features of hybridization at an individual rare-earth site, i.e., a mechanism of an incoherent or impurity nature. In the latter case the concept of anisotropic hybridization[29] becomes logically complete, and within this framework the role of the hybridization contribution to the crystal field of an intermediate-valence field becomes clear: using the assumption made in Ref. 28 that T_K and $-\Theta$ are close, we can link the values $-\Theta_\parallel = 26.5$ K and $-\Theta_\perp = 67$ K to different char-

TABLE III. Fourth- and sixth-order crystal fields and parameters of the superposition model.

Parameters of crystal field	Variant of crystal field		Parameters of superposition model	Variant of crystal field	
	1	2		1	2
B_4^0, meV	-0.21×10^{-1}	-0.20×10^{-1}	$\tilde{A}_4(Si)$, meV	-0.61×10^{-1}	-0.67×10^{-1}
B_4^4, meV	$+0.46 \times 10^{-1}$	$+0.45 \times 10^{-1}$	$\tilde{A}_4(Cu)$, meV	-5.49	-5.41
B_6^0, meV	-0.10×10^{-4}	-0.21×10^{-5}	$\tilde{A}_6(Si)$, meV	1.97	0.56
B_6^4, meV	$+0.17 \times 10^{-2}$	$+0.48 \times 10^{-3}$	$\tilde{A}_6(Cu)$, meV	1.47	0.42

acteristic hybridization energies: $\varepsilon_{K\parallel}=2.25$ eV in the direction of the c axis and $\varepsilon_{K\perp}=5.68$ meV in the aa plane.

It may also be that both the coherent and impurity mechanisms participate in the shaping of the anisotropic paramagnetic temperature Θ. Only experiments can provide the final answer to all the questions concerning the anisotropy of Θ and the parameters of the crystal field in $YbCu_2Si_2$. Two experiments are needed. First, we must measure the anisotropy of the static magnetic susceptibility not of the Kondo lattice but of the Kondo impurity. This requires studying a single crystal in which Yb is partially replaced by a rare-earth element with zero magnetic moment, e.g., Lu. Second, in a neutron scattering experiment involving an $YbCu_2Si_2$ single crystal we must measure the dynamic magnetic susceptibility in the direction of the c axis and in the perpendicular direction with energy transfers ranging from ~ 1 meV to 10 meV. Such measurements will make it possible to directly observe the values of ε_K corresponding to different crystallographic directions, while in experiments with a polycrystal these values may remain unresolved.

5. CONCLUSION

We have applied inelastic thermal-neutron scattering to the intermediate-valence system $YbCu_2Si_2$. Measurements performed at different temperatures have made it possible to separate the nonoverlapping ranges of energy transfers where magnetic scattering takes place. The characteristic value of the Kondo temperature for this compound has been determined on the basis of an analysis of the scattering at low energy transfers. An analysis of scattering in the high-energy region has made it possible to establish the phenomenological parameters of the crystal field, primarily of the fourth-order crystal field. The results of a comparison of these fields with the crystal fields of other compounds belonging to the same series suggest that there is hybridization of f electrons and copper electrons. By comparing the crystal field found in this work and the results of measurements of the static magnetic susceptibility we conclude that the negative paramagnetic Curie temperature is anisotropic in the direction of the c axis and in the perpendicular plane.

As for the question of how hybridization and the crystal field are related, we have arrived at the following conclusion: the $k-f$ hybridization interaction does indeed fit into the crystal-field model fairly well. As long as line broadening does not prevent us from separating the components in the scattering spectrum, the crystal-field model will provide a meaningful description of the behavior of a variable-valence rare-earth ion. All possible renormalizations of the spectrum, the drift of levels with temperature, etc. are nothing more than "ripples" that distort but do not entirely break the "backbone" of the phenomenological crystal-field Hamiltonian from which the varying properties of rare-earth ions from high to low temperatures "hang." The presence of hybridization directly leads to the formation of an entirely new crystal field, which differs dramatically from the crystal field of a rare-earth ion with a stable f shell.

According to the results of the treatment of data on the intermediate-valence system $YbCu_2Si_2$, the most significant deviations from a purely crystal-field behavior in the magnetic responses of systems with strong hybridization are related to the "quasielastic" component. In addition to the fact that this component contains a clearly distinguishable Kondo-scattering part and cannot be described by the relaxation approximation, there is a high probability that the magnetic-scattering "deficit" due to reduction of the effective magnetic moment of the variable-valence rare-earth ion, primarily affects the component's intensity. Note that we were successful in incorporating the crystal field of the Yb-Cu system into the general picture of the series by assuming that the inelastic scattering component follows the crystal-field models exactly, while ascribing the general decrease in magnetic scattering to its "quasielastic" part alone has led to a plausible low-temperature value of the valence of the Yb ion. Such a distinct status of the "quasielastic" component was to be expected, since it is directly related to the ground state of the Kondo system.

One can easily see that the applicability of the crystal-field model to the analysis of the spectrum of the inelastic incoherent magnetic scattering of neutrons in a Kondo system depends on the ratio of the width of the "quasielastic" scattering component or, which is the same, the position of the Kondo scattering peak, i.e., the Kondo energy, to the energy of the first excited level of the effective crystal field. If this ratio is equal to, or larger than, unity, the magnetic response is a broad undifferentiated line. Conversely, if the ratio is smaller than unity, at least two components can be distinguished in the spectrum, viz., a "quasielastic" and an inelastic component, and the latter can then be analyzed using the crystal-field model.

*)E-mail: muzyk@nf.jinr.dubna.su

[1] K. C. Sales and R. Viswanathan, J. Low Temp. Phys. **23**, 449 (1976).

[2] M. N. Groshev, M. D. Koterlin, E. M. Levin, R. V. Lutsiv, N. M. Miftakhov, Yu. P. Smirnov, A. E. Sovestnov, A. V. Tyunis, V. A. Shaburov, R. I. Yasnitskiĭ, S. M. Kuz'mina, V. I. Petrova, and V. A. Tyukavin, Fiz. Tverd. Tela (Leningrad) **28**, 2711 (1986) [Sov. Phys. Solid State **28**, 1519 (1986)].

[3] J. Rohler, in *Handbook of the Physics and Chemistry of Rare Earths*, K. A. Gschneider, Jr. and L. Eyring (eds.), North-Holland, Amsterdam (1988), Vol. 10, p. 453.

[4] E. A. Goremychkin and R. Osborn, Phys. Rev. B **47**, 14 280 (1993).

[5] E. A. Goremychkin, A. Yu. Muzychka, and R. Osborn, Physica B **179**, 184 (1992).

[6] E. A. Goremychkin, R. Osborn, and A. Yu. Muzychka, Phys. Rev. B **50**, 13 863 (1994).

[7] E. A. Goremychkin and A. Yu. Muzychka, Zh. Éksp. Teor. Fiz. **110**, 1339 (1996) [JETP **83**, 738 (1996)].

[8] A. Yu. Muzychka, *Preprint P-14-97-154*, Joint Institute for Nuclear Research, Dubna (1997).

[9] E. Holland-Moritz, D. Wohlleben, and M. Loewenhaupt, J. Phys. (Paris), Colloq. **39**, C6-835 (1978).

[10] E. Holland-Moritz, D. Wohlleben, and M. Loewenhaupt, Phys. Rev. B **25**, 7482 (1982).

[11] T. Shimizu, H. Yasuoka, Z. Fisk, *et al.*, J. Phys. Soc. Jpn. **56**, 4113 (1987).

[12] R. Currat and A. P. Murani, Physica B **156–157**, 812 (1989).

[13] V. Zevin, G. Zwicknagl, and P. Fulde, Phys. Rev. Lett. **60**, 2331 (1988).
[14] Y. Kuramoto, Z. Phys. B **53**, 37 (1983).
[15] H. Kojima, Y. Kuramoto, and M. Tachiki, Z. Phys. B: Condens. Matter **54**, 293 (1984).
[16] Y. Kuramoto and H. Kojima, Z. Phys. B: Condens. Matter **57**, 95 (1984).
[17] N. E. Bickers, D. L. Cox, and J. W. Wilkins, Phys. Rev. B **36**, 2036 (1987).
[18] K. Tomala and E. Holland-Moritz, J. Magn. Magn. Mater. **89**, 143 (1990).
[19] G. Zwicknagl, V. Zevin, and P. Fulde, Z. Phys. B: Condens. Matter **79**, 365 (1990).
[20] E. A. Goremychkin (private communication).
[21] K. W. H. Stevens, Proc. Phys. Soc. London, Sect. A **65**, 209 (1952).
[22] U. Walter, J. Phys. Chem. Solids **45**, 401 (1984).
[23] C. M. Varma, W. Weber, and L. J. Randall, Phys. Rev. B **33**, 1015 (1986).

[24] T. V. Ramakrishnan and K. Sur, Phys. Rev. B **26**, 1798 (1982).
[25] P. F. de Chatel, Solid State Commun. **41**, 853 (1982).
[26] O. Gunnarsson and K. Schönhammer, Phys. Rev. B **28**, 4315 (1983).
[27] Y. Kuramoto and E. Muller-Hartman, J. Magn. Magn. Mater. **52**, 122 (1985).
[28] B. C. Sales and D. K. Wohlleben, Phys. Rev. Lett. **35**, 1240 (1975).
[29] P. M. Levy and S. Zhang, Phys. Rev. Lett. **62**, 78 (1989).
[30] D. J. Newman and B. Ng, Rep. Prog. Phys. **52**, 699 (1989).
[31] A. Szytula and J. Leciejewicz, *Handbook of Crystal Structures and Magnetic Properties of Rare Earth Intermetallics*, CRC Press, Boca Raton, Fla. (1994).

Influence of the degree of disorder of amorphous solids on the intensity of light scattering by acoustic phonons

N. N. Ovsyuk*[)]

United Institute of Geology, Geophysics, and Mineralogy, Siberian Branch of the Russian Academy of Sciences, 630090, Novosibirsk

V. N. Novikov

Institute of Automation and Electrical Measurements, Siberian Branch of the Russian Academy of Sciences, 630090, Novosibirsk

It is observed experimentally that in the low-frequency Raman light scattering spectrum of amorphous porous silicon the boson peak situated in the acoustic range is more sensitive to the structural order than the optical mode presently used to determine the degree of disorder. It is shown that this is because, unlike the coefficient of interaction with optical vibrations, the coefficient of interaction between light and acoustic vibrations contains an additional factor, the square of the reciprocal correlation length of the vibrational excitations, i.e., the intensity of light scattering by acoustic phonons has an additional dependence on the degree of disorder. © *1998 American Institute of Physics.* [S1063-7761(98)02107-6]

1. INTRODUCTION

Problems are currently being encountered in determining the ratio of the volumes of the amorphous and crystalline phases in solids using Raman scattering, which is highly sensitive to structural order. In the literature the ratio of the volumes of the amorphous and crystalline phases in thin silicon films has been determined using the Raman spectra from the ratio of the integrated areas below the broad amorphous-like and narrow crystalline peaks of the transverse optical (TO) phonon.[1,2] However, this method is inaccurate because the annealing of amorphous silicon microparticles and films has a different influence on the Raman spectra in the optical range. In Ref. 3, for example, when small silicon particles were annealed to 800 °C, the Raman spectra revealed no appreciable changes and continued to show only one amorphous-like TO peak although the films, judging by the spectra, had become completely crystalline. In addition, the spectra of Raman scattering by optical phonons obtained for microparticles (<10 nm) also always contain a very large contribution of the amorphous component, although high-resolution electron microscopy suggests that their structure is crystalline.[4–6] We observed that light scattering by acoustic phonons is more sensitive to structural order than scattering in the optical range and thus, the area below the acoustic peak corresponds more accurately to the volume of the amorphous phase.

In order to ensure that the changes in the Raman spectra of amorphous solids are more noticeable as the degree of order increases, the sample must have dimensions comparable with the correlation length of the vibrational excitations. In this case, the changes in the volume of the amorphous phase will be more abrupt than in unbounded samples since a larger volume of the ordered phase will participate in

the light scattering. From this point of view, it is convenient to use either amorphous microparticles or microporous media. We chose amorphous microporous silicon. This material has been attracting interest following the observation of photoluminescence similar to the luminescence from crystalline porous silicon,[7,8] i.e., it was found that crystallinity is not a necessary condition for the observation of high-intensity visible luminescence at room temperature for microporous silicon. In another study of Raman scattering,[9] observations with a microscope attachment revealed that the regions of porous silicon contributing to the visible luminescence necessarily always contain some amorphous phase, in addition to the crystalline, i.e., the amorphous phase is evidently always present in microporous silicon. These facts provided an additional stimulus for a more accurate determination of the ratio of the volumes of the amorphous and crystalline phases in microstructures.

2. SAMPLES AND EXPERIMENTAL SETUP

Layers of porous silicon were obtained by anodizing (100)-oriented p^+-type silicon substrates with a resistivity of 0.006 $\Omega \cdot$cm in a hydrofluoric acid solution (42.4% HF:H_2O:C_3H_7OH in the ratio 2:1:2) at a current density of 100 mA/cm². This produced a 2 μm thick layer of silicon with 70% porosity. The samples were bombarded with 100 keV $^{10}B^+$ ions to produce an amorphous layer. As a result of this ion implantation, the amorphization dose of the porous silicon was 5×10^{15} cm^{-2}, which is an order of magnitude lower than the similar value for ordinary silicon. The Raman spectra were recorded in a 90° scattering geometry using a DFS-52 double monochromator with a spectral slit of 2 cm^{-1} width and $\lambda = 488$ nm exciting radiation under conditions of
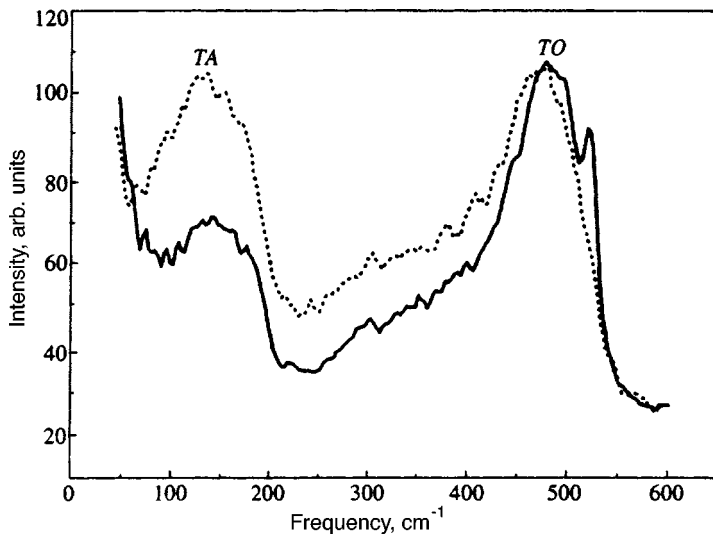
FIG. 1. Variation of the Raman light scattering spectrum of p^+-type amorphous porous (70% porosity) silicon as a result of annealing in a nitrogen atmosphere for 30 min at $T = 500\ °C$. The dashed and solid curves show the spectra before and after annealing, respectively.

doubly parallel polarization, i.e., when the exciting and scattered light beams were polarized in the scattering plane.

3. RESULTS AND DISCUSSION

It can be seen from Fig. 1 that the Raman spectrum of amorphous porous silicon consists of a broad TO mode at $480\ cm^{-1}$ and a broadened peak at $150\ cm^{-1}$ which is not observed for crystalline silicon. This type of peak in disordered materials is conventionally called a boson peak. In glasses it is usually observed at a frequency approximately 1/5 of the Debye frequency and is attributed to nanometer-scale structural correlations, reflecting an excess density of vibrational states in the low-frequency range ($20–80\ cm^{-1}$) compared with the Debye density. In tetragonal amorphous semiconductors, the correlation length is shorter than in glasses and is comparable with the interatomic distances. In particular, in silicon the boson peak coincides with the transverse acoustic TA mode and the density of vibrations near the boson peak is thus simply the density of the TA vibrations, but the reasoning put forward below will also apply to cases when the boson peak does not coincide with the TA mode. Since the boson peak does not appear in the Raman spectra of crystals, its intensity (like that of the broad TO mode) may also serve as a measure of the amorphousness of the material.

After the porous silicon had been converted to the amorphous state, we commenced isochronous stepwise annealing in a nitrogen atmosphere to study the influence of an increase in the degree of order on the Raman spectra in the optical and acoustic ranges. When the annealing temperature reached 500 °C, crystallization was initiated, this being observed as a very small spike at $520\ cm^{-1}$, which corresponds to the transverse optical phonon frequency in crystalline silicon, and the amplitude of the boson peak unexpectedly dropped sharply, by approximately half (Fig. 1). It is known that the size of the critical nucleus in bulk silicon is approximately 3–4 nm and the average size of the structural elements of porous silicon with 70% porosity, such as that being studied, is approximately 4–5 nm. Thus, the volume of the amorphous phase should decrease sharply at the instant

when crystal nuclei appear. It can be seen from Fig. 1 that the Raman spectrum only revealed an abrupt change in the amplitude of the boson peak at the instant of nucleation, while the amplitude and half-width of the amorphous-like TO peak remained almost unchanged, although its half-width should be reduced to approximately $30\ cm^{-1}$ as a result of annealing.[10] In our case, the half-width of the TO peak is $50\ cm^{-1}$, which suggests a high degree of structural disorder in the layers near the interface.

It is well known that stronger localization is observed[11] for optical vibrations, and thus when the structural order of an amorphous medium changes, the intensities of the Raman scattering by acoustic and optical vibrations may behave differently as a result of the different degree of violation of the wave vector selection rule. In the amorphous phase, as a result of the absence of translational invariance and the consequent nonconservation of the wave vector, the light scattering spectrum reveals the complete density of the acoustic or optical phonons with a weight proportional to the interaction constant $C(\omega)$ between the light and these vibrations (Ref. 12):

$$I(\omega) = C(\omega) g(\omega) (n(\omega) + 1)/\omega, \qquad (1)$$

where $n(\omega)$ is the Bose factor and $g(\omega)$ is the density of the acoustic or optical vibrational states.

We shall first analyze the case of scattering by acoustic vibrations. Since the density of the acoustic vibrations varies negligibly during annealing,[13] the observed decrease in the amplitude of the boson peak can only be attributed to a decrease in $C_{ac}(\omega)$ as a result of a change in the degree of violation of the wave vector selection rule. We shall use a method of analyzing the violation of the selection rule which is based on introducing a correlation length characterizing the spatial elongation of the normal vibrational mode and we shall show how $C_{ac}(\omega)$ is related to the structural order.

In general, the Raman light scattering intensity is determined by the fluctuations of the permittivity tensor $\delta\chi_{\alpha\beta}(\mathbf{r})$ (Refs. 14 and 15):

$$I_{\alpha\beta\gamma\delta}(\mathbf{q},\omega) \propto \int dt\, d\mathbf{r}_1 d\mathbf{r}_2 \exp(i\omega t - i\mathbf{q}\cdot(\mathbf{r}_1 - \mathbf{r}_2))$$

$$\times \langle \delta \chi_{\alpha\beta}^{+}(\mathbf{r}_1,t) \delta \chi_{\delta\gamma}(\mathbf{r}_2,0) \rangle, \tag{2}$$

where $\mathbf{q}=\mathbf{q}_1-\mathbf{q}_2$ is the difference between the initial and final photon momenta. The variations of the tensor $\delta \chi_{\alpha\beta}(\mathbf{r},t)$ under the action of the acoustic phonons are proportional to the deformation tensor $\partial u_\alpha/\partial r_\beta+\partial u_\beta/\partial r_\alpha$, where $u_\alpha(\mathbf{r},t)$ is the displacement of the atoms induced by the vibrations. Since the vibrations determining the boson peak lie in the acoustic range, it may be assumed that they also interact with light by means of the deformation tensor. Bearing in mind that $q \ll l^{-1}$, where $q \sim 2\pi/\lambda \sim 10^{-2}$ nm is the momentum imparted by the light ($\lambda=500$ nm) and l is the characteristic length of the vibrations near the boson peak, of the order of a nanometer,[15] we can neglect the term $\mathbf{q}\cdot(\mathbf{r}_1-\mathbf{r}_2)$ in the exponential function in Eq. (2) since this is much less than unity. For the Raman scattering intensity we then have

$$I_{ac}(\omega) \propto \int d\mathbf{r} \langle \nabla u_\omega^{+}(\mathbf{r})\cdot\nabla u_\omega(0) \rangle g_{ac}(\omega). \tag{3}$$

Here $u_\omega(\mathbf{r})$ is the amplitude of the vibrations at frequency ω and the angular brackets denote spatial and statistical averaging. The result of the averaging depends on how the vibrational displacements correlate at different spatial points. The correlation function of the gradients of the vibrational displacements at the frequency ω is conveniently expressed in terms of the normalized correlation function $F_{\omega,ac}(r)$:

$$\langle \nabla u_\omega^{+}(\mathbf{r})\cdot\nabla u_\omega(0) \rangle = F_{\omega,ac}(r) \langle |\nabla u_\omega(0)|^2 \rangle. \tag{4}$$

The form of the correlation function near the boson peak in amorphous solids is not known exactly, since the nature of these vibrations is not sufficiently well understood: it is normalized using the condition $F_\omega(r)\to 1$ when $r\to 0$. However, it is generally acknowledged that these are quasilocal vibrational excitations with a characteristic dimension of nanometer order.[14,15] When averaged over an ensemble of these vibrations localized in regions of a disordered solid having different configurations, the corresponding correlation function should decrease as a function of distance with a certain characteristic correlation length of nanometer order. In addition, an additional decay of the correlations may take place inside the region of localization as a result of the specific geometry of the vibrational modes which, as predicted, may have dimensions of less than three or fractal dimensions. This geometric factor leads to the appearance of an additional factor $r^{-\alpha}$ in the correlation function, where $\alpha<3$, which reflects the power decrease in the correlation. As a result, the correlation function has the form

$$F_{\omega,ac}(r)=(a/r)^\alpha f(r/l_{\omega,ac}), \tag{5}$$

where $l_{\omega,ac}$ is the correlation length of the acoustic vibrations, a is the interatomic distance, $f(r/l_{\omega,ac})$ is a decreasing function which reflects the decay of the correlations with distance as a result of localization of the vibrational modes, and the factor $r^{-\alpha}$ is associated with the internal geometry of the vibrations. The specific form of the function $f(r/l_{\omega,ac})$ is unimportant for the present study, as will become clear from the following analysis, although it may be postulated that this corresponds to exponential damping, $\exp(-r/l_{\omega,ac})$, or

Gaussian damping, $\exp(-r^2/l_{\omega,ac}^2)$. For example, for a damped plane acoustic wave which describes long-wavelength acoustic vibrations in an amorphous solid, the correlation function (5) is given by

$$\exp(i\mathbf{q}\cdot\mathbf{r}-r/l_\omega)(1+r/l_\omega+r^2/3l_\omega^2).$$

The oscillation factor naturally disappears from the boson peak for quasilocal vibrations because of the large variations in the form of the natural modes and the wide range of wave vectors required to form a localized mode. Bearing in mind that for harmonic vibrations and the Stokes component of the spectrum

$$\langle |u_\omega(0)|^2 \rangle \propto (n(\omega)+1)/\omega,$$

and that

$$\nabla u_\omega \sim u_\omega/l_{\omega,ac},$$

we have

$$\langle |\nabla u_\omega(0)|^2 \rangle \sim l_{\omega,ac}^{-2}\langle |u_\omega(0)|^2 \rangle \propto l_{\omega,ac}^{-2}(n(\omega)+1)/\omega.$$

As a result, Eq. (4) has the form

$$\langle \nabla u_\omega^{+}(\mathbf{r})\cdot\nabla u_\omega(0) \rangle \propto l_{\omega,ac}^{-2}F_{\omega,ac}(r)(n(\omega)+1)/\omega. \tag{6}$$

Substituting Eq. (6) into Eq. (3) and comparing with Eq. (1), we obtain

$$C_{ac}(\omega) \propto l_{\omega,ac}^{-2} \int d\mathbf{r}\, F_{\omega,ac}(r). \tag{7}$$

Unlike formula (3), for optical vibrations, the light scattering intensity is determined directly by the correlation function of the atomic vibrational displacements:

$$I_{opt}(\omega) \propto \int d\mathbf{r} \langle u_\omega^{+}(\mathbf{r})u_\omega(0) \rangle g_{opt}(\omega). \tag{8}$$

In this case, there is no need to impose the constraint that the Hamiltonian of the interaction between the vibrations and light be expressed in terms of derivatives of the displacement components with respect to the coordinates, since for optical vibrations the center of gravity of a unit cell remains constant (see, for example, Ref. 16). Here the correlation function of the vibrational displacements has the form

$$\langle u_\omega^{+}(\mathbf{r})u_\omega(0) \rangle \propto F_{\omega,opt}(r)(n(\omega)+1)/\omega. \tag{9}$$

Substituting Eq. (9) into Eq. (8) and comparing with Eq. (1), we obtain

$$C_{opt}(\omega) \propto \int d\mathbf{r}\, F_{\omega,opt}(r). \tag{10}$$

It can be seen from a comparison of Eqs. (7) and (10) that the interaction coefficient between light and optical phonons does not contain the additional square of the correlation length in the denominator which appears as a result of the gradients of the displacements for the acoustic phonons. Thus, as the correlation length increases under annealing, the acoustic part of the spectrum decays more rapidly than the optical part, i.e., is more sensitive to the degree of disorder.

We shall now explain the changes in the Raman spectrum observed when amorphous porous silicon is annealed

(Fig. 1). We shall first show how specific is the dependence of $C_{ac}(\omega)$ on the correlation length. Using Eq. (5), it is easy to show that

$$\int d\mathbf{r}\, F_{\omega,ac}(r) \propto b l_\omega^{3-\alpha},$$

where

$$b = \int_0^\infty x^{2-\alpha} f(x)\, dx$$

is a certain constant which depends on the specific form of the correlation function. It then follows from Eq. (7) that

$$C_{ac}(\omega) \propto l_{\omega,ac}^{1-\alpha}.$$

Since the vibrations near the boson peak are localized and satisfy the Ioffe–Regel criterion,[17] $l_{\omega,ac} \sim \lambda \propto \omega^{-1}$, where $\lambda = 2\pi c/\omega$ is the wavelength of the acoustic excitation, c is the velocity of light, and assuming that near the boson peak $C_{ac}(\omega) \propto \omega$ (Ref. 18), we find that

$$l_{\omega,ac}^{1-\alpha} = l_{\omega,ac}^{-1},$$

i.e.,

$$C_{ac}(\omega) \propto l_{\omega,ac}^{-1}.$$

If crystallites of diameter D appear within the illuminated region, the mean free path of an acoustic vibration in these crystallites should be equal to their size $l_{\omega,ac} \sim D$. In this case, the reciprocal correlation length averaged over the sample $l_{\omega,ac}^{-1}$ undergoes a jump since the upper limit of the correlation length for amorphous silicon is 1.2–1.5 nm and the size of a critical nucleus is 3–4 nm, and thus the intensity of the boson peak in Fig. 1 drops sharply, being approximately halved.

Thus, it can be concluded that when Raman scattering is used to determine the volume ratio of the amorphous component in solids consisting of mixed phases, the boson peak should be used in preference to the amorphous-like optical mode currently used since this peak is more sensitive to the order because the intensity of light scattering by acoustic phonons has an additional dependence on the degree of disorder.

[1] R. Tsu, J. Gonzalez-Hernandez, S. S. Chao, S. C. Lee, and K. Tanaka, Appl. Phys. Lett. **40**, 534 (1982).
[2] E. Bustarret, M. A. Hachicha, and M. Brunel, Appl. Phys. Lett. **52**, 1675 (1988).
[3] T. Okada, T. Iwaki, H. Kasahara, and K. Yamamoto, Solid State Commun. **52**, 363 (1984).
[4] S. Hayashi and H. Abe, Jpn. J. Appl. Phys., Part 1 **23**, 824 (1984).
[5] S. Hayashi and K. Yamamoto, Superlatt. Microstruct. **2**, 581 (1986).
[6] N. N. Ovsyuk, E. B. Gorokhov, V. V. Grishchenko, and A. P. Shebanin, JETP Lett. **47/298**, 1988 ().
[7] E. Bustarret, M. Ligeon, and I. Ortega, Solid State Commun. **83**, 461 (1992).
[8] E. Bustarret, E. Sauvain, M. Ligeon, and M. Rosenbauer, Thin Solid Films **276**, 134 (1996).
[9] J. M. Perez, J. Villalobos, P. McNeill, J. Prasad, R. Cheek, J. Kelber, J. P. Estrera, P. D. Stevens, and R. Gloser, Appl. Phys. Lett. **61**, 563 (1992).
[10] R. Tsu, J. G. Hernandez, and F. H. Pollak, J. Non-Cryst. Solids **66**, 109 (1984).
[11] P. Sheng and M. Zhow, Science **253**, 539 (1991).
[12] R. Shucker and R. W. Gammon, Phys. Rev. Lett. **25**, 222 (1970).
[13] T. Shimada, Y. Katayama, K. nakagawa, H. Matsuhara, M. Migitaka, and E. Murayama, J. Non-Cryst. Solids **59/60**, 783 (1983).
[14] J. Jäckle, in *Amorphous Solids: Low-Temperature Properties*, edited by W. A. Phillips (Springer-Verlag, Berlin, 1981), p. 135.
[15] A. J. Martin and W. Brenig, Phys. Status Solidi B **64**, 163 (1974).
[16] V. L. Bonch-Bruevich and S. G. Kalashnikov, *Semiconductor Physics* [in Russian], Nauka, Moscow (1977), p. 465.
[17] M. Foret, E. Courtens, R. Vasher, and J.-B. Suck, Phys. Rev. Lett. **77**, 3831 (1996).
[18] V. K. Malinovsky, V. N. Novikov, P. P. Parshin, A. P. Sokolov, and M. G. Zemlyanov, Europhys. Lett. **11**, 43 (1990).

Translated by R. M. Durham

High-frequency modulation of current in tunnel-coupled quantum wells by transverse radiation in the terahertz range

F. T. Vas'ko[*]

Institute of Semiconductor Physics, Ukrainian National Academy of Sciences, 252650 Kiev, Ukraine

The modulation of a longitudinal current in double quantum wells by a transverse voltage in the terahertz frequency range is calculated. A quantum kinetic equation which allows for the contributions of a pair of asymmetric tunnel-coupled levels to the resonant response is solved (the difference in effective mass or the asymmetric scattering in the left- and right-hand quantum wells is also taken into account). These structural features lead to conversion of the modulating voltage into a high-frequency longitudinal current. The calculated longitudinal-transverse nonlinear susceptibility determines the conditions needed for creating an efficient field-effect transistor circuit that would use pumping in the terahertz range instead of a transverse control voltage. © *1998 American Institute of Physics.* [S1063-7761(98)02207-0]

1. INTRODUCTION

The modulation of a longitudinal electric current in double quantum wells with asymmetric scattering by a static transverse voltage has been widely studied during the last decade (see Ref. 1 and the literature cited in Ref. 2). The possible applications of this effect, known as resistance resonance, in creating a transistor structure, were demonstrated by Patel *et al.*[2] and Ohno *et al.*,[3] while the frequency dispersion of resistance resonance (i.e., the response to a variable alternating voltage with a frequency of order of the collision rate) was discussed in Ref. 4. On the other hand, as the frequency grows, the response becomes dependent on resonant transitions between tunnel-coupled levels. The resonant response of electrons in double quantum wells to terahertz radiation whose electric field is polarized transversely to the structure was studied by Heyman *et al.*,[5] who measured the transmission and the induced transverse voltage (the rectification effect). On the basis of the experiments described in Ref. 1–3 and 5, we shall study the conditions for, and the features of, the modulation of a longitudinal current in a doped double quantum well by a resonantly pumped, alternating electric field in the terahertz range, i.e., we shall study a double-quantum-well field-effect transistor that uses modulation by submillimeter resonant radiation instead of a static (or microwave) transverse control voltage.

Here we shall calculate the nonlinear response of electrons in double quantum wells to a longitudinal constant field **F** and a transverse high-frequency field E_ω (Fig. 1), a response determined by the induced current

$$\delta \mathbf{j} = \omega \chi_\omega \mathbf{F} E_\omega, \tag{1}$$

where the component χ_ω of the second-order nonlinear susceptibility tensor determines the coupling of the longitudinal and transverse responses in double quantum wells. Conversion of a transverse voltage into a longitudinal current occurs only for an asymmetric double quantum well, since the rank-3 tensor $\hat{\chi}_\omega$ vanishes for a symmetric well, i.e., a het-

erostructure that does change in response to the replacement $z \rightarrow -z$. The calculations we are about to do take into account the asymmetry of the energy band diagram of a double quantum well and the difference in effective mass or the unequal scattering efficiency in the left- (l) and right-hand (r) quantum wells. The asymmetry of the energy spectrum of double quantum wells, which is determined by the splitting between the energy levels in the l and r quantum wells, can be controlled by applying a transverse voltage to the double-quantum-well structure. The dissimilar scattering in tunnel-coupled quantum wells comes into effect when the doping levels of the wells are different (see Refs. 1–3), while the difference in effective mass is due to the nonparabolicity of the energy spectra of the l and r quantum wells, the different subbarrier penetration of these wells by the wave functions, or by the difference in composition of the alloys forming the l and r quantum wells (the energy spectrum of double quantum wells with the l and r quantum wells obeying different dispersion laws was discussed in Ref. 6). A description of the nonlinear response of structures is given below on the basis of a kinetic equation written in the two-level approximation (a derivation of such an equation can be found in Ref. 7).

Section 2 gives the basic formulas needed to describe the nonlinear response of double quantum wells with different masses and asymmetric scattering. These formulas are based on the solution of the quantum kinetic equation. The results of calculating χ_ω, which specify the spectral relations and the dependence of the longitudinal-transverse conversion efficiency on the splitting between the energy levels (which can be controlled by applying a transverse field to the double quantum well), are presented in Sec. 3. Finally, a discussion of the various approximations and concluding remarks can be found in Sec. 4.

2. LONGITUDINAL-TRANSVERSE NONLINEAR RESPONSE

The quantum kinetic equation for the high-frequency contribution to the 2×2 density matrix, $\delta \hat{f}_\mathbf{p} e^{-i\omega t}$, which
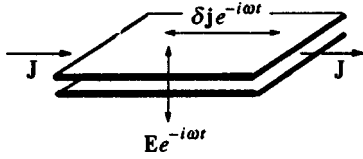
FIG. 1. Modulation of a longitudinal current by a transverse high-frequency voltage in double quantum wells.

describes the response of the electrons in an asymmetric double quantum well to a transverse perturbation $\delta \hat{h} e^{-i\omega t}$, can be written as follows:

$$-i\omega \delta \hat{f}_{\mathbf{p}} + \frac{i}{\hbar}[\hat{h}_p, \delta \hat{f}_{\mathbf{p}}] + \frac{i}{\hbar}[\delta \hat{h}, \hat{f}_{\mathbf{p}}]$$

$$+ e\mathbf{F} \cdot \frac{\partial \delta \hat{f}_{\mathbf{p}}}{\partial \mathbf{p}} = \hat{I}_{\text{sc}}(\delta \hat{f}|\mathbf{p}). \qquad (2)$$

Here we have used the approximation described in Ref. 4 of a pair of tunnel-coupled states in the l and r quantum wells, in which the Hamiltonian $\hat{h}_p$ allowing for the difference between the effective masses (m_l and m_r) is given by the matrix

$$\hat{h}_p = \begin{vmatrix} p^2/2m_l + \Delta/2 & T \\ T & p^2/2m_r - \Delta/2 \end{vmatrix}, \qquad (3)$$

where Δ is the splitting between the levels in the absence of tunneling, and T is the tunneling matrix element. The effect of the uniform longitudinal electric field $\mathbf{F}$ is taken into account in Eq. (1) through both the usual field contribution and the current contributions to the steady-state distribution $\hat{f}_{\mathbf{p}}$, and the collision integral $\hat{I}_{\text{sc}}$ is written below for the case of scattering on static inhomogeneities, the scattering efficiency being different in the l and r quantum wells (the respective expressions are given in Refs. 7 and 4). The operator $\delta \hat{h}$ describes intersubband transitions excited by the transverse electric field $E_\omega e^{-i\omega t}$ and is given by the following expression:

$$\delta \hat{h} = \frac{ie}{\omega} E_\omega \hat{v}_\perp, \quad \hat{v}_\perp = \frac{T}{\hbar} Z \hat{\sigma}_y, \qquad (4)$$

where the off-diagonal matrix $\hat{v}_\perp$ has the meaning of the interwell transition rate,[8] $\hat{\sigma}_i$ are the Pauli matrices ($i = x, y, z$), and Z is the distance between the maxima of the wave functions in the l and r quantum wells.

The longitudinal velocity operator is determined by the diagonal components $\mathbf{p}/m_{l,r}$, so that the high-frequency contribution to the longitudinal current can be written as

$$\delta \mathbf{j} = \frac{2e}{m} \int \frac{d\mathbf{p}}{(2\pi\hbar)^2} \mathbf{p}\, \text{Tr}\left(1 + \frac{m}{2M}\hat{\sigma}_z\right)\delta \hat{f}_{\mathbf{p}}, \qquad (5)$$

where Tr stands for the trace in a discrete variable related to the matrix nature of the Hamiltonian (2), and the average effective mass m and the effective-mass difference M are introduced as follows:

$$m^{-1} = \frac{m_l^{-1} + m_r^{-1}}{2}, \quad M^{-1} = m_l^{-1} - m_r^{-1}. \qquad (6)$$

If $m_l = m_r$, then (5) can be expressed in terms of $\text{Tr}\,\delta f_{\mathbf{p}}$, and applying the operator Tr to Eq. (2) we get $-i\omega \delta \hat{f}_{\mathbf{p}} = 0$, provided we ignore collisions and the dynamic field contribution to (2). This contribution is small if $eF\omega^{-1} \ll p_F$ (p_F is the Fermi momentum of the strongly degenerate electrons), but the distribution $\hat{f}_{\mathbf{p}}$ contains a current contribution (which is asymmetric in the $\mathbf{p}$ plane) that determines the effect for $\omega\bar{\tau} \gg 1$ ($\bar{\tau}$ is the average relaxation time). Thus, the fact that (5) is finite is due either to the difference in effective mass (and χ_ω is small in the parameter m/M) or to the allowance for collisions (and χ_ω near resonance contains the small factor $\hbar/T\bar{\tau}$).

It is convenient to solve Eq. (2) and calculate the current (5) by employing the basis set of wave functions of the eigenvalue problem $\hat{h}_p |jp\rangle = \varepsilon_{jp}|jp\rangle$. Here $j = \pm$ classifies the tunnel-coupled levels, and ε_{jp} specify the dispersion laws of the $\pm$ states, which are separated by an energy gap $\Delta_T = \sqrt{\Delta^2 + (2T)^2}$ when $p = 0$. If we neglect the off-diagonal components of the steady-state distribution $\hat{f}_{\mathbf{p}}$, which are small in the parameter $\hbar/\Delta_T\bar{\tau}$, we can rewrite Eq. (2) for the asymmetric contribution $\delta f_{jj'}(\mathbf{p})$ to the high-frequency response as

$$\frac{i}{\hbar}(\varepsilon_{jp} - \varepsilon_{j'p} - \hbar\omega)\delta f_{jj'}(\mathbf{p})$$

$$+ \frac{i}{\hbar}\delta h_{jj'}(p)(f_{j\mathbf{p}}^{\text{as}} - f_{j'\mathbf{p}}^{\text{as}}) = I_{jj'}(\delta f|\mathbf{p}). \qquad (7)$$

In the basis set of functions adopted, $\delta h_{jj'}(p) = \langle jp|\delta \hat{h}|j'p\rangle$, $f_{j\mathbf{p}}^{\text{as}}$ is the asymmetric part of the steady-state distribution, and the collision integral for scattering by a short-range static potential becomes

$$I_{jj'}(\delta f|\mathbf{p}) = \frac{i}{\hbar}\sum_{k=l,r} w_k \int \frac{d\mathbf{p}_1}{(2\pi\hbar)^2}$$

$$\times \sum_{j_1 j_1'}\left[\frac{\delta f_{jj_1}(\mathbf{p})}{\varepsilon_{jp} - \varepsilon_{j_1'p_1} - \hbar\omega - i\lambda}\langle j_1 p|\hat{P}_k|j_1'p_1\rangle \right.$$

$$\times \langle j_1'p_1|\hat{P}_k|j'p\rangle + \langle jp|\hat{P}_k|j_1'p_1\rangle$$

$$\left. \times \langle j_1'p_1|\hat{p}_k|j_1 p\rangle \frac{\delta f_{j_1 j_1'}(\mathbf{p})}{\varepsilon_{j_1'p_1} - \varepsilon_{j'p} - \hbar\omega - i\lambda}\right]. \qquad (8)$$

where $\hat{P}_k$ is the matrix for projection onto the kth quantum well, the w_k are related to the momentum relaxation time τ_k by the expression $w_k = \hbar^3/m\tau_k$, and $\lambda \to +0$.

The diagonal components of the solution of Eq. (7), $\delta f_{j\mathbf{p}} \equiv \delta f_{jj}(\mathbf{p})$, are small in the parameter $1/\omega\bar{\tau}$:

$$\delta f_{j\mathbf{p}} \simeq \frac{i}{\omega}I_{jj}(\delta\tilde{f}|\mathbf{p}), \qquad (9)$$

where the only term left on the right-hand side in the resonant approximation ($\omega \simeq \Delta_T/\hbar$) is the large off-diagonal

component $\delta \widetilde{f}_{\mathbf{p}} \equiv \delta f_{+-}(\mathbf{p})$. The induced current density (5) is expressed in terms of these corrections to the high-frequency response as follows:

$$\delta \mathbf{j} \simeq \frac{2e}{m} \int \frac{d\mathbf{p}}{(2\pi\hbar)^2} \mathbf{p} \left[\sum_j \delta f_{j\mathbf{p}} + \frac{m}{M} \langle -|\hat{\sigma}_z|+\rangle \delta \widetilde{f}_{\mathbf{p}} \right]. \tag{10}$$

To within first-order terms in m/M we can calculate $\delta f_{j\mathbf{p}}$ and $\delta \widetilde{f}_{j\mathbf{p}}$ by neglecting the effective-mass difference. Here the resonant contribution to the distribution $\delta \widetilde{f}_{j\mathbf{p}}$ can be obtained from (7) in the form

$$\delta \widetilde{f}_{\mathbf{p}} = \frac{\langle +|\delta \hat{h}|-\rangle (f^{as}_{+\mathbf{p}} - f^{as}_{-\mathbf{p}})}{\Delta_T - \hbar\omega - \gamma_p}, \tag{11}$$

where the collisional renormalization γ_p of the interlevel transition energy is introduced as

$$\gamma_p = \sum_{k=l,r\ j=\pm} w_k \int \frac{d\mathbf{p}_1}{(2\pi\hbar)^2}$$

$$\times \left[\frac{\langle -|\hat{P}_k|j\rangle\langle j|\hat{P}_k|-\rangle}{\varepsilon_{+p} - \varepsilon_{jp_1} - \hbar\omega - i\lambda} + \frac{\langle +|\hat{P}_k|j\rangle\langle j|\hat{P}_k|+\rangle}{\varepsilon_{jp_1} - \varepsilon_{-p} - \hbar\omega - i\lambda} \right]. \tag{12}$$

The real contributions to (12) diverge logarithmically at large $|\mathbf{p}_1|$, and the integral must be truncated at momenta of order $\hbar/l_c$, where l_c is the characteristic size of short-wavelength inhomogeneities. As a result $\mathrm{Re}\,\gamma_p$ renormalizes the interlevel splitting, and the denominator in (11) acquires the term

$$\widetilde{\Delta}_T \simeq \Delta_T - \frac{\Lambda}{\pi} \left(\frac{\hbar}{\tau_l} - \frac{\hbar}{\tau_r} \right) \frac{\Delta}{\Delta_T}, \quad \Lambda = \ln\frac{\hbar}{l_c\bar{p}}, \tag{13}$$

where $\bar{p}$ is the characteristic momentum, which determines the leading contribution to (10) and is equal to the maximum value of p_F: $\sqrt{2m\Delta_T}$ or $\sqrt{2m\hbar\omega}$. The imaginary contribution to (11), $i\bar{\gamma}_p$, becomes

$$\bar{\gamma}_p = \sum_{kj} \frac{\hbar}{2\tau_k} [|\langle -|\hat{P}_k|j\rangle|^2 \theta(\varepsilon_{+p} - \varepsilon_j - \hbar\omega)$$

$$+ |\langle +|\bar{P}_k|j\rangle|^2 \theta(\varepsilon_j - \varepsilon_{-p} - \hbar\omega)], \tag{14}$$

so that the denominator in (11) can be rewritten as $\widetilde{\Delta}_T - \hbar\omega - i\bar{\gamma}_p$. Similarly, the contribution of $\Sigma_j I_{jj}$ [found from (8)] to the first term becomes

$$\sum_{j=\pm} I_{jj}(\delta \widetilde{f}|\mathbf{p}) = -\delta \widetilde{f}_{\mathbf{p}} \sum_{kj'} (2\tau_k)^{-1} \langle -|\hat{P}_k|j\rangle\langle j|\hat{P}_k|+\rangle$$

$$\times [\theta(\varepsilon_{+p} - \varepsilon_{j'} - \hbar\omega) + \theta(\varepsilon_{j'} - \varepsilon_{-p} - \hbar\omega)]. \tag{15}$$

Thus, the calculation of χ_ω is reduced to finding the integral in (10) with (11) and (15) plugged into $\delta \mathbf{j}$ [with allowance for (14)]. Here the asymmetric correction $f^{as}_{j\mathbf{p}}$ to the distribution can be written as

$$f^{as}_{j\mathbf{p}} = \frac{e}{m} (\mathbf{F}\cdot\mathbf{p}) \tau_j \delta(\varepsilon_F - \varepsilon_{jp}),$$

$$\tau_j^{-1} = \sum_{kj'} \tau_k^{-1} |\langle j|\hat{P}|j'\rangle|^2 \theta(\varepsilon_F - \varepsilon_{j'}), \tag{16}$$

where $\varepsilon_F = p_F^2/2m$ is the Fermi energy, and $\tau_\pm$ is the momentum relaxation time for the $\pm$ tunnel-coupled states. The use of the weakly anisotropic distribution (16) presupposes that $eF\bar{\tau} \ll p_F$.

3. SPECTRAL RELATIONS AND DEPENDENCE OF χ_ω ON Δ

According to (16), only the electrons on the Fermi surface contribute to $\delta \mathbf{j}$, so that the integration over the angle and energy in (10) can easily be performed, and closed expressions are obtained for χ_ω. Here it is convenient to separate the contributions of the first and second terms in (10) by writing χ_ω as $\chi_\omega^{(1)} + \chi_\omega^{(2)}$. The contribution $\chi_\omega^{(1)}$, which is specified by (15), becomes

$$\chi_\omega^{(1)} \simeq i\frac{e^3 Z}{m\omega} \left(\frac{T}{\Delta_T} \right)^2 \sum_{kj} \frac{\hbar}{2T\tau_k} \langle -|\hat{P}_k|j\rangle\langle j|\hat{P}_k|+\rangle$$

$$\times \Bigg\{ \frac{n_-\tau_-}{\widetilde{\Delta}_T - \hbar\omega - i\bar{\gamma}_-} [\theta(\varepsilon_F + \Delta_T - \varepsilon_j - \hbar\omega)$$

$$+ \theta(\varepsilon_j - \varepsilon_F - \hbar\omega)] - \frac{n_+\tau_+}{\widetilde{\Delta}_T - \hbar\omega - i\bar{\gamma}_+}$$

$$\times [\theta(\varepsilon_F - \varepsilon_j - \hbar\omega) + \theta(\varepsilon_j - \varepsilon_F + \Delta_T + \hbar\omega)] \Bigg\}, \tag{17}$$

while the contributions proportional to m/M are given by the expression

$$\chi_\omega^{(2)} \simeq \frac{e^3 Z}{m\omega} \left(\frac{T}{\Delta_T} \right)^2 \frac{2m}{M}$$

$$\times \left[\frac{n_-\tau_-}{\widetilde{\Delta}_T - \hbar\omega - i\bar{\gamma}_-} - \frac{n_+\tau_+}{\widetilde{\Delta}_T - \hbar\omega - i\bar{\gamma}_+} \theta\left(\varepsilon_F - \frac{\Delta_T}{2} \right) \right]. \tag{18}$$

Equations (17) and (18) contain the electron concentrations at the $\pm$ levels, $n_\pm = \rho_{2D}(\varepsilon_F \pm \Delta_T/2)$ for $n > \rho_{2D}\Delta_T$, while at lower concentrations $n_+ = 0$ and $n_- = n$. The relaxation times $\tau_\pm$ are given by (16), and the broadening energies $\bar{\gamma}_\pm$ are obtained by calculating (14) with $\varepsilon_{\pm p} = \varepsilon_F$.

At frequencies close to resonance for $\chi_\omega^{(1,2)}$ we have Lorentzian spectral relations [here we set $\hbar\omega \simeq \Delta_T$ in the arguments of the θ functions in (17)], and the dependence of χ_ω on the splitting Δ can be calculated by using the matrix elements of the projection operators:

$$\langle j|\hat{P}_l|j\rangle = \frac{\Delta_T + j\Delta}{2\Delta_T}, \quad \langle j|\hat{P}_r|j\rangle = \frac{\Delta_T - j\Delta}{2\Delta_T},$$

$$\langle j|\hat{P}_r|j'\rangle = -\langle j|\hat{P}_l|j'\rangle = \frac{T}{\Delta_T}, \quad j \neq j'. \tag{19}$$

We present explicit expressions for the resonant nonlinear susceptibilities at low concentrations ($\varepsilon_F < \Delta_T/2$, when $n_- = n$ and $n_+ = 0$), intermediate concentrations ($\Delta_T/2 < \varepsilon_F < 3\Delta_T/2$), and high concentrations ($\varepsilon_F > 3\Delta_T/2$) when the concentrations of the $\pm$ electrons participating in interband transitions are equal. For ε_F close to $\Delta_T/2$ or $3\Delta_T/2$ the dependences of χ_ω become more complicated, but they are realized only in narrow concentration ranges (or in small ranges of Δ), and we shall not discuss them further.

At high concentrations the peak broadening energies coincide, $\bar{\gamma}_\pm \equiv \bar{\gamma} = (\hbar/2\tau) \times [1 - (T/\Delta_T)^2 - \mu(\Delta/\Delta_T)]$, and the relaxation times are

$$\tau_\pm = \frac{\bar{\tau}}{1 \pm \mu \Delta/\Delta_T}, \tag{20}$$

where the average relaxation time $\bar{\tau}$ and the scattering asymmetry parameter μ are defined by the relations:

$$\bar{\tau}^{-1} = \frac{\tau_l^{-1} + \tau_r^{-1}}{2}, \quad \mu = \frac{\tau_r - \tau_l}{\tau_r + \tau_l}. \tag{21}$$

In this concentration range, $n_\pm = n \pm \rho_{2D}\Delta_T/2$, and (17) and (18) become

$$\chi_\omega = \frac{e^3 Z}{m\omega} \frac{n_+ \tau_+ - n_- \tau_-}{\tilde{\Delta}_T - \hbar\omega - i\bar{\gamma}} \left[i\frac{\hbar}{T\bar{\tau}}\mu\left(\frac{T}{\Delta_T}\right)^3 - \frac{2m}{M}\left(\frac{T}{\Delta_T}\right)^2 \right], \tag{22}$$

so that the susceptibility is due either to the scattering asymmetry ($\mu \neq 0$) or to the effective-mass difference ($m/M \neq 0$). At low concentrations (when $n < \rho_{2D}\Delta_T$) we have

$$\tau_- = \frac{\bar{\tau}}{1 - 2(T/\Delta_T)^2 - \mu\Delta/\Delta_T}, \tag{23}$$

and the resonance width $\bar{\gamma}_-$ and τ_-^{-1} are related by the equality $\bar{\gamma}_- = \hbar/2\tau_-$. The expression for χ_ω becomes

$$\chi_\omega = \frac{e^3 Z}{m\omega} \frac{n\tau_-}{\tilde{\Delta}_T - \hbar\omega - i\bar{\gamma}_-}$$

$$\times \left[i\frac{\hbar}{T\bar{\tau}}\left(\frac{T}{\Delta_T}\right)^3\left(\frac{\Delta}{\Delta_T} - \mu\right) + \frac{2m}{M}\left(\frac{T}{\Delta_T}\right)^2 \right], \tag{24}$$

and a nonzero effect is also present for double quantum wells with symmetric scattering and coinciding effective masses, but with $\Delta \neq 0$. Since the general formula (17) for intermediate concentrations is cumbersome, we give the expression for χ_ω in the case of a double quantum well with $\mu = 0$ and $m_l = m_r$, where the effect is also present if $\Delta \neq 0$:

$$\chi_\omega = -i\frac{e^3 Z}{m\omega} \frac{n_+ \bar{\tau}}{\tilde{\Delta}_T - \hbar\omega - i\bar{\gamma}_+} \frac{\hbar}{2T\bar{\tau}}\left(\frac{T}{\Delta_T}\right)^3 \frac{\Delta}{\Delta_T}. \tag{25}$$

In the region under consideration $\bar{\gamma}_+ = (\hbar/2\tau)[1 - 2(T/\Delta_T)^2]$. Comparing (25) with (22) and (24), we see

that in the approximations adopted χ_ω changes abruptly when ε_F is close to the limits of the range, $\Delta_T/2$ and $3\Delta_T/2$.

To obtain numerical estimates, it is convenient to isolate the size factor

$$\frac{e^3 Z}{m\tilde{\Delta}_T/\hbar} \frac{n\bar{\tau}}{\bar{\gamma}} \tag{26}$$

in the formulas. This factor determines the maximum value of the nonlinear susceptibility after it is multiplied by a dimensionless factor, $\hbar/2T\bar{\tau}$ or m/M, and a numerical factor (not exceeding 0.1 for $\Delta \leq 2T$ and rapidly decreasing with increasing Δ). For GaAs/AlGaAs with a double-quantum-well structure and the standard parameters equal to $Z = 100$ Å, $n = 2 \times 10^{11}$ cm^{-2}, $\bar{\gamma} = 1$ meV, $\Delta_T = 15$ meV, and $\bar{\tau} = 10^{-11}$ s (which correspond to an electron mobility of 2.5×10^5 cm^2/(V·s) and resonant pumping by radiation with a wavelength of 100 μm), the maximum value of χ_ω is found to be of order $(1-5) \times 10^{-7}$ esu. This is in good agreement with the resonant value of the second-order nonlinear susceptibility found by Sherwin *et al.*[9] in a double-quantum-well structure from second-harmonic generation measurements with similar values for the parameters. However, in our case the χ_ω vs Δ and χ_ω vs ω curves differ significantly from those measured by Sherwin *et al.*[9] (although the spectra are Lorentzian in both cases). It is convenient to study the dependence of χ_ω on bias voltage (which determines Δ) at a fixed pump frequency; below we limit our discussion to the cases of low, intermediate, and high concentrations.

For double quantum wells with thin barriers, the condition $\Delta_T/2 > \varepsilon_F$ is also met at tunneling resonance, $\Delta = 0$, so that Eq. (24) can be used for all values of Δ. For a 30 Å Ga$_{0.3}$Al$_{0.7}$As barrier (this value of the barrier width, like the values of the other parameters, is close to the one used in Ref. 9) we find that $2T \simeq 15$ meV, and the low-concentration regime is observed when $n \leq 4.2 \times 10^{11}$ cm^{-2}. The values of the dimensionless factors $\hbar/2T\bar{\tau}$ and $2m/M$ prove to be close to 5×10^{-3} and 2×10^{-2}, respectively (due to the nonparabolicity of the spectrum, the effective-mass difference $|m_l - m_r|/m$ is estimated here, according to Ref. 10, as Δ_T/ε_g, where ε_g is the band gap), so that the contributions $\chi_\omega^{(1)}$ and $\chi_\omega^{(2)}$ to the nonlinear susceptibility are of the same order. As a result, Eq. (24) can be rewritten as

$$\chi_\omega = \bar{\chi}\left[\frac{\hbar}{2T\bar{\tau}}f_\mu(\Delta) + \frac{2m}{M}g(\Delta) \right], \tag{27}$$

the real and imaginary parts of the functions $f_\mu(\Delta)$ and $g(\Delta)$ being depicted in Fig. 2 with the above values of the double-quantum-well parameters. Here we used the phenomenological value $\bar{\gamma} \simeq 1$ meV, which agrees with the experimental data of Refs. 5 and 9, but is significantly larger than the value obtained from the relation between $\bar{\gamma}_-$ and τ_- (probably, the inhomogeneous broadening of the resonant transition, which contributes nothing to mobility, is the main mechanism of spectrum broadening). Figures 2a and 2b show that the nature of the Δ dependence changes significantly because of the asymmetry in scattering: the function $f_0(\Delta)$ is odd, while Re $f_1(\Delta)$ has two resonance peaks for
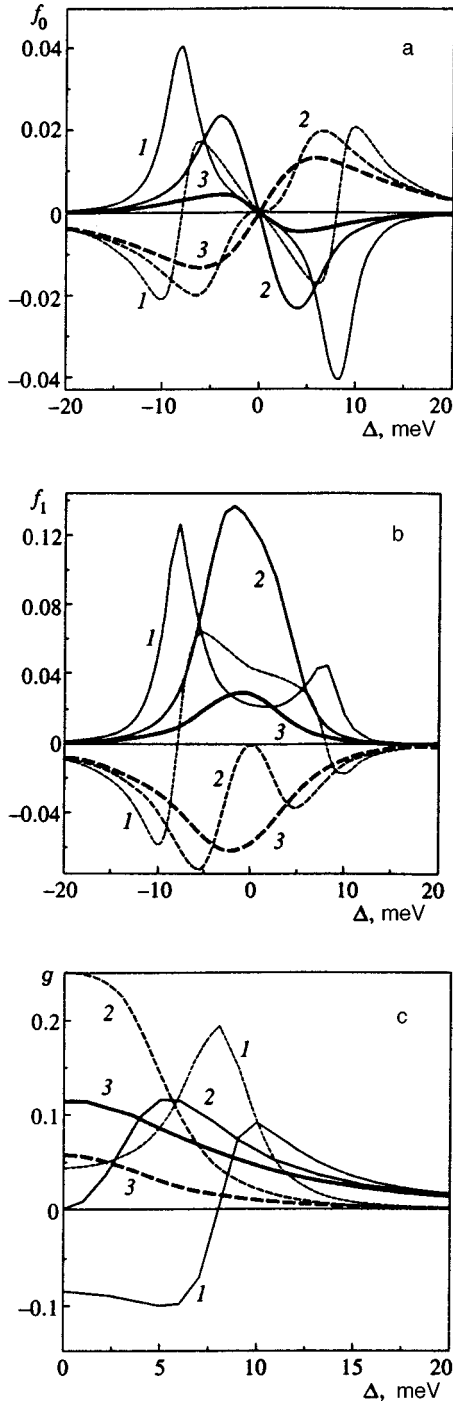
FIG. 2. Real and imaginary parts (solid and dashed curves, respectively) of the function $f_\mu(\Delta)$ with (a) $\mu=0$ and (b) $\mu=1$ and the g vs Δ (c) at pump energies $\hbar\omega$ equal to 17 meV (curve *1*), 15 meV (curve *2*), and 13 meV (curve *3*).
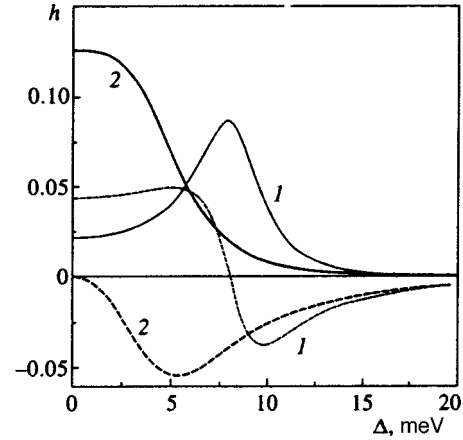

FIG. 3. Real and imaginary parts (solid and dashed curves, respectively) of the function $h(\Delta)$ at pump energies $\hbar\omega$ equal to 17 meV (curve *1*) and 15 meV (curve *2*).

$\hbar\omega>2T$. The function $g(\Delta)$ in Fig. 2c is even, and resonance peaks also appear on Re $g(\Delta)$ for $\hbar\omega>2T$. When $\Delta>2T$, the nonlinear response rapidly decreases in all cases due to the suppression of tunnel coupling.

In the concentration range $(4.2-8.4)\times10^{11}$ cm^{-2} (for double quantum wells with the above parameter values at $\Delta=0$; as Δ increases, this range moves toward higher values of n), the intermediate case described by (25) comes into play. Here the value of $\bar\chi$ increases and so does the value of

the dimensionless factor $\hbar/2T\bar\tau$ (since the concentration increases and mobility decreases). Figure 3 shows the h vs Δ curves determined by the formula $\chi_\omega/\bar\chi=(\hbar/2T\bar\tau)h(\Delta)$ for the same values of the double-quantum-well parameters and the same pump energies as in Fig. 2. Here, as in the low-concentration case, we have a peak on Re $h(\Delta)$ when $\hbar\omega>2T$, while the Im h vs Δ curve passes through zero in the vicinity of the peak [if $\mu=0$, the function $h(\Delta)$ is even]. When $\hbar\omega<2T$, the peak on Re $h(\Delta)$ shifts to $\Delta=0$, while for larger values of Δ the effect is suppressed.

The high-concentration case described by (22) is realized in double quantum wells with wider barriers, where the resonant nonlinear-response regime can be realized under the conditions of strong tunnel coupling, $\Delta\leqslant2T$, in a spectral range with a wavelength of about 300 μm. For a 45 Å Al$_{0.3}$Ga$_{0.7}$As barrier we obtain $2T=2.4$ meV, and the condition $n>3T/\rho_{2D}$ is met when the concentration is higher than 2.7×10^{11} cm^{-2} (at $\Delta=0$; the concentration increases with Δ). Let us rewrite (22) as (27): $\chi_\omega/\bar\chi=(\hbar/2T\bar\tau)v(\Delta)+(2m/M)w(\Delta)$. Figure 4 depicts the functions $v(\Delta)$ and $w(\Delta)$ for an electron concentration of 6×10^{11} cm^{-2} (the other double-quantum-well parameters coincide with those used in Figs. 2 and 3). Since the functions $v(\Delta)$ and $w(\Delta)$ rapidly decrease in weakly asymmetric scattering ($|\mu|\ll1$), the v vs Δ and w vs Δ curves are shown for the case of scattering in the l quantum wells ($\mu=1$). Due to the smallness of the fraction $(n_- - n_+)/n$, the curves representing $v(\Delta)$ and $w(\Delta)$ are nearly asymmetric. The behavior of these curves is similar to that of the curves in Figs. 2 and 3, but because of a decrease in splitting between the levels, the resonance effect is rapidly suppressed as Δ increases. However, for highly asymmetric scattering, one of the relaxation times $\tau_\pm$ in (20) diverges when $\Delta\gg2T$, and the function $w(\Delta)$ approaches a finite value.

4. CONCLUSION

We have examined the conversion of a transverse modulating voltage in the terahertz range into a longitudinal current due to electron transitions between tunnel-coupled levels
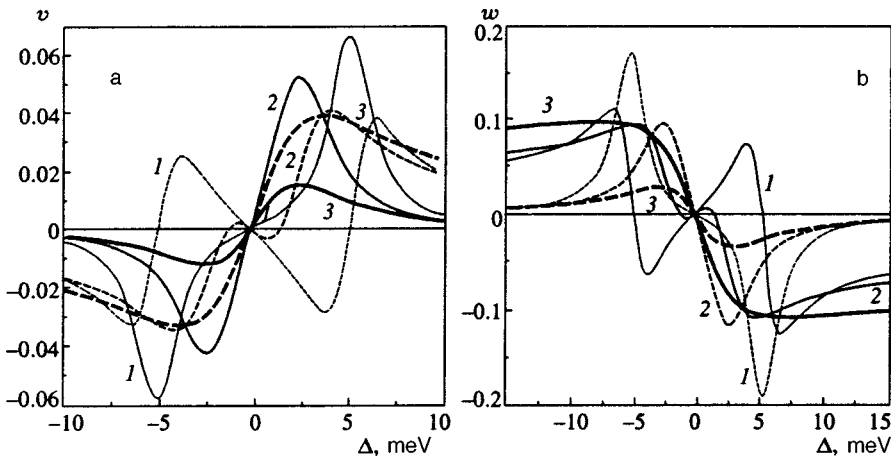
FIG. 4. The real and imaginary parts (solid and dashed curves, respectively) of the functions $v(\Delta)$ (a) and $w(\Delta)$ (b) at pump energies $\hbar\omega$ equal to 7 meV (curve *1*), 5 meV (curve *2*), and 3 meV (curve *3*).

in double quantum wells with a constant current. We have found that large values of the nonlinear susceptibility defined in (1) are realizable in an asymmetric double quantum well, where the asymmetry is due to a difference in effective mass or asymmetry in the scattering in the *l* and *r* quantum wells and/or asymmetry of the energy spectrum. Since the induced high-frequency response is proportional to the constant current, gain can be achieved at high currents. To estimate the effectiveness of this gain mechanism, we need to know not only the characteristics of the nonlinear response, but also the geometry of the system (in the same way as the gain of an ordinary field-effect transistor depends on the circuit parameters).

Let us discuss the approximations used in our calculation scheme. What is left out when double quantum wells are described as a pair of tunnel-coupled levels is the contribution of the excited states of the *l* and *r* quantum wells and the self-consistent transverse electric field; in our case Δ and T were only approximately estimated for the model of rectangular quantum wells and a barrier. Values of these parameters that are more exact can be found in papers on longitudinal transport[1-4] or on interlevel transitions with pumping in the terahertz range.[5,9] The value of m/M can be refined in multiband calculations of double quantum wells, and the degree of scattering asymmetry μ (which depends on the 2D momentum for nonpointlike scatterers) can be considered a phenomenological parameter, whose value can be estimated using resistance-resonance data.[1-3] Moreover, allowing for the depolarization shift of an interlevel transition can alter $\widetilde{\Delta}_T$ considerably already at a concentration of 2×10^{11} cm^{-2} (Ref. 9). These parameters can be refined by comparing them with the experimental data at hand. The approximation of a weakly anisotropic distribution of 2D electrons is also valid for heating electric fields F, so that the high-current regime is described by the above formulas. Thus, the approximations adopted do not alter the nonlinear-response mechanism, the magnitude of this response, or the nature of the above relations; therefore, further refinements are needed only for more detailed calculations.

The mechanism of longitudinal-transverse conversion and radiation amplification by an electron current in the terahertz range in double quantum wells is related to the applications of intersubband transitions to the detection[11] and generation[12-15] of radiation in the submillimeter range which are being actively studied. Although the mechanism of monopolar lasing[16] has been realized in several ways,[17] such devices operate only in the mid-IR range. Hence the gain mechanism considered in this paper may be of interest in developing a 2D electron source of radiation in the terahertz range. Moreover, the possibility of converting transverse pump radiation into a longitudinal response has never before been studied in 2D systems; therefore, the experimental realization of such a scheme, as well as the study of other mechanisms of such conversion for different heterostructures and transition types, could be of great interest.

*)E-mail: zinovi@lab2.Kiev.ua

[1] A. Palevski, F. Beltram, F. Capasso, L. N. Pfeiffer, and K. W. West, Phys. Rev. Lett. **65**, 1929 (1990).

[2] N. K. Patel, A. Kurobe, I. M. Castelton, E. H. Linfield, K. M. Brown, M. P. Grimshaw, D. A. Ritchie, G. A. C. Jones, and M. Pepper, Semicond. Sci. Technol. **11**, 703 (1996).

[3] Y. Ohno, M. Tsuchiya, and H. Sakaki, Appl. Phys. Lett. **62**, 1952 (1993).

[4] F. T. Vasko, Phys. Rev. B **47**, 2410 (1993); F. T. Vas'ko and O. É. Raĭchev, Zh. Éksp. Teor. Fiz. **107**, 951 (1995) [JETP **80**, 539 (1995)].

[5] J. N. Heyman, K. Unterrainer, K. Craig, B. Galdrikian, M. S. Sherwin, K. Campman, P. F. Hopkins, and A. C. Gossard, Phys. Rev. Lett. **74**, 2682 (1995).

[6] F. T. Vas'ko, Fiz. Tekh. Poluprovodn. **26**, 825 (1992) [Sov. Phys. Semicond. **26**, 464 (1992)]; F. T. Vas'ko and O. É. Raĭchev, Zh. Éksp. Teor. Fiz. **104**, 3103 (1993) [JETP **77**, 452 (1993)].

[7] F. T. Vas'ko and O. É. Raĭchev, Zh. Éksp. Teor. Fiz. **108**, 2103 (1995) [JETP **81**, 1146 (1995)].

[8] L. Zheng and A. H. MacDonald, Phys. Rev. B **47**, 10 619 (1993); O. E. Raichev and F. T. Vasko, Phys. Rev. B **55**, 2321 (1997).

[9] M. S. Sherwin, K. Craig, B. Galdrikian, J. Heyman, A. Markelz, K. Campman, S. Fafard, P. F. Hopkins, and A. Gossard, Physica D **83**, 229 (1995); J. N. Heyman, K. Craig, B. Galdrikian, M. S. Sherwin, K. Campman, P. F. Hopkins, S. Fafard, and A. C. Gossard, Phys. Rev. Lett. **72**, 2183 (1994).

[10] I. M. Tsidil'kovskiĭ, *Band Structure of Semiconductors* [in Russian], Nauka, Moscow (1978).

[11] *Intersubband Transitions in Quantum Wells*, E. Rosencher, B. Vinter, and B. Levine (eds), Plenum Press, New York (1992); *Quantum Well Intersubband Transitions: Physics and Devices*, H. C. Liu, B. F. Levine, and J. Y. Anderson, Kluwer Academic, Dordrecht (1994).

[12] S. I. Borenstain and J. Katz, Appl. Phys. Lett. **55**, 654 (1989).

[13] A. Kastalsky, V. J. Goldman, and J. H. Abeles, Appl. Phys. Lett. **59**, 2636 (1991).

[14] V. Berger, Semicond. Sci. Technol. **9**, 1453 (1994).

[15] F. T. Vasko and Yu. N. Soldatenko, Appl. Phys. Lett. **66**, 544 (1995).

[16] P. F. Kazarinov and R. A. Suris, Fiz. Tekh. Poluprovodn. **5**, 797 (1971) [Sov. Phys. Semicond. **5**, 707 (1971)].

[17] J. Faist, F. Capasso, D. L. Sivko, C. Sirtori, A. L. Hutchinson, and A. Y. Cho, Science **264**, 553 (1994); J. Faist, F. Capasso, C. Sirtori, D. L. Sivko, A. L. Hutchinson, and A. Y. Cho, Nature (London) **387**, 777 (1997); G. Scamarcio, F. Capasso, C. Sirtori, J. Faist, A. L. Hutchinson, D. L. Sivko, and A. Y. Cho, Science **276**, 773 (1997).

Microscopic calculations of ferroelectric instability in perovskite crystals

O. V. Ivanov,*) D. A. Shport, and E. G. Maksimov

P. N. Lebedev Physical Institute, Russian Academy of Sciences, 117924 Moscow, Russia
(Submitted 22 October 1997)
Zh. Éksp. Teor. Fiz. **114**, 333–358 (July 1998)

First-principles calculations are performed relating to the stability of a series of perovskite crystals with respect to transition to the ferroelectric and the antiferroelectric state. The calculations employ the generalized Gordon–Kim method, in which the total charge density of an ionic crystal is represented as a superposition of the densities of the individual ions. In the spirit of the nonequilibrium thermodynamics of Leontovich the charge density of an individual ion is calculated in the presence of external auxiliary fields which deform this density. Multipole deformations up to quadrupole are taken into account. The actual magnitude of the deformation is found by minimizing the total energy of the crystal in the Thomas–Fermi–Dirac approximation. The calculated values of the ion shifts in the ferroelectric phase for $BaTiO_3$, and also the electron contribution to the dielectric constant ε_∞ and the dynamic Born effective charges Z^{eff} are found to be in good agreement with the experimental data. The proposed method allows one to obtain an analytical expression for ε_∞, Z^{eff}, and the dynamic vibration matrix. It is shown that these expressions formally coincide with the expressions arising in the phenomenological models of the polarized and deformed ion. Analysis of the expressions obtained confirms the validity of the classical theory of ferroelectrics of displacement type for perovskite crystals. © *1998 American Institute of Physics.* [S1063-7761(98)02307-5]

1. INTRODUCTION

Perovskite crystals with the general formula ABO_3 at high temperatures as a rule possess cubic structure. Atoms of the metal A occupy sites at the vertices of a cube, while atoms of the metal B are found at the center of the cube. Atoms of the nonmetal O are found at the face centers of the cube and form an octahedron surrounding the B atom. Of greatest interest from the point of view of studying lattice instabilities are oxygen-containing perovskites, where the O atom actually is oxygen. A significant number of these compounds are ferroelectrics.[1] They often exhibit a sequence of ferroelectric phases, transforming from the cubic phase to the tetragonal, then to orthorhombic, and finally, to rhombohedral. Roughly speaking, this sequence of phases can be considered as shifts of the central B ion in the (100) direction in the tetragonal phase, in the (110) direction in the orthorhombic phase, and in the (111) direction in the rhombohedral phase. The actual ion shifts form a much more complicated picture, involving elastic lattice deformations in a number of cases.

Besides ferroelectric phase transitions, perovskite crystals reveal the presence of a large number of other lattice instabilities and structural transformations. Some of them are associated with different kinds of rotations of the oxygen octahedron about one of the axes of the cube. The most thoroughly studied transition of this type occurs in the compound $SrTiO_3$ at $T = 150$ K (Ref. 1). This transition is often called antiferroelectric in analogy with magnetic transformations. It may be noted that this analogy is quite arbitrary. Thus, the ferroelectric transition, apart from the fact that it is a structural lattice transformation, entails a nontrivial electrodynamic consequence. In this instance, a spontaneous electrical polarization arises in the crystal even in the absence of an external electric field and, correspondingly, the dielectric constant diverges at the transition point. In this sense the transition in $SrTiO_3$ at $T = 105$ K does not differ in any important way from the vast majority of other structural transitions. There are other examples of structural transformations in perovskites. These include, for example, the transition at $T = 400$ K in $BaBiO_3$, in which successive compression and dilation of the oxygen octahedra takes place about the central Bi atom along the (111) axis.[2]

The determination of the nature of the structural transformations in perovskites and, in particular, the ferroelectric transition has been one of the most pressing problems of solid state theory for over 50 years[1,3] ever since the discovery of the ferroelectric transition in $BaTiO_3$. A completely clear answer to this question is still lacking even though, in our opinion, the basic, underlying physical ideas about the nature about the ferroelectric instability in perovskite crystals can be regarded as well established. Even in the early works of Skanavi[4] and Slater[5] attention was directed at the important role played in the ferroelectric instability in perovskite crystals by local field and electron polarizability effects. The works of Ginzburg,[6] Anderson,[7] and Cochran[8] established a connection between the ferroelectric transition and the dynamic lattice and introduced the concept of a soft mode, according to which the ferroelectric instability is due to vanishing at the transition point of one of the optically active transverse modes at the center of the Brillouin zone. Since this mode possesses a dipole moment and is accompanied by an electric polarization wave, spontaneous polarization arises

in the crystal at the phase transition point when this mode ''freezes.''

Using the polarizable ion model, Cochran showed[8] for binary ionic crystals that the frequency of the optically active phonon mode at $\mathbf{q}=0$ in the harmonic approximation can be written as

$$\omega_{\mathrm{TO}}^2 = \omega_{sr}^2 - \omega_{dd}^2, \tag{1}$$

where ω_{dd}^2 is the contribution of the long-range dipole–dipole interaction:

$$\omega_{dd}^2 = \frac{4\pi e^2 \widetilde{Z}^2}{\mu v_0 (\varepsilon_\infty + 2)}. \tag{2}$$

Here ε_∞ is the electron contribution to the dielectric constant of the ionic crystal, and $\widetilde{Z}$ is the dynamic Born effective charge, μ is the reduced mass, and v_0 is the volume of the unit cell. The total polarizability of the crystal under the corresponding static ion shift $\mathbf{u}_n$ is expressed in terms of this charge as

$$\mathbf{P} = \frac{e}{v} \sum_n \widetilde{Z}_n \mathbf{u}_n. \tag{3}$$

The quantity $\omega_{sr}^2 > 0$ is the repulsive contribution of the close-range forces to the frequency of the TO mode. Kvyatkovskiĭ[9] (see also the review in Ref. 10 for more details) proved the validity of expressions (1) and (2) in a general model-independent approach for binary crystals within the framework of a rigorous microscopic theory of lattice dynamics based on a description of the electron contribution employing the dielectric matrix $\varepsilon(\mathbf{q}+\mathbf{K}, \mathbf{q}+\mathbf{K}', 0)$, where $\mathbf{K}$ and $\mathbf{K}'$ are inverse lattice vectors. Later he proved[11] that these formulas remain valid for more complicated crystals, including perovskites. In this case $\widetilde{Z}$ is the Born charge of the corresponding soft mode.

In ordinary ionic insulators the Born charges are close to the values of the nominal ion charges or, more accurately, the Szigetti effective charges[12] defining the ionicity of the crystal. In ferroelectrics, however, the dynamic Born charges, as a rule, have substantially larger values. Thus, for example, in ferroelectric IV–VI compounds the effective ionic charges satisfy $Z_s \lesssim 1$, while the Born charges reach values $Z \gtrsim 10$ (Ref. 10).

Thus, expressions (1) and (2) reduce the problem of establishing the nature of the ferroelectric instability to a search for answers to the following questions. Why are the Born charges in ferroelectrics substantially larger than the Szigetti ionic charges? What is the situation with the short-range repulsion forces? What in the series of analogous compounds, e.g., in $BaTiO_3$ and $SrTiO_3$, determines whether a ferroelectric transition is present: the corresponding change in ω_{sr}^2 or ω_{dd}^2? Unfortunately, until recently there have been no direct calculations without the use of adjustable parameters which would provide answers to these questions. The problem is further complicated by the fact that the values of ω_{sr}^2 and ω_{dd}^2 in ferroelectrics are almost an order of magnitude larger than the resulting value of ω_{TO}^2.

First-principles calculations of ferroelectric instability in perovskite crystals have appeared recently[13–15] based on

band calculations of the electronic spectrum and its contribution to the total crystal energy within the framework of the density functional theory (DFT). These calculations have confirmed that it is possible to achieve the accuracy needed for a valid description of the energetics of the structural distortions in perovskite crystals. In particular, the above assertion that ω_{dd}^2 is almost an order of magnitude larger than ω_{TO}^2 follows from the values of the dynamic Born charges and the quantities ω_{TO}^2 obtained in Refs. 16 and 17. Unfortunately, it is difficult to obtain unambiguous answers to the above questions from these works. In particular, from these results it is even difficult to say why the Born charges in perovskites at the B ion and one of the oxygen atoms exceed the nominal charges of these ions while the Born charges on the A ion and the other oxygen ion are very close in most cases to their nominal values. The explanations given in Refs. 16 and 17, tied up with hybridization effects of electron d states of the B ion and p states of the oxygen ion, are unconvincing and, on top of this, the connection between these explanations and the results of phenomenological calculations of these charges in the earlier work of Axe[18] remains unclear. This earlier work was based on a treatment of the model of polarizable ions and did not include any hybridization effects, but nevertheless led to practically the same values of the Born charges. Recently, very successful calculations[19] of some other instabilities in perovskites have appeared, including the so-called antiferroelectric instability in $SrTiO_3$.

In the present paper we carry out simple first-principles calculations based on a generalization of the Gordon–Kim model developed earlier by two of us.[20–22] The aim of these calculations is to facilitate a microscopic study of the energetics of perovskite crystals and the nature of different kinds of lattice distortions, including the ferroelectric phase. Earlier this method was successfully applied to calculate the electron polarizability and lattice dynamics of a large number of ionic crystals.[20–23] An even simpler variant of this method was employed with success to examine the nature of the structural distortions in $BaBiO_3$.

The organization of this paper is as follows. Section 2 presents a brief exposition of the new version of our generalized Gordon–Kim (GGK) approach, which is needed to maintain the required accuracy of the calculations in perovskite crystals. Section 3 present results of numerical calculations of the energetics of a series of perovskite crystals, taking account of distortions of their cubic structure, and also contains a discussion of these results. In the last section we briefly discuss the outlook for future possible studies together with refinements of the proposed method needed toward this end.

2. GENERALIZED GORDON–KIM METHOD

The essence of the Gordon–Kim method[25,26] for calculating the static and dynamic properties of ionic crystals consists in the following. The crystal is treated as a set of isolated ions and its total charge density is written in the form

$$\rho(\mathbf{r}) = \sum_i \rho_i(\mathbf{r} - \mathbf{R}_i), \tag{4}$$

where the sum is taken over all the ions in the crystal. Next, in the spirit of the density functional theory (DFT) the total crystal energy is written as a functional of the charge density

$$E^{cr} = \left[E\left\{ \sum_i \rho_i(\mathbf{r} - \mathbf{R}_i) \right\} - \sum_i E\{\rho_i(\mathbf{r} - \mathbf{R}_i)\} \right] + E^N$$

$$+ \sum_i E\{\rho_i(\mathbf{r} - \mathbf{R}_i)\}. \tag{5}$$

In this formula the sum of energies of the individual ions is added and subtracted in accordance with the Gordon–Kim method. Here E^N is the interaction energy of the nuclei, which has the form

$$E^N = \frac{1}{2} \sum_{i,j} \frac{Z_i^N Z_j^N}{|\mathbf{R}_i - \mathbf{R}_j|}, \tag{6}$$

and Z_i^N is the nuclear charge of the ith ion. The quantity $E\{\rho(\mathbf{r})\}$ is the charge density functional, which according to Hohenberg and Kohn[28] can be represented in the form

$$E\{\rho\} = \int d\mathbf{r}\rho(\mathbf{r})V^{ext}(\mathbf{r}) + \frac{1}{2} \int d\mathbf{r}\, d\mathbf{r}'\, \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}$$

$$+ \int d\mathbf{r}\, F\{\rho(\mathbf{r})\}. \tag{7}$$

Here $V^{ext}(\mathbf{r})$ is the external potential, which in the case of crystals is simply the Coulomb potential of the nuclei:

$$V^{ext}(\mathbf{r}) = \sum_i \frac{Z_i^N}{|\mathbf{r} - \mathbf{R}_i|}. \tag{8}$$

The quantity $F\{\rho(\mathbf{r})\}$ is a universal functional describing the contributions of the kinetic energy and the exchange–correlation energy to the total electron energy. In their original papers,[25,26] Gordon and Kim used a significant number of additional simplifications in their analysis of expression (5) for the crystal energy. First, for the functional $F\{\rho(\mathbf{r})\}$ they used a simple local approximation in the spirit of Thomas–Fermi. Second, the first term in expression (5) neglects contributions from the overlap of more than two ions at a time. As a result, the problem reduces to a calculation of pairwise interactions V_{ij} between the ions,

$$V_{ij} = \int [E\{\rho_i(\mathbf{r} - \mathbf{R}_i) + \rho_j(\mathbf{r} - \mathbf{R}_j)\} - E\{\rho_i(\mathbf{r} - \mathbf{R}_i)\}$$

$$- E\{\rho_j(\mathbf{r} - \mathbf{R}_j)\}]d\mathbf{r} \tag{9}$$

and to subsequent calculation of the total crystal energy with the corresponding pairwise interaction taken into account. The charge density of an individual ion $\rho_i(\mathbf{r})$ was taken to be spherically symmetric and was calculated by the Hartree–Fock method for free ions.

In its original form the Gordon–Kim method has a quite limited region of applicability. This has to do, mainly, with the use of the spherically symmetric free ion density for $\rho_i(\mathbf{r})$. Their approach can be considered as a microscopic realization of the rigid ion model. In real crystals the charge density distribution of an ion is obviously not rigid. As follows from the experimental data,[27] even in the case of the simplest ionic crystals of the NaCl type, where the charge

density can be taken with good accuracy to be spherically symmetric, noticeable compression of the charge density of the negative ion is observed. In general, static and dynamic distortions of the ion charge density are possible, possessing multipole symmetry of any order. In particular, an account of dipole distortions is absolutely necessary to describe the electron polarization of an ion.

Following the early works of Gordon and Kim, a large number of efforts were undertaken to improve the given method (see the review in Ref. 29). In particular, significant attention was given to monopole, i.e., spherically symmetric, distortions of the charge density. With this aim, it was proposed to calculate not the charge density of a free ion, but that of an ion located in a charged Watson sphere.[30] The charge of the sphere was taken to be equal to the charge of the corresponding ion, but of opposite charge. To determine the radius of the sphere, a number of variational procedures have been proposed. Such a modification of the Gordon–Kim method is necessary for the following reasons. First, a spherically symmetric distortion of the ion charge density is observed experimentally, and use of the Watson sphere indeed leads to spherically symmetric compression of the charge density of the negative ion. Second, use of the Watson sphere makes it possible to describe crystals containing ions that do not exist in the free state, such as the O^{2-} ion. Efforts were also undertaken to allow for dipole distortions of the charge density, but all of them were based on the use of phenomenological ion models of a different kind, of the shell-model type (for more details, see the review in Ref. 29).

In these efforts to improve the Kim–Gordon model, another problem arose which was not consistently resolved, in our opinion, in previous treatments. This is the so-called ion self-energy problem. The point here is that in the use in the Gordon–Kim model of ion charge densities calculated using the Watson sphere, the energy of each ion begins to depend on the radius of the Watson sphere. This radius, in turn, depends on the properties of the entire crystal as a whole; therefore the sum of energies of the individual ions should be included in the expression for the total energy and be taken into account in the course of the minimization defining this radius. On the question of the expression for this energy, especially for unstable ions, the opinions of different authors differ widely.

Earlier we developed a method[20,21] which makes it possible within the framework of the pairwise interaction approximation to allow for distortions of the ion charge density of dipole type and introduce a self-consistent definition of the self-energy of the ion. Using this approach, we successfully calculated the electron polarization and lattice dynamics of many binary ionic crystals. The accuracy of this method turned out, however, to be insufficient for calculating the ferroelectric instability in perovskite crystals. This is due both to the necessity in perovskites of taking account of changes in the charge density due to higher multipoles and to the necessity of departing from the framework of pairwise interactions. We recently developed a corresponding generalization of the Gordon–Kim method; a brief exposition of it is given in Ref. 22.

The essence of our GGK method consists in the following. First, in line with ideas of nonequilibrium thermodynamics developed by Leontovich,[30] the nonequilibrium state of an individual ion is prepared, with charge density distribution possessing any multipole symmetry. Toward this end, we solve the quantum-mechanical problem for the individual ion in the presence of external auxiliary fields possessing the corresponding necessary multipole symmetry and calculate the charge density distribution of the given ion and also its self-energy. After this, we place the ion in the crystal and using formula (5) for the crystal energy we determine the actual deformation of the ion from the minimum of this functional. The minimization is carried out for different values of the volume of the unit cell and different crystalline structures. The state corresponding to the deepest minimum is the true ground state of the crystal. This approach also makes it possible with the help of the frozen phonon method to find the frequencies of phonon vibrations for a series of wave vectors. With this aim, we calculate the energy difference between an ideal crystalline lattice and a lattice in which the ions are displaced just as is actually the case in a phonon mode.

To solve the quantum-mechanical problem of one ion, we use the density functional theory. The corresponding Kohn–Sham equation[31] in the presence of auxiliary fields has the form

$$\left(-\frac{\nabla^2}{2} + \frac{Z_i^N}{|\mathbf{r}|} + \int d\mathbf{r} \frac{\rho_i(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + V^{xc}(\mathbf{r}) \right.$$
$$\left. + V^{ext}(\mathbf{r},K_n) \right) \psi_\alpha(\mathbf{r}) = \varepsilon_\alpha \psi_\alpha(\mathbf{r}) \qquad (10)$$

and

$$\rho_i(\mathbf{r}) = \sum_\alpha \psi_\alpha^*(\mathbf{r}) \psi_\alpha(\mathbf{r}), \qquad (11)$$

where the sum is taken over all filled states. Here $V^{xc}(\mathbf{r})$ is the exchange–correlation potential, defined as

$$V^{xc}(\mathbf{r}) = \frac{\delta E^{xc}\{\rho_i(\mathbf{r})\}}{\delta \rho_i(\mathbf{r})}. \qquad (12)$$

For the exchange–correlation energy E^{xc} we use, as usual, the local approximation. The quantity $V^{ext}(\mathbf{r},K_n)$ is the potential of the auxiliary fields, which possesses multipole symmetry. At this point we should mention the substantial difference in the calculations when the auxiliary potential has spherical symmetry, i.e., is a monopole, and when the auxiliary potentials have any other multipole symmetry. For a monopole external field we are forced to solve Eq. (10) exactly in the presence of this potential. This is due to the necessity, mentioned above, of calculating ions not existing in the free state. This means that for them a solution of Eq. (10) that does not take the auxiliary monopole potential into account simply does not exist. Moreover, in the local approximation for the exchange–correlation energy the solution of Eq. (11) absent an auxiliary monopole potential is lacking even for the majority of negative ions.

In the present paper we use an auxiliary monopole potential in the form of a smoothed Watson sphere:

$$V_{ext}^0 = (\mathbf{r},K_0) = \begin{cases} K_0 + br^2, & r < R_w, \\ -Z_i^{ion}/r, & r > R_w, \end{cases} \qquad (13)$$

where Z_i^{ion} is the nominal charge of the ion. The parameter b and the radius R_w of the Watson sphere are chosen from the condition of smooth joining of the potential and its first derivative at $r = R_w$. The choice of a more smoothed potential than the initial Watson sphere potential is motivated mainly by the need to refine the procedure of the numerical calculations. The quantity K_0, characterizing the depth of the potential at the nucleus, is a variational parameter and is determined from the minimum of the total energy of the crystal. The quantity $V^{ext}(\mathbf{r},K_l)$ for the higher multipoles is chosen in the following form, also smoothed:

$$V^{ext}(\mathbf{r},K_l) = -r^l K_l P_l(\cos \theta) e^{-\beta_l r}. \qquad (14)$$

Here the Z axis is chosen in the direction of the external field, $P_l(\cos \theta)$ are the ordinary Legendre polynomials, and K_l and β_l are the corresponding variational parameters. It is important here that there is no need to solve Eq. (10) exactly in the presence of multiple potentials. It was sufficient to find the changes in the charge density distribution and the ion energy from perturbation theory in the potential $V^{ext}(\mathbf{r},K_l)$. The point here is that the amplitudes of the multipole distortions in the charge density distribution are determined, in the final analysis, by the crystal fields or their changes in response to changes in the crystalline structure, but these fields are always much smaller than the intra-atomic fields.

For the calculations we have used perturbation theory in the form suggested by Sternheimer. As is customary in perturbation theory, the wave function $\psi_\alpha(\mathbf{r})$ is represented in the form

$$\psi_\alpha(\mathbf{r}) = \psi_\alpha^0(\mathbf{r}) + \delta\psi_\alpha(\mathbf{r}) \qquad (15)$$

and

$$\delta\rho(\mathbf{r}) = \sum_\alpha \psi_\alpha^0(\mathbf{r}) \delta\psi_\alpha(\mathbf{r}). \qquad (16)$$

The unperturbed function $\psi_\alpha^0(\mathbf{r})$ satisfies the equation

$$\left(-\frac{\nabla^2}{2} + V_{eff}(\mathbf{r}) \right) \psi_\alpha^0(\mathbf{r}) = \varepsilon_\alpha^0 \psi_\alpha^0(\mathbf{r}), \qquad (17)$$

where

$$V_{eff}(\mathbf{r}) = \frac{Z_i^N}{|\mathbf{r}|} + \int d\mathbf{r}' \frac{\rho_i^0(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + V^{xc}(\mathbf{r}) + V_{ext}^0(\mathbf{r},K_0), \qquad (18)$$

and the variation of the wave function $\delta\psi_\alpha(\mathbf{r})$ is defined by the equation

$$\left(-\frac{\nabla^2}{2} + V_{eff}(\mathbf{r}) - \varepsilon_\alpha \right) \delta\psi_\alpha(\mathbf{r})$$
$$= -\left[V_{ext}(\mathbf{r},K_l) + \frac{\delta V_{eff}}{\delta\rho(\mathbf{r})} \delta\rho(\mathbf{r}) \right] \psi_0(\mathbf{r}). \qquad (19)$$

In the Sternheimer method both equations for $\psi_\alpha^0(\mathbf{r})$ and $\delta\psi_\alpha(\mathbf{r})$ are solved simultaneously with allowance for the self-consistency procedures for $V_{\text{eff}}(\mathbf{r},\rho)$ and $\delta V_{\text{eff}}(\mathbf{r},\rho)$ $=(\delta V_{\text{eff}}(\mathbf{r})/\delta\rho)\delta\rho$. We used a numerical method to solve these equations numerically. The wave functions were represented as series in Chebyshev polynomials which match smoothly at large distances r with the corresponding asymptotic limit for the functions $\psi_\alpha^0(\mathbf{r})$ and $\delta\psi_\alpha(\mathbf{r})$.

Solving Eqs. (17) and (19), we obtain the variation of the wave function, $\delta\psi_\alpha(\mathbf{r})$, which depends implicitly on the parameters K_0 and β_l and depends linearly on the parameter K_l.

The variation of the charge density $\delta\rho^l(\mathbf{r},K_0,\beta_l)$ can be represented in the form

$$\delta\rho^l(\mathbf{r})=K_l\delta\widetilde{\rho}(\mathbf{r},K_0,\beta_l)P_l(\cos\,\theta). \quad (20)$$

According to standard electrostatics it is possible to determine the values of the multipole moments corresponding to the charge distribution (20)

$$P_l=\int d\mathbf{r}\ r^l P_l(\cos\,\theta)\delta\rho^l(\mathbf{r},K_0,\beta_l). \quad (21)$$

If we take (20) into account, Eq. (21) gives a linear relationship between the parameters K_l and P_l, which allows us to express the ion energy in terms of the multipole moments and perform an additional variation of the total crystal energy not in the parameters of the auxiliary potential but in the multipole moments of the ions.

The expression for the total energy of an ion in DFT in the presence of the auxiliary fields has the form

$$E^{\text{ion}}\{\rho(\mathbf{r})\}=\int d\mathbf{r}\ \frac{Z\rho(\mathbf{r})}{|\mathbf{r}|}+\frac{1}{2}\int d\mathbf{r}'\ \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}$$
$$+\sum_l\int d\mathbf{r}\ V_{\text{ext}}^l(\mathbf{r},K_l,\beta_l)\rho(\mathbf{r})+F\{\rho(\mathbf{r})\}. \quad (22)$$

Since we calculated the charge densities of the ions by varying the external auxiliary potentials, expression (22) is in fact a functional of the auxiliary potential $V_{\text{ext}}^l(\mathbf{r})$. To determine the self-energy E_i^{self} of an ion with nonequilibrium charge density distribution as a functional of the density it is necessary according to the ideas of nonequilibrium thermodynamics[30] to proceed as follows.

First, from the relation

$$\frac{\delta E^{\text{ion}}}{\delta V_{\text{ext}}(\mathbf{r})}=\rho(\mathbf{r}) \quad (23)$$

it follows that we should express $V_{\text{ext}}(\mathbf{r})$ in terms of $\rho(\mathbf{r})$ and substitute them in the formula for E_{ion}. After this, it is necessary to subtract from the energy E_{ion} the work A associated with the interaction with the external field, i.e., the quantity

$$A=\int d\mathbf{r}\ V^{\text{ext}}(\rho(\mathbf{r}))\ \rho(\mathbf{r}). \quad (24)$$

All these procedures in DFT are actually trivial since the energy of the ion (22) is always expressed as the corresponding density functional, and for $E_{\text{ion}}^{\text{self}}$ we obtain

$$E_{\text{ion}}^{\text{self}}=\int\ \frac{Z\rho(\mathbf{r})}{|\mathbf{r}|}+\frac{1}{2}\int d\mathbf{r}\ d\mathbf{r}'\ \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}+F\{\rho(\mathbf{r})\}. \quad (25)$$

After this, expression (5) for the total crystal energy can be rewritten in the form

$$E_{\text{cr}}=\left[E\left\{\sum_i\rho_i(\mathbf{r}-\mathbf{R}_i)\right\}-\sum_i E_i^{\text{self}}\{\rho_i(\mathbf{r}-\mathbf{R}_i)\}\right]+E^N$$
$$+\sum_i E_i^{\text{self}}\{\rho_i(\mathbf{r}-\mathbf{R}_i)\}. \quad (26)$$

We point out here at the outset that in all the refinements of the Gordon–Kim method the first term on the right-hand side of Eq. (26) coincides exactly with our result. This means that from the electron energy, written as a function of all the densities, the sum of energies of the individual ions was subtracted in a form corresponding exactly to our expression (25). In this light the source of the differences in the previous generalizations of this method in the form of the last term in Eq. (26) is not entirely clear to us. Nevertheless, we will not use the last term of Eq. (26) in our calculations in the form of expression (25). The quantity E_i^{self} as defined by expression (25) contains very large contributions from the ion energy, which in fact do not depend on the crystalline environment, and taking them into account can lead to large errors in the numerical calculations. Toward this end, we will make use of the well-known properties of the density functional and first calculate the necessary derivatives of the self-energy with respect to the variational parameters K_l. Toward this end, we write E^{self} in the form

$$E_{\text{ion}}^{\text{self}}=E_{\text{ion}}-\int d\mathbf{r}\ V_{\text{ext}}(\mathbf{r},K_l)\rho(\mathbf{r}) \quad (27)$$

and, correspondingly,

$$\frac{\partial E_{\text{ion}}^{\text{self}}}{\partial K_l}=\frac{\partial E_{\text{ion}}}{\partial K_l}-\int d\mathbf{r}\ \frac{\partial V_{\text{ext}}}{\partial K_l}\rho(\mathbf{r})-\int d\mathbf{r}\ V_{\text{ext}}(\mathbf{r},K_l)\ \frac{\partial\rho}{\partial K_l}. \quad (28)$$

From the condition for the minimum of the functional E_{ion} we have

$$\frac{\partial E_{\text{ion}}}{\partial K_l}=\int d\mathbf{r}\ \frac{\partial V_{\text{ext}}(\mathbf{r},K_l)}{\partial K_l}\ \rho(\mathbf{r}), \quad (29)$$

which at once gives us

$$\frac{\partial E_{\text{ion}}^{\text{self}}}{\partial K_l}=-\int d\mathbf{r}\ V_{\text{ext}}(\mathbf{r},K_l)\ \frac{\partial\rho}{\partial K_l}. \quad (30)$$

Moreover, we can also easily calculate the variation of the self-energy with variation of the parameters after integrating expression (30) with respect to K_l:

$$\Delta E_{\text{ion}}^{\text{self}}=-\int dK\int d\mathbf{r}\ V_{\text{ext}}(\mathbf{r},K_l)\ \frac{\partial\rho}{\partial K_l}. \quad (31)$$

Taking into account the linear relationship (21) between the parameters K_l and the multipole moments, we easily ob-

tain a final expression for the self-energy of the ion for multipole distortions of the density with $l>0$ by integrating in (31) from zero to K_l:

$$E_l^{\text{self}}(l)=K_l^2/2\alpha_l, \tag{32}$$

where α_l is the corresponding multipole polarizability. For monopole distortions we cannot calculate for any value of K_0, including $K_0=0$, for the reasons discussed above. To determine the energy difference arising for small lattice distortions, it is sufficient to integrate with respect to K_0 in (31) over an interval including the quite small region of actual variations of K_0.

To wrap up this section, we briefly describe the transformations of expression (26) for the total energy of the crystal to a form suitable for numerical analysis, following Ref. 34.

First we note that calculation of the self-energy does not present any special difficulty in light of the results for it obtained above. The remaining terms in expression (26) are very large and should be grouped correspondingly. First let us consider the contribution to the crystal energy associated with the kinetic and exchange–correlation energy of the electrons. It has the form

$$E_{el}=\int_V d\mathbf{r}\, F\left\{\sum_i \rho_i(\mathbf{r}-\mathbf{R}_i)\right\}$$
$$-\sum_i \int_V d\mathbf{r}\, F\{\rho_i(\mathbf{r}-\mathbf{R}_i)\}, \tag{33}$$

where the integration is over all space. By virtue of the periodicity of the total charge density of the crystal, the first integral in (33) can at once be represented as an integral over one unit cell. In the second integral it is possible to transform from an integration over all space to an integration over the lth unit cell and a sum over unit cells. As a result we have

$$E_{el}=\int_\Omega d\mathbf{r}\, F\left\{\sum_i \rho_i(\mathbf{r}-\mathbf{R}_i)\right\}$$
$$-\sum_{i,l} \int_{\Omega_l} d\mathbf{r}\, F\{\rho_i(\mathbf{r}-\mathbf{R}_{il})\}. \tag{34}$$

With the help of the change of variables $\mathbf{r}=\mathbf{r}'-\mathbf{r}_l$, where $\mathbf{r}_l$ is the position vector of the lth unit cell, it is easy to reduce the second integral to an integration over one unit cell. As a result, we obtain

$$E_{el}=\int_\Omega d\mathbf{r}\left[F\left\{\sum_i \rho_i(\mathbf{r}-\mathbf{R}_i)\right\}-\sum_{i,l} F\{\rho_i(\mathbf{r}-\mathbf{R}_{il})\}\right], \tag{35}$$

where $\mathbf{R}_{il}=\mathbf{R}_i+\mathbf{r}_l$.

Because of the exponential falloff of the charge density, the summation over lattice sites is carried out quite simply. The integration over a unit cell is performed numerically using the method of special points. Details of the calculation are contained in Ref. 35.

The greatest difficulty in the calculation comes from the Coulomb part of the total crystal energy due to the presence of long-range forces. The main problem in the transformation of this contribution consists in separating the long-range

interactions of the corresponding point multipoles and the short-range interactions of the extended multipoles.

The Coulomb part of the crystal energy E_c can be represented in the form

$$E_c=\sum_{i\neq j}\left(Z_i^N V_j(\mathbf{R}_i-\mathbf{R}_j)+\frac{1}{2}\int d\mathbf{r}\, \rho_i(\mathbf{r}-\mathbf{R}_i)\right.$$
$$\left.\times V_j(\mathbf{r}-\mathbf{R}_j)\right)+\mathbf{E}_N, \tag{36}$$

where $V_i(\mathbf{r})$ is the Coulomb potential created by the electrons of the individual ion.

The expansion of the charge density $\rho(\mathbf{r})$ into multipoles is written as follows:

$$\rho(\mathbf{r})=\sum_{m,l} \rho_l(r)Q_{lm}Y_{lm}^*(\hat{\mathbf{r}}), \tag{37}$$

where Q_{lm} are the multipole charges:

$$Q_0^{(0)}=\sqrt{4\pi}Z^{el}, \quad Q_0^{(1)}=i\sqrt{\frac{4\pi}{3}}\, P_z,$$

$$Q_{\pm1}^{(1)}=\mp\sqrt{\frac{2\pi}{3}}\, (P_x\pm iP_y)\ldots,$$

and $\rho_l(r)$ are the multipole components of the charge density, normalized by the following condition:

$$\frac{4\pi}{2l+1}\int_0^{+\infty} r^{l+2}\rho_l(r)dr=1. \tag{38}$$

The multipole expansion of the Coulomb potential $V_i(\mathbf{R})$ is written as follows:

$$V(\mathbf{R})=\int d\mathbf{r}\, \frac{\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{R}|}=\sum_{l,m} Q_{lm}\int d\mathbf{r}\, \frac{\rho_l(r)Y_{lm}^*(\hat{\mathbf{r}})}{|\mathbf{r}-\mathbf{R}|}$$
$$=\sum_{l,m} Q_{lm}V_l(R)Y_{lm}^*(\hat{\mathbf{R}}), \tag{39}$$

$$V_l(R)=\frac{4\pi}{2l+1}\left(\frac{1}{R^{l+1}}\int_0^R r^{l+2}\rho_l(r)dr\right.$$
$$\left.+R^l\int_R^{+\infty} \rho_l(r)r^{1-l}dr\right). \tag{40}$$

From the last expression we can separate out the long-range part after rewriting the potential in the following form:

$$V_l(R)=\frac{4\pi}{2l+1}\left(-\frac{1}{R^{l+1}}\int_R^{+\infty} r^{l+2}\rho_l(r)dr\right.$$
$$\left.+R^l\int_R^{+\infty} \rho_l(r)r^{1-l}dr\right)+\frac{1}{R^{l+1}}. \tag{41}$$

Thus, the Coulomb potential is represented in the form of a short-range part $\tilde{V}(\mathbf{R})$ and the long-range potential $V^p(\mathbf{R})$ of point multipoles

$$V(\mathbf{R})=\tilde{V}(\mathbf{R})+V^p(\mathbf{R}). \tag{42}$$

Substituting the expansion (42) into the expression for the Coulomb energy (36), we obtain

$$E_c = \sum_{i \neq j} \left(Z_i^N \widetilde{V}_j(\mathbf{R}_i - \mathbf{R}_j) + Z_i^N V_j^p(\mathbf{R}_i - \mathbf{R}_j) \right.$$

$$\left. + \frac{1}{2} \int d\mathbf{r} \, \rho_i(\mathbf{r} - \mathbf{R}_i) \widetilde{V}_j(\mathbf{r} - \mathbf{R}_j) + \frac{1}{2} \int d\mathbf{r} \, \rho_i \right.$$

$$\left. \times (\mathbf{r} - \mathbf{R}_i) V_j^p(\mathbf{r} - \mathbf{R}_j) \right) + E_N. \tag{43}$$

By virtue of the exponential falloff of the potential $\widetilde{V}(\mathbf{R})$ only the sum over nearest neighbors contributes to the first and third terms. These two terms together describe the short-range interaction of the extended multipoles. The sum of the three remaining terms, including the Coulomb interaction of the nuclei, describes the interaction of the point multipoles with each other and the interaction of the point multipoles with the extended multipoles. The procedure for separating the point objects from the interactions while taking account of only monopole and dipole distortions of the charge density is described in detail in our previous papers[21,22] using the standard description of dipoles in Cartesian coordinates. The account of the quadrupole distortions introduces an additional matrix, the quadrupole moment matrix, and working with it in Cartesian coordinates becomes extremely tiresome. In the present paper we represent the multipole distortions of the charge density in spherical coordinates. Using the transformation formulas for spherical functions, this makes it quite simple to formulate the calculation of the Coulomb interactions between point and extended multipoles. An analogous technique is used in electron band theory[36] to calculate the structure constants. Details of the calculations of expression (43) are given in Ref. 35.

3. RESULTS AND DISCUSSION

In our numerical calculations of the energy of perovskite crystals in distorted phases as well as in the cubic phase we have taken into account multipole components of the ion charge density up to quadrupole. The necessity in perovskite crystals of allowing for quadrupole distortions of the charge density of the O^{2-} ion even in the cubic phase follows directly from the comparisons made above[14,37] of the charge density distributions in the $BaTiO_3$ crystal obtained from first-principles calculations and from a superposition of spherically symmetric ion densities. The difference between these densities clearly demonstrates the presence of a quadrupole distribution of the charge density on the oxygen ion. By virtue of the crystalline symmetry of the cubic phase dipole components of the density are forbidden on all ions, and quadrupole components are forbidden on all ions except O^{2-}. In the distorted phases, on the other hand, corresponding, for example, to the ferroelectric state of the crystal, it is necessary to allow for the existence of both these components on all ions. The importance of quadrupole components of the ion density for stabilization of the low-symmetry phases of ionic crystals was noted in a recent paper by Wilson *et al.*[38] We have carried out calculations of the energies of the cubic phases and distorted phases of $BaTiO_3$, $KNbO_3$, $CaTiO_3$, and $BaZrO_3$. For each value of the lattice parameters we minimized the total crystal energy over the five

variational parameters K_0, $P_{1,(2)}$, $\beta_{1,(2)}$. Here we used the conditions for an energy minimum in the form

$$\partial E_{\mathrm{cr}} / \partial P = 0, \tag{44}$$

where by P we mean any of the above five parameters. We examined the minimization over the cutoff parameters of the multipole potentials β_l in detail for the case of $BaTiO_3$. We showed that the crystal energy possesses an extremely shallow minimum with respect to these parameters, and the exact values of β_l^{min} differ only slightly from one other for the different phases and unit-cell volumes. In line with this, most of our calculations for the parameters β_l were based on the use of the constant values $\beta_l = 0.15,\dots,\beta_2 = 0.25$. The effective distances $r_{\mathrm{eff}}^l = 1/\sqrt{\beta_l}$ corresponding to the indicated values of β_l correspond roughly to the interatomic distance for the dipole potential and are slightly less for the quadrupole potential.

Let us comment, first of all, on the results of our calculations for $BaTiO_3$. Carrying out the minimization of the total crystal energy in the cubic phase for different values of the lattice constant, we determined its equilibrium value a. It turned out to be equal to 7.41 Bohr radii. This value is less than the experimental value, which is equal to 7.58, but is only 0.5% less than the value obtained in previous calculations[13–15] based on the electron band description. We obtained exactly the same accuracy for the equilibrium parameters in the cubic phase for the other perovskite crystals. Taking the quadrupole distortion of the charge density on the O^{2-} ions into account leads to a slight increase of about 0.5% in the value of the lattice constant. As is well known from theoretical calculations,[13–15] the ferroelectric instability in perovskite crystals is very sensitive to the volume of the unit cell. First-principles calculations[13–15] have demonstrated that this instability is either very weakly manifested or is generally absent for a volume corresponding to the theoretical value of the lattice constant equal to 7.45. Hence all of the foregoing calculations were performed for the experimental value of the lattice constant. Our work is in this sense not an exception—we also performed all calculations of distorted phases for experimental values of a.

As we already noted earlier, the actual ion shifts during the transition to the ferroelectric phase do not reduce to just a displacement of the central Ti ion from its equilibrium position. A complete study of the possible ion shifts and a determination of the dependence of the energy on these displacements, as noted in Ref. 14, would be an extremely arduous undertaking even using simple pairwise interatomic potentials. Such a complete study was not the goal of this work. In our study of the $BaTiO_3$ instability, we employed experimental values of the eigenvectors of the soft ferroelectric mode and performed calculations of the dependence of the energy on just the amplitude of this mode. The results are shown in Fig. 1. As can be seen from the figure, the calculated values of the energy and the ferroelectric shift which do not allow for quadrupole distortions of the charge density overestimate the shift by a factor of two and the resulting energy decrease by an order of magnitude. The calculated values which take the quadrupole distortions into account, on the other hand, agree quite well with the values obtained in
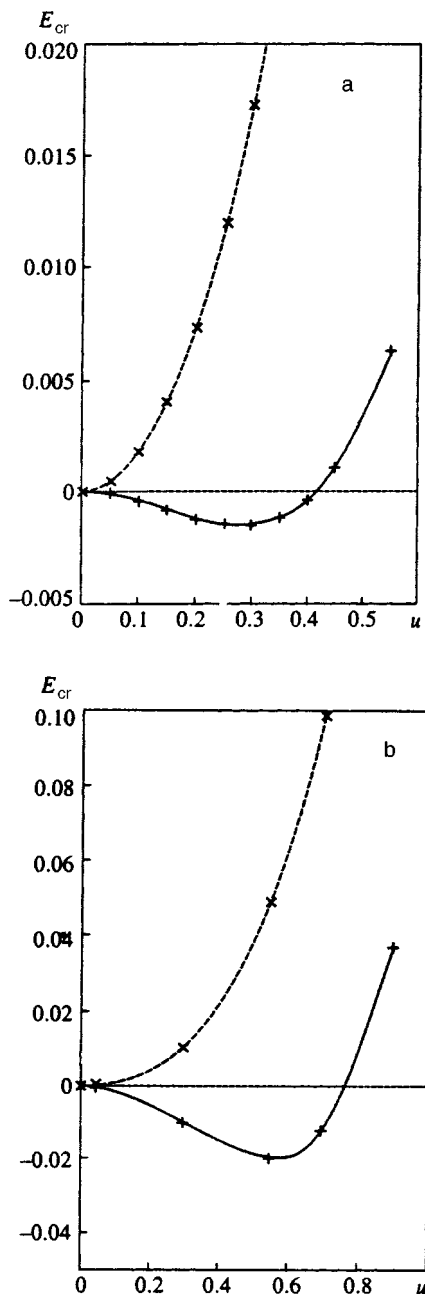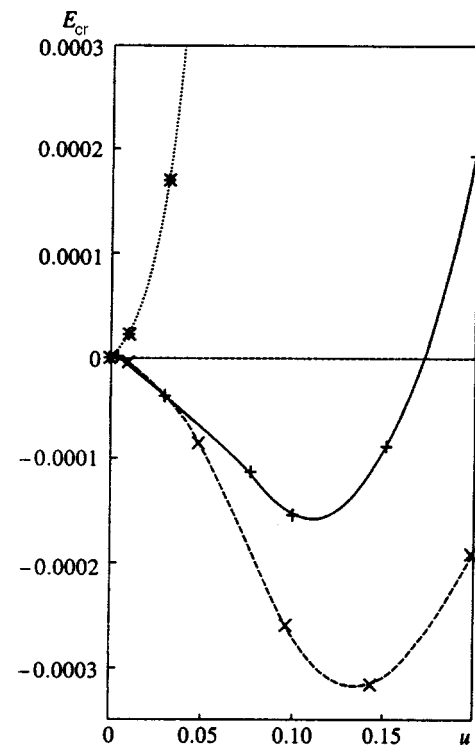
FIG. 1. Crystal energy E_{cr} versus amplitude of the ferroelectric shift u in BaTiO$_3$ (zero energy corresponds to the equilibrium value E_{cr}^0 in the cubic phase) in the ''rigid ion'' model ($\times$) and in the self-consistent calculation ($+$); a—monopole, dipole, and quadrupole distortions of the charge density of the ions taken into account, b—only monopole and dipole distortions taken into account.

FIG. 2. Crystal energy E_{cr} versus amplitude of the ferroelectric shift u in KNbO$_3$ (zero energy corresponds to the equilibrium value E_{cr}^0 in the cubic phase): $+$ corresponds to the Γ point, $\times$ corresponds to the X point (longitudinal mode), $*$ corresponds to the X point (transverse mode).

previous works.[13–15] Figure 1 also demonstrates that when the dipole distortions of the ions are neglected, i.e., in the rigid ion model, the cubic phase remains stable.

This result serves as explicit confirmation of the validity of ideas going back to Slater[5] and Skanavi[4] on the importance of the electron polarizability of the ions in the determination of the ferroelectric instability. Earlier, the ferroelectric instability in BaTiO$_3$ was examined in Ref. 39 within the framework of the Gordon–Kim model without allowance for the dipole polarization of the ions. As was shown in this work, the cubic phase is stable in this model. We can only

emphasize that the absence of a ferroelectric instability in this approach has nothing to do with whether hybridization of the electron states of the Ti^{4+} and O^{2-} ions is or is not taken into account, as was supposed in Ref. 13, but is explained simply by the absence in this approach of an account of the electron dipole polarization of the ions.

We also examined lattice instabilities for the KNbO$_3$ crystal. The results of these calculations are shown in Fig. 2. In this case we calculated not the change in the energy corresponding to an actual ferroelectric ion shift in the unit cell but the change in the energy due to a shift of the one central Nb ion in the unit cell. As can be seen from this figure, the lattice is unstable with respect to this shift. Moreover, this mode is unstable not only at the center of the Brillouin zone, but also on its boundary. Thus, there exist branches of the optical vibrations in perovskite crystals that are unstable in the cubic phase over the entire Brillouin zone. This result is in complete agreement with the results obtained recently within the framework of first-principles calculations based on the electron band structure method.[40] Such behavior of the phonon modes can have a substantial effect on many characteristics of the cubic phase of perovskite crystals. Figure 3 plots the results of calculations of the so-called antiferroelectric instability for BaTiO$_3$ and CaTiO$_3$. These results, in complete agreement with the available experimental data and the results of recent calculations,[16] demonstrate the absence of the corresponding instability in BaTiO$_3$ and its presence in CaTiO$_3$. They also show that this instability is almost completely determined by the Coulomb interaction of the spherically symmetric ions and by the close-range forces.

FIG. 3. Crystal energy E_{cr} versus rotation angle of the oxygen octahedron ϕ in BaTiO$_3$ and CaTiO$_3$ (zero energy corresponds to the equilibrium value E_{cr}^0 in the cubic phase): $\times$ corresponds to the ''rigid ion'' model, $+$ corresponds to the self-consistent calculation.

TABLE I. Born effective charges.

Ion	Charge	$\bar{Z}_{eff}$	$\bar{Z}_{eff}^{exp}$
Ba	2	2.71	2.9
Ti	4	5.67	6.7
O$_1$	−2	−4.72	−4.8
O$_2$	−2	−1.83	−2.4
O$_2$	−2	−1.83	−2.4

$$\varepsilon_\infty = 1 + \frac{4\pi}{v_0 E} \sum_s \mathbf{P}_s. \qquad (48)$$

Here the sum is taken over the limits of one unit cell and v_0 is its volume. The results of numerical calculations lead to the following value of the dielectric constant:

$$\varepsilon_\infty = 4.86.$$

It is also possible to determine the magnitudes of the Born effective charges that describe the polarization of the crystal arising upon an ion shift equal to $\mathbf{u}_s$

$$\delta\mathbf{P} = \frac{1}{v_0} \sum_s \hat{Z}_s^{eff} \mathbf{u}_s = \frac{1}{v_0} \sum_s \hat{Z}_s^{ion} \mathbf{u}_s + \frac{1}{v_0} \sum_s \mathbf{P}_s. \qquad (49)$$

The effective charge tensors $Z_{\alpha\beta}^{eff}$ in perovskite crystals for the A and B ions are isotropic and have two different values for the O^{2-} ion for nonequivalent directions: Z_1^{eff} in the direction toward the B ion and $\hat{Z}_2^{eff}$ perpendicular to it. Performing four ion shifts (one each for the A and B ions and two for the O^{2-} ion) and calculating from the energy minimum the corresponding values of the electron polarizability $\mathbf{P}_n$, we can determine all four values of the Born effective charges. Results of the corresponding calculations for BaTiO$_3$ are listed in Table I.

We note first that the Born effective charges we calculated satisfy exactly the acoustic sum rule

$$\sum_s \widetilde{Z}_{eff}^s = 0. \qquad (50)$$

The dielectric constant ε_∞ and the Born effective charges we obtained slightly underestimate the experimental data. This has to do, to a significant extent, with the substantial lowering of the polarizability of the individual ions due to the strong cutoff factor $e^{-\beta_l r^2}$ in the perturbing dipole potential. Choice of smoother cutoff functions can lead to better agreement with experiment. In any case, considering the complete absence of adjustable parameters in our calculations, the agreement obtained with the experimental data can be considered to be quite good.

After our demonstration that the GGK method we have developed leads to results for the cubic phase and distorted phases of perovskite crystals, whose accuracy does not differ very much from that of other methods, it would be desirable to address the questions posed in the Introduction. The most important of these is the question, what is the nature of the ferroelectric instability and what sort of interaction forces are responsible for this instability. This problem, of course, was discussed at length in all earlier calculations. It should be

Electron polarizability effects have a very small effect on this instability. Our calculations for BaZrO$_3$ demonstrate stability of the cubic lattice in complete agreement both with previous calculations[15] and with the experimental data.

Our method allows one to calculate not only the energy of the cubic phase and distorted phases, but also the total and electron polarizations of the distorted phases. First of all, let us consider the behavior of a cubic crystal in an external electric field $\mathbf{E}_n$. Toward this end, we add a term to the crystal energy allowing for the interaction of the electric dipoles with this external electric field:

$$W = -\sum_n \mathbf{P}_n \mathbf{E}_n \qquad (45)$$

and restrict discussion to the case of a uniform field $\mathbf{E}_n = \mathbf{E}$ = const. Minimizing the total energy of the crystal for given $\mathbf{E}$, we can find, in the linear approximation, the values of $\mathbf{P}_n$ as functions of $\mathbf{E}$:

$$\mathbf{P}_n = \widetilde{\alpha}_n \mathbf{E}_n, \qquad (46)$$

where the quantities $\widetilde{\alpha}_n$ are, of course, not equal to the quantities α_n defined for the isolated ions. Next we can find the electronic contribution ε_∞ to the dielectric constant of the crystal from the relation

$$\mathbf{D} = \varepsilon_\infty \mathbf{E} = \mathbf{E} + 4\pi\mathbf{P} = \mathbf{E} + \frac{4\pi}{v_0} \sum_s \mathbf{P}_s \qquad (47)$$

or

noted at once, though, that they gave entirely different answers to these questions. In order to get to the bottom of these differences, we use the standard expression for the total crystal energy in the density functional theory

$$E_{\mathrm{cr}} = E_{el}\{\rho(\mathbf{r})\} + E^N, \tag{51}$$

where $E_{el}\{\rho(\mathbf{r})\}$ is the electron energy functional and E^N is the energy of the Coulomb interaction of the nuclei (or interaction energy of the corresponding ions using the pseudopotential method). We write the functional $E_{el}\{\rho(\mathbf{r})\}$ in the two equivalent forms most frequently used for calculations in the band description of the electrons:

$$E_{el}\{\rho(\mathbf{r})\} = T_0\{\rho(\mathbf{r})\} + \sum_n \int d\mathbf{r}\, \frac{Z_n\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{R}_n|}$$
$$+ \frac{1}{2}\int d\mathbf{r}'\, \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + E_{xc}\{\rho(\mathbf{r})\} \tag{52}$$

and

$$E_{el}\{\rho(\mathbf{r})\} = \sum_{\mathbf{k},\lambda} \varepsilon_{\mathbf{k},\lambda} - \frac{1}{2}\int d\mathbf{r}'\, \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + E_{xc}\{\rho(\mathbf{r})\}$$
$$- \int d\mathbf{r}\, V_{xc}\{\rho(\mathbf{r})\}\rho(\mathbf{r}). \tag{53}$$

In formulas (52) and (53) we have used the standard notation: $T_0\{\rho(\mathbf{r})\}$ is the kinetic energy of the noninteracting ions, $E_{xc}\{\rho(\mathbf{r})\}$ is the exchange–correlation energy, and $V_{xc}\{\rho(\mathbf{r})\}$ is the exchange–correlation potential in the corresponding Kohn–Sham equation for the band electrons. The first term on the right-hand side of Eq. (53) is the sum of the single-particle electron energies described by the Kohn–Sham equation.

It would be desirable to turn our attention now to the change of sign of the contribution of the Coulomb interelectron interaction in Eq. (53) in comparison with Eq. (52) and to the appearance of an additional term (the last term) in Eq. (53). Both of these changes arise as a result of double counting of the corresponding interactions in the sum of the single-particle energies. A change in any of the terms (in the series of cases and even for E_{xc}) in the expression for $E_{el}\{\rho(\mathbf{r})\}$ for a small distortion of the crystal structure is much larger than the resulting change in the total crystal energy E_{cr}.

In this sense, any of these changes leading to a decrease in the energy of the distorted phase in comparison with the cubic phase can, if desired, be considered as the reason for the appearance of the distorted phase itself. Thus, Rabe and Jonnopoulos,[41] in the context of calculations of the ferroelectric instability in IV–VI semiconductors using the form (52) for E_{el}, cited the interaction of the electrons and ions, i.e., the second term on the right-hand side of Eq. (52), as the reason for the ferroelectric instability. This term is generally absent in explicit form in formula (53). Hidaka,[42] who used formula (53) for E_{el}, cited as the reason for the ferroelectric instability the sum of single-electron energies, which decreases substantially during the transition to the ferroelectric phase. Precisely this reason, i.e., the decrease in the sum of single-particle energies due to the electron–phonon interaction, was considered as the main reason for the ferroelectric instability in the so-called vibronic model (see the recent review in Ref. 43).

The previously cited recent calculations of perovskite crystals[13–15] cited as the reason for the ferroelectric instability a ''delicate balance between the long-range Coulomb forces and short-range repulsion forces.'' In reality, this conclusion, which is, in our opinion, absolutely valid, in no way follows directly from the calculations performed in Refs. 13–15. In these calculations, no redistribution of contributions in the expression for the total crystal energy was attempted which would have led to a separation of the short-range and long-range forces.

This discussion can be shown to be completely academic as long as it concerns a clarification of the question, which of the crystalline structures has the lowest energy at $T=0$. The problem acquires a clear physical meaning in the consideration of phase transitions in perovskites at temperatures $T \neq 0$, and also in a discussion of the properties of the various phases at finite temperatures. As is well known,[1,3] ferroelectrics are usually divided into two classes, those of order–disorder type and those of displacement type. In ferroelectrics of order–disorder type there are initially several nonequivalent sites in the unit cell for one of the ions, i.e., there exists a so-called multiwell potential. In this case, the existence of an ion at any of these sites leads to the presence of a dipole moment in the unit cell. At high temperatures the dipole moments are disordered either because of ion hops from one site to another or because of disorder in the arrangement of the various dipoles. At the phase transition point ordering of the dipole moments of the unit cells takes place. Ferroelectrics of displacement type are characterized by the absence of a multiwell potential for the ions at temperatures above the phase transition temperature T_c.

The very concept of a multiwell potential arises only for $T<T_c$ due to the appearance of real disordered ion shifts throughout the entire crystal. On this plane, the various conclusions about the nature of ferroelectricity in perovskite crystals reached via first-principles calculations at $T=0$ can lead to quite different conclusions about the nature of the phase transition and the properties of the cubic phase. Thus, the vibronic model assumes the possible presence of a multilevel potential for the central B ion in each unit cell. The polarizable ion model, naturally, leads to a ferroelectric transition of displacement type. If the only reason for the ferroelectric instability is the long-range dipole–dipole interaction, then no multiwell potential for the ions can exist until ordered dipoles exist throughout the entire crystal or in some large part of it. It is clear that such an ordering is absent in the polarizable ion model for $T>T_c$, outside the region of strong ferroelectric fluctuations. Perovskite crystals have usually been considered as classical examples of ferroelectrics of displacement type.[1,3]

During the last 20 years a large number of experimental works have appeared (see Ref. 44 and the references therein), whose authors have interpreted their results as a manifestation of order–disorder effects in perovskite crystals. Theoretical works have also appeared[44] in which along with a treatment of the soft ferroelectric mode due to the

dipole–dipole interaction, additional phenomenological mechanisms are introduced leading to the existence of a multiwell potential also for $T > T_c$. As was shown in a recent paper by Hüler,[45] many of these phenomena, which are reminiscent of order–disorder effects, can in reality be explained within the framework of the standard quasiharmonic approximation for the lattice dynamics of a perovskite crystal. In particular, Hüler showed that the strong anisotropy of the correlation of the B atom displacements due to this instability of entire branches of the optical vibrations in the cubic phase makes the diffusive x-ray scattering highly anisotropic. This has often been linked with disordered static displacements of the B ion from its equilibrium position. Another example, in our view confirming the validity of the polarizable ion model and its ability to explain many phenomena similar to order–disorder effects, is the recent Monte Carlo simulation[46] of the phase transition in BaTiO$_3$ for a system of interacting dipoles. The parameters of the model Hamiltonian were determined from first-principles calculations of the energies of the cubic and ferroelectric phases at $T = 0$.

Now let us go on to a more detailed discussion of questions associated with factors leading to the ferroelectric instability, and also with the problem of the anomalously large values of the Born effective charges in perovskites. Toward this end, we take advantage of the possibility of an analytical representation of the total crystal energy within the framework of the GGK method. But first of all, we simplify this expression guided by the results of the numerical calculations. First, we neglect the change in the quadrupole moment of the ions at the transition from the cubic to the ferroelectric phase and, minimizing with respect to it, we exclude it from further consideration. This procedure leads only to a corresponding renormalization of the short-range interactions. Next, we restrict the discussion to the approximation of pairwise interactions between the deformed ions. This approximation, as was already noted, is insufficient in a number of cases to maintain the required accuracy for the numerical calculations, but it does not lead to any qualitative changes in the calculated results. In final form the expression for the total crystal energy can be written in the form

$$E_{\mathrm{cr}} = \widetilde{E}_{sr}\{K_0\} + \frac{1}{2} \sum_{n,n'} \frac{Z_n^{\mathrm{ion}} Z_{n'}^{\mathrm{ion}}}{|\mathbf{R}_n - \mathbf{R}_{n'}|} + \frac{1}{2} \sum_n \frac{P_n^2}{2\alpha_n}$$

$$+ \frac{1}{2} \sum_{n,n'} \mathbf{P}_n \hat{\Phi}_{n,n'} \mathbf{P}_{n'} + \frac{1}{2} \sum_{n,n'} \mathbf{P}_n \hat{\Gamma}_{n,n'} \mathbf{P}_{n'}$$

$$- \sum_{n,n'} \mathbf{P}_n \hat{S}_{n,n'} - \sum_n \mathbf{P}_n \mathbf{E}_n . \tag{54}$$

In the form of this expression we have changed over from the representation of the dipole moment vectors $\mathbf{P}_n$ of the ions in spherical coordinates (which is convenient for the numerical calculations) to the standard form in Cartesian coordinates. Formally, expression (54) completely coincides with the expression we obtained earlier[21] allowing only for dipole distortions of the charge density. But in fact there is a difference. First, the contribution from the short-range inter-

action forces $\widetilde{E}_{sr}\{K_0\}$ takes into account the renormalization of these forces resulting from elimination of the quadrupole moment. Second, the matrix $\hat{S}_{n,n'}$ of the extended spherically-symmetric ions also includes quadrupole renormalization effects. The remaining terms in the linear approximation in the dipole and quadrupole distortions do not differ from those obtained earlier. The quantity α_n is the dipole polarizability of the ion, and $\hat{\Phi}_{n,n'}$ is the interaction matrix of the point dipoles:

$$\Phi^{\alpha\beta}(\mathbf{R}) = \frac{\delta_{\alpha\beta}}{R^3} - \frac{3 R_\alpha R_\beta}{R^5}; \tag{55}$$

the matrix $\hat{\Gamma}_{n,n'}$ describes the short-range interaction between the extended dipoles. The last term in expression (54) represents the interaction between the point dipoles and the electric field $\mathbf{E}_n$. The matrices $\hat{S}_{n,n'}$ and $\hat{\Gamma}_{n,n'}$ can be obtained from the numerical calculations described in the foregoing section. Using formula (54), it is possible to obtain an analytical expression for the electronic contribution to the static dielectric constant of the crystal ε_∞ and the Born effective charges Z_n^{eff} for perovskite crystals. Let us consider, first, the quantity ε_∞, taking the electric field $\mathbf{E}_n$ in formula (5) to be external. Then, from the condition

$$\partial E / \partial \mathbf{P}_n = 0 \tag{56}$$

we have

$$\mathbf{P}_n = \alpha_n \mathbf{E}_n - \alpha_n \sum_{n'} \hat{\Phi}_{n,n'} \mathbf{P}_{n'} - \alpha_n \sum_{n'} \hat{\Gamma}_{n,n'} \mathbf{P}_{n'} . \tag{57}$$

The last term in expression (57) can be eliminated by introducing the nonlocal polarizability

$$\hat{\alpha}_{n,n'} = (\delta_{n,n'} + \alpha_n \hat{\Gamma}_{n,n'})^{-1} \hat{\alpha}_{n'} . \tag{58}$$

We in fact are considering a uniform external field $\mathbf{E}_n = \mathrm{const}$, and the polarization of each unit cell is

$$\sum_s \mathbf{P}_s = \mathrm{const}.$$

In this case, the dipole–dipole interaction matrix in cubic crystals can be written as[9,10]

$$\Phi_{n,n'}^{\alpha\beta} = -\frac{4\pi}{3} \delta_{\alpha\beta} - \gamma_{n,n'}^{\alpha\beta} . \tag{59}$$

Here $\hat{\gamma}_{n,n'}$ is the matrix of internal field constants, characterizing the departure of the external field factor from a Lorentzian (equal to $4\pi/3$). The sum of products with the matrix $\hat{\gamma}_{n,n'}$ is taken only over nonequivalent sublattices. Expression (57) can now be rewritten in the form

$$\mathbf{P}_n = \sum_{n'} \hat{\alpha}_{n,n'} \mathbf{E} + \frac{4\pi}{3} \sum_{n'} \hat{\alpha}_{n,n'} \mathbf{P} + \sum_{n',n''} \hat{\alpha}_{n,n'} \hat{\gamma}_{n',n''} \mathbf{P}_{n''}, \tag{60}$$

where $\mathbf{P}$ is the total polarization of the crystal

$$\mathbf{P} = \sum_n \mathbf{P}_n = \frac{1}{v_0} \sum_s \mathbf{P}_s . \tag{61}$$

It is easy to solve Eq. (60) completely and to obtain an expression for the total polarization **P**. Noting that for crystals with cubic symmetry ε_∞ is a scalar, we have

$$\varepsilon_\infty = 1 + \frac{4\pi P}{E}, \tag{62}$$

$$\varepsilon_\infty = 1 + \frac{4\pi\alpha_c}{1 - (4\pi/3)\alpha_c}. \tag{63}$$

Here α_c is the polarizability of the unit cell:

$$\alpha_c = \frac{1}{v_0} \sum_{s,n,n',n''} (\delta_{s,n} - \hat{\alpha}_{s,n}\hat{\gamma}_{n,n'})^{-1} \hat{\alpha}_{n',n''}. \tag{64}$$

The s summation in Eq. (64) is over one unit cell, and over the remaining indices in accordance with the falloff radius of the matrices $\hat{\alpha}_{n,n'}$ and $\hat{\gamma}_{n,n'}$. For binary crystals, where $\hat{\gamma}_{n,n'} = 0$ holds, we have in the point dipole approximation $\Gamma_{n,n'} = 0$, and the polarizability of the unit cell reduces to the sum of the polarizabilities of the two ions:

$$\alpha_c = \alpha_1 + \alpha_2. \tag{65}$$

Formally, it is possible to introduce an effective polarizability of each ion α_s, which in this case is no longer a scalar, but a tensor in Cartesian coordinates:

$$\hat{\alpha}_s^{ik} = \sum_{n,n',n''} (\hat{I} - \hat{\alpha}_{s,n}\hat{\gamma}_{n,n'})^{-1} \hat{\alpha}_{n',n''}. \tag{66}$$

The total polarizability of the unit cell is expressed, in accordance with Eq. (64), by the sum of effective polarizabilities.

Let us first discuss the influence of non-Lorentzian internal field constants $\gamma_{n,n'}$ on the polarizability of the ions in perovskite crystals, neglecting for the time being, the matrix $\hat{\Gamma}_{n,n'}$:

$$\alpha_s^{ik} = \sum_{n,n'} (\hat{I} - \alpha_s\hat{\gamma}_{s,n})^{-1} \alpha_n. \tag{67}$$

We consider by way of an example the effective polarizability of the oxygen atom in the direction corresponding to the position of the Ti atom. As is well known,[11] the local field factors $\hat{\gamma}_{n,n'}$ coupling different sublattices differ greatly in magnitude. The constant $\gamma_{\mathrm{Ti,O}}$ is several times greater than all the other constants, and

$$\sum_{\mathrm{Ti}} \gamma_{\mathrm{O,Ti}} \approx 30. \tag{68}$$

This circumstance leads to a very large anisotropy of the effective polarizability of the oxygen atoms in a perovskite lattice. The polarization in the direction of the Ti ion is several times greater than its value in perpendicular directions. This has nothing to do with the properties of the oxygen atom itself, nor with hybridization of the electron states of Ti^{4+} and O^{2-}, rather it is a specific, geometrical property of the perovskite lattice, due to the existence of parallel chains of Ti^{4+} and O^{2-} ions. This circumstance, as was shown in Ref. 40, leads to the appearance of an instability on entire branches of the TO vibrations. In turn, the nonlocal polariz-

ability matrix $\hat{\alpha}_{n,n'}$ is also anisotropic due to the anisotropy of the matrices $\hat{\Gamma}_{n,n'}$ and $\hat{\gamma}_{n,n'}$. The nonlocal polarizability $\hat{\alpha}_{n,n'}$ most strongly hinders, in particular, the appearance of dipole moments on the oxygen ions in the Ti–O direction. Thus, in perovskites there arises a mutual compensation of the effect of non-Lorentzian corrections to the local field and the effect of short-range repulsion forces of the dipoles.

Formula (54) also allows one to obtain an analytical expression for the Born effective charge $\hat{Z}_n^{\mathrm{eff}}$ defined by formula (49). For the electron polarizability we have from formula (54)

$$\mathbf{P}_s = \alpha_s \mathbf{E}_s - \alpha_s \sum_{n'} \hat{\Phi}_{s,n'} \mathbf{P}_{n'} - \alpha_s \sum_{n'} \hat{\Gamma}_{s,n'} \mathbf{P}_{n'}$$
$$+ \sum_{n'} \hat{m}_{n,n'} Z_{n'} \mathbf{u}_{n'}. \tag{69}$$

Here we have introduced the nonlocal deformability of the ion $\tilde{m}_{n,n'}$, writing the matrix $\hat{S}_{n,n'}$ for small displacements $\mathbf{u}_n$ as

$$\hat{S}_{n,n'} = \hat{m}_{n,n'} Z_{n'} \mathbf{u}_{n'}. \tag{70}$$

The electric field $\mathbf{E}_s$ at a site has the form

$$\mathbf{E}_s = -\sum \hat{\bar{\Phi}}_{n,n'} Z_{n'} \mathbf{u}_{n'} \tag{71}$$

and

$$\hat{\bar{\Phi}}_{n,n'} = \hat{\Phi}_{n,n'} - \delta_{n,n'} \sum_l \hat{\Phi}_{n,l}. \tag{72}$$

Using the definition (58) for the nonlocal polarizability $\alpha_{n,n'}$, it is possible to rewrite formula (69) as

$$\mathbf{P}_s = \sum_{n'} \alpha_{n,n'} \mathbf{E}_n - \sum_{n',n''} \alpha_{n,n'} \hat{\Phi}_{n',n''} \mathbf{P}_{n''}$$
$$- \sum_{n'} Z_{n'} \hat{\bar{m}}_{n,n'} \mathbf{u}_{n'}, \tag{73}$$

where the quantity $\hat{\bar{m}}_{n,n'}$ is equal to

$$\hat{\bar{m}}_{n,n'} = \sum_{n''} \alpha_{n,n''} \hat{m}_{n'',n'}. \tag{74}$$

Formally, the expression we have just obtained for $\mathbf{P}_s$ coincides exactly with the expression derived for the general type of phenomenological model of lattice dynamics, taking nonlocal polarizability and nonlocal deformability of the ion into account.[47] Now, using formula (73), it is possible to express the electron dipole moment in terms of the ion shifts and, substituting it in formula (49) for the total dipole moment, to find an expression for the Born effective charge Z_s^{eff}. As a result of straightforward, but lengthy calculations similar to those performed above in the derivation of ε_∞, we obtain

$$\hat{Z}_s^{\mathrm{eff}} = \left(\hat{I} - \frac{1}{v_0} \frac{4\pi}{3} \alpha_c \right)^{-1} \hat{\xi}(s) = \frac{\varepsilon_\infty + 2}{3} \hat{\xi}(s), \tag{75}$$

$$\hat{\xi}(s) = \sum_t \left[\hat{I} - \frac{1}{v_0}\, \hat{\alpha}\, \hat{\gamma} \right]^{-1}_{ts} \left[\hat{I} - \frac{1}{v_0}\, \hat{\alpha}\, \hat{\gamma} \right]^{sn}$$

$$\times [\hat{I} - \hat{\tilde{m}}_{n,n'}]^{n,n'} Z_{n'}. \qquad (76)$$

The factor $(\varepsilon_\infty + 2)/3$ in the expression for the effective charge characterizes the well-known[48] growth of the external dipole moment in a polarizable medium with dielectric constant ε_∞. Formally, formula (76) for $\hat{\xi}(s)$ almost coincides with the formula obtained earlier[10] in the polarizable point ion model. The difference is that instead of a nonlocal scalar polarizability of one ion, $\alpha_n \delta_{n,n'} \delta^{\alpha\beta}$, formula (76) includes the nonlocal polarization matrix $\alpha_{n,n'}$. Also, expression (76) has an additional factor $\hat{I} - \hat{\tilde{m}}$ characterizing the diminution of the nominal charge, more accurately the dipole moment arising upon displacement of the ion due to its concomitant deformation. Using formulas (70) and (74) and the condition of translational invariance for $\hat{m}_{n,n'}$,

$$\sum_n \hat{\tilde{m}}_{n,n'} = 0, \qquad (77)$$

we can easily prove the acoustic sum rule for the Born charges.

Expanding the crystal energy (54) in a series in ion displacements $\mathbf{u}_n$ readily yields an expression in the harmonic approximation for the matrix of force constants of the crystal $\Phi^{\alpha\beta}_{s\tau}$ which depends on the coordinates of the atoms in the unit cell s and τ and the wave vector $\mathbf{q}$. For small $\mathbf{q}$ we have

$$\Phi^{\alpha\beta}_{s\tau}(\mathbf{q}+\beta) = \Phi^{\alpha\beta}_{s\tau}(0) + \tilde{C}^{\alpha\beta}_{s\tau}(0) - \frac{4\pi}{V_0}\, \tilde{Z}^{\text{eff}}_s\, \frac{I}{\varepsilon_\infty+2}\, \tilde{Z}^{\text{eff}}_\tau$$

$$+ \frac{4\pi}{V_0 \varepsilon_\infty}\, \tilde{Z}^{\text{eff}}_s\, \frac{q_i q_j}{\varepsilon_\infty q^2}\, \tilde{Z}^{\text{eff}}_\tau. \qquad (78)$$

Expression (78), like formulas (75) and (76) for the Born effective charge, formally coincides completely both with the expressions obtained earlier in the polarizable ion model[10] and with the expressions arising in the exact theory of lattice dynamics based on the electron dielectric matrix $\varepsilon(\mathbf{q}+\mathbf{K}, \mathbf{q}+\mathbf{K}', 0)$ (Ref. 9). The first term in formula (78) is the contribution to the matrix of force constants from the short-range repulsion forces of the spherically symmetric ions [the first term in the expression for the total energy (54)]. The fourth and third terms, respectively, describe the contributions of the non-analytical and analytical parts of the long-range dipole–dipole interaction in a polarizable medium with dielectric constant ε_∞. $\tilde{C}^{\alpha\beta}_{s\tau}(0)$ is an additional contribution to the short-range forces from the dipole–dipole interactions. The specific form of this matrix, like for $\hat{Z}^{\text{eff}}_{\alpha\beta}$ and the entire matrix $\Phi^{\alpha\beta}_{s\tau}$, coincides formally with the form obtained earlier in the polarizable point ion model.[10] Because of its length, we do not reproduce it here.

Let us now briefly discuss the analytical results obtained for $\hat{Z}^{\text{eff}}_s$ and $\Phi^{\alpha\beta}_{s\tau}$. As can be seen from formulas (75), (76), and (78), not only the ferroelectric instability itself, but also the Born effective charge $\hat{Z}^{\text{eff}}_s$, is a result of a ''delicate balance between the long-range dipole forces and short-range

repulsion forces.'' First of all, as can be seen from formula (75), ε_∞ increases the effective charges of all the ions of the crystal. Further, as was already discussed in the calculation of ε_∞, $\hat{Z}^{\text{eff}}_s$ increases substantially becasue of the non-Lorentzian local field constants $\hat{\gamma}_{s\tau}$. This growth is strongest for the B and O_I ions, which are bound by the largest (in magnitude) constant $\gamma_{B,O}$. The constant $\gamma_{s\tau}$ binding the A and O_{II} ions is also quite large.

This circumstance explains the correlation between Z^{eff} observed in the calculations precisely for these ions for a large number of perovskites. The corresponding growth of Z^{eff} in the polarizable point ion models is so large that to explain the magnitude of the total polarizability of the crystal it is necessary, following in the footsteps of Slater,[5] to radically decrease the values of the nominal charges of the ions. Expression (76) shows that this decrease arises automatically in real crystals due to the deformability of the ions as well as the short-range dipole–dipole interaction. Both these factors lead to a decrease in the electron dipole moment arising as a result of the ion displacements. Expression (78) for $\Phi^{\alpha\beta}_{s\tau}$ shows, in complete agreement with the polarizable ion models, that it is precisely the long-range dipole–dipole interaction [third term in expression (78)] that is the reason for the ferroelectric instability. All the short-range forces (i.e., the first and second term) favor stabilization of the cubic phase. The fourth term contributes only to the longitudinal optical modes and leads to splitting of the LO and TO modes.

4. CONCLUSION

Let us briefly sum up this work. First, we have presented a simple and clear scheme for calculating the static and dynamic properties of ionic crystals. This approach is based on the generalized Gordon–Kim model and allows one to take account of the distortions of the ion charge density corresponding to any multipole symmetry. In itself this approach, based on the density functional theory, is absolutely rigorous and does not contain any arbitrary approximations or adjustable parameters. Unfortunately, in its most general form it cannot be applied to real calculations. First, it is necessary to restrict the calculation to a finite number of multipole components of the charge density. Second, an exact expression for the energy as a functional of the density $\mathbf{E}\{\rho\}$ is not known, and it is necessary to use an approximation for it.

As was shown in this work, very simple approximations not relying on any adjustable parameters are able to describe very slight energy differences between the cubic phase and various types of distorted phases in perovskite crystals. For this it is sufficient to take into account multipole distortions up to quadrupole and use a simple local Thomas–Fermi–Dirac approximation for the density functional.

One of the merits of the proposed method consists in the possibility of representing the total crystal energy analytically as a function of the ion shifts and the electron dipole moments. Formally, the expression obtained coincides exactly with the most general type of phenomenological model of a polarizable and deformable ion. Moreover, this expression demonstrates that all the short-range forces, including those arising from the dipole–dipole interaction, stabilize the

cubic phase. The only contribution leading to a ferroelectric instability is the one associated with the analytical part of the long-range dipole–dipole interaction. This conclusion is in complete agreement with the classical theory of ferroelectrics of displacement type developed in the early works of Skanavi, Slater, Ginzburg, and Anderson. Our results also indicate that no mechanisms for phenomena of order–disorder type are present in perovskite crystals except for those that can arise within the framework of quasiharmonic or anharmonic dynamics.

*)E-mail: ivanov@td.lpi.ac.ru

[1] M. E. Lines and A. M. Glass, *Principles and Applications of Ferroelectrics and Related Materials* (Clarendon Press, Oxford, 1997).

[2] D. E. Cox and A. W. Seight, Solid State Commun. **19**, 969 (1976).

[3] V. G. Vaks, *Introduction to the Microscopic Theory of Ferroelectrics* [in Russian] (Nauka, Moscow, 1973).

[4] F. I. Skanavi, Zh. Éksp. Teor. Fiz. **17**, 399 (1947).

[5] J. C. Slater, Phys. Rev. **78**, 748 (1950).

[6] V. L. Ginzburg, Usp. Fiz. Nauk **38**, 490 (1949).

[7] F. Anderson, in *Physics of Dielectrics* [in Russian] (published by the Academy of Sciences of the USSR, Moscow, 1960).

[8] W. Cochran, Adv. Phys. **9**, 387 (1960).

[9] O. E. Kvyatkovskiĭ, Fiz. Tverd. Tela **27**, 63 (1985) [Sov. Phys. Solid State **27**, 35 (1985)].

[10] O. E. Kvyatkovskiĭ and E. G. Maksimov, Usp. Fiz. Nauk **154**, 3 (1988) [Sov. Phys. Usp. **31**, 1 (1988)].

[11] O. E. Kvyatkovskiĭ, Fiz. Tverd. Tela **35**, 2154 (1993) [Phys. Solid State **35**, 1071 (1993)].

[12] B. Szigeti, Trans. Faraday Soc. **45**, Pt. 2, 155 (1949).

[13] R. E. Cohen and H. Krakauer, Phys. Rev. B **42**, 6416 (1990).

[14] R. E. Cohen and H. Krakauer, Ferroelectrics **136**, 65 (1992).

[15] R. D. King-Smith and D. Vanderbilt, Phys. Rev. B **49**, 5828 (1994).

[16] W. Zhong, R. D. King-Smith, and D. Vanderbilt, Phys. Rev. Lett. **72**, 3618 (1994).

[17] R. Resta, M. Pasternak, and A. Baldereschi, Phys. Rev. Lett. **70**, 1010 (1993).

[18] J. X. Axe, Phys. Rev. **157**, 429 (1967).

[19] W. Zhong and D. Vanderbilt, Phys. Rev. Lett. **74**, 2587 (1995).

[20] O. V. Ivanov and E. G. Maksimov, Phys. Rev. Lett. **69**, 108 (1992).

[21] O. V. Ivanov and E. G. Maksimov, Zh. Éksp. Teor. Fiz. **108**, 1841 (1995) [JETP **81**, 1008 (1995)].

[22] O. V. Ivanov and E. G. Maksimov, Solid State Commun. **97**, 163 (1996).

[23] E. G. Maksimov, D. A. Shport, and O. V. Ivanov, Solid State Commun. **101**, 393 (1997).

[24] O. V. Ivanov, E. G. Maksimov, and I. I. Mazin, Solid State Commun. **76**, 1267 (1990).

[25] R. G. Gordon and Y. S. Kim, J. Chem. Phys. **56**, 3122 (1972).

[26] Y. S. Kim and R. G. Gordon, Phys. Rev. B **9**, 3548 (1974).

[27] P. Hohenberg and W. Kohn, Phys. Rev. B **136**, 864 (1964).

[28] D. R. Yoder and R. Colella, Phys. Rev. B **25**, 2545 (1982).

[29] R. G. Gordon and R. Lesar, Adv. Quantum Chem. **21**, 341 (1990).

[30] R. E. Watson, Phys. Rev. **111**, 1108 (1958).

[31] M. A. Leontovich, *Introduction to Thermodynamics. Statistical Physics* [in Russian] (Nauka, Moscow, 1983).

[32] W. Kohn and L. J. Sham, Phys. Rev. A **140**, 1113 (1965).

[33] R. M. Sternheimer, Phys. Rev. **96**, 951 (1954).

[34] C. Muhlhausen and R. G. Gordon, Phys. Rev. B **23**, 900 (1981).

[35] O. V. Ivanov, *POSIM User's Guide, Technical Notes of the Lebedev Physical Institute* (Moscow, 1996).

[36] O. K. Andersen, Phys. Rev. B **12**, 3060 (1975).

[37] K. H. Wierich and R. Siems, Z. Phys. B **61**, 63 (1985).

[38] M. Wilson, U. Schönberger, and M. W. Finnis, Phys. Rev. B **54**, 9147 (1996).

[39] L. L. Boyer, M. J. Mehl, J. W. Flocken and J. R. Hardy, J. Appl. Phys. **24**, Suppl. 24-2, 204 (1985).

[40] R. Yu and H. Krakauer, Phys. Rev. Lett. **74**, 4067 (1995).

[41] K. M. Rabe and J. D. Jonnopoulos, Phys. Rev. B **32**, 2302 (1985).

[42] T. Hidaka, Phys. Rev. B **20**, 2769 (1979).

[43] N. Kristoffel and P. Konsin, Phys. Status Solidi B **149**, 1 (1988).

[44] Y. Girshberg and Y. Yacoby, Solid State Commun. **103**, 425 (1997).

[45] A. Hüler, Solid State Commun. **7**, 589 (1969).

[46] W. Zhong, D. Vanderbilt, and K. M. Rabe, Phys. Rev. B **52**, 6301 (1995).

[47] K. Kunc, M. Balkanski, and M. A. Nusimovici, Phys. Rev. B **12**, 4346 (1975).

[48] G. D. Mahan, Phys. Rev. **153**, 983 (1967).

Translated by Paul F. Schippnick

The Ulam problem and the ionization of Rydberg atoms by microwave radiation

A. S. Roshchupkin

Moscow Engineering-Physics Institute, 115409 Moscow, Russia

V. P. Kraĭnov*⁾

Moscow Institute of Physics and Technology, 141700 Dolgoprudnyĭ, Moscow Region, Russia

We study the Ulam problem for long times (several million collisions) by numerical methods. We show that in the diffusion regime, which is valid for moderate times, this problem is mathematically equivalent to the problem of the diffusive ionization of atomic Rydberg states by microwave radiation. It is concluded that the diffusion regime sets in only for a very small number of initial conditions (field phases). It is theorized that the analogy between the two problems can be extrapolated to times longer than the diffusion time. We show in the Ulam problem that after the diffusional buildup of energy has finished, the quasistationary regime does not continue indefinitely: after several million particle-wall collisions the energy rapidly drops to zero. On the basis of this extrapolation we examine the possibility that an electron which has reached the continuous spectrum will not fly off to infinity (ionization), but will return to bound Rydberg states of the atoms (if the field acts for a sufficiently long time). This can make the diffusive ionization probability much lower than the value given by the known estimates. © *1998 American Institute of Physics.* [S1063-7761(98)00307-2]

1. INTRODUCTION

The ionization of atomic Rydberg states (primarily, of the hydrogen atom) by microwave radiation has been studied both experimentally and theoretically (see the monograph by Delone and Kraĭnov[1] and the reviews by Delone *et al.*[2] and Gasati *et al.*[3]). Delone *et al.*[4] showed for the first time, via the semiclassical approximation, that when an electron in a Rydberg state with the principal quantum number n is subjected to microwave radiation with a frequency of order of the Keplerian frequency of revolution of the electron about the atomic core, it is diffusively transferred to orbits with ever increasing quantum numbers until its spectrum becomes continuous, i.e., until ionization occurs. The same process takes place in purely classical mechanics, but only when the field strength exceeds a certain threshold value, $F > F_{\mathrm{th}}$ (Ref. 5). Thus, diffusive ionization is a threshold process in classical mechanics, according to the Kolmogorov–Arnold–Moser theorem. For instance, according to numerical calculations,[6] the threshold field strength for a linearly polarized field with a frequency $\omega = (1-4)/n^3$ is $F_{\mathrm{th}} = 1/40n^4$. (Here and below we use the atomic system of units). The threshold value is small compared to the atomic field $F_a = 1/16n^4$ (Ref. 7), in which the ionization of an atomic Rydberg state by a low-frequency field takes place in the course of a single Keplerian period $2\pi n^3$. The ionization of Rydberg states of the hydrogen atom by microwave radiation was observed in the experiments of Sauer *et al.*[8] The experimental data on the threshold field strength agree with the above theoretical estimate.

The diffusion equation for the quantum-number distribution function $f(n)$ in the excitation of the electron is of the standard form

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial n}\left[D(n)\frac{\partial f}{\partial n}\right]. \tag{1}$$

The nonlinear diffusion coefficient is given by the following equation:[1]

$$D(n) = 0.27\frac{F^2}{\omega^{4/3}}n^3. \tag{2}$$

In one-dimensional diffusion the variance of the principal quantum number in the time Δt is proportional to this time interval:

$$\langle \Delta n^2 \rangle = 2D(\langle n \rangle)\Delta t. \tag{3}$$

Combining (2) and (3), we arrive at a law that governs the variation of the mean quantum number:

$$-\frac{d}{dt}\left(\frac{1}{\langle n \rangle}\right) = 0.27\frac{F^2}{\omega^{4/3}} = \mathrm{const}. \tag{4}$$

This law yields the time it takes an electron to reach the boundary of the continuous spectrum ($n \to \infty$), i.e., the ionization time:

$$t_{\mathrm{ion}} = \frac{\omega^{4/3}}{0.27F^2 n_0}, \tag{5}$$

where n_0 is the principal quantum number of the initial Rydberg state of the atom. For field strengths $F \sim 1/40n^4$ and frequencies $\omega \sim 1/n^3$, we find that although analytically the
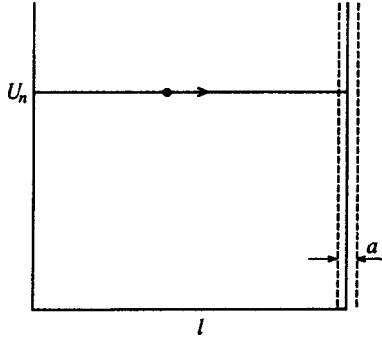
FIG. 1. Rectangular box with infinitely high walls that contains a classical particle. The right-hand wall oscillates with a small amplitude (the Ulam problem).

ionization time is of order of the Keplerian revolution period of the electron, $2\pi n^3$, numerically it is much longer than this period.

The aim of the present paper is to find an analogy between the classical ionization of atomic Rydberg states by a microwave field and the well-known Ulam problem[9] (the details will follow). We shall see that the problem is that as the principal quantum number n increases, the conditions for diffusive ionization deteriorate due to a decrease in the frequency of Keplerian revolution of the electron along the corresponding orbit as compared to the frequency of the microwave field. Hence the electron may never reach the boundary of the continuous spectrum or it may reach it and ''bounce back.'' We study the dynamics of electron motion using an analogy with the Ulam problem for short times, where the diffusion equations for both problems are mathematically similar, as we shall shortly see. Of course, extrapolation of this analogy to long times is an hypothesis that our calculations in the case of atoms do not corroborate directly. In view of the mathematical difficulties involved in solving the problem in the case of atoms, we attempted to discern how the diffusion process is terminated in the Ulam problem for very long times. It would seem that a quasistationary regime should set in. However, numerical calculations produced quite unexpected results, which we believe are the most interesting aspect of the present investigation.

2. THE ULAM MAP

As noted earlier, from the mathematical viewpoint it is convenient to reduce the problem of the diffusive ionization of atoms by a microwave field to what is known as the Ulam map[9] in such a way that the diffusion equations in both cases would have the same appearance, although they describe different physical processes. This map originated from Fermi's idea (1949) that a cosmic particle collides more often with interstellar clouds that are moving toward it than with those moving away from it. For this reason, on the average the particle is diffusively accelerated after a large number of collisions.

The Ulam problem examines the classical motion of a particle moving with a high velocity inside a one-dimensional, infinitely high rectangular potential well, or box (see Fig. 1). One of the walls oscillates with time (to be

definite, we have chosen it to be the right-hand wall). It is this oscillatory motion that ensures the diffusive Fermi acceleration of the particle.

Let l be the width of the box. The velocity V of the right-hand wall is assumed to vary according to a harmonic law with a frequency ω and an amplitude V_0: $V(t) = V_0 \sin(\omega t + \varphi)$. The box width is assumed to be large compared to the amplitude $a = V_0/\omega$ of oscillations of the right-hand wall: $l \gg a$. We also assume that at all times the particle velocity u is high compared to the oscillation velocity of the right-hand wall: $u \gg V_0$. This allows us to exclude double collisions of the particle with the right-hand wall of the box (such collisions would occur if the particle velocity is low).

The law of momentum conservation in the nth elastic collision of the particle with the oscillating wall is simple:

$$u_{n+1} = u_n + 2V_0 \sin(\omega t_n + \varphi), \tag{6}$$

where t_n is the time of the nth collision took place, u_n is the particle velocity just before the nth collision, and u_{n+1} is the particle velocity prior to the next, $(n+1)$th, collision (after the particle is reflected from the stationary left-hand wall of the box).

The time of free particle motion between two successive particle collisions with the oscillating wall is

$$t_{n+1} - t_n = \frac{2l}{u_{n+1}}. \tag{7}$$

Introducing the dimensionless phase $\psi_n \equiv \omega t_n + \varphi$, the dimensionless velocity $U_n \equiv u_n/2V_0$, and the dimensionless parameter $M \equiv l\omega/2\pi V_0 \gg 1$, we arrive at the Ulam map in dimensionless variables (for details see Ref. 10):

$$U_{n+1} = U_n + \sin\psi_n,$$

$$\psi_{n+1} = \psi_n + \frac{2\pi M}{U_{n+1}} \bmod(2\pi). \tag{8}$$

This is a particular case of the Poincaré map.

Note that if the particle velocity is

$$U_n = \frac{M}{m}, \tag{9}$$

where m is an integer, then, according to the first relation in (8), the particle velocity begins to increase or decrease regularly by the same amount $\sin\psi_n$ in each collision until (9) is violated substantially. At other points the variation of the particle velocity is irregular. The points indicated are called periodic points of map (8) (see Ref. 10).

The phase expansion modulus (according to Chirikov[11]) is defined as follows:

$$K \equiv \left| \frac{(\psi_{n+1} - \psi_n) - (\psi_n - \psi_{n-1})}{\psi_n - \psi_{n-1}} \right|. \tag{10}$$

Dynamic chaos, i.e., irregular variation of the particle velocity with the passage of time, sets in if $K > 1$, i.e., when the phase difference changes not very strongly, but irregularly in each collision. Introducing the average estimates $|\psi_n - \psi_{n-1}| \sim \pi/2$ and $|\sin\psi_n| \sim 1$, we see that (8) and (10) yield

$$K \approx \frac{2\pi M}{(\pi/2)U_n^2}|U_{n+1} - U_n| \approx \frac{4M}{U_n^2}. \tag{11}$$

The condition $K > 1$ means that in dynamic chaos we have $M \gg U_n \gg 1$.

3. DIFFUSIVE INCREASE IN VELOCITY

According to (6), after each collision the particle velocity u_n changes very little in comparison to the velocity itself. A large change in velocity can occur only after a very large number of collisions. Let us introduce the concepts of running velocity and mean velocity (over many collisions):

$$u_{n+1} - u_n = u - \langle u \rangle = \Delta u.$$

Averaging the square of this difference in a single collision with allowance for (6) and (7), we get

$$\langle \Delta u^2 \rangle = 4V_0^2 \langle \sin^2 \psi_n \rangle = 2V_0^2 = 2V_0^2 \frac{\Delta t}{2l/u}. \tag{12}$$

This relationship, which is valid for a single collision, can be generalized to the case of any number of collisions and the corresponding time interval Δt. According to the law of one-dimensional diffusion, we have $\langle \Delta u^2 \rangle = 2D\Delta t$. Hence, due to (12), the variable diffusion coefficient increases linearly with velocity:

$$D(u) = \frac{V_0^2}{2l}u. \tag{13}$$

Plugging this expression into the diffusion equation (1), multiplying the result by the current velocity u, and integrating the product over all velocities, we obtain, after integrating by parts twice,

$$\frac{d}{dt}\langle u \rangle = \frac{V_0^2}{2l}, \quad \langle u \rangle = u_0 + \frac{V_0^2}{2l}t. \tag{14}$$

This equation describes uniformly accelerated particle motion in the Ulam problem.

Equation (14) has the same form as Eq. (4) in the problem of the diffusive ionization of a Rydberg atom if we make the following substitutions:

$$\langle u \rangle \rightarrow u_0 - \frac{1}{\langle n \rangle} + \frac{1}{n_0}, \quad \frac{V_0^2}{2l} \rightarrow 0.27\frac{F^2}{\omega^{4/3}}. \tag{15}$$

Thus, we have found the Ulam map (8) with parameters (15) for the problem of the diffusive ionization of a Rydberg atom.

4. REGULAR MOTION

In the Ulam problem, the particle velocity increases only to a certain limit. Dynamic chaos ceases when the phase expansion modulus K is smaller than unity. Equation (11) then suggests that this occurs when $U_{max} = 2\sqrt{M}$. The numerical calculations of Lieberman and Lichtenberg[12] yielded a somewhat more accurate estimate:

$$U_{max} = 2.5\sqrt{M}. \tag{16}$$

This corresponds to the condition where there is no overlap of neighboring resonances.[11]

After the particle velocity reaches its maximum value, it remains constant in the course of several thousand collisions, varying only slightly. When the velocity value corresponds to a periodic point [see Eq. (9)] with an index

$$m = \frac{M}{U_n} = \frac{U_n}{6.25} \gg 1, \tag{17}$$

we have, in accordance with the ideas discussed in Sec. 2, a small upward or downward jump in velocity to a new value that only slightly differs from the old (by a fraction of a percent), after which the velocity stabilizes at the new value with very weak oscillations.

In our numerical calculations we give the results for $M = 10\,000$, $U_0 = 25$, so that, according to (11), the initial phase expansion modulus K_0 amounts to 64. According to (16), the maximum velocity is $U_{max} = 250$. By varying the phase φ of the oscillations of the right-hand wall of the box we found that the diffusion regime described in Sec. 2 is actually achieved only for a negligibly small number of phases. In most cases the particle velocity varies irregularly, and the velocity exceeds the initial value of the particle velocity only slightly. Here, in the course of only several thousand collisions, instead of increasing diffusively, the particle velocity drops abruptly to zero. After this the count is terminated because double collisions of particle with the right-hand wall of the box become possible, in which case the original simplified Ulam map (8) becomes invalid.

By specially selecting the phase φ we were able to establish the rare cases where the diffusion regime sets in. One is illustrated in Fig. 2, which depicts the dependence of the dimensionless particle velocity U_n on the number of collisions n. Clearly, there is a diffusive increase in particle velocity with time, the velocity finally reaching a plateau that corresponds to the maximum velocity (16), as predicted by theory. The diffusion regime continues only for approximately 30 000 collisions, while the plateau extends to several hundred thousand collisions and more. An analytic expression for the diffusion curve can be obtained directly from (14):

$$\frac{dU_n}{dn} = \frac{1}{4U_n}. \tag{18}$$

The curve in Fig. 2 is reproduced by (18) fairly accurately.

5. ELECTRON DIFFUSION AMONG THE RYDBERG ORBITS OF AN ATOM

The first condition in (15) suggests the following. When the maximum velocity u_{max} in the Ulam problem is reached, does the corresponding value of n_{max} become infinite? If the answer is yes, diffusive ionization has been achieved. Otherwise, n_{max} remains finite and can be found by the formula
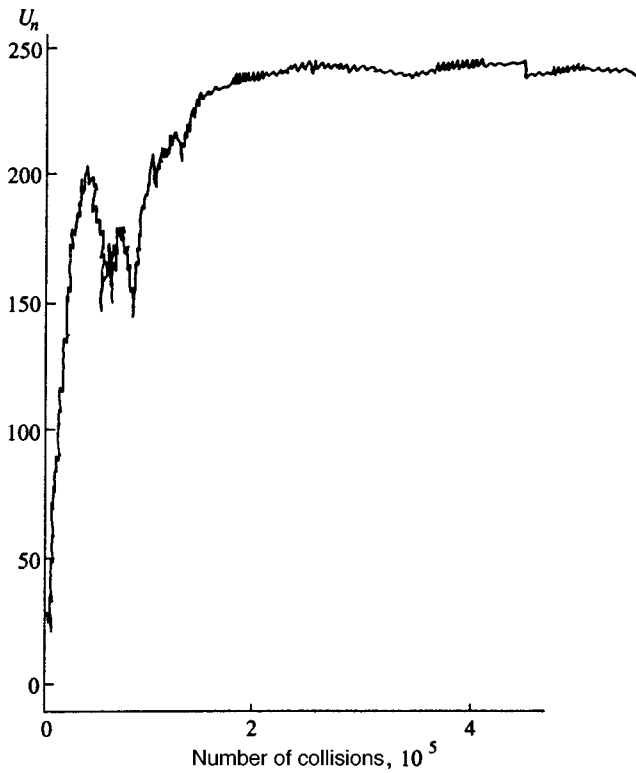
$$\frac{1}{n_{max}} = \frac{1}{n_0} - u_{max} \tag{19}$$

FIG. 2. Dependence of the dimensionless particle velocity in the Ulam problem on the number of collisions in the diffusion regime.

FIG. 3. Drop in particle velocity in the Ulam problem after a large number of collisions.

(here we have neglected the initial velocity u_0 in comparison to u_{max}). In this case diffusive excitation of the atom to a certain level would occur instead of diffusive ionization. From (15) and (16), for M we find that

$$u_{max} = 5.4\pi M^{3/2} \frac{F^2}{\omega^{7/3}}. \tag{20}$$

Plugging in the above values $M = 10\,000$, $F = 1/40n_0^4$, and $\omega = 2/n_0^3$, we find that $u_{max} = \text{const}/n_0$, where const $\sim 10^3$. According to (19), this means that diffusive ionization is sure occur in at such field strengths and frequencies.

Another problem arises when the observation times are long. The fact that an electron has reached the continuous spectrum does not necessarily mean ionization, since the microwave field continues to act on the atom and may return the electron to bound states. A similar situation indeed occurs in the tunneling ionization of atoms by strong low-frequency laser light.[13] An electron that has left an atom may return to the atom during a half-period of the field or even go back to its initial bound state after emitting a spontaneous high-frequency photon.

In the Ulam problem this scenario is formulated in the following way. The stationary value of the maximum velocity of a particle in a box with an oscillating wall is either maintained indefinitely or changes radically after a fairly large number of collisions. But what actually happens?

In the above example (see Fig. 2) we established numerically that the maximum quasistationary value of particle velocity obtained $U_n = 250$ is maintained for as many as three million collisions. However, after that the velocity's behavior changes radically (Fig. 3). In the interval between 3

$\times 10^6$ and 3.3×10^6 collisions the particle velocity drops practically to zero! Such behavior is observed for all initial conditions for which the diffusion process takes place and the maximum velocity is attained. For an atom this means that the electron may return to bound states from the continuous spectrum (if the assumption that the analogy between the two problems extends to times longer than diffusion times is valid) and that the particle velocity rapidly decreases after a certain moment in time.

In conclusion we note that numerical solutions of the Ulam problem for long times suggest that two facts may hinder the classical diffusive ionization of a Rydberg atom:

1) for most classical paths, there is no diffusion of a highly excited electron among Rydberg orbits, and only stochastic variation of the electron energy near its initial value occurs;

2) even if diffusion occurs, during the long time that the system is subjected to microwave radiation the electron returns to its initial bound state, although intermediate values of the electron energy belong to the continuous spectrum; actually, diffusive ionization takes place only if the pulse of microwave radiation ends at the necessary point in time [of the order given in (5)].

We stress once again that outside the diffusion region there is no rigorous justification for extrapolating the analogy between the problem of hydrogen-atom ionization and the Ulam problem. We simply assume that such extrapolation is valid.

*⁾E-mail: krainov@theory.mipt.ru

[1] N. B. Delone and V. P. Krainov, *Multiphoton Processes in Atoms*, Springer, Berlin-Heidelberg (1994).

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

[3] G. Gasati, B. V. Chirikov, D. L. Shepelyansky, and I. Guarneri, Phys. Rep. **154**, 78 (1987).

[4] N. B. Delone, B. A. Zon, and V. P. Kraĭnov, Zh. Éksp. Teor. Fiz. **75**, 445 (1978) [Sov. Phys. JETP **48**, 223 (1978)].

[5] B. I. Meerson, E. A. Oks, P. V. Sasorov, JETP Lett. **29**, 72 (1979).

[6] R. V. Jensen, S. M. Susskind, and M. M. Sanders, Phys. Rev. Lett. **62**, 1476 (1989).

[7] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Systems*, Springer, Berlin (1957); Fizmatgiz, Moscow (1960).

[8] B. E. Sauer, S. Yoakum, L. Moorman, P. M. Koch, D. Richards, and P. A. Dando, Phys. Rev. Lett. **68**, 468 (1992).

[9] G. M. Zaslavskiĭ and R. Z. Sagdeev, *Introduction to Nonlinear Physics* [in Russian], Nauka, Moscow (1988).

[10] A. J. Lichtenberg and M. A. Lieberman, *Regular and Stochastic Motion*, Springer, Berlin (1983); Mir, Moscow (1984).

[11] B. V. Chirikov, Phys. Rep. **52**, 265 (1979).

[12] M. A. Lieberman and A. J. Lichtenberg, Phys. Rev. A **5**, 1852 (1972).

[13] P. B. Corkum, Phys. Rev. Lett. **71**, 1994 (1993).

Exciton magnetotransport in two-dimensional systems: weak-localization effects

P. I. Arseev

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

A. B. Dzyubenko*)

Institute of General Physics, Russian Academy of Sciences, 117942 Moscow, Russia
(Submitted 30 October 1997)
Zh. Éksp. Teor. Fiz. **114**, 359–378 (July 1998)

The paper considers the effect of a magnetic field B on the transport of neutral composite particles, namely excitons, in weakly disordered two-dimensional (2D) systems. In the case of classical transport (when the interference of different paths is neglected), the magnetic field suppresses exciton transport, and the static diffusion constant $D(B)$ monotonically drops with B. When quantum-mechanical corrections due to weak localization are taken into account, $D(B)$ becomes a nonmonotonic function of B. In weak magnetic fields, where the magnetic length is much larger than the exciton Bohr radius, $\ell_B = (\hbar c/eB)^{1/2} \gg a_B = \varepsilon \hbar^2/\mu e^2$, a positive magnetodiffusion effect is predicted, i.e., the exciton mobility should increase with B. © *1998 American Institute of Physics.* [S1063-7761(98)02407-X]

1. INTRODUCTION

In two-dimensional (2D) systems, all states are localized, no matter how weak the disorder is.[1–3] This phenomenon is universal for all processes of wave propagation and is associated with the constructive interference of paths subjected to the time-reversal operation. For particles with nonzero mass, this is a quantum-mechanical effect, which cannot be described in terms of classical mechanics. The quantum statistics of particles in this case does not play a crucial role (see, for example, works on the weak localization of phonons[4] and light[5]). While localization has been thoroughly investigated on the base of the one-particle approach, many questions concerning the interplay between localization and Coulomb effects remain unanswered. The variety of physical situations requires the application of different techniques suitable for describing the respective class of phenomena. For instance, it was predicted by the weak-localization theory that the electron-electron interaction should weaken the interference effects and lead to a higher conductivity (see, e.g., the review by Lee and Ramakrishnan[3]). A numerical calculation for two interacting electrons in a random potential has also predicted a correlated-propagation length larger than the localization length of an isolated particle.[6] The issue discussed in this paper, namely the propagation of an exciton, which consists of an electron and a hole interacting with one another, in a magnetic field and in a random potential is also one of the aspects of the general problem. The weak localization of excitons in the absence of a magnetic field was investigated previously.[7]

Introduction of a magnetic field B generates new features in the physical picture of the weak localization of electrons. Formally, a magnetic field B breaks time-reversal symmetry. The physical consequence is negative magnetoresistance in electron systems.[8,9] This effect is caused by the fact that charged particles acquire different phase shifts in magnetic fields when they travel along closed paths in opposite directions.[10] As a result, the field B breaks the constructive interference between time-reversed paths and thereby suppresses the weak localization of electrons. If we take into account the electron spin, four different channels for interference between two electronic waves are possible: one of them is singlet ($S=0$), and three are triplet ($S=1$, $S_z = \pm 1,0$). The interference in the triplet (singlet) channels gives a positive (negative) contribution to the conductivity.[9,11] Fast spin-flip processes can change the relation between the contributions of the singlet and triplet channels, thus resulting in either negative or positive magnetoresistance. Various mechanisms of spin-orbit coupling that are important for electrons in quasi-two-dimensional semiconductor quantum wells and heterojunctions were taken into account in Ref. 12. Note also that in systems with strongly localized electron states (the hopping conductivity regime) magnetically induced changes in the phase relations between different transition amplitudes can lead to either negative or positive magnetoresistance.[13,14]

An important question in the case of excitons, which are composite and, as a whole, electrically neutral particles, is whether the time-reversal symmetry for an $e-h$ pair is broken by a magnetic field B. One may assume that the $t \rightarrow -t$ symmetry for a pair should be broken by magnetic field since it is broken for an electron or hole taken separately, and this is true in a general case. There is, however, an exceptional case. Consider the Hamiltonian

$$H = \frac{1}{2m_e}\left(-i\hbar\nabla_e + \frac{e}{c}\mathbf{A}_e\right)^2 + \frac{1}{2m_h}\left(-i\hbar\nabla_h - \frac{e}{c}\mathbf{A}_h\right)^2 + U_{eh}(\mathbf{r}_e - \mathbf{r}_h) + V_e(\mathbf{r}_e) + V_h(\mathbf{r}_h), \quad (1)$$

which describes the motion of an $e-h$ pair in a uniform magnetic field B and external (random) potentials V_e and V_h. When the particle masses are equal, $m_e = m_h$, and the

scattering potentials are identical, $V_e \equiv V_h$, the e and h components transform into one another after time reversal.[1)] In this case Hamiltonian (1) commutes with the time-reversal operator:

$$[H, \hat{T}] = 0. \qquad (2)$$

This means that the $t \rightarrow -t$ symmetry is not broken, and 2D excitons should remain localized even in the presence of a magnetic field.

In the general case, one should analyze how a magnetic field B suppresses the weak localization of excitons, which are electroneutral as a whole, and how their internal structure manifests itself. In a magnetic field, the center-of-mass motion and relative motion of an $e - h$ pair are coupled. Therefore, the scattering of an exciton as a whole is affected by the magnetic field B and the internal $e - h$ interaction.

Recently, the transport of quasi-two-dimensional excitons in quantum wells under a magnetic field has attracted a lot of attention on the part of experimentalists (see Refs. 16–18 and references therein). Butov *et al.*[17] reported intriguing low-temperature anomalies in exciton magnetotransport. In particular, they found that the exciton diffusion constant D is a nonmonotonic function of B and increases considerably in the range of intermediate fields $B \approx 6$ T. This fact was interpreted as evidence in favor of Bose–Einstein condensation and a manifestation of the superfluidity of excitons. It seems interesting to check whether the localization effects of excitons can give rise to such features of the $D(B)$ curve in the normal phase. In this paper we investigate theoretically the magnetotransport of 2D excitons in the presence of weak disorder in the limiting case when the magnetic length is much larger than the Bohr radius of the exciton, $\ell_B = (\hbar c / eB)^{1/2} \gg a_B = \varepsilon \hbar^2 / \mu e^2$. Fields that satisfy this condition will be dubbed weak. Taking these results together with those for the cases of classical[19] and quantum transport[20,21] in the opposite limit, $\ell_B \ll a_B$, we shall suggest an approximate form of the diffusion constant D as a function of B at all fields, including the intermediate range, where $\ell_B \sim a_B$. A brief account of some results of this paper was reported previously.[21]

2. EXCITON TRANSPORT IN A MAGNETIC FIELD *B*

2.1. Problem statement

In the weak-localization regime, the interaction with an isolated defect does not give rise to a bound state, and localization is possible only at large distances due to the interference of scattered waves. This localization regime takes place in the case of weak scattering, in which

$$\gamma_0(p) \ll \epsilon(p), \qquad (3)$$

where γ_0 is the damping coefficient (the reciprocal of the momentum relaxation time) of an exciton with energy $\epsilon(p)$. The scatterers in this case are the random potentials $V_e(\mathbf{r})$ and $V_h(\mathbf{r})$ in Eq. (1), which act on the electron and hole separately. They can be, e.g., potentials generated by charged impurities, effective potentials due to irregularities on quantum-well interfaces, etc. At low temperatures, the dominant scattering mechanism in quantum wells is that due to
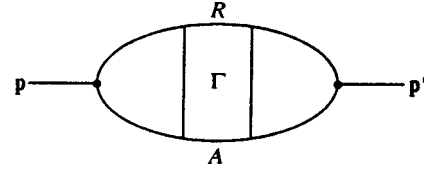
FIG. 1. Diagrammatic representation of Eq. (5): Γ is the irreducible vertex corresponding to scattering by a random potential; the lines labeled by A and R represent the advanced and retarded propagators G^A and G^R of excitons averaged with respect to disorder in a magnetic field. The current vertices in the diagram correspond to the exciton center-of-mass velocity $V(\mathbf{p})$.

irregularities on the interfaces ("surface roughness"), and our attention will be mainly focused on this mechanism. In the case of 2D excitons in a quantum well of width d in the presence of interface irregularities with a characteristic amplitude Δ and a correlation length Λ (see Ref. 19 and references therein), there is a characteristic exciton momentum defined as follows:[2)]

$$p_{\min} \sim \frac{1}{a_B} \left(\frac{\Delta \Lambda a_B}{d^3} \right), \qquad \ell_B \gg a_B. \qquad (4)$$

For long-wavelength excitons with momenta $p < p_{\min}$ we have $\gamma_0 \gtrsim \epsilon$, and such excitons are strongly localized. An analysis of the strong localization of composite particles in a magnetic field is beyond the scope of this paper. However, if the parameter $\Delta \Lambda a_B / d^3 \ll 1$ is sufficiently small, the range of strong localization of excitons is narrow compared with the characteristic momentum a_B^{-1} in our problem (ℓ_B^{-1} in high magnetic fields), and the theory developed in this paper has a region of applicability.

It is essential that the scattering of two-particle $e - h$ states can be described diagrammatically in terms of effective one-particle (exciton) scattering (this approach is justified in Appendix A). This approximation allows us to treat excitons at low densities as Bose-particles. Their internal structure manifests itself in changes in the effective scattering potential $V_{\mathbf{p}, \mathbf{p}'}$ and the dispersion law $\epsilon(p)$ due to the magnetic field B (see Sec. 2.2). The potentials $V_e(\mathbf{r})$ and $V_h(\mathbf{r})$ may be uncorrelated, for example, when the particles e and h are spatially separated,[19] or fully correlated, e.g., when both particles are in the same spatial domain. We assume that the distribution of random fields is Gaussian and use standard diagram techniques[23] with two-particle [retarded (R) and advanced (A)] excitonic propagators in a magnetic field averaged with respect to disorder: $G_\omega^{R(A)}(p) = [\omega - \epsilon(p) \pm i \gamma_0(p)]^{-1}$ (see, e.g., Ref. 20).

For the case of elastic scattering, one can introduce a diffusion constant $D(\omega, \epsilon)$ for excitons with a given energy ϵ at a frequency ω, which can be derived from the expression for the generalized "conductivity" (Fig. 1) $\sigma(\omega, \epsilon) = D(\omega, \epsilon) \mathcal{N}(\epsilon)$:

$$\sigma(\omega, \epsilon) = \frac{1}{2\pi} \int d\mathbf{p} \int d\mathbf{p}' \langle\langle V_x(\mathbf{p}) G^R(\mathbf{p}, \mathbf{p}', \epsilon + \omega)$$

$$\times G^A(\mathbf{p}', \mathbf{p}, \epsilon) V_x(\mathbf{p}') \rangle\rangle, \qquad (5)$$

where $\mathbf{V}(\mathbf{p})$ is the velocity of the exciton's center of mass, $\mathcal{N}(\epsilon)$ is the exciton density of states, and $\langle\langle \ldots \rangle\rangle$ denotes

averaging with respect to disorder. Note that the diffusion constant $D(\omega,\epsilon)$ is a parameter included in the ''diffusion'' pole of the exciton ''density-density'' correlation function and, therefore, determines the propagation characteristics of particles with a given energy ϵ in the long-wavelength limit. The localization of quantum states with energy ϵ means that the diffusion constant as a function of frequency, $D(\omega)$, tends to zero in the static limit, $\omega \to 0$. If inelastic scattering it taken into account, $D(\omega)$ turns out to be finite (see, for example, the review by Lee and Ramakrishnan[3]). The static limit $D(\epsilon) = D(\epsilon,\omega = 0)$ is determined by the time τ_ϕ of the loss of phase coherence (dephasing). The *total* static diffusion constant $D = D(T)$ corresponding to fluctuations in the density of excitons with all allowed energies can be obtained from the microscopic values of $D(\epsilon)$ using the generalized Einstein relation:

$$D = \frac{\int d\epsilon \, \mathcal{N}(\epsilon)[-\partial f/\partial\epsilon] D(\epsilon)}{\int d\epsilon \, \mathcal{N}(\epsilon)[-\partial f/\partial\epsilon]}, \qquad (6)$$

where $f = f(\mu_X, T)$ is the distribution function and μ_X is the chemical potential of the excitons.

2.2. Effective scattering potential

The Hamiltonian H_0 of relative motion of an $e-h$ pair with center-of-mass momentum $\hbar\mathbf{p}$ (where $\mathbf{p}$ is the wave vector) in a perpendicular magnetic field B has the form[24,25]

$$H_0 = -\frac{\hbar^2}{2\mu}\nabla_{\mathbf{r}}^2 - \frac{i\hbar eB}{2c}\left(\frac{1}{m_h} - \frac{1}{m_e}\right)(\mathbf{r}\times\nabla_{\mathbf{r}})_z + \frac{e^2 B^2}{8\mu c^2}r^2$$
$$+ \frac{e\hbar}{Mc}\mathbf{B}\cdot(\mathbf{r}\times\mathbf{p}) - \frac{e^2}{\varepsilon|\mathbf{r}|}, \qquad (7)$$

where $\mathbf{r} = \mathbf{r}_e - \mathbf{r}_h$ is the relative $e-h$ coordinate, and $\mu^{-1} = m_e^{-1} + m_h^{-1}$. In writing this expression, we have utilized the existence of an exact integral of motion, namely the magnetic center-of-mass momentum[24] defined by the operator

$$\hbar\hat{\mathbf{p}} = -i\hbar\nabla_{\mathbf{R}} - \frac{e}{c}\mathbf{A}(\mathbf{r}),$$

where $\mathbf{R} = (m_e\mathbf{r}_e + m_h\mathbf{r}_h)/M$ is the center-of-mass coordinate, $M = m_e + m_h$, and the vector potential is taken in the symmetrical gauge $\mathbf{A} = \mathbf{B}\times\mathbf{r}/2$. The exciton wave function in a magnetic field B has the form

$$\Psi_{\mathbf{p}}(\mathbf{R},\mathbf{r}) = \exp\left\{ i\mathbf{R}\left[\mathbf{p} + \frac{e}{c}\mathbf{A}(\mathbf{r})\right]\right\}\Phi_{\mathbf{p}}(\mathbf{r}). \qquad (8)$$

An important point is that the wave function $\Phi_{\mathbf{p}}$ of relative motion of an $e-h$ pair depends on the center-of-mass momentum $\mathbf{p}$,[24] i.e., the relative motion and the center-of-mass motion are coupled. The scattering matrix elements between the exciton states with the center-of-mass momenta $\mathbf{p}$ and $\mathbf{p}'$ in an external potential $\hat{V} = V_e(\mathbf{r}_e) + V_h(\mathbf{r}_h)$ have the form (see Appendix A)

$$V_{\mathbf{p},\mathbf{p}'} = \langle\Psi_{\mathbf{p}}|\hat{V}|\Psi_{\mathbf{p}'}\rangle. \qquad (9)$$

In this work we use an approximation that ignores transitions to excited states of internal motion.[19,20] In the weak-field

limit, $\ell_B \gg a_B$, the problem can be treated analytically.[3)] We calculate the ground-state wave function $\Phi_{\mathbf{p}}(\mathbf{r})$ in a magnetic field using perturbation theory with respect to terms containing the magnetic field in the Hamiltonian (7) of the relative motion of an $e-h$ pair, and then we obtain the scattering matrix elements $V_{\mathbf{p},\mathbf{p}'}$. They can be expressed as

$$V_{\mathbf{p},\mathbf{p}'} = F^e_{\mathbf{p},\mathbf{p}'}\widetilde{V}_e(\Delta\mathbf{p}) + F^h_{\mathbf{p},\mathbf{p}'}\widetilde{V}_h(\Delta\mathbf{p}), \qquad (10)$$

where $\widetilde{V}_j(\mathbf{p})$ are two-dimensional Fourier transforms of the potentials $V_j(\mathbf{r})$ $(j=e,h)$, $\Delta\mathbf{p} = \mathbf{p}' - \mathbf{p}$ is the momentum transfer,

$$F^{e(h)}_{\mathbf{p},\mathbf{p}'} = \int d\mathbf{r}\,\Phi_{\mathbf{p}}^*(\mathbf{r})\Phi_{\mathbf{p}'}(\mathbf{r})\exp\left\{ \pm i\frac{m_{h(e)}}{M}(\mathbf{p}' - \mathbf{p})\mathbf{r}\right\} \qquad (11)$$

are the form factors related to the wave function of the internal motion of the exciton. In the weak-field limit, we must calculate the wave functions up to the second order in B and then substitute them into Eqs. (11) and (10) (see Appendix B). Note that the exponential function in Eq. (11) can be expanded in powers of its argument when p, $p' \ll a_B^{-1}$, and only terms of the lowest orders need be included. The limitation on the momenta is essential if we do not take into consideration transitions to excited states. Indeed, if p, $p' \sim a_B^{-1}$, the exciton kinetic energy is sufficient for transitions to excited states of internal motion, which are excluded from our analysis.

Taking the essential terms of up to the second order in B and the lowest orders in pa_B of interest to us, we obtain

$$V_{\mathbf{p},\mathbf{p}'} = \bar{V}_e(\Delta\mathbf{p})\left[1 + \beta_e(\Delta\mathbf{p})^2 a_B^2\left(\frac{a_B}{\ell_B}\right)^4 - i\alpha_e[\mathbf{p}\mathbf{p}']_z a_B^2 \right.$$
$$\left. \times\left(\frac{a_B}{\ell_B}\right)^2\right] + \bar{V}_h(\Delta\mathbf{p})\left[1 + \beta_h(\Delta\mathbf{p})^2 a_B^2\left(\frac{a_B}{\ell_B}\right)^4 \right.$$
$$\left. + i\alpha_h[\mathbf{p}\mathbf{p}']_z a_B^2\left(\frac{a_B}{\ell_B}\right)^2\right]. \qquad (12)$$

Here $\bar{V}_i(\Delta\mathbf{p}) = \widetilde{V}_i(\Delta\mathbf{p})F^i_{\mathbf{p},\mathbf{p}'}(B=0)$,

$$F^{e(h)}_{\mathbf{p},\mathbf{p}'}(B=0) = \left\{ 1 + \frac{1}{16}\left[\frac{m_{h(e)}(\mathbf{p}-\mathbf{p}')a_B}{M}\right]^2\right\}^{-3/2}$$

is the form factor corresponding to the ground-state wave function of the 2D-exciton at $B=0$.

An important point is that time-reversal symmetry is broken for this effective scattering potential:

$$V_{\mathbf{p},\mathbf{p}_1} \neq V_{-\mathbf{p}_1,-\mathbf{p}}, \qquad (13)$$

the only exception being the case of $V_e = V_h$ and $m_e = m_h$ [see Eq. (2)]. Equation (12) contains the dimensionless constants

$$\beta_{e(h)} = -\frac{\mu^2}{2M^2}\frac{\hbar^4}{\mu^2 a_B^6}\sum_n{}'\frac{|\langle 0|x|n\rangle|^2}{(\epsilon_0 - \epsilon_n)^2}$$
$$+ \frac{m_{h(e)}^2}{8M^2}\frac{\hbar^2}{\mu a_B^6}\sum_n{}'\frac{|\langle 0|r^2|n\rangle|^2}{\epsilon_n - \epsilon_0} \qquad (14)$$
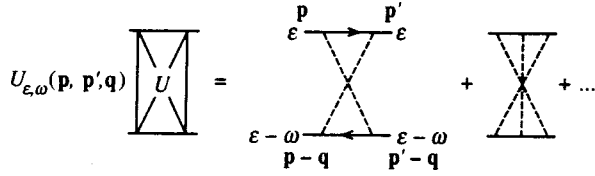
FIG. 2. Sum of maximally crossed diagrams $U_{\epsilon,\omega}(\mathbf{p},\mathbf{p}',\mathbf{q})$. The upper (lower) line corresponds to the retarded (advanced) propagator $G^{R(A)}$ of an exciton averaged with respect to disorder.
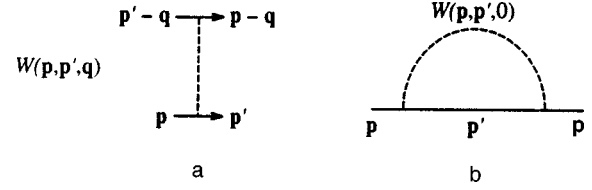

FIG. 3. (a) Simplest impurity vertex $W(\mathbf{p},\mathbf{p}',\mathbf{q})$ [Eq. (19)]; (b) diagram of lowest order for the self-energy part of the excitonic propagator. The dashed line correspond to the correlation function $W(\mathbf{p},\mathbf{p}',0)$.

and

$$\alpha_{e(h)} = -\frac{2m_{e(h)}}{M}\kappa, \quad \kappa = \frac{\hbar^2}{Ma_B^4}\sum_n{}' \frac{|\langle 0|x|n\rangle|^2}{\epsilon_0 - \epsilon_n}. \quad (15)$$

Here n denotes the exciton excited states. Exact calculations of the dimensionless constants α and β for a 2D Wannier–Mott exciton are given in Appendix B. Note that β_e, β_h >0 are positive; therefore, exciton scattering increases with B when $\ell_B \gg a_B$.

Using perturbation theory, one can also obtain the exciton density of states in a magnetic field. The exciton spectrum is given by the formula

$$\epsilon(p) = -\epsilon_0\left[1 - \left(\frac{l_2}{\ell_B}\right)^4\right] + \frac{\hbar^2 p^2}{2M}\left[1 - \kappa\left(\frac{a_B}{\ell_B}\right)^4\right], \quad (16)$$

where the parameter $l_2 = 3a_B/8$ determines the diamagnetic shift. The 2D-exciton density of states derives from the second term on the right-hand side of Eq. (16):

$$\mathcal{N}(\epsilon) = \frac{2M/\hbar^2}{1 - \kappa(a_B/\ell_B)^4}. \quad (17)$$

The exciton mass and, hence, the density of states $\mathcal{N}(\epsilon)$, increase with the magnetic field B. As will be shown below, it is this tendency that generally determines the change in the classical diffusion constant in weak magnetic fields.

2.3. Cooperon: weak-field limit

The approximation in which the complete vertex Γ (Fig. 1) is replaced by a sum of ladder diagrams (a diffuson) corresponds to the description of transport based on the Boltzmann equation (see, e.g., Refs. 23 and 26). This approximation yields the "classical" diffusion constant, which does not take into account the interference of different paths. If the random potential is weak, all other diagrams with crossed impurity lines have smallness[3] of order $\gamma/\epsilon \ll 1$. The only exception is the class of maximally crossed diagrams in the electron-hole channel,[2] which determines quantum weak-localization corrections to the diffusion constant. The complete sum of such diagrams (the cooperon) is shown in Fig. 2. The exceptional role of these diagrams is due to the following fact: when the total momentum $\mathbf{p}+\mathbf{p}'-\mathbf{q}\approx 0$, the Green's functions G^R and G^A for the maximally crossed diagrams are always grouped in pairs with close poles by virtue of momentum conservation. As a result, they provide a "resonant" contribution after integration.

The damping coefficient for an exciton with a momentum $\mathbf{p}$ is determined by the imaginary part of the self-energy part (Fig. 3b):

$$\gamma_0(p) = -\,\text{Im}\int \frac{d\mathbf{p}'}{(2\pi)^2}\frac{W(\mathbf{p},\mathbf{p}',0)}{\epsilon - \epsilon(p') + i\gamma_0(p')}. \quad (18)$$

Here $\epsilon(p)$ is the dispersion law (16), and $W(\mathbf{p},\mathbf{p}_1,\mathbf{q})$ is the correlation function of the scattering potential (Fig. 3a):

$$W(\mathbf{p},\mathbf{p}_1,\mathbf{q}) \equiv \langle\langle V_{\mathbf{p},\mathbf{p}_1}V_{\mathbf{p}_1-\mathbf{q},\mathbf{p}-\mathbf{q}}\rangle\rangle. \quad (19)$$

In the weak-field limit discussed in this paper, it has the form[4)]

$$W(\mathbf{p},\mathbf{p}',0) = B_{ee}(\Delta\mathbf{p})\left[1 + \left(\frac{a_B}{\ell_B}\right)^4(2\beta_e(\Delta\mathbf{p})^2 a_B^2 + \alpha_e^2\right.$$

$$\times (\mathbf{p}\times\mathbf{p}')_z^2 a_B^4\big] + B_{hh}(\Delta\mathbf{p})\left[1 + \left(\frac{a_B}{\ell_B}\right)^4\right.$$

$$\times (2\beta_h(\Delta\mathbf{p})^2 a_B^2 + \alpha_h^2(\mathbf{p}\times\mathbf{p}')_z^2 a_B^4)\big] + B_{eh}(\Delta\mathbf{p})$$

$$\times \left[1 + \left(\frac{a_B}{\ell_B}\right)^4((\beta_e+\beta_h)(\Delta\mathbf{p})^2 a_B^2 - \alpha_e\alpha_h\right.$$

$$\times (\mathbf{p}\times\mathbf{p}')_z^2 a_B^4\big] + B_{he}(\Delta\mathbf{p})\left[1 + \left(\frac{a_B}{\ell_B}\right)^4\right.$$

$$\times ((\beta_e+\beta_h)(\Delta\mathbf{p})^2 a_B^2 - \alpha_e\alpha_h(\mathbf{p}\times\mathbf{p}')_z^2 a_B^4\big], \quad$$

$$(20)$$

where $B_{ij}(\mathbf{p}) = \langle\langle\bar{V}_i(\mathbf{p})\bar{V}_j(-\mathbf{p})\rangle\rangle$. As usual, if $\gamma_0 \ll \epsilon$, we have

$$\gamma_0(p) = \pi\mathcal{N}(\epsilon)\int\frac{d\phi_{p_1}}{2\pi}W(\mathbf{p},\mathbf{p}_1,0), \quad (21)$$

where $|\mathbf{p}_1|$ lies on the mass surface $\epsilon(p_1) = \epsilon$, so that only averaging over angles remains in Eq. (18). The effect of the magnetic field B on the damping coefficient $\gamma_0(p)$ can be approximately estimated as follows (we assume that all correlators of random fields B_{ij} are comparable):

$$\gamma_0(p) \approx \gamma_0 \frac{1 + 4(\beta_e+\beta_h)(pa_B)^2(a_B/\ell_B)^4}{1 - \kappa(a_B/\ell_B)^4}, \quad (22)$$

where γ_0 is the damping coefficient in a zero magnetic field. The numerator on the right-hand side of Eq. (22) contains the additional small parameter $(pa_B)^2 \ll 1$ in comparison with

the denominator. This means that the main effect of the magnetic field B is due to the growing exciton density of states (increase in exciton mass) with increasing B [see Eq. (16)], whereas the changes in the scattering matrix elements play a minor role.

In the weak-field limit, as in the case of strong magnetic fields,[20] the diffusion pole in the cooperon is absent owing to the broken time-reversal symmetry for the effective potential (12). Let us prove this statement. As usual, it is convenient to write the equation for the cooperon U in variables $\mathbf{p}$, $\mathbf{p}'$, and $\mathbf{K} = \mathbf{p} + \mathbf{p}' - \mathbf{q}$, where $\mathbf{K}$ is the total (conserved) momentum, and $\mathbf{q}$ is the momentum corresponding to density fluctuations. For U we obtain the Bethe–Salpeter equation in the usual manner:

$$U_{\epsilon,\omega}(\mathbf{p},\mathbf{p}',\mathbf{K}) = U^0_{\epsilon,\omega}(\mathbf{p},\mathbf{p}',\mathbf{K})$$

$$+ \int \frac{d\mathbf{p}_1}{(2\pi)^2} \widetilde{W}(\mathbf{p},\mathbf{p}_1,\mathbf{K}) G^R_\epsilon(\mathbf{p}_1) G^A_{\epsilon-\omega}$$

$$\times (\mathbf{K}-\mathbf{p}_1) U_{\epsilon,\omega}(\mathbf{p}_1,\mathbf{p}',\mathbf{K}), \qquad (23)$$

where

$$U^0_{\epsilon,\omega}(\mathbf{p},\mathbf{p}',\mathbf{K}) = \int \frac{d\mathbf{p}_1}{(2\pi)^2} \widetilde{W}(\mathbf{p},\mathbf{p}_1,\mathbf{K}) G^R_\epsilon(\mathbf{p}_1) G^A_{\epsilon-\omega}$$

$$\times (\mathbf{K}-\mathbf{p}_1) \widetilde{W}(\mathbf{p}_1,\mathbf{p}',\mathbf{K}) \qquad (24)$$

and we have introduced the correlation function:

$$\widetilde{W}(\mathbf{p},\mathbf{p}_1,\mathbf{K}) \equiv \langle\langle V_{\mathbf{p},\mathbf{p}_1} V_{\mathbf{K}-\mathbf{p},\mathbf{K}-\mathbf{p}_1}\rangle\rangle. \qquad (25)$$

In contrast to the conventional theory, the system is characterized by two correlation functions W [Eq. (20)] and $\widetilde{W}$.[5] The difference between the correlation functions is caused by the broken time-reversal symmetry for the effective scattering potential (13). As a result, the terms with the vector product $(\mathbf{p}\times\mathbf{p}_1)_z$ in $\widetilde{W}(\mathbf{p},\mathbf{p}_1,\mathbf{K}=0)$ have signs opposite to those of the terms in $W(\mathbf{p},\mathbf{p}_1,\mathbf{q}=0)$.

In the limit of weak disorder, we have $G^R G^A \sim \delta(\epsilon(p) - \epsilon)$, so that the integration in Eq. (23) is reduced to averaging over angles. In the usual case the isotropic (with respect to $\mathbf{p},\mathbf{p}'$) part of $U_{\epsilon,\omega}(\mathbf{p},\mathbf{p}_1,\mathbf{K})$ diverges as $\mathbf{K},\omega \to 0$. This happens because the following relation holds:

$$\int \frac{d\mathbf{p}_1}{(2\pi)^2} W(\mathbf{p},\mathbf{p}_1,0) G^R_\epsilon(\mathbf{p}_1) G^A_\epsilon(-\mathbf{p}_1) = 1. \qquad (26)$$

Then it follows from Eqs. (23) and (26) (if $\widetilde{W} = W$) that $\int d\phi_{\mathbf{p}}\int d\phi_{\mathbf{p}_1} U_{\epsilon,\omega}(\mathbf{p},\mathbf{p}_1,0) \to \infty$. In the case under consideration, however, the isotropic part of U is finite in the limit $\mathbf{K},\omega \to 0$. In fact, using the identity $\widetilde{W} \equiv W + (\widetilde{W} - W)$, we obtain

$$\int \frac{d\mathbf{p}_1}{(2\pi)^2} \widetilde{W}(\mathbf{p},\mathbf{p}_1,0) G^R_\epsilon(\mathbf{p}_1) G^A_\epsilon(-\mathbf{p}_1) = 1 - \frac{\gamma_B}{\gamma_0}, \qquad (27)$$

where $\gamma_B(p) = \gamma_0(p) - \widetilde{\gamma}_0(p) \geq 0$,

$$\widetilde{\gamma}_0(p) = \pi\mathcal{N}(\epsilon) \int \frac{d\phi_{\mathbf{p}_1}}{2\pi} \widetilde{W}(\mathbf{p},\mathbf{p}_1,0). \qquad (28)$$

If, however, $\gamma_B \ll \gamma_0$, the isotropic part of U still makes the principal contribution. A solution of Eq. (23) in the region of low frequencies ω and small momenta $\mathbf{K}$, which is discussed in the paper, can be obtained using an expansion in terms of angular momenta, and was described in detail in Ref. 20 (see also Ref. 27). Ultimately, the expression for the cooperon has the form

$$U(\mathbf{K},\omega) = \frac{2\widetilde{\gamma}_0\gamma_0/\pi\mathcal{N}(\epsilon)}{D^c K^2 - i\omega + 2\gamma_B\gamma_0/\widetilde{\gamma}_0}. \qquad (29)$$

Here

$$D^c = p^2/4M^2\widetilde{\gamma}_{\mathrm{tr}}, \quad \widetilde{\gamma}_{\mathrm{tr}} = \gamma_0 - \widetilde{\gamma}_1 \geq 0, \qquad (30)$$

$$\widetilde{\gamma}_1 = 2\int \frac{d\phi_{\mathbf{p}}}{2\pi} \int \frac{d\phi_{\mathbf{p}_1}}{2\pi} (\hat{\mathbf{p}}\hat{\mathbf{K}})\widetilde{W}(\mathbf{p},\mathbf{p}_1,0)(\hat{\mathbf{p}}_1\hat{\mathbf{K}}), \qquad (31)$$

where $\hat{\mathbf{p}} = \mathbf{p}/|\mathbf{p}|$. The special feature of this solution is that it contains a finite dephasing time γ_B^{-1} for a neutral composite particle in a magnetic field B, and this dephasing time eliminates a singularity, namely the diffusion pole. Formally, this case is similar to that of electron scattering by magnetic impurities.[8,9,11,12]

In weak magnetic fields, γ_B can be estimated using explicit expressions for W and $\widetilde{W}$:

$$\gamma_B(p) \simeq (pa_B)^4\left(\frac{a_B}{\ell_B}\right)^4 \gamma_0(p). \qquad (32)$$

The emergence of the characteristic dephasing time $\tau_B = \hbar/\gamma_B$ estimated by Eq. (32) can be interpreted in qualitative terms as follows. An exciton acquires a random phase in a magnetic field only as a result of impurity scattering. When an exciton with momentum $\mathbf{p}$ is scattered by an impurity, its kinetic energy $E_{\mathrm{kin}} = \hbar^2 p^2/2M$ can be treated as a perturbation to the internal electron-hole motion with an energy $E_{\mathrm{exc}} = \epsilon_0$. This results in fluctuations in the mean square distance between e and h: $\langle\langle\Delta r^2\rangle\rangle \sim (E_{\mathrm{kin}}/E_{\mathrm{exc}})a_B^2$. This additional separation between the electron and hole orbits due to a scattering act leads to an increase in the magnetic flux passing "through" the exciton $\Delta\Phi \sim \langle\langle\Delta r^2\rangle\rangle B$, which corresponds to a (random) phase shift $\Delta\phi \sim \langle\langle\Delta r^2\rangle\rangle B/\Phi_0$ in the wave function, where Φ_0 is the magnetic flux quantum. Therefore, the random phase shift in a single scattering act, which takes place during the time interval $\tau = \hbar/\gamma$, is $\sim(\Delta\Phi/\Phi_0) = (pa_B)^2(a_B/\ell_B)^2 \ll 1$. Since the phase shifts of the wave function are random, the total phase shift becomes comparable to unity and coherence is lost only after $(\Phi_0/\Delta\Phi)^2$ scattering acts. The corresponding characteristic time $\tau_B \sim (\Phi_0/\Delta\Phi)^2\tau \sim \tau(pa_B)^{-4}(\ell_B/a_B)^4$, which is consistent with Eq. (32).

2.4. Diffusion constant

In order to obtain quantum corrections to the diffusion constant, one should include Eq. (29) for the cooperon $U(\mathbf{K},\omega)$ together with the first-order impurity vertex
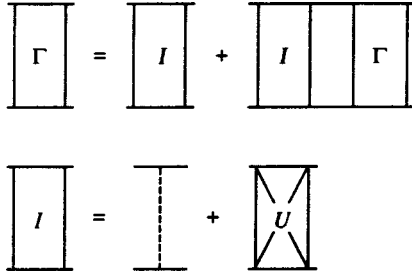
FIG. 4. Diagrammatic representation of the approximation for the vertex Γ including quantum corrections to the diffusion constant.
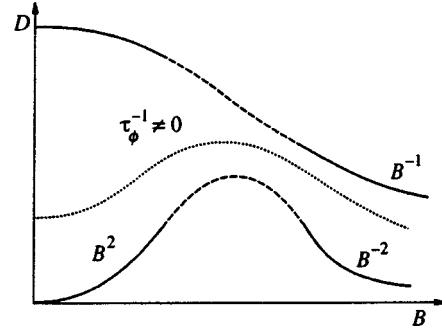

FIG. 5. Static diffusion constant D of excitons as a function of the field B classical transport (upper curve), with consideration of quantum corrections (lower curve), and with weak inelastic scattering (middle curve).

$W(\mathbf{p},\mathbf{p}',\mathbf{K})$ in the ladder diagrams (Figs. 1 and 4) for the effective conductivity.[3] In the ladder approximation, we have a transport coefficient γ_{tr} instead of γ for a random field with a finite correlation length. Technical details of the diagram treatment for 2D excitons are given in Appendix C. The diffusion constant for an exciton of energy ϵ takes the form

$$D(\epsilon)=D_0(\epsilon)\left[1+\frac{\widetilde{\gamma}_0}{4\pi^2\gamma_{\text{tr}}\mathcal{N}(\epsilon)D^c}\ln\left(\frac{D^cK_0^2\widetilde{\gamma}_0}{2\gamma_B\gamma_0}\right)\right]^{-1},\tag{33}$$

where $K_0\simeq\gamma(p)/V(p)$ is the cut-off momentum and $D_0=p^2/4M^2\gamma_{\text{tr}}$ is the conventional ("classical") diffusion constant for an exciton.[19,21]

Before discussing the quantum corrections (33), let us derive the classical diffusion constant D_0 as a function of the magnetic field B. Using general expressions (22) and (16), we obtain

$$D_0(\epsilon,B)\simeq D_0(\epsilon)\left[1-3\kappa\left(\frac{a_B}{\ell_B}\right)^4\right]\equiv D_0\left[1-\left(\frac{B}{B_0}\right)^2\right],\tag{34}$$

where the characteristic magnetic field B_0 is determined by the expression $B_0a_B^2\simeq\Phi_0$ and D_0 is the diffusion constant at $B=0$.[19] The diffusion constant D_0 monotonically decreases as the magnetic field increases in accordance with Eq. (34).

The inclusion of quantum corrections drastically changes the dependence of D on B. Indeed, γ_B tends to zero as $B\to0$, and D vanishes as a result [see Eq. (33)]; this is the weak localization of excitons in the absence of B (excitons, like ordinary 2D particles, are localized in a random potential[7]). A self-consistent approach may be used in this situation.[27] In fact, the approximation for the complete vertex Γ including only ladder diagrams for Γ_0 and maximally crossed diagrams for U applies only to the case of weak scattering, where the resulting diffusion constant is large. When the complete vertex corresponds to strong scattering and the diffusion constant D is small, one cannot, strictly speaking, select a preferential class of diagrams. The underlying idea of the self-consistent approach[27] is the existence of a relation between Γ_0 and U in the presence of time-reversal symmetry (maximally crossed diagrams in the $e-h$ channel are ladder diagrams in the $e-e$ channel). One consequence of this relation is that the diffusion pole, which exists in the diffuson at small momentum transfers, is "transmitted" to the cooperon (where it exists at small total momenta $\mathbf{K}$). Since the vertex Γ is directly related to the

"density-density" correlation function, it is physically clear that it is the diffusion constant D that should appear in Γ, and the constant D_0 in the cooperon should consequently also be replaced by D. The mathematical basis of this approach was discussed in detail by Suslov.[28]

In the specific case under consideration, time-reversal symmetry is broken, and, strictly speaking, there is no duality between the diffuson and cooperon. We can use, however, the self-consistent approximation in order to obtain the leading terms in the B expansion of the total diffusion constant. The point is that D_0 and D^c behave similarly in the leading orders in B. Therefore, the diffusion constant D^c in the cooperon can be replaced by the total diffusion constant D so that $D(B)$ could be calculated in a self-consistent manner using Eq. (33). In the case $D(B)\ll D_0$ the magnetic-field dependence is given by

$$D(\epsilon,B)=(pa_B)^4\left(\frac{B}{B_0}\right)^2D_0(\epsilon)\exp[\mathcal{N}(\epsilon)D_0(\epsilon)],\tag{35}$$

where $B_0a_B^2\simeq\Phi_0$ [see Eq. (34)]. Thus, the static diffusion constant at $B=0$ is zero and increases proportionally to B^2 at small B. This behavior of D corresponds to the suppression of the weak localization of excitons in a magnetic field.

In the strong-field limit, $\ell_B\ll a_B$, the exciton diffusion constant D drops as B^{-2}.[20] Thus, it is clear that, if weak-localization effects are taken into account, D is a nonmonotonic function of the magnetic field. Note that the classical diffusion constant D_0 decreases monotonically with the magnetic field in both the strong-field limit, $\ell_B\ll a_B$, and in the weak-field limit, $\ell_B\gg a_B$. These results are illustrated by Fig. 5. Weak inelastic processes characterized by the dephasing time $\tau_\phi=\hbar/\gamma_\phi$ can be included in our scheme phenomenologically. If τ_ϕ is finite, the diffusion constant is finite even at $B=0$. The value $D(B=0)$ is controlled by γ_ϕ, which should be added to γ_B in Eq. (33). The appearance of γ_ϕ affects D_0 and D differently. If the condition $\gamma_\phi\ll\gamma_0$ holds, this addition has little effect on D_0. But, since $D_0\mathcal{N}\gg1$ in the weak-scattering limit, the relation $D_0\mathcal{N}\gamma_\phi\gtrsim\gamma_0$ can hold even when $\gamma_\phi\ll\gamma_0$. In this case, the weak-localization corrections are small, and we have $D(B)\simeq D_0(B)$.

Analytical calculations are impossible in intermediate fields, where $\ell_B\sim a_B$. It is quite natural to assume that in

this range the magnetic field dependence of the diffusion constant (either the classical constant, D_0, or the constant which takes into account the quantum corrections, D) has the form shown by the dashed lines in Fig. 5. The increase in D with the magnetic field B (*a positive magnetodiffusion effect*) is due to the suppression of the weak localization of excitons in magnetic fields. This effect is similar to the negative magnetoresistance in 2D electron systems.[8]

3. CONCLUSIONS

We have shown that a magnetic field B eliminates divergence of the maximally crossed diagrams in the ''exciton-antiexciton'' channel (the exciton analogue of the cooperon). Unlike charged particles, an exciton acquires a phase in the field B not during free motion, but only upon scattering by defects. As a result, the diffusion constant of 2D excitons in magnetic fields remains finite as $\omega \to 0$ (under the assumption that the random potential is weak). The static diffusion constant $D(B)$ is a decreasing function of B in strong magnetic fields, $\ell_B \ll a_B$, whereas in weak magnetic fields, $\ell_B \gg a_B$ (and, probably, in intermediate fields, $\ell_B \sim a_B$) $D(B)$ increases with the magnetic field, i.e., a positive magnetodiffusion effect takes place for excitons. The self-consistent approximation yields $D \propto B^2$ in weak magnetic fields, which indicates that weak localization is suppressed in a magnetic field B. Quantum corrections are also important in the strong-field limit, $\ell_B \ll a_B$, and lead to a faster power-law decrease in the diffusion constant with the magnetic field, $D \propto B^{-2}$,[20] in comparison to the classical diffusion constant, $D_0 \propto B^{-1}$.[19,21] This is because the characteristic internal length scale of the magnetoexciton $\ell_B \ll a_B$ decreases with increasing B as $\ell_B \propto B^{-1/2}$, and its internal structure has a lesser impact on the scattering process, so that the magnetoexciton becomes similar to a structureless neutral boson. Thus, for neutral $e-h$ systems, crossover to the exciton weak-localization regime takes place in the strong-field limit (unlike electron systems, which contain delocalized states in the quantum Hall effect regime).

Although the calculated function $D(B)$ is a nonmonotonic function of B, it does not reproduce all the details of the experimental findings for $D(B)$.[17] For instance, the observed suppression of exciton magnetotransport in the range of relatively low fields[17] is in agreement with the theoretical predictions for the behavior of the *classical* diffusion constant (Fig. 5). Our calculations, however, demonstrate that the increase in $D(B)$ in the range $B > 6$ T observed in Ref. 17 cannot be interpreted in terms of the suppression of weak localization of excitons in a magnetic field. Note that the localization regime in double quantum wells used in experiments[16-18] is closer to the strong-localization regime of excitons.[19] Also, we have not considered the effects of the Bose–Einstein condensation of excitons. The investigation of the effects of a magnetic field on the strong localization of excitons and of Bose–Einstein condensation on the transport of neutral composite particles (excitons) is a very interesting problem, which has not been solved. Note also that, as in the case of electrons in quasi-two-dimensional semiconductor structures,[12] the effects of fast transitions between different spin states may be important for excitons.

Our theoretical prediction of an increase in exciton mobility with increasing B in the weak-localization regime can be tested experimentally at low temperatures (where inelastic scattering is suppressed and the dephasing time τ_ϕ is large) in magnetic fields for which $\ell_B \gtrsim a_B$. Such experiments require quantum wells with a weak random potential, for example, wide quantum wells with smooth interfaces.

APPENDIX A: DIAGRAMMATIC REPRESENTATION OF EXCITON SCATTERING

To the best of our knowledge, no approximation which reduces the scattering of excitons to an effectively one-particle process has been rigorously developed using diagram techniques. An approach similar to that developed in this Appendix can be applied to other problems, such as investigations of the role of transitions to excited states, the effects of a finite exciton density, and strong localization in terms of the effective exciton scattering.

In the electron-hole representation, a Wannier–Mott exciton is described in terms of a sum of ladder diagrams, which include the $e-h$ Coulomb interaction. The corresponding two-particle Green's function can be expanded in terms of the exciton eigenfunctions $\Psi_\lambda(\mathbf{r}_e, \mathbf{r}_h)$:

$$G_2(\mathbf{r}_1, \mathbf{r}_2, t; \mathbf{r}_3, \mathbf{r}_4, t')$$
$$\equiv -i \langle T \hat{\Psi}_e(\mathbf{r}_1, t) \hat{\Psi}_h(\mathbf{r}_2, t) \hat{\Psi}_h^\dagger(\mathbf{r}_4, t') \hat{\Psi}_e^\dagger(\mathbf{r}_3, t') \rangle$$
$$= \int dE \sum_\lambda \frac{\Psi_\lambda^*(\mathbf{r}_1, \mathbf{r}_2) \Psi_\lambda(\mathbf{r}_3, \mathbf{r}_4)}{E - \epsilon_\lambda} \exp[-iE(t-t')],$$
$$\text{(A1)}$$

where $\hat{\Psi}_e^\dagger(\mathbf{r}, t)$ and $\hat{\Psi}_h^\dagger(\mathbf{r}, t)$ are the electron and hole creation operators in the Heisenberg representation. For simplicity, we consider the case of zero magnetic field, $B = 0$, where ϵ_λ are the ordinary eigenvalues of the exciton energies, $\epsilon_\lambda = \epsilon(p) + \epsilon_n$, ϵ_n is the energy (in either the discrete or continuous spectrum) of relative motion, $\epsilon(p)$ is the center-of-mass kinetic energy, $\Psi_\lambda(\mathbf{r}_1, \mathbf{r}_2) = \exp(i\mathbf{p}\mathbf{R}) \Phi_n(\mathbf{r})$ are the exciton wave functions, $\mathbf{R} = (m_e \mathbf{r}_1 + m_h \mathbf{r}_2)/M$, and $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$. Our aim is to replace the two-particle $e-h$ Green's function by an effective ''one-particle'' Green's function of an exciton defined by the formula

$$G(\mathbf{R}, t; \mathbf{R}', t') = -i \langle T B(\mathbf{R}, t) B^\dagger(\mathbf{R}', t') \rangle, \quad \text{(A2)}$$
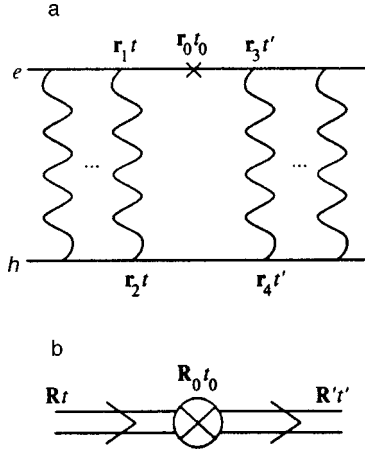
where the exciton creation operator is defined as

FIG. 6. (a) Impurity vertex in the electron line of the two-particle $e-h$ propagator. The wavy lines correspond to the Coulomb $e-h$ interaction. (b) Impurity vertex (10) in the excitonic propagator.

$$B_\lambda^\dagger(\mathbf{R},t)=\int d\mathbf{r}\hat{\Psi}_e^\dagger\left(\mathbf{R}+\frac{m_h}{M}\mathbf{r},t\right)$$

$$\times\hat{\Psi}_h^\dagger\left(\mathbf{R}-\frac{m_e}{M}\mathbf{r},t\right)\Psi_\lambda^*(\mathbf{R},\mathbf{r}).\qquad(A3)$$

The function G_2 satisfies the following Bethe–Salpeter equation:

$$G_2(\mathbf{r}_1,\mathbf{r}_2,t;\mathbf{r}_3,\mathbf{r}_4,t')$$

$$=G_e(\mathbf{r}_1,\mathbf{r}_3,t-t')G_h(\mathbf{r}_2,\mathbf{r}_4,t-t')$$

$$+\int d\mathbf{r}_3'd\mathbf{r}_4'dt_1'G_e(\mathbf{r}_1,\mathbf{r}_3',t-t_1')G_h(\mathbf{r}_2,\mathbf{r}_4',t-t_1')$$

$$\times U(\mathbf{r}_3'-\mathbf{r}_4')G_2(\mathbf{r}_3',\mathbf{r}_4',t_1';\mathbf{r}_3,\mathbf{r}_4,t').\qquad(A4)$$

Applying the operator (A4)

$$[G_e(\mathbf{r}_1,t)G_h(\mathbf{r}_2,t)]^{-1}=i\frac{\partial}{\partial t}+\frac{\nabla_1^2}{2m_e}+\frac{\nabla_2^2}{2m_h}$$

to both sides, we obtain the Schrödinger-like equation:

$$\left[i\frac{\partial}{\partial t}+\frac{\nabla_1^2}{2m_e}+\frac{\nabla_2^2}{2m_h}-U(\mathbf{r}_1-\mathbf{r}_2)\right]G_2(\mathbf{r}_1,\mathbf{r}_2,t;\mathbf{r}_3,\mathbf{r}_4,t')$$

$$=\delta(\mathbf{r}_1-\mathbf{r}_3)\delta(\mathbf{r}_2-\mathbf{r}_4)\delta(t-t').\qquad(A5)$$

Now let us consider a diagram with only one impurity vertex corresponding to the external potential V_e in the electron line (Fig. 6a).

The set of external coordinates $(\mathbf{r}_e,\mathbf{r}_h,t)$ will be symbolically denoted by $\bar{X},\bar{X}'$. The analytical expression for $\widetilde{G}_2(\bar{X},\bar{X}')$ in the case of the diagram in Fig. 6a has the form

$$\widetilde{G}_2(\bar{X},\bar{X}')=\int d\mathbf{r}_0d\mathbf{r}_1d\mathbf{r}_2d\mathbf{r}_3d\mathbf{r}_4dtdt_0dt'G_2(\bar{X};\mathbf{r}_1,\mathbf{r}_2,t)$$

$$\times U(\mathbf{r}_1-\mathbf{r}_2)G_e(\mathbf{r}_1-\mathbf{r}_0,t-t_0)V_e(\mathbf{r}_0)$$

$$\times G_e(\mathbf{r}_0-\mathbf{r}_3,t_0-t')G_h(\mathbf{r}_2-\mathbf{r}_4,t-t')$$

$$\times U(\mathbf{r}_3-\mathbf{r}_4)G_2(\mathbf{r}_3,\mathbf{r}_4,t';\bar{X}').\qquad(A6)$$

Using Eq. (A5), we can replace the function G_2U by a differential operator acting on G_e and G_h. Given that

$$\left(i\frac{\partial}{\partial t}+\frac{\nabla^2}{2m_{e(h)}}\right)G_{e(h)}(\mathbf{r}-\mathbf{r}',t-t')=\delta(t-t')\delta(\mathbf{r}-\mathbf{r}'),$$

we obtain the following expression for function (A6):

$$\widetilde{G}_2(\bar{X},\bar{X}')=\int d\mathbf{r}_1d\mathbf{r}_2d\mathbf{r}_3d\mathbf{r}_4dt\,dt'\,G_2(\bar{X};\mathbf{r}_1,\mathbf{r}_2,t)$$

$$\times V_e(\mathbf{r}_1)G_e(\mathbf{r}_1-\mathbf{r}_3,t-t')G_h(\mathbf{r}_2-\mathbf{r}_4,t-t')$$

$$\times U(\mathbf{r}_3-\mathbf{r}_4)G_2(\mathbf{r}_3,\mathbf{r}_4,t';\bar{X}')+\int d\mathbf{r}_0d\mathbf{r}_1d\mathbf{r}_2$$

$$\times d\mathbf{r}_3dt\,dt_0G_2(\bar{X};\mathbf{r}_1,\mathbf{r}_2,t)G_e(\mathbf{r}_1-\mathbf{r}_0,t-t_0)$$

$$\times V_e(\mathbf{r}_0)G_e(\mathbf{r}_0-\mathbf{r}_3,t_0-t)U(\mathbf{r}_3-\mathbf{r}_2)$$

$$\times G_2(\mathbf{r}_3,\mathbf{r}_2,t;\bar{X}').\qquad(A7)$$

Note that the second term contains the product $G_e(t)\times G_e(-t)$, which contributes a factor $\sim n_e(1-n_e)$ and can, therefore, be neglected in the low-density limit. Thus only the first term remains on the right-hand side of Eq. (A7). Let us also take into account that the expansion for G_2 begins with a term of zero order in G_eG_h [corresponding to δ functions on the right-hand side of Eq. (A5)] and add it to Eqs. (A6) and (A7). Then we can see that the first term of the right-hand side of Eq. (A7) contains the Coulomb ladder diagrams on both sides of the impurity vertex $V_e(\mathbf{r}_1)$. In addition, the temporal and spatial coordinates coincide in such a manner that, using representations (A2) and (A5), we can represent expression (A3) in the form of a diagram corresponding to scattering of an exciton as a whole (Fig. 6b). After adding the analogous term for scattering of the hole, we can see that the effective scattering potential in the excitonic representation is indeed determined by Eqs. (9) and (10).

APPENDIX B: CALCULATION OF PERTURBATION SERIES

In a perturbative analysis of systems with 2D excitons in a magnetic field, sums like those in Eqs. (14) and (15) appear frequently. Therefore, it seems useful to calculate these sums exactly for the case of a two-dimensional hydrogenic exciton. If the operator approach[29] is applied, the explicit form of the ground-state wave function is sufficient. In the intermediate calculations we set $a_B=\hbar=1$ and return to dimensional quantities in the final expressions. Let us start with the constant κ in Eq. (15). If we can find the explicit form of the operator $\hat{b}$ that satisfies the quantum equation of motion $\mu\partial\hat{b}/\partial t=i\mu[H_0,\hat{b}]=x$, where H_0 is the Hamiltonian of a 2D hydrogen atom (7) in the absence of a field B, for the matrix elements we have

$$i\mu(\epsilon_0-\epsilon_n)\langle 0|\hat{b}|n\rangle=\langle 0|x|n\rangle,\qquad(B1)$$

[where $\epsilon_n=\epsilon_0/(n+1/2)^2$ and $\epsilon_0=-\mu e^4/2\varepsilon^2\hbar^2$], and the sum in Eq. (15) is reduced to the diagonal matrix element

$$\kappa = i\frac{\mu}{M}\sum_n{}' \langle 0|x|n\rangle\langle n|\hat{b}|0\rangle = i\frac{\mu}{M}\langle 0|x\hat{b}|0\rangle, \qquad (B2)$$

where the prime means that state $n=0$ is not included in the summation. We have used the completeness condition for the states $\Sigma_n|n\rangle\langle n|=1$ and the equality $\langle 0|x|0\rangle=0$. In the coordinate representation, we introduce the notation $\hat{b}\phi_0(r) \equiv b(\mathbf{r})\phi_0(r)$, where $\phi_0(r)=\sqrt{8/\pi}\,\exp(-2r)$ is the ground-state wave function. Using the explicit form of the Hamiltonian, we obtain a differential equation for $b(\mathbf{r})$:

$$\frac{1}{2}[\nabla^2 b(\mathbf{r})]\phi_0(r)+(\nabla b(\mathbf{r})\cdot\nabla\phi_0(r))=-ix\phi_0(r), \quad (B3)$$

whence it follows that $b(\mathbf{r})=-ib(r)\cos\phi$, and the unknown function $b(r)$ is given by the equation

$$b''(r)+b'(r)\left(\frac{1}{r}-4\right)-\frac{b(r)}{r^2}-2r=0. \qquad (B4)$$

Solving Eq. (B4), we obtain

$$b(\mathbf{r})=-i\left(\frac{1}{4}r^2+\frac{3}{16}r\right)\cos\,\phi. \qquad (B5)$$

Finally, the matrix element in Eq. (B2) is expressed as

$$\kappa=\frac{\mu}{M}\int_0^\infty dr\left(\int_0^{2\pi}d\phi\,\frac{8}{\pi}\,\cos^2\phi\right)$$

$$\times\exp(-4r)\left(\frac{1}{4}r^2+\frac{3}{16}r\right)r^2=\frac{21}{16^2}\frac{\mu}{M}, \qquad (B6)$$

and the coefficient α in Eq. (15) is given by the expression

$$\alpha_{e(h)}=2\frac{21}{16^2}\frac{m_{e(h)}\mu}{M^2}.$$

The same operator $\hat{b}$ can be used in calculating the first sum in Eq. (14):

$$I_1=\sum_n{}'\,\frac{|\langle 0|x|n\rangle|^2}{(\epsilon_0-\epsilon_n)^2}. \qquad (B7)$$

Using Eq. (B1), we can also reduce I_1 to a diagonal matrix element: $I_1=\mu^2\langle 0|\hat{b}\hat{b}|0\rangle$. In combination with the explicit expression (B5) for $b(r)$, this equation yields

$$I_1=\mu^2\int_0^\infty dr|b(r)|^2\phi_0^2(r)=\frac{159}{4^6}\mu^2. \qquad (B8)$$

In order to calculate the second sum in Eq. (14),

$$I_2=\sum_n{}'\,\frac{|\langle 0|r^2|n\rangle|^2}{\epsilon_n-\epsilon_0}, \qquad (B9)$$

we must find an operator $\hat{b}_2$ such that $i\mu[H_0,\hat{b}_2]=r^2$. We set $\hat{b}_2\phi_0(r)=b_2(r)\phi_0(r)$. Then

$$b_2''(r)+b_2'(r)\left(\frac{1}{r}-4\right)-2ir^2=0.$$

The solution is the function

$$b_2(r)=-\frac{i}{2}\left(\frac{1}{3}r^3+\frac{3}{8}r^2+\frac{3}{8}r+\frac{3}{32}\,\ln\,r+c_1\right) \qquad (B10)$$

with an undetermined constant c_1. One feature of calculations of sums like I_2 is that, using the completeness condition for the intermediate states, we should eliminate the matrix element of the ground state $n=0$, which does not automatically equal zero, unlike the coordinate matrix element $\langle 0|x|0\rangle=0$. Therefore, we have [cf. Eq. (B2)]

$$I_2=i\mu(\langle 0|\hat{b}_2r^2|0\rangle-\langle 0|\hat{b}_2|0\rangle\langle 0|r^2|0\rangle). \qquad (B11)$$

We see that, as a result of subtraction, the final expression does not include the constant c_1 introduced in Eq. (B10). This allows us to obtain the exact expression $I_2=105\mu/2^9$, and for β we have

$$\beta_{e(h)}=\frac{1}{4^6M^2}\left(105m_{e(h)}^2-\frac{159}{2}\mu^2\right). \qquad (B12)$$

Note that the coefficients $\beta_{e(h)}>0$ are always positive [since $\mu=m_em_h/(m_e+m_h)<m_e,m_h$], but numerically small: $\beta\lesssim 0.02$.

APPENDIX C: CALCULATION OF THE DIFFUSION CONSTANT D

This section gives details of the calculation of the diffusion constant D. The calculation of $D(\epsilon)$ should include, in addition to the diagrams of Fig. 4, the zero-order diagram G^RG^A. Therefore, the diffusion constant $D(\epsilon)$ is given by the expression

$$D(\epsilon)=\frac{1}{2\pi\mathcal{N}(\epsilon)}\int d\mathbf{p}\int d\mathbf{p}'\frac{\mathbf{pp}'}{M^2}|G^R(\mathbf{p})|^2[\,\delta(\mathbf{p}-\mathbf{p}')$$

$$+\Gamma(\mathbf{p},\mathbf{p}')|G^R(\mathbf{p}')|^2]. \qquad (C1)$$

If the cooperon is included in the irreducible part, the complete vertex Γ satisfies the Bethe–Salpeter equation shown in Fig. 4. Note that the cooperon (as a function of the variables $\mathbf{p}$ and $\mathbf{p}'$ at $\mathbf{q}=0$) can be expressed approximately as

$$U_{\epsilon,\omega}(\mathbf{p},\mathbf{p}';\mathbf{p}',\mathbf{p})\simeq\int\frac{d\mathbf{K}}{(2\pi)^2}U(\mathbf{K},\omega)\,\delta(\mathbf{p}+\mathbf{p}')$$

$$\equiv U\,\delta(\mathbf{p}+\mathbf{p}'). \qquad (C2)$$

This approximation can be used because there are essentially different momentum scales in the problem. Indeed, in integrals like

$$\int d\mathbf{p}\int d\mathbf{p}'|G^R(\mathbf{p})|^2|G^R(\mathbf{p}')|^2U_{\epsilon,\omega}(\mathbf{p},\mathbf{p}';\mathbf{p}',\mathbf{p})$$

$$=\int d\mathbf{p}\int d\mathbf{K}|G^R(\mathbf{p})|^2|G^R(\mathbf{K}-\mathbf{p})|^2U_{\epsilon,\omega}(\mathbf{p},\mathbf{K}-\mathbf{p};\mathbf{K})$$

$$(C3)$$

only small $\mathbf{K}$ are important owing to the presence of the diffusion pole in U. In this case, we can assume in an approximation that $\mathbf{K}-\mathbf{p}\approx-\mathbf{p}$ and perform integration over $\mathbf{p}$

and **K** independently. This yields Eq. (C2) for the cooperon. Then the equation for the vertex $\Gamma(\mathbf{p},\mathbf{p}')$ shown in Fig. 4 takes the form

$$\Gamma(\mathbf{p},\mathbf{p}') = W(\mathbf{p},\mathbf{p}',0) + U\delta(\mathbf{p}+\mathbf{p}')$$

$$+ \int \frac{d\mathbf{p}_1}{(2\pi)^2}[W(\mathbf{p},\mathbf{p}_1,0) + U\delta$$

$$\times(\mathbf{p}+\mathbf{p}_1)]G^R(\mathbf{p}_1)G^A(\mathbf{p}_1)\Gamma(\mathbf{p}_1,\mathbf{p}'). \qquad (C4)$$

The quantity needed for the calculation of $D(\epsilon)$ has the form [see Eq. (C1)]

$$\Gamma_1 = \int \frac{d\phi_{\mathbf{p}}}{2\pi} \int \frac{d\phi_{\mathbf{p}'}}{2\pi}(\hat{\mathbf{p}}\cdot\hat{\mathbf{p}}')\Gamma(\mathbf{p},\mathbf{p}'),$$

where the integration is performed on the mass surface $\epsilon(\mathbf{p}) = \epsilon(\mathbf{p}') = \epsilon$. For the term corresponding to the allowance for the first angular momentum in Γ_1, Eq. (C4) gives

$$\Gamma_1 = \frac{\gamma_1}{\pi\mathcal{N}_\epsilon} - \frac{U}{2\pi\mathcal{N}_\epsilon\gamma_0} + \frac{\gamma_1}{\gamma_0}\Gamma_1 - \frac{U}{2\gamma_0^2}\Gamma_1, \qquad (C5)$$

where

$$\gamma_1 = \int \frac{d\phi_{\mathbf{p}}}{2\pi} \int \frac{d\phi_{\mathbf{p}'}}{2\pi}(\hat{\mathbf{p}}\cdot\hat{\mathbf{p}}')W(\mathbf{p},\mathbf{p}',0).$$

The solution of Eq. (C5) is

$$\Gamma_1 = \frac{\gamma_0}{\pi\mathcal{N}_\epsilon}\frac{\gamma_1 - U/2\gamma_0}{\gamma_{\mathrm{tr}} + U/2\gamma_0}, \qquad (C6)$$

where, as usual, $\gamma_{\mathrm{tr}} = \gamma_0 - \gamma_1$. Using Eq. (C6), from Eq. (C1) we obtain

$$D(\epsilon) = D_0\left[1 + \frac{U}{2\gamma_{\mathrm{tr}}\gamma_0}\right]^{-1}. \qquad (C7)$$

This allows us to perform the last step of the calculation: by substituting the expressions (C2) and (29) into Eq. (C7), we obtain Eq. (33).

---

*)E-mail: dzyub@gpi.ac.ru

[1])Hereafter we assume that the valence band is nondegenerate and holes have spin 1/2. We ignore the effects due to the different spin states of the exciton, which should be taken into account if the relaxation between them is fast. In the case of III–V semiconductors, this analysis should take account of the real valence band spectrum (see, e.g., Ref. 15), which is beyond the scope of the present study.

[2])When excitons are scattered by charged impurities with a 2D density n_{imp}, we have[19] $\gamma_0/\epsilon \simeq \nu_{\mathrm{imp}}$, where $\nu_{\mathrm{imp}} = 2\pi a_B^2 n_{\mathrm{imp}}$ ($\nu_{\mathrm{imp}} = 2\pi \ell_B^2 n_{\mathrm{imp}}$) is the dimensionless density of impurities for the limiting case $\ell_B \gg a_B$ ($\ell_B \ll a_B$). The smallness of parameter ν_{imp} ensures the applicability of the weak-scattering approximation to this scattering mechanism.

[3])In a strong field, $\ell_B \ll a_B$, the term for the Coulomb interaction in Eq. (7) is treated as a perturbation,[19,20] and the wave functions $\Phi_{\mathbf{p}}(\mathbf{r})$ obtained in this limit describe 2D magnetoexcitons.[25]

[4])Equation (20) may include, in principle, terms linear in the magnetic field B [see Eq. (12)]. This is possible, however, only under the peculiar condition that there be a preferential direction in space in the system so that terms linear in B would not vanish in calculating the correlators of $V_{e(h)}$. We do not consider such a possibility in this paper. Note that a preferential direction can be assigned in a system, for example, by an applied electric field.

[5])Otherwise (as in the case of exciton localization in a zero magnetic field[7]), we would simply have the conventional weak-localization theory in an

effective potential with a finite correlation length comparable to the exciton radius.

---

[1] E. P. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan, Phys. Rev. Lett. **42**, 673 (1979).
[2] L. P. Gor'kov, A. I. Larkin, and D. E. Khmel'nitskii, JETP Lett. **30**, 228 (1979).
[3] P. A. Lee and T. V. Ramakrishnan, Rev. Mod. Phys. **57**, 287 (1985).
[4] S. John and M. J. Stephen, Phys. Rev. B **28**, 6358 (1983).
[5] M. P. Van Albada and A. Langedijk, Phys. Rev. Lett. **55**, 2692 (1985); E. L. Ivchenko, G. E. Pikus, B. S. Razbirin, and A. I. Starukhin, Zh. Éksp. Teor. Fiz. **72**, 2230 (1977) [Sov. Phys. JETP **45**, 1172 (1977)].
[6] D. L. Shepelyansky, Phys. Rev. Lett. **73**, 2607 (1994); Y. Imry, Europhys. Lett. **30**, 405 (1995).
[7] Zh. S. Gevorkyan and Yu. E. Lozovik, Fiz. Tverd. Tela **27**, 1800 (1985) [Sov. Phys. Solid State **27**, 1079 (1985)].
[8] B. L. Altshuler, D. E. Khmel'nitskii, A. I. Larkin, and P. A. Lee, Phys. Rev. B **22**, 5142 (1980).
[9] S. Hikami, A. I. Larkin, and Y. Nagaoka, Prog. Theor. Phys. **63**, 707 (1980).
[10] D. E. Khmel'nitskii, Physica B (Amsterdam) **126**, 235 (1984).
[11] B. L. Al'tshuler, A. G. Aronov, A. I. Larkin, and D. E. Khmel'nitskii, Zh. Éksp. Teor. Fiz. **81**, 768 (1981) [Sov. Phys. JETP **54**, 411 (1981)].
[12] S. V. Iordanskii, Yu. B. Lyanda-Geller, and G. E. Pikus, JETP Lett. **60**, 206 (1994); F. G. Pikus and G. E. Pikus, Phys. Rev. B **51**, 16 928 (1995); W. Knap, C. Skierbiszewski, A. Zduniak, *et al.*, Phys. Rev. B **53**, 3912 (1996).
[13] V. L. Nguen, B. Z. Spivak, and B. I. Shklovskii, Zh. Éksp. Teor. Fiz. **89**, 1770 (1985) [Sov. Phys. JETP **62**, 1021 (1985)].
[14] U. Sivan, O. Entin-Wohlman, and Y. Imry, Phys. Rev. Lett. **60**, 1566 (1988).
[15] M. Z. Maialle, E. A. Andrada e Silva, and L. J. Sham, Phys. Rev. B **47**, 15 776 (1993).
[16] L. V. Butov, A. Zrenner, G. Abstreiter, G. Böhm, and G. Weimann, Phys. Rev. Lett. **73**, 304 (1994).
[17] L. V. Butov, A. Zrenner, M. Hagn, G. Abstreiter, G. Böhm, and G. Weimann, Surf. Sci. **362**, 243 (1996).
[18] M. Hagn, A. Zrenner, and G. Weimann, Appl. Phys. Lett. **67**, 232 (1995).
[19] A. B. Dzyubenko and G. E. W. Bauer, Phys. Rev. B **51**, 14 524 (1995).
[20] P. I. Arseyev and A. B. Dzyubenko, Phys. Rev. B **52**, R2261 (1995).
[21] P. I. Arseyev, A. B. Dzyubenko, and G. E. W. Bauer, in *Proceedings of the 12th International Conference ''High Magnetic Fields in the Physics of Semiconductors II''*, G. Landwehr and W. Ossau (eds.), World Scientific, Singapore (1997), p. 729.
[22] A. B. Dzyubenko and P. I. Arseyev, in *Proceedings of the 23rd International Conference on Physics of Semiconductors*, M. Scheffler and R. Zimmermann (eds.), World Scientific, Singapore (1996), p. 2063.
[23] A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyaloshinskii, *Methods of Quantum Field Theory in Statistical Physics*, Prentice-Hall, Englewood Cliffs, N.J. (1963).
[24] L. P. Gor'kov and I. E. Dzyaloshinskii, Zh. Éksp. Teor. Fiz. **53**, 717 (1967) [Sov. Phys. JETP **26**, 449 (1968)].
[25] I. V. Lerner and Yu. E. Lozovik, Zh. Éksp. Teor. Fiz. **78**, 1167 (1980) [Sov. Phys. JETP **51**, 588 (1980)].
[26] G. D. Mahan, *Many-Particle Physics*, Plenum Press, New York (1990), Chap. 7.
[27] D. Vollhardt and P. Wölfle, Phys. Rev. B **22**, 4666 (1980); Phys. Rev. Lett. **48**, 699 (1982).
[28] I. M. Suslov, Zh. Éksp. Teor. Fiz. **108**, 1686 (1995) [JETP **81**, 925 (1995)].
[29] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics. Non-Relativistic Theory*, 3rd ed., Pergamon Press, Oxford (1977).

Symmetries and causes of the coincidence of the emission spectra of mirrors and charges in 1+1 and 3+1 spaces

V. I. Ritus*)

P. N. Lebedev Physical Institute, Russian Academy of Sciences, 117924 Moscow, Russia
(Submitted 22 December 1997)
Zh. Éksp. Teor. Fiz. **114**, 46–62 (July 1998)

This paper discusses the symmetry of the wave field that lies to the right and left of a two-sided accelerated mirror in 1+1 space and satisfies a single condition on it. The symmetry is embodied in the Bogolyubov matrix coefficients α and β that connect the two complete sets of solutions of the wave equations. The amplitudes of the quantum processes in the right and left half-spaces are expressed in terms of α and β and are related to each other by the transformation (12). The coefficient $\beta^{*}_{\omega'\omega}$ plays the role of the source amplitude of a pair of oppositely directed particles with frequencies ω and ω' of which one is in the left half-space and the other is in the right half-space because one of them has undergone reflection. Such an interpretation makes $\beta^{*}_{\omega'\omega}$ observable and explains why, as shown by Eq. (1) and found earlier by Nikishov and Ritus [Zh. Éksp. Teor. Fiz. **108**, 1121 (1995); transl. JETP **81**, 615 (1995)] and by Ritus [Zh. Éksp. Teor. Fiz. **110**, 526 (1996); transl. JETP **83**, 282 (1996)], the emission spectra of a mirror in 1+1 space coincide with those of charges in 3+1 space. The reason is that the angular momentum of the pair emitted by the mirror coincides with the angular momentum of the single particle emitted by the charge. © *1998 American Institute of Physics.* [S1063-7761(98)00407-7]

1. INTRODUCTION

It was found in Refs. 1 and 2 that the spectra of bosons and fermions emitted by an accelerated mirror in 1+1 space coincide with the spectra of photons and scalar quanta emitted by electric and scalar charges in 3+1 space when the latter move along the same trajectory as does the mirror. Namely, the Bogolyubov coefficients $\beta^{B,F}_{\omega'\omega}$ that describe the spectra of the Bose and Fermi radiations of an accelerated mirror and the Fourier transforms of the density of the 4-current $j_\alpha(k_+,k_-)$ and the scalar charge density $\rho(k_+,k_-)$ that describe the spectra of the photons and scalar quanta emitted by electric and scalar charges are connected by the relationships

$$|\beta^{B}_{\omega'\omega}|^2 = \frac{1}{e^2}|j_\alpha(k_+,k_-)|^2,$$

$$|\beta^{F}_{\omega'\omega}|^2 = \frac{1}{e^2}|\rho(k_+,k_-)|^2. \qquad (1)$$

It is assumed here that the components $k_\pm = k^0 \pm k^1$ of the wave 4-vector k^α of the quantum emitted by the charge are identified with the doubled frequencies ω and ω' of the quanta emitted by the mirror:

$$2\omega = k_+, \quad 2\omega' = k_-, \qquad (2)$$

and e is the electrical or scalar charge in Heaviside units.

However, there is a substantial physical difference between the right-hand and left-hand quantities in Eqs. (1), i.e., between the emission spectra of the charges and of the mirror. Whereas the former are the distribution of the mean number of radiated quanta over the two independent compo-

nents k_+ and k_- of the wave vector of the quantum (as a consequence of the azimuthal symmetry of the radiation, there is no dependence on the third independent variable), the latter have a more complicated interpretation. Actually, they will be the spectra of the mean number of quanta emitted by the mirror to the right only after integration over frequency ω' (Ref. 3):

$$d\bar{n}_\omega = \frac{d\omega}{2\pi}\int_0^\infty \frac{d\omega'}{2\pi}|\beta_{\omega'\omega}|^2. \qquad (3)$$

If the mirror is two-sided and infinitely thin, then, besides the quanta emitted to the right with the spectrum given by Eq. (3), it will (as we can see) also emit quanta to the left with the spectrum

$$d\bar{n}_{\omega'} = \frac{d\omega'}{2\pi}\int_0^\infty \frac{d\omega}{2\pi}|\beta_{\omega'\omega}|^2. \qquad (4)$$

The question naturally arises whether it is not possible to regard the quantity

$$|\beta_{\omega'\omega}|^2\frac{d\omega\,d\omega'}{(2\pi)^2} \qquad (5)$$

as the mean number of pairs of quanta, one of which, with frequency ω in the interval $d\omega$, is emitted by the mirror to the right, while the other, with frequency ω' in the interval $d\omega'$, is emitted to the left. In this case, two frequencies ω and ω' would be observable, characterizing one event: the mirror emits a pair of quanta, in the same way as two components k_+ and k_- also characterize the emission of one

quantum by a charge. As we can see, with certain nontrivial complications, such a treatment is actually valid. In any case, the mirror emits quanta in pairs.

Clearly, this circumstance helps to understand another difference between the coincident spectra of a charge and a mirror. While the bosons and fermions emitted by a mirror have spins of 0 and 1/2, the photons and scalar quanta emitted by electrical and scalar charges have spins of 1 and 0. Even though the quanta have different spin, the emission spectra of the charges coincide with the boson and fermion spectra of the mirror.

This coincidence is explained by the fact that, unlike charges, the mirror emits particles in pairs, and a pair of spinless bosons can have a total angular momentum of 1, while a pair of fermions can have a total angular momentum of 0. Then the angular momentum of the pair emitted by the mirror coincides with the spin of the particle emitted by the charge. The fact that, upon reflection, $\beta^B_{\omega'\omega}$ behaves like a pseudoscalar while $\beta^F_{\omega'\omega}$ behaves like a scalar can serve as an indirect confirmation of this (see Secs. 2 and 4).

It is shown in Sec. 2 that the system of Bogolyubov coefficients obtained for a right-sided mirror (i.e., for the field to the right of a mirror with a boundary condition on it), because of the properties of mirror symmetry, also describes the processes in the field to the left of a mirror with the same boundary condition. In other words, the same system of Bogolyubov coefficients characterizes the behavior of the field in all of space—both to the right and to the left of a two-sided mirror.

Section 3 treats the connection between the integral quantities that characterize the radiation of a two-sided mirror, their behavior under certain space–time transformations, and the symmetry (or asymmetry) of the space–time regions where they are formed.

The symmetry of the Bogolyubov coefficients reflects the symmetry of two inequivalent total systems of solutions of the wave equation, definite and smooth in all of $1+1$ space, satisfying inside it—on the trajectory of the mirror—a single condition and characterized by propagation of a monochromatic component of each solution toward the right in one system and toward the left in the other. When the field is quantized and when the usual comparison of monochromatic plane waves to particles is made, these two systems of solutions form ingoing and outgoing systems for the field to the right of the trajectory and outgoing and ingoing systems for the field to the left of it. Therefore, the quantum processes in the field to the right and to the left of the mirror are independent, even though they are described by a single system of Bogolyubov coefficients. In particular, the particle-production amplitudes to the right and to the left of the mirror, the single-particle scattering amplitudes, etc. are associated with the transformation (12). These amplitudes, certain frequency distributions, and also the distribution of pair-production probabilities over the number of pairs, which is invariant under the transformation (12), are computed in Section 4. It is shown that $\beta^*_{\omega'\omega}$ plays the role of the source amplitude of a pair of particles potentially emitted to the right and to the left with frequencies ω and ω', with the spin

of a boson pair equaling 1, while that of the fermion pair equals 0.

In the last section, Sec. 5, a similar method is used to treat the emission by an accelerated mirror of pairs the particle and antiparticle of which are not identical.

A system of units in which $\hbar = c = 1$ is used in this article. To simplify the formulas in Secs. 4 and 5, the frequencies are considered discrete, integration over $d\omega/2\pi$ is replaced by summation over ω, and the delta function $2\pi\delta(\omega - \omega'')$ is replaced by the Kronecker symbol $\delta_{\omega\omega''}$.

2. SYMMETRY OF THE BOGOLYUBOV COEFFICIENTS AND RADIATION FROM AN ACCELERATED TWO-SIDED MIRROR

Let us consider the connection between emission spectra and other quantities in two problems in which the mirror trajectories $x = \xi_1(t)$ and $x = \xi_2(t)$ differ by reflection: $\xi_1(t) = -\xi_2(t)$. Then, if the first trajectory is described on the plane of the variables $u = t - x$, $v = t + x$ by the function $v = v_1 = f(u)$, the second trajectory will be described by the function inverse to it: $v = v_2 = g(u)$, $g(f(u)) = u$.

The Bogolyubov coefficients, defined as in Refs. 1 and 2 for the field to the right of the mirror [see also Eqs. (33) and (34)],

$$\alpha^B_{\omega'\omega}[f], \beta^{B*}_{\omega'\omega}[f] = \pm \sqrt{\frac{\omega}{\omega'}} \int_{-\infty}^{\infty} du \, \exp(\mp i\omega u + i\omega' f(u))$$
(6)

$$= \sqrt{\frac{\omega'}{\omega}} \int_{-\infty}^{\infty} dv \, \exp(i\omega' v \mp i\omega g(v)),$$
(7)

being functionals of the trajectory, when $f(u)$ is replaced by $g(u)$ and consequently $g(v)$ is replaced by $f(v)$, transform to

$$\alpha^B_{\omega'\omega}[g] = \alpha^{B*}_{\omega\omega'}[f], \quad \beta^B_{\omega'\omega}[g] = -\beta^B_{\omega\omega'}[f]. \quad (8)$$

Similarly, the Bogolyubov coefficients for a fermion field $\alpha^F_{\omega'\omega}[f], \beta^{F*}_{\omega'\omega}[f]$

$$= \int_{-\infty}^{\infty} du \, \sqrt{f'(u)} \, \exp(\mp i\omega u + i\omega' f(u)) \quad (9)$$

$$= \int_{-\infty}^{\infty} dv \, \sqrt{g'(v)} \, \exp(i\omega' v \mp i\omega g(v)) \quad (10)$$

when the trajectory is replaced by its mirror reflection, transform to

$$\alpha^F_{\omega'\omega}[g] = \alpha^{F*}_{\omega\omega'}[f], \quad \beta^F_{\omega'\omega}[g] = \beta^F_{\omega\omega'}[f]. \quad (11)$$

The matrix notations for the Bogolyubov coefficients make it possible to write the transformations of Eqs. (8) and (11), i..e, the transition from trajectory $f(u)$ to $g(u)$, in the form

$$\alpha \to \alpha^+, \quad \beta \to \mp\tilde{\beta}, \quad (12)$$

where the upper and lower signs here and subsequently correspond to Bose and Fermi fields.

At the same time, according to the expansions given by Eqs. (33) and (34), $\alpha_{\omega'\omega}$ and $\beta_{\omega'\omega}$ are the amplitudes of the waves with frequencies ω' and $-\omega'$ contained in the incident part of the outgoing wave with frequency ω, while $\alpha^*_{\omega'\omega}$ and $\mp\beta_{\omega'\omega}$ are the amplitudes of the waves with frequencies ω and $-\omega$ contained in the reflected part of the ingoing wave with frequency ω'. Therefore, amplitudes $\alpha^*_{\omega'\omega}[g]$ and $\mp\beta_{\omega'\omega}[g]$ describe the generation by a right-sided mirror on trajectory $g(u)$ of waves escaping to the right with frequencies ω and $-\omega$ when waves with frequency ω' incident from right to left are absorbed. From purely geometrical considerations, they must coincide with the amplitudes for the mirror-symmetric process—the generation by a left-sided mirror on trajectory $f(u)$ of waves escaping to the left with frequencies ω and $-\omega$ when a wave incident from left to right with frequency ω' is absorbed. Then, according to Eqs. (8) and (11), these last are also equal to $\alpha_{\omega\omega'}[f]$ and $\beta_{\omega\omega'}[f]$ or equal to $\alpha_{\omega'\omega}[f]$ and $\beta_{\omega'\omega}[f]$ if the frequencies of the monochromatic waves propagating to the right and to the left are denoted as ω and ω', as was assumed for the field to the right of the trajectory.

Thus, for the field to the left of a mirror moving along trajectory $f(u)$, $\alpha_{\omega'\omega}[f]$ and $\beta_{\omega'\omega}[f]$ are the amplitudes of waves with frequencies ω' and $-\omega'$ contained in the reflected part of the ingoing wave with frequency ω, while $\alpha^*_{\omega'\omega}[f]$ and $\mp\beta_{\omega'\omega}[f]$ are the amplitudes of waves with frequencies ω and $-\omega$ contained in the incident part of the outgoing wave with frequency ω'. Therefore, the matrix that connects the ingoing and outgoing waves of the field to the left of the mirror differs from the analogous matrix for the field to the right of it by the transformation (12).

The transition from the trajectory $f(u)$ to the mirror-symmetric $g(u)$ is thus equivalent to considering the field in the part of the Minkowski plane not to the right but to the left of the trajectory $f(u)$ with the previous boundary condition on the mirror.

The mean number of particles formed by a two-sided infinitely thin mirror on the left part of the Minkowski plane is the same as on the right, since the integral

$$N=\int\int_0^\infty \frac{d\omega\, d\omega'}{(2\pi)^2}\, |\beta_{\omega'\omega}|^2 \qquad (13)$$

does not change when $\beta_{\omega'\omega}$ is replaced by $\mp\beta_{\omega'\omega}$. At the same time, the energy

$$\mathscr{E}'=\int\int_0^\infty \frac{d\omega\, d\omega'}{(2\pi)^2}\, \omega'|\beta_{\omega'\omega}|^2, \qquad (14)$$

emitted by the mirror to the left, general speaking, is not equal to the energy

$$\mathscr{E}=\int\int_0^\infty \frac{d\omega\, d\omega'}{(2\pi)^2}\, \omega|\beta_{\omega'\omega}|^2, \qquad (15)$$

emitted to the right.

The equality of the mean numbers of particles emitted by a two-sided accelerated mirror to the right and to the left suggests that the particles are generated in pairs and fly off in opposite directions. The quantity (5) is usually considered the mean number of actual quanta with frequency ω in the interval $d\omega$, emitted to the right when a quantum with frequency ω' in the interval $d\omega'$ is absorbed from the vacuum from the right. The question arises whether it is not possible to regard the same quantity as the mean number of pairs of quanta emitted to the right and to the left with frequencies ω and ω' in the intervals $d\omega$ and $d\omega'$, respectively. In other words, is $N^{-1}|\beta_{\omega'\omega}|^2$ the two-dimensional probability distribution of frequencies ω and ω' of two quanta escaping to the right and to the left with angular momenta ω and $-\omega'$?

Such an interpretation of the frequency distribution of bosons (fermions) emitted by a mirror in $1+1$ space would make the coincidence of this distribution with the emission spectrum of an electric (scalar) charge in $3+1$ space detected in Refs. 1 and 2 less formal. In the case of mirror emission, the frequencies ω and ω' of the bosons (fermions) escaping in different directions are random quantities, whereas in the case of charge emission, the components k_+ and k_- of the wave vector of the vector (scalar) quanta emitted to the right and to the left, corresponding to the $k_+ - k_- \gtrless 0$, are random quantities.

Let us give two more evidences of left–right symmetry of the wave field of an accelerated mirror that are reflected by the Bogolyubov coefficients.

First, Eqs. (6) and (7) or (9) and (10), obtained for the field to the right of the mirror for the Bogolyubov coefficients, represent $|\beta_{\omega'\omega}|^2$ by a double integral over the entire uv plane, as shown by Ref. 2. Thus

$$|\beta^B_{\omega'\omega}|^2=-\,\mathrm{Re}\int\int_{-\infty}^{\infty} du\, dv$$

$$\times \exp[i\omega u+i\omega' f(u)-i\omega' v-i\omega g(v)], \quad (16)$$

while $|\beta^F_{\omega'\omega}|^2$ differs from Eq. (16) by an additional factor of $-\sqrt{f'(u)g'(v)}$ under the integral. Similarly, in the double integral for the mean number of particles emitted to the right,

$$N^{B,F}=\frac{1}{4\pi^2}\int_{-\infty}^{\infty} du\, K^{B,F}(u),$$

$$K^B(u)=\!\!\!\fint_{-\infty}^{\infty}\frac{dv}{v-f(u)}\left[\frac{1}{g(v)-u}-\frac{f'(u)}{v-f(u)}\right], \qquad (17)$$

$$K^F(u)=-\sqrt{f'(u)}\int_{-\infty}^{\infty}\frac{dv}{v-f(u)}\left[\frac{\sqrt{g'(v)}}{g(v)-u}-\frac{\sqrt{f'(u)}}{v-f(u)}\right], \qquad (18)$$

the integration is carried out over the entire uv plane, i.e., over all of Minkowski space, and not over the part of it lying to the right of the trajectory. The wave fields to the right and to the left of the trajectory that satisfy the same condition on it are described by a single analytical function and therefore

are not independent. Therefore, the frequencies of the quanta emitted the right and to the left are also not independent.

Second, the mean energies $\mathscr{E}$ and $\mathscr{E}'$ emitted to the right and to the left, according to Ref. 2, can be represented as integrals over the proper time τ of the mirror:

$$\mathscr{E}^B = \frac{1}{12\pi} \int_{-\infty}^{\infty} [d\tau \, a^2 \sqrt{f'} - d(a\sqrt{f'})], \qquad (19)$$

$$\mathscr{E}'^B = \frac{1}{12\pi} \int_{-\infty}^{\infty} \left[d\tau \, \frac{a^2}{\sqrt{f'}} + d\left(\frac{a}{\sqrt{f'}}\right) \right]. \qquad (20)$$

Here a is the acceleration of the mirror in its proper system.

The first terms under the integral in Eqs. (19) and (20) represent the energy irreversibly emitted by the mirror respectively to the right and to the left of the section $d\tau$ of the trajectory. In the mirror's proper system, these portions of the energy are identical and equal $a^2 d\tau/12\pi$, whereas the portions of irreversibly emitted momentum equal $\pm a^2 d\tau/12\pi$. In the laboratory system, these portions of the energy, because they are moving oppositely relative to the velocity β of the source, acquire Doppler factors $\sqrt{f'}$ and $1/\sqrt{f'}$. We recall that $\sqrt{f'} = \sqrt{(1+\beta)/(1-\beta)}$. The second, Schott terms under the integrals in Eqs. (19) and (20) "smear" the region where the radiation develops, as a result of which, for the formation of the radiation of the energy, the intervals $\Delta\tau$ on which the irreversibly emitted energy exceeds the change of the Schott energy are important; i.e.,

$$\Delta\tau a^2 \sqrt{f'} > |a\sqrt{f'}|, \quad \Delta\tau a^2/\sqrt{f'} > |a/\sqrt{f'}|, \qquad (21)$$

or $\Delta\tau > a^{-1}$; the proper time interval must be greater than the inverse proper acceleration. The proper acceleration determines the characteristic frequency of the radiation in the proper system and its scatter, $\omega \sim \Delta\omega \sim a$. Therefore, the condition $\Delta\tau a > 1$ is equivalent to the uncertainty relation $\Delta\tau \Delta\omega > 1$.

3. SYMMETRY AND THE RELATIONS OF CERTAIN INTEGRAL QUANTITIES

The following representations for the mean number N of emitted particles and the mean emitted energies $\mathscr{E} = \mathscr{E}_+$ and $\mathscr{E}' = \mathscr{E}_-$ are convenient for explaining their properties with respect to certain space–time transformations:

$$N^B = \int_{-\infty}^{\infty} \int du \, dv \, S(u,v),$$

$$N^F = -\int_{-\infty}^{\infty} \int du \, dv \, \sqrt{f'(u)g'(v)} S(u,v), \qquad (22)$$

$$\mathscr{E}_{\pm}^B = \int_{-\infty}^{\infty} \int du \, dv \, A_{\pm}(u,v),$$

$$\mathscr{E}_{\pm}^F = -\int_{-\infty}^{\infty} \int du \, dv \, \sqrt{f'(u)g'(v)} A_{\pm}(u,v), \qquad (23)$$

where S and $A_{\pm}$ are singular functions of ($\varepsilon, \delta \to +0$):

$$S(u,v) = \frac{1}{8\pi^2} \left[\frac{1}{(v-f(u)-i\varepsilon)(g(v)-u-i\delta)} + \text{c.c.} \right], \qquad (24)$$

$$A_+(u,v) = \frac{1}{8\pi^2 i} \left[\frac{1}{(v-f(u)-i\varepsilon)(g(v)-u-i\delta)^2} - \text{c.c.} \right], \qquad (25)$$

$$A_-(u,v) = \frac{1}{8\pi^2 i}$$

$$\times \left[\frac{1}{(v-f(u)-i\varepsilon)^2(g(v)-u-i\delta)} - \text{c.c.} \right]. \qquad (26)$$

1. *Lorentz transformations.* The quantities $S(u,v)$, $\sqrt{f'(u)g'(u)}$, and $du \, dv$ are scalars with respect to the Lorentz transformations, while $A_{\pm}(u,v)$ transform as the $\pm$ components of a vector. Therefore, $N^{B,F}$ are Lorentz invariants, while $\mathscr{E}_{\pm}^{B,F}$ are the $\pm$ components of a vector.

2. *Mirror symmetry.* When the trajectory is replaced by a mirror-symmetric trajectory, $f(u) \to g(u)$, $g(v) \to f(v)$, the integrals $N[f]$ and $\mathscr{E}_{\pm}[f]$ transform, respectively, to

$$N[g] = N[f], \quad \mathscr{E}_{\pm}[g] = \mathscr{E}_{\mp}[f], \qquad (27)$$

since, for such a replacement,

$$S(u,v) \to S(v,u), \quad A_{\pm}(u,v) \to A_{\mp}(v,u),$$

$$\sqrt{f'(u)g'(v)} \to \sqrt{g'(u)f'(v)},$$

after which the transformed integrals N and $\mathscr{E}_{\pm}$ differ from the untransformed N and $\mathscr{E}_{\mp}$ only in the designation of the variables of integration. Thus, the mean numbers of particles emitted from the same trajectory to the right and to the left are identical and do not change when the trajectory is replaced by the mirror-symmetric one, while the mean energies emitted to the right and to the left are different and transform into each other when such replacement is made.

3. *Synchromirror transformation.* This discrete transformation consists of replacing coordinates u and v with the coordinates

$$\tilde{u} = g(v), \quad \tilde{v} = f(u), \quad \text{so that } u = g(\tilde{v}), \quad v = f(u). \qquad (28)$$

Points (u,v) and $(\tilde{u}, \tilde{v})$, which are related by the transformation (28), lie on the Minkowski plane on different sides of the trajectory of the mirror on the intersection of the light cones whose vertices are found on the trajectory at points $A(u,f(u))$ and $B(g(v),v)$. The points of any compact region lying to the right of the trajectory are mapped one-to-one into the points of a compact region lying to the left of the trajectory.

Functions $S(u,v)$ and $A_{\pm}(u,v)$ are form-invariant with respect to the transformation (28); i.e., their functional dependences on the new and old variables are identical:

$$S(u,v) \equiv S(g(\widetilde{v}), f(\widetilde{u})) = S(\widetilde{u}, \widetilde{v}),$$

$$A_{\pm}(u,v) = A_{\pm}(\widetilde{u}, \widetilde{v}). \qquad (29)$$

Since the area element $du\, dv\, \sqrt{f'(u) g'(v)}$ appearing in the Fermi integrals N^F and $\mathscr{E}_{\pm}^F$ is also form-invariant with respect to the transformation (28); i.e.,

$$du\ dv\ \sqrt{f'(u) g'(v)} = d\widetilde{u}\ d\widetilde{v}\ \sqrt{f'(\widetilde{u}) g'(\widetilde{v})}, \qquad (30)$$

the contributions to the Fermi intervals from any two regions on the uv plane related by the symmetry transformation (28) are identical. In particular, the contributions from the entire region to the right and the entire region to the left of the trajectory are identical.

In the Bose integrals N^B and $\mathscr{E}_{\pm}^B$, the contributions of the right-hand and left-hand regions related by the transformation (28) are, generally speaking, different, since the area element $du\, dv$ that appears in these integrals, unlike the functions S and $A_{\pm}$ being integrated, is mapped by the transformation (28) into the unequal element $d\widetilde{u}\, d\widetilde{v}$:

$$du\ dv = d\widetilde{u}\, d\widetilde{v}\, f'(\widetilde{u}) g'(\widetilde{v}),$$

$$d\widetilde{u}\, d\widetilde{v} = du\ dv\, f'(u) g'(v). \qquad (31)$$

Therefore, the contributions to the Bose integrals from these two elementary areas are proportional to their areas; i.e., their ratio equals the Jacobian of the transformation.

The transformation (28) of the variables of integration of course does not change the values of the integrals N and $\mathscr{E}_{\pm}$. Its meaning is that the local contributions to N and $\mathscr{E}_{\pm}$ from any pair of right-hand and left-hand regions associated by the transformation (28) have a definite symmetry or asymmetry. Namely, for Fermi integrals, this symmetry consists of the equality of such contributions, whereas, for Bose integrals, there is left–right asymmetry of the contributions, determined by the Jacobian of the transformation.

4. RADIATION OF A TWO-SIDED MIRROR, QUANTUM APPROACH

For a consistent description of the quantum wave field lying both to the right and to the left of the mirror and satisfying a single condition at the mirror, it is convenient to use the two complete sets $\{\phi_{\text{out}\,\omega}, \phi_{\text{out}\,\omega}^*\}$ and $\{\phi_{\text{in}\,\omega'}, \phi_{\text{in}\,\omega'}^*\}$ of solutions of the wave equation, given in Refs. 1 and 2. Possessing in the right-hand Minkowski plane the physical meaning of the out and in sets and satisfying the boundary condition at the mirror, these solutions can be smoothly extended into the left half-plane with no change of their functional form. However, in the left half-plane, these sets acquire the physical meaning of the in and out sets, respectively, and they must be designated as $\{\phi_{\text{in}\,\omega}, \phi_{\text{in}\,\omega}^*\}$ and $\{\phi_{\text{out}\,\omega'}, \phi_{\text{out}\,\omega'}^*\}$.

Each such solution is actually unambiguously characterized by the frequency ω or ω' of its monochromatic component travelling to the right or to the left and by the condition at the mirror. For a Lorentzian transformation with velocity β along the x axis, the frequencies ω and ω' transform into $\widetilde{\omega}$ and $\widetilde{\omega}'$ according to the mutually inverse laws

$$\widetilde{\omega} = D^{-1}(\beta)\omega, \quad \widetilde{\omega}' = D(\beta)\omega', \quad D(\beta) = \sqrt{\frac{1+\beta}{1-\beta}}, \qquad (32)$$

where $D(\beta)$ is the Doppler factor. Thus, ω and ω' possess opposite covariance. Below, frequencies that transform like ω will be labelled with an even number of primes, while those that transform like ω' will have an odd number. Then the subscript in or out, in addition to the frequency, will simply indicate the side of the Minkowski plane on which the solution is considered.

We have written the expansion of the solutions of the first set in the solutions of the second set and the inverse expansion (in the right-hand half-plane) in the form

$$\phi_{\text{out}\,\omega} = \alpha_{\omega'\,\omega}\phi_{\text{in}\,\omega'} + \beta_{\omega'\,\omega}\phi_{\text{in}\,\omega'}^*, \qquad (33)$$

$$\phi_{\text{in}\,\omega'} = \alpha_{\omega'\,\omega}^*\phi_{\text{out}\,\omega} \mp \beta_{\omega'\,\omega}\phi_{\text{out}\,\omega}^*, \qquad (34)$$

or, if matrix notation is used,

$$\begin{pmatrix} \phi_{\text{out}} \\ \phi_{\text{out}}^* \end{pmatrix} = \begin{pmatrix} \widetilde{\alpha} & \widetilde{\beta} \\ \beta^+ & \alpha^+ \end{pmatrix} \begin{pmatrix} \phi_{\text{in}} \\ \phi_{\text{in}}^* \end{pmatrix},$$

$$\begin{pmatrix} \phi_{\text{in}} \\ \phi_{\text{in}}^* \end{pmatrix} = \begin{pmatrix} \alpha^* & \mp\beta \\ \mp\beta^* & \alpha \end{pmatrix} \begin{pmatrix} \phi_{\text{out}} \\ \phi_{\text{out}}^* \end{pmatrix}. \qquad (35)$$

As a consequence of the orthogonality and normalization of the solutions in both sets, the matrices that appear in Eqs. (35) are mutually inverse. This means that the Bogolyubov coefficients satisfy four independent matrix relations:

$$\alpha^+\alpha \mp \beta^+\beta = 1, \quad \beta^+\alpha^* \mp \alpha^+\beta^* = 0,$$

$$\alpha\alpha^+ \mp \beta^*\widetilde{\beta} = 1, \quad \alpha\beta^+ \mp \beta^*\widetilde{\alpha} = 0. \qquad (36)$$

On the left-hand half-plane, Eqs. (33)–(35) are conserved, but a new physical meaning requires the interchange of the subscripts in $\rightleftarrows$ out in the functions, which is equivalent to the transformation (12).

For a quantized field in the right half-plane, the connection of the in and out creation operator a^+ and absorption operator a is given by the Bogolyubov transformations

$$\begin{pmatrix} a_{\text{in}} \\ a_{\text{in}}^+ \end{pmatrix} = \begin{pmatrix} \alpha & \beta^* \\ \beta & \alpha^* \end{pmatrix} \begin{pmatrix} a_{\text{out}} \\ a_{\text{out}}^+ \end{pmatrix},$$

$$\begin{pmatrix} a_{\text{out}} \\ a_{\text{out}}^+ \end{pmatrix} = \begin{pmatrix} \alpha^+ & \mp\beta^+ \\ \mp\widetilde{\beta} & \widetilde{\alpha} \end{pmatrix} \begin{pmatrix} a_{\text{in}} \\ a_{\text{in}}^+ \end{pmatrix}. \qquad (37)$$

For a field in the left-hand half-plane, an interchange of the subscripts in $\rightleftarrows$ out is required on operators a and a^+ in transformations (37). This again is equivalent to the transformation (12).

Following the work of DeWitt[3] and its notation, we represent the vector of the vacuum state of the field in the distant past in the form of an expansion in the vectors of the n-particle states of the field in the distant future:

$$|\text{in}\rangle = e^{iW} \sum_{n=0}^{\infty} \frac{i^{n/2}}{n!} \sum_{i_1 i_2 \ldots i_n} V_{i_1 i_2 \ldots i_n} |i_1 i_2 \ldots i_n \text{out}\rangle. \qquad (38)$$

In our case, the quantum numbers $i_1 i_2 \ldots i_n$ of the out states of the individual particles stand for frequencies that transform like ω or like ω' if one is dealing with the field, respectively, to the right or to the left of the mirror.

Using the equation $a_{\rm in} |{\rm in}\rangle = 0$, transformations (37), and the expansion given by Eq. (38), it is easy to show[3,4] that the relative production amplitudes $V_{i_1 i_2 \ldots i_n}$ of n particles equal zero for odd n, whereas, for even n, they are expressed in terms of the production amplitude of a pair of particles:

$$V_{i_1 i_2 \ldots i_n} = \sum_p \delta_p V_{i_1 i_2} V_{i_3 i_4} \ldots V_{i_{n-1} i_n}. \tag{39}$$

Here Σ_p denotes summation over $n!/2^{n/2}(n/2)!$ different pairings of subscripts $i_1 i_2 \ldots i_n$, while $\delta_p = 1$ for bosons and $\delta_p = \pm 1$ for fermions when the permutation leading to the given pairing is, respectively, even or odd. The production amplitudes of a pair of particles with frequencies ω'' and ω in the right-hand region and frequencies ω''' and ω' in the left-hand region equal

$$V_{\omega'' \omega} = i(\alpha^{-1} \beta^*)_{\omega'' \omega}, \quad V_{\omega''' \omega'} = -i(\beta \alpha^{-1})^*_{\omega''' \omega'}. \tag{40}$$

They are related to each other by the transformation (12), which is symmetric for a Bose field and antisymmetric for a Fermi field, as follows from Eqs. (36).

The indicated number of terms in the amplitude of Eq. (39) appears in connection with its symmetrization (antisymmetrization) and equals the number $n!$ of permutations of its subscripts, reduced by a factor of $2^{n/2}$ because of the already existing symmetry (antisymmetry) of the two-particle amplitudes and by a factor of $(n/2)!$ because the permutations of these amplitudes are negligible.

Particle production in pairs is explained by the linearity of the Bogolyubov transformations in the operators a and a^+. The operator $a_{\rm in}$, when it acts on the n-particle out state, transforms it into a superposition of $n-1$-particle and $n+1$-particle out states. Therefore, in the expansion of the null vector $a_{\rm in} |{\rm in}\rangle$ in the n-particle out states, the expansion coefficients equal to zero represent the linear relation between the amplitudes of the $n+1$-particle and $n-1$-particle creations. Since $n \geq 0$, the amplitude of the single-particle production V_{i_1} is equal to zero, and, along with it, all the formation amplitudes of an odd number of particles.

The absolute amplitudes of the n-particle production are determined and are related to the relative amplitudes by

$$\langle {\rm out} \, i_1 i_2 \ldots i_n | {\rm in} \rangle \equiv \langle {\rm out} | a_{{\rm out} \, i_n} \ldots a_{{\rm out} \, i_2} a_{{\rm out} \, i_1} | {\rm in} \rangle$$

$$= e^{iW} i^{n/2} V_{i_1 i_2 \ldots i_n}. \tag{41}$$

The vacuum-conservation amplitude $\langle {\rm out} | {\rm in} \rangle = e^{iW}$ is determined to within a phase factor by the fact that the total probability of the transition from the initial vacuum state is equal to unity:

$$1 = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{i_1 i_2 \ldots i_n} |\langle {\rm out} \, i_1 i_2 \ldots i_n | {\rm in} \rangle|^2$$

$$= e^{-2 \, {\rm Im} \, W} \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{i_1 i_2 \ldots i_n} |V_{i_1 i_2 \ldots i_n}|^2. \tag{42}$$

The sum of the relative probabilities

$$q_n = \frac{1}{n!} \sum_{i_1 i_2 \ldots i_n} |V_{i_1 i_2 \ldots i_n}|^2 \tag{43}$$

of the production of n particles (or of $n/2$ pairs) on the right-hand side of Eq. (42) we shall call the sum over states. It can be shown that, in the case considered here, in which pairs of identical particles and antiparticles are formed, the sum over states equals

$$\sum_{n=0}^{\infty} \frac{1}{n!} \sum_{i_1 i_2 \ldots i_n} |V_{i_1 i_2 \ldots i_n}|^2 = \det(1 \mp M)^{\mp 1/2}$$

$$= \exp\left(\mp \frac{1}{2} \, {\rm Tr} \, \ln(1 \mp M) \right), \tag{44}$$

where $M = VV^+$ is a Hermitian positive-semidefinite matrix formed from the matrices in Eqs. (40). In particular, the first three terms of the sum over states, determined by the relative amplitudes

$$1, \quad V_{i_1 i_2}, \quad V_{i_1 i_2} V_{i_3 i_4} \pm V_{i_1 i_3} V_{i_2 i_4} + V_{i_1 i_4} V_{i_2 i_3}, \tag{45}$$

and by Eq. (43), are equal, respectively, to

$$q_0 = 1, \quad q_2 = \frac{1}{2} \, {\rm Tr} \, M, \quad q_4 = \frac{1}{8} \, ({\rm Tr} \, M)^2 \pm \frac{1}{4} \, {\rm Tr} \, M^2. \tag{46}$$

The absolute probabilities of forming n pairs are equal to $p_{2n} = p_0 q_{2n}$, where p_0 is the vacuum-conservation probability:

$$p_0 = e^{-2 \, {\rm Im} \, W}, \quad 2 \, {\rm Im} \, W = \mp \frac{1}{2} \, {\rm Tr} \, \ln(1 \mp M). \tag{47}$$

Since the relative probabilities $q_{2n}(M)$ of producing n pairs are homogeneous functions of degree n, $q_{2n}(\lambda M) = \lambda^n q_{2n}(M)$, it is convenient to compute the mean number of pairs from

$$\bar{n} = \sum_{n=0}^{\infty} n p_{2n} = p_0 \lambda \frac{\partial}{\partial \lambda} \sum_{n=0}^{\infty} \lambda^n q_{2n}(M) \bigg|_{\lambda=1}$$

$$= \lambda \frac{\partial}{\partial \lambda} 2 \, {\rm Im} \, W(\lambda M) \bigg|_{\lambda=1} = \frac{1}{2} \, {\rm Tr} \, \frac{M}{1 \mp M}. \tag{48}$$

The matrices M are different for the right-hand and left-hand regions:

$$M = VV^+ = \begin{cases} \beta^+ \beta (1 \pm \beta^+ \beta)^{-1}, & \text{(49)} \\ \beta^* \tilde{\beta} (1 \pm \beta^* \tilde{\beta})^{-1}, & \text{(50)} \end{cases}$$

but are related to each other by the transformation (12). However, the positive-definite quantities $\,{\rm Tr} \, M^n$, $n = 1,2\ldots$, are invariants of this transformation. Therefore, the total probabilities given above for conservation of the vacuum,

p_0, and of the production of n pairs, p_{2n}, and the mean number of pairs, $\bar{n}$, are identical for the right-hand and left-hand regions. In particular, the quantities

$$p_0 = e^{-2\,\mathrm{Im}\,W}, \quad 2\,\mathrm{Im}\,W = \pm\frac{1}{2}\,\mathrm{Tr}\,\ln(1 \pm \beta^+\beta), \quad (51)$$

$$p_2 = e^{-2\,\mathrm{Im}\,W}\frac{1}{2}\,\mathrm{Tr}\,\beta^+\beta(1 \pm \beta^+\beta)^{-1}, \quad (52)$$

$$\bar{n} = \frac{1}{2}\,\mathrm{Tr}\,\beta^+\beta \quad (53)$$

do not change under the transformation (12) or $\beta^+\beta \to \beta\beta^+$.

Nevertheless, the frequency distributions of the probabilities and of the mean number of particles possess no left–right symmetry. Thus, the production probability of one pair, one particle of which has a definite frequency while the other has any frequency, equals

$$p_{2\omega} = e^{-2\,\mathrm{Im}\,W}\left(\frac{\beta^+\beta}{1 \pm \beta^+\beta}\right)_{\omega\omega} \quad (54)$$

for the right-hand region and equals

$$p_{2\omega'} = e^{-2\,\mathrm{Im}\,W}\left(\frac{\beta\beta^+}{1 \pm \beta\beta^+}\right)_{\omega'\omega'} \quad (55)$$

for the left-hand region. The frequency distributions of the mean number of particles emitted by the mirror to the right and to the left are also functionally different from each other:

$$N_\omega = (\beta^+\beta)_{\omega\omega}, \quad N_{\omega'} = (\beta\beta^+)_{\omega'\omega'}. \quad (56)$$

Along with the amplitudes given by Eq. (41) for particle production from vacuum by the mirror, it is necessary to consider the amplitudes of single-particle scattering by the mirror

$$\langle\mathrm{out}\,\omega|\omega'\,\mathrm{in}\rangle = \langle\mathrm{out}\,|a_{\mathrm{out}\,\omega}a_{\mathrm{in}\,\omega'}^+|\mathrm{in}\rangle = e^{iW}\alpha_{\omega\omega'}^{-1}, \quad (57)$$

$$\langle\mathrm{out}\,\omega'|\omega\,\mathrm{in}\rangle = \langle\mathrm{out}|a_{\mathrm{out}\,\omega'}a_{\mathrm{in}\,\omega}^+|\mathrm{in}\rangle = e^{iW}\alpha_{\omega\omega'}^{-1*} \quad (58)$$

for the right-hand and left-hand regions, respectively. These amplitudes differ only in their phases. Of course, they are related to each other by the transformation (12), but we shall be interested in their relation to the corresponding pair-production amplitudes:

$$\langle\mathrm{out}\,\omega''\omega|\mathrm{in}\rangle = -e^{iW}(\alpha^{-1}\beta^*)_{\omega''\omega}$$

$$= -\sum_{\omega'}\langle\mathrm{out}\,\omega''|\omega'\,\mathrm{in}\rangle\beta_{\omega'\omega}^*, \quad (59)$$

$$\langle\mathrm{out}\,\omega'\omega'''|\mathrm{in}\rangle = e^{iW}(\beta\alpha^{-1})_{\omega'\omega'''}^*$$

$$= \sum_\omega \beta_{\omega'\omega}^*\langle\mathrm{out}\,\omega'''|\omega\,\mathrm{in}\rangle. \quad (60)$$

Since the pair-production amplitudes and the single-particle scattering amplitudes are quantities that can in principle be experimentally measured from the corresponding probabilities, Eqs. (59) and (60) make it possible to experimentally measure $\beta_{\omega'\omega}^*$. Moreover, these relationships make

it possible to regard $\beta_{\omega'\omega}^*$ as the amplitude of the source of a pair of particles potentially emitted to the right and to the left with frequencies ω and ω', respectively. In this case, if a particle with frequency ω actually escaped to the right, a particle with frequency ω' does not escape to the left, but experiences an internal reflection and is actually emitted to the right with altered frequency ω''. Conversely, if a particle with frequency ω' actually escaped to the left, a particle with frequency ω cannot escape to the right, but, after internal reflection, is actually emitted to the left with another frequency ω'''.

For fermions, the amplitude $\beta_{\omega'\omega}^F$ is diagonal in the projection of the spin of the ingoing and outgoing waves (see Ref. 2). But one of the waves forming $\beta_{\omega'\omega}^F$ has a negative frequency and therefore describes an antiparticle with frequency and spin projection opposite in sign to the frequency and spin projection of this wave (see §26 in Ref. 5 or §9 of chap. 2 in Ref. 6). Thus, the spin of a pair of generated fermions equals zero. This is confirmed by the scalar nature of the identically equal integrals in Eqs. (9) and (10), in which $du\sqrt{f'(u)}$ and $dv\sqrt{g'(v)}$ are elements of proper time $d\tau$, and by their coincidence,

$$\beta_{\omega'\omega}^{F*} = \frac{1}{e}\,\rho(k_+, k_-), \quad (61)$$

with the Fourier component of the scalar-charge density in $3+1$ space.

The amplitude $\beta_{\omega'\omega}^{B*}$ of the source of a boson pair, according to Eqs. (6) and (7), is linearly expressed in terms of the Fourier components $j_\pm(k)$ of the current density of an electric charge in $3+1$ space:

$$\beta_{\omega'\omega}^{B*} = -\sqrt{\frac{k_+}{k_-}}\frac{j_-}{e} = \sqrt{\frac{k_-}{k_+}}\frac{j_+}{e}, \quad (62)$$

$$j_- = e\int_{-\infty}^{\infty} du\,\exp\left[\frac{i}{2}(k_+u + k_-f(u))\right],$$

$$j_+ = e\int_{-\infty}^{\infty} dv\,\exp\left[\frac{i}{2}(k_-v + k_+g(v))\right] \quad (63)$$

[see also Eqs. (1) and (2) in this paper and Eqs. (43) and (44) in Ref. 1]. The last equality in Eq. (62) is none other than the current-transverseness condition, $k_+j_- + k_-j_+ = 0$. It can also be seen from Eq. (62) that $\beta_{\omega'\omega}^B$ is a pseudoscalar, since, with the reflection $k_\pm \to k_\mp$, $j_\pm \to j_\mp$, and β^B changes sign. Vector $j_\alpha(k)$ is spacelike and, in a system where $k_+ = k_-$ (or $\omega = \omega'$), has only a spatial component, precisely equal to $e\beta_{\omega'\omega}^B$. In covariant form,

$$e\beta_{\omega'\omega}^{B*} = \varepsilon_{\alpha\beta}k^\alpha j^\beta/\sqrt{k_+k_-}.$$

Thus, the source of a boson pair is the conserved current vector given by Eqs. (63), and this means that its spin equals 1.[7]

The fact that the spin of a boson pair equals 1 while that of a fermion pair equals 0 is essential for understanding the coincidence of the spectra of a mirror and of a charge.

If $\beta^*_{\omega'\omega}$ is small, i.e., if the mean number of emitted quanta is small, then, as is easy to obtain from Eqs. (6) and (9),

$$\alpha_{\omega'\omega} \approx 2\pi\delta(\tilde{\omega}'-\tilde{\omega}), \quad \alpha^{-1}_{\omega\omega'} \approx 2\pi\delta(\tilde{\omega}-\tilde{\omega}'), \qquad (64)$$

where $\tilde{\omega}$ and $\tilde{\omega}'$ are related to ω and ω' by transformation (32), in which β is the effective velocity of the mirror on the emission section. In this approximation, the emission amplitudes given by Eqs. (59) and (60) for pairs of particles with frequencies ω and ω'' to the right and pairs of particles with frequencies of ω' and ω''' to the left equal, respectively,

$$\langle \text{out } \omega''\omega | \text{in} \rangle \approx -e^{iW} D^{-1}(\beta) \beta^*_{\omega'\omega}, \quad \omega' = D^{-2}(\beta)\omega'', \qquad (65)$$

$$\langle \text{out } \omega'\omega''' | \text{in} \rangle \approx e^{iW} D(\beta) \beta^*_{\omega'\omega}, \quad \omega = D^2(\beta)\omega'''. \qquad (66)$$

These formulas, including the connection between the frequencies of the waves incident on the mirror and reflected from it, confirm the interpretation of $\beta^*_{\omega'\omega}$ given above.

We now turn our attention to interference effects in the production of Bose and Fermi particles. They become most substantial when the matrices M for bosons and fermions satisfy the conditions

$$\mp \frac{1}{2} \text{Tr} \ln(1\mp M) = \mp \ln\left(1\mp\frac{1}{2}\text{Tr } M\right),$$

i.e.,

$$\frac{1}{2}\text{Tr } M^n = \left(\frac{1}{2}\text{Tr } M\right)^n, \quad n=2,3,\dots. \qquad (67)$$

Then the sum over states given by Eq. (44) for Bose and Fermi particles reduces, respectively, to

$$\frac{1}{1-(1/2)\text{Tr } M} \quad \text{and} \quad 1+\frac{1}{2}\text{Tr } M. \qquad (68)$$

This means that the probabilities of producing n pairs of bosons form the geometrical progression

$$p^B_{2n} = p^B_0 q^{Bn}_2, \quad p^B_0 = 1-\frac{1}{2}\text{Tr } M, \quad q^B_2 = \frac{1}{2}\text{Tr } M, \qquad (69)$$

while the probabilities of emitting two or more pairs of fermions disappear; i.e., only the production of one fermion pair is possible:

$$p^F_0 = \left(1+\frac{1}{2}\text{Tr } M\right)^{-1}, \quad p^F_2 = p_0 \frac{1}{2}\text{Tr } M,$$

$$p^F_{2n} = 0, \quad n \geq 2. \qquad (70)$$

In other words, the conditions given by Eqs. (67) denote the most constructive interference of bosons and the most destructive interference of fermions. In these cases, the mean number of boson pairs is always greater than 1, while that of the fermion pairs is less than 1:

$$1 < \bar{n}^B = \frac{(1/2)\text{Tr } M}{1-(1/2)\text{Tr } M} < \infty,$$

$$0 \leq \bar{n}^F = \frac{(1/2)\text{Tr } M}{1+(1/2)\text{Tr } M} < 1. \qquad (71)$$

We are less interested in the case in which interference effects can be neglected. In this case,

$$\text{Tr } M^k \ll \text{Tr } M, 1; \quad k \geq 2, \qquad (72)$$

and the probability distribution over the number of generated pairs coincides with the Poisson distribution:

$$p_{2n} = e^{-\bar{n}} \frac{(\bar{n})^n}{n!}, \quad \bar{n} = \frac{1}{2}\text{Tr } \beta^+\beta. \qquad (73)$$

5. EMISSION OF PAIRS CONSISTING OF NONIDENTICAL PARTICLES AND ANTIPARTICLES

In the case of pair production of nonidentical particles and antiparticles (ab pairs), the direct and inverse Bogolyubov transformations (37) are replaced by

$$\begin{pmatrix} a_{\text{in}} \\ b^+_{\text{in}} \end{pmatrix} = \begin{pmatrix} \alpha_{aa} & \beta^*_{ab} \\ \beta_{ba} & \alpha^*_{bb} \end{pmatrix} \begin{pmatrix} a_{\text{out}} \\ b^+_{\text{out}} \end{pmatrix},$$

$$\begin{pmatrix} a_{\text{out}} \\ b^+_{\text{out}} \end{pmatrix} = \begin{pmatrix} \alpha^+_{aa} & \mp\beta^+_{ba} \\ \mp\tilde{\beta}_{ab} & \tilde{\alpha}_{bb} \end{pmatrix} \begin{pmatrix} a_{\text{in}} \\ b^+_{\text{in}} \end{pmatrix}. \qquad (74)$$

These transformations contain not two but four matrices α_{aa}, α_{bb}, β_{ab}, and β_{ba}, which satisfy not the four Eqs. (36) but the six equations

$$\alpha^+_{aa}\alpha_{aa} \mp \beta^+_{ba}\beta_{ba} = 1, \quad \alpha^+_{bb}\alpha_{bb} \mp \beta^+_{ab}\beta_{ab} = 1,$$

$$\beta^+_{ba}\alpha^*_{bb} \mp \alpha^+_{aa}\beta^*_{ab} = 0, \quad \alpha_{aa}\alpha^+_{aa} \mp \beta^*_{ab}\tilde{\beta}_{ab} = 1,$$

$$\alpha_{bb}\alpha^+_{bb} \mp \beta^*_{ba}\tilde{\beta}_{ba} = 1, \quad \alpha_{aa}\beta^+_{ba} \mp \beta^*_{ab}\tilde{\alpha}_{bb} = 0. \qquad (75)$$

However, these relationships can be written in the form of Eqs. (36) if α and β stand for the 2×2 matrices consisting of the indicated quartets:

$$\alpha = \begin{pmatrix} \alpha_{aa} & 0 \\ 0 & \alpha_{bb} \end{pmatrix}, \quad \beta = \begin{pmatrix} 0 & \beta_{ab} \\ \beta_{ba} & 0 \end{pmatrix}. \qquad (76)$$

As can be seen from Eqs. (74), the interchange in$\rightleftarrows$out is now equivalent to the interchange

$$\alpha_{aa} \rightarrow \alpha^+_{aa}, \quad \alpha_{bb} \rightarrow \alpha^+_{bb}, \quad \beta_{ab} \rightarrow \mp\tilde{\beta}_{ba},$$

$$\beta_{ba} \rightarrow \mp\tilde{\beta}_{ab}, \qquad (77)$$

which can be represented in the form of the transformation (12) if α and β stand for the matrices of Eqs. (76).

Using for the in vacuum state an expansion of the type of Eq. (38) and the equations $a_{\text{in}}|\text{in}\rangle = b_{\text{in}}|\text{in}\rangle = 0$, it can be shown that all the emission amplitudes of an odd number of particles equal zero, while the production amplitudes of an even number of particles are products of the production amplitudes of ab pairs:

$$V^{ab}_{\omega''\omega} = i(\alpha^{-1}_{aa}\beta^*_{ab})_{\omega''\omega}, \quad V^{ab}_{\omega'''\omega'} = -i(\beta_{ab}\alpha^{-1}_{bb})^*_{\omega'''\omega'}, \qquad (78)$$

respectively for the right-hand and left-hand regions. As follows from Eqs. (75), the amplitudes given by Eqs. (78) possess intrinsic Bose symmetry or Fermi antisymmetry:

$$V^{ab}_{\omega''\omega} = \pm V^{ba}_{\omega\omega''} \equiv \pm i(\alpha^{-1}_{bb}\beta^*_{ba})_{\omega\omega''},$$

$$V^{ab}_{\omega'''\omega'} = \pm V^{ba}_{\omega'\omega'''} \equiv \mp i(\beta_{ba}\alpha^{-1}_{aa})^*_{\omega'\omega'''}. \qquad (79)$$

Thus, the formation amplitude of an ab pair can be denoted via $V_{i_1 i_2}$, where the subscript i_1 characterizes the state of the particle and i_2 that of the antiparticle. The production of two ab pairs is described by the amplitude

$$V_{i_1 i_2 i_3 i_4} = V_{i_1 i_2} V_{i_3 i_4} \pm V_{i_3 i_2} V_{i_1 i_4}, \qquad (80)$$

symmetric (antisymmetric) separately with respect to states i_1 and i_3 of the particles and separately with respect to states i_2 and i_4 of the antiparticles. We also write the production amplitude of three pairs:

$$V_{i_1 i_2 \dots i_6} = V_{i_1 i_2} V_{i_3 i_4} V_{i_5 i_6} \pm V_{i_3 i_2} V_{i_1 i_4} V_{i_5 i_6}$$
$$+ V_{i_3 i_2} V_{i_5 i_4} V_{i_1 i_6} \pm V_{i_1 i_2} V_{i_5 i_4} V_{i_3 i_6}$$
$$+ V_{i_5 i_2} V_{i_1 i_4} V_{i_3 i_6} \pm V_{i_5 i_2} V_{i_3 i_4} V_{i_1 i_6}. \qquad (81)$$

In the general case, the production amplitude of $n/2$ pairs has the form

$$V_{i_1 i_2 \dots i_n} = \sum_p \delta_p V_{i_1 i_2} V_{i_3 i_4} \dots V_{i_{n-1} i_n}, \qquad (82)$$

where the sum is taken over all $(n/2)!$ terms that differ by a permutation of the odd subscripts (or, what is the same thing, by a permutation of the even subscripts), with $\delta_p = \pm 1$ in the case of fermions for an even or odd permutation, respectively, while $\delta_p = 1$ in the case of bosons. Then amplitude $V_{i_1 i_2 \dots i_n}$ will be symmetric (antisymmetric) both over particle states $i_1 i_3 \dots i_{n-1}$ and over antiparticle states $i_2 i_4 \dots i_n$.

The relative probability

$$q_n = \frac{1}{(n/2)!(n/2)!} \sum_{i_1 i_2 \dots i_n} |V_{i_1 i_2 \dots i_n}|^2 \qquad (83)$$

of producing $n/2$ pairs consisting of nonidentical particles and antiparticles contains the factor $1/(n/2)!(n/2)!$, which, along with the symmetry (antisymmetry) of amplitude $V_{i_1 i_2 \dots i_n}$ separately for even and separately for odd subscripts, makes it possible to sum over the particle and antiparticle states, considering the ranges of variation of the quantum numbers of these states to be independent. Without this factor, the sum over $i_1 i_2 \dots i_n$ would have had to contain only physically different states. In our case, for example, it would be unambiguous that the frequencies of the particles must satisfy the condition $\omega_1 \geq \omega_3 \geq \dots \geq \omega_{n-1}$, while the frequencies of the antiparticles must satisfy the condition $\omega_2 \geq \omega_4 \geq \dots \geq \omega_n$.

It is easy to construct the first four terms of the sum over states in terms of the relative amplitudes shown above:

$$q_0 = 1, \quad q_2 = \mathrm{Tr}\, M, \quad q_4 = \frac{1}{2}(\mathrm{Tr}\, M)^2 \pm \frac{1}{2}\mathrm{Tr}\, M^2,$$

$$q_6 = \frac{1}{6}(\mathrm{Tr}\, M)^3 \pm \frac{1}{2}\mathrm{Tr}\, M\,\mathrm{Tr}\, M^2 + \frac{1}{3}\mathrm{Tr}\, M^3. \qquad (84)$$

For the sum over states as a whole, we get

$$\sum_{n=0}^{\infty} \frac{1}{(n/2)!(n/2)!} \sum_{i_1 i_2 \dots i_n} |V_{i_1 i_2 \dots i_n}|^2$$
$$= \det(1 \mp M)^{\mp 1} = \exp(\mp \mathrm{Tr}\, \ln(1 \mp M)). \qquad (85)$$

Here, as in Eq. (44), $M = VV^+$ is a Hermitian positive-semidefinite matrix. It is given by Eqs. (49) and (50), in which by β is meant, respectively, β_{ba} and β_{ab}.

Just as above, the absolute probabilities of the formation of n pairs of nonidentical particles and antiparticles equal $p_{2n} = p_0 q_{2n}$, where p_0 is the vacuum-conservation probability:

$$p_0 = e^{-2\,\mathrm{Im}\, W}, \quad 2\,\mathrm{Im}\, W = \mp \mathrm{Tr}\,\ln(1 \mp M). \qquad (86)$$

The mean number of pairs, computed according to the rule given in Eq. (48), equals

$$\bar{n} = \mathrm{Tr}\, \frac{M}{1 \mp M}. \qquad (87)$$

It can be seen that these formulas differ from the corresponding Eqs. (47) and (48) for pair production of identical particles by replacing $(1/2)\mathrm{Tr}$ by Tr in the latter equations. Because of the $a \rightleftarrows b$ symmetry in the matrices, under the Tr sign, β can stand for both β_{ba} and β_{ab}.

It is easy to see that this rule connects all the formulas for the integral characteristics of pair production of identical particles with the formulas of the corresponding characteristics of ab-pair production. Thus, in order to obtain from Eqs. (51)–(53) and (67)–(73) the analogous expressions for ab-pair production, it is sufficient to replace $(1/2)\mathrm{Tr}$ in these formulas with Tr and by β to understand β_{ba} or β_{ab}.

As far as the spectral characteristics shown, for example, in Eqs. (54)–(56) are concerned, they undergo no changes when the transition is made to the case under consideration, if β represent β_{ba} (β_{ab}) for the spectrum of particles (antiparticles) emitted to the right and β_{ab} (β_{ba}) for the spectrum of particles (antiparticles) emitted to the left.

In fact, for the differential probability $p_{2\omega}$ shown in Eq. (54), the original integral

$$p_{2\omega} = \int_0^{\infty} \frac{d\omega''}{2\pi} |\langle \text{out}\ \omega\omega''\text{in}\rangle|^2 \qquad (88)$$

represents it as the sum of the probabilities of physically different events regardless of whether the particles are identical or not. However, the total pair-formation probability p_2 as a sum of probabilities of physically different events for identical particles is represented by the integral

$$p_2 = \int_0^{\infty} \frac{d\omega}{2\pi} \int_0^{\omega} \frac{d\omega''}{2\pi} |\langle \text{out}\ \omega\omega''|\text{in}\rangle|^2$$
$$= \frac{1}{2} \int_0^{\infty} \frac{d\omega}{2\pi} p_{2\omega}, \qquad (89)$$

since the states in this case differ only by the values of the large frequency ω and the small frequency ω'' of two identical particles. At the same time, for an ab pair,

$$p_2 = \int_0^\infty \frac{d\omega}{2\pi} \int_0^\infty \frac{d\omega''}{2\pi} \, |\langle \text{out } \omega\omega''|\text{in}\rangle|^2 = \int_0^\infty \frac{d\omega}{2\pi} \, p_{2\omega},$$

(90)

since the states differ in the frequencies ω'' and ω of the particle and antiparticle independently of each other, while the particle and antiparticle differ in turn in that they interact differently with the counters.

Turning to the amplitude of ab-pair production,

$$\langle \text{out}|b_{\text{out } \omega''} a_{\text{out } \omega}|\text{in}\rangle \equiv \langle \text{out } \omega\omega''|\text{in}\rangle$$

$$= -e^{iW}(\alpha_{aa}^{-1}\beta_{ab}^*)_{\omega\omega''}$$

$$= \mp e^{iW}(\alpha_{bb}^{-1}\beta_{ba}^*)_{\omega''\omega},$$

(91)

we note that it reduces to a product of the source amplitude β_{ab}^* or β_{ba}^* of oppositely directed a and b particles and the backscattering amplitude α_{aa}^{-1} or α_{bb}^{-1} of one of them, as a result of which both particles of the pair move in the same direction. The symmetry of Eqs. (79) makes it impossible to establish which of the particles of the ab pair experiences backscattering.

*)E-mail: ritus@lpi.ac.ru

[1] A. I. Nikishov and V. I. Ritus, Zh. Éksp. Teor. Fiz. **108**, 1121 (1995) [JETP **81**, 615 (1995)].

[2] V. I. Ritus, Zh. Éksp. Teor. Fiz. **110**, 526 (1996) [JETP **83**, 282 (1996)].

[3] B. S. DeWitt, Phys. Rep. C **19**, 295 (1975).

[4] R. M. Wald, Commun. Math. Phys. **45**, 9 (1975).

[5] V. B. Berestetskiĭ, E. M. Lifshitz, and L. P. Pitaevskiĭ, *Relativistic Quantum Theory* (Nauka, Moscow, 1989; Pergamon Press, Oxford, 1971).

[6] A. I. Akhiezer and V. B. Berestetskiĭ, *Quantum Electrodynamics* (Nauka, Moscow, 1969; Wiley, New York, 1965)

[7] J. Schwinger, *Particles, Sources, and Fields* (Addison-Wesley, Reading, Mass, 1970; Mir, Moscow, 1976).

Translated by W. J. Manthey

Hund's Case (a)-case (b) transition in the singlet–triplet absorption spectrum of pyrazine in a supersonic jet

E. S. Medvedev*)

Institute of Chemical Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

D. W. Pratt

Department of Chemistry, University of Pittsburgh, Pittsburgh, PA 15260, USA
(Submitted 6 January 1998)
Zh. Éksp. Teor. Fiz. **114**, 63–90 (July 1998)

An analytic expression is derived for calculating the intensities of individual spin–rovibronic lines in the fully resolved gas phase electronic spectrum of a polyatomic molecule, in which one of the zero-order electronic states is a triplet state. The expression is employed to calculate the effect of fine structure splitting on the singlet–triplet absorption spectrum of pyrazine using the parameters available from experiment. A transition from Hund's coupling Case (a) to Case (b) on going from low J to high J rotational levels is predicted to occur at a moderate resolution of a few hundred MHz. The effect is more pronounced in pyrazine-d_4 and the pyrazine–argon van der Waals complex owing to their larger mass. © *1998 American Institute of Physics.* [S1063-7761(98)00507-1]

1. INTRODUCTION

The pyrazine molecule continues to attract much attention due to its unique photophysical properties. Owing to a moderate $S_1 - T_1$ energy gap (4056 cm^{-1}), the density of T_1 vibrational levels at the S_1 origin is relatively low, which puts pyrazine in the class of so-called intermediate-case polyatomic molecule[1] with respect to $S_1 - T_1$ intersystem crossing (ISC). In intermediate-case molecules, the coupling between the S_1 and T_1 states results in the formation of mixed singlet–triplet levels. These mixed levels have been revealed in ultrahigh-resolution spectra and are believed to be responsible for observed nonexponential fluorescence decay behavior.[2–7] While the singlet states that participate in this mixing are well characterized, little is known about the zero-order triplet states. The important point, in particular, is that the effect of the fine structure splitting on the triplet state energy level structure of the gas-phase molecule is not well understood, even at the T_1 origin.

In solid matrices at liquid helium temperatures, the energy of the lowest triplet state is split into three spin sublevels T_x, T_y and T_z with very different phosphorescence lifetimes τ_x, τ_y, and τ_z. The separations and lifetimes of these sublevels in pyrazine have been measured by the microwave-induced delayed phosphorescence (MIDP) technique.[8–12] Figure 1 shows the sublevel ordering and microwave transitions observed[9] in a benzene crystal at 1.2 K together with the appropriate choice of molecular axes. The lifetimes measured in various media are listed in Table I.

In the MIDP experiments, the T_1 state is prepared in an indirect way via $S_1 \leftarrow S_0$ excitation followed by $S_1 \rightarrow T_1$ ISC. The direct $T_1 \leftarrow S_0$ electronic transition in pyrazine has also been observed in the absorption spectra of solids[13–15] and vapors.[16–18] The measured oscillator strengths were determined to be 10^{-7} and 3×10^{-8}, respectively,[18] which is typical of a spin-forbidden transition. In supersonic jets, the $T_1 \leftarrow S_0$ transition was first detected using the multiphoton ionization (MPI) method[19] and then by phosphorescence excitation methods,[20–27] in some cases using SEELEM (surface electron ejection by laser excited metastables) detection. Valuable information about the vibrationally excited T_1 state has also been obtained with the pulse-field-ionization technique.[28]

As can be seen from Fig. 1 and Table I, the T_y and T_z sublevels in the solid host are nearly degenerate and have similar lifetimes, whereas the T_x sublevel is set apart by about 10 GHz and has a much shorter lifetime. While this phenomenon is well established in the condensed phase, no conclusive evidence for its manifestation in the gas phase molecule is available. Theoretically, it was explained in terms of the properties of spin–orbit coupling in aromatics and heteroaromatics, with their characteristic π-electronic structure and planar geometry.[29–34] Since the planarity of the two participating states has been shown to survive in the gas phase,[23] this phenomenon can be expected to occur in the gas phase, too.[25,26] However, previous studies were performed at relatively low resolution (~ 2 GHz),[25,26] not high enough to address the question definitively.

Two additional factors complicate the interpretation of a gas phase experiment. The first is the lack of a condensed phase environment. If τ_x, τ_y, and τ_z are assumed to be the same in the solid and gas phases, then a mean lifetime τ_{ph} arising from rotational mixing of the spin sublevels would be observed in jets, which at least would not be shorter than the shortest lifetime in solids, i.e., 6.3 ms.[9,11] The measured gas-phase phosphorescence lifetime at the T_1 origin in a supersonic expansion is only 1.45 ms, with an upper limit of 2.5 ms[27] having been established in a separate experiment.
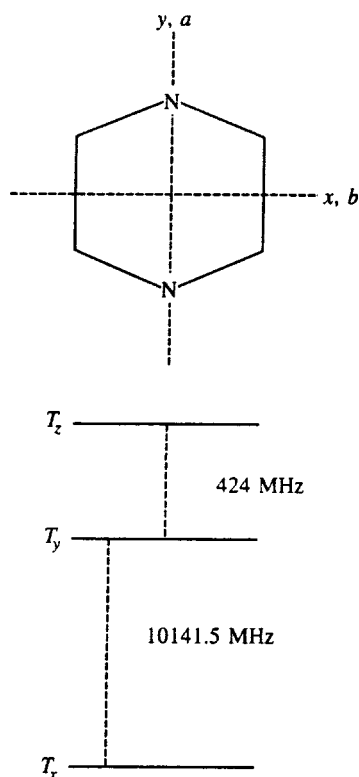
FIG. 1. Ordering and energy level separations of the lowest triplet state of pyrazine in the solid state, as obtained from MIDP experiments. The smallest moment of inertia in the gas phase is about the N–N line (the a axis) in both the S_0 and T_1 states. The largest moment of inertia is about the axis perpendicular to the molecular plane, the z (c) axis of the top.

Clearly, then, the environment plays a role. With the quantum yield of phosphorescence emission as large as 0.3 in a solid solution at 77 K,[35] this implies that acceleration of radiative and nonradiative transitions to the ground state might both be responsible for the lifetime shortening in the jet. If, further, we accept the absolute values of the $T_1 \leftarrow S_0$ oscillator strengths cited above, which tell us that the oscillator strength is a factor of three lower in the gas phase than in a solid, then it is the nonradiative transition rate that is responsible for the lifetime shortening.[27] The reasons for this behavior are completely unknown.

The second factor complicating the interpretation of a gas phase spectrum are rotations, which produce extensive mixing of the three spin sublevels of the triplet state. In the ground state, as well as in the lowest excited singlet and triplet states, the pyrazine molecule is a nearly symmetric top with rotational constants of about 3 and 6 GHz[26] (see Table II). If in the gas phase the spin splittings are assumed to be of the same order of magnitude, 10 GHz, as in solids (which

TABLE I. Triplet spin sublevel lifetimes of the lowest triplet state of pyrazine.

	τ_x	τ_y	τ_z	Ref.
Benzene host crystal at 1.2 K	6.3	284	163	9
Para-dioxane host crystal at 1.6 K	6.5	400	200	11
Supersonic jet	$\tau_{ph} = 1.45$ (upper limit 2.5)			27

Note. All values in msec.

TABLE II. Rotational and fine structure splitting constants of pyrazine in cm^{-1} (values in parentheses are for pyrazine-d_4).

	S_0	T_1	Ref.
A	0.21285	0.212	26 and 44
	(0.18389)		44
B	0.19767	0.196	26 and 44
	(0.17502)		44
C	0.10249	0.101	26 and 44
	(0.08967)		44
D		0.3455	9 and 44
E		−0.00705	8 and 43

seems to be the case for glyoxal, where they change from 0.6 and 2.2 GHz to 1.1 and 2.4 GHz between the solid and gas phases, respectively[22]), then the nature of the spectrum will be quite different at low and high rotational quantum numbers.

For high values of the rotational angular momentum N, Hund's Case (b) should apply, in which case the spectrum should consist of rotational bands corresponding to transitions between states with definite values of the asymmetric top quantum numbers N, K_{+1} and K_{-1}, each of which, except for $N=0$, is split into states with different values of $J=N$, $N\pm 1$, where $\mathbf{J}=\mathbf{N}+\mathbf{S}$ is the total angular momentum and $\mathbf{S}$ is the spin angular momentum. This is the analog of the rotational structure in the Σ states of diatomic molecules, which typically belong to Case (b) because of a weak spin-axis interaction.[36,37] However, for low values of N, the observed bands will be shifted away from the purely rotational band positions, since they are now spin–rotational bands. Again an analogy comes from diatomics, whose Π, Δ, etc. states usually belong to Case (a), apart from some light diatomics. The transition from Hund's Case (a) to Case (b) with increasing rotational quantum number is called spin uncoupling, of which several examples are given by Herzberg.[36]

A general approach to calculating band intensities in the singlet–triplet spectra of polyatomic molecules was developed by Hougen,[38] Creutzberg and Hougen,[39] and di Lauro.[40] Energies and wavefunctions of spin–rotational levels of a triplet state were derived from an effective Hamiltonian given by Van Vleck[41] and Raynes.[42] The intensities of transitions in nearly symmetric tops were represented in the form of Hougen and di Lauro factors (rather than Hönl–London factors), tabulated for some limiting cases. The Hougen factors[38] are appropriate for a Hund's Case (b) molecule with no multiplet splitting in the triplet state. Creutzberg and Hougen further extended this approach to near-symmetric top molecules of symmetry C_{2v}, D_2, and D_{2h}, defining a new "limiting" Case (ab), which was further subdivided into types I, II, and III. Their Case (ab) corresponds to a situation, not unlike that of pyrazine in its lowest triplet state, in which two of the spin components of the nonrotating molecule are separated by a small energy, and the third is separated by a large energy, compared to the rotational interval. (In this respect, T_1 pyrazine is a Case (ab), type II molecule; thus we are actually dealing with a Case (ab)–Case (b) transition here.) The di Lauro factors are appropriate for all types

of Hund's Case (ab) molecules; they account for the multiplet splittings by using Raynes' effective Hamiltonian[42] for asymmetric tops, which includes contributions from various magnetic interactions.

　　Here we take a new approach to this problem, one that relies on the derivation of a closed-form analytic expression for the intensities, assuming arbitrary relations between rotational intervals, asymmetry, and fine structure splittings. This approach is motivated by the likely success of future high-resolution experiments on the singlet–triplet transitions of a wide variety of molecules, requiring a more general method for their interpretation. In the present instance, we make two simplifying assumptions in order to make the problem tractable. Only the spin–spin interaction is taken into account,[43,44] and only one spin sublevel of the nonrotating molecule is assumed to be radiatively active.[31] A more general expression for the intensities is also obtained and compared with the Hougen factors.[38] We test our approach by computing the $T_1 \leftarrow S_0$ spectra of pyrazine, pyrazine-d_4, and the pyrazine–Ar van der Waals complex, for comparison with existing and future experimental spectra. The approach described here can be extended to other molecules, thereby anticipating new experiments in high-resolution spectroscopy.

2. THEORY

2.1. Rotational states in S_0

　　Pyrazine has D_{2h} symmetry in its ground (S_0, $^1A_{1g}$) and first excited triplet (T_1, $^3B_{1u}$) states. The smallest moment of inertia is about the axis passing through the two nitrogen atoms.[23,44] Therefore, the rotational Hamiltonian for both S_0 and T_1 is of the form

$$H_r = BN_x^2 + AN_y^2 + CN_z^2, \tag{1}$$

where $A > B > C$ are the rotational constants (nearly the same in the two states; see Table II) and $N_{x,y,z}$ are the projections of the rotational angular momentum vector on the axes of the molecular coordinate frame (MCF), the inertial axes. For a symmetric top ($A = B$) without spin, the operators $\mathbf{N}^2$, N_z, and $N_{\tilde{z}}$, i.e., the rotational angular momentum squared and its projections on the z axis of the MCF and on the $\tilde{z}$ axis of the laboratory coordinate frame (LCF), respectively, constitute a complete set of commuting operators.

　　The electronic–vibrational–rotational (EVR) wavefunctions can be written as

$$|\Gamma_S; NK\tilde{K}\rangle \equiv |\Gamma_S\rangle|NK\tilde{K}\rangle, \tag{2}$$

where Γ_S ($=A_{1g}$ for the vibrationless S_0 state) is the symmetry species (IR, irreducible representation) of the electronic–vibrational wavefunction $|\Gamma_S\rangle$, N is the rotational angular momentum quantum number, and K and $\tilde{K}$ are the eigenvalues of N_z and $N_{\tilde{z}}$, respectively. As a basis set for an asymmetric top without spin, we use the symmetrized rotational wavefunctions that transform (under rotations by an angle π about the MCF axes) according to the IR's $\Gamma_r = A$, B_1, B_2, B_3 of the D_2 group,

$$|\Gamma_r \lambda \tau NK\tilde{K}\rangle = f_K 2^{-1/2}[|NK\tilde{K}\rangle + \tau|N, -K, \tilde{K}\rangle]$$

TABLE III. K and λ for irreducible representations of the D_2 group.

Γ_r	λ	K
A	$+1$	even
B_1	-1	even
B_2	-1	odd
B_3	$+1$	odd

$$\text{for } K \geqslant 0 \quad \text{and} \quad \tau = \pm 1, \tag{3}$$

where

$$f_K = 1 \quad \text{for } K > 0 \quad \text{and} \quad f_K = 2^{-1/2} \quad \text{for } K = 0. \tag{4}$$

In Eq. (3), two additional ''quantum numbers,'' τ and

$$\lambda = \tau(-1)^N = \pm 1, \tag{5}$$

are defined for the characterization of the symmetrized basis functions. The reason for using λ is that this is a ''good quantum number'' for the rotational, spin–spin, and spin–orbit Hamiltonians (see below). The relations between Γ_r and λ for different K are shown in Table III.

　　Since the asymmetric top Hamiltonian (1) has no matrix elements between the basis states (3) differing in either Γ_r, λ, or τ, its eigenfunctions expanded in this basis can also be characterized by Γ_r, λ, and τ (in addition to N and $\tilde{K}$),

$$|\Gamma_S \Gamma_r \lambda \tau Ni\tilde{K}\rangle = \sum_K C_K^{(\Gamma_S \Gamma_r \lambda \tau Ni)}|\Gamma_r \lambda \tau NK\tilde{K}\rangle, \tag{6}$$

where i labels the eigenvalues of H_r for a given Γ_r, λ, τ, and N. Here K runs from 0 to N over all odd or even values depending on Γ_r (see Table III). The asymmetric top functions are labeled with Γ_S because, unlike the case of the symmetric top, they are dependent on the rotational constants of the given electronic state. The eigenvalues and the coefficients $C_K^{(\Gamma_S \Gamma_r \lambda \tau Ni)}$ are found by diagonalizing H_r in the basis (3). The Hamiltonian (1) conserves all quantum numbers except for K. Using the notation

$$\langle \Gamma_r \lambda \tau NK\tilde{K}|H_r|\Gamma_r \lambda \tau NK'\tilde{K}\rangle \equiv \langle K|H_r|K'\rangle,$$

we obtain the following nonvanishing matrix elements:

$$\langle K|H_r|K\rangle = \frac{1}{2}(A+B)[N(N+1) - K^2] + CK^2 \tag{7.1}$$

for $K \neq 1$,

$$\langle 1|H_r|1\rangle = \frac{1}{2}(A+B)[N(N+1) - 1]$$

$$+ C + \frac{1}{4}\tau(B-A)N(N+1) \tag{7.2}$$

for $K = 1$,

$$\langle K|H_r|K+2\rangle = \langle K+2|H_r|K\rangle$$

$$= \frac{1}{4}(B-A)[(N-K)(N-K-1)$$

$$\times (N+K+1)(N+K+2)]^{1/2} \tag{7.3}$$

for $K \neq 0$, and

$$\langle 0|H_r|2\rangle = \langle 2|H_r|0\rangle = 2^{-3/2}(B-A)$$

$$\times [N(N-1)(N+1)(N+2)]^{1/2} \qquad (7.4)$$

for $K=0$ and $\tau=1$.

They obey the selection rules

$$\Delta\lambda \equiv \lambda-\lambda'=0, \quad \Delta\tau = \Delta N=0, \quad \Delta K=0,\pm 2. \qquad (8)$$

2.2. Spin–rotational states in T_1

For a symmetric top with spin, three sets of commuting operators can be formed (see Appendix in Ref. 1 for commutation relations involving all relevant momenta and their projections on the MCF and LCF axes).

The first set is obtained by adding $\mathbf{S}^2$ and $S_{\tilde{z}}$ to the above rotational angular momenta. The EVR wavefunctions $|\Gamma_T;NK\tilde{K}\rangle$ ($\Gamma_T=B_{1u}$ for the T_1 state vibrationless wavefunction $|\Gamma_T\rangle$) are multiplied by the spin functions $|S\tilde{\sigma}\rangle$, where S is the total spin quantum number and $\tilde{\sigma}$ is an eigenvalue of $S_{\tilde{z}}$. We call this Representation I, the uncoupled-spin representation in the LCF.

The second set comprises the operators $\mathbf{S}^2$, $\mathbf{J}^2$, J_z, $J_{\tilde{z}}$, and S_z. The rotational part of their common eigenfunctions is $|JP\tilde{P}\rangle$ and the spin part is $|S\sigma\rangle$. Here, J is the total angular momentum quantum number and P, $\tilde{P}$, and σ are eigenvalues of J_z, $J_{\tilde{z}}$, and S_z, respectively. This is Representation II, the uncoupled-spin representation in the MCF.

The third set of commuting operators is $\mathbf{S}^2$, $\mathbf{J}^2$, $\mathbf{N}^2$, N_z, and $J_{\tilde{z}}$, and the corresponding mixed spin–rotational wavefunctions are $|SJNKP\rangle$. This is Representation III, the coupled-spin representation. This notation also will be used for the singlet-state rotational wavefunction, in which case one has

$$|SJNK\tilde{P}\rangle = \delta_{S0}\delta_{JN}\delta_{\tilde{P}\tilde{K}}|NK\tilde{K}\rangle. \qquad (9)$$

Then the symmetrized function (3) is denoted

$$|\Gamma_r\tau SJNK\tilde{P}\rangle = f_K 2^{-1/2}[|SJNK\tilde{P}\rangle + \tau|SJN,-K,\tilde{P}\rangle] \qquad (10)$$

for $K \geqslant 0$ and $\tau=\pm 1$, and the expansion (6) for the ground state takes the form

$$|\Gamma_S\Gamma_r''\lambda''\tau''S''J''N''i''\tilde{P}''\rangle$$

$$= \sum_{K''} C_{K''}^{(\Gamma_S\Gamma_r''\lambda''\tau''S''J''N''i'')}|\Gamma_r''\lambda''\tau''S''J''N''K''\tilde{P}''\rangle, \qquad (11)$$

where the ground state quantities are double-primed.

The spin–rotational functions can be expanded in terms of products of the pure rotational and spin functions using both uncoupled Representations I and II. According to the momentum-addition rule,[37] which for Representations I and II can be written $\mathbf{J}=\mathbf{N}+\mathbf{S}$ and $\mathbf{N}=(-\mathbf{S})+\mathbf{J}$, respectively, the expansions have the form[1,45]

$$|SJNK\tilde{P}\rangle = (-1)^{N-S+\tilde{P}}\sqrt{2J+1}$$

$$\times \sum_{\tilde{K}\tilde{\sigma}} \begin{pmatrix} N & S & J \\ \tilde{K} & \tilde{\sigma} & -\tilde{P} \end{pmatrix}|NK\tilde{K};S\tilde{\sigma}\rangle, \qquad (12)$$

$$|SJNK\tilde{P}\rangle = (-1)^{S-J+K}\sqrt{2N+1}$$

$$\times \sum_{P\sigma} \begin{pmatrix} S & J & N \\ \sigma & P & -K \end{pmatrix}|JP\tilde{P}\rangle\theta|S\sigma\rangle, \qquad (13)$$

where θ is the time-reversal operator,[37]

$$\theta|S\sigma\rangle = (-1)^{S-\sigma}|S,-\sigma\rangle. \qquad (14)$$

With the use of Eq. (14) the expansion in Eq. (13) can be recast as

$$|SJNK\tilde{P}\rangle = \sqrt{2N+1} \sum_{P\sigma} (-1)^{J+P}$$

$$\times \begin{pmatrix} J & S & N \\ -P & \sigma & K \end{pmatrix}|JP\tilde{P};S\sigma\rangle, \qquad (15)$$

where $(...)$ denotes a Wigner 3-J symbol.

We use the symmetrized spin–rotational wavefunctions in the form of Eq. (10) as the basis set for an asymmetric top with spin–spin coupling. Again, the rotational Hamiltonian (1) conserves all quantum numbers except K. Its nonvanishing matrix elements in the symmetrized basis (10) calculated with the use of expansion (12) are

$$\langle \Gamma_r\lambda\tau SJNK\tilde{P}|H_r|\Gamma_r\lambda\tau SJNK'\tilde{P}\rangle \equiv \langle K|H_r|K'\rangle,$$

where $\langle K|H_r|K'\rangle$ is given by (7). The selection rules are the same as in Eq. (8), plus $\Delta J=0$.

2.3. Fine structure of the spin–rotational states in T_1

The Hamiltonian of the spin–spin interaction in the triplet state has the form[43]

$$H_{SS} = D\left(S_x^2 - \frac{1}{3}\mathbf{S}^2\right) + E(S_z^2-S_y^2), \qquad (16)$$

where $S_{x,y,z}$ are projections of the spin angular momentum operator on the MCF axes, and D and E are the fine structure splitting constants given in Table II. In a nonrotating molecule H_{SS} has three eigenfunctions,

$$H_{SS}|T_{x,y,z}\rangle = T_{x,y,z}|T_{x,y,z}\rangle, \qquad (17)$$

where

$$T_z = -\frac{2}{3}D, \quad T_y = \frac{1}{3}D+E, \quad T_z = \frac{1}{3}D-E. \qquad (18)$$

The eigenfunctions obey the relationships

$$S_x|T_x\rangle = S_y|T_y\rangle = S_z|T_z\rangle = 0. \qquad (19)$$

These functions transform as the x, y, z components of a vector under the D_2 group rotations of the electronic spin variables according to $\Gamma_\sigma = B_3$, B_2, and B_1, respectively. Since these rotations change only the sign of S_x, S_y, and S_z, the Hamiltonian (16) is invariant, and hence has no off-diagonal matrix elements. Similarly, the products in Eq. (19) transform by $\Gamma_\sigma \times \Gamma_\sigma = A$, and hence identically vanish. Thus, Eqs. (18) and (19) are consequences of the symmetry properties of the system.

The above wavefunctions can also be expressed in terms of the spin functions $|S\sigma\rangle$ with a definite spin projection σ,[38]

$$|T_x\rangle = -i2^{-1/2}(|11\rangle - |1,-1\rangle),$$

$$|T_y\rangle = 2^{-1/2}(|11\rangle + |1,-1\rangle),$$

$$|T_z\rangle = -i|10\rangle. \tag{20}$$

Let us define Cartesian components of two traceless tensors,

$$T_{ij} = \frac{1}{2}(S_i S_j + S_j S_i) - \frac{1}{3}\mathbf{S}^2 \delta_{ij}, \quad Q_{ij} = Q_{ii}\delta_{ij}, \quad Q_{xx} = \frac{2}{3}D,$$

$$Q_{yy} = -\frac{1}{3}D - E, \quad Q_{zz} = -\frac{1}{3}D + E, \tag{21}$$

and the corresponding spherical tensors $T_\nu^{(k)}$ and $Q_\nu^{(k)}$ of the second rank $k=2$[37] by

$$T_0^{(2)} = -\sqrt{\frac{3}{2}}T_{zz}, \quad T_{\pm 1}^{(2)} = \pm(T_{xz} \pm iT_{yz}),$$

$$T_{\pm 2}^{(2)} = -\frac{1}{2}(T_{xx} - T_{yy} \pm 2iT_{xy}),$$

$$Q_0^{(2)} = 6^{-1/2}(D - 3E), \quad Q_{\pm 1}^{(2)} = 0,$$

$$Q_{\pm 2}^{(2)} = -\frac{1}{2}(D + E), \tag{22}$$

with all projections in Eqs. (21) and (22) being taken on the MCF axes. Then Eq. (16) takes the form

$$H_{SS} = \sum_{\nu=-k}^{k} (-1)^{k-\nu} Q_\nu^{(k)} T_{-\nu}^{(k)}, \tag{23}$$

which enables us to invoke the Wigner–Eckart theorem to calculate the relevant matrix elements for a rotating molecule.

Consider first the matrix elements of H_{SS} in the nonsymmetrized basis $|SJNK\widetilde{P}\rangle$. Since H_{SS} commutes with $\mathbf{S}^2$, $\mathbf{J}^2$, and $J_{\tilde{z}}$, and does not commute with $\mathbf{N}^2$ and N_z, it is diagonal in S, J, and $\widetilde{P}$, but not in N and K. Introducing the notation

$$\langle SJNK\widetilde{P}|H_{SS}|S'J'N'K'\widetilde{P}'\rangle$$

$$\equiv \delta_{SS'}\delta_{JJ'}\delta_{\widetilde{P}\widetilde{P}'}\langle NK|H_{SS}|N'K'\rangle$$

and inserting expansion (15), we obtain

$$\langle NK|H_{SS}|N'K'\rangle \equiv \langle N'K'|H_{SS}|NK\rangle = \sqrt{(2N+1)(2N'+1)}$$

$$\times \sum_{\nu P \sigma \sigma'} (-1)^{2-\nu} Q_\nu^{(2)} \begin{pmatrix} J & S & N \\ -P & \sigma & K \end{pmatrix}$$

$$\times \begin{pmatrix} J & S & N' \\ -P & \sigma' & K' \end{pmatrix} \langle S\sigma|T_{-\nu}^{(2)}|S\sigma'\rangle. \tag{24}$$

The advantage of using (15) rather than (12) is that both $Q_\nu^{(2)}$ and $\langle S\sigma|T_{-\nu}^{(2)}|S\sigma'\rangle$ are independent of the Euler angles, so that the rotational factor is merely $\delta_{JJ'}\delta_{PP'}\delta_{\widetilde{P}\widetilde{P}'}$. The remaining pure spin matrix element in Eq. (24) is calculated by applying the Wigner–Eckart theorem,[37]

$$\langle S\sigma|T_{-\nu}^{(2)}|S\sigma'\rangle = (-1)^{S-\sigma-1}$$

$$\times \begin{pmatrix} S & 2 & S \\ -\sigma & -\nu & \sigma' \end{pmatrix} \langle S\|T^{(2)}\|S\rangle, \tag{25}$$

where the reduced matrix element is

$$\langle S\|T^{(2)}\|S\rangle = \frac{1}{2\sqrt{6}}[(2S-1)2S(2S+1)(2S+2)(2S+3)]^{1/2} \tag{26}$$

$(=\sqrt{5}$ for $S=1$). Substituting Eqs. (25) and (26) into Eq. (24) and summing, we obtain for $S=1$

$$\langle NK|H_{SS}|N'K'\rangle \equiv \langle N'K'|H_{SS}|NK\rangle$$

$$= (-1)^{J-K}\sqrt{5(2N+1)(2N'+1)}$$

$$\times \begin{pmatrix} 2 & N' & N \\ K'-K & -K' & K \end{pmatrix}$$

$$\times \begin{Bmatrix} N & N' & 2 \\ 1 & 1 & J \end{Bmatrix} Q_{K-K'}^{(2)}, \tag{27}$$

where $\{...\}$ stands for a Wigner $6-J$ symbol. The following selection rules stem from Eqs. (22) and (27):

$$\Delta S = \Delta J = 0, \quad \Delta N = 0, \pm 1, \pm 2, \quad \Delta K = 0, \pm 2. \tag{28}$$

Matrix elements (27) also obey the relation

$$\langle N, -K|H_{SS}|N', -K'\rangle \equiv (-1)^{N+N'}\langle NK|H_{SS}|N'K'\rangle, \tag{29}$$

which is easy to deduce from the symmetry properties of the $3-J$ symbol.[37]

The matrix elements in the symmetrized basis (10) are off-diagonal in τ, N, and K. Introducing the notation

$$\langle \Gamma_r \lambda \tau SJNK\widetilde{P}|H_{SS}|\Gamma_r \lambda' \tau' S'J'N'K'\widetilde{P}'\rangle$$

$$\equiv \delta_{\lambda\lambda'}\delta_{SS'}\delta_{JJ'}\delta_{\widetilde{P}\widetilde{P}'}\langle \tau NK|H_{SS}|\tau'N'K'\rangle \tag{30}$$

and applying (29), we find

$$\langle \tau NK|H_{SS}|\tau'N'K'\rangle = f_K f_{K'}[\langle NK|H_{SS}|N'K'\rangle$$

$$+ \tau\langle N, -K|H_{SS}|N'K'\rangle].$$

It is easy to verify that the H_{SS} matrix is symmetric in this basis, as well as in the original basis (see Eq. (24)). As is evident from Eq. (30), the following selection rule, in addition to (28), applies to spin–spin coupling:

$$\Delta\lambda \equiv \lambda - \lambda' = 0. \tag{31}$$

Explicitly, the nonvanishing matrix elements of H_{SS} in terms of the matrix elements (27) are

$$\langle 1N0|H_{SS}|\tau'N'2\rangle = \sqrt{2}\langle N0|H_{SS}|N'2\rangle,$$

$$\langle \tau N1|H_{SS}|\tau'N'1\rangle = \langle N1|H_{SS}|N'1\rangle + \tau\langle N, -1|H_{SS}|N'1\rangle,$$

$$\langle \tau NK|H_{SS}|\tau'N'K'\rangle = \langle NK|H_{SS}|N'K'\rangle, \quad K+K' \neq 2. \tag{32}$$

The total Hamiltonian for the triplet state can be written

$$H_T = H_r + \varepsilon H_{SS}, \tag{33}$$

where ε is a parameter to be used in drawing a correlation diagram. Since H_T has no matrix elements between symmetrized basis states (10) differing in either Γ_r or λ, its eigenfunctions can also be characterized by Γ_r and λ (in addition to the conserved quantum numbers S, J, and $\widetilde{P}$). Putting a prime on all quantities relating to the triplet state, we obtain the following representation for the wavefunctions:

$$|\Gamma_T\Gamma_r'\lambda'S'J'i'\widetilde{P}'\rangle = \sum_{\tau'N'K'} C_{\tau'N'K'}^{(\Gamma_T\Gamma_r'\lambda'S'J'i')}$$

$$\times |\Gamma_r'\lambda'\tau'S'J'N'K'\widetilde{P}'\rangle, \qquad (34)$$

where i' labels the eigenvalues of H_T for a given Γ_r', λ', S', J'. For any given J', Γ_r' runs over the IR's of the D_2 group, and, for each Γ_r', the λ' assumes a definite value from Table III. The summation on the right-hand side of Eq. (34) is taken over $N' = J'$, $J' \pm 1$; K' runs from 0 to N' over even or odd values for a given Γ_r', according to Table III, and the sum over τ' involves only a single term, since τ' is uniquely fixed by λ' and N' according to the definition in Eq. (5).

2.4. Singlet contamination of the triplet state

The triplet state is assumed to be mixed, via spin–orbit coupling, with an excited singlet state S_n having $\Gamma_{S_n} = B_{2u}$ symmetry,[26,31] giving it radiative character. (We do not consider a mixing of the ground state with an excited triplet state T_n, which may be of importance in pyrazine.[46]) The intensities in the $T_1 \leftarrow S_0$ absorption spectra depend on the degree of this mixing. We first consider general constraints, due to the Pauli principle, on the symmetry species of levels that can mix with one another. These restrictions are independent of the form of the particular mixing operator.

Let ψ and ψ' stand for the total wavefunctions of the states that are coupled via some interaction Hamiltonian H_{int}. They are represented as products of the electronic, vibrational, rotational, and nuclear-spin wavefunctions,

$$\psi = \psi_e\psi_v\psi_r\psi_{ns}, \quad \psi' = \psi_e'\psi_v'\psi_r'\psi_{ns}',$$

whose symmetry properties depend on their behavior with respect to the feasible permutations[47] of the nuclei. Every such permutation P can be performed by a rotation $R \in D_2$. This is done in three steps. First, the MCF is rotated with respect to the LCF, which results in the transformation $\psi_r \rightarrow \Gamma_r\psi_r$, where Γ_r is an IR of D_2. Second, the electronic coordinates and nuclear displacements must be returned to their initial values by the inverse rotation R^{-1}; in D_2, it is the same as R. This leads to $\psi_e\psi_v \rightarrow (\Gamma_e\psi_e)(\Gamma_v\psi_v)$, where Γ_e and Γ_v are again IR's of D_2. Third, the permutation P is applied to the nuclear spin variables, resulting in $\psi_{ns} \rightarrow \Gamma_{ns}\psi_{ns}$. Since the set of three feasible permutations plus the identity (no rotation) operator constitute a group isomorphic to D_2, Γ_{ns} is an IR of D_2 as well.

The symmetry species of the total wavefunction with respect to the feasible permutations is therefore the product of the symmetry species of each of the components,

$$\Gamma = \Gamma_e\Gamma_v\Gamma_r\Gamma_{ns}, \quad \Gamma' = \Gamma_e'\Gamma_v'\Gamma_r'\Gamma_{ns}'. \qquad (35)$$

TABLE IV. Symmetry species Γ_r and $\Gamma_{r'}$ of the rotational sublevels of the vibrationless B_{1u} and B_{2u} states, respectively, that will mix via an arbitrary coupling independent of nuclear spin. The symmetry species of the corresponding nuclear spin wavefunctions, Γ_{ns}, are also given.

Γ_r	Γ_r'	Γ_{ns}
A	B_3	B_1
B_1	B_2	A
B_2	B_1	B_3
B_3	A	B_2

These permutations involve spinless carbon nuclei, spin-one nitrogen nuclei, and an even number (two or four) of protons whose total spin is again an integer. Then, according to the Pauli principle, the total wavefunction must be invariant under the feasible permutations of the nuclei, i.e.,

$$\Gamma = \Gamma' = A, \qquad (36)$$

where A is the totally symmetric representation of D_2.

Now, the interaction Hamiltonian H_{int} of interest is independent of nuclear spin. This means that $\psi_{ns} = \psi_{ns}'$ and $\Gamma_{ns} = \Gamma_{ns}'$. Combining this with Eqs. (35) and (36) yields the final relation

$$\Gamma_e\Gamma_v\Gamma_r = \Gamma_e'\Gamma_v'\Gamma_r' = \Gamma_{ns}. \qquad (37)$$

For the mixing of the vibrationless S_n and T_1 states one has $\Gamma_e = B_{2u}$, $\Gamma_e' = B_{1u}$, and $\Gamma_v = \Gamma_v' = A_g$. Then Eq. (37) reduces to $B_2\Gamma_r = B_1\Gamma_r' = \Gamma_{ns}$, where the D_2 group IR's are used instead of D_{2h}. Table IV shows the symmetry species of the rotational levels that are allowed to mix according to this result.

Next, we derive an expression for matrix elements of the spin–orbit Hamiltonian H_{so}, which couples T_1 to the singlet manifold S_n. In the Hartree–Fock self-consistent field approximation, H_{so} is represented by a sum of one-electron operators, each acting on the coordinates of a single electron,[1,34]

$$H_{\text{so}} = \sum_a \mathbf{b}_a \cdot \mathbf{s}_a, \qquad (38)$$

where $\mathbf{b}_a$ is a body-fixed vector whose components in the MCF depend on the Cartesian coordinates of the ath electron and on the nuclear displacements, and $\mathbf{s}_a$ is the spin operator of the ath electron. Equation (38) can be further simplified providing that the electronic states of interest are the ground state and only one-electron excited states. For instance, the $^1B_{2u}$ and $^3B_{1u}$ states of pyrazine are formed by promoting a single electron from the filled π_2 and n_+ orbitals, respectively, to the empty orbital π_4^*.[31,48] Then the summation in Eq. (38) can be taken over two electrons,

$$H_{\text{so}} = \mathbf{b}_1 \cdot \mathbf{s}_1 + \mathbf{b}_2 \cdot \mathbf{s}_2. \qquad (39)$$

The equivalence of Eqs. (38) and (39) can be proved by a straightforward calculation, e.g., using the method of second quantization.[37]

In order to calculate the $S_n - T_1$ vibronic matrix element of H_{so}, Eq. (39), we note that the spatial part of the elec-

tronic wavefunction is symmetric under electron permutations for the singlet state and antisymmetric for the triplet state. Then, the vector

$$\mathbf{c} \equiv \langle \Gamma_{S_n} | \mathbf{b}_1 | \Gamma_T \rangle = -\langle \Gamma_{S_n} | \mathbf{b}_2 | \Gamma_T \rangle \qquad (40)$$

can be defined for each participating singlet state, and the desired matrix element takes the form

$$\langle \Gamma_{S_n} | H_{so} | \Gamma_T \rangle = \mathbf{cG} = \sum_{\sigma=0,\pm 1} (-1)^{1-\sigma} c_\sigma^{(1)} G_{-\sigma}^{(1)}, \qquad (41)$$

where $\mathbf{G} = \mathbf{s}_1 - \mathbf{s}_2$. The components of the first-rank spherical tensors are defined as[37]

$$c_0^{(1)} = i c_z, \quad c_{\pm 1}^{(1)} = \mp i 2^{-1/2}(c_x \pm i c_y), \qquad (42)$$

and similarly for $G_\sigma^{(1)}$. The matrix elements of $G_\sigma^{(1)}$ over the spin functions $|S\sigma\rangle$ are easily calculated to give

$$\langle 00 | (-1)^{1-\sigma} G_{-\sigma}^{(1)} | 1\sigma' \rangle = -i \delta_{\sigma\sigma'}. \qquad (43)$$

In the particular case of pyrazine, the spatial parts of the electronic wave functions can be written in terms of one-electron orbitals as

$$|\Gamma_{S_n}\rangle = 2^{-1/2}[\pi_2(\mathbf{r}_1)\pi_4^*(\mathbf{r}_2) + \pi_2(\mathbf{r}_2)\pi_4^*(\mathbf{r}_1)],$$

$$|\Gamma_T\rangle = 2^{-1/2}[n_+(\mathbf{r}_1)\pi_4^*(\mathbf{r}_2) - n_+(\mathbf{r}_2)\pi_4^*(\mathbf{r}_1)],$$

where $\mathbf{r}_{1,2}$ are the coordinates of the two electrons. Then,

$$\mathbf{c} = \frac{1}{2}\langle \pi_2 | \mathbf{b} | n_+ \rangle.$$

Since π_2 and n_+ transform as B_3 and A, respectively, the vector $\mathbf{c}$ has only one nonvanishing component,

$$c_x = \frac{1}{2}\langle \pi_2 | b_x | n_+ \rangle \equiv v_{so}\sqrt{2}, \quad c_y = c_z = 0, \qquad (44)$$

where the spin–orbit coupling parameter is v_{so}.

Using Eq. (44), we recast Eq. (42) as

$$c_\sigma^{(1)} = -i\sigma v_{so}. \qquad (45)$$

From Eqs. (20), (43), and (45) we deduce that only the T_x triplet sublevel is contaminated with the S_n singlet via spin–orbit coupling (41). This result is a consequence of symmetry.[38] Indeed, H_{so} is invariant under the D_2 group transformations of the electronic space and spin coordinates and nuclear displacements. Hence, a vibrationless triplet level $|\Gamma_T\Gamma_\sigma\rangle$ can be mixed with a vibrationless singlet $|\Gamma_{S_n}\rangle$ only when $\Gamma_T \times \Gamma_\sigma = \Gamma_{S_n}$. For $\Gamma_T = B_1$ and $\Gamma_{S_n} = B_2$, it follows that only the T_x level with $\Gamma_\sigma = B_3$ mixes with the singlet.

Next, we calculate the spin–rotational matrix elements of Eq. (41),

$$\langle SJNK\widetilde{P} | H_{so} | S'J'N'K'\widetilde{P}' \rangle$$

$$\equiv \delta_{JN}\delta_{JJ'}\delta_{\widetilde{P}\widetilde{P}'} \langle K | H_{so} | J'N'K' \rangle, \qquad (46)$$

where $S = 0$ and $S' = 1$. Now H_{so} denotes $\langle \Gamma_{S_n} | H_{so} | \Gamma_T \rangle$, and standard notation is used for the singlet and triplet wavefunc-

tions (see Eq. (9)). Inserting Eq. (41) and the spin–rotational functions in the form of expansion (15), we obtain after simple algebra[1]

$$\langle K | H_{so} | J'N'K' \rangle = v_{so}(-1)^{J'-K+1}\sqrt{2N'+1}$$

$$\times \sum_{\sigma'=\pm 1} \sigma' \begin{pmatrix} J' & 1 & N' \\ -K & \sigma' & K' \end{pmatrix}, \qquad (47)$$

where use was made of Eqs. (43) and (45). These matrix elements have the obvious property

$$\langle -K | H_{so} | J'N', -K' \rangle \equiv (-1)^{N+N'} \langle K | H_{so} | J'N'K' \rangle, \qquad (48)$$

where $N = J'$ according to Eq. (46). The matrix elements of H_{so} in the symmetrized basis (10),

$$\langle \Gamma_r \lambda \tau SJNK\widetilde{P} | H_{so} | \Gamma_r' \lambda' \tau' S'J'N'K'\widetilde{P}' \rangle$$

$$\equiv \delta_{\lambda\lambda'} \delta_{JN} \delta_{JJ'} \delta_{\widetilde{P}\widetilde{P}'} \langle \tau K | H_{so} | \tau'J'N'K' \rangle, \qquad (49)$$

can be expressed in terms of (47):

$$\langle \tau K | H_{so} | \tau'J'N'K' \rangle = f_K f_{K'} [\langle K | H_{so} | J'N'K' \rangle$$

$$+ \tau' \langle K | H_{so} | J'N', -K' \rangle], \qquad (50)$$

where use was made of Eq. (48). The selection rules for the matrix elements (49) are

$$\Delta\lambda = 0, \quad \Delta S = \pm 1, \quad \Delta J = 0,$$

$$\Delta N = 0, \pm 1, \quad \Delta K = \pm 1. \qquad (51)$$

We see that H_{so} conserves λ and changes K by one. According to Table III, this means that H_{so} mixes the states of the following rotational symmetries Γ_r: $B_1 \leftrightarrow B_2$ and $B_3 \leftrightarrow A$. In other words, H_{so} satisfies the general requirements for a mixing operator as displayed in Table IV. Substituting Eq. (47) into (50), we obtain

$$\langle \tau K | H_{so} | \tau'J'N'K' \rangle = v_{so} f_K f_{K'}(-1)^{J'-K+1}\sqrt{2N'+1}$$

$$\times \sum_{\sigma'=\pm 1} \sigma' \left[\begin{pmatrix} J' & 1 & N' \\ -K & \sigma' & K' \end{pmatrix} \right.$$

$$\left. + \tau' \begin{pmatrix} J' & 1 & N' \\ -K & \sigma' & -K' \end{pmatrix} \right]. \qquad (52)$$

Penner *et al.*[25] introduced the conjecture of a so-called $M_{J'}$ selectivity of the spin–orbit coupling, assuming that, out of a manifold of triplet states with given J', a single state with $M_{J'} = 0$ borrows oscillator strength from an excited singlet. This is shown to be incorrect by our Eq. (52). It is easy to verify that for every $J' \geq 1$, all $3(2J' + 1)$ triplet sublevels borrow oscillator strength from S_n.

The next step is to calculate the singlet-contaminated triplet wave function,

$$|\Gamma_T; \Gamma_r' \lambda' \tau' S'J'N'K'\widetilde{P}' \rangle$$

$$\rightarrow \sum \frac{\langle \Gamma_r \lambda \tau SJNK\widetilde{P} | H_{so} | \Gamma_r' \lambda' \tau' S'J'N'K'\widetilde{P}' \rangle}{E_T - E_{S_n}}$$

$$\times |\Gamma_{S_n}; \Gamma_r \lambda \tau SJNK\widetilde{P} \rangle, \qquad (53)$$

where the unperturbed triplet function is omitted and the sum is taken over all quantum numbers of the S_n state. The energy denominator depends on the quantum numbers of the participating states, and in general Eq. (53) depends on the rotational constants of the contaminating singlet state. However, in pyrazine the S_n state is assumed to lie at a much higher energy than the T_1 state, in which case the purely electronic energy gap ΔE between S_n and T_1 can be substituted for the energy denominator in Eq. (53). We also note that, strictly speaking, the linear combinations (34) found after diagonalizing the total triplet-state Hamiltonian (33) rather than the basis functions (10) should be corrected for the contamination with S_n. However, after making the above approximation to the energy denominator, we obtain the same result by calculating (53) and inserting it into (34). Finally, the symmetric top wavefunctions for the S_n state are used in Eq. (53), thus neglecting the S_n asymmetry effect. As a result, the mixing of the triplet state is independent of the rotational constants of the contaminating singlet state.

Inserting Eqs. (49) and (52) into (53) yields

$$|\Gamma_T;\Gamma'_r\lambda'\tau'S'J'N'K'\widetilde{P}'\rangle$$

$$\rightarrow \frac{1}{\Delta E}\sum \delta_{\lambda\lambda'}\delta_{JN}\delta_{JJ'}\delta_{\widetilde{P}\widetilde{P}'}$$

$$\times\langle\tau K|H_{\mathrm{so}}|\tau'J'N'K'\rangle|\Gamma_{S_n};\Gamma_r\lambda\tau SJNK\widetilde{P}\rangle$$

$$= \frac{v_{\mathrm{so}}}{\Delta E}\sum_{P'=0}^{J'} f_{P'}f_{K'}(-1)^{J'-P'+1}\sqrt{2N'+1}$$

$$\times\sum_{\sigma'=\pm 1}\sigma'\left[\begin{pmatrix} J' & 1 & N' \\ -P' & \sigma' & K' \end{pmatrix}\right.$$

$$\left.+\tau'\begin{pmatrix} J' & 1 & N' \\ -P' & \sigma' & -K' \end{pmatrix}\right]|\Gamma_{S_n};\Gamma_r\lambda'\tau SJ'J'P'\widetilde{P}'\rangle,$$

(54)

where we changed the notation K to P' following the convention that K is a projection of **N** and P is a projection of **J**. In the singlet function, τ is fixed by the condition that $\lambda=\lambda'$, which yields (see the definition of λ in Eq. (5))

$$\tau=\tau'(-1)^{J'+N'}.$$

(55)

Equation (54) can be derived in a different way. First, we calculate the contaminated triplet spin wavefunction neglecting rotations and using Eqs. (41), (43), and (45),

$$|\Gamma_T;S'\sigma'\rangle\rightarrow\frac{1}{\Delta E}\langle\Gamma_{S_n};00|H_{\mathrm{so}}|\Gamma_T;S'\sigma'\rangle|\Gamma_{S_n};00\rangle$$

$$= -\frac{v_{\mathrm{so}}}{\Delta E}\sigma'|\Gamma_{S_n};00\rangle.$$

(56)

Then, we insert this into expansions (10) and (15) for the symmetrized spin–rotational function,

$$|\Gamma_T;\Gamma'_r\lambda'\tau'S'J'N'K'\widetilde{P}'\rangle$$

$$\rightarrow -\frac{v_{\mathrm{so}}}{\Delta E}f_{K'}2^{-1/2}\sqrt{2N'+1}\sum_{P'=-J'}^{J'}(-1)^{J'+P'}$$

$$\times\sum_{\sigma'=\pm 1}\sigma'\left[\begin{pmatrix} J' & 1 & N' \\ -P' & \sigma' & K' \end{pmatrix}\right.$$

$$\left.+\tau'\begin{pmatrix} J' & 1 & N' \\ -P' & \sigma' & -K' \end{pmatrix}\right]|\Gamma_{S_n};J'P'\widetilde{P}';00\rangle.$$

(57)

Here, the singlet wavefunction must also be symmetrized. We rename the summation variables, $P'\rightarrow K$ and $\sigma'\rightarrow\sigma$, and recast the sum over K as

$$\sum_{K=0}^{J'} f_K^2(-1)^{J'+K}\sum_{\sigma=\pm 1}\sigma\left[\begin{pmatrix} J' & 1 & N' \\ -K & \sigma & K' \end{pmatrix}\right.$$

$$\left.+\tau'\begin{pmatrix} J' & 1 & N' \\ -K & \sigma & -K' \end{pmatrix}\right]|\Gamma_{S_n};J'K\widetilde{P}';00\rangle$$

$$+\sum_{K=0}^{J'} f_K^2(-1)^{J'-K}\sum_{\sigma=\pm 1}\sigma\left[\begin{pmatrix} J' & 1 & N' \\ K & \sigma & K' \end{pmatrix}\right.$$

$$\left.+\tau'\begin{pmatrix} J' & 1 & N' \\ K & \sigma & -K' \end{pmatrix}\right]|\Gamma_{S_n};J',-K,\widetilde{P}';00\rangle.$$

(58)

In the second term of Eq. (58) we change the sign of the summation variable, $\sigma\rightarrow-\sigma$. Using the symmetry properties of the 3-J symbols[37] and the identity $(\tau')^2=1$, we rewrite Eq. (58) in the form

$$\sum_{K=0}^{J'} f_K^2(-1)^{J'+K}\sum_{\sigma=\pm 1}\sigma\left[\begin{pmatrix} J' & 1 & N' \\ -K & \sigma & K' \end{pmatrix}\right.$$

$$\left.+\tau'\begin{pmatrix} J' & 1 & N' \\ -K & \sigma & -K' \end{pmatrix}\right][|J'K\widetilde{P}'\rangle$$

$$+\tau|J',-K,\widetilde{P}'\rangle]|\Gamma_{S_n};00\rangle,$$

(59)

where τ is given by Eq. (55). By definitions (3)–(5), (9), and (10), the total wavefunction in Eq. (59) can be written

$$2^{1/2}f_K^{-1}|\Gamma_{S_n};\Gamma_r\lambda\tau SJ'J'K\widetilde{P}';00\rangle$$

(60)

with $\lambda\equiv\tau(-1)^{J'}=\tau'(-1)^{N'}=\lambda'$. Then inserting Eqs. (58)–(60) into (57) again leads to Eq. (54).

2.5. Intensities of individual singlet–triplet transitions

Contamination of the triplet state with the singlet state S_n results in a nonzero transition moment $\langle\Gamma_S|\boldsymbol{\mu}|\Gamma_T\Gamma_\sigma\rangle$, which is proportional to $\langle\Gamma_S|\boldsymbol{\mu}|\Gamma_{S_n}\rangle$, where $\boldsymbol{\mu}$ is the electric dipole moment operator. The only nonvanishing component of the transition moment in the MCF for $\Gamma_S=A_{1g}$ and $\Gamma_{S_n}=B_{2u}$ is $\langle\Gamma_S|\mu_y|\Gamma_{S_n}\rangle$ since the symmetry species of μ_y is B_{2u}. The corresponding Hougen intensity parameter $\mu(B_{3g})$[38] can be introduced by the relationship

$$\mu(B_{3g})=\sqrt{2}\frac{v_{\mathrm{so}}}{\Delta E}\langle\Gamma_S|\mu_y|\Gamma_{S_n}\rangle\equiv\frac{c_x\mu_y}{\Delta E},$$

(61)

where μ_y now stands for the matrix element and the notation $\mu(B_{3g})$ reflects the fact that the triplet sublevel T_x borrowing the $B_{2u} \leftarrow A_{1g}$ oscillator strength belongs to the symmetry species $\Gamma_\sigma = B_{3g}$.

The intensity of an individual transition is given by

$$I(\Gamma_T \Gamma'_r \lambda S' J' i' \leftarrow \Gamma_S \Gamma''_r \lambda'' \tau'' S'' J'' N'' i'')$$

$$= 3 \sum_{\tilde{P}' \tilde{P}''} |\langle \Gamma_T; \Gamma'_r \lambda' S' J' i' \tilde{P}' | \boldsymbol{\mu} \cdot \mathbf{e} | \Gamma_S; \Gamma''_r \lambda'' \tau'' S'' J'' N'' i'' \tilde{P}'' \rangle|^2, \tag{62}$$

where $\mathbf{e}$ is a unit vector in the direction of the electric field. Defining the first-rank spherical tensor components of μ and $\mathbf{e}$ both in the LCF and MCF by the relations given in Eq. (42), e.g.,

$$e_0^{(1)} = i e_{\tilde{z}}, \quad e_{\pm\tilde{1}}^{(1)} = \mp i 2^{-1/2}(e_{\tilde{x}} \pm i e_{\tilde{y}}), \tag{63}$$

and similarly for $\boldsymbol{\mu}$ where a tilde indicates the LCF axes, we have

$$\boldsymbol{\mu} \cdot \mathbf{e} = \sum_{\tilde{\sigma} = 0, \pm 1} (-1)^{1-\tilde{\sigma}} \mu_{\tilde{\sigma}}^{(1)} e_{-\tilde{\sigma}}^{(1)}, \tag{64}$$

which is similar to Eq. (41), except that projections are taken in the LCF rather than MCF. The LCF projections can be expressed in terms of the MCF projections[37]:

$$\mu_{\tilde{\sigma}}^{(1)} = \sum_{\sigma = 0, \pm 1} D_{\sigma\tilde{\sigma}}^{(1)} \mu_\sigma^{(1)}. \tag{65}$$

In Eqs. (64) and (65), $\mu_\sigma^{(1)}$ and $e_{-\tilde{\sigma}}^{(1)}$ are invariant with respect to rotations, whereas the Wigner functions $D_{\sigma\tilde{\sigma}}^{(1)}$ depend on the Euler angles. We substitute Eqs. (64), (65) and expansions (11) and (34) into Eq. (62), obtaining

$$I(\Gamma_T \Gamma'_r \lambda' S' J' i' \leftarrow \Gamma_S \Gamma''_r \lambda'' \tau'' S'' J'' N'' i'')$$

$$= 3 \sum_{\tilde{P}', \tilde{P}''} \left| \sum_{N'K'K''} C_{N'K'}^{(i')} C_{K''}^{(i'')} M_1 \right|^2, \tag{66}$$

where we omitted some indices of the coefficients of expansions (11) and (34), and

$$M_1 = \langle \Gamma_T; \Gamma'_r \lambda' \tau' S' J' N' K' \tilde{P}' | \sum_{\sigma\tilde{\sigma}} (-1)^{1-\tilde{\sigma}} D_{\sigma\tilde{\sigma}}^{(1)} \mu_\sigma^{(1)} e_{-\tilde{\sigma}}^{(1)} | \Gamma_S; \Gamma''_r \lambda'' \tau'' S'' J'' N'' K'' \tilde{P}'' \rangle. \tag{67}$$

The triplet function is approximated by Eq. (54),

$$M_1 = \frac{v_{so}}{\Delta E} \sum_{P'} f_{P'} f_{K'} (-1)^{J'-P'+1} \sqrt{2N'+1}$$

$$\times \sum_{\sigma'=\pm 1} \sigma' \left[\begin{pmatrix} J' & 1 & N' \\ -P' & \sigma' & K' \end{pmatrix} \right.$$

$$\left. + \tau' \begin{pmatrix} J' & 1 & N' \\ -P' & \sigma' & -K' \end{pmatrix} \right] M_2, \tag{68}$$

where

$$M_2 = \langle \Gamma_{S_n}; \Gamma_r \lambda' \tau S J' J' P' \tilde{P}' | \sum_{\sigma\tilde{\sigma}} (-1)^{1-\tilde{\sigma}} D_{\sigma\tilde{\sigma}}^{(1)} \mu_\sigma^{(1)} e_{-\tilde{\sigma}}^{(1)} | \Gamma_S; \Gamma''_r \lambda'' \tau'' S'' J'' N'' K'' \tilde{P}'' \rangle \tag{69}$$

and τ is given by Eq. (55). The electronic factor in Eq. (69) is

$$\langle \Gamma_{S_n} | \mu_\sigma^{(1)} | \Gamma_S \rangle = |\sigma| \mu_y / \sqrt{2}. \tag{70}$$

We will choose the LCF axis $\tilde{z}$ along the electric field, since the result is independent of the field direction. Then from Eq. (63) we have $e_0^{(1)} = i$, $e_{\pm\tilde{z}}^{(1)} = 0$, and Eq. (69) becomes $M_2 = -i\mu_y M_3 / \sqrt{2}$, where

$$M_3 = \langle \Gamma_r \lambda' \tau S J' J' P' \tilde{P}' | \sum_{\sigma=\pm 1} D_{\sigma 0}^{(1)} | \Gamma''_r \lambda'' \tau'' S'' J'' N'' K'' \tilde{P}'' \rangle \tag{71}$$

is the rotational factor. This matrix element can be expressed

in terms of the integral of a product of three Wigner functions.[37] With the use of Eqs. (9), (10), and (55) the rotational factor becomes

$$M_3 = f_{P'} f_{K''} \delta_{\lambda', -\lambda''} \left[\langle J' K \tilde{P} | \sum_{\sigma=\pm 1} D_{\sigma 0}^{(1)} | J'' K'' \tilde{P}'' \rangle \right.$$

$$\left. + \tau'' \langle J' K \tilde{P}' | \sum_{\sigma=\pm 1} D_{\sigma 0}^{(1)} | J'', -K'', \tilde{P}'' \rangle \right]$$

$$= f_{P'} f_{K''} \delta_{\lambda', -\lambda''} (-1)^{P'-\tilde{P}} \sqrt{(2J'+1)(2J''+1)}$$

$$\times \sum_{\sigma=\pm 1} \left[\begin{pmatrix} J' & 1 & J'' \\ -P' & \sigma & K'' \end{pmatrix} \right.$$

$$+\tau''\begin{pmatrix} J' & 1 & J'' \\ -P' & \sigma & -K'' \end{pmatrix}\Bigg]\begin{pmatrix} J' & 1 & J'' \\ -\tilde{P}' & 0 & \tilde{P}'' \end{pmatrix}.$$

After inserting Eqs. (67)–(72) into Eq. (66), the last factor in Eq. (72) squared can be summed to give unity. Thus, we arrive at the final expression for the intensity of an individual transition:

$$I(\Gamma_T \Gamma_r' \lambda' S' J' i' \leftarrow \Gamma_S \Gamma_r'' \lambda'' \tau'' S'' J'' N'' i'')$$

$$= \frac{1}{4} \, \delta_{\lambda',-\lambda''} |\mu(B_{3g})|^2 (2J'+1)(2J''+1)$$

$$\times \Bigg| \sum_{N'=|J'-1|}^{J'+1} \sqrt{2N'+1}$$

$$\times \sum_{\substack{K'=0 \\ \text{(only even or odd)}}}^{N'} f_{K'} C_{N'K'}^{(i')}$$

$$\times \sum_{\substack{K''=0 \\ \text{(only even or odd)}}}^{J''} f_{K''} C_{K''}^{(i'')} \sum_{P'=0}^{J'} f_{P'}^2$$

$$\times \sum_{\sigma'\sigma=\pm 1} \sigma' F_{P'\sigma'\sigma} \Bigg|^2, \tag{73}$$

where

$$F_{P'\sigma'\sigma} = \Bigg[\begin{pmatrix} J' & 1 & N' \\ -P' & \sigma' & K' \end{pmatrix} + \tau'\begin{pmatrix} J' & 1 & N' \\ -P' & \sigma' & -K' \end{pmatrix}\Bigg]$$

$$\times \Bigg[\begin{pmatrix} J' & 1 & J'' \\ -P' & \sigma & K'' \end{pmatrix} + \tau''\begin{pmatrix} J' & 1 & J'' \\ -P' & \sigma & -K'' \end{pmatrix}\Bigg], \tag{74}$$

and, according to Eq. (5),

$$\tau' = \lambda'(-1)^{N'}, \quad \tau'' = \lambda''(-1)^{J''}. \tag{75}$$

For given Γ_r' and Γ_r'' the evenness or oddness of K' and K'', as well as the values of λ' and λ'', are found from Table III. The following selection rules are evident from the properties of the Wigner 3-J symbols in Eq. (74):

$$\Delta J = 0, \pm 1, \quad \Delta N = 0, \pm 1, \pm 2, \quad \Delta K = 0, \pm 2. \tag{76}$$

For instance, P' differs from both K' and K'' by one, therefore K' and K'' differ by 0 or 2. Since the K's are both even or odd, and since the λ's are opposite in sign, the allowed transitions (see Table III) are $A \leftrightarrow B_1$ and $B_2 \leftrightarrow B_3$.

Equation (73) is the main result of the present paper. It can be applied to molecules for which only the spin–spin interaction is important, and in which only one spin sublevel of the nonrotating molecule is radiatively active, as in the $T_1 \leftarrow S_0$ absorption spectrum of pyrazine. A more general formula, valid for other molecules, is derived in the Appendix.

In Eq. (73), the intensities are given for transitions between the true molecular eigenstates, represented by the symmetrized rotational wavefunctions in Eqs. (3) and (10),
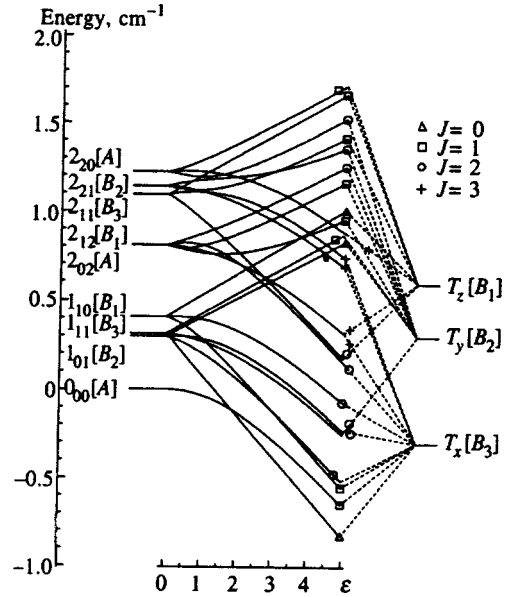
FIG. 2. Correlation diagram for the lowest spin–rotational levels of the T_1 state of pyrazine in the gas phase.

rather than between the states represented by the functions in Eqs. (2) and (15). Consequently, the Boltzmann averaging necessitated by a given experimental arrangement does not involve any additional complications associated with using the signed angular momentum projections.[38] Instead, each intensity calculated from Eq. (73) need only be multiplied by $g \exp(-E/kT)$, where g is the nuclear spin statistical weight.

The singlet–triplet spectrum of a polyatomic molecule is a very complicated matter, even for a symmetric top without multiplet splitting. Therefore, our focus in this paper is to develop a general approach accompanied by a detailed derivation for a specific molecule, which can easily be extended to any other one. A more general formula is given in the Appendix.

3. RESULTS AND DISCUSSION

3.1. Correlation diagram

A correlation diagram for the spin–rotational levels of the Hamiltonian (33) is shown in Fig. 2. For $\varepsilon = 0$ (no fine structure (FS) splitting) the levels of an asymmetric top are shown as horizontal bars on the left-hand side of the diagram. The level labeling is $N_{K_{-1}K_{+1}}$, where the asymmetric-top quantum number K_{+1} reduces to the projection of the angular momentum $\mathbf{N}$ on the top axis c, K_c, in the limit of an oblate symmetric top $(A = B)$. The asymmetry splitting for a nearly oblate symmetric top is first-order in the small asymmetry parameter $A - B$ for $K_{+1} = 1$, and second-order for $K_{+1} > 1$. Therefore, the splitting of the 2_{12} and 2_{02} levels is not seen at the energy scale of the figure. The symmetry species of the levels are shown in brackets.

Turning on the FS interaction $(\varepsilon > 0)$ produces a splitting of all levels other than $N_{K_{-1}K_{+1}} = 0_{00}$, as shown in Fig. 2. Typically, three components are observed, corresponding to three possible values of $\mathbf{J} = \mathbf{N} + \mathbf{S}$. In general, the separation of the three components decreases with increasing N
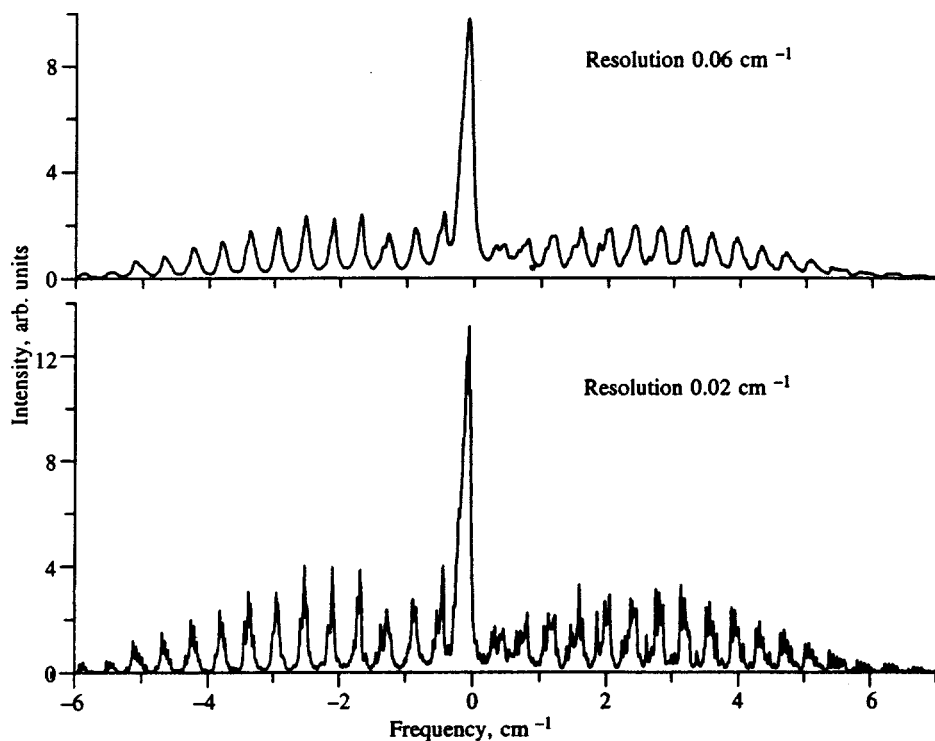
FIG. 3. The calculated singlet–triplet spectrum of pyrazine-h_4 at two different resolutions. The rotational temperature is $T_{rot} = 10$ K, $J_{max} = 12$.

(and J). The $J = 3$ components of the 2_{12} and 2_{02} levels remain nearly degenerate, since their coupling is to the 3_{12} and 3_{22} levels, which lie at significantly higher energy. Each set of FS components belongs to the same symmetry species as the levels from which they are derived at $\varepsilon = 0$.

The spin levels of the nonrotating molecule and their correlation with the spin–rotational levels are shown on the right-hand side of the figure. The energy separations of the $T_{x,y,z}$ are arbitrary. A full correlation diagram would be two-dimensional, depending upon two parameters of the FS interaction Hamiltonian, D and E (see Eq. (16)). Instead, we show a one-dimensional section of the diagram, in which D and E are both proportional to ε, while their ratio, E/D, remains constant (see Eq. (33)). In this case the correspondence between the symmetry species of the spin–rotational levels and the pure spin levels is not uniquely defined, being dependent upon the value of E/D. Indeed, each spin–rotational wavefunction involves contributions from all three spin levels $T_{x,y,z}$, as can be seen from the expansion (15) in the limit of $\varepsilon = 0$. When $\varepsilon > 0$, the same is true for the wavefunctions (34), where the relative contributions of $T_{x,y,z}$ depend on ε. In the limit of large ε, one of them dominates, but which one cannot be predicted without calculations, since this depends on the E/D ratio.

As an example, let us consider the $J = 1$ components of 2_{20}, 2_{02}, and 0_{00}, which correlate with T_z, T_y, and T_x, respectively. If we change the sign of the FS constant E, the ordering of the levels T_z and T_y is reversed (see Eq. (18)). However, because of the non-crossing rule for terms of the same symmetry (A in the present case), the previous correlation is not preserved. Instead, the $J = 1$ components of 2_{20}, 2_{02}, and 0_{00} now correlate with T_y, T_z, and T_x, respectively.

The number of levels with a given J in Fig. 2 is equal to

$6J + 3$, being defined by the usual momentum addition rule.

3.2. Singlet–triplet spectrum of pyrazine-h_4

We next used Eq. (73) to calculate the singlet–triplet absorption spectrum of the 0_0^0 band of pyrazine-h_4. The spectrum was calculated by Boltzmann averaging the intensities of the individual transitions using a rotational temperature of 7 cm^{-1} (10 K). The summation over the transitions was truncated at a maximum value of the total angular momentum $J_{max} = 12$, at which point the level populations were less than one percent of the total population. Nuclear spin statistical weights, $g(B_1) = g(B_3) = 9$, $g(B_2) = 13$, and $g(A) = 17$, were also taken into account. Each line was dressed with a Lorentzian whose full width at half maximum represents the resolution of the calculated spectra.

The spectra calculated with the parameters of Table II (and $\varepsilon = 1$ in Eq. (33)) are shown in Fig. 3. The top spectrum has the same resolution (~ 2 GHz) as the published experimental spectra.[25,26] The spectrum consists of a strong central Q branch due to the $\Delta N = 0$ ($\Delta N \equiv N' - N''$) transitions, an R-form branch involving $R(\Delta N = 1)$ and $S(\Delta N = 2)$ transitions on the high-frequency side of the Q branch, and a P-form branch due to $O(\Delta N = -2)$ and $P(\Delta N = -1)$ transitions[25,26] on the low-frequency side. At this level of resolution, most of the individual bands in the R- and P-form branches of the calculated spectrum are structureless. Yet some splittings are seen, similar to those observed in the experimental spectra.[25,26] For example, the $P(1)$ member has a shoulder on its red side, whereas $P(2)$ is structureless. $R(0)$ and $R(2)$ are weaker than $P(1)$ and exhibit more obvious splitting.

The FS splitting becomes more pronounced in the spectrum calculated with a higher resolution of about 600 MHz
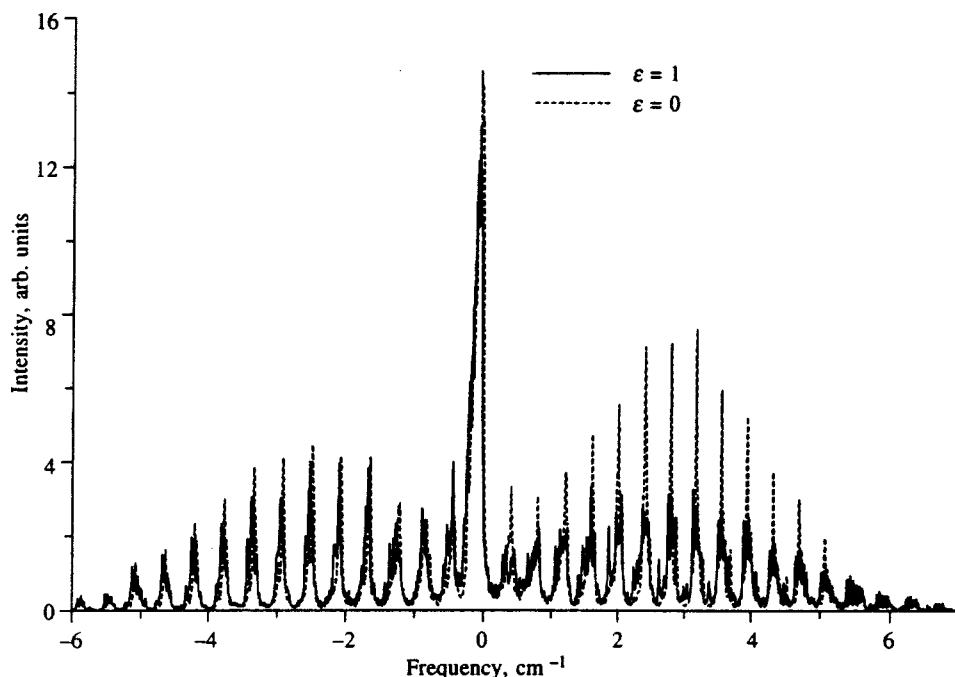
FIG. 4. Effect of fine structure splitting on the singlet–triplet spectrum of pyrazine. Resolution 0.02 cm^{-1}.

(bottom panel in Fig. 3). The line splittings are more extensive in the *R*-form branch than in the *P*-form branch because transitions in the *R*-form branch access higher *J* values in the triplet state.

The effect of the FS interaction on the spectrum is made clearer in Fig. 4 by the comparison of two spectra, one with $\varepsilon = 0$ and the other with $\varepsilon = 1$. With $\varepsilon = 0$, the *R*-form branch is stronger than the *P*-form branch. Introducing the FS splitting reduces the overall intensity of each branch, but the effect on the *R*-form branch is greater, since the degeneracy of the triplet levels for *R* and *S* transitions is higher than that for *P* and *O* transitions. Hence, the *R*- and *S*-type transitions are split into a larger number of components sharing the total intensity of a given transition. As a result, the intensities of the *P*- and *R*-form branches become more similar, and the spectrum acquires a sort of mirror-image symmetry with respect to the *Q* branch. Such uniformity of high-*J* and low-*J* parts of the spectrum is expected from the fact that the rotational spacing is proportional to $2J$ and the rotational level degeneracy to $2J + 1$, so the density of states is essentially independent of *J*.

Figure 4 also clearly demonstrates the transition from Case (a) to Case (b) behavior. In both *P*-form and *R*-form branches, the low *J* bands are split significantly, the splitting being of the order of the rotational band spacing. This corresponds to Hunds's Case (a) coupling (or, more precisely, Case (ab), as noted earlier). With *J* increasing, the splitting decreases and ultimately disappears in the high-*J* bands, characteristic of Case (b). Thus, we predict that Hund's Case (a)–Case (b) transition can be observed in as large a polyatomic molecule as pyrazine under quite moderate resolution conditions. The effect will also be observable in larger molecules, at higher resolution. Apart from the diatomic molecules mentioned in the Introduction, there have been no previous observations of this phenomenon in a polyatomic molecule. Recently, a pure Case (ab) spectrum was reported in H$_2$CSe.[49]

Figure 5 shows the *R*-form branch on an expanded scale at higher resolution (150 MHz), to illustrate the relative importance of the asymmetry and FS splittings in pyrazine. The asymmetry splitting is more important for transitions terminating in $J \leqslant 4$, whereas the FS splitting is more important at higher *J*. Note, in Fig. 5b, that the asymmetry split bands still preserve their identity as rotational bands since the splittings are smaller than the band separations. In contrast, the FS splitting in Panel c is large enough to fill the gaps between the low *J* bands. Thus, these bands are actually spin–rotational rather than rotational in nature. Clearly, a high resolution of about 100 MHz will be required to observe this behavior. Even higher resolution (~10 MHz) will be required to observe the true molecular eigenstates.

3.3. Singlet–triplet spectra of pyrazine-*d*$_4$ and the pyrazine–Ar van der Waals complex

We also used Eq. (73) to calculate the singlet–triplet spectra of pyrazine-*d*$_4$ and the pyrazine–Ar van der Waals complex. Our objective was to compare these spectra with those of pyrazine-*h*$_4$, thereby exploring the effects of varying the magnitudes of the rotational constants on the Case (a)–Case (b) transition. In the calculation on pyrazine-*d*$_4$, we used the known rotational constants of the S_0 state (see Table II) and assumed, for the T_1 state, that the rotational constants are reduced upon deuteration in the same proportion as in the S_0 state. The smallest moment of inertia in both states is now the moment about the *x* axis.[44] Thus, the rotational constants *A* and *B* are exchanged in the Hamiltonian (1). The nuclear spin statistical weights are $g(B_1) = g(B_3) = 6$, $g(B_2) = 7$, and $g(A) = 8$ in this case.

Taking these changes into account, we obtain the spectra
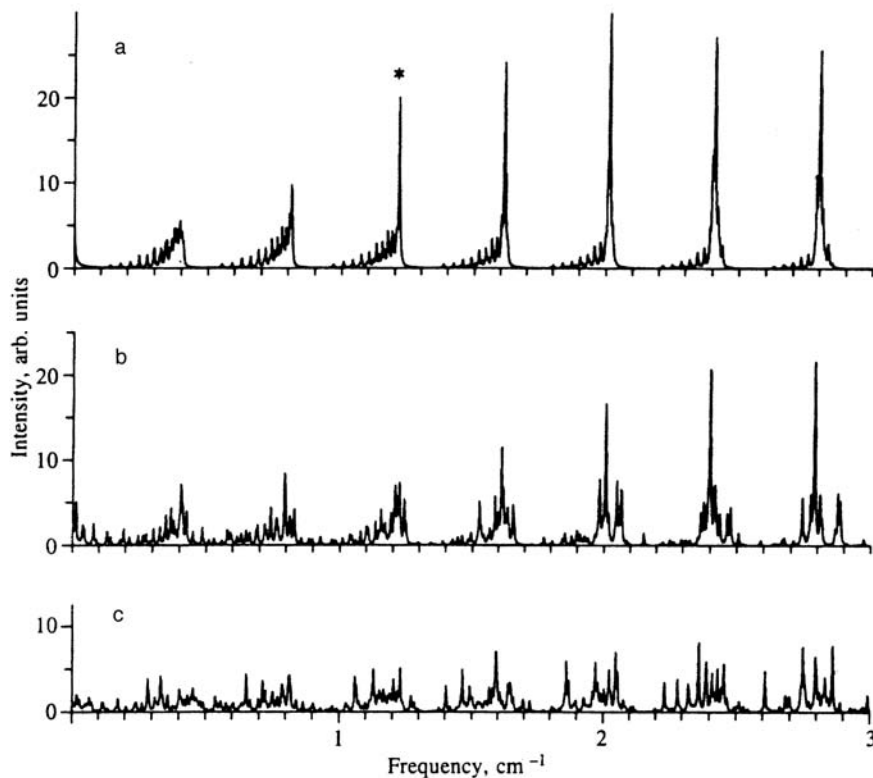
FIG. 5. Evolution of the R-form branch from a symmetric top with no fine structure splitting (a, $\varepsilon = 0$) to an asymmetric top with no fine structure splitting (b, $\varepsilon = 0$) to an asymmetric top with fine structure splitting (c, $\varepsilon = 1$). The starred band in panel a involves degenerate $R(2)$ and $S(1)$ lines; these are split in panel b due to asymmetry and further split in panel c due to the spin–spin interaction. Resolution 0.005 cm^{-1}.

shown in Fig. 6, with and without FS splitting. Comparing the two spectra, we again see that turning on the FS splitting reduces the intensity of the R-form branch relative to the P-form branch. We also note that numerous lines appear in the gaps between the rotational bands at low J, and that these lines disappear at high J. This behavior is again a manifestation of the Case (a)–Case (b) transition; turning on the

rotation decouples the spin from the molecular frame and distributes the oscillator strength more uniformly, as noted in Sec. 3.2.

The singlet–triplet spectrum of the pyrazine–Ar van der Waals complex has been observed in a supersonic jet, using the MPI technique[19] at low resolution. No rotationally resolved spectra have been reported to date. To model such a
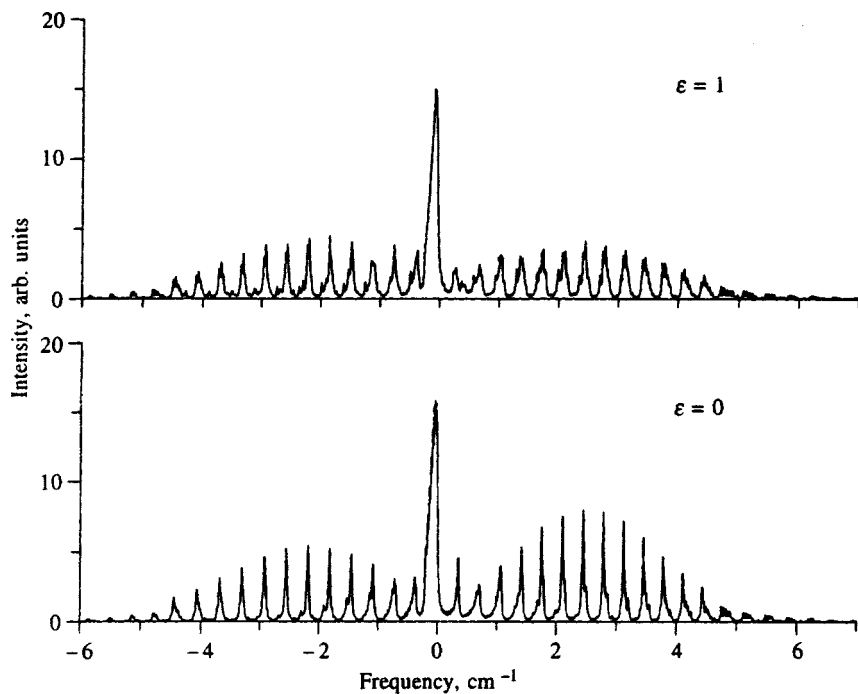
FIG. 6. The calculated singlet–triplet spectrum of pyrazine-d_4. Top panel, with fine structure splitting; bottom panel, without fine structure splitting. Resolution 0.01 cm^{-1}, $T_{\text{rot}} = 10$ K.

E. S. Medvedev and D. W. Pratt
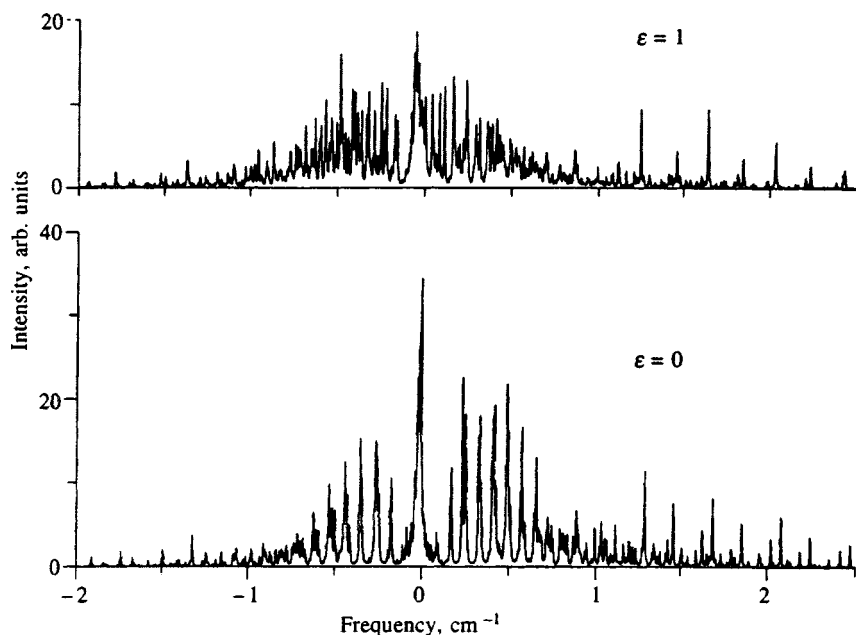
FIG. 7. The calculated singlet–triplet spectrum of the pyrazine–Ar van der Waals complex. Top panel, with fine structure splitting; bottom panel, without fine structure splitting. Resolution 0.005 cm^{-1}, $T_{rot} = 1.5$ K.

spectrum, we assumed that the Ar atom lies on the z axis perpendicular to the ring plane, at a distance of 3.5 Å, as it is in benzene–Ar[50] and s-tetrazine–Ar.[51] With this model, the C rotational constant remains unchanged, whereas the A and B constants are reduced by a factor of 4.8 in both states. The nuclear spin statistical weights are $g(A) = g(B_1) = 13$ and $g(B_2) = g(B_3) = 11$. The spectra calculated using these modified parameters are shown in Fig. 7.

Comparing the spectra for these species (Figs. 3, 6, and 7), we see above all that the singlet–triplet spectrum of pyrazine–Ar is considerably more congested than that of the bare molecule, owing to the significant decrease in A and B. Still, the complex exhibits well-defined branches in its spectrum in the absence of the FS interaction. However, turning on this interaction has a dramatic effect on the spectrum. The low-J rotational transitions, within ± 1 cm^{-1} of the Q branch, are extensively mixed by the spin–spin coupling. Still, a defined level structure exists, although extremely high resolution (~ 1 MHz) will be required to expose the individual eigenstates. The long lifetime of the triplet state should permit such experiments in the near future, raising the intriguing possibility of seeing still more underlying structure, including that due to hyperfine interaction and/or couplings to nearly isoenergetic ground-state levels. The possibility of observing such effects is enhanced at high J (cf. Fig. 7), where the structure with $\varepsilon \neq 0$ is even simpler than that with $\varepsilon = 0$, the transition to pure Case (b) being "complete."

4. CONCLUSIONS

A closed-form analytic expression has been derived for calculating the intensities of individual spin–rovibronic lines in the singlet–triplet absorption spectrum of a polyatomic molecule. This expression takes into account both the intra-manifold spin–spin coupling within the triplet state and the intermanifold spin–orbit coupling of the triplet to an excited singlet state. It also includes asymmetry splittings, and therefore can be applied to asymmetric tops as well as symmetric

tops. With this expression, we have calculated the spectra of three species: pyrazine-h_4, pyrazine-d_4, and the pyrazine–Ar van der Waals complex, using the available experimental values of the gas-phase rotational and solid-state fine structure parameters of pyrazine-h_4 and model parameters for the remaining molecules. Comparison of the predictions of the theory with the available data for pyrazine-h_4 shows good agreement with experiment. The remaining molecules have yet to be examined at high resolution.

The computed spectra exhibit a number of interesting properties, the most notable being the transition from Case (a) to Case (b) with increasing J and/or increasing molecular size. Inclusion of the fine structure interaction in the zero-order triplet state results in a decrease in the intensity of the R- and S-transitions, compared to P- and O-transitions branches, and a more symmetric spectrum. This effect has been observed in moderate resolution experiments.[25,26] Additionally, spin–rotational transitions appear in the gaps between "pure" rotational transitions at low J, fragmenting the spectrum, but disappear at high J. This effect is more pronounced in pyrazine-d_4 and the pyrazine–Ar van der Waals complex, owing to their smaller rotational constants. All of the predicted behavior should be observable in experiments performed with a resolution of $\leqslant 100$ MHz.

The analytic expression derived here can be used to interpret the singlet–triplet spectrum of any polyatomic molecule with arbitrarily large (or small) rotational and fine structure constants.

APPENDIX

General formula for the intensities and its relation to the Hougen factors

A general formula for the intensities can be derived from Eq. (62) if we represent the triplet function in the most general form, Eqs. (10) and (15), not invoking Eq. (53). Introducing the notation

$$\Omega_{\sigma'\sigma} = \langle \Gamma_T; 1\sigma' | \mu_\sigma^{(1)} | \Gamma_S \rangle, \tag{A1}$$

we obtain

$$I(\Gamma_T \Gamma_r' \lambda' S' J' i' \leftarrow \Gamma_S \Gamma_r'' \lambda'' \tau'' S'' J'' N'' i'')$$

$$= \frac{1}{4} (2J'+1)(2J''+1)$$

$$\times \left| \sum_{N'=|J'-1|}^{J'+1} \sqrt{2N'+1} \right.$$

$$\times \sum_{\substack{K'=0 \\ \text{(only even or odd)}}}^{N'} f_{K'} C_{N'K'}^{(i')}$$

$$\times \sum_{\substack{K''=0 \\ \text{(only even or odd)}}}^{J''} f_{K''} C_{K''}^{(i'')}$$

$$\times \left. \sum_{P'=-J'}^{J'} \sum_{\sigma'\sigma=0,\pm 1} \Omega_{\sigma'\sigma} F_{P'\sigma'\sigma} \right|^2, \tag{A2}$$

where $F_{P'\sigma'\sigma}$ is given by Eq. (74). The selection rules are the same as in Eq. (75), plus $\Delta K = \pm 1$. The rule $\lambda' = -\lambda''$ does not apply anymore.

A general expression for $\Omega_{\sigma'\sigma}$ found from Eqs. (41), (43), and (56) is

$$\Omega_{\sigma'\sigma} = \sum \frac{i}{\Delta E} c_{\sigma'}^{(1)*} \mu_\sigma^{(1)}, \tag{A3}$$

where the star denotes the complex conjugate, $\mu_\sigma^{(1)}$ now stands for $\langle \Gamma_{S_n} | \mu_\sigma^{(1)} | \Gamma_S \rangle$, and the summation is over all contributing singlet states S_n. The following identity can be easily derived either by applying the time reversal operator (14) or from the definition of the spherical tensors in Eq. (42):

$$\Omega_{-\sigma',-\sigma} = (-1)^{\sigma'+\sigma+1} \Omega_{\sigma'\sigma}^*. \tag{A4}$$

For pyrazine, retaining a single term of the sum (A3) and inserting Eqs. (45), (61), and (70), we find

$$\Omega_{\sigma'\sigma} = -\frac{1}{2} \sigma' |\sigma| \mu(B_{3g}). \tag{A5}$$

In this case Eq. (A4) reduces to

$$\Omega_{-\sigma',-\sigma} = -\Omega_{\sigma'\sigma}. \tag{A6}$$

Denoting the last sum in Eq. (A2) by $A(P')$ and using Eq. (A6) and the properties of 3-J symbols in Eq. (74), we obtain

$$A(-P') = -\lambda'\lambda'' A(P'). \tag{A7}$$

Inserting this into the identity

$$\sum_P A(P) \equiv \sum_{P \geq 0} f_P^2 [A(P) + A(-P)] \tag{A8}$$

following from Eq. (4), we immediately obtain the selection rule $\lambda' = -\lambda''$. Now, it is easy to verify that Eq. (A2) reduces to Eq. (73) in the particular case of pyrazine.

The intensities of individual rotational lines in the singlet–triplet spectrum of a symmetric top molecule without fine structure splitting were given by Hougen[38] in tabular form for the D_{2h} point symmetry group, together with directions for the use of these tables for other groups. Our general formula (A2) enables us to derive a comprehensive expression embracing all cases considered by Hougen. In Eq. (A2), one must put

$$C_{N'K'}^{i'} = \delta_{i',N'K'} \quad \text{and} \quad C_{K''}^{i''} = \delta_{i'',K''},$$

and then convert it to yield the intensities for transitions between nonsymmetrized states (with K' and K'' taking both positive and negative values). This is performed by a simple transformation of the transition amplitude from basis (3) to basis (2). The result reads

$$I(\Gamma_T J'N'K' \leftarrow \Gamma_S J''K'') = (2N'+1)(2J'+1)(2J''+1)$$

$$\times \left| \sum_{P'\sigma'\sigma} \Omega_{\sigma'\sigma} \begin{pmatrix} J' & 1 & N' \\ -P' & \sigma & K' \end{pmatrix} \right.$$

$$\times \left. \begin{pmatrix} J' & 1 & J'' \\ -P' & \sigma' & K'' \end{pmatrix} \right|^2, \tag{A9}$$

where the actual summation is performed only over σ, since $P' = \sigma + K'$ and $\sigma' = \sigma + \Delta K$ ($\Delta K \equiv K' - K''$). The 3×3 matrix (A1) can be expressed in terms of nine real intensity parameters, Ω_{ik}, $i,k = x,y,z$, in the same manner as each member of the sum of Eq. (A3) can be expressed in terms of $c_i \mu_k / \Delta E$ using the definition (42). For instance,

$$\Omega_{00} = i\Omega_{zz}, \quad \Omega_{+1,-1} = \frac{1}{2} (-iQ_{xx} + i\Omega_{yy} - \Omega_{xy} - \Omega_{yx}), \tag{A10}$$

etc. For the D_{2h} group, our intensity parameters can be related to the Hougen parameters $\mu(B_{1g})$, $\mu(B_{2g})$, and $\mu(B_{3g})$ as shown in the following examples.

For $\Delta K = \Delta N = \Delta J = 0$, Eq. (A9) yields

$$I = \frac{2J+1}{J^2(J+1)^2} \left| \frac{i}{2} (\Omega_{xx} + \Omega_{yy})[K^2 - J(J+1)] \right.$$

$$\left. + i\Omega_{zz} K^2 - \frac{1}{2} (\Omega_{xy} - \Omega_{yx})K \right|^2, \tag{A11}$$

where $K = K''$ and $J = J''$. Thus, in general, five intensity parameters govern this particular transition. In case of the $^3A_u \leftarrow {}^1A_{1g}$ transition in the D_{2h} group, the possible contaminating singlets having nonzero transition moments $\mu_{x,y,z}$ from the ground state are $^1B_{3u}$, $^1B_{2u}$, and $^1B_{1u}$, respectively. The corresponding triplet sublevels acquiring the os-

cillator strength via the spin–orbit coupling parameters $c_{x,y,z}$ are $T_{x,y,z}$. Since the transformation rules for Ω_{ik} are the same as for $c_i\mu_k$, we obtain the following three nonvanishing parameters, assigning them Hougen's notations:

$$\mu(B_{1g}) = \Omega_{zz}, \quad \mu(B_{2g}) = -\Omega_{yy}, \quad \mu(B_{3g}) = \Omega_{xx}. \tag{A12}$$

These definitions apply to all transitions with $\Gamma_S \times \Gamma_T = A_u$. Equations (A9) and (A12) entirely reproduce Hougen's Table I, apart from an insignificant phase factor of the transition amplitude. Similarly, for $\Gamma_S \times \Gamma_T = B_{1u}$, two nonvanishing parameters are

$$\mu(B_{2g}) = -\Omega_{yx}, \quad \mu(B_{3g}) = \Omega_{xy}. \tag{A13}$$

Equations (A9) and (A13) reproduce Hougen's Table II. Pyrazine belongs here, with $\mu(B_{2g}) = 0$ (cf. Eqs. (A5), (A10), and (A13)). For $\Gamma_S \times \Gamma_T = B_{2u}$ or B_{3u}, all five parameters vanish.

For $\Delta K = \Delta N = \Delta J = 1$ Eq. (A9) yields

$$I = \frac{(J+K+1)(J+K+2)}{4(J+1)^2(J+2)} \left| (i\Omega_{xz} - \Omega_{yz})(J-K+1) \right.$$
$$\left. + (i\Omega_{zx} - \Omega_{zy})(K+1) \right|^2. \tag{A14}$$

For $\Gamma_S \times \Gamma_T = B_{2u}$ in the D_{2h} group, the nonvanishing parameters are

$$\mu(B_{1g}) = \Omega_{zx}, \quad \mu(B_{3g}) = \Omega_{xz}. \tag{A15}$$

Equations (A9) and (A15) reproduce Hougen's Table III. Application of Eq. (A9) to any other point group is straightforward; one must calculate the intensity of a particular ΔK, ΔN, ΔJ transition and then determine, by the usual rules, which of the intensity parameters vanish. For a molecule with no symmetry, all nine intensity parameters Ω_{ik} are present, but not all of them are relevant to a given transition. Thus, five parameters govern the intensity of the $\Delta K = \Delta N = \Delta J = 0$ branch, Eq. (A11), and the other four govern the intensity of the $\Delta K = \Delta N = \Delta J = 1$ branch, Eq. (A14).

*)E-mail: medvedev@landau.ac.ru

[1] E. S. Medvedev and V. I. Osherov, *Radiationless Transitions in Polyatomic Molecules*, Springer, Berlin (1995).
[2] O. Sekiguchi, N. Ohta, and H. Baba, Laser Chem. **7**, 213 (1987).
[3] J. Kommandeur, W. A. Majewski, W. L. Meerts, and D. W. Pratt, Annu. Rev. Phys. Chem. **38**, 433 (1987).
[4] J. Kommandeur, Adv. Chem. Phys. **70**, 1 (1988).
[5] A. Amirav, J. Phys. Chem. **92**, 3725 (1988).
[6] E. S. Medvedev, Usp. Fiz. Nauk **161**, 31 (1991) [Sov. Phys. Usp. **34**, 16 (1991)].
[7] E. S. Medvedev and D. W. Pratt, J. Chem. Phys. **105**, 3366 (1996).
[8] *The Triplet State*, A. B. Zahlan (Ed.), Cambridge University Press, Cambridge (1963).
[9] D. M. Burland and J. Schmidt, Mol. Phys. **22**, 19 (1971).
[10] M. A. El-Sayed, in *MTP International Review of Science*, Physical Chemistry Series 1, D. A. Ramsay (Ed.), Butterworths, London (1972), Vol. 3, p. 119.

[11] A. A. Gwaiz and M. A. El-Sayed, Chem. Phys. Lett. **19**, 11 (1973).
[12] U. P. Wild, in *Triplet States*, Springer, Berlin (1975), Vol. 2.
[13] H. Hirt, Spectrochim. Acta **12**, 114 (1958).
[14] L. Goodman and M. Kasha, J. Mol. Spectrosc. **2**, 58 (1958).
[15] R. M. Hochstrasser and C. Marzzacco, J. Chem. Phys. **49**, 971 (1968).
[16] K. K. Innes and L. E. Giddings, Jr., Discuss. Faraday Soc. **35**, 192 237 (1963).
[17] A. E. Douglas and M. Milton, Discuss. Faraday Soc. **35**, 235 (1963).
[18] G. Fisher, Chem. Phys. Lett. **79**, 573 (1981).
[19] E. Villa, M. Terazima, and E. C. Lim, Chem. Phys. Lett. **129**, 336 (1986).
[20] L. H. Spangler, Y. Matsumoto, and D. W. Pratt, J. Phys. Chem. **87**, 4781 (1983).
[21] L. H. Spangler and D. W. Pratt, J. Chem. Phys. **84**, 4789 (1986).
[22] L. H. Spangler, D. W. Pratt, and F. W. Birss, J. Chem. Phys. **85**, 3229 (1986).
[23] J. L. Tomer, K. W. Holtzclaw, D. W. Pratt, and L. H. Spangler, J. Chem. Phys. **88**, 1528 (1988).
[24] A. Penner and A. Amirav, J. Phys. Chem. **92**, 5079 (1988).
[25] A. Penner, Y. Oreg, E. Villa, E. C. Lim, and A. Amirav, Chem. Phys. Lett. **150**, 243 (1988).
[26] K. W. Holtzclaw, L. H. Spangler, and D. W. Pratt, Chem. Phys. Lett. **161**, 347 (1989).
[27] O. Sneh, D. Dünn-Kittelplon, and O. Cheshnovsky, J. Chem. Phys. **91**, 7331 (1989); I. Becker and O. Cheshnovsky, J. Chem. Phys. **101**, 3649 (1994).
[28] S. Hillenbrand, L. Zhu, and P. M. Johnson, J. Chem. Phys. **95**, 2237 (1991).
[29] J. W. Sidman, J. Mol. Spectrosc. **2**, 33 (1958).
[30] E. Clementi and M. Kasha, J. Mol. Spectrosc. **2**, 297 (1958).
[31] L. Goodman and V. G. Krishna, Rev. Mod. Phys. **35**, 541,735 (1963).
[32] M. A. El-Sayed, in *Molecular Luminescence*, E. C. Lim (Ed.), Benjamin, New York (1969), p. 715.
[33] M. A. El-Sayed, J. Chem. Phys. **38**, 2834 (1963).
[34] D. S. McClure, J. Chem. Phys. **20**, 682 (1952).
[35] B. J. Cohen and L. Goodman, J. Chem. Phys. **46**, 713 (1967).
[36] G. Herzberg, *Molecular Spectra and Molecular Structure. I. Spectra of Diatomic Molecules*, Van Nostrand, Princeton (1964), p. 231.
[37] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, 3rd ed., Pergamon, Oxford (1981), p. 319.
[38] J. T. Hougen, Can. J. Phys. **42**, 433 (1964).
[39] F. Creutzberg and J. T. Hougen, Can. J. Phys. **45**, 1363 (1967).
[40] C. di Lauro, J. Mol. Spectrosc. **35**, 461 (1970).
[41] J. H. Van Vleck, Rev. Mod. Phys. **23**, 213 (1951).
[42] W. T. Raynes, J. Chem. Phys. **41**, 3020 (1964).
[43] H. Sternlicht, J. Chem. Phys. **38**, 2316 (1963).
[44] K. K. Innes, I. G. Ross, and W. R. Moomaw, J. Mol. Spectrosc. **132**, 492 (1988).
[45] W. E. Howard and E. W. Schlag, in *Radiationless Transitions*, S. H. Lin (ed.), Academic, New York (1980), p. 81.
[46] L. Goodman and V. G. Krishna, J. Chem. Phys. **37**, 2721 (1962).
[47] P. R. Bunker, *Molecular Symmetry and Spectroscopy*, Academic, New York (1979).
[48] D. A. Kleier, R. L. Martin, W. R. Wadt, and W. R. Moomaw, J. Am. Chem. Soc. **104**, 60 (1982).
[49] D.-L. Joo, D. J. Clouthier, R. H. Judge, and D. C. Moule, J. Chem. Phys. **102**, 7351 (1995).
[50] Th. Weber, A. von Bargen, E. Riedle, and H. J. Neusser, J. Chem. Phys. **92**, 90 (1990); Th. Weber and H. J. Neusser, J. Chem. Phys. **94**, 7689 (1991).
[51] P. M. Weber and S. A. Rice, J. Chem. Phys. **88**, 6120 (1988).

Radiation of a charged particle in a random stack of plates

Zh. S. Gevorkian*)

Institute of Radiophysics and Electronics, National Academy of Sciences of Armenia, 378410 Ashtarak-2, Armenia

Radiation from a charged particle moving in a system of randomly spaced plates is considered. It is shown that the dominant radiation mechanism is diffusion. The total intensity of radiation is investigated, and its quadratic dependence on particle energy is noted in the optical region. A comparison with Cherenkov radiation is carried out. © *1998 American Institute of Physics.* [S1063-7761(98)00607-6]

1. INTRODUCTION

More than 50 years ago, Ginzburg and Frank[1] showed that radiation is produced when a charged, uniformly moving particle passes through the interface between two media with different dielectric constants. Since then, much research has been done on this problem (for a review, see for example Ginzburg and Tsytovich[2]). It turns out that the dependence of total intensity of radiation at an isolated interface on particle energy is logarithmic in the optical region. To be able to use transition radiation to detect relativistic charged particles, it is desirable to have a stronger energy dependence. In this context, the x-ray region turns out to be more promising, because in this region the energy dependence of the radiation intensity is linear.[3] However, the number of photons emitted at the interface is small. To increase this number, systems of many plates are used. Earlier, when investigating radiation in a stack of plates, mainly the x-ray region was considered (see for example Garibian and Yang[4]). In this region the interaction of the electromagnetic field with each plate is weak, so multiple scattering effects can be neglected.

The objective of our paper is to take these effects into account when charged particles radiate while traversing a random stack of plates. Having considered three-dimensional random media,[5] we know that multiple scattering effects in the electromagnetic field play a crucial role in the radiation of a charged particle. Below we show that in the one-dimensional case, these effects play an even more important role, particularly in the optical region.

2. FORMULATION OF THE PROBLEM

The system which we want to study is a stack of plates randomly spaced in a homogeneous medium. Let us assume that the plates fill the regions $z_i - a/2 < z < z_i + a/2$ (where a is the plate thickness and z_i are random coordinates). The permittivity of the system can be represented in the form

$$\varepsilon(z,\omega) = \varepsilon_0(\omega) + \sum_i \left[b(\omega) - \varepsilon_0(\omega) \right]$$

$$\times \left[\left| \theta(z - z_i - a/2) - \theta(z - z_i + a/2) \right| \right], \quad (1)$$

where $\varepsilon_0(\omega)$ and $b(\omega)$ are the permittivity of the homogeneous medium and the plate, respectively, and $\theta(x)$ is the Heaviside step function. It is convenient to represent the permittivity as a sum of average and varying parts:

$$\varepsilon(z,\omega) = \varepsilon + \varepsilon_r(z,\omega), \quad \langle \varepsilon_r(z,\omega) \rangle = 0, \quad (2)$$

where $\varepsilon = \langle \varepsilon(z,\omega) \rangle$, and averaging over random coordinates of plates is defined as follows:

$$\langle f(z,\omega) \rangle = \int \prod_i \frac{dz_i}{L_z} f(z, z_i, \omega), \quad (3)$$

where L_z is the system size in the z-direction.

In the Fourier representation, Maxwell's equations for the vector potential $\mathbf{A}$ of the electromagnetic field has the form

$$\nabla^2 \mathbf{A} + \frac{\omega^2}{c^2} \varepsilon(\mathbf{r},\omega) \mathbf{A}(\mathbf{r},\omega) = \mathbf{j}(\mathbf{r},\omega), \quad (4)$$

where

$$\mathbf{j}(\mathbf{r},\omega) = -\frac{4\pi e}{c} \frac{\mathbf{v}}{v} \delta(x) \delta(y) e^{i\omega z/v}$$

is the current of a charged particle moving uniformly in the z-direction with velocity v. The electric field $\mathbf{E}$ is related to the potentials by

$$\mathbf{E}(\mathbf{r},\omega) = \frac{i\omega}{c} \mathbf{A}(\mathbf{r},\omega) - \nabla \varphi(\mathbf{r},\omega). \quad (5)$$

Finally we write the condition relating the vector and scalar potentials of the electromagnetic field:

$$\nabla \cdot \mathbf{A} - \frac{i\omega}{c} \varepsilon(\mathbf{r},\omega) \varphi(\mathbf{r},\omega) = 0. \quad (6)$$

One needs the relations (4)–(6) to calculate the intensity of radiation. It follows from the symmetry of the problem that the vector potential $\mathbf{A}$ points in the z-direction, so $A_i = \delta_{zi} A(\mathbf{r},\omega)$.

We separate the electric field into two parts, $\mathbf{E} = \mathbf{E}_0 + \mathbf{E}_r$, to determine the radiation intensity. Here $\mathbf{E}_0$ is the electric field of a charge moving in a homogeneous medium

with dielectric constant ε, and $\mathbf{E}_r$ is the radiation field associated with fluctuations of the dielectric constant. The radiation tensor is

$$I_{ij}(\mathbf{R}) = E_{ri}(\mathbf{R})E_{rj}^*(\mathbf{R}), \qquad (7)$$

where $\mathbf{R}$ is the radius vector of the observation point, which is far from the system ($R \gg L$).

The vector potential can be split in a similar manner, $\mathbf{A} = \mathbf{A}_0 + \mathbf{A}_r$, where $\mathbf{A}_0$ and $\mathbf{A}_r$ satisfy the equations

$$\nabla^2 \mathbf{A}_0 + \frac{\omega^2}{c^2}\,\varepsilon \mathbf{A}_0 = \mathbf{j}(\mathbf{r},\omega),$$

$$\nabla^2 \mathbf{A}_r + \frac{\omega^2}{c^2}\,\varepsilon \mathbf{A}_r + \frac{\omega^2}{c^2}\,\varepsilon_r \mathbf{A}_r = -\frac{\omega^2}{c^2}\,\varepsilon_r \mathbf{A}_0. \qquad (8)$$

The first equation is easily solved, and for the background field $\mathbf{A}_0$ one has

$$A_0(\mathbf{q}) = -\frac{8\pi^2 e}{c}\,\frac{\delta(q_z - \omega/v)}{k^2 - q^2}. \qquad (9)$$

It is convenient to express the radiation tensor (7) in terms of the radiation potential $\mathbf{A}_r$. Using (5)–(7), we obtain

$$\langle I_{ij}(\mathbf{R})\rangle = \frac{\omega^2}{c^2}\,\delta_{zi}\delta_{zj}\langle A_r(\mathbf{R},\omega)A_r^*(\mathbf{R},\omega)\rangle$$

$$+ \frac{\delta_{zi}}{\varepsilon}\left\langle A_r(\mathbf{R},\omega)\frac{\partial^2}{\partial R_j \partial z}A_r^*(\mathbf{R},\omega)\right\rangle$$

$$+ \frac{\delta_{zj}}{\varepsilon}\left\langle A_r^*(\mathbf{R},\omega)\frac{\partial^2}{\partial R_i \partial z}A_r(\mathbf{R},\omega)\right\rangle$$

$$+ \frac{c^2}{\omega^2 \varepsilon^2}\left\langle \frac{\partial^2}{\partial R_i \partial z}A_r(\mathbf{R},\omega)\frac{\partial^2}{\partial R_j \partial z}A_r^*(\mathbf{R},\omega)\right\rangle. \qquad (10)$$

We express the radiation potential A_r in terms of the Green's function of the second equation in (8) for averaging in (10):

$$A_r(\mathbf{R}) = -\frac{\omega^2}{c^2}\int \varepsilon_r(\mathbf{r})A_0(\mathbf{r})G(\mathbf{R},\mathbf{r})\,d\mathbf{r}, \qquad (11)$$

where the Green's function satisfies

$$\left[\nabla^2 + k^2 + \frac{\omega^2}{c^2}\,\varepsilon_r(z)\right]G(\mathbf{r},\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}'), \qquad (12)$$

and $k = \omega\sqrt{\varepsilon}/c$.

3. GREEN'S FUNCTION

The bare ($\varepsilon_r = 0$) Green's function can easily be obtained from (12):
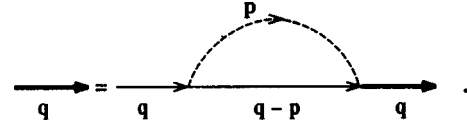
$$G_0(\mathbf{q}) = \frac{1}{k^2 - q^2 + i\delta}. \qquad (13)$$

In the coordinate representation one has from (13)

$$G_0(r) = -\frac{1}{4\pi r}\,e^{ikr}. \qquad (14)$$

To perform the averaging, we use the impurity-diagram method.[6] Summing the diagrams in the independent-scatterer

approximation, we obtain the Dyson equation for the average Green's function:

$$(15)$$

The dashed line denotes the Fourier component $B(\mathbf{p}) = (2\pi)^2 \delta(\mathbf{p}_\rho)B(|p_z|)$ of the correlation function of the one-dimensional random field

$$B(|z-z'|) = \frac{\omega^4}{c^4}\langle \varepsilon_r(z)\varepsilon_r(z')\rangle, \qquad (16)$$

where $\mathbf{p}_\rho$ is the transverse component of $\mathbf{p}$. The solution of Eq. (15) can be represented in the form

$$G(\mathbf{q}) = \frac{1}{k^2 - q^2 + i\,\mathrm{Im}\,\Sigma(\mathbf{q})}, \qquad (17)$$

in which the imaginary part of the self-energy is given by Ward's identity:

$$\mathrm{Im}\,\Sigma(\mathbf{q}) = \int \frac{d\mathbf{p}}{(2\pi)^3}\,B(\mathbf{p})\,\mathrm{Im}\,G_0(\mathbf{q}-\mathbf{p})$$

$$= \frac{1}{4\sqrt{k^2-q_\rho^2}}\left[B(|q_z - \sqrt{k^2 - q_\rho^2}|)\right.$$

$$\left. + B(|q_z + \sqrt{k^2 - q_\rho^2}|)\right], \quad |\mathbf{q}_\rho| < k. \qquad (18)$$

The dephasing length of a pseudophoton in the z-direction is determined by the imaginary part of the self-energy:

$$l(\mathbf{q}) = \frac{\sqrt{k^2 - q_\rho^2}}{\mathrm{Im}\,\Sigma(\mathbf{q})}. \qquad (19)$$

As expected, the dephasing length depends on the direction of the pseudophoton momentum. When the momentum is directed along z, one obtains from (18) and (19)

$$l(\vartheta = 0) = \frac{4k^2}{B(0) + B(2k)}. \qquad (20)$$

From this point on, we call this quantity the pseudophoton mean free path.

Using (1)–(3) and (16), one can find the correlation function

$$B(q_z) = \frac{4(b-\varepsilon)^2 n\,\sin^2(q_z a/2)}{q_z^2}\,\frac{\omega^4}{c^4}. \qquad (21)$$

Here $n = N/L_z$ is the density of plates in the system. Using (21), it is easy to see that $B(2k)/B(0) \sim 1/(ka)^2 \ll 1$ when $ka \gg 1$. Therefore, the photon mean free path is

$$l \equiv l(\vartheta = 0) \approx \begin{cases} 4k^2/B(0), & ka \gg 1 \\ 2k^2/B(0), & ka \ll 1. \end{cases} \qquad (22)$$

The foregoing only holds in the weak scattering regime, for which $\mathrm{Im}\,\Sigma(\mathbf{q})/(k^2 - q_\rho^2) \ll 1$. Substituting (18) into this condition, we obtain

$$\frac{B(0)+B(2k|\cos\vartheta|)}{4k^3|\cos\vartheta|^3}\ll 1. \qquad (23)$$

It follows from (23) that at $\vartheta\approx\pi/2$, the weak-scattering condition is not satisfied. This is natural, because in this case the pseudophoton moves parallel to the plates. Taking $\vartheta=\pi/2-\delta$ and using (22) and (23), one has $\delta\gg(1/kl)^{1/3}$.

4. RADIATION INTENSITY

We now turn to a close examination of radiation intensity. First we consider the single-scattering approximation. In this approximation, the Green's function in (11) is simply replaced by the bare one. Substituting (11) into (10) and using the relations

$$G_0(\mathbf{R},\mathbf{r})\approx-\frac{1}{4\pi R}\,e^{ik(R-\mathbf{n}\cdot\mathbf{r})},$$

$$\frac{\partial^2 G_0(\mathbf{R},\mathbf{r})}{\partial R_i\partial z}\approx\frac{k^2 n_i n_z}{4\pi R}\,e^{ik(R-\mathbf{n}\cdot\mathbf{r})},\quad R\gg r \qquad (24)$$

and (14), after simple transformations we obtain the following expression for the single-scattering contribution to radiation intensity $I(\mathbf{n})=(c/2)R^2 I_{ii}(\mathbf{R})$:

$$I^0(\mathbf{n})=\frac{\pi e^2}{c}\,\delta(0)\,\frac{B(|k_0-kn_z|)n_\rho^2}{[k^2 n_z^2-k_0^2]^2}\,\frac{\omega^2}{c^2}. \qquad (25)$$

Here $\mathbf{n}=\mathbf{R}/R$ is the unit vector in the direction of the observation point $\mathbf{R}$ and $k_0=\omega/v$; the δ-type singularity of (25) results from the infinite path of the charged particle in the medium. If one takes into account the finite size of the system, $\delta(0)$ must be replaced by $L_z/2\pi$. To analyze the angular dependence of (25), it is convenient to represent it in the form

$$I^0(\theta)=\frac{e^2}{2c}\frac{L_z B(|k_0-k\cos\vartheta|)\sin^2\vartheta}{[\gamma^{-2}+(k^2/k_0^2)\sin^2\vartheta]^2}\frac{\omega^2}{k_0^4 c^2}, \qquad (26)$$

where $\gamma=(1-\varepsilon v^2/c^2)^{-1/2}$ is the Lorentz factor of the particle in the medium, $n_z=\cos\vartheta$, and $n_\rho=\sin\vartheta$.

Note the key features of the single-scattering contribution I^0. It follows from (26) and the form (21) of the correlation function B that at relativistic energies ($\gamma\gg 1$, $k_0\to k$), the maximum of radiation lies in the range of angles $\vartheta\sim\gamma^{-1}$ in the forward direction. Integrating (21) over the angles, it is easy to see that the dependence of total intensity on particle energy is logarithmic, $I^0\propto\ln\gamma$. As $B\propto n$, the dependence of radiation intensity on the number of plates is linear. All of these results are consistent with previous results.[2,4]

We now consider the diffusion contribution to the radiation intensity. Using (10), (11) and (22), one can represent the diffusion contribution to the radiation tensor in the form
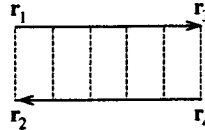
$$I_{ij}^D(\mathbf{R})=\frac{k^2}{16\pi^2 R^2\varepsilon}\int d\mathbf{r}\,d\mathbf{r}'B(\mathbf{r}-\mathbf{r}')A_0(\mathbf{r})A_0^*(\mathbf{r}')$$

$$\times\int d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 d\mathbf{r}_4\,\exp[-ik\mathbf{n}(\mathbf{r}_1-\mathbf{r}_2)]$$

$$\times P(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\mathbf{r}_4)G(\mathbf{r}_3,\mathbf{r})G^*(\mathbf{r},\mathbf{r}_4)$$

$$\times[\delta_{zi}\delta_{\hat z j}+n_i n_j n_z^2-\delta_{\hat z i}n_j n_z-\delta_{\hat z j}n_i n_z], \qquad (27)$$

where $P(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\mathbf{r}_4)$ is the diffusion propagator:

$$(28)$$

We find the diffusion propagator as in the three-dimensional case.[5] As follows from (28), P can be represented in the form

$$P(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\mathbf{r}_4)=B(\mathbf{r}_1-\mathbf{r}_2)B(\mathbf{r}_3-\mathbf{r}_4)$$

$$\times P(\mathbf{R}',\mathbf{r}_1-\mathbf{r}_2,\mathbf{r}_3-\mathbf{r}_4), \qquad (29)$$

where $\mathbf{R}'=(1/2)(\mathbf{r}_3+\mathbf{r}_4-\mathbf{r}_1-\mathbf{r}_2)$ and P satisfies the equation

$$\int\frac{d\mathbf{p}}{(2\pi)^3}\left[1-\int\frac{d\mathbf{q}}{(2\pi)^3}f(\mathbf{q},\mathbf{K})B(\mathbf{p}-\mathbf{q})\right]P(\mathbf{K},\mathbf{p},\mathbf{q}')$$

$$=f(\mathbf{q}',\mathbf{K}), \qquad (30)$$

where

$$f(\mathbf{q},\mathbf{K})=G(\mathbf{q}+\mathbf{K}/2)G^*(\mathbf{q}-\mathbf{K}/2). \qquad (31)$$

As will be seen, one has to know P when $\mathbf{K}\to 0$. In this limit, the diffusion propagator has the form[5]

$$P(\mathbf{K}\to 0,\mathbf{p},\mathbf{q})=\frac{\text{Im }G(\mathbf{p})\text{Im }G(\mathbf{q})}{\text{Im }\Sigma(\mathbf{q})}A(\mathbf{K}), \qquad (32)$$

where

$$A(\mathbf{K})=\left[\int\frac{(\mathbf{q}\cdot\mathbf{K})^2\,\text{Im }G(\mathbf{q})}{\text{Im}^2\,\Sigma(\mathbf{q})}\frac{d\mathbf{q}}{(2\pi)^3}\right]^{-1}. \qquad (33)$$

Substituting (17) and (18) into (33), choosing $\mathbf{K}\|\hat z$, and calculating the integral, we obtain

$$A(K)=\frac{1}{k}\frac{20\pi}{K^2 l^2}. \qquad (34)$$

When we know the form of the diffusion propagator, we can calculate the diffusion contribution to the radiation intensity. Transforming variables in (27) and going to the Fourier representation, we obtain

$$I^D(\mathbf{n})=\frac{k^2 c}{32\pi^2\varepsilon}(1-n_z^2)\int\frac{d\mathbf{q}_1 d\mathbf{q}_2 d\mathbf{q}_3 d\mathbf{q}_4}{(2\pi)^{12}}$$

$$\times|A_0(\mathbf{q}_1)|^2 B(\mathbf{q}_2)B(\mathbf{q}_3)B(\mathbf{q}_4)P(\mathbf{K}\to 0,$$

$$-\mathbf{q}_3-k\mathbf{n},\mathbf{q}_1+\mathbf{q}_2+\mathbf{q}_4)|G(\mathbf{q}_1+\mathbf{q}_2)|^2. \qquad (35)$$

Substituting (9) into (35) and integrating using Ward's identity (18), we have

$$I^D(\mathbf{n}) = \frac{e^2}{2c\varepsilon}(1-n_z^2)A(\mathbf{K})k^2$$

$$\times \text{Im}\,\Sigma(k\mathbf{n})L_z \int \frac{d\mathbf{q}_\rho}{(2\pi)^2}\frac{1}{(q_\rho^2+k_0^2-k^2)^2}$$

$$\times \frac{B(|k_0+\sqrt{k^2-q_\rho^2}|)+B(|k_0-\sqrt{k^2-q_\rho^2}|)}{B(0)+B(2\sqrt{k^2-q_\rho^2})}. \qquad (36)$$

The singularity of the radiation intensity results from the diffusion pole. When one takes into account the finite size of the system, the diffusion paths of the photon are cut off at the system size, and therefore $1/K^2$ can be replaced by L_z^2 as $K\to 0$ (we assume that $L_z < L_x, L_y$). It follows from (36) that for particle energies $k_0\to k$, $\gamma\gg 1$, the main contribution to the integral over q_ρ comes from $q_\rho\to 0$. The correlation function B in (36) varies slowly when $q_\rho\to 0$ provided that $\gamma^2\gg ak_0$ (we discuss this condition in more detail in the next section); therefore, taking $q_\rho\approx 0$ (under the condition $\gamma\gg ak$) and substituting (34) and (16) into (36), we obtain

$$I^D(\omega,\vartheta) = \frac{5}{2}\frac{e^2\gamma^2}{\varepsilon c}\left(\frac{L_z}{l(\omega)}\right)^3\frac{\sin^2\vartheta}{|\cos\vartheta|}. \qquad (37)$$

Note the main features of the diffusion contribution (37). Comparing (37) with the single-scattering contribution (26), we see that $I^D/I^0\sim L_z^2/l^2\gg 1$. This means that in the wavelength range $\lambda\ll l(\lambda)\ll L_z$, the dominant radiation mechanism is diffusion. Note the strong dependence of spectral intensity on the particle energy, which also holds for the total intensity (integrated over frequencies and angles). Recall that this dependence in conventional transition radiation is logarithmic in the optical region. When $L_z\sim N$, then from (37) the radiation intensity has a strong dependence on the number of plates, $I\propto N^3$.

We now discuss a reason for the strong dependence of radiation intensity on particle energy. It is convenient to represent the background field in the form

$$A_0(\mathbf{q})\propto\frac{\delta(q_z-k_0)}{q_\rho^2+k_0^2\gamma^{-2}}. \qquad (38)$$

It follows from (38) that at relativistic energies $\gamma\gg 1$, most pseudophotons have momentum with transverse component $q_\rho\to 0$. It is easy to see from (38) that the total number of pseudophotons $N_{ps}\propto\int A_0^2(\mathbf{q})d\mathbf{q}$ is proportional to γ^2. Each pseudophoton must be scattered to be converted into a real photon. The probability of large-angle scattering of pseudophotons is low in single scattering. Therefore, only a small contigent of pseudophotons is converted into photons.

This picture changes dramatically in multiple scattering, for which almost all pseudophotons are converted into photons via multiple scattering by the plates. As the total number of pseudophotons is proportional to γ^2, the radiation intensity (total number of photons) is also proportional to γ^2.

5. COHERENCE LENGTH

It is known (see for example Ref. 1) that the coherence length (or radiation formation zone) is the distance at which

the intrinsic field of charged particle separates from the radiation field. In other words, it is the length at which the interference term becomes small. The interference term consists of expressions like $I^i\sim A_0^*(\mathbf{R},\omega)<A_r(\mathbf{R},\omega)$. Using (9) and going to the coordinate representation, one obtains the following expression for the background field:

$$A_0^*(\mathbf{R}) = \frac{2e}{\pi^2 c}e^{-ik_0z}K_0\left(\frac{k_0\rho}{\gamma}\right), \qquad (39)$$

where K_0 is the modified Bessel function (see for example Ref. 7).

Using (11), we obtain the averaged radiation potential

$$\langle A_r(\mathbf{R})\rangle = -\frac{\omega^2}{c^2}\int d\mathbf{r}A_0(\mathbf{r})\langle\varepsilon_r(\mathbf{r})G(\mathbf{R},\mathbf{r})\rangle. \qquad (40)$$

Using the impurity diagrams (15), one can represent the average in (40) in the form

$$\langle\varepsilon_r(\mathbf{r})G(\mathbf{R},\mathbf{r})\rangle = \int d\mathbf{r}_1 G_0(\mathbf{R},\mathbf{r}_1)B(\mathbf{r}_1-\mathbf{r})G(\mathbf{r}_1-\mathbf{r}). \qquad (41)$$

Using the approximations (24) for an observation point R far from the system, $R\gg r_1$, we finally obtain

$$I^i(\mathbf{R})\sim\frac{\exp(ikR-ik_0z)}{4\pi R}K_0\left(\frac{k_0\rho}{\gamma}\right)\int d\mathbf{r}\,d\mathbf{r}_1 A_0(\mathbf{r})$$

$$\times B(\mathbf{r}_1-\mathbf{r})G(\mathbf{r}_1-\mathbf{r})\exp(-ik\mathbf{n}\cdot\mathbf{r}_1). \qquad (42)$$

For our purposes, it suffices to consider only the oscillating part of (42),

$$I^i(\mathbf{R})\propto\exp(ikR-ik_0R\cos\vartheta). \qquad (43)$$

The interference term will be small when the oscillations are strong, $R(k-k_0\cos\theta)\gg 2\pi.$[1] In this case any integration will make the interference contribution negligible. Consequently, the coherence length in our case has the form

$$l_c(\vartheta) = \frac{2\pi}{|k-k_0\cos\vartheta|}. \qquad (44)$$

Now consider some special cases. For relativistic energies $k_0\to k$ and small angles $\vartheta\approx 0$, taking into account the definition of k and k_0, one finds from (44)

$$l_c\equiv l_c(0)\approx\frac{4\pi\gamma^2}{k_0}. \qquad (45)$$

For angles $\theta\approx\pi$, the coherence length has the form

$$l_c(\pi)\approx\frac{2\pi}{k+k_0}. \qquad (46)$$

As expected, the coherence length in the direction of particle motion ($\vartheta\approx 0$) is much greater than in the backward direction ($\vartheta\approx\pi$), where it is of the order of the wavelength.

Now the meaning of the condition $\gamma^2\gg ak_0$, which we used in the previous section, becomes clearer. It means that many plates must be placed at the coherence length, $l_c\gg a$, in order for multiple scattering effects of the pseudophoton to play an important role.

6. CONCLUSIONS

We have considered the diffusion contribution to the radiation intensity of a relativistic particle traversing a stack of randomly spaced plates. It was shown that for a large number of plates ($N \gg 1$), in the wavelength range $\lambda \ll l$, for angles $|\cos \vartheta| \gg (1/kl)^{1/3}$, and coherence length much greater than the plate thickness ($l_c \gg a$), the diffusion contribution is dominant. Note that the backward and forward intensities of relativistic charged particle radiation intensity are equal, whereas in a regular stack, a relativistic particle radiates mainly in the forward direction.

Note that we did not take photon absorption into account. This is correct provided that $l \ll l_{in}$, where l_{in} is the photon inelastic mean free path in the medium. In the theory of diffusive propagation of waves, weak absorption ($l \ll l_{in}$) is taken into account in the following way.[8] If the absorption is so weak that $L_z < \sqrt{l l_{in}}$, then the expression (37) remains unchanged. When $L_z > \sqrt{l l_{in}}$, one must replace L_z^2 with $l l_{in}$ in (37):

$$I^D(\omega, \vartheta) = \frac{5}{2} \frac{e^2 \gamma^2}{\varepsilon c} \frac{L_z l_{in}(\omega)}{l^2(\omega)} \frac{\sin^2 \vartheta}{|\cos \vartheta|}. \tag{47}$$

It follows from (47) that in this case the dependence of radiation intensity on the number of plates is weaker, $I \propto N$.

Note that absorption changes the frequency dependence of the spectral intensity.

Now compare the radiation considered above with Cherenkov radiation for the corresponding values of particle energy, which, however, are on opposite sides of the critical value $c/\sqrt{\varepsilon}$. The intensity of Cherenkov radiation has the form[8]

$$I^{Ch}(\omega) = \frac{e^2 \omega d}{c^2} \left(1 - \frac{c^2}{v^2 \varepsilon} \right), \tag{48}$$

where d is the path of the charged particle, which traverses a medium with dielectric constant ε. Comparing (48) with (37), we have

$$\frac{I^D(\omega)}{I^{Ch}(\omega)} \sim \frac{\gamma^2}{kl} \left(\frac{L_z}{l} \right)^2. \tag{49}$$

Note that the Cherenkov intensity is greater than the single-scattering contribution, $I^0/I^{Ch} \sim 1/kl \ll 1$.[9] From (49), the diffusion radiation, in contrast, can be stronger than the Cherenkov radiation.

[1] V. L. Ginzburg and I. M. Frank, Zh. Eksp. Teor. Fiz. **26**, 15 (1946).

[2] V. L. Ginzburg and V. N. Tsytovich, *Transition Radiation and Transition Scattering* [in Russian], Nauka, Moscow (1984); M. L. Ter-Mikaelyan, *High-Energy Electromagnetic Processes in Condensed Media*, Wiley, New York (1972).

[3] G. M. Garibian, Zh. Eksp. Teor. Fiz. **37**, 527 (1960) [Sov. Phys. JETP **10**, 372 (1960)].

[4] G. M. Garibian and C. Yang, *X-Ray Transition Radiation* [in Russian], Yerevan (1983); X. Artru, G. Yodh, and G. Menassier, Phys. Rev. D **12**, 1289 (1975).

[5] Zh. S. Gevorkian, Phys. Lett. A **162**, 187 (1992); Radiofizika **36**, 36 (1993).

[6] A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyloshinskii, *Methods of Quantum Field Theory in Statistical Physics*, Prentice-Hall, Englewood Cliffs (1969).

[7] I. S. Gradshtein and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, New York (1965).

[8] P. W. Anderson, Philos. Mag. B **52**, 505 (1985).

[9] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media*, Pergamon, New York (1984).

l mixing and dissociation of Rydberg molecules accompanying slow collisions with inert-gas atoms

G. V. Golubkov, G. K. Ivanov, E. M. Balashov,*) and M. G. Golubkov

N. N. Semenov Institute of Chemical Physics, Russian Academy of Sciences, 117977 Moscow, Russia
(Submitted 21 November 1997)
Zh. Éksp. Teor. Fiz. **114**, 100–113 (July 1998)

Processes involving a change of orbital angular momentum *l* and the dissociation of Rydberg molecules when they collide with inert-gas atoms are considered, using a method based on analyzing the terms of the interacting systems. The proposed method makes it possible to take into account the perturbations in a large group of Rydberg states that weakly penetrate into the ion core and to go beyond the limits of the two-level approximations [R. F. Stebbings and F. B. Dunning, eds., *Rydberg States of Atoms and Molecules* (Cambridge University Press, Cambridge, 1983; Mir, Moscow, 1985)] widely used in the theory of *l* mixing with the participation of Rydberg atoms. Using the Na**(nd)+Xe reaction as an example, it is demonstrated that this method gives good results in a wide range of variation of the principal quantum number *n* of the Rydberg electron ($n \approx 10$–40). Features of the *l* mixing and dissociation of Rydberg molecules are studied for the H_2^{**}+Xe system. It is shown that collisional dissociation occurs by the formation of an intermediate Rydberg state of the molecules, followed by self-decay into a dissociative continuum. © *1998 American Institute of Physics.*
[S1063-7761(98)00707-0]

1. INTRODUCTION

Elementary processes in which highly excited (Rydberg) atoms and molecules participate play an important role in atmospheric and astrophysical phenomena, and are also of interest for a number of applied areas of physics (low-temperature plasma, MHD generators, gas-laser systems, etc.). These processes are extremely diverse and are characterized, as a rule, by large cross sections and largely determine the properties of the upper atmosphere.[1] They include, most importantly, inelastic vibronic interactions, ionization, and charge exchange, which, with relative velocities of the colliding particles satisfying

$$u_c \ll 1/n \tag{1}$$

(*n* is the principal quantum number of the Rydberg state, and $\hbar = m_e = e = 1$), are characterized by local transition regions.[2–4]

However, the most efficient processes are those in which the angular momentum *l* of the Rydberg electron changes (*l* mixing), induced by slow collisions with a perturbing neutral particle. In the region in which the principal quantum number varies, $n \approx 10$–40, the total *l*-mixing cross section substantially exceeds the gas-kinetic cross section, attaining values of $\sigma \sim 10^{-12}$–10^{-11} cm^2.[1] These processes correspond to a small energy transfer into translational degrees of freedom and have no pronounced transition regions. Their treatment requires more complex dynamic models (by comparison with those used in Refs. 2–4).

This paper is devoted to the solution of this problem. The proposed approach is based on a study of the features of the potential energy surfaces of the interacting system and makes it possible to go beyond the framework of concepts concerning local transitions. It takes into account the distortion of the level system of the Rydberg states, created by the field of perturbing atom A, which is completely ignored in the existing theory.

Dynamical models used to describe processes with the participation of Rydberg atoms are covered in detail in the review article by Hickman, Olson, and Pascal.[1] We shall not discuss all possible theoretical approaches here and shall indicate only those that are currently most often and most effectively used. These include the Born or the momentum approximation,[5] which are essentially a generalization of Fermi's pseudopotential model.[6] The approximations are introduced on the basis of information concerning the scattering length *a* and the static polarizability β of the atom in the form in which it is reproduced in the expression for the $e^- - A$ scattering amplitude,[7]
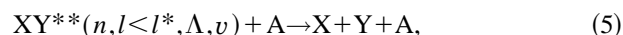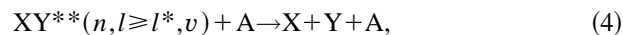
$$f_e = a + \frac{\pi}{3}\beta k, \tag{2}$$

where *k* is the momentum of the incident electron.

We shall apply the method developed below to the study of the processes of *l* mixing

$$XY^{**}(n, l<l^*, \Lambda, v) + A \rightarrow XY^{**}(n', l' \geqslant l^*, v') + A \tag{3}$$

and of collisional dissociation

$$XY^{**}(n, l \geqslant l^*, v) + A \rightarrow X + Y + A, \tag{4}$$

$$XY^{**}(n, l < l^*, \Lambda, v) + A \rightarrow X + Y + A, \tag{5}$$
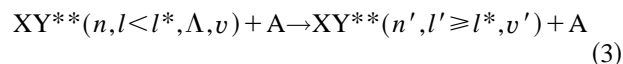
where Λ is the absolute magnitude of the projection of the electronic angular momentum onto the axis of the molecule, and *v* is the vibrational quantum number. The effective value
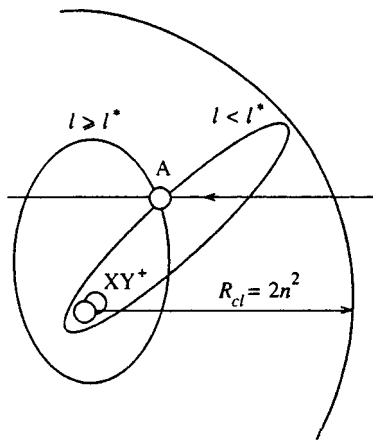
FIG. 1. Schematic image of classical trajectories of an electron and of atom A relative to ion XY$^+$. Radius $R_{cl}=2n^2$ denotes the classically resolved region of motion of the electron.

l^* of the orbital angular momentum makes it possible to divide the Rydberg configurations into two inequivalent groups of states, strongly and weakly interacting with the ion core.[2] The states of the first group are characterized by small values of the orbital angular momentum of the electron ($l<l^*$) and in classical language correspond to convex elliptical trajectories adjacent to the ion core. As is well known, they possess quantum defects that determine how much their levels deviate from the Coulomb levels. The states of the second group ($l \geqslant l^*$) correspond to trajectories far from the ion core and virtually coincide with the Coulomb states; i.e., they are characterized by high degeneracy multiplicity. The number of these states $p_n \approx n^2$ is rather high (in particular, the angular momentum l^* separating the two indicated groups of states equals $l^*=3$ for Na** and H$_2^{**}$).

Since states with large $l \geqslant l^*$ are energetically unresolvable, it is impossible to treat them separately. They interact with the entire set. The involvement of large groups of states in the interaction process is a specific property of Rydberg systems. To explain the role of these states in the dynamics of an elementary interaction event, it is necessary to study the potential-energy surface of the combined XY**+A system. When this is done, two goals are achieved simultaneously: The distortion of the levels that occurs when neutral atom A falls into the sphere of influence of Rydberg molecule XY** is found, and the mixing of states that results from the transition from certain trajectories to others is taken into account. The direction of the classical momentum of the electron is altered by the elastic scattering at perturbing atom A. A schematic diagram of this physical situation is shown in Fig. 1.

The position of the Rydberg levels depends on the distance R between the XY** and A particles. The character of their perturbation for sufficiently large R is determined by the features of the interaction of a free electron with atom A. Since the cross sections of processes (3)–(5) are large, the main contribution comes from the region of large distances, where the asymptotic method can be used to solve the problem. This makes it possible to construct the potential-energy surface of the combined quantum system XY**+A in ana-

lytical form and to study the dynamics of the processes considered here, using a formal scheme to calculate the distortion of the Rydberg terms. The subsequent analysis is carried out in the adiabatic approximation (with respect to rotation), in which the orientation of the molecule during collision is considered fixed. In this case, the potential-energy surfaces are constructed by using a vibronic basis, in which the vibrational quantum numbers are given instead of the interatomic distances in the XY** molecule. Such a basis is convenient when the dynamic behavior of the XY**+A system is being studied and most clearly reflects the specifics of the problem under consideration.

The general concepts concerning the terms or concerning the potential-energy surfaces of the X**+A and XY**+A systems, in which atom A is considered a structureless particle (as a rule, atoms of the inert gases possess this property) are discussed in Sec. 2. The dynamics of the l-mixing process are investigated in Sec. 3. Collisional dissociation of diatomic Rydberg molecules XY** is considered in Sec. 4. The last section, Sec. 5, presents the main conclusions and analyzes ways of further developing the theory.

2. POTENTIAL-ENERGY SURFACES OF THE XY**+A SYSTEM

We now pass to a consideration of the potential-energy surfaces of the combined system XY**+A corresponding to states of the Rydberg electron that strongly ($l<l^*$) and weakly ($l \geqslant l^*$) penetrate into the ion core.

Information on the behavior of these terms in the field of a perturbing neutral particle can be obtained by analyzing the equations for the level-shift operator τ. The form in which these equations are written depends on the basis of the states that is used for the unperturbed XY** molecule and the number of $|LM\rangle$ states for the scattering of the free electron e^- at particle A that are taken into account (indices L and M here denote the angular momentum of the electron relative to the scattering center and its projection in the direction of the vector **R**). Below we shall restrict ourselves to S scattering in the $e^- - A$ system, assuming that the angular moment of the electron is $L=0$. Including states with $L \neq 0$, as can be shown, does not substantially change the results obtained for the energy region under consideration ($n \geqslant 10$).

In choosing the basis of the states of the XY** molecules, one should be guided by the two limiting cases, corresponding to two different relationships between the rotational period $\tau_{\text{rot}} \sim 1/B$ (where B is the rotational constant of the XY$^+$ ion) of the molecules and the time $\tau_c \sim n^2/u_c$ for particles A with relative velocity u_c to pass through the characteristic region $\sim n^2$. For slow rotation, $\tau_{\text{rot}} \gg \tau_c$, when

$$u_c \gg Bn^2,$$

the orientation of the XY** molecules remains unchanged during collision, and the potential-energy surface must be determined for the given position of the axis of the molecular ion XY$^+$. It is this case that is studied in the present paper.

Slow vibrational motion of the ion with frequency ω_v is characterized by the similar relationship $u_c \gg \omega_v n^2$, which,

under the conditions considered here, $u_c \ll 1/n$, is equivalent to the requirement $\omega_v \ll 1/n^3$. This relationship is not satisfied for the majority of diatomic molecules (of the type H_2, N_2, O_2, CO, OH, etc.). Therefore, in constructing the terms, it is necessary to fix the vibrational states of the XY^+ ion and to go from the adiabatic to the vibronic potential surfaces of the system. In the classically resolved region of motion of a Rydberg electron, these terms for the most important case $L=0$ are determined by analogy with Ref. 2 from the following equation:

$$\tau = 2\pi \sum_{nl<l^*,\Lambda,v} \mathbf{K} \frac{\Psi^*_{nl\Lambda v}(\mathbf{R},\rho)\Psi_{nl\Lambda v}(\mathbf{R},\rho')}{E - E_{n_v l \Lambda v}} \tau$$

$$+ \sum_v k_v(R)\mathbf{K} \cot \pi \nu_v \chi_v(\rho)\chi_v(\rho')\tau, \qquad (6)$$

where $\Psi_{nl\Lambda v}(\mathbf{R},\rho)$ is the wave function for the Rydberg $nl\Lambda v$ state at point $\mathbf{R}$ (which in general is a superposition of states with different vibrational wave functions χ_v of the XY^+ ion, dependent on the interatomic coordinate ρ), k_v is the classical momentum of the Rydberg electron, $E_{n_v l \Lambda v}$ is the energy of the vibronic level, and $\nu_v = [2(E_v - E)]^{-1/2}$. The elements of the $\mathbf{K}$ scattering matrix[2] which are diagonal with respect to the vibrational quantum number,

$$K_{v,v'} = K_v \delta_{vv'}, \qquad (7)$$

are determined in Eq. (6) when the energy is $\varepsilon_v = k_v^2/2$. For S scattering and taking into account the first two leading terms, they coincide to within the sign with the expression for the $e^- - A$ scattering amplitude, Eq. (2). Since we are studying the process of l mixing in pure form, the terms in Eq. (6) that correspond to nonadiabatic vibronic coupling are left out. This makes it possible to use the ordinary representation $\Psi_{nl\Lambda v}(\mathbf{R},\rho) = \Phi_{nl\Lambda}(\mathbf{R})\chi_v(\rho)$ and to omit an analysis of the narrow regions of pseudo-crossing of terms corresponding to different values of vibrational quantum number v, i.e., to eliminate vibronic transitions from consideration (for definiteness, we restrict ourselves to the case $v=0$).

It is also assumed in writing Eq. (6) that l in isolated Rydberg states is a "good" quantum number. The sum over l in the first term of Eq. (6) corresponds to the contribution of the strongly penetrating orbitals ($l<l^*$). The second term, containing $\cot \pi \nu_v$ (whose poles reconstruct the position of the Coulomb levels) describes the contribution of the weakly penetrating orbitals. The choice of critical value l^* is determined by the smallness of the characteristic Massey parameter

$$\frac{\Delta E_{n_v l \Lambda} n^2}{u_c} \ll 1 \qquad (8)$$

with the value of $\Delta E_{n_v l \Lambda}$ that characterizes the deviation of the position of the Rydberg $n_v l \Lambda$ level from the Coulomb level.

States with $l < l^*$, as is well known, are weakly perturbed by particle A. Therefore, the positions of the Rydberg vibronic terms virtually coincide with the levels of the isolated molecule.[2] Their energies, measured from the ground state of the XY^+ ion, are defined as

$$E_{n_v l \Lambda} = E_v - \frac{1}{2(n_v - \mu_{l\Lambda v})^2}, \qquad (9)$$

where E_v is the energy of the vibrational perturbation of the XY^+ ion, and $\mu_{l\Lambda v}$ is the quantum defect of the $n_v l \Lambda$ level.

As a consequence of strong $l\Lambda$ mixing, the states with large $l \geq l^*$ form covalent terms split off from the Coulomb levels ($\mu_{l\Lambda v} = 0$) and characterized by the type of scattering of the Rydberg electron at particle A. These terms, denoted below by subscript L, in the case under consideration of $L=0$, are described by the transcendental equation

$$U_{n_v L}(R) = E_v - \frac{1}{2[n_v - \mu_{0v}(R)]^2}. \qquad (10)$$

The quantity μ_{0v} has the meaning of the quantum Lv-defect of the level induced by the $e^- - A$ interaction and equals

$$\mu_{0v}(R) = -\frac{1}{\pi} \arctan(k_v K_v(R)). \qquad (11)$$

It is important to point out that, with large values of $l \geq l^*$, the electron density close to the ion core is small because of the presence of the orbital barrier. Therefore, the strong field of the core, like its orientation in the $XY^{**}+A$ system, has no appreciable effect on these states and can be neglected in first approximation.

Equation (6) completely determines the potential-energy surfaces that describe the groups of states indicated above and also makes it possible to calculate the interaction between them. Actually, using Eq. (6) as a basis for setting up the characteristic (secular) equation for the eigenvalues of the electron energy, it is easy to obtain

$$V^2_{n_v l \Lambda, n_v l' \Lambda'} = 2\pi K_v^2 |\Phi_{n_v l \Lambda}(\mathbf{R})\Phi_{n_v l' \Lambda'}(\mathbf{R})|^2, \qquad (12)$$

$$V^2_{n_v l \Lambda v, n_v 0 v} = \frac{B_{l\Lambda 0 v} k_v K_v^2 \cos^2 \pi \nu_v}{\pi \nu_v^3}. \qquad (13)$$

Equations (12) and (13) describe, respectively, the interactions between the states within the $l<l^*$ group and between the $l<l^*$ and $l \geq l^*$ groups. Moreover, the former corresponds to the case of two interacting parallel terms:

$$U_{n_v l \Lambda} = E_{n_v l \Lambda} + 2\pi K_v |\Phi_{n_v l \Lambda}(\mathbf{R})|^2,$$

$$U_{n_v l' \Lambda'} = E_{n_v l' \Lambda'} + 2\pi K_v |\Phi_{n_v l' \Lambda'}(\mathbf{R})|^2. \qquad (14)$$

The latter corresponds to the interaction of two intersecting l and L terms. The quantity $B_{l\Lambda 0 v}$ characterizes the probability of finding a Rydberg electron close to the perturbing atom A and equals

$$B_{l\Lambda 0 v} = 2\pi |\Phi_{n_v l \Lambda}(\mathbf{R})|^2. \qquad (15)$$

When S scattering in the $e^- - A$ system is taken into account, the group with $l \geq l^*$ is represented, essentially, by one state, which is the consequence of intense $l\Lambda$ mixing. It should also be pointed out that, in the case of atoms (X^{**} instead of XY^{**}), when the electron energy is independent of Λ, the sum over Λ in Eq. (6) for the states $l<l^*$ is cut short. As a result, a term appears that is independent of the direction of the vector $\mathbf{R}$; i.e. we essentially operate with one
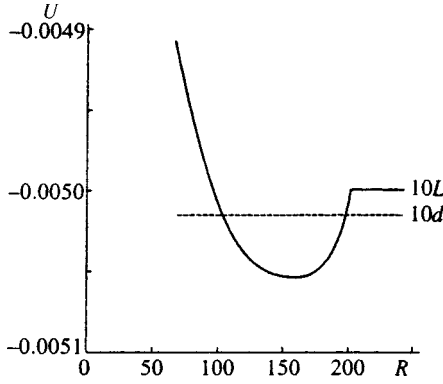
FIG. 2. Diabatic terms of the Na**(nd)+Xe system. The dashed curve shows the Rydberg $10d$ term, and the solid curve shows the covalent $10(L=0)$ term.

state (instead of $2l+1$ for the given index l). This means that complete mixing of the Λ components of the nl orbitals of the Rydberg electron occurs at a preliminary stage of the collision process.

The interaction between such states in the X**+A system is written similarly to Eqs. (12) and (13), with $\Phi_{nl\Lambda}(\mathbf{R})$ replaced by the function $\widetilde{\Phi}_{nl}(R)$, which is independent of the direction of the vector $\mathbf{R}$ and which can be formally obtained by using function $\Phi_{nl\Lambda}(R)$, where the quantization axis is directed along the vector $\mathbf{R}$. The use of such functions substantially simplifies the study of the dynamic behavior of the system.

3. DYNAMICS OF THE MIXING

The investigation of the terms of the X**+A system is a necessary stage of the study of the dynamics of elementary processes. This stage is completely ignored in existing theoretical approaches. Thus, the approximation of two undistorted states with an interaction of the type of Eq. (12) is ordinarily used to describe l mixing, even though it is actually necessary to take into account n^2 Rydberg states.

A theory based on term concepts gives an effective method of solving this complex quantum problem. The principal novelty that it introduces into the theory of collisions with the participation of highly excited atoms is associated with taking into account the distortion of the levels of the Rydberg electron and the mixing of large groups of degenerate states when X** and A particles interact.

Using the experimentally and theoretically[8] well-studied system Na**(nd)+Xe as an example (see Fig. 2), we can see that the relative position of the terms substantially depends on the distance R between the colliding particles. In order to check the validity of the method proposed below for calculating the l-mixing cross section in Rydberg molecules, we have carried out a preliminary analysis of this system. In terms of an approach using complete information concerning the terms of the X**+A system, the process of l mixing can be regarded as a dynamic transition from a group of states mixed with regard to projections m but with a fixed value of l (with $l<l^*$) to a group of states with $l\geq l^*$, strongly mixed

in m and in l. The interaction of the indicated groups of states is given by

$$V_{nlL}=\frac{\sqrt{2l+1}}{\pi\nu_l^3}\frac{K_0}{R}\sin\Theta_{nl}(R)\cos\pi\nu_l,\qquad(16)$$

which is independent of the direction of the vector $\mathbf{R}$ because of the strong mixing of the m components of the nl level. The argument of the sine of $\Theta_{nl}(R)$ in Eq. (16) is formed from the radial part of the wave function $\Phi_{nl\Lambda=0}(R)$ described in the quasi-classical representation ($\nu_l=n-\mu_l$ is the effective principal quantum number).

We can use the approximation of straight-line trajectories to compute the transition amplitude A_{nlL}, restricting ourselves to first-order perturbation theory in the interaction V_{nlL}:[9]

$$A_{nlL}(b)=\int_{-\infty}^{\infty}V_{nlL}(R(t))\exp\left[i\int_0^t\omega_{nlL}(R(t'))dt'\right]dt,$$

$$\omega_{nlL}(R)=E_{nl}-U_{nL}(R),\quad R(t)=\sqrt{b^2+u_c^2t^2}.\qquad(17)$$

The condition for perturbation theory to be applicable, $|A_{nlL}|^2\ll1$, breaks down in the region of small impact parameters b, which have an insignificant effect on the cross section of the process. Short-range interaction of atom A with the ion core is neglected in this case. According to Eqs. (10) and (11), it is also necessary to introduce an additional limitation into the relative velocity u_c of the colliding X** and A atoms, i.e.,

$$u_c>\left(\frac{1+2\mu_l}{M_cn^3}\right)^{1/2},\qquad(18)$$

which means that the kinetic energy of the particles exceeds the barrier height formed by the covalent L term, and that there are no branch points of the trajectories (M_c is the reduced mass of the interacting atoms). For the given velocity u_c (or the temperature T_c of the colliding atoms), the condition given by inequality (18) is, in essence, a lower bound on the region of variation of principal quantum number n. Taking into account the limitations given here, the transition amplitude of Eqs. (17) takes the following form:

$$A_{nlL}(b)=\sqrt{2(2l+1)}\,\frac{\cos\pi\nu_l}{\pi\nu_l^3u_c}$$

$$\times\int_0^{z_0}\frac{dz}{R(z)}K_0(z)\sin\Theta_{nl}(z)\cos\Omega_{nlL}(z,b),$$

$$\Omega_{nlL}(z,b)=\frac{1}{u_c}\int_0^{z(b)}\omega_{nlL}(R(z'))dz',\qquad(19)$$

with the limits of integration

$$z_0=2n^2,\quad z(b)=\sqrt{z_0^2-b^2}.$$

We write the cross section of the l-mixing process as[10]

$$\sigma_{nlL}=\pi g_lb_0^2|A_{nlL}(b_0)|^2+2\pi g_l\int_{b_0}^{\infty}b\,db|A_{nlL}(b)|^2,$$
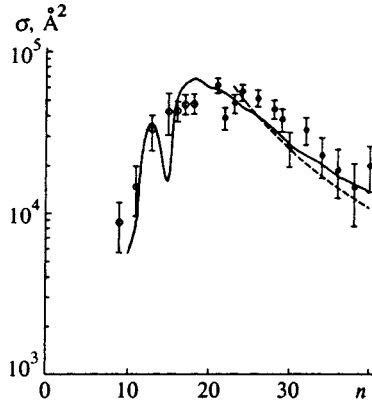
$$(20)$$

FIG. 3. Comparison of the theoretical and experimental *l*-mixing cross sections of the *nd* levels of Na** when it collides with Xe atoms. The solid curve shows the calculation from Eqs. (19) and (20) for $u_c = 2 \times 10^{-4}$, the dashed curve shows the results of the calculation of Ref. 8, ● shows experimental values from Ref. 13, and ○ shows experimental values from Ref. 14.

where $g_l = 1/(2l+2)$ is the statistical weight of the initial state. The choice of the impact parameter b_0 that separates the region of integration into the inner, $b \leqslant b_0$ (strong coupling) region and the outer $b > b_0$ (weak coupling) region is, generally speaking, arbitrary. Actually, in the inner region, where perturbation theory is not applicable, the expression $W = |A|^2 g_l$ strongly oscillates as a function of b and can appreciably exceed unity at individual points. However, starting with a certain value b_0, the probability W strictly is less than 1; $W(b > b_0) < 1$ in the outer region of variation of b, which is a necessary criterion for choosing parameter b_0. A detailed discussion of this question can be found, for example, in Ref. 11. Below, to estimate the cross sections of processes (3)–(5), we restrict ourselves to the condition $W(b_0) = 1$.

Equation (19) makes it possible to analyze how the cross section given by Eq. (20) depends on principal quantum number n. For small n, the cross section must increase because of the increase of the scale factor $\propto n^4$, proportional to the square of the geometrical size of the Rydberg particle. For large n, ω_{nlL} becomes small and depends weakly on R. It can be shown that, in the limit $\omega_{nlL} = 0$, Eq. (20) transforms to the well-known result[12]

$$\sigma_{nlL} = 2\pi \frac{a^2}{n^3 u_c^2},$$

which is closely obeyed in the momentum and Born approximations.[1] Consequently, the n dependence of the cross section given by Eq. (20) must have a characteristic bell shape with a pronounced maximum.

The above remarks are illustrated in Fig. 3, which shows the results of a calculation of the *l*-mixing cross section in the system Na**(*nd*) + Xe, using Eqs. (19) and (20) with $L = 0$, and gives a comparison with experiment.[13,14] The calculation is carried out for a relative velocity of the colliding particles of $u_c = 2 \times 10^{-4}$ (corresponding to a temperature of $T_c \approx 300$ K), with parameters $\mu_d = 0.015$,[15] $a = -6.0$, and

$\beta = 27.06$.[16] In the calculations here and below, we have used the long-wavelength expansion of the elements of the **K** matrix, written in the form[3]

$$K_0 = a + \frac{\pi\beta}{3} k + \left(\frac{4}{3}\alpha\beta \ln k + \gamma\right)k^2, \quad (21)$$

where the parameter for the Xe atom is $\gamma = -210$. The filled circles on the right-hand wing of the curve correspond to room temperature, $T_c = 300$ K.[13] On the left-hand wing of the curve, the open circles show the results obtained in Ref. 14 at a temperature of $T_c = 430$ K. The theoretical estimate agrees well with the features of the experimental dependence. The presence of a dip on the left-hand wing is evidently associated with the fact that the region of transitions that makes the main contribution to the cross section of the process comes at a node of the radial wave function, whose position depends on the quantum defect of the level. The dashed curves on same figure show the dependence obtained in the momentum approximation.[8]

However, it should be pointed out that, for large $n \gg (M_c u_c^2)^{-1/3}$, along with quasi-elastic processes, inelastic processes accompanied by a change of the principal quantum number n must also occur. A group of Rydberg states will participate here, from which it follows that the cross section of the process must be calculated from the corresponding summation. As a result, the n dependence at the periphery of the right-hand wing is $\propto 1/n$;[17] i.e., the total cross section here must significantly exceed that calculated from Eqs. (19) and (20). For example, for a relative velocity of u_c corresponding to the thermal temperature, this will occur starting with $n \approx 40$, as is clearly demonstrated in Fig. 3.

Rydberg molecules XY** are a more complicated quantum object of study. Actually, the strong field of diatomic molecules, unlike that of X** atoms, is not spherically symmetric and possesses only axial symmetry. Complete mixing of the m components of the electronic angular momentum l also does not occur, since states with different projections of Λ onto the axis of the molecule are energetically different.

It can be shown that, for strongly penetrating orbitals $(l < l*)$, the angular dependence of the wave functions of a Rydberg electron is rather weak and does not substantially affect the dynamics of the process. The main physical factors here are the interaction modulations created by quasi-classical oscillations of the radial wave functions and the energy transferred into the translational degrees of freedom (i.e., the quantum defects of the states combined in the collisional transitions). Therefore, it is sufficient to use the simple equations (19) and (20) to analyze the structure of the *l*-mixing cross sections in Rydberg molecules.

As a specific application, let us consider the process

$$\text{H}_2^{**}(np\Lambda, v=0) + \text{Xe} \rightarrow \text{H}_2^{**}(n, l \geqslant 3, v=0) + \text{Xe}, \quad (22)$$
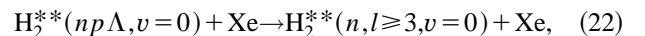
where Λ takes the values 0 and 1 (corresponding to the σ and π states, the adiabatic quantum defects of which are equal, and $\mu_\sigma = 0.191$ and $\mu_\pi = -0.078$, Ref. 18). The optical $np\Lambda$ series of the H_2^{**} molecule are not predissociated and possess radiative lifetimes $r_{\text{rad}} \geqslant 10^{-6}$ sec, much greater than the characteristic collision times of $\tau_c \sim 10^{-11}$ sec. Figure 4 shows the dependence of the cross section of process (22) on
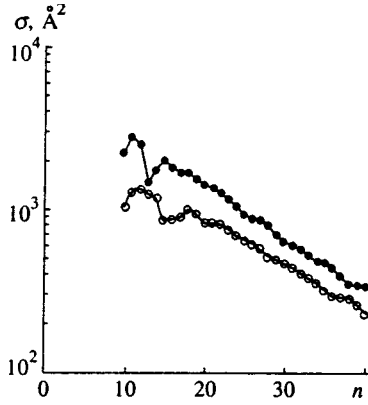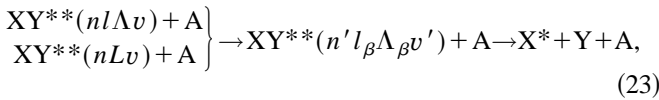
FIG. 4. Cross section of the $H_2^{**}(np\Lambda, v=0)+Xe \to H_2^{**}(n,l\geq 3, v=0)$ $+Xe$ process, calculated from Eqs. (19) and (20) for $u_c = 9 \times 10^{-4}$ ($\bullet$—$p\pi$ state, $\circ$—$p\sigma$ state).

the principal quantum number n, calculated for a relative velocity of $u_c \approx 9 \times 10^{-4}$. The resulting curves have the typical shape characteristic of atomic systems.

4. COLLISIONAL DISSOCIATION OF RYDBERG MOLECULES

We now turn to a consideration of the collisional dissociation of XY** molecules in highly excited states that are not predissociated in the absence of atom A. They include two types of excitation, i.e., states with a small orbital angular momentum $l < l^*$ and states with $l \geq l^*$ that weakly interact with the ion core, correlating with covalent L configurations. The former, as a rule, are populated during resonance photoabsorption and are characterized by radiative lifetimes τ_{rad}. They decay because of l mixing between states corresponding to strongly penetrating orbitals ($l < l^*$). The latter are formed during threshold excitation by electron impact and do not predissociate, since they are restrained by a high orbital barrier ($l \geq l^*$). Here predissociation occurs by reverse $L \to l$ transitions. In both cases, collisional dissociation occurs by a two-stage "state-to-state" mechanism

$$\left. \begin{array}{l} XY^{**}(nl\Lambda v)+A \\ XY^{**}(nLv)+A \end{array} \right\} \to XY^{**}(n'l_\beta\Lambda_\beta v')+A \to X^*+Y+A,$$
(23)

where the first stage is induced by interaction with atom A and is accompanied by a transition to the predissociated β state. The subsequent decay occurs in a time that substantially exceeds τ_c, when the colliding particles no longer actually interact. Moreover, the first stage can occur either as a result of a transition with no change of the initial vibrational state v or with excitation of the ion core. Note that for $n \approx 10-40$ the quasi-elastic transition cross sections are at least an order of magnitude larger than the inelastic cross sections. Since predissociation at the second stage of the process occurs with a probability equal to unity, the quasi-elastic transitions ($v = v'$) must dominate in this region of variation of n.

For optically excited XY** molecules, dissociation results from a direct transition from a stable state of the group ($l < l^*$) into a predissociated $nl_\beta\Lambda_\beta v$ state belonging to the

same group. The collisional dissociation cross section can be calculated here from Eq. (12) with a transition amplitude

$$A_{n_v l\Lambda, n_v l_\beta \Lambda_\beta}(b) = 2\int_0^\infty dt \ V_{n_v l\Lambda, n_v l_\beta \Lambda_\beta}(R(t))\cos(\omega_{l\Lambda, l_\beta \Lambda_\beta}t),$$
(24)

$$\omega_{l\Lambda v, l_\beta\Lambda_\beta v} = (\mu_{l\Lambda v} - \mu_{l_\beta\Lambda_\beta v})n_v^{-3},$$
(25)

in which the interaction V of two parallel terms, taking into account the remarks in Sec. 3 relating to the angular dependence of the wave functions, has the form

$$V_{n_v l\Lambda, n_v l_\beta \Lambda_\beta}$$

$$= \sqrt{\frac{(2l+1)(2l_\beta+1)}{2}} \frac{K_v \ \sin \ \Theta_{n_v l\Lambda}(R)\sin \ \Theta_{n_v l_\beta\Lambda_\beta}(R)}{k_v R^2(\pi\nu_{n_v l\Lambda}\nu_{n_v l_\beta\Lambda_\beta})^{3/2}}.$$
(26)

The distortion of the nl terms can be neglected (as in the theoretical approaches used by Ref. 11).

If a Rydberg electron is formed with large orbital moments ($l \geq l^*$), Eqs. (19) and (20) can be used to estimate the partial cross section of the reaction, but the latter equation must be supplemented with the additional factor
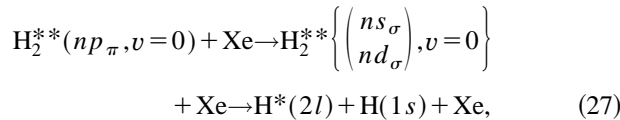
$$g_l(n_v) = \frac{1}{n_v^2}.$$

This is because the given process involves a transition to the predissociated $nl_\beta\Lambda_\beta v$ state from a group of states strongly mixed in terms of l and m, in which each separate l segment is represented with statistical weight $g_l(n_v)$. As a result, the cross section is reduced, even though it remains substantially greater than the gas-kinetic cross section. To find the total cross section, Eq. (20) must be summed over all possible predissociated Rydberg l_β configurations.
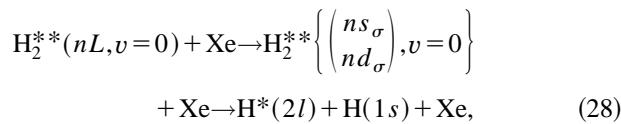
For threshold excitation of molecules by electron impact, it is impossible to fix states with a definite value of l, since this involves forming a group of states with a rather smooth distribution function whose maximum comes at $l \sim n^{1/2}$.[19,20] The total cross section must therefore be the corresponding convolution.

Turning to the hydrogen molecule, we briefly recall the most characteristic features of the behavior of its terms close to the bottom of the ion potential. The H_2^{**} molecule, as is well known, possesses one low-lying doubly excited electronic state $^1\Sigma_g^+(2p\sigma_u)^2$, whose potential curve (see Fig. 5) intersects the ionic term $U_i(\xi)$ in the neighborhood of the point $\xi = 2.65$ a.u., located close to the right-hand classical turning point of the first excited vibrational level $v = 1$. The two Rydberg $ns\sigma$ and $nd\sigma$ series with $n \geq 4$ are liable to decay in this energy region because of configuration mixing with the dissociative continuum.[21] Thus, of the nine possible (s, p, and d) dynamically active states, only two are predissociated.
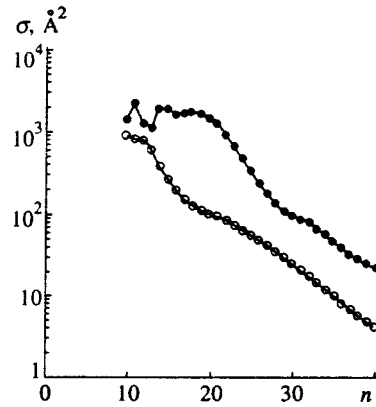
Figure 6 shows the n dependence of the partial cross sections of collisional dissociation of the first type:
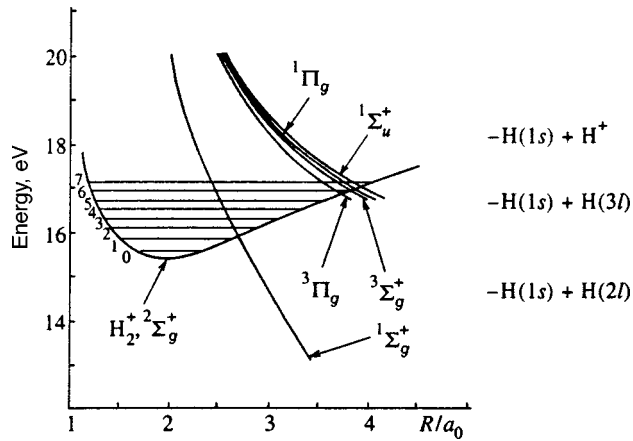
FIG. 5. Potential curves of the H_2^{**} molecule.[21]

FIG. 7. Partial cross sections of reaction (28), calculated from Eqs. (19) and (20) for $u_c = 9 \times 10^{-4}$ (●—intermediate $d\sigma$ state, ○—$s\sigma$ state).

$$H_2^{**}(np_\pi, v=0) + Xe \rightarrow H_2^{**}\left\{ \begin{pmatrix} ns_\sigma \\ nd_\sigma \end{pmatrix}, v=0 \right\}$$
$$+ Xe \rightarrow H^*(2l) + H(1s) + Xe, \qquad (27)$$

corresponding to transitions to the intermediate Rydberg $ns\sigma$ and $nd\sigma$ states. The calculation was carried out using Eqs. (24)–(26) with quantum defects $\mu_{s\sigma} = -0.12$ and $\mu_{d\sigma} = 0.022$,[22] for a relative velocity of the colliding particles of $u_c = 9 \times 10^{-4}$. Figure 7 also shows the corresponding curves for a reaction of the second type:

$$H_2^{**}(nL, v=0) + Xe \rightarrow H_2^{**}\left\{ \begin{pmatrix} ns_\sigma \\ nd_\sigma \end{pmatrix}, v=0 \right\}$$
$$+ Xe \rightarrow H^*(2l) + H(1s) + Xe, \qquad (28)$$

in which the hydrogen molecules are excited by electron impact. It can be seen that, in the region of variation of n considered here, the partial cross sections of reaction (27) are an order of magnitude less than those of reaction (28); i.e., the decay process of H_2^{**} molecules formed during electron impact is preferred. In both cases, the channel involving formation of an intermediate $nd\sigma$ configuration predominates.
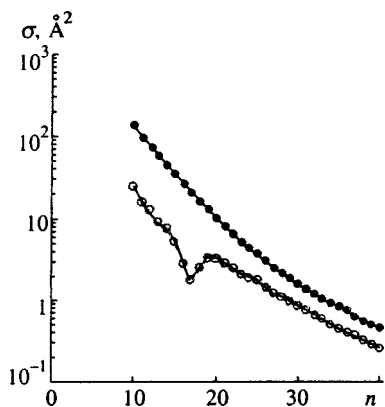
FIG. 6. Partial cross sections of reaction (27), calculated from Eqs. (24)–(26) for $u_c = 9 \times 10^{-4}$ (●—intermediate $d\sigma$ state, ○—$s\sigma$ state).

5. CONCLUSION

In conclusion, we present some results of the study that has been carried out. Most importantly, we should point out that, unlike the momentum and Born approximations, the proposed method is based on an analysis of the behavior of the potential energy surfaces of the combined system and makes it possible to describe l mixing, process (3), over a rather wide region of variation of n. The latter is especially important for small n, where the indicated approximations are not applicable.

The next important advantage of the approach developed here is that it can be used to describe collisional processes with the participation of Rydberg molecules, where l mixing can be accompanied by dissociation of the molecules.

Concerning further development of the theory, it is necessary to distinguish the following avenues of research. They include, first, the study of effects caused by the presence of strong nonadiabatic coupling with rotation, which is especially substantial for hydrogen-containing molecules. When $u_c \propto Bn^2$, this can result in a sharp irregular dependence of the reaction cross section on the initial excitation energy of the XY^{**} molecule (or on the principal quantum number n), the physical cause of which is associated with the specific behavior of the corresponding quantum defects $\mu(n)$.[23] The inclusion of nonadiabatic coupling with rotation in the formal scheme requires that the theory be generalized, with the introduction of rovibronic potential-energy surfaces of the combined system, which have not been discussed in the literature. The appearance of a stroboscopic effect, for example, can be expected here.[24] Moreover, a more detailed analysis of the behavior of the cross sections for large n, taking into account all possible channels of the motion, is necessary for a deep understanding of the nature of l mixing. Second, along with neutral perturbing atoms A, it is expedient to consider molecules that possess positive electron affinity and are capable of forming ionic configurations on the potential-energy surface of the combined system.[4]

*)E-mail: balashov@center.chph.ras.ru

[1] R. F. Stebbings and F. B. Dunning, eds., *Rydberg States of Atoms and Molecules* (Cambridge University Press, Cambridge, 1983; Mir, Moscow, 1985).

[2] G. V. Golubkov and G. K. Ivanov, Z. Phys. A **319**, 17 (1984).

[3] E. M. Balashov, G. V. Golubkov, and G. K. Ivanov, Zh. Éksp. Teor. Fiz. **103**, 374 (1993) [JETP **76**, 200 (1993)].

[4] G. V. Golubkov, G. K. Ivanov, and E. M. Balashov, Khim. Fiz. **14**, No. 8, 38 (1995).

[5] Matsuzava, J. Chem. Phys. **55**, 2685 (1971); **58**, 2678 (1971); Phys. Rev. A **18**, 1396 (1978).

[6] E. Fermi, Nuovo Cimento **11**, 157 (1934).

[7] T. F. O'Malley, Phys. Rev. **130**, 1963 (1963); **134**, 1198 (1963).

[8] L. Petitjean and F. Gounand, Phys. Rev. A **30**, 2946 (1984).

[9] E. E. Nikitin and S. Ya. Umanskii, *Theory of Slow Atomic Collisions* (Springer-Verlag, New York, 1984; Atomizdat, Moscow, 1979).

[10] J. I. Gerstein, Phys. Rev. A **14**, 1354 (1976).

[11] V. S. Lebedev and I. I. Fabrikant, Phys. Rev. A **54**, 2888 (1996).

[12] A. Omont, J. Phys. (Paris) **38**, 1343 (1977).

[13] M. Chapelet, J. Boulmer, J. C. Gauthier, and J. F. Delpech, J. Phys. B **15**, 3455 (1982).

[14] R. Karchu, T. F. Gallaher, F. Gounand, K. A. Safinya, and W. Sandner, Phys. Rev. A **27**, 795 (1983).

[15] C. F. Dybko, M. N. Efimenko, V. A. Efimov, S. V. Podios, Kvant. Elektron. (Moscow) **22**, 946 (1995) [Quantum Electron. **25**, 914 (1995)].

[16] B. M. Smirnov and A. A. Radtsig, *Handbook on Atomic and Molecular Physics* (Atomizdat, Moscow, 1980).

[17] I. L. Beigman and V. S. Lebedev, Phys. Rep. **250**, 95 (1995).

[18] G. Herzberg and Ch. Jungen, J. Mol. Spectrosc. **41**, 425 (1972).

[19] G. F. Drukarev, Zh. Éksp. Teor. Fiz. **83**, 946 (1982) [Sov. Phys. JETP **56**, 532 (1982)].

[20] A. R. P. Rau, J. Phys. B **17**, L75 (1984).

[21] G. V. Golubkov, G. K. Ivanov, and M. G. Golubkov, Khim. Fiz. **15**, 36 (1996).

[22] H. Takagi and H. Nakamura, Phys. Rev. A **27**, 691 (1983).

[23] G. V. Golubkov, G. K. Ivanov, and I. E. Cherlina, Opt. Spektrosk. **54**, 427 (1983) [Opt. Spectrosc. **54**, 251 (1983)].

[24] G. V. Golubkov, G. K. Ivanov, and E. M. Balashov, Opt. Spektrosk. **80**, 33 (1996) [Opt. Spectrosc. **80**, 27 (1996)].

Photoprocesses on a surface of nanoporous quartz under resonant laser radiation

Yu. A. Bykovskiĭ, G. E. Kotkovskiĭ, M. B. Kuznetsov, and A. A. Chistyakov

Moscow State Institute of Engineering Physics, 115409 Moscow, Russia

E. N. Lopukhina

"Polyus" Research Institute, 117342 Moscow, Russia

The photoexcitation of iodine molecules on surfaces of solid (nonporous) and nanoporous quartz by resonant laser radiation in the visible region has been studied. We have detected and studied the high-energy photodesorption of iodine molecules with a translational energy of 1.4 to 1.8 eV from nanoporous quartz surfaces at an exciting photon energy ranging between 1.9 and 2.3 eV, as well as the nonequilibrium surface dissociation of molecules. Unlike the photoprocesses, which are observed only on the surface of nanoporous quartz, the thermal desorption of I_2 molecules with a considerably lower kinetic energy has also been detected on the surface of solid quartz. We have suggested a physical mechanism of photodesorption, under which electronic excitation of an iodine molecule in the confined volume of a nanopore is accompanied by a Franck–Condon electronic transition of a molecule-surface complex to a state with a higher potential energy and subsequent release of this energy in the form of kinetic energy. It has been concluded that photoprocesses on a nanostructured surface are radically different from ordinary surface photoprocesses. © *1998 American Institute of Physics.*
[S1063-7761(98)00807-5]

1. INTRODUCTION

The current interest in photoprocesses on surfaces of condensed materials under laser radiation has stimulated intense research in photophysical processes and nonlinear optical effects, including those in nanostructured materials on developed surfaces with small characteristic dimensions.[1,2] Most experiments on photoprocesses in nanoporous materials have dealt with either the photoluminescence of such materials[3] or the diffusion and sorption of molecules driven by laser radiation (Ref. 4, p. 5). However, the small pore sizes (20–100 Å), well-developed surfaces, and other features of nanomaterials can be manifested in elementary photophysical and photochemical processes on their surfaces, such as photodesorption, photodissociation, light-induced surface diffusion, etc.

In the present work, we have used time-of-flight mass spectroscopy, laser-induced luminescence, and absorption spectroscopy[5] to study the effect of nanosecond resonant laser pulses in the 532–640 nm wavelength range on submonolayer coatings of molecular iodine submonolayers formed on surfaces of nanoporous and solid (nonporous) quartz.

2. EXPERIMENTAL APPARATUS AND TECHNIQUES

The experimental techniques included measurements of laser-induced luminescence, absorption spectroscopy, and laser-driven time-of-flight mass spectroscopy. The mass-spectroscopic system consisted of a dynamic time-of-flight mass spectrometer and a powerful YAG:Nd^{3+} laser operating in a single-mode regime with a pulse energy of 120 mJ

($\lambda = 532$ nm) and a pulse duration of 10 ns. Radiation with $\lambda = 532$ nm was used to pump a tunable solid-state dye laser with a tuning range 565–640 nm and a pulse energy of 50 mJ. The power density on the sample was varied in the range $(0.5 – 15) \times 10^6$ W/cm^2. The sample was placed in a vacuum chamber on a stage cooled by liquid nitrogen. The working pressure in the chamber during experiments was 10^{-8} Torr, and the base pressure was 4×10^{-9} Torr. Given a time-of-flight base $l = 31$ cm, we could record time-of-flight spectra, i.e., the amplitudes of mass peaks as functions of the time for the direct flight of the products to the ion source of the mass spectrometer. A specific feature of our technique is that time-of-flight spectra are recorded not in the conventional mode, where the arrival of particles of a single mass at the detector is recorded as a function of time, but the full mass spectrum is recorded as a function of the delay with respect to the laser pulse and then processed on a computer. This technique was employed because the maximum operating frequency of the mass spectrometer was 10 kHz, and time-of-flight spectra of high-energy particles with durations of 50–100 μs could not be recorded in the mode of time-resolved detection of a single mass.

The quartz (α-quartzoid) samples had pore sizes of 20–100 Å and a porosity parameter $V_{pore}/V_{bulk\ SiO_2} = 0.24$. Before experiments, the samples were outgassed in vacuum for a long time, and the process was monitored using the mass spectrometer. In order to measure the content of impurities and sorbed molecules, samples cooled to 77 K were exposed to radiation from an optical parametric oscillator ($\lambda = 2.82$, 2.65, and 3.2 μm) in the absorption band of pure quartz,

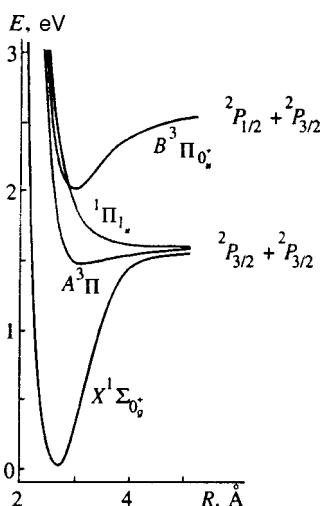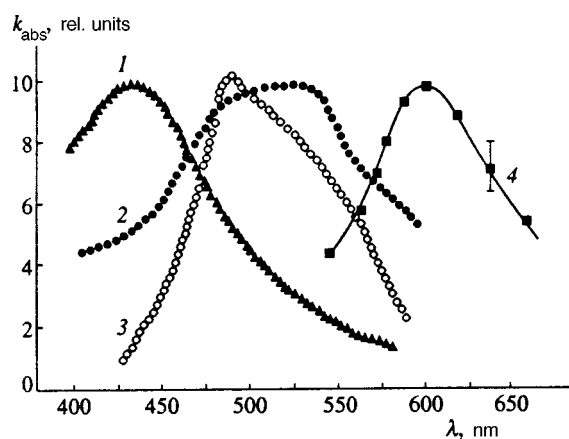FIG. 1. Potential-energy curves of a free iodine molecule.

FIG. 2. Absorption spectra (in rel. units) of chemisorbed iodine molecules in nanoporous quartz (*1*), physisorbed iodine molecules in nanoporous quartz (*2*), gaseous iodine (*3*), and a molecular crystal of iodine (*4*).

which was powerful enough to cause ablation of the material. The mass and time-of-flight spectra recorded did not show the presence of impurities and residual gas molecules at the level of the absolute detection limit of the mass-spectrometric system ($5 \times 10^4 \, cm^{-3}$).

Iodine submonolayers were produced on the quartz surface inside the vacuum chamber using a high-vacuum injection valve. The iodine coverage of the surface was less than 0.5 of a monolayer and was monitored using both the laser-driven thermal desorption technique and measurements of the absorption coefficient of samples of nanoporous quartz ($30 \, cm^{-1}$).

Our measurements indicate that iodine molecules sorbed in nanopores are of two types, namely chemisorbed molecules, which are tightly bound to the surface, and physisorbed ones, which readily leave the sample.

The adsorption energy of physisorbed iodine molecules was estimated from the desorption rate, assuming that it is an Arrhenius function of temperature:

$$\frac{dN}{dt} = N\omega \exp\left(-\frac{E_{ads}}{kT}\right),$$

where N is the number of adsorbed molecules and ω is the vibrational frequency of a molecule on the surface.

It follows from our measurements that the adsorption energy of physisorbed iodine molecules on a nanoporous quartz surface can be up to 0.7 eV, which means that their binding to the surface is fairly tight. The adsorption energy of chemisorbed molecules is even higher.

3. OPTICAL SPECTRA OF IODINE MOLECULES ADSORBED ON NANOPOROUS QUARTZ SAMPLES

Figure 1 shows the potential-energy curves of a free iodine molecule. Transitions between them are responsible for the absorption in the visible spectral range (Fig. V-21 on p. 228 in Ref. 6). The main photophysical process in free iodine molecules (Ref. 6, p. 225) is dissociation, which has a direct mechanism (direct photolysis) with a characteristic time of 5×10^{-12} s if a molecule is excited to the $^3\Pi_{0_u^+}$ state at λ

<499 nm. If the photon energy is below the dissociation threshold, a slower predissociation mechanism involving the $^1\Pi_{1_u}$ state may be activated.

The spectral and photophysical properties of iodine molecules adsorbed in pores are notably different. Figure 2 shows the absorption spectra of physi- and chemisorbed molecules in nanoporous quartz, gaseous iodine, and a molecular crystal of iodine.

It is remarkable that the peak of the absorption spectrum of the chemisorbed molecules has undergone a shift toward shorter wavelengths relative to the spectrum of free molecules, which is not typical of the changes in the state of aggregation from free molecules to an adsorbate and then to a crystal. This shift can be due to the increase in the internuclear distance in the molecule when a charge-transfer complex is formed and the molecule-surface binding is tightened. An important point is that the FWHM's in the absorption spectrum of both chemi- and physisorbed iodine molecules are larger than the widths not only for the free molecule, but also for the molecular crystal of iodine. In addition, according to the line shapes, there is a good reason to suppose that in the case of physisorbed molecules, which absorb at shorter wavelengths, as well as in the case of chemisorbed molecules, the binding to the surface is stronger and the internuclear distance is larger. An analysis of absorption spectra of physisorbed I_2 molecules in the long-wavelength range in comparison with the spectra of gaseous iodine indicates that the most probable cause of this broadening for some of the molecules is a decrease in the gap between the higher and lower terms along the energy scale owing to the negative correction to the energy levels characteristic of the Stark broadening due to a higher state of aggregation of the material. In this case, one should expect a shift of the dissociation edge of the $^3\Pi_{0_u^+}$ term toward longer wavelengths by more than 30 nm, which should lead to an onset of direct photolysis at $\lambda = 530$ nm (instead of 499 nm for gaseous I_2).

An estimate of the radiationless energy relaxation time[7] based on our measurements of absorption and fluorescence spectra and the fluorescence quantum efficiency ($B_{fl} = 5 \times 10^{-5}$) for samples of nanoporous quartz coated with io-
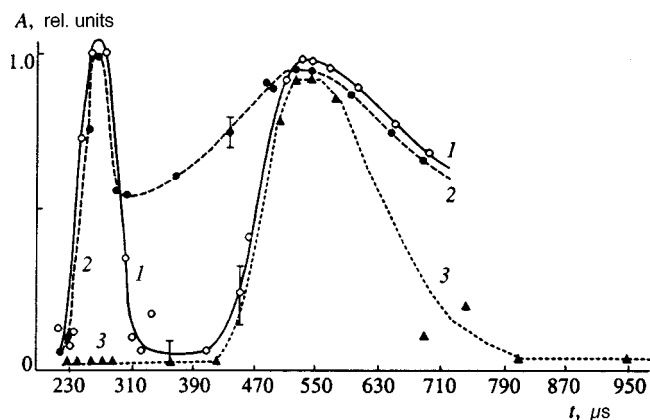
FIG. 3. Time-of-flight spectra under laser action with $\lambda = 532$ nm and $q = 1.5 \times 10^7$ W/cm^2: *1*) — iodine molecules from a nanoporous quartz surface; *2*) — iodine atoms from a nanoporous quartz surface; *3*) — iodine molecules from a solid (nonporous) quartz surface. The curves are only visual guides.
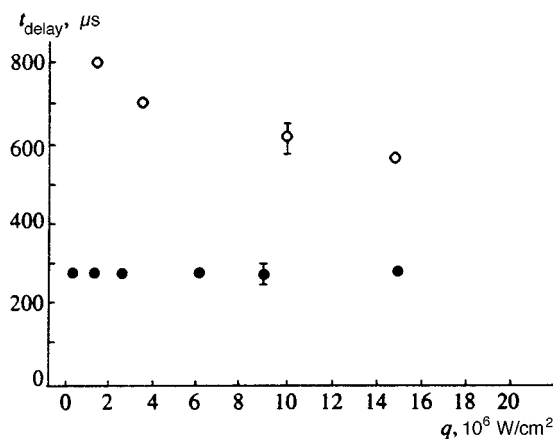

FIG. 4. Positions of the first (filled circles) and second (unfilled circles) peaks in the time-of-flight spectra of iodine molecules desorbed from a nanoporous quartz surface as functions of the laser power density at $\lambda = 532$ nm.

dine yielded $\tau_r \sim 10^{-12}$ s, which is comparable to the direct photolysis time ($\tau_{dis} = 5 \times 10^{-12}$ s) and notably shorter than the predissociation time.

Thus, our spectral and photophysical measurements indicate that the states of iodine molecules adsorbed in nanoporous quartz are qualitatively diverse (physisorbed and chemisorbed states and charge-transfer complexes) and that the radiationless relaxation for adsorbed iodine molecules is very fast, significantly hampering the initiation of nonequilibrium photoprocesses. The only channel that can compete with the radiationless relaxation is direct photolysis, which is possible under the conditions of our experiments only through the excitation of physisorbed molecules at a wavelength of 532 nm.

4. INVESTIGATION OF PROCESSES INDUCED BY RESONANT LASER RADIATION IN MOLECULAR IODINE SUBMONOLAYERS ADSORBED ON NANOPOROUS QUARTZ

4.1. General features

The experiments were performed using dynamic time-of-flight mass spectroscopy. In this case, more information can be obtained in the collisionless particle propagation regime, where the velocity distributions of species produced on the surface can be measured. The collisionless character of the particle propagation in our experiments was confirmed by numerical calculations of the rates of collisions between different molecules, which were performed for the three-dimensional case, unlike those reported by Harrison *et al.*[8]

Figure 3 shows curves of the number of particles arriving at the ion source of the mass spectrometer as functions of time (time-of-flight spectra) obtained by exposing iodine submonolayers physisorbed on surfaces of solid and nanoporous quartz to $\lambda = 532$ nm radiation at a power density $q = 1.5 \times 10^7$ W/cm^2. When iodine was adsorbed on the surface of solid quartz, only molecular iodine ($m/e = 254$) with a time-of-flight distribution peaking at 550 μs were detected. In experiments with the samples of nanoporous quartz coated with iodine, both molecular ($m/e = 254$) and atomic (m/e

$= 127$) iodine species were generated. In this case, the time-of-flight spectrum of I$_2$ is radically different from the case of solid quartz and contains two peaks at 270 and 510 μs, whereas no iodine was detected in the time interval of from 310 to 430 μs.

It is remarkable that the translational energy of iodine molecules desorbed from a nanoporous quartz surface is fairly high. The most probable energy in a bunch of high-energy particles is 1.8 eV when the energy of the pumping photons is 2.3 eV.

Even a preliminary analysis of the results indicates that photoprocesses on the nanoporous quartz surface are radically different from those on surfaces of solid (nonporous) materials.

Figure 4 shows the positions of the first and second peaks in the time-of-flight spectra of iodine molecules as functions of the laser power density at a wavelength $\lambda = 532$ nm. The dependence on the power density q for the iodine molecules of lower energy attests to their thermal nature. The absence of such a dependence for the bunch of high-energy molecules, along with their higher kinetic energy, indicates that their desorption is nonthermal (photodesorption).

4.2. Nonthermal one-photon photodesorption

Efficient populating of the $^3\Pi_{0_u^+}$ state in nanoporous quartz is impossible under conditions of radiationless relaxation with a characteristic time of 10^{-12} s at power densities of $(0.5-15) \times 10^6$ W/cm^2 and a molecular iodine absorption cross section of 10^{-19} cm^2; therefore, the photoprocess observed in experiments cannot be attributed to a multistage excitation mechanism.

The conclusion that this photoprocess has a one-photon mechanism is also supported by the linear dependence of the number of I$_2$ molecules in the high-energy peak of the time-of-flight spectrum on the energy (number of photons) in the laser pulse at $\lambda = 532$ nm (Fig. 5).

When the laser wavelength is switched from 532 to 575 nm and then to 640 nm, the time-of-flight spectrum of the
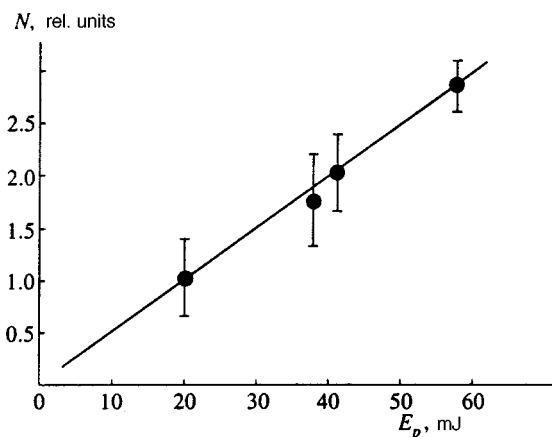
FIG. 5. Relative yield of high-energy iodine molecules desorbed from a nanoporous quartz surface by laser pulses with $\lambda = 532$ nm versus pulse energy.
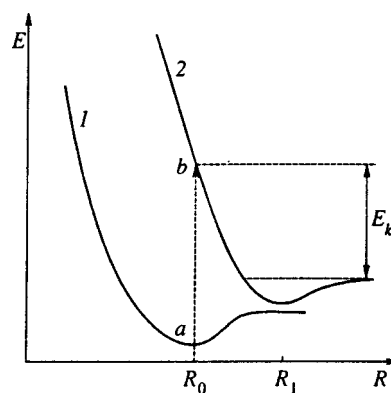

FIG. 7. Schematic potential-energy curves of the iodine-molecule/nanoporous-quartz-surface system. Curve *2* corresponds to the molecule in the excited electronic state, and curve *1* corresponds to the ground electronic state of the iodine molecule.

desorbed iodine molecules shifts (Fig. 6) from 265–300 μs to 295–325 μs. As can be seen in Fig. 6, the time-of-flight spectra for pump wavelengths of 575 nm and 640 nm are identical within the experimental error and have a more complex shape than the spectrum recorded at $\lambda = 532$ nm.

Thus, the experimentally detected nonthermal photodesorption of high-energy iodine molecules from the surface of nanoporous quartz has a one-photon character and is observed under the action of laser radiation of wavelengths $\lambda = 532$, 575, and 640 nm over a wide range of pump power densities $q = 5 \times 10^5 – 1.5 \times 10^7$ W/cm^2.

In most publications photodesorption is treated as a Franck–Condon transition to either a bound state of an adsorbate-ion/surface complex (Antonievich's mechanism)[9] or a dissociated state of an adsorbate/adsorbent complex (Menzel–Gomer–Readhead mechanism).[10] Antonievich's mechanism is characterized by small kinetic energies for the desorbed particles, in contradiction to our experimental data. Photodesorption by Menzel's mechanism is also impossible in our case because excited diatomic molecules in pores are typically in stable states of adsorbate/adsorbent complexes (Ref. 4, p. 8).

FIG. 6. Time-of-flight spectrum of high-energy iodine molecules desorbed from a nanoporous quartz surface. Laser radiation with $\lambda = 575$ nm (*1*) and $\lambda = 640$ nm (*2*); $q = 1.5 \times 10^7$ W/cm^2. The curves are visual guides.

In our experiments no high-energy component was detected in the time-of-flight spectra of I_2 molecules desorbed from nonporous quartz surfaces (Fig. 3). Therefore, we assume that the high-energy photodesorption is due to the nanostructure of the substrate and suggest the following interpretation.

The $^1\Sigma_{0_g^+}$ ground electronic state of the iodine molecule has the electronic configuration $\sigma_g^2 \pi_u^4 \pi_g^4$. As a result of the transition from the $^1\Sigma_{0_g^+}$ state to the $^3\Pi_{0^+}$ state, the electronic configuration changes from $\sigma_g^2 \pi_u^4 \pi_g^4$ to $\sigma_g^2 \pi_u^4 \pi_g^3 \sigma_u^1$, i.e., one of the π_g electrons moves into the σ_u orbital, whereas the effective radius of the σ_u orbital is considerably larger than that of the π_g orbital. For comparison, we note that the electronic radii of Cs and I are 2.62 Å and 1.36 Å, respectively (cesium was chosen because it is the first element after iodine in the periodic table in which the sixth electronic level is populated).

Figure 1 shows that when a free iodine molecule is excited to the $^3\Pi_{0^+}$ state, the equilibrium internuclear distance increases from 2.66 Å to 3.0 Å. The absorption spectra of iodine molecules physisorbed on a surface of nanoporous quartz (Fig. 2) indicate that some of the physisorbed molecules (the short-wavelength portion of the spectrum) have an even larger internuclear distance in the excited state than in the free molecule. It is estimated to be up to 3.5 Å. Thus, the effective diameter of an excited iodine molecule in a pore with allowance for the increases in both the internuclear distance and the effective radius of electronic states is 7–10 Å.

The possible dependence of the potential energy of an iodine molecule on its distance from the quartz surface is schematically depicted by curve *1* in Fig. 7, where R_0 is the equilibrium distance between the surface and molecule and the potential well depth specifies the adsorption energy. When the electronic system of the iodine molecule is excited, the effective diameter of the electronic state increases, hence the equilibrium distance from the surface to the molecule also rises. Since the nanopore volume is limited, this is equivalent to a considerable increase in the potential energy of the surface-molecule interaction. In Fig. 7 this change is illustrated by the vertical (Franck–Condon) $a–b$ transition

to potential-energy curve *2* with an equilibrium distance R_1 $> R_0$. Then the molecule is capable of desorbing with a high kinetic energy E_k.

Since the actual structure of pores is fairly complex, there can be several equilibrium positions of an excited molecules in a pore characterized by different potential-energy curves.

As was noted above, the positions of the time-of-flight spectra of desorbed particles at the pump wavelength $\lambda = 532$ nm, on the one hand, and at $\lambda = 575$ and 640 nm, on the other, are different. In terms of energy, this difference is 0.26 eV and can be ascribed to different vibrational states of the molecules. That the time-of-flight spectra of desorbed iodine molecules at $\lambda = 575$ and 640 nm are identical means that desorption occurs at the same vibronic energy of the adsorbate/adsorbent system in both cases, and this is possible only if the molecule undergoes preliminary vibrational relaxation to the zero-point level within the excited electronic term. Then we should assume that at $\lambda = 532$ nm, where molecules are excited to states near the dissociation threshold of the $^3\Pi_{0_u^+}$ term, desorption proceeds in the absence of equilibrium with respect to the vibrational states. In this case, the characteristic time of the process is comparable to the vibrational relaxation time (10^{-11}–10^{-12} s).

An analysis of experimental data also leads to a conclusion that the total number of high-energy particles is no greater than 4% of the desorbed molecules. This percentage is so small because, owing to the large pore size (20–100 Å), only a small part of the sorbed molecules exist under conditions that lead to the size effect in excitation in a confined volume followed by desorption by the mechanism described above.

4.3. Surface photodissociation of iodine molecules

Another significant difference between the photoprocesses on the surface of nanoporous quartz and the processes on the surface of solid (nonporous) quartz under laser radiation with $\lambda = 532$ nm is the presence of atomic iodine among the species emitted from the surface. The time-of-flight spectrum $A_{I_1}(t)$ is shown in Fig. 3. The spectrum has two peaks. This curve does not adequately represent the number of iodine atoms generated by the interaction and arriving at the ions source of the spectrometer since it does not take into account the fragmentation of molecular iodine. After an analysis based on the fragmentation relations, we found that the part of the $A_{I_1}(t)$ distribution peaking at 270 μs is due entirely to the fragmentation of high-energy desorbed iodine molecules in the spectrometer ion source. In addition, we could derive the time-of-flight spectrum of atomic iodine formed as a result of the photodissociation of I_2 molecules. This curve is shown in Fig. 8. The absence of atomic iodine among the species emitted when laser radiation acts on iodine molecules sorbed on solid quartz allows us to conclude that the photolysis of physisorbed iodine molecules occurs not in the gaseous phase, but on the nanostructured quartz surface.

Under laser radiation with $\lambda = 575$ nm and 640 nm focused on iodine molecules on the surface of nanoporous
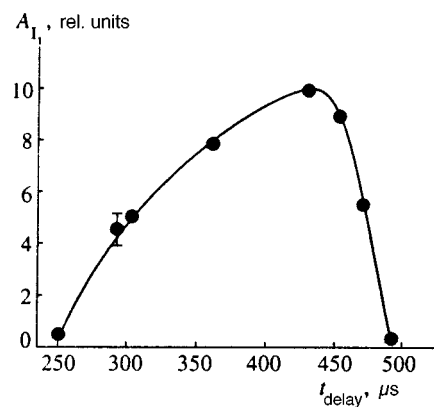

FIG. 8. Time-of-flight spectrum of atomic iodine generated in the photodissociation of I_2 molecules. Laser radiation with $\lambda = 532$ nm and $q = 1.5 \times 10^7$ W/cm^2. The curve is a visual guide.

quartz, no iodine atoms were detected among the species emitted. This fact can easily be interpreted on the basis of the spectra in Fig. 2. Only laser radiation with $\lambda = 532$ nm excites iodine molecules on a sample of nanoporous quartz near the dissociation threshold and can bring about direct photolysis of a molecule with a characteristic time of 5×10^{-12} s. Conversely, radiation with wavelengths of 575 and 640 nm has photon energies below the dissociation threshold of the $^3\Pi_{0_u^+}$ term and can lead only to the predissociation of iodine molecules via the $^1\Pi_{1_u}$ state with a characteristic time much longer than 10^{-12} s; therefore, predissociation cannot compete with either direct photolysis (when $\lambda = 532$ nm) or radiationless relaxation.

Thus, it should be concluded that excitation of the $^1\Sigma_{0_g^+} \rightarrow {}^3\Pi_{0_u^+}$ transition in I_2 submonolayers on the surface of nanoporous quartz driven by radiation at $\lambda = 532$ nm leads to the surface photodissociation of iodine molecules, which proceeds only via the direct channel. No predissociation of molecules excited to predissociation states below the photolysis threshold of the $^3\Pi_{0_u^+}$ term ($\lambda = 575$, 640 nm) was detected.

The photolysis of iodine molecules on the surface of nanoporous quartz and its absence on the surface of solid quartz can be attributed to the strong effect of the nanostructured surface on physisorbed iodine molecules, which broadens their absorption spectrum. Some of the molecules are excited at $\lambda = 532$ nm to states near the photolysis threshold of the $^3\Pi_{0_u^+}$ term (the peak of the absorption spectrum) and dissociate via the direct channel. Iodine molecules on solid quartz surfaces do not experience such a strong effect of the surface, and the action of radiation with $\lambda = 532$ nm excites them to states below the photolysis threshold on the $^3\Pi_{0_u^+}$ term (just as in the gaseous phase) and can result only in predissociation, which, as was shown above, is inefficient in the present case.

4.4. Desorption via the thermal channel

As was noted in Sec. 4.1 (Fig. 3), the second bunch of molecules desorbed from a nanoporous quartz surface with lower energies is due to the thermal effect of laser radiation.

An estimate of the maximum temperature increase in the case of uniform heating of the entire exposed area yields $\Delta T \approx 4$ K. At the same time, one can see in the time-of-flight spectra that the translational energy of the slowest desorbed iodine molecules is $\geqslant 10^3$ K.

Among the plausible mechanisms that can account for these results, we single out the mechanism of local heating and the mechanism due to the vibrational and rotational re-excitation of molecules in the upper electronic state.[11] In the latter case, the kinetic energy of the particles should not vary with the number of photons in the laser pulse, in contradiction to the experimental results (Fig. 4). Thus, the more probable mechanism of the desorption of iodine molecules with lower energies is local heating.

5. CONCLUSIONS

We have detected nonthermal one-photon desorption of iodine molecules with large kinetic energies from submonolayers on surfaces of nanoporous quartz under resonant laser radiation over a wide spectral range in the visible region, alongside the thermal desorption and surface photolysis of I_2 molecules. The character and mechanisms of the processes on nanostructured surfaces are essentially different from those on surfaces of bulk condensed materials. This has led us to the conclusion that surface photoprocesses are controlled by the specific properties of nanoporous materials.

[1] A. Belogorokhov, V. Karavanskiĭ, A. Obraztsov *et al.*, JETP Lett. **60**, 274 (1994).

[2] V. Dneprovskiĭ, V. Karavanskiĭ, V. Klimov *et al.* [JETP Lett. **57**, 406 (1993).

[3] A. Belogorokhov, V. Karavanskiĭ, and L. Belogorokhova, Fiz. Tekh. Poluprovodn. **30**, 1177 (1996) [Semiconductors **30**, 621 (1996)].

[4] V. Kravchenko, A. Orlov, Yu. Petrov *et al.*, Tr. IOFAN **11**, 1 (1988).

[5] Yu. Bykovskiĭ, M. Potapov, V. Ukraintsev, and A. Chistyakov, Khim. Vys. Energ. **21**, 361 (1987).

[6] H. Okabe, *Photochemistry of Small Molecules*, Wiley, New York (1978).

[7] S. Strickler and R. Berg, J. Chem. Phys. **37**, 814 (1962).

[8] I. Harrison, J. Polany, and P. Young, J. Chem. Phys. **89**, 1475 (1988).

[9] P. Antonievich, Phys. Rev. B **21**, 3811 (1980).

[10] D. Menzel and R. Gomer, J. Chem. Phys. **41**, 3311 (1964); P. Readhead, Can. J. Phys. **42**, 886 (1964).

[11] F. Dzegilenko and E. Herbst, J. Chem. Phys. **100**, 9205 (1994).

Tensor structure of the stationary point of the radiative relaxation operator of an atom

A. V. Taĭchenachev,*) A. M. Tumaĭkin, and V. I. Yudin

Novosibirsk State University, 630090 Novosibirsk, Russia

G. Nienhuis

Huygens Laboratorium, University of Leiden, P.O.B. 9504, 2300 RA Leiden, The Netherlands
(Submitted 22 December 1997)
Zh. Éksp. Teor. Fiz. **114**, 125–134 (July 1998)

We formulate the problem of the stationary point of the operator of radiative relaxation of an atom: the initial distribution among the sublevels of the excited state, whose nonzero eigenvalues (populations) coincide with the populations of the final distribution (after spontaneous decay) among the sublevels of the ground state. We show that these distributions can be expressed in terms of spherical functions of the complex direction. The results are then used to develop a compact analytical representation of the stationary density matrix of atoms interacting with an elliptically polarized monochromatic field. © *1998 American Institute of Physics.* [S1063-7761(98)00907-X]

1. INTRODUCTION

It is widely known that relaxation processes play a major role in the interaction of atoms with resonant radiation. A common reason for relaxation in an atomic system is the interaction (usually weak) of the system and the ambient, which drives the atoms toward an equilibrium state and causes the atomic levels to have a finite lifetime. In a low-density ensemble of atoms, where the interatomic interaction can be completely ignored, the main relaxation mechanism is radiative (the spontaneous emission of radiation). According to modern conceptions,[1] the radiative relaxation of atoms is determined by the interaction with the vacuum modes of the electromagnetic field and can be considered a QED effect. The properties of the relaxation operator reflect certain properties of the vacuum, e.g., its isotropy.

If relaxation is ignored, the interaction of an atom with light is described by a Schrödinger equation with a time-dependent Hamiltonian. This description, however, can be used only at extremely short times that are shorter than the reciprocal spontaneous relaxation rate $\gamma^{-1} \sim 10^{-8}$ s. When $t > \gamma^{-1}$, the wave function of the atom becomes a rapidly fluctuating quantity,[2] and the density-matrix formalism becomes more appropriate for analytical description. The density matrix satisfies the corresponding quantum kinetic equation, and the form of the radiative relaxation operator in the quantum kinetic equation for the atomic density matrix in the dipole approximation without allowance for strong-field effects[3] is well known (see, e.g., Ref. 4). However, its tensor properties and its effect on the formation of multipole moments in the system of Zeeman sublevels have to be studied in greater detail.

In Refs. 5 and 6 we studied the problem of the redistribution of atoms among the sublevels of the ground and excited states as a result of the closed (i.e., in which the total population is conserved) optical transitions $F_g = F \rightarrow F_e = F$ (F is a half-integer) and $F_g = F \rightarrow F_e = F + 1$ upon interac-

tion with elliptically polarized monochromatic radiation. Obviously, the stationary (nonequilibrium) distribution of the atoms among the sublevels of the ground (g) and excited (e) states is determined not only by the parameters of the external field (polarization, intensity, and frequency), but also by the properties of the radiative relaxation operator. One of the main results of Refs. 5 and 6 is that the problem of the stationary density matrix can be reduced to the problem of finding the stationary point of the radiative relaxation operator. In Ref. 5 we found that the stationary point for the $F_g = F \rightarrow F_e = F$ transitions corresponds to an isotropic distribution among the sublevels of the excited state, which made it possible to write a compact analytical expression for the stationary density matrix. In Ref. 6 the solution for the stationary point was found for the $F_g = F \rightarrow F_e = F + 1$ transitions with $F = 0, 1/2, \ldots, 5$ in the form of polynomials in the operator of the resonant interaction with the field. However, we were unable to derive a closed analytical expression for the expansion coefficients that would be valid for all values of F. Besides, the method used in Ref. 6 was based on a theorem whose validity in the general case ($F > 5$) has yet to be rigorously proved.

In the present paper the problem of the stationary point of the radiative relaxation operator of an atom is formulated explicitly, i.e., the stationary point refers to the initial distribution among the sublevels of the excited state whose non-zero eigenvalues (populations) coincide with the populations of the final distribution (after spontaneous decay) among the sublevels of the ground state. We find a particular solution of this problem in the form of convolutions of Wigner tensor operators and spherical functions of complex direction. The results are then used to develop a compact analytical representation of the stationary density matrix of atoms interacting with an elliptically polarized monochromatic field. We also prove the theorem of Ref. 6 in general form for all values of F.

70

2. STATEMENT OF THE PROBLEM

We consider an atom with two degenerate energy levels coupled by a dipole transition: a ground state (g) with angular momentum F_g and an excited state (e) with angular momentum F_e. In the quantum kinetic equation describing the evolution of the atomic density matrix[6] we can single out the terms that describe evolution as a consequence of radiative relaxation alone:

$$\left[\frac{\partial \hat{\rho}}{\partial t}\right]_{\text{rel}} = -\hat{\Gamma}\{\hat{\rho}\} = -\frac{\gamma}{2}(\hat{\Pi}_e \hat{\rho} + \hat{\rho}\hat{\Pi}_e)$$
$$+ \gamma \sum_{\beta=0,\pm 1} \hat{Q}_\beta^\dagger \hat{\rho}\hat{Q}_\beta, \qquad (1)$$

where γ is the radiative relaxation rate. The first term on the right-hand side describes the damping of the excited-state density matrix $\hat{\rho}^{ee}$ and of the off-diagonal elements $\hat{\rho}^{eg}$ and $\hat{\rho}^{ge}$, and the operator

$$\hat{\Pi}_e = \sum_{\mu_e=-F_e}^{F_e} |F_e, \mu_e\rangle\langle F_e, \mu_e| \qquad (2)$$

projects onto the excited state, $|F_e, \mu_e\rangle$ being the wave functions of the magnetic sublevels. The second term on the right-hand side of Eq. (1) describes the spontaneous passage of the atoms from the excited state to the ground state. The raising operator $\hat{Q}_\beta$ and the lowering operator $\hat{Q}_\beta^\dagger$ can be expressed in terms of the $3jm$ symbols as follows:

$$\hat{Q}_\beta = \sum_{\mu_e,\mu_g} |F_e,\mu_e\rangle\sqrt{2F_e+1}(-1)^{F_e-\mu_e}$$
$$\times \begin{pmatrix} F_e & 1 & F_g \\ -\mu_e & \beta & \mu_g \end{pmatrix}\langle F_g,\mu_g|. \qquad (3)$$

In view of the orthogonality of the vector addition coefficients[7] we have $\Sigma \hat{Q}_\beta \hat{Q}_\beta^\dagger = \hat{\Pi}_e$ and hence $\text{Tr}\{\hat{\Gamma}\{\hat{\rho}\}\} = 0$, which means that the total population is conserved.

We now assume that in some way an initial distribution $\hat{\rho}_{\text{in}}^{ee}$ is created at the excited level and that the off-diagonal elements $\hat{\rho}_{\text{in}}^{eg}$ and the ground-state density matrix $\hat{\rho}_{\text{in}}^{gg}$ are then equal to zero. Like any nonnegative definite Hermitian matrix, $\hat{\rho}_{\text{in}}^{ee}$ is given in a corresponding basis set by a set of nonnegative numbers $\{p_1^e, \ldots, p_{2F_e+1}^e\}_{\text{in}}$, the populations. After this initial distribution has been created, the atomic density matrix evolves according to Eq. (1). After a long time $t \gg \gamma^{-1}$ has elapsed, all the atoms find themselves in the ground state due to spontaneous relaxation: $\hat{\rho}_{\text{out}}^{ee} = \hat{\rho}_{\text{out}}^{eg} = 0$ and $\hat{\rho}_{\text{out}}^{gg} = \Sigma \hat{Q}_\beta^\dagger \hat{\rho}_{\text{in}}^{ee}\hat{Q}_\beta$. The corresponding eigenvalues $\{p_1^g, \ldots, p_{2F_g+1}^g\}_{\text{out}}$ differ from the initial ones $\{p_i^e\}_{\text{in}}$ in the general case. Among the possible distributions $\hat{\rho}_{\text{in}}^{ee}$ we look for those in which the nonzero populations belonging to the sets $\{p_i^e\}_{\text{in}}$ and $\{p_i^g\}_{\text{out}}$ coincide. Note that this leads, among other things, to equality between the entropies of the atomic subsystem in the initial and final states:

$$\text{Tr}\{\hat{\rho}_{\text{in}}^{ee} \ln(\hat{\rho}_{\text{in}}^{ee})\} = \text{Tr}\{\hat{\rho}_{\text{out}}^{gg} \ln(\hat{\rho}_{\text{out}}^{gg})\}.$$

We have thus completed the statement of the problem of the stationary point of the relaxation operator $\hat{\Gamma}$.

3. SPHERICAL FUNCTIONS OF COMPLEX DIRECTION

We found a particular solution of this problem. The solution, however, has important applications. To formulate our result we need to introduce spherical functions of complex direction, $Y_{LM}(\mathbf{a})$, which can be defined invariantly via a multiple tensor product.[7,8] For instance, for an arbitrary complex vector $\mathbf{a} = \mathbf{a}' + i\mathbf{a}''$ we have

$$Y_{LM}(\mathbf{a}) = \frac{1}{a^L}\sqrt{\frac{(2L+1)!!}{4\pi L!}}$$
$$\times \{\cdots\{\{\mathbf{a}\otimes\mathbf{a}\}_2\otimes\mathbf{a}\}_3\cdots\otimes\mathbf{a}\}_{LM}, \qquad (4)$$

where $a = \sqrt{\mathbf{a}\cdot\mathbf{a}}$ (not to be confused with $|\mathbf{a}| = \sqrt{\mathbf{a}^*\times\mathbf{a}}$). Spherical functions generalized in this way depend only on the direction in three-dimensional complex space,[8] i.e., the functions remain unchanged in response to the transformation $\mathbf{a}\rightarrow\nu\mathbf{a}$, where ν is an arbitrary complex number. For real vectors ($\mathbf{a}''=0$), the definition (4) leads to standard spherical functions.[7] If the quantization axis z is orthogonal to $\mathbf{a}'$ and $\mathbf{a}''$, the definition (4) yields

$$Y_{LM}(\mathbf{a}) = e^{\chi M}\sqrt{\frac{(2L+1)(L-M)!}{4\pi(L+M)!}}P_L^M(0), \qquad (5)$$

where $P_L^M(x)$ are associated Legendre functions, and the parameter χ can be expressed in terms of the cyclic components of $\mathbf{a}$ as follows: $e^\chi = -\sqrt{2}a_{+1}/a$ and $e^{-\chi} = \sqrt{2}a_{-1}/a$, since $-2a_{+1}a_{-1} = a^2$ in the respective basis set. Formula (5) can be regarded as a suitable analytic continuation of the standard definition of the spherical functions $Y_{LM}(\theta, \phi)$ (Ref. 7) to complex values of the angle ϕ. What is important here is that the function (4) satisfies the same basic group relationships as ordinary spherical functions. In particular, the Clebsch–Gordan expansion of the product of two spherical functions of the same argument holds, i.e.,

$$Y_{l_1 m_1}(\mathbf{a})Y_{l_2 m_2}(\mathbf{a}) = \sum_{LM}\sqrt{\frac{(2l_1+1)(2l_2+1)}{4\pi(2L+1)}}$$
$$\times C_{l_1 0\ l_2 0}^{L0}C_{l_1 m_1 l_2 m_2}^{LM}Y_{LM}(\mathbf{a}) \qquad (6)$$

($C_{l_1 m_1 l_2 m_2}^{LM}$ are the Clebsch–Gordan coefficients), as does the addition theorem for the scalar product of spherical functions of different arguments[8]

$$Y_L(\mathbf{a})Y_L(\mathbf{b}) = \frac{2L+1}{4\pi}P_L\left(\frac{\mathbf{a}\cdot\mathbf{b}}{ab}\right), \qquad (7)$$

where $P_L(x)$ are Legendre polynomials.

4. ALGEBRA OF THE OPERATORS $\hat{V}_L(\mathbf{a})$

We now define the operator convolutions

$$\hat{V}_L(\mathbf{a}) = \frac{a^L\sqrt{4\pi}}{2L+1}\sum_{M=-L}^{L}(-1)^M\hat{T}_{LM}^{eg}Y_{L-M}(\mathbf{a}) \qquad (8)$$

of spherical functions and Wigner tensor operators of rank L:

$$\hat{T}_{LM}^{ab} = \sum_{\mu_a,\mu_b} |F_a,\mu_a\rangle \sqrt{2L+1}(-1)^{F_a-\mu_a}$$

$$\times \begin{pmatrix} F_a & L & F_b \\ -\mu_a & M & \mu_b \end{pmatrix} \langle F_b,\mu_b|. \tag{9}$$

Employing the algebra of Wigner operators, we can write general expressions for products of $\hat{V}$ operators of different ranks:

$$\hat{V}_L(\mathbf{a})\hat{V}_J^\dagger(\mathbf{b}) = \sum_{K,q} \frac{(-1)^{K+q+F_e-F_g}4\pi a^L b^{*J}}{\sqrt{(2L+1)(2J+1)}}$$

$$\times \begin{Bmatrix} K & L & J \\ F_g & F_e & F_e \end{Bmatrix}$$

$$\times \{Y_L(\mathbf{a})\otimes Y_J(\mathbf{b}^*)\}_{K-q}\hat{T}_{Kq}^{ee},$$

$$\hat{V}_L^\dagger(\mathbf{a})\hat{V}_J(\mathbf{b}) = \sum_{K,q} \frac{(-1)^{K+q+F_g-F_e}4\pi a^{*L} b^{J}}{\sqrt{(2L+1)(2J+1)}}$$

$$\times \begin{Bmatrix} K & L & J \\ F_e & F_g & F_g \end{Bmatrix}$$

$$\times \{Y_L(\mathbf{a}^*)\otimes Y_J(\mathbf{b})\}_{K-q}\hat{T}_{Kq}^{gg}, \tag{10}$$

where we have used the standard notation for the $6j$ symbols.[7] Hence, in particular, we have the following commutation relations.

1) For the arbitrary ranks L and J and all dipole transitions of the type $F_e=F\to F_g=F,F\pm 1$ we have

$$\hat{V}_L(\mathbf{a})\hat{V}_J^\dagger(\mathbf{a}^*) = \hat{V}_J(\mathbf{a})\hat{V}_L^\dagger(\mathbf{a}^*),$$

$$\hat{V}_L^\dagger(\mathbf{a})\hat{V}_J(\mathbf{a}^*) = \hat{V}_J^\dagger(\mathbf{a})\hat{V}_L(\mathbf{a}^*), \tag{11}$$

since expansion of the right- and left-hand sides in the tensor operators $\hat{T}_{Kq}^{ee}(\hat{T}_{Kq}^{gg})$ leads to identities of the type $\{Y_L(\mathbf{a})\otimes Y_J(\mathbf{a})\}_{Kq} \equiv \{Y_J(\mathbf{a})\otimes Y_L(\mathbf{a})\}_{Kq}$, which in turn follow from the Clebsch–Gordan expansion (6) for spherical functions.

2) Depending on the type of transition, for the arbitrary vectors $\mathbf{a}$ and $\mathbf{b}$ we have:

a) for $F_g=F\to F_e=F$ transitions

$$\hat{V}_0^\dagger(\mathbf{a})\hat{V}_1(\mathbf{b}) = \hat{V}_1^\dagger(\mathbf{b}^*)\hat{V}_0(\mathbf{a}^*); \tag{12}$$

b) for $F_g=F\to F_e=F+1$ transitions

$$\hat{V}_{2F+1}^\dagger(\mathbf{a})\hat{V}_1(\mathbf{b}) = \hat{V}_1^\dagger(\mathbf{b}^*)\hat{V}_{2F+1}(\mathbf{a}^*); \tag{13}$$

c) for $F_g=F\to F_e=F-1$ transitions

$$\hat{V}_{2F-1}(\mathbf{a})\hat{V}_1^\dagger(\mathbf{b}) = \hat{V}_1(\mathbf{b}^*)\hat{V}_{2F-1}^\dagger(\mathbf{a}^*). \tag{14}$$

The property (12) is obvious if we recall that in this case the operator $\hat{V}_0$ is proportional to the identity matrix and $\hat{V}_1^\dagger(\mathbf{b}^*) = \hat{V}_1(\mathbf{b})$. To prove the validity of (13) it is sufficient to expand both sides of (13) in the operators $\hat{T}_{Kq}^{gg}$ and allow for the fact that all ranks except $K=2F$ are forbidden by the selection rules contained in the $6j$ symbols in (10). Equation (13) then reduces to the identity $\{Y_{2F+1}(\mathbf{a}^*)\otimes Y_1(\mathbf{b})\}_{2Fq}$

$\equiv \{Y_1(\mathbf{b})\otimes Y_{2F+1}(\mathbf{a}^*)\}_{2Fq}$, which holds since the number $(2F+1)+1-2F=2$ is even. Equation (14) can be proved in a similar way.

5. STATIONARY POINT OF THE RELAXATION OPERATOR

We shall seek the solution for the stationary point of the operator $\hat{\Gamma}$ in the form

$$\hat{\rho}_{\text{in}}^{ee} = \hat{V}_L(\mathbf{a})\hat{V}_L^\dagger(\mathbf{a}). \tag{15}$$

The above algebraic relationships imply that (a) for $F_g=F \to F_e=F-1$ transitions no such solution exists and that (b) for $F_g=F\to F_e=F$ transitions the matrix (15) is the solution for the lowest possible rank $L=0$.

Incidentally, these two statements are obvious corollaries of the isotropy of the relaxation operator. The thing is that since for $F_e=F_g-1$ the number of filled sublevels of the excited state cannot exceed $2F_g-1$, at least two more sublevels of the ground state are filled as a result of spontaneous decay, and for this reason alone there can be no stationary point. When $F_e=F_g$, the number of sublevels in both states is identical, and the isotropic initial distribution will simply become the isotropic final distribution, i.e., a stationary point exists and corresponds to a uniform distribution.[5]

We can also prove that a third statement is also true: for $F_g=F\to F_e=F+1$ transitions the matrix (15) is the solution for highest possible rank $L=2F+1$. This is a nontrivial fact.

Indeed, in this case the angular momentum of the excited state exceeds the angular momentum of the ground state by unity, and for a stationary point to exist two sublevels of the excited state must be unoccupied. Clearly, no matter how we choose these two sublevels, the uniform distribution among the remaining $2F_g+1$ sublevels of the excited state corresponds to some anisotropy, and spontaneous decay to the ground state produces an anisotropic, i.e., nonuniform, distribution among the sublevels of the ground state. Thus, the ''naïve'' generalization of the results for $F_e=F_g$ does not work here, and the form of the solution is unclear.

Now let us go back to the proof of the third statement. In (13) we set $\mathbf{b}=\mathbf{e}_\beta$, where $\mathbf{e}_\beta$ are the cyclic basis unit vectors. Then, bearing in mind that $\hat{Q}_\beta = \sqrt{2F_e+1}\,\hat{V}_1(\mathbf{e}_\beta)$, we obtain

$$\hat{\rho}_{\text{out}}^{gg} = \sum_\beta \hat{Q}_\beta^\dagger \hat{V}_{2F+1}(\mathbf{a})\hat{V}_{2F+1}^\dagger(\mathbf{a})\hat{Q}_\beta = \hat{V}_{2F+1}^\dagger(\mathbf{a}^*)$$

$$\times \left(\sum_\beta \hat{Q}_\beta \hat{Q}_\beta^\dagger\right)\hat{V}_{2F+1}(\mathbf{a}^*) = \hat{V}_{2F+1}^\dagger(\mathbf{a}^*)\hat{V}_{2F+1}(\mathbf{a}^*). \tag{16}$$

Note that the matrix $\hat{V}_{2F+1}(\mathbf{a})\hat{V}_{2F+1}^\dagger(\mathbf{a})$ has a two-dimensional null subspace (a subspace of vectors with zero eigenvalues), which coincides with the null subspace of the matrix $\hat{V}_1(\mathbf{a})\hat{V}_1^\dagger(\mathbf{a})$. This becomes evident in the system of coordinates in which the complex vector $\mathbf{a}$ is a linear combination of one linear and one circular component[5,6] (e.g., $\mathbf{a}=a^0\mathbf{e}_0+a^{+1}\mathbf{e}_{+1}$). Setting $\mathbf{b}=\mathbf{a}$ in (13) and employing

(11), we find that the matrices $\hat{V}_{2F+1}(\mathbf{a})\hat{V}_{2F+1}^{\dagger}(\mathbf{a})$ and $\hat{V}_{2F+1}^{\dagger}(\mathbf{a}^*)\hat{V}_{2F+1}(\mathbf{a}^*)$ satisfy the following commutation relation:

$$\hat{V}_{2F+1}(\mathbf{a})\hat{V}_{2F+1}^{\dagger}(\mathbf{a})\hat{V}_1(\mathbf{a}) = \hat{V}_1(\mathbf{a})\hat{V}_{2F+1}^{\dagger}(\mathbf{a}^*)\hat{V}_{2F+1}(\mathbf{a}^*). \tag{17}$$

Multiplying (17) on the right (left) by $\hat{V}_1^{\dagger}(\mathbf{a})$ and finding the Hermitian conjugate of the product, we obtain the important commutation relations

$$[\hat{V}_{2F+1}(\mathbf{a})\hat{V}_{2F+1}^{\dagger}(\mathbf{a}), \hat{V}_1(\mathbf{a})\hat{V}_1^{\dagger}(\mathbf{a})] = 0,$$
$$[\hat{V}_{2F+1}^{\dagger}(\mathbf{a}^*)\hat{V}_{2F+1}(\mathbf{a}^*), \hat{V}_1^{\dagger}(\mathbf{a})\hat{V}_1(\mathbf{a})] = 0. \tag{18}$$

From operator algebra we know that $\hat{V}_{2F+1}(\mathbf{a})\hat{V}_{2F+1}^{\dagger}(\mathbf{a})$ and $\hat{V}_{2F+1}^{\dagger}(\mathbf{a}^*)\hat{V}_{2F+1}(\mathbf{a}^*)$ are diagonal in basis sets of the Hermitian operators $\hat{V}_1(\mathbf{a})\hat{V}_1^{\dagger}(\mathbf{a})$ and $\hat{V}_1^{\dagger}(\mathbf{a})\hat{V}_1(\mathbf{a})$, respectively. Calculating the matrix elements of (17) using the eigenvectors of the operators $\hat{V}_1(\mathbf{a})\hat{V}_1^{\dagger}(\mathbf{a})$ and $\hat{V}_1^{\dagger}(\mathbf{a})\hat{V}_1(\mathbf{a})$, we conclude that $\hat{\rho}_{\text{in}}^{ee} = \hat{V}_{2F+1}(\mathbf{a})\hat{V}_{2F+1}^{\dagger}(\mathbf{a})$ and $\hat{\rho}_{\text{out}}^{gg} = \hat{V}_{2F+1}^{\dagger}(\mathbf{a}^*)\hat{V}_{2F+1}(\mathbf{a}^*)$ have coinciding nonzero eigenvalues, which is what we set out to prove.

6. STATIONARY ATOMIC DENSITY MATRIX

We can show that the above solution for the stationary point of the relaxation operator corresponds to a stationary distribution of atoms among the sublevels of the excited state in an external field. Actually, this fact was used in Ref. 5 for $F_g = F \rightarrow F_e = F$ transitions, (F is a half-integer). Similarly, the nontrivial solution for $F_g = F \rightarrow F_e = F+1$ transitions produces a compact analytical representation for the atomic density matrix.

Briefly, the statement of the problem is as follows. The quantum kinetic equation for atoms interacting with an elliptically polarized monochromatic field is obtained by adding the following Hamiltonian part to (1):

$$\frac{\partial \hat{\rho}}{\partial t} = -\frac{i}{\hbar}[\hat{H}, \hat{\rho}] - \hat{\Gamma}\{\hat{\rho}\}, \tag{19}$$

where

$$\hat{H} = -\hbar \delta \hat{\Pi}_e + \hbar(\Omega \hat{V}_1(\mathbf{e}) + \Omega^* \hat{V}_1^{\dagger}(\mathbf{e})) \tag{20}$$

is the atomic Hamiltonian in the rotating-wave approximation, $\delta = \omega - \omega_0$ is the detuning of the field frequency ω from the transition frequency ω_0, and Ω is the Rabi frequency. We see that the operators introduced in (8) for $L = 1$ coincide with the tensor factor in the resonant dipole interaction operator. The unit complex vector $\mathbf{e}$ of elliptic polarization of the field satisfies the conditions $\mathbf{e}^* \cdot \mathbf{e} = 1$ and $\text{Im}(\mathbf{e} \cdot \mathbf{e}) = 0$. The exact stationary solution $\hat{\rho}_{\text{st}}$ of Eq. (19) for $F_g = F \rightarrow F_e = F+1$ transitions can be expressed in terms of the operators $\hat{V}_L$ as follows:

$$\hat{\rho}_{\text{st}}^{ee} = \frac{S}{\alpha_0 + 2S\alpha_1} \hat{V}_{2F+1}(\mathbf{e})\hat{V}_{2F+1}^{\dagger}(\mathbf{e}),$$

$$\hat{\rho}_{\text{st}}^{gg} = \frac{1}{\alpha_0 + 2S\alpha_1}([\hat{V}_1^{\dagger}(\mathbf{e})\hat{V}_1(\mathbf{e})]^{-1} + S)$$
$$\times \hat{V}_{2F+1}^{\dagger}(\mathbf{e}^*)\hat{V}_{2F+1}(\mathbf{e}^*),$$

$$\hat{\rho}_{\text{st}}^{eg} = (\hat{\rho}_{\text{st}}^{ge})^{\dagger} = \frac{\Omega}{(\delta + i\gamma/2)(\alpha_0 + 2S\alpha_1)}\hat{V}_1(\mathbf{e}) \tag{21}$$
$$\times [\hat{V}_1^{\dagger}(\mathbf{e})\hat{V}_1(\mathbf{e})]^{-1}\hat{V}_{2F+1}^{\dagger}(\mathbf{e}^*)\hat{V}_{2F+1}(\mathbf{e}^*),$$

where $S = |\Omega|^2/(\gamma^2/4 + \delta^2)$ is the transition saturation parameter. The validity of this solution becomes evident if we plug (21) into (19) (or into the equations for the stationary density matrices $\hat{\rho}_{\text{st}}^{ee}$ and $\hat{\rho}_{\text{st}}^{gg}$, whose expanded form is given in Ref. 6) and allow for Eqs. (16)–(18) and if we bear in mind that the projection of the matrix $\hat{V}_{2F+1}(\mathbf{e})\hat{V}_{2F+1}^{\dagger}(\mathbf{e})$ onto the null subspace of the operator $\hat{V}_1(\mathbf{e})\hat{V}_1^{\dagger}(\mathbf{e})$ is zero. The uniqueness of the stationary solution for the present type of optical transition was proved earlier in Ref. 6. We stress that the distribution among the sublevels of the excited state coincides, to within a normalization constant, with the stationary point of the relaxation operator for the particular value $\mathbf{a} = \mathbf{e}$. The normalization constants α_i depend only on the ellipticity ε of the field ($\cos 2\varepsilon = \mathbf{e} \cdot \mathbf{e}$) and can be found explicitly for all values of F. The addition theorem (7) for spherical functions implies that

$$\alpha_1 = \text{Tr}\{\hat{V}_{2F+1}(\mathbf{e})\hat{V}_{2F+1}^{\dagger}(\mathbf{e})\}$$
$$= \frac{(\cos 2\varepsilon)^{2F+1}}{4F+3}P_{2F+1}\left(\frac{1}{\cos 2\varepsilon}\right). \tag{22}$$

Calculating the matrix elements (see the Appendix) in the coordinate system used in Refs. 5 and 6, where $\mathbf{e} = \mathbf{e}_0\sqrt{\cos 2\varepsilon} + \sqrt{2}\mathbf{e}_{+1}\sin\varepsilon$, we find that

$$\alpha_0 = \text{Tr}\{[\hat{V}_1^{\dagger}(\mathbf{e})\hat{V}_1(\mathbf{e})]^{-1}\hat{V}_{2F+1}^{\dagger}(\mathbf{e}^*)\hat{V}_{2F+1}(\mathbf{e}^*)\}$$
$$= \frac{(F+1)(2F+1)!(2F+3)!}{(4F+3)!}(\cos 2\varepsilon)^{2F}$$
$$\times \sum_{n=0}^{2F} A_n\left(\frac{\sin^2\varepsilon}{\cos 2\varepsilon}\right)^n, \tag{23}$$

$$A_n = \sum_{m=-F}^{F-n}\frac{(F+m+n)!(F-m-n)!}{(F+m)!(F-m)!}$$
$$\times\left[\sum_{\mu=0}^{n}(-1)^{\mu}\frac{(2F+1+\mu)!}{\mu!(F+1+\mu+m)!(F+1-\mu-m)!}\right]^2.$$

The commutation relations (18) immediately yield

$$[\hat{\rho}_{\text{st}}^{ee}, \hat{V}_1(\mathbf{e})\hat{V}_1^{\dagger}(\mathbf{e})] = 0, \quad [\hat{\rho}_{\text{st}}^{gg}, \hat{V}_1^{\dagger}(\mathbf{e})\hat{V}_1(\mathbf{e})] = 0. \tag{24}$$

Thus, we have proved the main theorem of Ref. 6 in general form.

Note that the solution (21) coincides with the result of Ref. 6 when $F = 0, 1/2, \ldots, 5$. To verify this it is enough to represent the operators $\hat{V}_{2F+1}(\mathbf{e})\hat{V}_{2F+1}^{\dagger}(\mathbf{e})$ and $\hat{V}_{2F+1}^{\dagger}(\mathbf{e}^*)\hat{V}_{2F+1}(\mathbf{e}^*)$ by power series in $\hat{V}_1(\mathbf{e})\hat{V}_1^{\dagger}(\mathbf{e})$ and $\hat{V}_1^{\dagger}(\mathbf{e})\hat{V}_1(\mathbf{e})$, respectively. Here the normalization factors cal-

culated in Ref. 6 (we denote them by $\tilde{\alpha}_1$ and $\tilde{\alpha}_0$) differ from (22) and (23) by a conversion factor C, i.e., $\tilde{\alpha}_i = C\alpha_i$, where

$$C = \frac{(4F+3)(2F+1)!^2}{[(F+1)(2F+1)(2F+3)]^{2F+1}}.$$

In Ref. 6 we presented concrete calculations (of the absorption coefficient and the radiative force) using the stationary solution for several values of angular momentum, $F = 0, 1/2, \ldots, 5$. Here we shall not consider applications of the results, since this constitutes a separate problem.

The solution (21) has an invariant form, which makes it possible to find the values of the elements of the stationary density matrix in both the ordinary (F, μ) representation (see the Appendix) and in the atomic-polarization-moments representation [the (K, q) representation]. In particular, from (21) and (10) it follows that the polarization moments of the excited state can be expressed in terms of tensor products of spherical functions:

$$\rho_{Kq}^{ee} = \mathrm{Tr}\{\hat{T}_{Kq}^{ee}\hat{\rho}_{st}^{ee}\} = \frac{4\pi S(\cos 2\varepsilon)^{2F+1}}{(\alpha_0 + 2S\alpha_1)(4F+3)}(-1)^{K+1}$$

$$\times \begin{Bmatrix} K & 2F+1 & 2F+1 \\ F & F+1 & F+1 \end{Bmatrix}$$

$$\times \{Y_{2F+1}(\mathbf{e}) \otimes Y_{2F+1}(\mathbf{e}^*)\}_{Kq}. \tag{25}$$

Note that the distribution among the sublevels of the excited state $\hat{\rho}_{st}^{ee}$ in the form (21) for the particular case of linear polarization of the field was first obtained by Macek and Hertel.[9]

7. CONCLUSION

The present work is a continuation of a series of papers[10,5,6] devoted to the fundamental problem of the stationary density matrix of atoms interacting with an elliptically polarized resonant field. The exact solutions obtained in a compact analytical form have an extremely broad range of applications, among which the most important are laser cooling and the trapping of atoms in nonuniformly polarized fields, the resonance fluorescence of atoms captured by a magneto-optical trap, and nonlinear polarization spectroscopy without Doppler broadening. On the other hand, in this paper we have developed a new approach for studying the tensor properties of the radiative relaxation operator, and the results may also prove useful in analyzing the temporal evolution of atoms subjected to an external elliptically polarized electromagnetic field and a vacuum electromagnetic field.

APPENDIX: CALCULATING THE MATRIX ELEMENTS

The simplest way to calculate the matrix elements of the stationary atomic density matrix (21) is to employ the system of coordinates used in Refs. 5 and 6, where the field polarization vector is a linear combination of one linear and one circular component:

$$\mathbf{e} = \mathbf{e}_0 \sqrt{\cos 2\varepsilon} + \sqrt{2}\mathbf{e}_{+1}\sin\varepsilon. \tag{A1}$$

To make things more compact, we introduce the notation $\hat{V}_1(\mathbf{e}) = \hat{V}$, $\hat{V}_{2F+1}(\mathbf{e}) = \hat{W}$, and $\hat{V}_{2F+1}(\mathbf{e}^*) = \hat{\tilde{W}}$ and rewrite the solution (21) as follows:

$$\hat{\rho}_{st}^{ee} = \frac{S}{\alpha_0 + 2S\alpha_1} \hat{W}\hat{W}^\dagger,$$

$$\hat{\rho}_{st}^{gg} = \frac{1}{\alpha_0 + 2S\alpha_1}(\hat{X}\hat{X}^\dagger + S\hat{\tilde{W}}^\dagger\hat{\tilde{W}}), \tag{A2}$$

$$\hat{\rho}_{st}^{eg} = (\hat{\rho}_{st}^{ge})^\dagger = \frac{\Omega}{(\delta + i\gamma/2)(\alpha_0 + 2S\alpha_1)} \hat{W}\hat{X}^\dagger,$$

where $\hat{X} = \hat{U}\hat{W}$ with $\hat{U}$ the pseudoinverse matrix[11] relative to $\hat{V}$: $\hat{V}\hat{U}\hat{V} = \hat{V}$. For the generalized spherical functions (4) in the basis set (A1) we have

$$Y_{L-M}(\mathbf{e}) = (-1)^M$$

$$\times \sqrt{\frac{(2L+1)(L+M)!}{4\pi(L-M)!}}\frac{1}{M!}\left(\frac{\sin\varepsilon}{\sqrt{\cos 2\varepsilon}}\right)^M \tag{A3}$$

for $M \geq 0$ and $Y_{L-M}(\mathbf{e}) = 0$ for $M < 0$. Plugging (A3) into the definition (8), we obtain ($F_g = F$, $F_e = F+1$, and $L = 2F+1$)

$$W_{\mu m} = (-1)^{F-m}\frac{(2F+1+\mu-m)!}{(\mu-m)!}$$

$$\times \sqrt{\frac{(2F+2)!(2F)!}{(4F+3)!(F+1+\mu)!(F+1-\mu)!(F+m)!(F-m)!}}$$

$$\times(\cos 2\varepsilon)^{F+1/2}\left(\frac{\sin\varepsilon}{\sqrt{\cos 2\varepsilon}}\right)^{\mu-m}, \tag{A4}$$

where $\mu = -F-1, -F, \ldots, F+1$, $m = -F, -F+1, \ldots, F$, and $\mu - m \geq 0$. The matrix $\hat{\tilde{W}}$ can be obtained from (A4) using the time-reversal operation $\tilde{W}_{\mu m} = (-1)^{\mu+m}W_{-\mu-m}$. The nonzero matrix elements of $\hat{V}$ are given by the following formulas:

$$V_{\mu,\mu} = \sqrt{\frac{(F+1-\mu)(F+1+\mu)}{(F+1)(2F+1)(2F+3)}}\sqrt{\cos 2\varepsilon},$$

$$V_{\mu,\mu-1} = \sqrt{\frac{(F+\mu)(F+1+\mu)}{(F+1)(2F+1)(2F+3)}}\sin\varepsilon. \tag{A5}$$

As the matrix pseudoinverse to $\hat{V}$ we take the matrix with the elements

$$U_{m\mu} = \sqrt{\frac{(F+1)(2F+1)(2F+3)}{(F+1+\mu)(F+1-\mu)\cos 2\varepsilon}}$$

$$\times\left(-\frac{\sin\varepsilon}{\sqrt{\cos 2\varepsilon}}\right)^{m-\mu}\prod_{\nu=\mu+1}^{m}\sqrt{\frac{F+\nu}{F+1-\nu}} \tag{A6}$$

for $\mu = -F, -F+1, \ldots, F$, supplemented by the zero columns $U_{m-F-1} = 0$ and $U_{mF+1} = 0$. Here we set $\Pi_{\nu=m+1}^{m}f_\nu$

$\equiv 1$ and $\Pi^{m}_{\nu=\mu+1} f_{\nu} \equiv 0$ for $\mu > m$. The matrix multiplication of (A6) by (A4) yields the following result for the elements of $\hat{X}$:

$$X_{mm'} = \sqrt{\frac{(F+1)(2F+1)!(2F+3)!(F+m)!(F-m)!}{(4F+3)!(F+m')!(F-m')!}}$$

$$\times (\cos 2\varepsilon)^{F} \left(-\frac{\sin \varepsilon}{\sqrt{\cos 2\varepsilon}} \right)^{m-m'}$$

$$\times \sum_{\mu=m'}^{m} \frac{(-1)^{F-\mu}(2F+1+\mu-m')!}{(\mu-m')!(F+1+\mu)!(F+1-\mu)!}. \quad (A7)$$

*⁾E-mail: llf@admin.nsu.ru

[1] G. S. Agarwal, *Quantum Statistical Theories of Spontaneous Emission and Their Relation to Other Approaches*, Springer, Berlin (1974); H. J. Kimble and L. Mandel, Phys. Rev. A **13**, 2123 (1976).

[2] K. Mølmer and Y. Castin, Quantum Semiclassic. Opt. **8**, 49 (1996).

[3] É. G. Pestov, Tr. Fiz. Inst. im. P. N. Lebedev, Akad. Nauk SSSR **187**, 60 (1988); An. V. Vinogradov, Tr. Fiz. Inst. im. P. N. Lebedev, Akad. Nauk SSSR **187**, 117 (1988) (and the literature cited therein).

[4] S. G. Rautian, G. I. Smirnov, and A. M. Shalagin, *Nonlinear Resonances in the Spectra of Atoms and Molecules* [in Russian], Nauka, Novosibirsk (1979), Chap. 2, Sec. 3.

[5] A. V. Taĭchenachev, A. M. Tumaĭkin, V. I. Yudin, and G. Nienhuis, Zh. Éksp. Teor. Fiz. **108**, 415 (1995) [JETP **81**, 224 (1995)].

[6] A. V. Taĭchenachev, A. M. Tumaĭkin, and V. I. Yudin, Zh. Éksp. Teor. Fiz. **110**, 1727 (1996) [JETP **83**, 949 (1996)].

[7] D. A. Varshalovich, A. I. Moskalev, and V. K. Khersonskiĭ, *Quantum Theory of Angular Momentum*, Nauka, Leningrad (1975); World Scientific, Singapore (1988).

[8] N. L. Manakov and A. V. Merem'yanin, Zh. Éksp. Teor. Fiz. **111**, 1984 (1997) [JETP **84**, 1080 (1997)].

[9] J. Macek and I. V. Hertel, J. Phys. B **7**, 2173 (1974).

[10] V. S. Smirnov, A. M. Tumaĭkin, and V. I. Yudin, Zh. Éksp. Teor. Fiz. **96**, 1613 (1989) [Sov. Phys. JETP **69**, 913 (1989)].

[11] F. R. Gantmacher, *The Theory of Matrices*, Chelsea, New York (1959) [Russ. original of newer addition, Nauka, Moscow (1988), Chap. 1, Sec. 5].

Formation mechanisms and structure of the luminescence spectra of a dense resonant medium

Yu. K. Zemtsov, A. Yu. Sechin, A. N. Starostin[*]

Troitsk Institute of Innovative and Thermonuclear Research, 142092 Troitsk, Moscow Region, Russia

A. G. Leonov, A. A. Rudenko, and D. I. Chekhov[†]

Moscow Institute of Physics and Technology, 141700 Dolgoprudnyĭ, Moscow Region, Russia

The purely thermal visible and infrared radiation emitted by a dense resonant medium (sodium vapor) heated nonuniformly to temperatures of 600–1200 K was investigated experimentally for the first time under conditions where the photon mean free path is comparable with the emission wavelength. The profile of the recorded spectra and the absolute luminescence intensities in the different spectral ranges show good agreement with the results of a numerical simulation using a previously developed theory of resonance radiation transport which assumes a Boltzmann spectral distribution of the resonant level population proportional to $\exp(-\hbar\omega/T)$. The self-reversed resonant sodium line exhibited strong asymmetry and it was shown that under certain conditions, the luminescence spectrum of the medium may exhibit an additional broad peak on the far ''red'' limb of the resonance line. Calculations and measurements demonstrated that the intensity of the thermal emission of sodium vapor at this red peak is several orders of magnitude higher than that obtained from the standard theory of resonance radiation transport. This effect is arbitrarily termed an infrared ''catastrophe.'' It is noted that in a solar corona plasma and in gas-discharge lamps, the far red limbs of the resonant lines may make a substantial contribution to the total luminescence intensity and in some cases, considerably exceed the intensity of the photorecombination and bremsstrahlung continuum. © *1998 American Institute of Physics.* [S1063-7761(98)01007-5]

1. INTRODUCTION

The formation of luminescence spectra and resonance radiation transport processes in dense gases and plasmas has recently attracted increasing interest.[1–7] This is mainly attributable to the development of research to study radiation processes in dense, multiply charged, ion plasmas created by laser radiation or by various types of discharges (*Z* and *X* pinches and so on) and work on the development of x-ray lasers. The considerable progress achieved in this field has revealed an urgent need to radically reexamine various established concepts in the standard theory of resonance radiation transport[8,9] developed for low gas densities, and in particular, has shown that some of the approximations used in this theory must be abandoned. This particularly applies to the well-known Biberman criterion[8] $\varepsilon' - 1 \ll 1$ for the validity of the standard theory (ε' is the real part of the permittivity $\varepsilon = \varepsilon' + i\varepsilon''$ of the medium) and the narrow-line approximation $\Delta\omega = \omega - \omega_0 \ll \omega_0$, where ω_0 is the frequency of the resonant transition and ω is the frequency of the emitted photon. Taking account of this observation, the authors of Refs. 10 and 11 developed for the first time a generalized theory of resonance radiation transport free from these constraints. The approach developed in these studies predicted the existence of fundamentally new effects in radiation transport processes in dense media. In particular, it was shown in Ref. 11 that in a dense strongly absorbing medium the inten-

sity of the equilibrium radiation J_ω within the line profile may substantially exceed the Planck value. It was also observed[10] that under certain conditions, the resonant line profile may exhibit a second broad peak shifted in the low-frequency direction relative to ω_0, where the emission intensity in this red limb may be many orders of magnitude higher than the intensity of its near-resonant component. This phenomenon was described in Ref. 12 as an infrared catastrophe.

These effects may show up to a considerable extent in observations of thermal radiation from a dense nonuniformly heated medium, so that the theory put forward in Refs. 10 and 11 can be checked experimentally. Since the highly time-dependent, highly inhomogeneous plasma of a laser spark, pinches, etc., is a fairly inconvenient object to study, it was interesting to check this theory experimentally using a simple model medium such as thermally heated dense alkali metal vapor. The present paper reports the experimental observation of the effects predicted in Refs. 10 and 11 using the thermal radiation from dense sodium vapor heated to temperatures of 600–1200 K under steady-state conditions (results were briefly reported in Refs. 12 and 13). The experimental data are also compared with the results of a numerical simulation. The structure of the luminescence spectrum of the sodium resonance doublet $3P - 3S$ was studied experimentally. We particularly note that purely thermal radiation was recorded from the vapor in the absence of any excitation

by an electric field or an external radiation source. Under these conditions, the electron density is negligible and the intensity of the bremsstrahlung and photorecombination continua is also low. The relatively low temperatures eliminate the appearance of any other lines in the thermal radiation spectrum.

This study is constructed as follows. Section 2 gives a qualitative analysis of the effects and Sec. 3 reports the results of a numerical simulation of the radiation processes in dense media under different conditions. Section 4 gives comparative estimates of the ratio of the radiation intensity on the resonant lines and the intensity of the photorecombination and bremsstrahlung continua under plasma conditions. Section 5 contains descriptions of the experimental apparatus and diagnostic techniques. Section 6 presents results of measurements of the thermal radiation emitted by dense sodium vapor in various spectral ranges, discusses these experimental results and compares them with the calculations. The main conclusions are put forward in the concluding section.

2. QUALITATIVE ANALYSIS

It has already been noted in the Introduction that at high pressures, the standard theory of resonance radiation transport[8,9] developed for low gas densities is no longer valid, since the condition for a small deviation of ε from unity ceases to be satisfied and the photon mean free path is comparable with the emission wavelength. It is easy to show that in a dense medium, a photon is a ''poor'' quasiparticle, for which the dispersion relation $\sqrt{\varepsilon'}\,\omega = ck$ is not satisfied, where c is the velocity of light and k is the wave vector. In this case, the problem of resonance radiation transport can be solved using an equation for the Fourier components of the correlation functions of the electromagnetic field which can be formulated in terms of kinetic Green's functions.[11] This allows us to introduce the generalized spectral ''intensity'' $J(\omega, \mathbf{k})$ of the radiation, where the frequency ω and the wave vector $\mathbf{k}$) are independent variables. In general the function $J(\omega, \mathbf{k}$ is not positive and several of its moments, i.e., integrals over k with different weighing functions, are physically meaningful. Thus, the spectral radiation intensity J_ω usually encountered in the measurements and standard transport theory is related to $J(\omega, \mathbf{k})$ by[11]

$$J_\omega(\mathbf{\Omega}) = \frac{2c^2}{\omega} \int_0^\infty kJ(\omega, \mathbf{k}) \frac{k^2 dk}{(2\pi)^5}, \tag{1}$$

where $\mathbf{\Omega}$ is the unit vector in the $\mathbf{k}$ direction.

This approach can predict fundamentally new effects in resonance radiation transport processes in dense media. For instance, under conditions of thermodynamic equilibrium in an unbounded medium the function $J(\omega, \mathbf{k})$ can be determined from the fluctuation–dissipation theorem for the spectral density of the fluctuations of the transverse electromagnetic field strength:

$$J(\omega, \mathbf{k}) = \frac{8\pi\hbar\omega^4 n(\omega)\varepsilon''}{|\omega^2\varepsilon - c^2 k^2|}, \tag{2}$$

where $\hbar$ is Planck's constant, and $n(\omega)$ are the equilibrium photon occupation numbers given by the Planck formula:

$$n(\omega) = (\exp(\hbar\omega/T) - 1)^{-1}. \tag{3}$$

Substituting Eq. (2) into Eq. (1) and integrating, assuming that $\varepsilon(\omega, \mathbf{k})$ depends weakly on the wave number (which is valid in dense media since the collisional width Γ under these conditions is much greater than the Doppler width), we obtain for an equilibrium highly absorbing medium

$$J_\omega \simeq \frac{\hbar\omega^3 n(\omega)}{4\pi^3 c^2}\frac{\varepsilon'}{2}\left(1 + \frac{2}{\pi}\arctan\frac{\varepsilon'}{\varepsilon''} + \frac{4\varepsilon''}{\varepsilon'}\ln\left(\frac{\Gamma}{\omega v_T/c}\right)\right)$$
$$\equiv J_\omega^{\text{Pl}}\psi(\omega)\varepsilon', \tag{4}$$

where J_ω^{Pl} is the Planck intensity for black body radiation in vacuum and v_T is the thermal atomic velocity. Since at high densities N the width Γ is determined by resonant collisions by the Vlasov–Fursov mechanism[14] ($\Gamma \propto N$), as a result of the presence of a logarithmic contribution in Eq. (4) which increases with N, the equilibrium intensity in an absorbing medium may be almost an order of magnitude higher than the radiation intensity in a transparent medium determined by the Clausius formula $J_\omega^{\text{Pl}}\varepsilon'$ within the spectral line profile for detuning $\Delta\omega \lesssim \Gamma$. Note that the numerical calculations of J_ω for sodium vapor assuming that ε depends on ω and k show that the approximation (4) is quite accurate.[11]

These effects may be observed to a considerable extent in observations of the thermal radiation emitted by a dense medium. Calculations of the intensity of this radiation should also be made using the function $J(\omega, k, \mathbf{\Omega}, \mathbf{r})$ introduced earlier. In the steady-state case, the generalized spectral intensity of the radiation $J(\omega, k, \mathbf{\Omega}, \mathbf{r})$ should satisfy two equations simultaneously. One of these has the form of the kinetic equation:

$$(\mathbf{\Omega}, \nabla)J = -k_\omega J + \tilde{\varepsilon}(\omega, k, \mathbf{\Omega}, \mathbf{r}), \tag{5}$$

where k_ω is the absorption coefficient and $\tilde{\varepsilon}$ is the generalized spectral intensity of the bulk spontaneous emission:[11]

$$k_\omega = \frac{\omega^2\varepsilon''(\omega, k)}{c^2 k} = \frac{4}{3}\frac{g_2}{g_1}\frac{(d\pi\omega)^2}{\hbar c^2 k}a(\omega, k)$$
$$\times\left(\tilde{N}_1 - \frac{g_1}{g_2}\tilde{N}_2\exp(-\hbar(\omega - \omega_0)/T)\right), \tag{6}$$

$$\tilde{\varepsilon} = \frac{4}{3}\frac{d^2\omega^3\hbar\omega}{\hbar c^2 k}\frac{\omega^2\varepsilon''}{|\omega^2\varepsilon - c^2 k^2|^2}(2\pi)^3\tilde{N}_2 a(\omega, k)$$
$$\times\exp\left(-\frac{\hbar(\omega - \omega_0)}{T}\right). \tag{7}$$

In Eqs. (6) and (7), $a(\omega, k)$ is the generalized line profile, $g_{1,2}$ are the statistical weights of the ground and excited states, $\tilde{N}_{1,2}$ are their effective populations (see below) which satisfy the Boltzmann relations in equilibrium and may be obtained from the kinetic equations given in Ref. 10 in the absence of equilibrium. Integrating Eq. (7) over k, neglecting the spatial dispersion (where $a(\omega, k) \simeq a(\omega)$), we can obtain a clear expression for the spectral intensity of the spontaneous emission in a highly absorbing medium:

$$\varepsilon_\omega = \frac{1}{4\pi} \hbar\omega A_0 \left[\frac{\omega}{\omega_0}\right]^3 \mathrm{Re}(\varepsilon^{1/2}) \tilde{N}_2 a(\omega)$$

$$\times \exp\left(-\frac{\hbar(\omega-\omega_0)}{T}\right). \tag{8}$$

It follows from Eq. (5) that in general there is no closed equation for the spectral intensity J_ω defined in Eq. (1) and in addition to the first-order partial differential equation (5), the function $J(\omega,k,\boldsymbol{\Omega},\mathbf{r})$ should also satisfy an inhomogeneous wave equation, which has the following form in the steady-state case

$$\left(-\frac{c^2}{2}\Delta + 2(c^2k^2-\omega^2\varepsilon')\right)J = \frac{8\pi\hbar\omega^4\varepsilon''}{|\omega^2\varepsilon-c^2k^2|}$$

$$\times \frac{2(c^2k^2-\omega^2\varepsilon')}{((\tilde{N}_1/\tilde{N}_2)(g_2/g_1)\exp(-\hbar(\omega-\omega_0)/T)-1)}. \tag{9}$$

In the homogeneous case, Eq. (9) yields the same result (2) as the kinetic equation (5). In the general case, Eqs. (5) and (9) must be solved jointly.

In order to close the system of equations (5) and (9), they must be supplemented by the equations for the populations $\tilde{N}_i$. In the limit $\tilde{N}_2 \ll \tilde{N}_1 \simeq N$ we can obtain the following equation for $\tilde{N}_2$ (assuming that complete frequency redistribution occurs in a dense medium):

$$\frac{2}{3}\frac{d^2}{\hbar^2}\frac{g_2}{g_1}\int \frac{d\omega\, d\boldsymbol{\Omega} k^2 dk}{(2\pi)^3} J(\omega,k,\boldsymbol{\Omega},\mathbf{r}) a(\omega,k)\tilde{N}_1$$

$$-\frac{8}{3}\frac{d^2}{\hbar}\int \frac{d\omega\, k^2 dk}{\pi} \frac{\omega^4\varepsilon''(\omega,k)}{|\omega^2\varepsilon-c^2k^2|^2} a(\omega,k)$$

$$\times \exp\left(-\frac{\hbar(\omega-\omega_0)}{T}\right)\tilde{N}_2 - W\left(\tilde{N}_2 - \tilde{N}_1 \frac{g_2}{g_1}\right.$$

$$\left.\times \exp\left(-\frac{\hbar\omega_0}{T}\right)\right) = 0. \tag{10}$$

The first term in Eq. (10) corresponds to the photoabsorption of radiation, the second corresponds to the spontaneous decay of excited atoms, and the third corresponds to collisional exchange between states 1 and 2 with the probability W. The boundary conditions satisfied by the intensity at the interface may be derived from an analysis of the emission from a heated medium into vacuum. In particular, for the radiation intensity $J_\omega^{\mathrm{ex}}(\boldsymbol{\Omega})$ from a heated medium having the temperature T_0, the reflection coefficient $R=R_0$, the permittivity $\varepsilon=\varepsilon_0$, and the function $\psi=\psi_0$ at the boundary (see Eq. (4)), we obtain

$$J_\omega^{\mathrm{ex}}(\boldsymbol{\Omega}) = \frac{J_\omega^{(-)}(\boldsymbol{\Omega})|_{z=0}}{\varepsilon_0'(\omega)\psi_0(\omega)}(1-R_0(\omega,\boldsymbol{\Omega})), \tag{11}$$

where $J_\omega^{(-)}$ is the intensity of the radiation emitted from the medium to the interface. Note that in a uniform half-space, relation (11) generalizes the Kirchhoff law for the radiation intensity emitted by an equilibrium medium:

$$J_\omega^{\mathrm{ex}} = J_\omega^{\mathrm{Pl}}(I-R). \tag{12}$$

The exponential factor $\exp(-\hbar(\omega-\omega_0)/T)$ appears in Eqs. (6)–(10) (the need to introduce this was noted in Refs. 15–17) as a result of generalizing the standard radiation transport theory[8,9] to the case of broad lines characteristic of a dense medium and was was justified in Ref. 10 in terms of the Keldysh theory using kinetic Green's functions.[18–21] We shall now consider the case of an equilibrium medium of two-level atoms emitting thermal radiation which, for simplicity, we shall describe as Planck.

The equilibrium condition implies that the rates of radiative decay of the spectral density of excited atoms and photoabsorption from the ground state are equal. We write this condition in terms of the photon and atomic occupation numbers taking into account the recoil momenta of the photons and the kinetic motion of the atoms, and also quantum statistics effects (see Ref. 10):

$$\tilde{N}_2(p)a_2(\varepsilon)\int d\omega_k d\boldsymbol{\Omega}\omega_k^3(1+n_k)[1-\tilde{N}_1(p-\hbar k)]$$

$$\times a_1(\varepsilon-\omega_k-\omega_0+E(\mathbf{p})/\hbar-E(\mathbf{p}-\hbar\mathbf{k})/\hbar)$$

$$=[1-\tilde{N}_2(p)]a_2(\varepsilon)\int d\omega_k d\boldsymbol{\Omega}\omega_k^3 n_k \tilde{N}_1(p-\hbar k)$$

$$\times a_1(\varepsilon-\omega_k+\omega_0+E(\mathbf{p})/\hbar-E(\mathbf{p}-\hbar\mathbf{k})\hbar). \tag{13}$$

Here we have written $\omega_k = c|\mathbf{k}|$, n_k is the photon occupation number:

$$n_k = (\exp(\hbar\omega_k/T)-1)^{-1}, \tag{14}$$

$\tilde{N}_i(p)$ are the generalized occupation numbers of the ith atomic state (to be specific we assume Fermi statistics for the atomic states), $E(p)=p^2/2m$ is the overall kinetic energy of the atom, and $a_i(\varepsilon)$ is the spectral density (profile) of the ith state ($i=1,2$),

$$a_i(\varepsilon) = \frac{\gamma_i/2\pi}{\varepsilon^2+(\gamma_i/2)^2}, \tag{15}$$

where γ_i is the width of this state. The following notation appears in Eq. (15)

$$\hbar\varepsilon = \hbar\omega - \hbar\omega_i - E(\mathbf{p}) + \mu, \tag{16}$$

$E_i = \hbar\omega_i$ is the energy of the ith state, μ is the chemical potential, $\omega_0 = \omega_2 - \omega_1$, i.e., ε is the frequency detuning from resonance in state i which is determined by the energy $\hbar\omega_i + E(\mathbf{p}) - \mu$.

Condition (13) is identically satisfied if $\tilde{N}_i(p)$ has the form (see Ref. 10)

$$\tilde{N}_i(p) = \frac{1}{\exp((\hbar\varepsilon+\hbar\omega_i+E(\mathbf{p})-\mu)/T)+1}$$

$$\equiv \frac{1}{\exp(\hbar\omega/T)+1}. \tag{17}$$

We shall consider the case where the ground state ($i=1$) is unbroadened, i.e., $a_1(\varepsilon)=\delta(\varepsilon)$. Then, allowing for Eq. (17), it follows from Eq. (13) that the rate of photon drift in the $\boldsymbol{\Omega}$ direction will be proportional to $R_\downarrow$:

$$R_\downarrow = \frac{1}{\exp((\hbar\varepsilon + \hbar\omega_2 + E(\mathbf{p}) - \mu)/T) + 1} a_2(\varepsilon)(1 + n_k)$$

$$\times \left[1 - \frac{1}{\exp((\hbar\varepsilon + \hbar\omega_1 + E(\mathbf{p}) - \hbar\omega_k + \hbar\omega_0)/T) + 1} \right]. \tag{18}$$

Here ω_k and ε, being the detuning from the position of level 2, are related by (see Eq. (16))

$$\hbar\varepsilon = \hbar\omega_k - \hbar\omega_0 - E(\mathbf{p}) + E(\mathbf{p} - \hbar\mathbf{k}). \tag{19}$$

Introducing the auxiliary population $\tilde{N}_i$, defined as

$$\tilde{N}_i = \exp(-\hbar\omega_i/T)[\exp(\mu/T)/\lambda_T^3] \tag{20}$$

(where for simplicity the statistical weights g_i were assumed to be unity, and λ_T is the thermal de Broglie wavelength of the atom), and neglecting degeneracy, the rate of radiative decay of an atom in state 2 with the momentum $\mathbf{p}$ and detuning ε will have the form

$$R_\downarrow \simeq \tilde{N}_2 \lambda_T^3 \exp(-E(\mathbf{p})/T - \hbar\varepsilon/T)(1 + n_k) a_2(\varepsilon). \tag{21}$$

Similarly, the rate of photoexcitation of an atom from state $i = 1$ with the momentum $\mathbf{p} - \hbar\mathbf{k}$, using the same notation, will have the form

$$R_\uparrow \simeq \tilde{N}_1 \lambda_T^3 \exp(-E(\mathbf{p} - \hbar\mathbf{k})/T) n_k a_2(\varepsilon). \tag{22}$$

Equating expressions (21) and (22), we obtain (taking account of Eqs. (19) and (20))

$$\frac{1 + n_k}{n_k} = \exp\left(\frac{\hbar\omega_k}{T} \right)$$

$$\equiv \exp\left[\frac{\hbar\varepsilon + \hbar\omega_0 + E(\mathbf{p}) - E(\mathbf{p} - \hbar\mathbf{k})}{T} \right]. \tag{23}$$

Relation (19) for the photon frequency ω_k implies that for the broadened atomic state 2, a photon emitted from the line limb ($\varepsilon \neq 0$) possesses an energy which is determined not only by the transition frequency ω_0 and the atomic recoil energy $E(\mathbf{p}) - E(\mathbf{p} - \hbar\mathbf{k})$ but also by the detuning ε.

If the rate of spontaneous decay is integrated over the momenta of the emitting atoms, after the term containing n_k is dropped Eq. (21) gives

$$\langle R_\downarrow \rangle_{\mathrm{sp}} = \tilde{N}_2 \exp(-\hbar(\omega_k - \omega_0)/T) \varphi(\omega_k - \omega_0), \tag{24}$$

where $\varphi(\omega_k - \omega_0) \equiv \varphi(\Delta)$ is the Voigt profile of the emission line, which in general is a convolution of the profiles a_2 and a_1. Expression (24) contains the correction factor $\exp(-\Delta/T)$ omitted in the standard theory of resonance radiation transport, which is usually close to unity near the line center, $\Delta \ll T$ but is essential to describe the radiation at the far limbs of the line ($\Delta \geqslant T$). In the resonance approximation, the quantity ε/T in the expression for the spectral populations (see Eqs. (17), (21) and (23)) is negligible compared with unity and the factor $\exp(-\Delta/T)$ automatically becomes unity. Note that in the derivation of these relations the specific homogeneous broadening mechanism is unimportant since γ_2 can be determined even by a purely radiative broadening mechanism.

The appearance of this exponential factor can be explained qualitatively as follows. In the narrow-line approximation, $\Delta\omega \ll \omega_0$ and in a low-density medium the spectral intensity of the bulk emission from a resonant medium is described by[8] (compare with expression (8))

$$\varepsilon_\omega = (1/4\pi)\hbar\omega A N_2(\omega), \tag{25}$$

where $N_2(\omega)$ is the spectral population density of the resonant level, i.e., the density of atoms capable of emitting a photon at the frequency ω (in the approximation of complete frequency redistribution of absorbed and emitted photons, it is assumed that $N_2(\omega) = N_2 a(\omega)$, where N_2 is the total resonant level population). It is known that in the standard theory of resonance radiation transport[8] the absorption coefficient is given by the following expression:

$$k_\omega = A \pi^2 (c^2/\omega^2)(g_2/g_1) a(\omega)(N_1 - (g_1/g_2)N_2). \tag{26}$$

Assuming that in equilibrium $N_2 = (g_2/g_1)N_1 \times \exp(-\hbar\omega_0/T)$ (Ref. 8), we obtain the following expression for the ratio $\varepsilon_\omega/k_\omega$:

$$\varepsilon_\omega/k_\omega = (\hbar\omega^3/4\pi^3 c^2)(\exp(\hbar\omega_0/T) - 1)^{-1}. \tag{27}$$

At equilibrium however, the ratio $\varepsilon_\omega/k_\omega$ should be equal and identical to the Planck intensity J_ω^{Pl}, which in turn is determined by an expression similar to (27) but whose exponential function contains the instantaneous frequency ω rather than the resonance frequency ω_0:

$$J_\omega^{\mathrm{Pl}} = (\hbar\omega^3/4\pi^3 c^2)(\exp(\hbar\omega/T) - 1)^{-1}. \tag{28}$$

In addition to this anomaly, there is also another. It is easy to see that the total radiation intensity integrated over all frequencies using expression (25) is infinite. Even if the dependence of the spontaneous emission probability on ω is neglected ($A = A_0(\omega/\omega_0)^3$, where A_0 is the classical probability of spontaneous emission in vacuum, see Refs. 10 and 11), and for example, for purely radiative broadening, the integral of ε_ω diverges logarithmically at high frequencies where $a(\omega) \propto \omega^{-2}$.

In order to obtain an accurate formula for J_ω^{Pl} and ensure convergence of the appropriate integral, the expression for $N_2(\omega)$ must be modified by introducing the correction factor

$$\tilde{N}_2(\omega) = \tilde{N}_2 a(\omega) \exp(-\hbar(\omega - \omega_0)/T), \tag{29}$$

where, as was shown in Ref. 10, expression (29) will not contain the true populations but the effective (auxiliary) populations $\tilde{N}_2$. In equilibrium we obtain the Boltzmann spectral distribution,

$$\tilde{N}_2(\omega) = \tilde{N}_1 a(\omega) \exp(-\hbar\omega/T) \tag{30}$$

(in the absence of equilibrium the effective populations $\tilde{N}_i(\omega)$ are obtained from the kinetic equations given in Ref. 10). In this case, the true total population of excited atoms is determined by the integral of Eq. (30) over frequency and for pure radiative broadening may be written in the form

$$N_2 \simeq \tilde{N}_1 \left\{ \exp\left(-\frac{\hbar\omega_0}{T} \right) + \frac{3!}{2\pi} \frac{A_0}{\omega_0} \left[\frac{T}{\hbar\omega_0} \right]^4 \right\}. \tag{31}$$

Interestingly, a similar sum of exponential and power contributions was obtained for the particle momentum distribution.[22] In Refs. 15 and 16, where an expression similar to (29) was used, it was assumed when integrating over frequency that the total population is determined by the first Boltzmann term in expression (31), which may lead to major errors. It follows from expression (31) that, strictly speaking, even in equilibrium the true population N_2 will not satisfy the Boltzmann relation and will only be the same as $\tilde{N}_2$ for a narrow line $\hbar\Gamma \ll T$. However, a detailed analysis[10] shows that Eqs. (6)–(10) actually contain $\tilde{N}_2$ and thus we define this as the effective population.

The introduction of the factor $\exp(-\hbar(\omega-\omega_0)/T)$ negligibly influences the nuclear line profile $\Delta\omega \propto \Gamma$ (although for a fairly broad line it increasingly enhances the asymmetry of the absorption and emission lines in a dense medium) but for $T \ll \omega_0$ may result in the appearance of a second peak on the far red limb of the line profile (where $\omega \ll \omega_0$) if the gas temperature is fairly low. Under these conditions we have

$$\varepsilon_\omega \propto \omega^4 a(\omega)\exp(-\hbar(\omega-\omega_0)/T)$$

(see Eq. (8)), and provided that the profile does not decrease too rapidly (power law) in the range of low-frequency detuning, the high value of the exponential factor ensures that an additional peak appears. We postulate that the resonant line broadening is created by the gas itself (and for large detuning the quasistatic profile $a(\omega) \propto \Delta^{-2}$) or by a buffer gas ($a(\omega) \propto \Delta^{-3/2}$), Ref. 14. It is then easy to show that the frequency ω_m corresponding to this maximum is determined in both cases by the temperature of the medium and is approximately $4T/\hbar$. This additional peak will occur in the temperature range $T \lesssim \hbar\omega_0/11$, and its intensity will be proportional to T^4. The intensity of the low-frequency peak will clearly differ appreciably in these two cases.

Under these conditions, detailed calculations (see below) show that the intensity of the far limb of the spectral line may appreciably exceed the intensity calculated by standard theory[8] and most of the energy emitted on the line may be assigned to its nonresonant red limb rather than to the central near-resonant component. This phenomenon was arbitrarily described in Ref. 13 as an infrared catastrophe. Moreover, even in the absence of a peak at relatively high temperatures, the emission intensity at the far limb will still be substantially higher than the intensity calculated by standard theory. Note that although the introduction of an exponential factor is not directly related to effects caused by the density of the medium, in reality its influence on the radiation intensity in various spectral ranges and on the ratio between them will nevertheless be determined by the concentration of interacting particles, since we have $\varepsilon_\omega \propto a(\omega)$ and the line profile in the impact and quasistatic cases is proportional to this concentration at a given radiation frequency.

The thermodynamic relations do not indicate the specific mechanism for the establishment of the Boltzmann distribution (30). However, it is easy to show that this distribution can be formed in various elementary events, in interaction between the atomic system and equilibrium radiation and in collisions. In this last case, for example, we can consider the collision between emitting and buffer atoms accompanied by virtual excitation and subsequent emission of a photon at frequency ω far from the resonance frequency ω_0 (an analog of polarizing radiation[23,24]). It is easy to see that the intensity of the spontaneous emission of photons in these collisions will be proportional to $\exp(-\omega/T)$. However, if heavy particles collide, as a result of the Massey criterion $a\omega/v \gg 1$ (where a is the characteristic scale of the particle interaction radius and v is their relative velocity), the probability of this process will be exponentially low and proportional to $\exp(-a\omega/v)$.

We shall analyze another radiation-collisional process, emission of a photon at the far quasistatic limb of the line of an excited particle whose interaction potential with a buffer atom is denoted by $U_2(R)$. For atoms in the ground state the interaction potential is denoted by $U_1(R)$. Using the Franck–Condon principle, we can then write the following condition for the frequency of the emitted photon:

$$\hbar\omega = \hbar\omega(R) = E_2 - E_1 + U_2(R) - U_1(R). \tag{32}$$

At the far limb of the line, we can neglect the Doppler shift and we can consider the line profile $a_2(\varepsilon)$ caused by interaction between quasistatic particles to be proportional to a δ-function. In this approximation, we obtain for the bulk intensity of the spontaneous emission (see Refs. 15 and 16)

$$\varepsilon_\omega \propto N_1(R)N_b \exp[-\hbar\omega(R)/T]|R^2 dR/d\omega|. \tag{33}$$

Here

$$N_1(R) \propto \exp(\mu/T)\exp[-(E_1 + U_1(R))/T]$$

is the density of ground-state atoms interacting with buffer particles (having the density N_b) located at the distance R, which corresponds to the emission of a photon of frequency $\omega(R)$. The derivative $dR/d\omega$ is obtained from relation (32). It can be seen that in this model example at the far quasistatic limb of the line, the emission probability is proportional to the Boltzmann factor $\exp(-\hbar\omega/T)$ and the quasistatic line profile determined by the last cofactor in Eq. (33).

3. NUMERICAL SIMULATION OF RESONANCE RADIATION TRANSPORT PROCESSES IN DENSE SODIUM VAPOR

A numerical solution of the system of equations (5), (9), and (10) with the boundary condition (11) was used to calculate the radiation intensity J_ω of the resonant sodium lines ($\omega_0 \simeq 3.2 \times 10^{15}$ s^{-1}) both inside and at the exit of a nonuniformly heated planar layer of sodium vapor of length L in the presence of an argon buffer gas at a pressure $P \simeq 1$ atm. In the calculations it was assumed that near resonance the broadening is determined by a collisional mechanism with resonant collisions predominating at high temperatures. The collisional widths for the D_1 and D_2 sodium lines are given in Ref. 25 and are

$$\Gamma_{D_1} = 2\pi \times 4.68 \times 10^{-8} N \text{ s}^{-1} \quad \text{and}$$

$$\Gamma_{D_2} = 2\pi \times 7.62 \times 10^{-8} N \text{ s}^{-1}$$

At the line limb it was assumed that the profile $a(\omega)$ is formed by quasistatic argon broadening for which
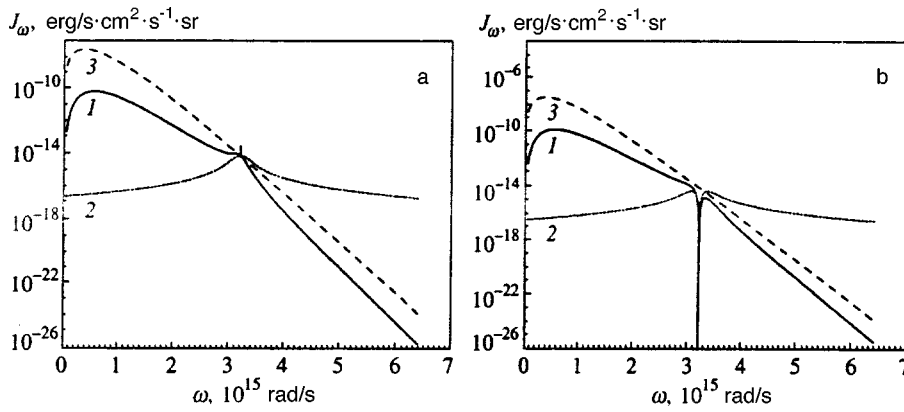
FIG. 1. Frequency dependence of the spectral radiation intensity at the center of a planar layer of sodium vapor (a) and of the radiation intensity leaving this layer (b); $T_c = 1000$ K, $L = 15$ cm; *1*—calculations using the theory put forward in Refs. 10 and 11, *2*—calculations using standard radiation theory,[8] and *3*—Planck intensity.

$$a(\omega) = (4\pi/6)(C_6)^{1/2} N_b \Delta \omega^{-1.5}$$

(see Ref. 26), where N_b is the buffer gas density. The van der Waals interaction constant $C_6 \simeq 1.9 \times 10^{-31}$ cm⁶/s was obtained from data on the argon collisional broadening of the sodium *D* lines.[27]

The results of the calculations are plotted in Figs. 1–3. Figure 1a gives the frequency dependence of the radiation intensity at the center of a layer of thickness $L = 15$ cm when the temperature at the center is $T_c = 1000$ K. In this case, the temperature profile in the layer corresponds to the experimental conditions (see below) and the Na atomic densities were assumed to correspond to the saturated vapor pressure at the given temperature.[28] Curve *1* was calculated using the theory put forward above, curve *2* was calculated using the standard theory of spectral line radiation,[8] and curve *3* gives the Planck intensity. The graph clearly shows that at the center of the line, curve *1* exhibits a narrow peak of intensity exceeding the Planck intensity, which is attributed to the strong dispersion and absorption in the dense sodium vapor, discussed in detail above. In the infrared (where $\omega = \omega_{max} \sim 5 \times 10^{14}$ s⁻¹) the calculations reveal a broad peak of intensity several orders of magnitude higher than that obtained in the standard theory at the same frequency. On the blue limb the opposite situation arises and, what is particularly important, the frequency integral of the intensity calculated using the standard theory diverges at high frequencies. The results show that in this theory the integrated intensity of the radiation spectrum at the given temperature is mainly determined by the red peak and the near-resonance component of the line

makes a relatively small contribution. Note that, as the calculations have shown, the optical layer thickness τ in the red peak is relatively small (for the conditions described above $\tau \sim 10^{-3}$) and "soft" photons can escape from the bounded medium.

Figure 1b shows the spectral dependence of the intensity of the thermal radiation from the cell under these conditions. It can be seen from a comparison of the curves that far from resonance the output radiation still shows the same dependence. For small detuning, however, the peak at the resonant frequency is replaced by a dip as a result of self-reversal although the calculations show that the intensity at the line center is still several orders of magnitude higher than that calculated by the standard theory.

Figure 2 gives a family of curves showing the change in the structure of the emission spectra at the center of a layer of thickness $L = 15$ cm and at the exit from this layer as a function of the temperature T_c (T_0 was taken to be 950 K in all cases). For the calculations the pressure of the sodium vapor and argon in the layer were assumed to be constant at $P_{Na} = 0.1$ atm and $P_{Ar} = 1$ atm. The temperature profile in the layer was assumed to be parabolic. It can clearly be seen that as the temperature increases, the red peak gradually becomes smoothed and disappears when $T_c > 2000$ K, although at high temperatures the contribution of the red limbs of the resonance line is still much greater than that calculated by the standard theory. To illustrate this, Fig. 3 gives the temperature dependence of the integrated luminescence intensity
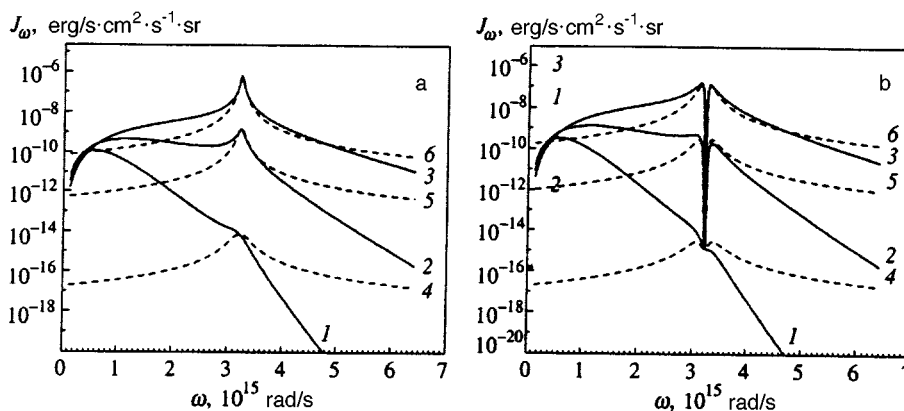

FIG. 2. Frequency dependence of the spectral intensities of the radiation at the center of a planar layer of sodium vapor (a) and radiation intensity leaving this layer (b), at various temperatures T_c; *1–3*—calculations using the theory put forward in Refs. 10 and 11, *4–6*—using standard radiation transport theory; $L = 15$ cm, $T_0 = 950$ K, $P_{Na} = 0.1$ atm, $P_{Ar} = 1$ atm, $T_c = 1000$ (*1,4*), 2000 (*2,5*), and 4000 K (*3,6*).

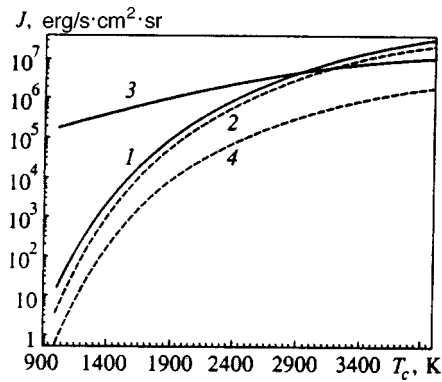FIG. 3. Dependence of the integrated intensity of the emission from a layer of sodium vapor in the visible (*1,2*), $0.4 < \lambda < 0.7$ μm and infrared (*3,4*), $\lambda > 0.7$ μm as a function of the temperature at the center of the layer; *1,3*—generalized radiation transport theory,[10,11] and *2,4*—standard theory: $L = 15$ cm, $T_0 = 950$ K, $P_{\mathrm{Na}} = 0.1$ atm, $P_{\mathrm{Ar}} = 0.1$ atm.
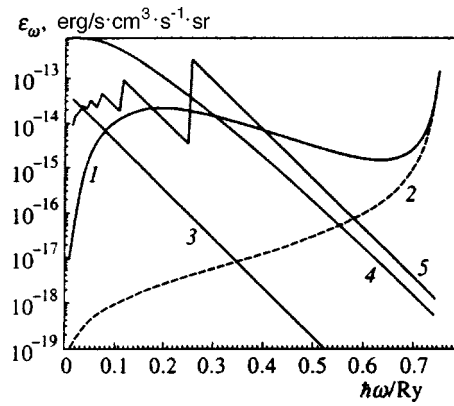

FIG. 4. Spectral emissivity of a hydrogen plasma: *1,2*—emission on L_α line, *3*—atomic bremsstrahlung, *4*—ion bremsstrahlung, *5*—photorecombination radiation: *1*—calculations using the theory put forward in Refs. 10 and 11, and *2*—using standard theory.

of the layer in the visible ($0.4 < \lambda < 0.7$ μm) and infrared ($\lambda > 0.7$ μm).

4. PLASMA RADIATION: COMPARATIVE ESTIMATES OF THE RATIO OF THE RADIATION INTENSITY OF RESONANCE LINES AND THE INTENSITY OF THE PHOTORECOMBINATION AND BREMSSTRAHLUNG CONTINUA

In the previous section we examined the structure of the spectra of pure thermal radiation from a resonant medium in the absence of ionization for the case of sodium vapor. In a plasma the luminescence intensity at the far limbs, although appreciable, will compete with the radiation intensity determined by other mechanisms, especially bremsstrahlung and photorecombination. The question therefore arises as to the ratio between these mechanisms and whether the radiation intensity at the limb of the resonant line can make a substantial contribution to the total plasma radiation. To solve this problem, we shall give fairly rough estimates of the contribution of the various mechanisms of plasma luminous excitance for specific examples.

As a first example we shall analyze a planar layer of equilibrium arc plasma in a sodium–mercury mixture with a characteristic length $L = 0.8$ cm having the following parameters. We shall assume that the gas temperature decreases parabolically from a maximum at the center of the layer $T_c(z=0) = 4000$ K to $T_0(z = \pm L/2) = 1500$ K at the boundary and the electron temperature is everywhere equal to the temperature of the heavy particles. The sodium (0.1 atm) and mercury (1 atm) vapor pressure is constant throughout and their atomic concentrations are determined by the gas temperature. (These parameters approximate to the conditions achieved in a high-pressure gas-discharge sodium lamp). Under these conditions the spectral intensity of the continuous spectrum comprising the sum of the photorecombination and bremsstrahlung continua, which was calculated using the Kramers formulas,[29] is constant for $\omega < \omega_0 \approx 3.2 \times 10^{15}$ rad/s and is approximately 1.5×10^{-10} erg/ s·cm^2·s^{-1}·sr. Calculations using the generalized theory showed that for the resonant sodium D line the intensity of

the far red limb begins to exceed the intensity of the continuum for $\omega > 0.6 \times 10^{15}$ rad/s. However, calculations using standard transport theory show that the line limb is only greater than the continuum for $\omega > 1.7 \times 10^{15}$ rad/s, and at this frequency the intensities of the line wing calculated using the standard and generalized theories differ by an order of magnitude.

As a second example, we shall consider the emission on the L_α line for a plasma containing hydrogen atoms or hydrogen-like ions and we shall compare the bulk emissivity ε_ω at the line limb with the continuum intensity. For multiply charged ions we also took into account the bremsstrahlung and recombination radiation (also in the Kramers approximation[29]) and the calculations were made for a plasma containing mainly bare nuclei, and hydrogen-like and helium-like ions (the ion composition was calculated using the Saha formula). Electron bremsstrahlung at atoms was also taken into account for hydrogen. The line profile for large detuning from resonance was calculated as the sum of the Stark and dispersion limbs.

The results of some of the calculations are plotted in Fig. 4 which gives the emissivity ε_ω of a hydrogen atom in an equilibrium isothermal plasma attributable to different mechanisms of emission at an electron density $N_e = 6.4 \times 10^{13}$ cm^{-3} and temperature 0.55 eV. These parameters correspond to the conditions achieved in the photosphere at the edge of the solar disk where the plasma is composed mainly of hydrogen (Harvard–Smithsonian standard model of the Sun's atmosphere[17]). On this graph, the ε_ω curve corresponding to emission on the line does not exceed the continuum at its maximum. However, at energies $\hbar\omega > \hbar\omega_1 = 0.412$ Ry ($\lambda_1 = 221.3$ nm) the intensity of the L_α line limb may appreciably exceed the continuum intensity. Note that at this point, the profile calculated by the standard theory gives an emissivity almost two orders of magnitude lower than that calculated using formula (8). However, even when the emissivity in the standard theory is comparable with the continuum (for $\hbar\omega_2 = 0.588$ Ry, $\lambda_2 = 155.1$ nm), the difference compared with formula (8) is still a factor of 20. Thus, even when the effects being discussed weakly influence the radia-
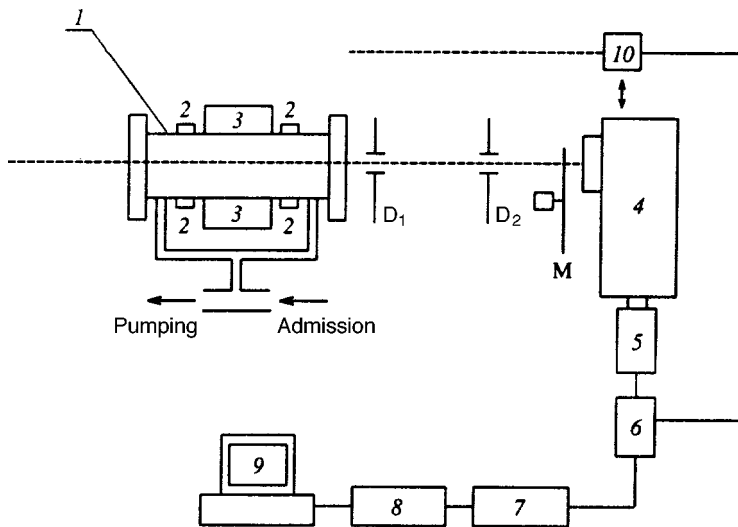
FIG. 5. Schematic of apparatus: *1*—thermal tube, *2*—water-cooled rings, *3*—heater, *4*—monochromator, *5*—radiation detector, *6*—amplifier, *7*—synchronous detector, *8*—analog–digital converter, *9*—computer, *10*—radiation detector to measure absolute integrated intensity in visible and infrared, $D_{1,2}$—aperture diaphragms, and M—modulator.

tion spectrum in the visible and infrared ranges, in the ultraviolet beyond the Balmer decrement $\hbar\omega > \mathrm{Ry}/4$, allowance for these effects can substantially alter the radiation balance, which may influence the output parameters of the semiempirical model of the solar atmosphere.

Calculations for multiply charged hydrogen ions for plasma parameters ($N_e = 10^8$ cm^{-3} and $T = 150$ eV) typical of the solar corona containing potassium ($Z = 20$) show that under these conditions the intensity in the red limb is much higher than the intensity calculated by the standard theory and the continuum intensity.

Thus, in many cases the mechanism for the formation of the spectral line limb examined in Sec. 2 can compete with the intensity determined by other radiation mechanisms (bremsstrahlung and photorecombination). This may prove to be an extremely important factor in calculations and interpretation of the luminescence measurements for various plasma objects. In addition, the appearance of a high-intensity red limb may strongly influence the processes of radiative filling of other levels and alter the plasma kinetics.

5. EXPERIMENTAL SETUP AND METHOD OF MEASURING THE SODIUM VAPOR LUMINESCENCE

Measurements of the spectra of pure thermal radiation made to check the theory developed above were carried out using a setup (see Fig. 5) consisting of a cylindrical ''thermal tube'' cell filled with sodium vapor and a highly sensitive recording apparatus. Prior to the experiment the cell (tube inner diameter 2.5 cm) containing metallic sodium was evacuated to a pressure of $\sim 10^{-5}$ Torr, filled with rare gas (argon or helium) at a pressure of 0.2–1 atm, and was then heated to the required temperature which was measured using various thermocouples positioned over the surface of the tube. For the experiments we used two cells with active zones (the section between the hydrogen-cooled flanges) of length $L = 15$ and 30 cm where the temperature distribution over the length was approximately parabolic. Note that at $T_c \sim 900$ K and above, yellow Na vapor luminescence is clearly visible through the exit window at the end of the cell.

The radiation in the axial zone of the tube, extracted via a sapphire window, was passed through a mechanical chopper which modulated the input light beam at a frequency of ~ 1000 Hz and was recorded directly by two detectors: an FÉU-84 photomultiplier in the visible range (in the 0.5–0.6 μm band cut out by a set of light filters) and a photodetector comprising a PbS photoresistor with a germanium window in the infrared (in the 2–3 μm band). The photoresistor was calibrated in terms of absolute intensity using radiation from a Globar heated to a known temperature, and the photomultiplier was calibrated using radiation from an SI-8-200 standard tungsten lamp. In both cases, the calibration accuracy was $\sim 50\%$. Note that the luminescence intensity integrated over these spectral ranges will subsequently be denoted as J_{ir} and J_v.

The luminescence spectra of the vapor in the 2–5 μm range were obtained by directing the radiation to the entrance slit of an infrared monochromator with an MG-30 pyroelectric detector positioned at the exit. For the spectral measurements in the visible and near infrared (in the wavelength range $\lambda \leqslant 1.2$ μm) we used another monochromator with another photomultiplier (FÉU-83 or FÉU-84) positioned at the exit, which was also calibrated over the spectrum using a tungsten lamp. The detector signals were amplified by narrow-band amplifiers and then passed to synchronous detectors from which the dc voltage was supplied to high-precision analog-digital converters connected to a computer. The spectral resolution in the experiments was generally 15 cm^{-1} in the visible and near infrared and ~ 300 cm^{-1} in the 2–5 μm range.

In the experiments, particular attention was paid to preventing direct radiation from the heated cell walls from entering the detector apertures. This was achieved by placing two aperture diaphragms between the detectors and the cell (see Fig. 5), their centers coinciding with the axis of the cell. These diaphragms, having aperture diameters $d_1 = 1 - 16$ and $d_2 = 5$ mm and separated by a distance of 43 cm, limited the solid angle of the radiation incident on the detectors. Measurements of the vapor luminescence intensity in the visible (J_v) and infrared (J_{ir}) showed that when d_1 was increased to
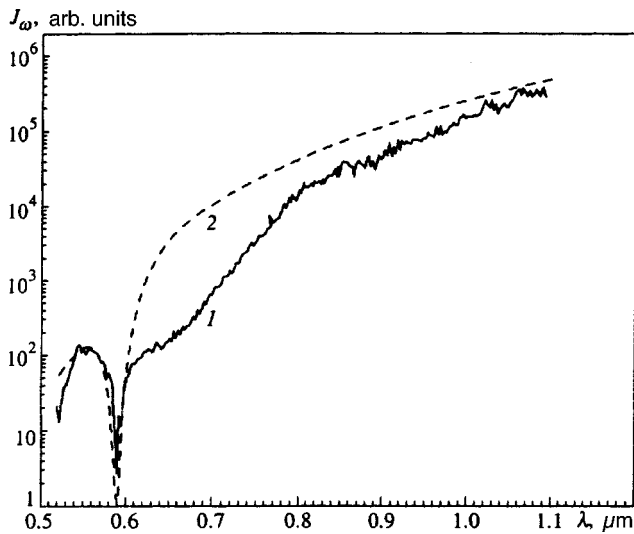
FIG. 6. Experimentally measured (*1*) and theoretically calculated (*2*) spectra of the luminescence of dense sodium vapor from a thermal tube in the range 0.5–1.1 μm; $L=15$ cm, $T_c=1140$.
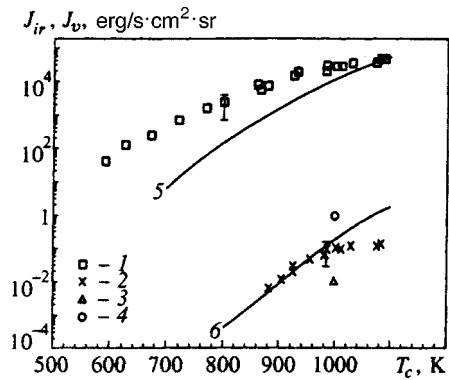

FIG. 7. Temperature dependence of the integrated luminescence intensities of sodium vapor in the infrared (*1,3,5*) and visible (*2,4,6*) at the center of the thermal tube: *1,2*—experiment, *3,4*—standard theory, *5,6*—theory from Refs. 10 and 11; $L=30$ cm.

10 mm, the values of J_v and $J_{\rm ir}$ are proportional to d_1^2 to a high degree of accuracy. In addition, in the absence of sodium in the cell, and regardless of whether buffer gas was present and its pressure, over the same range of variation of the diameter of the first diaphragm no signal was observed from the detectors (it was below the sensitivity of the recording apparatus) even at maximum heating temperatures. Thus, in all the experiments the value of d_1 was set below 6 mm.

6. RESULTS OF MEASUREMENTS AND COMPARISON BETWEEN EXPERIMENTAL DATA AND CALCULATED RESULTS

Figure 6 shows the characteristic self-reversed spectrum of thermal radiation from sodium vapor in the 0.5–1.1 μm range obtained using this apparatus. Also plotted is the theoretical curve calculated using Eqs. (5), (9), and (10). For the numerical simulation it was assumed, as was noted above, that the distribution of the sodium atomic density over the cell length is determined by the temperature distribution in accordance with the saturated vapor pressure curve, which is valid at least up to ~ 900 K (Ref. 30). On account of the high vapor and buffer gas density, and also the large ratio of cell length to diameter (~ 10), this approximation is completely accurate at higher temperatures. In the experimental measurements of the luminescence spectrum the relative luminescence intensity was recorded and thus, to permit comparison with the experimental data, the theoretical curve in Fig. 6 was normalized in terms of luminescence intensity to the short-wavelength peak of the measured spectrum. A comparison of the experimental data and the calculated results shows that the theoretical curve accurately describes the experimental data near resonance and quite satisfactorily, at long wavelengths. The appreciable discrepancy between the theoretical and experimental curves for $0.6<\lambda<0.8$ μm is evidently attributable to the calculations neglecting the influence of emission and absorption on electronic–vibrational transitions of molecular sodium Na$_2$, which is present in fairly large quantities in the bulk of the cell (on the order of a few percent of the atomic sodium density).

The most important factor observed theoretically and experimentally is the strong asymmetry of the thermal radiation spectrum of a resonant medium, which cannot be obtained in standard transport theory. Another predicted effect—an appreciable increase in the output intensity at the center of the line compared with that obtained from the standard theory—is difficult to check quantitatively using this experimental setup because of the inadequate sensitivity of the recording apparatus, since the radiation flux leaving the resonant medium at the center of the self-reversed dip, calculated using Eqs. (5), (9), and (10) remains extremely low.

Far from resonance, the effects associated with an excess of the radiation intensity above the Planck intensity become unimportant and the optical vapor density is much less than unity. However, for $\Delta\omega \gg \Gamma$ the exponential factor $\exp(-\hbar(\omega-\omega_0)/T)$ begins to play an important role in the radiation intensity in the infrared (see Sec. 2). The results of measurements of the absolute intensity of the thermal radiation of dense sodium vapor in the infrared and visible are plotted in Fig. 7 and the radiation in the 2–5 μm range for various T_c is shown in Fig. 8. Note that the radiation in such a far infrared limb of the spectral line has not been studied before (the measurements are usually confined to detuning $\leqslant 1000$ cm^{-1}—see Ref. 31—and in particular, no measurements have been made of the pure thermal radiation). The results plotted in Figs. 7 and 8 show that the intensity of the infrared radiation is several orders of magnitude higher than that near resonance and the infrared part of the spectrum exhibits a clearly defined peak which qualitatively confirms the theoretical reasoning. Figure 7 gives the integrated emission intensity calculated using the proposed theory in the spectral ranges investigated experimentally. It can be seen that the theoretical curve for the range 0.5–0.6 μm shows good agreement with the experiment. For the infrared the experimental and calculated data agree within an order of magnitude. Some discrepancy may be attributed to the imperfection of the theory, which describes the formation of the static limb extremely qualitatively at such large detuning.
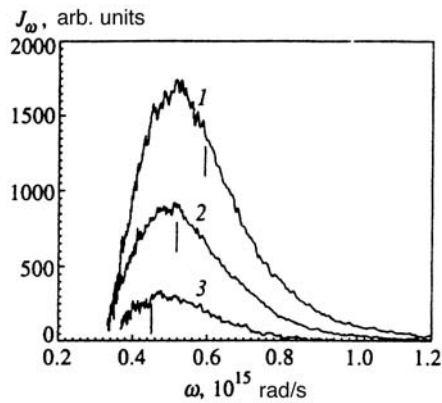
FIG. 8. Spectrum of thermal radiation from sodium vapor in the infrared at various temperatures in the center of the thermal tube: $T_c = 1080$ (*1*), 950 (*2*), and 830 (*3*) K. The vertical segments indicate the theoretical positions of the peaks, $L = 30$ cm.

Unfortunately, data on the potential curves of an Na–Ar system for interparticle distances corresponding to detuning $\Delta \sim 2 \times 10^{15}$ rad/s are not sufficiently reliable, and it is impossible to use more accurate data in the nearest-neighbor approximation. In addition to a knowledge of the potential curves at short distances, a quantitative description of the line profile for such large detuning also requires generalizing the broadening theory to the case of a multiparticle perturbation of the energy levels of the emitting particle by the buffer gas and itself.

It should also be noted that since the emissivity of a resonant medium is proportional to the line profile ($\varepsilon_\omega \propto a(\omega)$), the radiation intensity in the infrared should depend on th species of buffer gas and its pressure because at the quasistatic limb we find $a(\omega) \propto \sqrt{C_6}\, N_b$. The measurements showed that J_{ir} is approximately proportional to the buffer gas pressure and in the presence of helium at the same pressure, the radiation intensity in the infrared is an order of magnitude lower than that for argon. However, the value of $\sqrt{C_6}$ for helium is only 2.8 times lower than that for argon[32] which differs substantially from the experimentally measured ratio $J_{\mathrm{ir}}(\mathrm{Ar})/J_{\mathrm{ir}}(\mathrm{He}) \simeq 8$. Note that within measurement error the purity of the gas (which varied experimentally between 98 and 99.99%) did not influence the emission intensity of the sodium vapor in the visible and infrared. Moreover, the addition of air to the helium (in the ratio He:air $= 10:1$) had little influence on the measured radiation intensity.

The far limbs of the Na_2 molecular lines also make some contribution to the infrared emission of the vapor although these impurities are only present in small quantities, as has been noted above. Note that the limbs of the resonant transitions in the argon atoms make a negligible contribution because of their high excitation potential, which is confirmed by the disappearance of the luminescence signals in the absence of Na vapor. Another source of discrepancy between the experimental data and the calculated curves may be the inadequate accuracy in the calculations of the vapor density distribution over the cell length.

The integrated intensities obtained from the standard

theory differ markedly from those observed (see Fig. 7), being several times higher in the visible range and almost six orders of magnitude lower in the infrared. Thus, the experimental data fully agree with the theory developed in Refs. 10 and 11 in which the Boltzmann distribution of the spectral population of the resonant level is proportional to $\exp(-\hbar\omega/T)$ and confirms the existence of an infrared catastrophe in the radiation from a resonant medium.

7. CONCLUSIONS

A theory of the formation of the luminescence spectra of dense resonant media developed earlier[10,11] has been used for a numerical simulation of these effects and experimental data have been obtained on the structure of the spectra and the absolute intensity of the thermal radiation from dense sodium vapor in various spectral ranges. These data show fairly good qualitative agreement and in some cases, quantitative agreement with the results of the numerical calculations. The main results may be considered to be the experimental observation of a well-defined peak on the far red limb of the resonance sodium line, whose intensity is several orders of magnitude higher than the luminescence intensity near resonance, and also the observation that the self-reversed line is asymmetric. These data show that density effects play an important role in the energy balance of the emitting atomic and molecular systems and appreciably modify the established concepts adopted in the standard theory of resonance radiation transport[8] developed for low gas densities using the narrow line approximation.

The universal form of the spectral Boltzmann distribution, strictly substantiated theoretically in Refs. 10 and 11 and confirmed experimentally in the present paper, without claiming to provide a detailed description of the line profile for large detuning from resonance, may show up as an important factor in a wide range of different processes. For this reason, the results obtained here may prove important for various practical applications. For instance, we note that although the ratio of the integrated intensities in the visible and infrared changes with increasing temperature, at high temperatures the contribution of the red limbs of the resonant line still remains much higher than that calculated by the standard theory, as was noted earlier. Estimates made here show that this may be an important factor in calculations and measurements of the luminescence of various plasma formations (including gas-discharge light sources), in interpreting the results of measurements in the solar spectrum, and so on, since the luminescence intensity in the far red limbs may compete with the intensity determined by other radiation mechanisms (such as bremsstrahlung). Moreover, under certain conditions the high-intensity radiation in the red limb may substantially increase the radiative filling of other levels and thereby appreciably influence the kinetics of the plasma processes.

*⁾E-mail: staran@fly.triniti.troitsk.ru
†⁾E-mail: chekhov@post.mipt.rssi.ru

[1] R. Stamm, B. Talin, E. L. Pollock, and C. A. Iglesias, Phys. Rev. A **34**, 4144 (1986).

[2] A. Calisti, F. Khelfaoui, R. Stamm, and B. Talin, in *Spectral Line Shapes*, Vol. 6, edited by L. Frommhold and J. W. Kato (AIP Press, New York, 1990), p. 3.

[3] F. Khelfaoui, A. Calisti, R. Stamm, and B. Talin, *ibid*, p. 102.

[4] A. Calisti, R. Stamm, and B. Talin, Europhys. Lett. **4**, 1003 (1987).

[5] A. V. Anufrienko, A. L. Godunov, A. V. Demura *et al.*, Zh. Éksp. Teor. Fiz. **98**, 1304 (1990) [Sov. Phys. JETP **71**, 728 (1990)].

[6] A. V. Anufrienko, A. E. Bulyshev, A. L. Godunov *et al.*, Zh. Éksp. Teor. Fiz. **103**, 417 (1993) [JETP **76**, 219 (1993)].

[7] A. E. Bulyshev, A. V. Demura, V. S. Lisitsa *et al.*, Zh. Éksp. Teor. Fiz. **108**, 212 (1995) [JETP **81**, 113 (1995)].

[8] L. M. Biberman, V. S. Vorob'ev, and I. T. Yakubov, *Kinetics of Nonequilibrium Low-Temperature Plasmas* (Consultants Bureau, New York. 1987) [Russ. original, Nauka, Moscow, 1982].

[9] L. A. Apresyan and Yu. A. Kravtsov, *Theory of Radiative Transfer* [in Russian], Nauka, Moscow (1983).

[10] Yu. K. Zemtsov and A. N. Starostin, Zh. Éksp. Teor. Fiz. **103**, 345 (1993) [JETP **76**, 186 (1993)].

[11] Yu. K. Zemtsov, A. Yu. Sechin, and A. N. Starostin, Zh. Éksp. Teor. Fiz. **110**, 1654 (1996) [JETP **83**, 909 (1996)].

[12] Yu. K. Zemtsov, A. Yu. Sechin, A. N. Starostin *et al.*, JETP Lett. **65**, 13 (1997).

[13] Yu. K. Zemtsov, A. Yu. Sechin, A. N. Starostin *et al.*, JETP Lett. **65**, 839 (1997).

[14] I. I. Sobelman, *Introduction to the Theory of Atomic Spectra* (Pergamon Press, Oxford, 1973) [Russ. original, Fizmatgiz, Moscow, 1966].

[15] A. V. Phelps, *Tunable Gas Laser Utilizing Ground State Dissociation*, JILA Report No. 110 (University of Colorado, Boulder, CO, 1972).

[16] G. York and A. Gallagher, *Power Gas Laser on Alkali-Dimers A-X Band Radiation*, JILA Report No. 110 (University of Colorado, Boulder, CO, 1974).

[17] D. Mihalas, *Stellar Atmospheres*, 2nd ed. (Freeman, San Francisco, 1978) [Russ. transl., Mir, Moscow, 1982].

[18] L. V. Keldysh, Zh. Éksp. Teor. Fiz. **47**, 1515 (1964) [Sov. Phys. JETP **20**, 1018 (1964)].

[19] E. M. Lifshitz and L. P. Pitaevskiĭ, *Physical Kinetics* [Pergamon Press, Oxford, 1981) [Russ. original, Nauka, Moscow, 1979].

[20] B. A. Veklenko and G. B. Tkachuk, Opt. Spektrosk. **38**, 1132 (1975) [Opt. Spectrosc. **38**, 653 (1975)].

[21] V. A. Makhrov, A. Yu. Sechin, and A. N. Starostin, Zh. Éksp. Teor. Fiz. **97**, 1114 (1990) [Sov. Phys. JETP **70**, 623 (1990)].

[22] V. M. Galitskiĭ and V. V. Yakmets, Zh. Éksp. Teor. Fiz. **51**, 957 (1966) [Sov. Phys. JETP **24**, 637 (1967)].

[23] V. A. Kas'yanov and A. N. Starostin, Zh. Éksp. Teor. Fiz. **48**, 295 (1965) [Sov. Phys. JETP **21**, 193 (1965)].

[24] J. Borisov and L. Frommhold, in *Phenomena Induced by Intermolecular Interaction*, edited by G. Birnbaum (Plenum Press, New York, 1985), p. 67.

[25] C. G. Carrington, D. N. Stacey, and J. Cooper, J. Phys. B: Atom. Mol. Phys. **6**, 417 (1973).

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

[27] R. H. Chatham, A. Gallagher, and E. L. Levis, J. Phys. B: Atom. Mol. Phys. **13**, 47 (1980).

[28] *Handbook of Physical Quantities* [in Russian], Énergoatomizdat, Moscow (1986).

[29] Ya. B. Zel'dovich and Yu. P. Raizer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena* (Academic Press, New York, 1968) [Russ. original, Nauka, Moscow, 1966, Chap. 1].

[30] D. I. Chekhov, Dissertation for Candidate of PhysicoMathematical Science [in Russian], Moscow Institute of Physics and Technology, Moscow (1994).

[31] M. J. Jongerius, in *Spectral Line Shapes*, edited by B. Wende (Berlin, 1980), p. 963.

[32] A. A. Radtsig and B. M. Smirnov, *Reference Data on Atoms, Molecules, and Ions* (Springer-Verlag, Berlin, 1985) [Russ. original, later ed., Énergoatomizdat, Moscow, 1986].

Translated by R. M. Durham

Compression and heating of spherical fusion targets by indirect (x-ray) drive on the Iskra-5 facility

F. M. Abzaev, C. A. Bel'kov, A. V. Bessarab, S. V. Bondarenko, A. V. Veselov,
V. A. Gaĭdash, G. V. Dolgoleva, N. V. Zhidkov, V. M. Izgorodin, G. A. Kirillov,
G. G. Kochemasov,*) D. N. Litvin, E. I. Mitrofanov, V. M. Murugov, L. S. Mkhitar'yan,
S. I. Petrov, A. V. Pinegin, V. T. Punin, A. V. Senik, and N. A. Suslov

*Russian Federal Nuclear Center—All-Russia Research Institute of Experimental Physics, 607190 Sarov,
Nizhniĭ Novgorod Region, Russia*

Experiments have been carried out using the Iskra-5 facility in order to study the behavior of x-ray targets in response to a highly symmetric x-ray field. Results are presented of experiments using targets in the form of a spherical copper hohlraum coated with gold on the inside, with six laser entrance holes and a glass microtarget filled with DT gas located at the center. In some experiments the central capsule was coated with a plastic ablator layer of varying thickness. An analysis of the experimental results showed that on the whole, they are satisfactorily described by spherically symmetric gasdynamic calculations. © *1998 American Institute of Physics.* [S1063-7761(98)01107-X]

1. INTRODUCTION

Achieving highly symmetric irradiation of a fusion target is a key problem in the inertial-confinement fusion program. The difficulties which have been encountered in attempts to achieve uniform direct laser irradiation of targets have now caused many researchers to pin their hopes of achieving the required degree of symmetry on the conversion of laser radiation into x-rays.[1,2] In this type of indirect-drive system, a DT-filled target is compressed by its interaction with an x-ray field generated inside an almost closed cavity or hohlraum.

Indirect drive has been used in experiments on the largest laser facilities—Nova (USA),[3] GEKKO-XII (Japan),[4] and Phebus (France).[5] A particularly comprehensive research program has been carried out at the Lawrence Livermore National Laboratory[1] where laser energy was converted into x-rays in a cylindrical hohlraum. For this type of hohlraum however, an acceptable level of irradiation symmetry can only be achieved as an average over time. Numerous experiments and a wide range of precision diagnostics are required to analyze the beam input regime inside the hohlraum and to select the final coupling-in system.

A more natural method of generating a symmetric x-ray field involves using a spherical hohlraum. A spherical hohlraum is particularly convenient when the elements focusing the beam on the target chamber are not assembled in two clusters as in Nova but are distributed more uniformly over the surface of the sphere, as in Iskra-5 (Ref. 6). Using this characteristic feature of the twelve-beam Iskra-5 facility, we carried out experimental investigations of the compression and heating of fusion targets by x-ray radiation generated inside a spherical hohlraum. The total laser radiation energy ($\lambda = 1.315$ μm) fed into the cavity was 10 kJ with a pulse length of 0.3–0.4 ns. By varying the thickness of the shell of

the fusion target, we could vary the fuel compression ratio and thereby check the influence of the asymmetry of the x-ray field on the neutron generation.

The results of these investigations are reported here. In Sec. 2 we estimate the energy characteristics and the degree of symmetry of the x-ray radiation inside the hohlraum. Section 3 is devoted to the experimental setup and a brief description of the diagnostic complex. Section 4 contains the experimental results which are discussed in detail and compared with the one-dimensional spherically symmetric calculations which are used as the basis to estimate the level of symmetry attained.

2. BASIC PROPERTIES OF THE X-RAY FIELD INSIDE A SPHERICAL HOHLRAUM

2.1 Energy balance

We shall consider a spherical hohlraum with $N \gg 1$ laser entrance holes, distributed on average uniformly over the surface of the sphere.

For the wavelength $\lambda_L = 1.315$ μm and laser radiation intensities $I_L \cong 10^{14}$–10^{15} W/cm² the light absorption coefficient k_a is appreciably less than unity ($k_a \simeq 0.2$–0.5). The accompanying multiple reflection of light from the inner walls of the hohlraum leads to symmetrization of the laser field inside the cavity which means that averaged (over the surface of the sphere) values of the intensity can be used to calculate the laser energy balance in the hohlraum.

If $I_0 = E_L/4\pi R_0^2 \tau_L$ is the average intensity of the primary laser radiation injected through the holes (E_L and τ_L are the energy and duration of the laser pulse, R_0 is the hohlraum internal radius), the intensity acting on the walls of the hohlraum allowing for multiple reflection is

$$I_g = \frac{I_0}{1-(1-k_a)(1-\beta)}, \qquad (1)$$

where β is the relative area of the laser entrance holes. If $k_a \ll 1$ and $\beta \ll 1$, then $I_g \gg I_0$. For these experiments we used $E_L \simeq 9$ kJ, $\tau_L \simeq 0.3$ ns, $R_0 = 1$ mm, and $I_0 = 2.4 \times 10^{14}$ W/cm^2. Since the intensity satisfied $I_g > I_0$, and the average angle of incidence of the laser radiation on the surface for the conditions obtaining in Iskra-5 is $30-40°$, the absorption coefficient should lie in the range $k_a \simeq 0.2-0.3$ (Ref. 7) where most of the energy ($\sim 30-50\%$) is absorbed by a resonance mechanism. The total energy absorbed by the hohlraum is

$$E_a = E_L \frac{(1-\beta)k_a}{1-(1-k_a)(1-\beta)}. \qquad (2)$$

Under our conditions for $\beta = 0.13$, $k_a \simeq 0.3$ ($I_g/I_0 \simeq 2.6$), we have $E_a/E_L \simeq 0.66$. For $E_L = 9$ kJ the hohlraum walls absorb an energy $E_a \simeq 6$ kJ.

The release of energy results in the formation of a low-density ($n_e \sim n_{cr} \simeq 10^{21}$ cm^{-3}), hot ($T_e \simeq 2-5$ keV) laser corona at the walls, which is a source of primary, comparatively hard ($\overline{h\nu} \simeq 1-2$ keV) x-rays.

If the fraction of energy dissipated in the formation of fast ions by resonant absorption is k_{fi} and the coefficient of primary conversion of laser energy into x-rays is k_γ, the energy of the primary x-rays is

$$E_\gamma = E_a(1-k_{fi})k_\gamma. \qquad (3)$$

Under our conditions $k_{fi} \simeq 0.3$ and $k_\gamma \simeq 0.5$ (Ref. 8), so that $E_\gamma \simeq 2.1$ kJ.

The primary x-rays, being absorbed by the dense layers of the walls, form a comparatively dense ($\rho \simeq 0.5$ g/cm^2), cold ($T_e \approx T_i \simeq 0.15-0.2$ keV) x-ray corona which is a source of almost equilibrium ($\overline{h\nu} \simeq 0.5$ keV) x-rays.

If the albedo of the x-rays at the hohlraum, averaged over the entire spectrum, is α and the intensity of the primary x-rays is $Q_\gamma = E_\gamma/4\pi R_0^2 \tau_L$, the intensity of the x-rays from the hohlraum walls will be

$$Q_r \simeq Q_\gamma \frac{(1-\beta/2)\alpha}{1-(1-\beta)\alpha}. \qquad (4)$$

For $\alpha \simeq 0.6-0.7$, $\beta = 0.13$, and $E_\gamma = 2.1$ kJ we have $Q_\gamma \simeq (0.66-0.94) \times 10^{14}$ W/cm^2. This flux corresponds to equilibrium radiation at the temperature $T_\gamma \simeq 160-175$ eV.

2.2 Symmetry of the x-ray field

The symmetry of the x-ray field acting on the fusion target results from the geometric factor (each surface element of the central capsule is exposed to radiation from almost half the surface of the hohlraum) and also as a result of multiple reflections of photons and multiple rescattering of x-ray quanta. As a result of this last process, the temperature T_γ of the quasiequilibrium radiation emitted by the hohlraum walls may be considered, to a first approximation, to be the same over all sections of the inner surface.

Asymmetry of the x-ray field is caused first, by the presence of the laser entrance holes and second, by the asymme-

try of the primary x-ray flux $Q_\gamma/2$ generated by the laser corona and propagating away from the hohlraum walls. Another important factor is the spatial scale of the inhomogeneities and their distribution over the surface of the sphere. These determine the amplitudes of the spherical harmonics P_{lm} of the asymmetric x-ray field propagating from the hohlraum walls toward the capsule. An expansion in harmonis can be used to calculate the degree of smoothing of the x-ray field as it propagates from the surface of the hohlraum sphere to the central capsule.[1] The smoothing coefficients $C_l(r/R)$ depend on the number of the spherical harmonic l and the ratio of the radii of the target r_0 and the hohlraum R_0. In our case, this ratio is initially $r_0/R_0 \simeq 0.14$. The coefficients of smoothing of harmonics with $l = 1,2,3,4,5,6$ are 0.73, 0.37, 0.099, 0.023, 0.018, and 0.008, respectively.

The harmonic spectrum of the radiation field is determined by the configuration and diameter of the entrance holes and the primary laser illumination spots. Imbalance of the beam energy (power) and also targeting misses relative to the nominal positions of the spots impair the symmetry of the field. The centers of symmetry of the six holes are distributed over the surface of the hohlraum near the positions corresponding to the vertices of a cube inscribed in the hohlraum. The maximum angular deflection is $13°$. Loss of symmetry is caused by the desire to avoid coupling of parasitic generation between different amplifier channels. The penalty for this is the appearance of second and third harmonics in the spectrum of the radiation field, in addition to the fourth harmonic characteristic of the symmetry of the cube.

The primary illumination spots on the inner surface of the hohlraum are positioned so that they also contribute to the second, third, and fourth harmonics, in phase with the contribution from the laser entrance holes. On the one hand, this increases the asymmetry and on the other, ensures that the inhomogeneity pattern varies negligibly with time.

In addition, the level of asymmetry obtained is matched with the level caused by the energy imbalance in the beams. For typical conditions, the energy imbalance in the beams exceeds the level of inhomogeneity by no more than a factor of 2.

A quantitative estimate of the degree of inhomogeneity was made using the following simple model. The absorbed energy distribution over the inner surface of the hohlraum was calculated by the Monte Carlo method. It was assumed that the critical surface is spherically symmetric (except for the entrance holes) and that the reflected radiation has both specular and diffuse components; the dependence of the absorption coefficient on the angle of incidence was also taken into account.

The intensity of the primary (nonequilibrium) radiation I_γ of the laser corona was assumed to be proportional to the intensity of the absorbed laser radiation I_a with the coefficient $k_\gamma = 0.5$. The quantity $I_a k_\gamma/2$ determined the intensity of the primary x-ray flux propagating from the walls to the capsule. The intensity of the quasiequilibrium radiation from the x-ray corona was determined using formula (4). For the calculations it was assumed that $k_\gamma = 0.5$ and $\alpha = 0.65$.

Table I gives various quantities characterizing the non-

TABLE I. Estimate of the nonuniformity of the capsule x-ray irradiation.

	ε_{max}, %	ε_{max}^I, %	ε_{rms}, %	$\bar{\gamma}_j (j=1,...,6)$					
Same laser radiation energy in channels	3.5	3.5	2.0	4.0×10^{-4}	1.4×10^{-2}	1.2×10^{-2}	6.0×10^{-3}	1.1×10^{-3}	6.3×10^{-4}
Spread over channels $\varepsilon_{rms}\simeq20\%$	6.0	6.6	2.8	8.8×10^{-3}	2.3×10^{-2}	1.2×10^{-2}	6.1×10^{-3}	1.2×10^{-3}	6.4×10^{-4}
Spread over channels $\varepsilon_{rms}\simeq20\%$	7.2	6.2	2.8	1.1×10^{-2}	2.2×10^{-2}	1.2×10^{-2}	6.1×10^{-3}	1.2×10^{-3}	5.9×10^{-4}

uniformity of the x-ray irradiation of a DT target: the maximum nonuniformities

$$\varepsilon_{max}=\frac{I_{max}-I_{min}}{2I}, \quad \varepsilon_{max}^I=\frac{\max|I-\bar{I}|}{\bar{I}}$$

(where $\bar{I}$ is the average radiation intensity over the surface of the sphere) and the rms nonuniformity ε_{rms}, and also the amplitudes $\tilde{\gamma}_j$ of the expansion of the normalized distribution $I(\theta,\varphi)$ in terms of spherical harmonics for equal energies in the channels and for the spread over the channels recorded in two experiments. The calculations are made for a ratio of capsule radius to hohlraum radius of 0.14 and a relative area of laser entrance holes of 0.13.

It can be seen from Table I that in the balanced case, the rms nonuniformity is 2%, and when the real energy imbalance in the beams is taken into account this increases to $\simeq3\%$. Significantly, these figure vary little with time. For example, for calculations with a reduced absorption coefficient such that the fraction of the laser radiation energy lost through the holes increased from 0.34 to 0.6 and the albedo was reduced from 0.65 to 0.35, which simulates the initial situation, the nonuniformity was almost the same ($\varepsilon_{rms}=1.9\%$) as in the nominal calculations. The reason for this is that the lower degree of x-ray symmetrization is balanced by a higher degree of equalization of the laser intensity.

Calculated distributions of the absorbed laser radiation at the inner wall of the hohlraum for equal energies in the channels are plotted in Fig. 1a. The corresponding pattern for the x-ray radiation incident on the central capsule is shown in Fig. 1b. The shades of gray show the drop in intensity from maximum (white) to minimum (black). It can be seen that unlike the absorbed laser intensity, the x-ray distribution on the capsule exhibits no higher harmonics. Allowance for the energy spread over the channels results in a loss of symmetry between the front and back of the target (the appearance of $l=1$ harmonics).

2.3. Laser illumination and fast ions

X-rays are not the only carrier of laser radiation energy to the DT capsule. A characteristic feature of specular reflection from an internal, ideally spherical, critical surface is that an inner spherical portion, of diameter $\varnothing_{sh}$ is in shadow (under our conditions $\varnothing_{sh}/\varnothing_{box}\simeq1/5$, where $\varnothing_{box}=2R_0$ is the diameter of the hohlraum). Large-scale perturbations of the critical surface caused by the presence of entrance holes and irradiation nonuniformity reduce the shadow volume.

Scattering at small-scale plasma inhomogeneities (of the order of the laser wavelength), resembling diffuse reflection from the surface, results in unavoidable illumination of the capsule in the diameter $\varnothing_{sh}$. However, calculations (see below) show that even when the energies of the x-ray and laser radiation incident on the target are the same, this process has little influence.

Another method of supplying energy is achieved by movement of the gold plasma toward the center. Since the velocity of the material ($V\simeq10^7-10^8$ cm/s) is substantially lower than the speed of light, there is a time interval $\tau_2\approx\varnothing_{box}/2V\simeq1$ ns (where $\varnothing_{box}\simeq2$ mm) during which the shell is not subjected to the action of the material flux from the hohlraum walls. For a laser pulse duration $\simeq0.3$ ns and a convergence time of $\simeq0.5$ ns the action of the material flux can be neglected. However, the comparatively long wavelength of iodine laser radiation ($\lambda=1.315$ μm) with laser radiation intensities $I_L\geqslant3\times10^{14}$ W/cm^2 gives rise to hot electrons with temperature $T_h\simeq20-50$ keV, which accelerate some of the Au ions to velocities $V\geqslant(3-5)\times10^8$ cm/s. Thus, the time τ_2 is reduced to $\simeq0.3$ ns so that for shells with a comparatively long convergence time ($\tau_f\geqslant0.5$ ns) the possibility of an additional energy input to the target shell as a result of energy transfer by fast ions must be taken into account.

3. AIM OF THE EXPERIMENTS AND EXPERIMENTAL SETUP

As was shown in the previous section, an x-ray flux acting on a central DT-filled capsule is not ideally symmetrical. Estimates indicate that the asymmetry is large-scale (the spherical harmonics in its spectral expansion have numbers $l\leqslant1-4$). The characteristic level of asymmetry is 3%. The asymmetry of the x-ray field causes asymmetry of the accelerating pressures and shell implosion velocities: $\Delta V/V\leqslant\Delta I/I$. As a result of compression, the deviation of the contact boundary from spherical is

$$\frac{\delta R}{R}\simeq\frac{\delta V}{V}\frac{r_0}{r_f},$$

where r_0 is the initial radius of the DT-filled capsule and r_f is the radius of the DT gas at the instant of maximum compression. If we impose the constraint that this deviation does not exceed 1/3, we find that this asymmetry should not be noticeable up to compressions

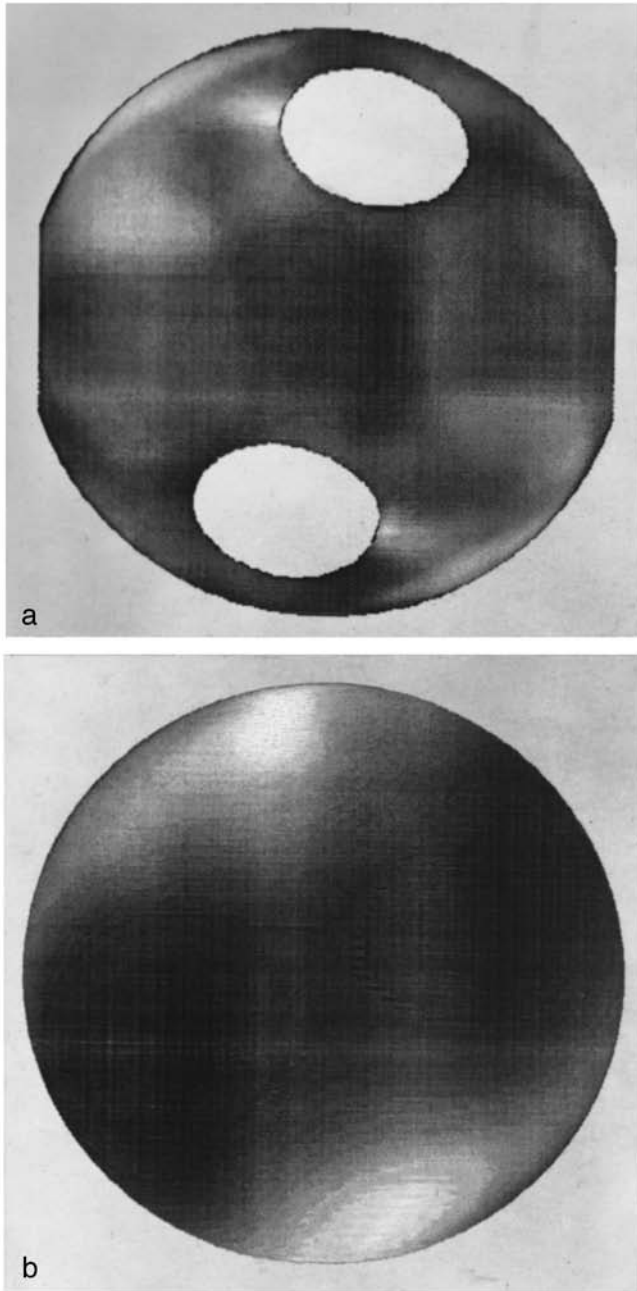$$\frac{r_0}{r_f}\leqslant\frac{1}{3}\frac{I}{\Delta I}\simeq11,$$

FIG. 1. Calculated laser illumination of the inner surface of the hohlraum (a) (the white spots correspond to the position of the laser entrance holes) and x-ray illumination of the central capsule (b) for typical experimental conditions.

which corresponds to volume compressions of $\simeq 10^3$.

This estimate is extremely approximate. In fact, the evolution of perturbations as a result of Richtmyer–Meshkov and Rayleigh–Taylor instabilities leads to a greater increase in the amplitude of the perturbations than that obtained from the above estimate. However, for those shells so thin that they can be completely heated under the action of the x-rays, the symmetry of the energy release may be better than 3%. For this scenario the heating pattern of the shells may be represented by a unilateral flux of such intensity that it can heat (in the supersonic regime) the entire shell during the pulse.

The influence of the residual asymmetry may be detected via the dependence of the neutron yield on the compression ratio when a certain critical compression δ_{cr} is attained. A catastrophic reduction in the neutron yield may be predicted when $\delta \gg \delta_{cr}$. Thus, experiments should be carried out with the compression ratio increasing from one experiment to the next. Since the compression is mainly determined by the ratio of the shell mass M to the mass m of the DT gas, the compression ratio may be varied by varying the thickness of the shell or the initial pressure of the DT gas. The shell diameter may also be varied, although an increase in the diameter results in the shell moving out of the shadow region (i.e., direct contact between the laser beams and the shell) while a reduction in the diameter impedes the diagnostics of the plasma parameters.

The targets in the Iskra-5 facility[6] are irradiated by twelve laser beams with an energy of up to $\simeq 1$ kJ at the channel exit, which corresponds to a total laser energy of $\simeq 9$ kJ deposited inside the target chamber. The laser pulse duration is $\simeq 0.35 \pm 0.05$ ns which corresponds to a radiation power up to $\simeq 30$ TW incident on the target. The pulse profile is close to Gaussian. The divergence of the radiation is stable at $\simeq (0.6-0.8) \times 10^{-4}$ rad. The contrast of the laser pulse is sufficiently high ($K > 10^6$) that the interior of the hohlraum did not fill up with plasma prior to the time of arrival of the main pulse. The precision of the transverse focusing on the target is ± 30 μm and the precision of the longitudinal matching of the foci of the aligning and power radiation is ± 100 μm. The diameter of the laser beam constrictions is $\leqslant 100$ μm.

The target (hohlraum) consists of a thin-walled copper spherical enclosure (similar to an inverted-corona target[6]) with six laser entrance holes. The inner surface is coated with a 1 μm thick layer of gold and the outer surface is coated with a layer of bismuth 0.1–0.3 μm thick. A glass microsphere filled with a gaseous equimolar mixture of deuterium and tritium at a pressure $P_{DT} \simeq 5-20$ atm is positioned at the center. The diameter of the enclosure (1.3–4 mm), the diameter of the laser entrance holes (0.4–0.7 mm), and the parameters of the glass microshell were all varied experimentally. In order to increase the thickness of the shell above 7 μm, a layer of polyparaxylylene up to 40 μm thick was deposited on the surface. The optimum design had an enclosure diameter of 2 mm, six holes of diameter 0.6 mm, a shell diameter of 280 μm, a glass shell thickness of 5 μm, and $P_{DT} \simeq 10$ atm.

The diagnostic apparatus of the Iskra-5 system was described in Refs. 6 and 9. It incorporates an energy balance system, an interference schlieren method for optical recording of the laser radiation contrast and the dynamics of the plasma expansion, a system for time-resolved, spatially resolved, and spectrally resolved recording of the x-ray emission, and neutron measurement facilities. The main experimental results are reported below. Where necessary, additional features of the measuring methods are noted.

4. EXPERIMENTAL RESULTS AND DISCUSSION

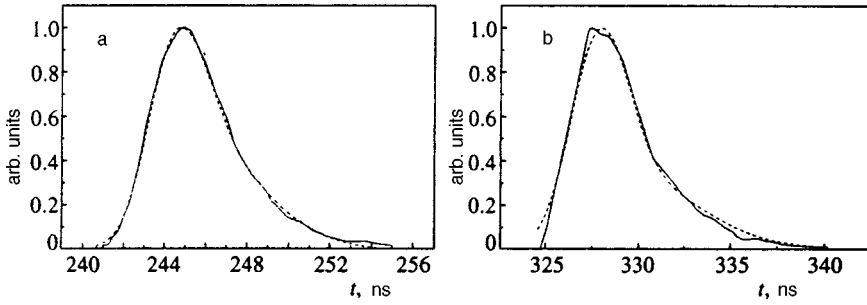Results of seventeen experiments carried out using 2 mm diameter targets coated with gold on the inside are presented

FIG. 2. Neutron pulse recorded using a time-of-flight method (solid curve) and calculated signal at a given ion temperature T_i (dashed curve): a—flight length 12.5 m, $T_i = 2.5$ keV; b—flight length 16.7 m, $T_i = 2.9$ keV.

below. The diameters of the entrance holes were $\varnothing_{hole} = 0.6$ mm except for three experiments with $\varnothing_{hole} = 0.4$–0.45 mm. The average shell diameter was $\varnothing_{sh} = 280$ μm with limiting values of 271 and 290 μm. The thickness of the glass shells (without coatings) varied between 3 and 7 μm. The thickness of the shells d_{sh} on which coatings were applied was 1.2–5 μm. The coatings were either plastic, magnesium, or SiO$_2$. The maximum thickness of the plastic layer was $d_{CH} = 41$ μm (in this case the thickness of the glass was 4.7 μm). The pressure of the DT gas was varied between 5 and 16 atm, except for one experiment in which the pressure was around 1 atm. The laser pulse energy E_L was varied between 7.4 and 9.6 kJ, except for one experiment where the energy was 3.9 kJ. The neutron yield varied between 6.5×10^9 (for $E_L = 9$ kJ, $P_{DT} = 13$ atm, $\varnothing_{sh} = 280$ μm, $d_{sh} = 4.8$ μm) and 7×10^7 ($E_L = 8.2$ kJ, $P_{DT} = 5$ atm, $\varnothing_{sh} = 277$ μm, $d_{sh} = 4.7$ μm, and $d_{CH} = 41$ μm).

The interference schlieren method indicates that in these experiments the single pulse contrast was fairly high and the target plasma only formed at the instant of arrival of the single pulse.

A typical oscilloscope trace, which can be used to determine the time of arrival of the neutron pulse at the scintillation detector relative to the time of generation of the hard ($h\nu \simeq 0.1$ MeV) x-ray pulse inside the hohlraum, was given in Ref. 2. The delay of the neutron pulse relative to the x-ray pulse corresponds to 14.1 MeV neutrons.

The temperature of the ion component of the DT plasma, obtained using a time-of-flight method, varied in the range $T_i = 1.5$–3 keV. Figure 2 shows the profiles of the neutron pulses recorded at 12.5 and 16.7 m. The dashed curves show the calculated neutron pulses for the given ion temperature T_i taking into account the instrumental function of the appropriate recording channel. The temperatures at which the calculated and experimental neutron pulses showed the smallest differences are given in the caption to Fig. 2. The accuracy of the ion temperature measurements was approximately 1.7 keV at 12.5 m and approximately 0.3 keV at 16.7 m.

The convergence time of the shells was recorded using a slit time scan of the x-ray image of the microtarget in the photon energy range $h\nu \simeq 4$ keV using an x-ray streak camera with $\simeq 50$ ps time resolution. An example of such a recording is shown in Fig. 3. Two emission peaks can be observed. The first corresponds to the maximum emission intensity of the plasma expanding from the inner surface of the hohlraum and the second occurs as a result of the com-

pression and heating of the shell containing the DT fuel. The shell convergence time $\tau_{\gamma\gamma}$, determined as the time between the points of intersection of lines linearly approximating the leading edges of the appropriate pulses with the time axis, varies between 0.25–0.73 ns and increases with increasing thickness (strictly speaking this is only valid for thicknesses $\geqslant 7$ μm). The convergence time can be used to determine the shell flight velocity and to estimate the energy input to the shell. A convergence time $\tau_{\gamma\gamma} \simeq 0.35$ ns corresponds to the velocity

$$V \simeq \frac{R_0 - R_{min}}{\tau_{\gamma\gamma}} \simeq 3 \times 10^7 \text{ cm/s}.$$

Figure 4 shows an example of an x-ray image of the central capsule obtained through a laser entrance hole using a pinhole camera. The initial position of the outer boundary and the spherically compressed core can be seen. A spherically compressed image (apart from distortions caused by the graininess of the photographic film) is observed in all experiments in which this image was obtained. In experiments using thick-walled targets no compressed core can be observed.

In these experiments the neutron yield varied by two orders of magnitude: between 6.5×10^9 for a shell thickness of 5 μm and 7×10^7 when a layer of polyparaxylylene 41 μm thick was deposited on the shell. The neutron yield data are described by the simple scaling law (see Fig. 5)
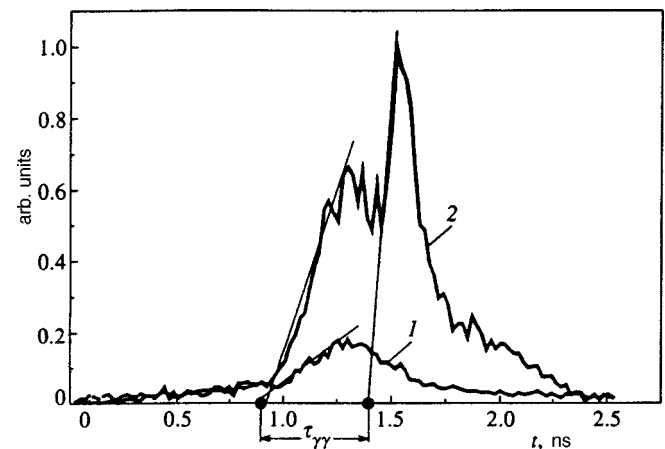
$$N \sim m^{2/3} M^{4/3} f(E_\gamma/M), \tag{5}$$

FIG. 3. X-ray emission from central target recorded using a spatially resolving x-ray streak camera: *1*—emission of laser corona of hohlraum, *2*—emission of central capsule.
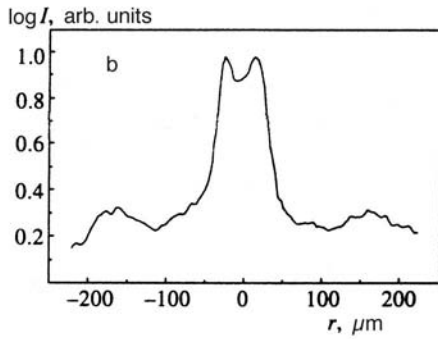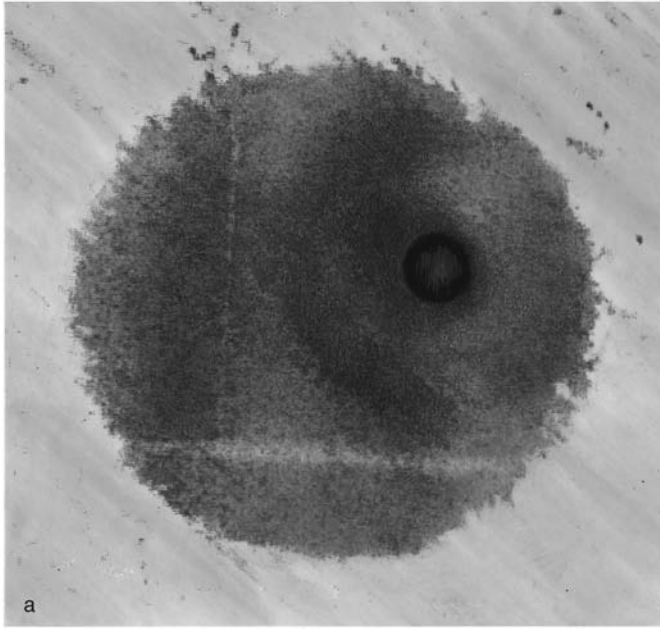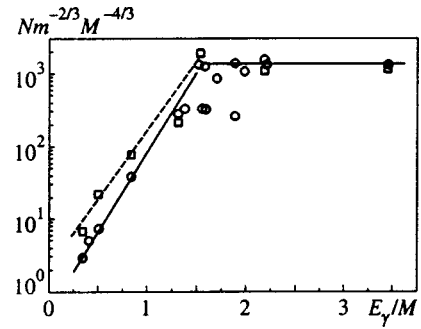
FIG. 5. Normalized neutron yield versus specific x-ray energy input to central capsule (arbitrary units): circles—experimental values, squares—calculations using SNDP program.

$$N \sim m^{1/3} \rho_0^{2/3} \delta_V^{2/3} \left(\frac{T}{T_0} \right)^n f(T_0). \qquad (8)$$

Here $\delta_V = \rho / \rho_0$ is the volume compression.

The compression ratio is determined to a considerable extent by the ratio M/m: $\delta_V \sim (M/m)^\alpha$. In the simplest exploding shell models we have $\alpha \simeq 1 - 2$ (Refs. 10 and 11) and for adiabatic compression $T = T_0 \delta_V^{2/3}$. In these experiments, the losses of heat as a result of electron heat conduction are substantial and thus the effective adiabatic exponent is $\gamma_{\text{eff}} < 5/3$ so that

$$T = T_0 \delta_V^{\gamma_{\text{eff}} - 1} \simeq T_0 \left(\frac{M}{m} \right)^{\gamma_{\text{eff}} - 1},$$

and this implies that

$$N \sim m^{4/3} \frac{M^{2/3}}{R_0^2} T_0^{n - 0.5} \left(\frac{M}{m} \right)^{(\gamma_{\text{eff}} - 1)(n - 0.5)}.$$

For $n \simeq 4$ and $\gamma_{\text{eff}} = 1.2$ we obtain the scaling law (5).

It can be seen from Fig. 5 that on the whole the scaling law reasonably accurately describes the experimental neutron yield (except for three experiments). In this case, as a result of a change in the technology used to fill the glass capsule with DT gas (a hole was drilled in the shell to admit argon and was then sealed), for two experiments there is reason to believe that the DT pressure at the time of the experiment was appreciably less than the initial pressure. It was impossible to monitor the pressure after depositing a layer of polyparaxylylene on the shell. We are unaware of any major reasons for failure of the third experiment in the series. However, the emission pattern of the target obtained using a pinhole camera clearly reveals emission from the edges of the light entrance holes. In this experiment, there was probably an alignment error for these channels, resulting in loss of laser input energy.

Figure 5 shows two sections of curve with different slopes. For $E_\gamma / M > 1.5$ (arbitrary units) the neutron yield does not depend on E_γ / M, whereas for $E_\gamma / M < 1.5$ it decreases rapidly with decreasing E_γ / M. The following treatment of this behavior seems justified. For $E_\gamma / M > 1.5$ the target operates as a completely heated (exploding) shell. The interface at $E_L = 8$ kJ corresponds to a shell mass $M \simeq 3.5$ μg. The thickness of a 280 μm diameter glass shell

FIG. 4. Pinhole photograph of central target (a) and its densitogram in one cross section (b).

where m and M are the mass of the gas and the shell, and E_γ is the x-ray energy in the hohlraum. We shall give a brief substantiation of formula (5). The neutron yield is given by

$$N = \int dt \int dV n_D n_T \langle \sigma N \rangle_{\text{DT}},$$

where n_D and N_T are the concentrations of deuterium and tritium nuclei, and $\langle \sigma V \rangle_{\text{DT}}$ is the rate constant of the thermonuclear reaction. We can write approximately

$$N \sim \rho^2 V \langle \sigma V \rangle_{\text{DT}} \frac{r}{\sqrt{T}} \simeq m \rho r f(T), \qquad (6)$$

where ρ and r are the density and radius of the compressed DT, and $f(T) = \langle \sigma V \rangle_{\text{DT}} T^{-1/2}$. Using the conservation of the DT mass during compression, we can write $\rho r \sim m^{1/3} \rho^{2/3}$ so that

$$N \sim m^{4/3} / \rho^{2/3} f(T). \qquad (7)$$

Approximating the rate constant of the DT reaction by the power function $f(T) \propto (T/T_0)^n T_0^n$, we have
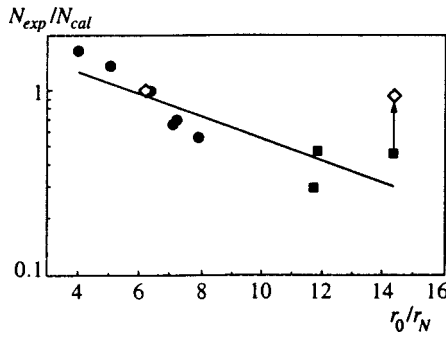
FIG. 6. Ratio of experimental and calculated neutron yields as a function of the calculated degree of radial convergence of the central capsule at the instant of neutron generation: circles—shell without coating; filled squares—shell with polyparaxylylene coating, open squares—calculations allowing for turbulent mixing.

FIG. 7. Comparison between experimental pinhole trace (dashed curve) and that calculated using the SNDP program (solid curve).

with this mass is ~ 7 μm. The fact that the normalized yield of neutrons is independent of E_γ/M where $E_\gamma/M \geqslant 1.5$ holds, indicates that the shell is completely heated to the same temperature. If $E_\gamma/M \leqslant 1.5$ holds, the stored thermal energy is not sufficient for complete heating and ablation results.

We shall consider for a moment E_γ, which we determined with empirical allowance for the losses of energy absorbed by fast ions. It was assumed that the fraction of absorbed energy that goes into accelerating them is

$$\eta = x(3.5+x)^{-1}, \quad x = I_g/I_1 - 1, \quad I_1 = 2 \times 10^{14} \text{ W/cm}^2$$

(for the definition of I_g see Eq. (1)).

Figure 5 also gives the results of calculations of the neutron yield using the SNDP nonequilibrium radiative gasdynamics program.[12] In this case, the calculated value of E_γ/M was used as a marker in the appropriate experiment. It can be seen that to within $\simeq 1.5$ the calculations reproduce the experimental data in the exploding shell regime. In the ablation regime the difference is $\simeq 2-3$.

The calculations were made using a spherically symmetric formulation with effective allowance for the losses of laser light and x-rays through the laser entrance holes. Allowance was only made for bremsstrahlung absorption with an average angle of incidence of the laser radiation on the inner surface of the hohlraum. The energy losses to the fast ions were neglected. The equations of state and the spectral ranges were calculated using the average ion approximation while the radiation transport was calculated using the multi-group diffusion approximation. The chemical composition of the glass shell corresponded to pure silicon oxide SiO_2. The chemical composition of the plastic layer corresponded to CH.

It is interesting to compare the experimental and calculated neutron yields, $N_{\exp}$ and N_{cal}, as a function of the calculated degree of radial convergence r_0/r_N, where r_N is the inner radius of the shell at the instant when the highest rate of neutron generation is observed (Fig. 6). As in the previous figure, it can be seen that as the compression ratio $\delta_V = (r_0/r_N)^3$ increases, the ratio $N_{\exp}/N_{\text{cal}}$ decreases, reaching $\sim 1/3$ for $r_0/r_N = 12$. Is this a consequence of the asymmetry of the x-ray field? In order to partly answer this question, we
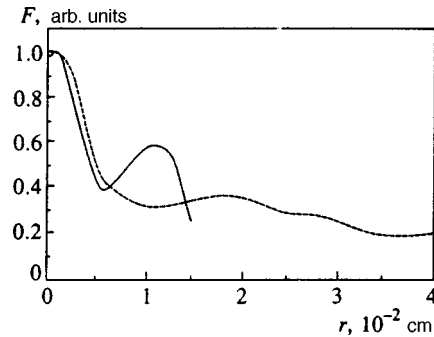
made calculations using the SND-TUR program[13] in which the Nikiforov model[14] is used to take into account the evolution of a turbulent mixing zone at the shell–fuel interface. It can be seen that for $r_0/r_N \leqslant 6$ ($\delta_V \leqslant 200$), turbulent mixing has little influence on the neutron yield, whereas for $r_0/r_N \simeq 14$ this mixing halved the neutron yield according to the calculations (see arrow in Fig. 6). This decrease is mainly caused by a drop in the fuel temperature as a result of the heat lost to heating the cold shell material entering the mixing zone. Thus, it seems justifiable to conclude that the symmetry of the x-ray irradiation of the target has little effect in this range of compressions. For clarification, we note that the accuracy of the calculated modeling of the neutron yield does not exceed $\simeq 2$, as is indicated by the two points corresponding to the lowest values of r_0/r_N in Fig. 6.

A close correspondence between the experimental and calculated data is also observed for other quantities. Figure 7 gives a comparison between the experimental and calculated spatial emission patterns of the central capsule obtained using a pinhole camera. Good agreement is observed in the central part of the core. The difference at the wings of the distribution is caused by the inadequate gasdynamic description of the collision process of the glass and gold plasmas which forms the emission profile at the wings in the calculations.

Figure 8 compares the dynamics of the experimental (obtained using an x-ray streak camera) and calculated emis-
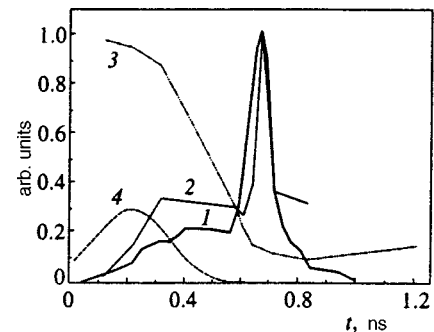

FIG. 8. Comparison between experimental (*1*) and calculated (*2*) luminous exitance of the central capsule in the range of x-ray quantum energies 3–4 keV; *3*—calculated $R-t$ diagram of the motion of the gas–glass interface; *4*—laser pulse.
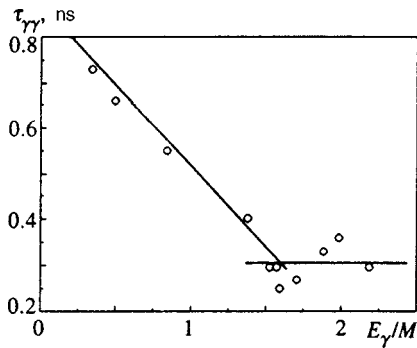
FIG. 9. Convergence time of central capsule versus specific x-ray energy input (arbitrary units).

sion patterns of the target in the 3–4 keV spectral range. Figure 9 gives the experimental dependences of the collapse time on the parameter E_γ/M.

5. CONCLUSIONS

Experiments have been carried out on the Iskra-5 laser facility using composite x-ray targets comprising a spherical hohlraum with six laser entrance holes and a centrally positioned DT-filled glass capsule coated with a layer of plastic (polyparaxylylene) between 10 and 40 μm thick. This capsule design was used to study the operation of targets with fuel compressions up to $\delta_V \lesssim 3 \times 10^3$ at the instant of neutron generation. The results showed that in this range of compressions the behavior of the targets closely corresponds to the one-dimensional spherically symmetric calculations, with turbulent mixing halving the neutron yield at the upper limit ($\delta_V \approx 3 \times 10^3$).

This behavior of the target indicates that the x-ray irradiation of the glass microsphere is highly uniform. Calculated estimates of the upper limit of the x-ray field asymmetry at the target yield values around 3%.

*)E-mail: kochemasov@otd13.vniief.ru

[1] J. Lindl, Phys. Plasmas **2**, 3933 (1995).
[2] F. M. Abzaev, V. I. Annenkov, V. G. Bezuglov *et al.*, JETP Lett. **58**, 28 (1995).
[3] E. M. Campbell, J. D. Kilkenny, D. R. Kania *et al.*, Laser Part. Beams **9**, 209 (1991).
[4] H. Nishimura, H. Shiraga, H. Takabe *et al.*, in *Proceedings of the 14th International Conference on Plasma Physics and Controlled Nuclear Fusion Research*, *Wurzburg, Germany*, *1992*, Vol. **3 (IAEA, Vienna, 1993), p.** 97.
[5] CEL-V Laser Team, in *Proceedings of the 12th International Conference on Laser Interaction and Related Plasma Phenomena*, *Osaka, Japan*, *1995*, AIP Conf. Proc. **369**, 40 (1995).
[6] A. V. Bessarab, V. A. Gaĭdash, G. V. Dolgoleva *et al.*, Zh. Éksp. Teor. Fiz. **102**, 1800 (1992) [Sov. Phys. JETP **75**, 970 (1992)].
[7] M. D. Rosen, D. W. Phillion, V. C. Rupert *et al.*, Phys. Fluids **22**, 2020 (1979).
[8] L. S. Mkhitarjan and G. G. Kochemasov, in *Proceedings of the 23rd European Conference on Laser Interaction with Matter*, *Oxford*, *1994*, Inst. Phys. Conf. Ser. **140**, 351 (1995).
[9] F. M. Abzaev, A. V. Bessarab, G. A. Kirillov *et al.*, Vopr. At. Nauki Tekh. Ser. Mat. Model. Fiz. Prots. No. 4, 68 (1992).
[10] M. D. Rosen and J. H. Nuckols, Phys. Fluids **22**, 1393 (1979).
[11] S. A. Bel'kov, A. V. Bessarab, A. V. Veselov *et al.*, Zh. Éksp. Teor. Fiz. **97**, 834 (1990) [Sov. Phys. JETP **70**, 467 (1990)].
[12] S. A. Bel'kov and G. V. Dolgoleva, Vopr. At. Nauki Tekh. Ser. Mat. Model. Fiz. Prots. No. 1, 59 (1992).
[13] V. A. Andronov, S. A. Bel'kov, G. V. Dolgoleva *et al.*, Vopr. At. Nauki Tekh. Ser. Mat. Model. Fiz. Prots. No. 3, 1 (1997).
[14] V. A. Andronov, E. E. Meshkov, V. V. Nikiforov *et al.*, Dokl. Akad. Nauk SSSR **264**, 76 (1982) [Sov. Phys. Dokl. **27**, 393 (1982)].

Translated by R. M. Durham

The role of helicity in turbulent MHD flows

E. Golbraikh*)

Center for MHD Studies, Ben Gurion University of the Negev, P.O.B. 653, Beer-Sheva 84105, Israel

O. G. Chkhetiani and S. S. Moiseev

Space Research Institute, Russian Academy of Sciences, 117810 Moscow, Russia
(Submitted 15 January 1998)
Zh. Éksp. Teor. Fiz. **114**, 171–181 (July 1998)

We have studied the behavior of a helical homogeneous small-scale MHD turbulent flow under the influence of a weak inhomogeneous large-scale disturbance. We have shown that turbulent energy redistribution in the presence of nonzero helicity occurs mainly over large scales. Helicity increases correlation time, leading to the weakening of a direct cascade and to the formation of steep spectra over small scales, with simultaneous turbulent energy growth over large scales. Furthermore, an expression for the effective viscosity of the mean flow is derived. It is shown that the magnetic field, in addition to the helicity, reduces the effective viscosity of the medium. This may be important in the study of MHD flow around obstacles in the presence of an external magnetic field. © *1998 American Institute of Physics.* [S1063-7761(98)01207-4]

1. INTRODUCTION

The problem of self-organization of a turbulent MHD flow with magnetic Reynolds number $\mathrm{Re_m} \ll 1$ in an external homogeneous magnetic field has long been under discussion (see, for instance, Refs. 1–4, with references therein). Obviously, an external magnetic field causes a rearrangement of the topological structure of a turbulent flow. Specifically, if the original turbulence (in absence of a magnetic field) is isotropic, it becomes anisotropic in the presence of a magnetic field. Furthermore, in presence of a magnetic field, the spectral and dynamical properties of turbulence can change.

As demonstrated by numerous experiments (see, e.g., Refs. 4–6), the turbulence spectrum varies with the magnetic field. It should be emphasized, however, that turbulence essentially always remains three-dimensional, although there exists a tendency to quasi-two-dimensionality. Over small scales, the spectral dependence of the turbulent energy E_t on wave number k is of the form $E_t \sim k^{-\alpha}$, where the exponent α varies with increasing magnetic field from $-5/3$ (at $\mathbf{B}=0$) to between -2 and $-7/3$ (at low $\mathbf{B}$ values).[1,4] With growing magnetic field, α ranges from $-11/3$ to -4, and the turbulence becomes highly intermittent.[4]

It is noteworthy that such magnetic field-dependent behavior of the turbulence spectrum is observed only in those experiments where turbulent flow is generated either by drawing a grid through the medium[1] or in the presence of a honeycomb.[4]

For the entire subsequent analysis, it is important to note that turbulence becomes helical in an external magnetic field. This means that the one-point correlation function $H = \langle \mathbf{v} \cdot (\nabla \times \mathbf{v}) \rangle \neq 0$, where $\mathbf{v}$ is the flow velocity. On the other hand, purely helical turbulence is characterized by the quantity $\alpha = -7/3$,[7] which is in agreement with experimental results.[1,4] In the general case, helicity, along with energy, is the most important feature of a turbulent flow. Helicity, being the second invariant of Euler's equation, just like energy,[8] exerts a significant influence on the evolution and stability of turbulent and laminar flows.[9] Helicity is probably one of the main sources of magnetic field generation and maintenance in astrophysical objects.[10] In the absence of a magnetic field, helical turbulence is unstable against large-scale disturbances.[11] This leads to energy redistribution between large-scale and small-scale fluctuations. On the other hand, helicity leads to an efficient viscosity decrease in the mean flow, i.e., to a decrease in Reynolds stresses.[12]

The present paper deals with the behavior of small-scale helical turbulence in an external homogeneous magnetic field and under a weak large-scale disturbance. We also examine the effect of a magnetic field on the viscosity of such turbulence.

2. PRINCIPAL EQUATIONS

Let us write the system of MHD equations in dimensionless form for an incompressible fluid:

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} = -\nabla P + \frac{1}{\mathrm{Re}} \nabla^2 \mathbf{u} + \frac{N}{\mathrm{Re_m}}(\nabla \times \mathbf{B}) \times \mathbf{B} + \mathbf{F}, \tag{1}$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}) + \frac{1}{\mathrm{Re_m}} \nabla^2 \mathbf{B}, \tag{2}$$

$$\nabla \cdot \mathbf{u} = \nabla \cdot \mathbf{B} = 0,$$

where $\mathbf{F}$ is an external non-electromagnetic force and P is pressure. The problem is characterized by three dimensionless numbers: the Reynolds number $\mathrm{Re} = UL/\nu$, the magnetic Reynolds number $\mathrm{Re_m}$, and the magnetic interaction parameter $N = \sigma B^2 L/\rho U$ (here ρ, ν are fluid density and viscosity, U and L are characteristic velocity and dimension).

We represent all fields as a sum of averaged and fluctuating values:

$$\mathbf{u} = \langle \mathbf{u} \rangle + \mathbf{u}', \quad \mathbf{B} = \mathbf{B}_0 + \mathbf{h}, \quad P = \langle P \rangle + P', \quad \mathbf{F} = \langle \mathbf{F} \rangle + \mathbf{F}',$$

$$\mathbf{B}_0 \gg \mathbf{h}, \quad \langle \mathbf{u}' \rangle = \langle \mathbf{h} \rangle = \langle P' \rangle = \langle \mathbf{F}' \rangle = 0,$$

where $\langle ... \rangle$ denotes averaging over an ensemble.

Assuming first that $\langle \mathbf{u} \rangle = 0$ and $\mathrm{Re_m} \ll 1$, one can easily derive an equation for $\mathbf{u}'$ up to second order in $\mathbf{h}$:

$$\frac{\partial \mathbf{u}'}{\partial t} + (\mathbf{u}' \cdot \nabla)\mathbf{u}' - \langle (\mathbf{u}' \cdot \nabla)\mathbf{u}' \rangle = -\nabla P' + \frac{1}{\mathrm{Re}} \nabla^2 \mathbf{u}'$$

$$+ \frac{N}{\mathrm{Re_m}} (\nabla \times \mathbf{h}) \times \mathbf{B}_0 + \mathbf{F}'.$$

Let us examine the stability of a small-scale turbulent flow under the magnetic field with respect to weak large-scale nonuniform disturbances. In this case, we represent the turbulent field as a sum of the initial turbulent field $\mathbf{u}^{(0)}$ and its disturbance $\mathbf{u}^{(1)}$:

$$\mathbf{u}' = \mathbf{u}^{(0)} + \mathbf{u}^{(1)},$$

where $\mathbf{u}^{(0)} \gg \mathbf{u}^{(1)}$.

We introduce the correlation functions

$$Q_{ij}^{00}(\xi, \tau) = \langle u_i^{(0)}(\mathbf{x}, t) u_j^{(0)}(\mathbf{x} + \xi, t + \tau) \rangle,$$

$$Q_{ij}^{10}(\mathbf{x}, \xi, t, \tau) = \langle u_i^{(1)}(\mathbf{x}, t) u_j^{(0)}(\mathbf{x} + \xi, t + \tau) \rangle,$$

and, in the same manner, correlators of higher order.

Assuming that the external force $\mathbf{F}'$ maintains the initial small-scale helical turbulence, we can derive an equation for $Q_{ij}^{00} = Q_{ij}^{00}(\xi, \tau)$:

$$\frac{\partial Q_{ij}^{00}}{\partial t} + \nabla_k Q_{kij}^{000} = -\nabla_i \langle P^{(0)} u_j^{(0)} \rangle + \frac{1}{\mathrm{Re}} \nabla^2 Q_{ij}^{00}$$

$$+ \frac{N}{\nabla^2} (\nabla_i \nabla_3 Q_{3j}^{00} - \nabla_3^2 Q_{ij}^{00}) + \langle F_i' u_j^{(0)} \rangle,$$

where the magnetic term is obtained from Eq. (2) with allowance for the above assumptions, and $\mathbf{B}_0$ is directed along the third coordinate.

In a similar way, we can write an equation for $Q_{ij}^{10} = Q_{ij}^{10}(\mathbf{x}, \xi, t, \tau)$:

$$\frac{\partial Q_{ij}^{10}}{\partial t} - \frac{1}{\mathrm{Re}} \nabla^2 Q_{ij}^{10} + \nabla_k (Q_{ikj}^{100} + Q_{kij}^{100})$$

$$= -\nabla_i \langle P^{(1)} u_j^{(0)} \rangle + \frac{N}{\nabla^2} (\nabla_i \nabla_3 Q_{3j}^{10} - \nabla_3^2 Q_{ij}^{10}).$$

Eliminating the pressure from this equation, we obtain the final equation for Q_{ij}^{10}:

$$\frac{\partial Q_{ij}^{10}}{\partial \tau} - \frac{1}{\mathrm{Re}} \nabla^2 Q_{ij}^{10} + \Pi_{im} \nabla_k (Q_{mkj}^{100} + Q_{kmj}^{100})$$

$$= -N \cos^2 \theta Q_{ij}^{10}, \qquad (3)$$

where $\Pi_{im} = (\delta_{im} - \nabla_i \nabla_m / \nabla^2)$ is a projection operator, and the operator $\cos^2 \theta = \nabla_3^2 / \nabla^2$.

Similarly to Eq. (3), we obtain an equation for the third moments $Q_{ijk}^{100} = Q_{ij}^{10}(\mathbf{x}, \xi, \xi', t, \tau, \tau')$:

$$\left(\frac{\partial}{\partial \tau} - \frac{1}{\mathrm{Re}} \nabla^2 + N \cos^2 \theta \right) Q_{ijn}^{100}$$

$$+ \Pi_{im} \nabla_k (Q_{mkjn}^{1000} + Q_{kmjn}^{1000} - Q_{mk}^{10} Q_{jn}^{00} - Q_{mk}^{10} Q_{jn}^{00}) = 0. \qquad (4)$$

To complete the system of Eqs. (3) and (4), we use the results of Ref. 12, where the finiteness of the correlation time is taken into account via the two-scale analog of the Orszag eddy-damped quasinormal Markovian (EDQNM) approximation.[13] This approximation consists in the replacement of the fourth-order moment cumulant in the equation for the third moment by an effective damping term proportional to the square-law combination of pair moments. This approach is analogous to other traditional turbulent second-order closures.[14] The EDQNM approximation for strong MHD turbulence was considered in detail in Ref. 15.

In this approximation, we obtain from Eq. (4)

$$Q_{ipj}^{100} = -\tilde{\tau} \Pi_{im} \nabla_k Q_{mkpj}^{1000},$$

where the three-point correlator

$$Q_{mkpj}^{1000}(x,t,x,t,\xi,\tau,\xi',\tau') = Q_{mk}^{10}(x,t,x,t) Q_{pj}^{00}(\xi - \xi', \tau - \tau')$$

$$+ Q_{mp}^{10}(x,t,\xi,\tau) Q_{kj}^{00}(\xi', \tau')$$

$$+ Q_{mj}^{10}(x,t,\xi',\tau') Q_{kp}^{00}(\xi, \tau) \quad (5)$$

and

$$\tilde{\tau} = \left(\frac{1}{\tau^*} + N \cos^2 \theta \right)^{-1},$$

and the correlation time $\tau^* \simeq L_{\mathrm{tur}} / E_{\mathrm{tur}}^{1/2}$ (L_{tur} and E_{tur} being the characteristic scale and average energy of turbulent flow, respectively).

To substitute the expression (5) for the third moment into Eq. (3), we pass to the limit $\xi' \to \xi$, $\tau' \to \tau$. Here (see Appendix) we take into account that $A \delta_{ij} = Q_{ij}^{00}(0,0)$ and $C \varepsilon_{ijp} = \partial Q_{ij}^{00}(\xi, \tau) / \partial \xi_p |_{\xi, \tau \to 0}$ depend on B (or N), i.e., on the magnetic field. This results in

$$\left(\frac{\partial}{\partial \tau} - \left(\frac{1}{Re} + A \tilde{\tau} \right) \Delta + N \cos^2 \theta \right) Q_{ij}^{10}(\mathbf{x}, t, \xi, \tau) - \tilde{\tau} C H_{ij}^{10}$$

$$= \tilde{\tau} \nabla_p \left[\frac{\partial}{\partial x_k} Q_{ip}^{10}(\mathbf{x}, t, 0, 0) Q_{kj}^{00}(\xi, \tau) \right.$$

$$+ \frac{\partial}{\partial x_s} Q_{kp}^{10}(\mathbf{x}, t, 0, 0) \nabla_k \xi_s Q_{lj}^{00}(\xi, \tau)$$

$$\left. + Q_{kp}^{10}(\mathbf{x}, t, 0, 0) \nabla_k Q_{ij}^{00}(\xi, \tau) \right], \qquad (6)$$

where $A = A(N, 0, 0)$ and $C = C(N, 0, 0)$ are scalar functions of N, and $H_{ij} = \varepsilon_{ikl} \nabla_k Q_{lj}^{10}(\mathbf{x}, t, \xi, \tau)$.

3. INSTABILITY OF THE SECOND MOMENTS

To study the stability of the system (6), we apply the operator $\varepsilon_{lmi} \nabla_m$, and write the resulting system in homogeneous form:

$$\left(\nabla_\tau-\frac{1}{\mathrm{Re}}\nabla^2-\widetilde{\tau}A\nabla^2+N\cos^2\theta\right)Q_{ij}^{10}-\widetilde{\tau}CH_{ij}^{10}=0,$$

$$\left(\nabla_\tau-\frac{1}{\mathrm{Re}}\nabla^2-\widetilde{\tau}A\nabla^2+N\cos^2\theta\right)H_{ij}^{10}+\widetilde{\tau}C\nabla^2Q_{ij}^{10}=0. \tag{7}$$

An equation for Q_{ij}^{10} follows from (7):

$$((\nabla_\tau-\nu_H\nabla^2+N\cos^2\theta)^2+\widetilde{\tau}^2C^2\nabla^2)Q_{ij}^{10}=0, \tag{8}$$

where the effective viscosity $\nu_H=1/\mathrm{Re}+\widetilde{\tau}A$. Passing into k-space in (8), we derive an expression for the decay factor $\gamma=-i\omega$:

$$\gamma=-\nu_H k^2-N\cos^2\theta+\widetilde{\tau}|Ck|, \tag{9}$$

where θ denotes the angle between $\mathbf{k}$ and $\mathbf{B}_0$.

It follows from the form of Eq. (9) that helicity increases relaxation time. In other words, helicity prolongs vortex lifetime. However, the helicity is essentially imperceptible in case of $\nu_H k^2+N\cos^2\theta\gg\widetilde{\tau}|Ck|$. Hence, at sufficiently weak magnetic fields the helicity effect is most important over large scales, i.e., at low k values. Under the condition $\gamma>0$, we obtain from Eq. (9) the instability condition

$$\frac{1}{2}k_0\left(1-\left(1-\frac{4N\cos^2\theta}{k_0^2\nu_H}\right)^{1/2}\right)$$
$$<k<\frac{1}{2}k_0\left(1+\left(1-\frac{4N\cos^2\theta}{k_0^2\nu_H}\right)^{1/2}\right), \tag{10}$$

where $k_0=\widetilde{\tau}|C|/\nu_H$.

We study two limiting cases. Let $\cos\theta=0$, i.e., consider modes for which $\mathbf{k}\perp\mathbf{B}_0$. In this case, (10) acquires the form of the purely hydrodynamic limit:[11]

$$0<k<k_0. \tag{11}$$

Here

$$k_0=\frac{\tau^*|C|}{1/\mathrm{Re}+A\tau^*},$$

and the dependence of the scale k_0 on the magnetic field is contained only in the coefficients A and C. As a rule, $A\tau^*\gg1/\mathrm{Re}$, which results in $k_0=C/A$. Consequently, the k_0 dependence on the magnetic field is determined by the ratio of helicity to turbulence intensity.

On the other hand, at $\cos\theta=1$, i.e., in modes for which $\mathbf{k}\|\mathbf{B}_0$, we obtain the condition (10) in the form

$$\frac{1}{2}\left(k_0-\left(k_0^2-\frac{4N}{\nu_H}\right)^{1/2}\right)<k<\frac{1}{2}\left(k_0+\left(k_0^2-\frac{4N}{\nu_H}\right)^{1/2}\right), \tag{12}$$

where

$$k_0=\frac{\widetilde{\tau}|C|}{1/Re+A\widetilde{\tau}},\quad\widetilde{\tau}=\frac{\tau^*}{1+N\tau^*}.$$

It is evident from (12) that at $\mathbf{B}_0=0$, this interval coincides with that of the modes with $\mathbf{k}\perp\mathbf{B}_0$. With growing magnetic field, $\widetilde{\tau}$ decreases, i.e., the effective correlation time decreases. Simultaneously, the instability interval is reduced and vanishes in fields described by $4N/\nu_H=k_0^2$.

It follows from (11) and (12) that at scales

$$0<k<\frac{1}{2}\left(k_0-\left(k_0^2-\frac{4N}{\nu_H}\right)^{1/2}\right),$$
$$\frac{1}{2}\left(k_0+\left(k_0^2-\frac{4N}{\nu_H}\right)^{1/2}\right)<k<k_0, \tag{13}$$

modes with $\mathbf{k}\perp\mathbf{B}_0$ are unstable against large-scale disturbances, whereas those with $\mathbf{k}\|\mathbf{B}_0$ are attenuated.

Thus, the energy of a large-scale disturbance is redistributed among the scales, so that modes transverse to the magnetic field predominantly grow.

4. TURBULENT VISCOSITY

To study the influence of the magnetic field on viscosity, we examine the variations in hydrodynamic viscosity (in the absence of a magnetic field) in an external homogeneous magnetic field.

We consider the case $\langle\mathbf{u}\rangle\neq0$ and $\langle\mathbf{u}\rangle\ll\mathbf{u}^{(0)}$. In this case, the equation for $\langle\mathbf{u}\rangle$,

$$\frac{\partial\langle\mathbf{u}\rangle}{\partial t}+\langle\mathbf{u}\rangle(\nabla\cdot\langle\mathbf{u}\rangle)-\langle(\mathbf{u}'\cdot\nabla)\mathbf{u}'\rangle=-\nabla\langle p\rangle$$
$$+\frac{1}{\mathrm{Re}}\nabla^2\langle\mathbf{u}\rangle+\frac{N}{\mathrm{Re}_m}\langle(\nabla\cdot\mathbf{B})\times\mathbf{B}\rangle+\langle\mathbf{F}\rangle, \tag{14}$$

involves Reynolds stresses $\langle(\mathbf{u}'\cdot\nabla)\mathbf{u}'\rangle$ that depend only on Q_{ij}^{10} for homogeneous turbulence, as shown in Refs. 12 and 17. On the other hand, additional terms appear on the right-hand side of Eq. (6) for $Q_{ij}^{10}(\xi,\tau,\mathbf{x},t)$:

$$\left[\left(\left(\frac{\partial}{\partial\tau}-\frac{1}{\mathrm{Re}}\Delta-\widetilde{\tau}A\nabla^2+N\cos^2\theta\right)\delta_{il}-\widetilde{\tau}C\varepsilon_{ikl}\nabla_k\right]Q_{lj}^{10}(\xi,\tau,\mathbf{x},t)\right.$$
$$=-\langle u_p\rangle\nabla_p Q_{ij}^{00}-\frac{\partial}{\partial x_p}\langle u_i\rangle Q_{pj}^{00}-\frac{\partial\langle u_p\rangle}{\partial x_s}\xi_s\nabla_p Q_{ij}^{00}$$
$$+\widetilde{\tau}\nabla_p\left[\frac{\partial Q_{ip}^{10}(\mathbf{x},t,0,0)}{\partial x_k}Q_{kj}^{00}+\frac{\partial Q_{kp}^{10}(\mathbf{x},t,0,0)}{\partial x_s}\nabla_k(\xi_s Q_{ij}^{00})\right.$$
$$\left.+Q_{kp}^{10}(\mathbf{x},t,0,0)\nabla_k Q_{ij}^{00}\right], \tag{15}$$

where $u_p=u_p(\mathbf{x},t)$ and $Q_{ij}^{00}=Q_{ij}^{00}(\xi,\tau,N)$.

Taking into account that j is a dummy index in Eq. (15), which allows us to write this equation in the vector form

$$\frac{\partial y_i}{\partial\tau}=A_{ij}y_j+f_i,$$

its formal solution being of the form

$$y_i(t)=Y_{ij}(t)y_i(0)+\int_0^t Y_{ik}(Y_{kj}^{-1}(t')f_j(t'))dt',$$

with the matrix Y_{ij} satisfying the homogeneous equation

$$\frac{\partial Y_{ij}}{\partial\tau}=A_{ik}Y_{kj}. \tag{16}$$

The solution of Eq. (14) has the form

$$Y_{ij}(\tau) = \exp(-(\nu_H k^2 + N\cos^2\theta))\Big[\cosh(\tilde{C}k\tau)$$

$$\times\Big(\delta_{ij} - \frac{k_i k_j}{k^2}\Big) + \frac{k_i k_j}{k^2} + i\varepsilon_{ikj}\frac{k_k}{k}\sinh(\tilde{C}k\tau)\Big], \quad (17)$$

where $\tilde{C} = \tilde{\tau}C$.

In this case, the expression for Q_{ij}^{10} takes the form

$$Q_{ij}^{10} = \int_0^\infty \exp(-(\nu_H k^2 + N\overline{\cos^2\theta}\tau))\Big\{\Big[\Big(-\frac{4\pi k^2}{3}\tilde{A}_0\Big.$$

$$-\frac{4\pi k^3}{15}\frac{\partial\tilde{A}_0}{\partial k}\Big)\cosh(\tilde{C}k\tau) + \frac{4\pi k^4}{15}\frac{\partial\hat{C}}{\partial k}\sinh(\tilde{C}k\tau)\Big]$$

$$\times\Big(\frac{\partial\langle u_i\rangle}{\partial x_j} + \frac{\partial\langle u_j\rangle}{\partial x_i}\Big) - \frac{4\pi}{15}\tilde{\tau}(k^2\tilde{A}_0\cosh(\tilde{C}k\tau)$$

$$-k^3\tilde{C}_0\sinh(\tilde{C}k\tau))(Q_{ij}^{10} + Q_{ji}^{10})\Big\}dk\,d\tau + Z_{ij},$$

$$(18)$$

where $Q_{ij}^{00}(k,\tau) = \hat{Q}'_{ij}$ (see Appendix), Z_{ij} are terms unrelated to the viscosity, and

$$\tilde{A}_0 = \frac{A_0}{1 + N\dfrac{2k^2\mathrm{Re}\,\overline{\cos^2\theta} + N\,\overline{\cos^4\theta}}{k^4\mathrm{Re}^{-2} + \omega^2}},$$

$$\tilde{C}_0 = \frac{C_0}{1 + N\dfrac{2k^2\mathrm{Re}\,\overline{\cos^2\theta} + N\,\overline{\cos^4\theta}}{k^4\mathrm{Re}^{-2} + \omega^2}}. \quad (19)$$

For the sake of clarity, we have applied the mean-value theorem when integrating over θ, resulting in the appearance of $\cos^2\theta$. As $N\to 0$, the expression (18) passes to the hydrodynamic limit.[12] In this case, noting that Reynolds stress appears in the equation for $\langle\mathbf{u}\rangle$ in the form

$$-\frac{\partial}{\partial x_j}(Q_{ij}^{10} + Q_{ji}^{10}),$$

we obtain the expression for the viscosity ν_t^m in a magnetic field:

$$\nu_t^m = \Big[1 + \frac{8\pi\tilde{\tau}}{15}\int_0^\infty (k^4\tilde{A}_0(k,\tau)\cosh(\tilde{C}k\tau) - k^5\tilde{C}_0(k,\tau)$$

$$\times\sinh(\tilde{C}k\tau))\exp(-(\nu_H k^2\tau + N\,\overline{\cos^2\theta})\tau)dk\,d\tau\Big]^{-1}$$

$$\times\int_0^\infty \exp(-(\nu_H k^2\tau + N\,\overline{\cos^2\theta})\tau)$$

$$\times\Big[\Big(\frac{8\pi k^2}{3}\tilde{A}_0(k,\tau) + \frac{8\pi k^3}{15}\frac{\partial\tilde{A}_0(k,\tau)}{\partial k}\Big)\cosh(\tilde{C}k\tau)$$
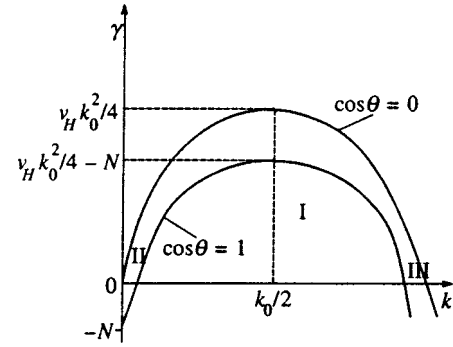

FIG. 1. Dependence of γ on k for two limiting values, $\cos\theta = 0$ and $\cos\theta = 1$.

$$-\frac{8\pi k^4}{15}\frac{\partial\tilde{C}_0(k,\tau)}{\partial k}\sinh(\tilde{C}k\tau)\Big]dk\,d\tau. \quad (20)$$

It follows from (20) that in MHD flows with $\mathrm{Re_m}\ll 1$, the turbulent viscosity decreases with increasing magnetic field.

5. DISCUSSION

As demonstrated above, a magnetic field alters the properties of homogeneous turbulence in a most significant manner. The existence of nonzero mean helicity results in instability of turbulent MHD flow against weak large-scale disturbances. However, the instability of helical MHD turbulence has distinctive features in comparison with the instability of helical turbulent hydrodynamic flow in the absence of a magnetic field.

Figure 1 schematically shows the behavior of the damping factor γ for two limiting values of $\cos^2\theta$. The regions I, II, and III correspond to the conditions (10) and (13), respectively.

Let us examine the behavior of the components of the correlation tensor in these regions. The incompressibility condition in k-space,

$$k_i Q_{ij}^{10} = 0, \quad (21)$$

leads to the following relation between the components:

$$\Big|\frac{k_3}{k_\perp}\Big| = 2\Big|\frac{Q_\perp}{Q_{3j}^{10}}\Big|, \quad (22)$$

where for the sake of simplicity and without any loss of generality, we have assumed that $k_1 Q_{1j}^{10} = k_2 Q_{2j}^{10} = k_\perp Q_\perp$. On the other hand,

$$k_i k_i = k^2. \quad (23)$$

Multiplying (23) by $Q_\perp$ and assuming that $Q_1\approx Q_2$, we find with the help of (22) that

$$Q_{3j}^{10} = \sqrt{2}\tan\theta\cdot Q_\perp. \quad (24)$$

Taking into consideration the fact that the energy density at fixed k is

$$E(k) = 2Q_\perp + Q_3, \quad (25)$$

where $Q_3 = Q_{33}^{10}$, we finally obtain

$$E(k)\cos\theta = (2\cos\theta + \sqrt{2}\sin\theta)Q_\perp. \quad (26)$$

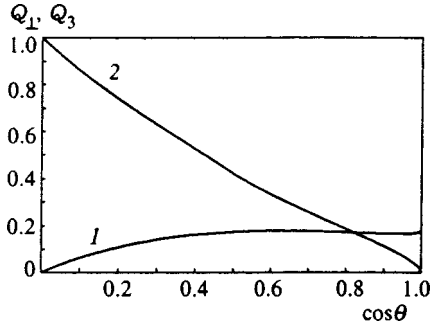FIG. 2. Angle dependence of $Q_\perp$ (1) and Q_3 (2) (normalized by $\widetilde{E}$) for $Nt=1$.

Thus, $\cos\theta$ is a measure of energy distribution among components along and across the magnetic field at specified k.

At $\cos\theta=0$ we obtain $Q_\perp=0$, and all the energy of the given mode is concentrated in the component parallel to the magnetic field, i.e., $E(k)=Q_3$. Here $k^2=2k_\perp^2$, i.e., in this mode fluctuations are normal to the magnetic field.

At $\cos\theta=1$ we observe the opposite situation: all the energy of these modes with $k^2=k_3^2$ is concentrated in fluctuations normal to the field, i.e., $E(k)=2Q_\perp$, which oscillate along the magnetic field.

If we take into account the form of $E(k)=Q_{ii}^{10}(k)$ and (8) for larger scales,

$$E(k,t)=\widetilde{E}(k,t)\exp(-Nt\,\cos^2\theta),$$

we obtain at some fixed time t

$$Q_\perp=\widetilde{E}\,\exp(-Nt\,\cos^2\theta)\,\frac{\cos\theta}{2\cos\theta+\sqrt{2}\,\sin\theta},$$

$$Q_3=\widetilde{E}\,\exp(-Nt\,\cos^2\theta)\,\frac{\sqrt{2}\,\sin\theta}{2\cos\theta+\sqrt{2}\,\sin\theta}. \quad (27)$$

Figure 2 represents the behavior of $Q_\perp/\widetilde{E}$ and $Q_3/\widetilde{E}$ for $Nt=1$ as a function of angle. One can easily see that modes with the same k behave differently, depending on $\cos\theta$.

Returning to the instability of the second moments, we have the following. In regions II and III, corresponding to the conditions (13), modes with $\cos\theta=1$ are attenuated, whereas modes with $\cos\theta=0$ grow. Hence, energy must be transferred from modes with $\cos\theta=0$ to modes with $\cos\theta=1$. In region I, energy growth is observed in all modes (but at different growth rates). In this case energy transfer between modes at fixed k probably proceeds such that at $\theta<\pi/4$, energy is transferred from Q_3 to $Q_\perp$, and conversely at $\theta>\pi/4$, from $Q_\perp$ to Q_3. However, in these cases the fluctuation amplitude will grow in a different manner. This is related to Joule dissipation, which is greatest at $\cos\theta=1$, and vanishes at $\cos\theta=0$.

The role of helicity reduces to the following. By increasing vortex lifetime at large scales, helicity slows down a direct Obukhov cascade from larger to smaller scales. Thus, it leads to incoming energy redistribution over large scales, i.e., an increase in vortex lifetime increases the probability of vortex mergers. On the other hand, at large k, helicity plays essentially no role, and at these scales turbulence is dissi-

pated. The joint action of these two processes results in energy growth at large scales, and in an efficient "erosion" of smaller scales that are weakly supplied with energy "from above" due to the presence of nonzero one-point helicity in the system. Since helicity grows with magnetic field,[18] at the same time, the connection between large and small scales is disturbed more strongly. Experimentally,[4] this leads to the energy spectrum steepening over the small-scale range, up to $\alpha=11/3-16/3$ with increasing magnetic field. Thus, the energy of a weak large-scale disturbance is redistributed among modes with differing $\cos\theta$, the cascade along the spectrum being weak.

In contrast, region I decreases with growing magnetic field, and fluctuation growth at $0<k<k_0$ is mainly associated with a rise in the energy of longitudinal fluctuations. Here, however, one must bear in mind that turbulence remains three-dimensional, but energy transfer from one component to another at a given fixed k results from rapidly occurring processes (instabilities). This leads to a quasi-two-dimensional fluctuation pattern (symmetric about the magnetic field).

It should be noted that in the intermediate range $0<\cos\theta<1$, fluctuations exist along all three components; for instance, at $\theta=\pi/4$ the intensity of modes along and across the magnetic field is the same, and they have the same instability growth rate. In the vicinity of this point, namely at $\theta\cong\pi/4$, energy exchange between components with given k is probably absent.

The authors are deeply grateful to Dr. A. Eidelman for useful comments on the results of the present paper.

APPENDIX

The influence of an external uniform magnetic field on correlations in a turbulent medium has been studied in Ref. 2. In a magnetic field, the second moment of the velocity field acquires the form

$$\hat{Q}_{ij}(\mathbf{k},\omega)=\frac{\hat{Q}'_{ij}(\mathbf{k},\omega)}{1+\dfrac{(\mathbf{k}\cdot\mathbf{B}_0)^2}{\mu\rho}\dfrac{2\eta\nu k^4-2\omega^2+(\mathbf{k}\cdot\mathbf{B}_0)^2/\mu\rho}{(\eta^2k^4+\omega^2)(\nu^2k^4+\omega^2)}}, \quad (A1)$$

where $\hat{Q}'_{ij}(\mathbf{k},\omega)$ is the correlation function in the absence of the magnetic field, and η,ν,ρ are magnetic and hydrodynamic viscosities and fluid density, respectively. Here we retain, for convenience, the notation of Ref. 2. Assuming that turbulence is helical and isotropic in the absence of the magnetic field, we can write $\hat{Q}'_{ij}(\mathbf{k},\omega)$ as follows (whether or not $H^0(0,t)=0$):

$$\hat{Q}'_{ij}(\mathbf{k},\omega)=\hat{A}_0(\mathbf{k},\omega)\left(\delta_{ij}-\frac{k_ik_j}{k^2}\right)+i\hat{C}_0(\mathbf{k},\omega)\varepsilon_{ijs}k_s. \quad (A2)$$

Passing to x-space, $Q_{ij}^{00}=\int\hat{Q}_{ij}(\mathbf{k},\omega)d^3k\,d\omega$ in the limit $\xi=x-x'\to0$, $\tau=t-t'\to0$.

When $\mathrm{Re_m} \ll 1$ we obtain from (A1)

$$\hat{Q}_{ij}(\mathbf{k},\omega) = \frac{\hat{Q}'_{ij}(\mathbf{k},\omega)}{1 + \dfrac{(\mathbf{k}\cdot\mathbf{B})^2}{\mu\rho} \dfrac{2\eta\nu k^4 + (\mathbf{k}\cdot\mathbf{B})^2/\mu\rho}{\eta^2 k^4(\nu^2 k^4 + \omega^2)}}$$

$$= \frac{\hat{Q}'_{ij}(\mathbf{k},\omega)}{1 + \dfrac{k^2\cos^2\theta B_0^2}{\mu\rho\,\eta^4} \dfrac{2\eta\nu + \cos^2\theta\,B_0^2/\mu\rho k^2}{\nu^2 k^4 + \omega^2}},$$

or in dimensionless form

$$\hat{Q}_{ij}(\mathbf{k},\omega) = \frac{\hat{Q}'_{ij}(\mathbf{k},\omega)}{1 + N\,\dfrac{2k^2\mathrm{Re}\cos^2\theta + N\cos^4\theta}{k^4\mathrm{Re}^{-2}+\omega^2}}. \qquad (A3)$$

Consequently,

$$A(N,0,0) = \frac{1}{2}\int \frac{\delta_{ij}\hat{Q}'_{ij}(\mathbf{k},\omega)}{1 + N\,\dfrac{2k^2\mathrm{Re}\cos^2\theta + N\cos^4\theta}{k^4\mathrm{Re}^{-2}+\omega^2}}\,d^3k\,d\omega,$$

$$C(N,0,0) = -\frac{i}{3}\,\varepsilon_{ijk}$$

$$\times \int \frac{k_k\hat{Q}'_{ij}(\mathbf{k},\omega)}{1 + N\,\dfrac{2k^2\mathrm{Re}\cos^2\theta + N\cos^4\theta}{k^4\mathrm{Re}^{-2}+\omega^2}}\,d^3k\,d\omega.$$

Clearly, both $A(N,0,0)$ and $C(N,0,0)$ decrease with increasing N (or magnetic field).

*)E-mail: golbref@bgumail.bgu.ac.il

[1] R. Moreau and A. Alemany, ''MHD homogeneous turbulence for small Reynolds numbers'', in *MHD Flows and Turbulence*, H. Branover (Ed.), John Wiley & Sons, New York (1976), p. 51.

[2] F. Krause and K.-H. Radler, *Mean-Field Magnetohydrodynamics and Dynamo Theory*, Academic-Verlag, Berlin (1980).

[3] B. Lehnert, Q. Appl. Math. **12**, 321 (1955).

[4] H. Branover, A. Eidelman, M. Nagorny, and M. Kireev, in *Progress in Turbulence Research*, H. Branover and Y. Unger (Eds.) (1994), Vol. 162, p. 64.

[5] A. Alemany, R. Moreau, P. L. Sulem, and U. Frish, J. Mech. **18**, 277 (1979).

[6] I. Platnieks and S. F. Seluto, in *Liquid Metal MHD*, J. Lielpeteris and R. Moreau (eds.), Kluwer Acad. Publ., Dordrecht (1989).

[7] S. S. Moiseev and O. G. Chkhetiani, Zh. Éksp. Teor. Fiz. **110**, 357 (1996) [JETP **83**, 192 (1996)].

[8] H. K. Moffatt, J. Fluid Mech. **35**, 117 (1969).

[9] H. K. Moffatt and A. Tsinober, Ann. Rev. Fluid Mech. **24**, 281 (1992).

[10] S. I. Vainshtein, Ya. B. Zeldovich, and A. A. Ruzmaikin, *Turbulent Dynamo in Astrophysics* [in Russian], Nauka, Moscow (1980).

[11] A. V. Belian, E. I. Golbraikh, S. S. Moiseev, and O. G. Chkhetiani, Pis'ma Zh. Tekh. Fiz. **22**(3), 20 (1996) [Tech. Phys. Lett. **22**, 187 (1996)].

[12] A. V. Belyan, S. S. Moiseev, and O. G. Chkhetiani, Preprint Pr-1845, Space Research Inst., Russian Academy of Sciences, Moscow (1992).

[13] S. A. Orszag, J. Fluid Mech. **41**, 363 (1970).

[14] A. S. Monin and A. M. Yaglom, *Statistical Hydromechanics* [in Russian], Gidrometeoizdat, St. Petersburg (1992), Vol. I.

[15] A. Pouquet, U. Frisch, and J. Lèorat, J. Fluid Mech. **77**, 321 (1976).

[16] R. Moreau, in *Proc. Symp. Turb. Fluids and Plasma*, Brooklyn Polytech. Inst. Press (1968).

[17] F. Krause and G. Rudiger, Astron. Nachr. **295**, 93 (1974).

[18] E. Golbraikh, O. G. Chkhetiani, and S. S. Moiseev, Zh. Éksp. Teor. Fiz. **114** (1998) to be published.